



**UNIVERSIDAD AUTÓNOMA DEL ESTADO DE  
MORELOS**

**INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS**

**CENTRO DE INVESTIGACIÓN EN INGENIERÍA Y CIENCIAS APLICADAS**

**RECOCIDO SIMULADO Y APRENDIZAJE  
AUTOMÁTICO APLICADO A LA OPTIMIZACIÓN DE  
FACTORES DE RIESGO EMPRESARIAL**

**TESIS PARA OBTENER EL GRADO DE:  
DOCTORADO EN INGENIERÍA Y CIENCIAS APLICADAS**

**MARTA LILIA ERAÑA DÍAZ**

**DIRECTOR: DR. MARCO ANTONIO CRUZ CHÁVEZ**

**CUERNAVACA, MORELOS**

**JUNIO, 2020**

Cuernavaca, Morelos, a 03 de junio de 2020.

**DR. ROSENBERG JAVIER ROMERO DOMÍNGUEZ**  
**COORDINADOR DEL POSGRADO EN**  
**INGENIERÍA Y CIENCIAS APLICADAS**  
**P R E S E N T E**

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada "RECOCIDO SIMULADO Y APRENDIZAJE AUTOMÁTICO APLICADO A LA OPTIMIZACIÓN DE FACTORES DE RIESGO EMPRESARIAL" que presenta la estudiante **MARTHA LILIA ERAÑA DÍAZ**, para obtener el título de **DOCTORADO EN INGENIERÍA Y CIENCIAS APLICADAS**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
DR. DAVID JUÁREZ ROMERO		
DRA. MARGARITA TECPOYOTL TORRES		
DR. ÁLVARO ZAMUDIO LARA		
DR. MARTÍN HERIBERTO CRUZ ROSALES (FCAeI-UAEM)		
DR. MARTÍN GERARDO MARTÍNEZ RANGEL (FCAeI-UAEM)		
DR. RAFAEL RIVERA LÓPEZ (Instituto Tecnológico de Veracruz)		
DR. MARCO ANTONIO CRUZ CHÁVEZ		

**PLAZO PARA LA REVISIÓN 20 DÍAS HÁBILES (A PARTIR DE LA FECHA DE RECEPCIÓN DEL DOCUMENTO)**

**NOTA. POR CUESTION DE REGLAMENTACIÓN LE SOLICITAMOS NO EXCEDER EL PLAZO SEÑALADO, DE LO CONTRARIO LE AGRADECEMOS SU ATENCIÓN Y NUESTRA INVITACIÓN SERÁ CANCELADA.**

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

### Sello electrónico

**MARTIN HERIBERTO CRUZ ROSALES** | Fecha:2020-06-10 20:41:30 | Firmante

TzrlvkLr75E6yPZlvBGCmgyrpAoOokkB1JJB7VsAjuJHErNwliwVR85Q79weacC5e4g5le5UIH0X3aAe3IEsdOPrDHsTC106auDO8fKky9IUWTTriN60kUoRCDn7Cq7qg8fWwoUjcmj  
GSO9ooG6K/Qd8RjtnhVNz9ctR5Kv/iCctoXVIQF8H4JbW8audVYQ8U3nVurVHD21xPu8AsB97tWQs2fdSj8wyC1Uuql+ajjQLjl/XoPxP0lwHPCC5FNABkVlKlckJwFY4IS41sjob  
Lr2KEnc1WRYNRlqxHZP98r933Disgs+NaJhdceKb3JYlks3XF2sot3T8T/U1MbA==

**MARGARITA TECPOYOTL TORRES** | Fecha:2020-06-12 00:20:03 | Firmante

H5ljj31su8OebZwou/PQaKQALc2Te0xp1zTbLc0E7knox00/4D7zeCCm0s4VySRiOfEJ981SKHeLB0SKz+2/VsNoRxygW4y6lg8g/8cJePaRzgWurEp5jq4PPwBogBRf9uVZqQiz6  
fYS1r9VafUbSfzG10Myv9VcxQo/vjyE9rQZrL+EuKYERqZYviCCHL/diYOk4h2lvZC2OpcAcJzb7XU4ZnMZ02ZfsOfkZQEa9KK5+dTvnptqUjwewlWdEsONJm+AiHsgp2G0lgijnk  
xlapaxXM5eyAmyrJ+RIW4s87iHohwz7olbUcu9DAd+45Q/SOQbuW0AlY2RHpbIA==

**MARCO ANTONIO CRUZ CHAVEZ** | Fecha:2020-06-13 18:44:41 | Firmante

pSv6T/ZFdiBAmN3+Jdsfs+KVsWCpFHLgdJ8RM+hGyT545cYDXfNwK09mP25zL67FwXhtnp39f2juL4AYzOJpm/Onl8fmJBhaW+lwIABKsvCsfIYecShiDmKgAKWzAq8Ndk99d  
cLkv3rR9CeoIT8UIRPeedz789O86rP0zRRrc5UhtPNMyWqFf0kDFsrICWLFrKhZZvHMRjdPwVPNEdNSuFDF9UCDG06DJOGp2/BTMwLtDxchi8TEVZKTBGh3PwuMivBJRqP4v3  
j3ccjt1NSyclvAO9QdEzB91z4BGzO+qUTHPUP96OcVtd7+ibfANb/ti856ffncWjQwC2+cA==

**DAVID JUAREZ ROMERO** | Fecha:2020-06-15 13:19:06 | Firmante

MJBUYEiFy3MrGaxstdAYnlyVOG45Ukr1I3QcTNoKe1F7zyTDJSlzyyJGD82K6MhIzeqleoDGicFORBQoHjB8WgERvICTYw3RHZI6db91KakesZtwPMEMIoH2IV3plwIMVv2zGF  
cOvYREUfPlxOe1DZIHWMDF3hkDAGq7Pof2zdlDcO+XAerbByEza2hdDGbHhSzGqYm5oWkrkctocq2kf+5NWHmu9mZMu9gaaw5+Obq3H5mUrjG+ZCtWMMNOE+tr3u//xXrx8tr  
hqegRwmtWYPGjaQCsrBFwPWyHSaDr+Xw5Ywjnt7u4GrnuH7fQsWzD8znX2SOfypyt+86Sg==

**MARTIN GERARDO MARTINEZ RANGEL** | Fecha:2020-06-15 14:28:54 | Firmante

SVTEdaVLeOvDjesofcbuDq86KjbxLURg2x8u8cUZbwW+VqOJO6I2DcSnpP7wtngT5WVXpX3uVoRdi5OdiPlo4hFq9sqN0Qx97I4ahMUj+Zyto7nAQhTmTrSNvSuA40813ljj2vCT  
czVU4Be9R9GlvHgpacPGQZJUJy9dfhamdRqGT3CFbjfGcS1AmPitLoEwg/JEK/5e9M2Pfxo34uaoTT7MyCwVID887xAbWrs3kpnuuq8evocurrvOvsZGpXxO2FRuSbh0QMOr5  
+PRIhZBdsBCR7zimfi8v7mzpxqkP1/0X4r09kVnWKiqZwRd2oWSVdYz9sbkrzX9r5w==

**ALVARO ZAMUDIO LARA** | Fecha:2020-06-16 13:18:13 | Firmante

syblx8asAEeYwVuK9tKph2Y/wt3D36B+NwaJ1T+OJ7gJv714R+zByhlc2CgeWKTnllyJ9a07coEZYIsAbA9XiZQG3TYogvjut6mXEGbgrZwUUAzA33efX1kChrjJRKp1UpH/Ns9CtaR  
g1cAZip/EuXPFLdKpjiWRfVnAHyogr0R9Fqh1b4wyFb9eUFGoXhrZ/x8q5yxp65pHu4Bzq4XkJrNR0eitAVGTGvXHu+Y5kGjmoyJtQcJmzq2nYrU8bEdir4mnUdhtF2YgRJuJyNmU  
6sHuZgpE4EBoTS/jmkyJoguzFFIYvbgZ9T1OY9ZCSENFkaDKfJVL63FxoXpxMo8A==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o  
escaneando el código QR ingresando la siguiente clave:



Nv9T7R

<https://efirma.uaem.mx/noRepudio/9llyMBJXmoCS3KKROODAvptlkhBzMQGE>

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

#### Sello electrónico

RIVERA LÓPEZ RAFAEL | Fecha:2020-06-22 19:52:09 | Firmante

U2RrRoQo9e9HS9CklEZvR1sb0mJtWlavk+Bdw6oOia4wSmD0NdecoY3NHcQv74Oh+tlpR4EoO8abOb65RVDEGob41Kvtj9JmrgnwPtVyt/Y02thTA7Q4oYNa+OLYB+jS4E78B8  
OGZZarbgMrtBnR1pfu0Nfir5XeyfexgimHT3OLVTTrFPMYivXNSc4mllYuMUvmAsnwcmmzQlsiaFyZL7OMrTAKNmJmwZGs+0GJiAkQmFF9tHU7m2o39m7GFILKb3Vp9+h2kv+VS  
1EYUCvgPGwYSVj7icryMTeKwAAw62M2gO40IX3sQXjKAfldkvnihGOfdmisCFHfluGKSmuZw==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o  
escaneando el código QR ingresando la siguiente clave:



[gyMOzV](#)

<https://efirma.uaem.mx/noRepudio/RdY8kXnKGyeEW8rl8nXiPjXRvKQJIMbt>

## Resumen

El riesgo empresarial, es un tipo de condición y/o situación que está presente en toda actividad empresarial, y que, dependiendo de sus características, puede llegar a impedir el cumplimiento eficiente de la misión, visión y objetivos estratégicos de la empresa. Cuando el nivel de riesgo es relevante se requiere de acciones que permitan resolver eficientemente el problema (plan de acción de riesgos), resultando para cada uno, dependiendo de las características propias de cada empresa, un conjunto de posibles soluciones. En este trabajo, se presenta una metodología computacional para la optimización del tratamiento de riesgo empresarial, se identifica el nivel de riesgo y se obtienen salvaguardas o contramedidas con las cuales es posible minimizar, o en algunos casos, evitar el riesgo.

La metodología se integra de una estructura de modelado matemático basado en la combinación de salvaguardas seleccionadas en función a las respuestas dadas por la empresa y su disposición para atender cada uno de los factores de riesgo detectados y limitados por un presupuesto asignado, de manera que la solución óptima deberá tener un costo mínimo y máximo beneficio. Para la determinación preliminar de soluciones óptimas locales, se utilizó la heurística Recocido Simulado, RS, la cual generó una serie de soluciones cercanas al óptimo, que optimizan el costo de las salvaguardas aplicadas a los factores de riesgo y que son analizadas y clasificadas mediante técnicas de aprendizaje automático (Machine Learning) para obtener una diversidad de soluciones que consuman el presupuesto asignado por la empresa bajo diferentes configuraciones. Finalmente, para evaluar en qué medida las soluciones están cercanas al óptimo, se desarrolló y resolvió el modelo matemático en el software R con el solver CPLEX.

## Abstract

Business risk is a type of condition and/or a situation that is present in all business activities, and which, depending on its characteristics, can impede the efficient fulfillment of the company's mission, vision and strategic objectives. When the level of risk is relevant, actions are required to efficiently solve the problem (risk action plan), resulting for each one, depending on the characteristics of each company, a set of possible solutions. In this work, a computational methodology for the optimization of business risk treatment is presented, the level of risk is identified and safeguards or countermeasures are obtained with which it is possible to minimize, or in some cases, avoid risk.

The methodology is made up of a mathematical modeling structure based on the combination of safeguards selected based on the responses given by the company and its willingness to address each of the risk factors detected and limited by an assigned budget, so that the Optimal solution should have a minimum cost and maximum benefit. For the preliminary determination of local optimal solutions, the Simulated Annealing, SA heuristic was used, which generated a series of near-optimal solutions that are analyzed and classified using machine learning techniques. Finally, to evaluate to what extent the solutions are close to optimal, the mathematical model was developed and solved in the R software with the CPLEX solver. The results show that the SA algorithm obtains high quality solutions in a reasonable time, and that other local optimal feasible solutions are detected.

## Agradecimientos

Agradezco a Dios, mi Padre celestial, que es mi fuente de amor, paz y bondad, quien me ha dado las habilidades y capacidades para culminar esta hermosa etapa de mi vida. A mis brillantes asesores sin quienes este trabajo hubiera sido posible. A mi familia, amigos y colegas, gracias por los momentos gratos y los no tan gratos que hemos experimentado juntos, y que han enriquecido tanto mi vida. Agradezco también, al Consejo de Ciencia y Tecnología (CONACyT) y a la Universidad Autónoma del Estado de Morelos por su apoyo y patrocinio para la realización de este proyecto de tesis.

## Dedicatoria

*A mi gran familia.*



# Nomenclatura

## Nomenclatura General

ISO 31000	Siglas en inglés para la Organización Internacional para Estandarización (International Organization for Standardization). Es la norma que proporciona principios y directrices para la gestión de riesgos.
COSO	Siglas en inglés para el Instituto de Administración de Proyectos (Project Management Institute)
PMI	Siglas en inglés para Comité de Organizaciones Patrocinadoras (Committee of Sponsoring Organizations of the Treadway)
Método William Fine	Método conocido también como "Riesgo Matemático" desarrollado por William Fine.
RS	Siglas para el algoritmo de Recocido Simulado.
CPLEX	Nombre de herramienta computacional derivado del método simplex y programado inicialmente en lenguaje C.
K-means	Nombre de método computacional de aprendizaje automático.
Árbol de Decisión	Nombre de método computacional de aprendizaje automático.
Software R	Nombre de un entorno y lenguaje de programación con un enfoque al análisis estadístico.
Script	Documentos de texto con la extensión de archivo del software correspondiente. Este lee y ejecuta el código que contienen.

### Nomenclatura del Tratamiento de Riesgo y Método de Solución

$a$	Número de factores de riesgo de una instancia.
$V_a$	Vector de $a$ factores de riesgo.
$k$	Número de tipos de respuesta al riesgo. Tipos de Salvaguardas.
$ES_{axk}$	Matriz de estructura de salvaguardas para $a$ riesgos con $k$ tipos de respuesta.
$J_{axk}$	Matriz de valores de justificación del beneficio de cada salvaguarda.
$CS_{axk}$	Matriz del costo monetario de implementación de cada salvaguarda.
$M_{axk}$	Matriz de valores de salvaguardas con $J_{axk} + CS_{axk}$ .
$JT$	Justificación total del conjunto de salvaguardas seleccionado.
$CT$	Costo total del conjunto de salvaguardas seleccionado.
$S$	Solución del conjunto de salvaguardas seleccionadas, con valores $JT$ y $CT$ .

## Índice

Resumen .....	5
Abstract.....	6
Agradecimientos .....	7
Dedicatoria.....	8
Nomenclatura.....	9
Lista de Tablas.....	8
Lista de Figuras .....	9
<b>Capítulo 1. Introducción</b> .....	11
1.1 Antecedentes.....	12
1.2. Estado del Arte .....	19
1.3 Planteamiento del Problema.....	21
1.4. Hipótesis.....	22
1.5. Justificación .....	22
1.6. Objetivos.....	23
1.6.1. Objetivo General.....	23
1.6.2. Objetivos Especificos .....	23
1.7. Aportaciones del Proyecto al Estado del Arte.....	24
<b>Capítulo 2. Marco Teórico</b> .....	26
2.1 Riesgo Empresarial. ....	26
2.1.1 ISO 31000 .....	26
2.1.2 Modelo de Altman .....	27
2.1.3 Riesgo de Quiebra o Insolvencia.....	30
2.2. Respuesta a los factores de riesgo.....	33
2.2.1 Tipos de salvaguardas (preventivas, correctivas, de control).....	34
2.3. Escalas de Valoración del Riesgo .....	34
2.3.1. Metodología PMI .....	34
2.3.2. COSO.....	36
2.3.3. Método William Fine. Riesgo Matemático .....	40
2.4. Modelo de Optimización .....	44

2.4.1. Optimización.....	44
2.4.2. Formulación Matemática del Modelo de Optimización .....	46
2.5. Heurísticas.....	48
2.5.1 Heurísticas de Búsqueda Local.....	48
2.5.2. Heurística Recocido Simulado .....	49
2.6. Aprendizaje Automático (Machine Learning) .....	51
2.6.1. K-means .....	52
2.6.2. Árboles de decisión .....	54
2.7. Complejidad computacional.....	55
2.8. Errores en la Toma de decisiones .....	56
<b>Capítulo 3. Metodología para la Optimización del Tratamiento de Riesgos Empresariales .....</b>	<b>58</b>
Algoritmo 1. Recocido Simulado para tratamiento factores de riesgo. ...	63
Algoritmo 2. Estructura de Vecindad $\Omega$ del RS.....	64
Algoritmo 3. R Script para clasificación de soluciones RS con k-means.	65
Algoritmo 4. R Script para factores relevantes de grupos de soluciones mediante Árbol de Decisión.....	66
<b>Capítulo 4. Aplicación de la metodología (caso de estudio).....</b>	<b>67</b>
4.1 Caso de Estudio .....	67
4.2. Aplicación de la Metodología.....	68
<b>Capítulo 5. Resultados Experimentales.....</b>	<b>92</b>
5.1 Caso de Estudio. ....	92
5.2 Desempeño Computacional del Recocido Simulado.....	93
<b>Capítulo 6. Discusión de los Resultados.....</b>	<b>99</b>
<b>Capítulo 7. Conclusiones y Trabajos Futuros .....</b>	<b>101</b>
7.1 Conclusiones .....	101
7.2 Trabajos Futuros.....	103
<b>Referencias.....</b>	<b>104</b>
Apéndice A. Cuestionario sobre Factores de Riesgo Empresarial.....	118
Apéndice B. Código fuente de la metaheurística Recocido Simulado. ....	120
Apéndice C. Productos Generados.....	133

## Lista de Tablas

<i>Número</i>	<i>Página</i>
2.1. Tabla de Resultados . . . . .	29
2.2. Factores de Riesgo por Dimensión. . . . .	30
2.3 Técnicas para análisis de riesgo del PMBoK. . . . .	34
2.4 Escala del Método W. Fine para valorar el grado de riesgo. . . . .	42
2.5 Escala del W. Fine para cálculo del beneficio de la salvaguarda. . . . .	43
2.6 Enfoques de solución de problemas computacionales. . . . .	45
4.1. Categoría de los Factores de Riesgo. . . . .	71
4.2. Matriz de Justificación y de Costo de Implementación . . . . .	75
4.3. Conjunto de Soluciones Vecinas. . . . .	77
4.4. Variables RS . . . . .	78
4.5. Parámetros RS . . . . .	78
4.6. Promedios, Des. Est, Min, Max por clúster . . . . .	83
4.7. Tipos de respuesta más relevantes para cada factor. . . . .	85
4.8. Solución Elegida a partir de la interfaz figura 4.17. . . . .	88
5.1. Solución Clúster 1 . . . . .	91
5.2. Valor de la FO obtenida por las 10 instancias . . . . .	92

## Lista de Figuras

<i>Número</i>	<i>Página</i>
1.1. Administración de los riesgos. Fuente COSO,2014 . . . . .	14
1.2. Tono del riesgo en la empresa. Fuente COSO,2014 . . . . .	15
1.3. Clasificación de los riesgos, bancos de Corea. [10] . . . . .	16
2.1. Gráfica Zonas Riesgo Modelo Altman . . . . .	28
2.2. Dimensiones y Factores de Riesgo en la Empresa. . . . .	36
2.3. Optima Toma de Riesgo Fuente COSO. . . . .	37
3.1. Diagrama de metodología Tratamiento de Factores de Riesgo.. . . .	58
3.2. Factores en seis áreas empresariales. . . . .	59
3.3. Matriz ES <sub>axk</sub> , estructura de los valores de las Salvaguardas. . . . .	61
4.1. Necesidades de la empresa para continuar operando . . . . .	67
4.2. Cuestionario sobre Actividad y Desempeño Empresarial . . . . .	68
4.3. Nivel de riesgo 32 factores. Riesgo - Rojo, SIN Riesgo – Verde . .	69
4.4. Tablas de Fine en interfaz electrónica para ponderar los factores .	69
4.5. Porcentaje del Riesgo por Categoría. . . . .	70
4.6. Salvaguardas ordenadas por los tipos de respuesta proporcionada.	73
4.7. Interfaz automatizada con los valores de Fine . . . . .	74
4.8. Interfaz automatizada para elección de corrección y coste . . . . .	75
4.9. Diagrama de Flujo del RS adaptado. . . . .	79
4.10. Ciclo Metròpoli del RS . . . . .	80
4.11. Convergencia del RS . . . . .	81
4.12. Tres pruebas de 30 soluciones casi óptimas de RS . . . . .	82
4.13. Grupos clasificados por la Justificación Total. . . . .	83
4.14. Árbol con la clasificación sobre el Total Justificación por Clúster. .	84
4.15. Respuestas Relevantes a los factores de riesgo por clúster . . . . .	85
4.16. Interfaz con Tablas Dinámicas. . . . .	86
4.17. Interfaz de usuario para seleccionar y ver una solución. . . . .	87
5.1 Consumo de memoria utilizado (Mb) por RS y CPLEX solver . . . . .	93
5.2 Prueba de Normalidad para los datos de CPLEX y RS . . . . .	94

5.3: Análisis de Homocedasticidad de los datos de CPLEX y RS . . . . . 90

## Capítulo 1. Introducción

Una empresa por lo general trabaja en un entorno cambiante, por lo que está expuesta a todo tipo de riesgos al trabajar en todas sus áreas tanto internas, como externas (administrativas, operativas, dentro del sector al que pertenece el negocio, etc.). Si estos riesgos no se gestionan adecuadamente, pueden afectar a la empresa en su rentabilidad y a largo plazo pueden desviarla de su misión y visión. Proteger estratégicamente el valor de sus activos actuales y crear nuevos (activos/valor) para el futuro, es uno de los principales objetivos de éstas, de manera que, el riesgo empresarial puede ser definido como el impacto y la probabilidad de que una situación con resultados negativos, o serie de eventos no favorables, puedan afectar de manera adversa la consecución de los objetivos [1]. De acuerdo con el informe presentado en [2,3], 8 de cada 10 empresas fracasan en los primeros dos años, los resultados de encuestas a empresarios muestran que esto se debe principalmente a dos motivos; una planeación defectuosa y la falta de habilidad para las ventas, sin embargo, el punto de vista del sector gubernamental, a través del Instituto Nacional del Emprendedor, INADEM, indica que el fracaso se debe a la falta de aptitudes empresariales y escaso conocimiento del ecosistema, e indican que la gran mayoría de los emprendedores mexicanos no están preparados para administrar una empresa y no conocen todas las herramientas empresariales que tienen a su disposición. Según la Organización para la Cooperación y el Desarrollo Económicos (OCDE) [4], la capacidad gerencial de los emprendedores en México se beneficia con la implementación de indicadores de desempeño para medir la efectividad del modelo de negocio, pero sólo el 30.2% de los negocios medianos monitorean de 3 a 5 indicadores. En contraste, el 65% de las micro organizaciones y el 35.7% de los pequeños emprendimientos no usaban ningún tipo de guía de rendimiento. Así, podemos afirmar que las empresas adquieren grandes ventajas al generar un plan, ya que, además de dar dirección al negocio, al incluir un plan de administración



de riesgo, se podrá anticipar a las amenazas, fortalecer sus vulnerabilidades y evolucionar sanamente evitando la quiebra y/o el fracaso empresarial [5–8].

### 1.1 Antecedentes

Los factores de riesgo empresarial son acciones o elementos que representan una amenaza o que son una vulnerabilidad de la empresa que pueden complicar su evolución y alejarla de sus metas estratégicas [9]. Desde la perspectiva de los inversionistas, los riesgos se agrupan en cuatro categorías relevantes: estratégicos, operativos, financieros y de cumplimiento. Los cuales se definen a continuación.

1. Riesgos Estratégicos - emergen de los objetivos estratégicos. La alta gerencia identifica a los riesgos más importantes a través del proceso de planificación, posteriormente se validan y aprueban por la junta directiva.
2. Riesgos Operativos – grandes riesgos que afectan la habilidad de la organización para ejecutar al plan estratégico.
3. Riesgos Financieros – incluyen información financiera, valoración, cobertura, riesgos de mercado - liquidez y riesgos de crédito en instituciones financieras.
4. Riesgos de Cumplimiento – riesgos no compensados, generalmente son el foco principal para las actividades de gestión de riesgo empresarial.

El tratamiento del riesgo implica atender los factores que los detonan mediante la selección e implementación de una o varias acciones, de manera que se disminuya o, en la medida de lo posible, se evite el riesgo. Estas acciones son llamadas salvaguardas o contramedidas las cuales deben implementarse con estrategias adecuadas ya que promueven adecuaciones en procesos y controles de la empresa [11]. Entre las diversas metodologías para el tratamiento de riesgo, está la generada por la Universidad Carnegie Mellon

con el PMBoK (Project Management Body of Knowledge) del PMI (Project Management Institute) [12-14] y la ISO 31000 [15] que presentan las mejores prácticas en la gestión del riesgo empresarial.

Estos principios y lineamientos fueron considerados en el desarrollo de la metodología del COSO [9, 16] (Committee of Sponsoring Organizations of the Treadway), la cual surgió en 1985 con el objeto de ayudar a las organizaciones a evaluar y mejorar sus sistemas de control interno. Apoyándonos en las estrategias y técnicas del COSO para la identificación y tratamiento del riesgo empresarial, el riesgo se considera: a) por forma de gobierno: la manera en que se marca los lineamientos de control de operaciones, código ético, etc.; b) por operaciones e Infraestructura; c) por finanzas: como el manejo de portafolio de inversión, acciones, etc.; d) de cumplimiento: regulaciones fiscales, de medio ambiente, etc. y e) por estrategias: con clientes, proveedores, empleados, etc. De manera que según esta metodología COSO, todo lo que amenace el valor para la empresa, es decir, lo que impida el crecimiento de ganancias, limite el margen operativo e interfiera en la eficiencia de activos, todo esto, es lo que genera los principales riesgos empresariales. En la figura 1.1 se muestra como la administración del riesgo, se da a través del desarrollo e implementación de estrategias que requieren mantenimiento y actualización hacia la mejora continua. La clasificación realizada se muestra en la base de la pirámide en cinco categorías: el gobierno o grupo directivo, las estrategias y/o la planeación, las operaciones e infraestructura, el cumplimiento y los reportes financieros. Continuando con la figura 1.1, una base común del riesgo generalmente está en los procesos, las personas y la tecnología, esto se representa como la infraestructura del riesgo. Con respecto a la cúspide, las acciones de vigilancia y el reconocimiento del tono ante los riesgos por parte de la dirección, son elementos que facilitan el ciclo de retroalimentación para avanzar hacia la mejora continua. El tono del riesgo, implica conocer el tipo de estrategias que se deben utilizar para su

manejo, como el hecho de estar abiertos a correr riesgo o ser conservadores frente a éste, es determinar la línea de pensamiento sobre el riesgo que regirá a la empresa, ver figura 1.2.

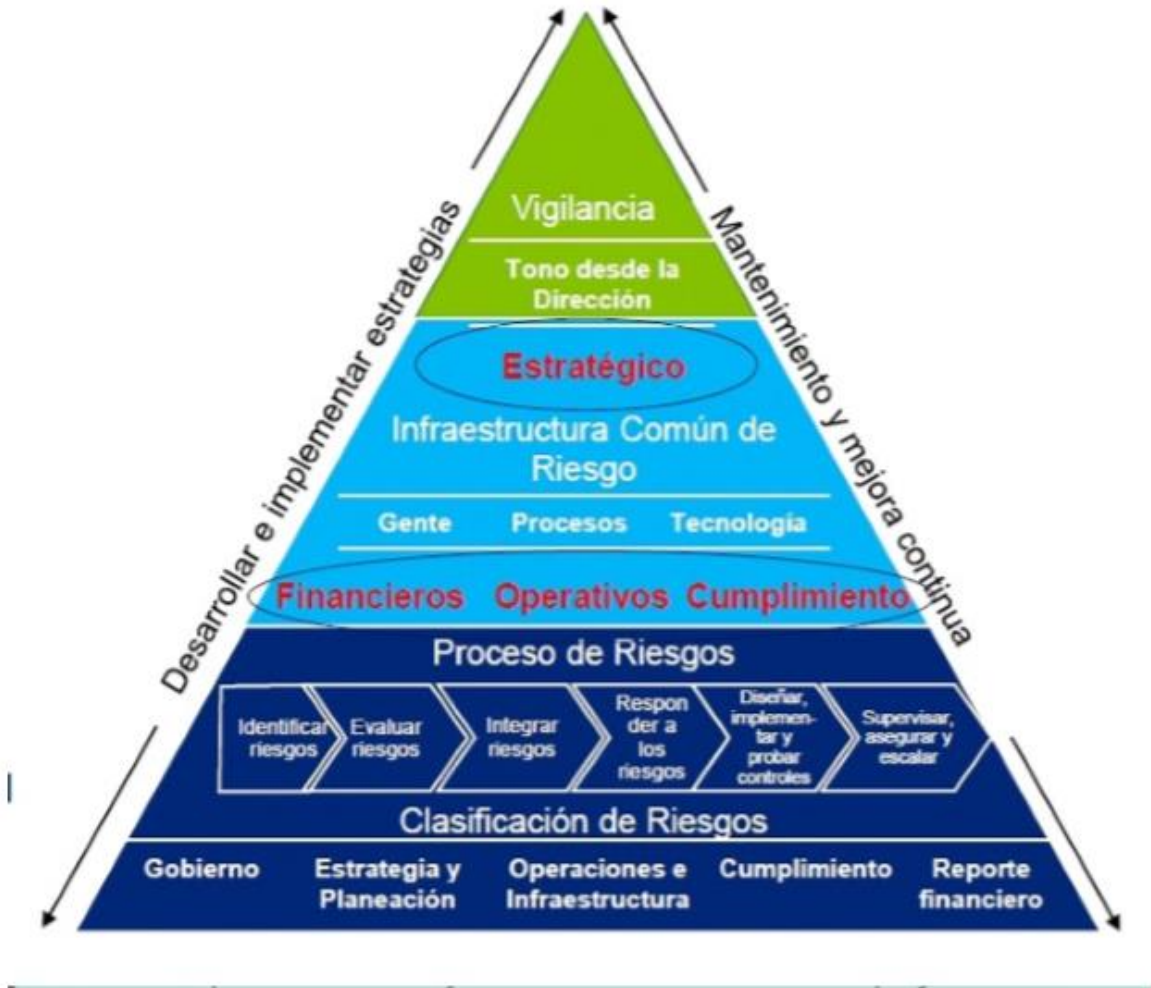


Figura 1.1 Administración de los riesgos. Fuente COSO,2014.



Figura 1.2 Tono del riesgo en la empresa. Fuente COSO,2014.

En Kim et al., [10], se realizó un estudio sobre los factores a considerar que permiten determinar si una empresa tiene un buen desempeño, de manera que garantice su estabilidad en el mercado. Con el propósito de extraer el conocimiento cualitativo de los expertos en determinar si una empresa estaba avanzando hacia una quiebra empresarial, se desarrolló un algoritmo genético con una muestra de 250 compañías de manufactura y servicios, 143 están clasificadas en el grupo de quiebra y 107 en el grupo de No-quiebra y fueron calificados cualitativamente, en 1) Positivo (sin riesgo), 2) Neutro y 3) Negativo (riesgo). Los expertos que calificaron estas compañías son oficiales de crédito experimentados de importantes bancos de Corea, que cuentan con 9 años en promedio de experiencia en esta área. Realizaron una evaluación de seis áreas de riesgo, en cada una de estas categorías, se encuentran los factores de riesgo, figura 1.3. El algoritmo genético fue contrastado con una red neuronal artificial, determinando con 11 reglas las combinaciones que producen la quiebra empresarial, con el propósito de conocer si la empresa es sujeta a crédito o no.

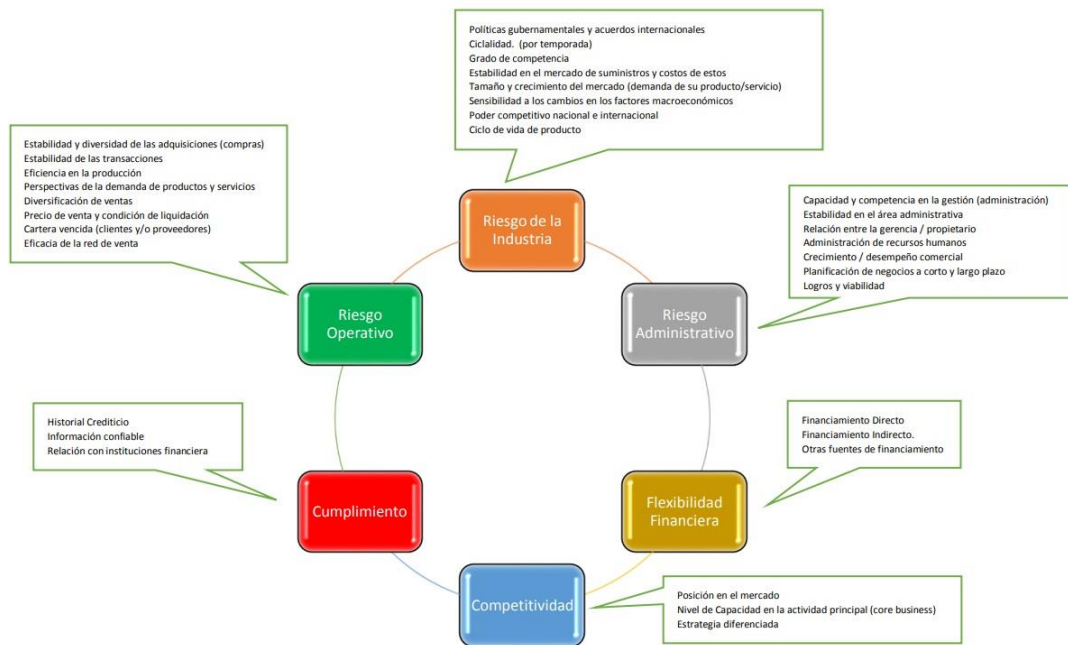


Figura 1.3 Clasificación de los factores de riesgos en seis categorías. Fuente: Bancos de Corea Kim et al., [10].

Para la valoración del riesgo, William Fine [17] presenta una metodología, conocida como el riesgo matemático para evaluar el grado de peligrosidad (GP), la cual consiste en valorar tres criterios: las Consecuencias (C), la Exposición (E) y la Probabilidad (P), evaluando con ello si está justificada la aplicación de salvaguardas. Esta metodología ha sido ampliamente utilizada para realizar el análisis de riesgos en diferentes sectores e industrias [18–21].

Con respecto a la selección de salvaguardas, que son las medidas preventivas o correctivas que atienden a los factores de riesgo, este proceso representa todo un desafío, ya que es deseable reducir las amenazas muchas veces con un presupuesto limitado en el corto plazo. El principio “As Low As Reasonably Practicable”, ALARP [22,23], que significa “tan bajo como sea razonablemente posible”, utiliza análisis de costo-beneficio que explícitamente refleja el peso que la toma de decisiones otorga a los diversos atributos, ¿pero?, ¿se están considerando las consecuencias posibles y asociadas? Las incertidumbres

están vinculadas a las diversas alternativas disponibles, por lo tanto, en la selección de un conjunto de salvaguardas para atender al factor de riesgo, el riesgo es ALARP, significa tomar un principio de precaución. El costo por sí solo nunca debe ser la justificación, además del costo financiero directo, el impacto económico indirecto, como lo es una interrupción comercial más prolongada y/o un impacto en la reputación de la empresa, pueden ser causa de un riesgo mayor debido a que a la larga su posible alcance financiero lleve al fracaso [24].

La mayoría de estas metodologías para evaluar al riesgo, se limita a aplicar matrices de probabilidades y consecuencias, dependiendo completamente de las habilidades, experiencia y pensamiento crítico de los evaluadores. La metodología propuesta en esta investigación concuerda en esas cualidades de los tomadores de decisiones, pero propone la realización y ponderación de más de una respuesta a cada factor, es decir, designar  $k$  tipos de respuestas, aunque muchas veces parecen no viables, es importante utilizar ese pensamiento crítico y experiencia para identificar varias soluciones a dichos factores de manera que se eviten sesgos en los juicios durante la toma de decisión. Posteriormente, mediante tablas que uniforman las respuestas, se califican de acuerdo con el grado de corrección o mejora del entorno organizacional, y, el costo que se generaría al implementarla en la empresa. Este costo debe incluir no solo al aspecto económico, sino al grado de esfuerzo que requiere la implementación de esa salvaguarda (tiempo de capacitación, resistencia al cambio, etc.). El PMI clasifica las respuestas al riesgo en cuatro categorías, evitar, transferir, mitigar y aceptar. En [25] analizan el riesgo económico y competitivo y las clasifican en naturales, operacionales, físicas y electrónicas. Utilizan la selección del conjunto de salvaguardas para determinar las pólizas de seguridad y los presupuestos. Diseñan también el programa de mitigación en los riesgos. En ambos casos, se genera una matriz de posibles salvaguardas, como las opciones de un portafolio, y, la selección

de la respuesta óptima a cada uno de los factores de riesgo, genera un problema de programación lineal entera.

El CPLEX es una herramienta tecnológica de alto desempeño para resolver modelos de programación lineal, entera mixta y cuadrática [26–29], CPLEX es el nombre asignado del método simplex implementado inicialmente en el lenguaje de programación C, su desarrollo en la actualidad ofrece múltiples interfaces. Fue desarrollado originalmente por Robert E. Bixby en 1988. El software R [30] es una de las herramientas más utilizadas en estadística computacional y en la ciencia de datos donde una amplia comunidad de usuarios desarrolla paquetes para este sistema. El paquete “Rglpk” [31] es una interfaz de alto nivel de R al CPLEX\_LP. Así que, entre las múltiples opciones para resolver al modelo matemático, independientemente del lenguaje y herramienta que se utilice, se requiere considerar tres puntos que todo software necesita [32]

- Construcción del modelo: ¿Cómo construyo al modelo de optimización?, ¿Se realiza con una restricción o con una matriz de restricciones?
- Ajustar los parámetros del solver: ¿Cuáles parámetros del programa deben cambiarse? ¿Qué efectos intentan generar esos cambios?
- Calcular y extraer la solución: ¿Se busca una solución óptima o solo una buena solución factible? ¿Cómo se extrae la solución que genera el solver? ¿Se extrae una por una las variables de la solución o se extrae al vector completo?

## 1.2. Estado del Arte

Las técnicas heurísticas de optimización son estructuralmente simples y confiables [33–38], se presentan a menudo como la mejor estrategia disponible para la solución de problemas complejos. En la revisión bibliográfica sobre la aplicación de heurísticas en el tratamiento al riesgo empresarial se encontró en [39] el desarrollo de un algoritmo genético para búsqueda de la mejor combinación entre costo de amenazas, de salvaguardas y el impacto en los activos mediante un sistema para soporte en la toma de decisiones sobre seguridad en informática. Se comparó con un algoritmo de búsqueda exhaustiva donde se demostró que este tipo de búsqueda es, computacionalmente, mucho más costosa en recursos como en tiempo, además, la heurística implementada, obtuvo el 30% de soluciones óptimas, concluyendo que el algoritmo hace que la toma de decisiones sea más rápida. En [40], Falsai propone una metodología que considera las interconexiones del riesgo e identifica la mejor combinación de estrategias de mitigación bajo restricciones de presupuesto, se analiza un caso de estudio de una empresa de manufactura. En [41–43] presentan al problema de optimizar el modelo de negocios de las empresas, Business Process Optimization (BPO). Genera un modelo de optimización mediante una serie de algoritmos evolutivos de optimización multiobjetivo, Evolutionary Multi-Objective Optimization Algorithms (EMOAs), una de las principales aportaciones de esta investigación, es un conjunto de soluciones alternativas del BPO dentro de un rango de escenarios experimentales y de la vida real. Para validar al modelo de optimización y la evaluación de los resultados, se probó un conjunto de escenarios de la vida real, se realizó mediante la comparación directa de ellos con los producidos por un algoritmo de fuerza- bruta, la heurística arrojó resultados satisfactorios, logrando reducir drásticamente el conjunto de datos de todos los escenarios examinados. Tupa et al., [44], realiza un estudio sobre los riesgos empresariales bajo el concepto de Industria 4.0 con la idea de encontrar aspectos involucrados en la gestión del riesgo, ya que la



interconexión de humanos, objetos y sistemas bajo este enfoque, genera interacciones en una dinámica auto-organizada y en tiempo real. Este impacto presenta nuevos riesgos, el artículo hace énfasis en que su gestión cambia de una manera relevante, generando la oportunidad de encontrar nuevos métodos para su tratamiento, incluyendo heurísticas bajo entornos "big data".

De manera que existen múltiples metodologías para evaluar el riesgo [14, 16, 45-46], sin embargo, para su atención, se limita a aplicar matrices de probabilidades y consecuencias, dependiendo completamente de las habilidades, experiencia y pensamiento crítico de los emuladores. La metodología propuesta en esta investigación concuerda con estas cualidades de la toma de decisiones, pero propone la realización y ponderación de más de una respuesta a cada factor, es decir, designar muchos tipos de respuestas, aunque a menudo parecen no viables, es importante usar ese pensamiento crítico y experiencia para identificar varias soluciones a los factores de riesgo. Posteriormente, la justificación de la salvaguarda es la evaluación por el método William Fine [17], los califica en el grado de corrección o mejora de los beneficios en el entorno organizacional y, en el costo que se generaría al implementarlo en la empresa, este costo no debe incluir solo el aspecto económico, pero el grado de esfuerzo que requiere la implementación de esta salvaguarda (tiempo de entrenamiento, resistencia al cambio, etc.).

Otro aspecto importante es que los métodos existentes no consideran resolver un modelo de optimización con respuestas de diversos tipos para tratar los factores de riesgo de la empresa mediante heurísticas computacionales como lo es la heurística RS [47], además, con las técnicas de aprendizaje automático, se agrupan las soluciones factibles cercanas al óptimo, que proveen a los tomadores de decisiones de alternativas muy eficaces para elaborar los planes de atención a los riesgos. Esto la hace más confiable y eficiente, complementándose con la interfaz gráfica que facilita la visualización

de los resultados. De acuerdo con [48–52], este tipo de implementaciones tecnológicas con el manejo de ETL (extracción, transformación y carga) de datos, serán en los próximos años las herramientas más utilizadas, esto es, la metodología planteada es actual, versátil y tecnológicamente de punta.

Utilizar herramientas computacionales para llevar a cabo el análisis de riesgos en una empresa, permite realizar acciones de manera más sistemática, facilitando que la información resultante sea reutilizable y comparable con resultados de sucesivos análisis. Además, permite una óptima toma de decisiones que contribuye en un mejor tratamiento a los factores de riesgo, incrementando así, la seguridad empresarial

### 1.3 Planteamiento del Problema

Una empresa que desconoce el grado de riesgo en el que trabaja, y por ende no lo trata para eliminarlo, minimizarlo o sobrellevarlo, es propensa a sufrir una situación que le impida crecer, mantenerse o que entre en recesión [53]. El tratamiento adecuado de los riesgos empresariales potencializa la posibilidad de una empresa a mantenerse en operación, al menos, bajo las mismas condiciones actuales. Existen en el estado del arte muchos métodos con los que actualmente tratan el riesgo empresarial, algunos optimizan la solución, sin embargo, no se cuenta a la fecha con una metodología con la cual, a través de la identificación, caracterización de los riesgos y con la consideración de restricciones de inversión, se genere una configuración de salvaguardas eficientes que muestren el costo-beneficio para que se realice la toma de decisiones con la mejor combinación, permitiendo que las salvaguardas seleccionadas puedan incorporarse en los planes empresariales que atienden las necesidades de la empresa.

Sobre las diversas acciones a seguir en atención a los factores de riesgo empresarial, no se cuenta con evidencias de que las metodologías utilizadas, apliquen procesos de optimización en el tratamiento de riesgos y en la búsqueda de soluciones, lo cual debe ser muy relevante para la toma de decisiones sujetas al presupuesto asignado ya que, al optimizar el costo de la solución de seguridad a implementar, se justifica que las salvaguardas seleccionadas, generen un grado mayor de corrección del riesgo con el que opera la empresa.

#### 1.4. Hipótesis

Con la aplicación de heurísticas computacionales, como el Recocido Simulado y los algoritmos de aprendizaje automático y la incorporación de las mejores prácticas para el tratamiento de los factores de riesgo empresarial, es posible diseñar y validar una metodología para la obtención de soluciones cercanas al óptimo las cuales reducen el costo de las salvaguardas que se aplican en el tratamiento de los factores de riesgo.

#### 1.5. Justificación

Dada la complejidad de una buena selección de salvaguardas para minimizar los factores de riesgo empresarial sujetos a un presupuesto asignado por la gerencia, el propósito de este estudio es analizar las áreas del riesgo empresarial y los factores de cada una de ellas, posteriormente cuantificar el nivel de riesgo y determinar las posibles salvaguardas a cada factor en los cuatro tipos de respuesta que son: aceptar, mitigar, transferir o evitar. Así, la principal contribución de este estudio es la aplicación de la heurística Recocido

Simulado, RS, para ofrecer soluciones cercanas al óptimo resolviendo el modelo de optimización que optimiza el costo de la solución de seguridad a implementar sin rebasar el presupuesto que se tiene y que justifica un alto grado de protección en la empresa. De manera que tenemos un vector  $\mathbf{V}_a$  de factores de riesgo y una matriz de  $M(a \times s)$  salvaguardas, donde  $a$  es el número de factores y  $s$  es el número de salvaguardas posible para cada factor, y donde cada salvaguarda tiene un valor  $J_{i,k}$  ( $i=1, \dots, a$ ) que es la justificación por reducción del riesgo del factor  $i$  utilizando la salvaguarda tipo  $k$  ( $k=1, \dots, s$ ), con un costo de implementación de  $CS_{i,k}$ . El problema es identificar un subconjunto de todas las salvaguardas que generen el menor costo con una mayor justificación de protección y que no exceda el presupuesto máximo asignado  $P$ .

## 1.6. Objetivos

### 1.6.1. Objetivo General

Diseñar y desarrollar una metodología computacional para el tratamiento óptimo de los factores de riesgo empresarial bajo el presupuesto con el que cuenta una empresa, mediante heurísticas y un modelo de optimización, con el propósito de identificar y reducir los efectos clasificados como perjudiciales en la empresa.

### 1.6.2. Objetivos Específicos

- Identificar y definir los parámetros y variables que son relevantes en el problema de riesgos empresariales.
- Adaptar para los factores de riesgo, la escala de valoración conforme a los objetivos empresariales.

- Implementar las técnicas de análisis de datos basadas en un modelo de optimización y heurísticas computacionales (recocido simulado, k-means, árboles decisión) a los elementos de riesgo ponderado.
- Desarrollar una herramienta tecnológica que valide la metodología para una fácil interpretación de los resultados.
- 

### 1.7. Aportaciones del Proyecto al Estado del Arte

La metodología presentada en este trabajo ha demostrado ser efectiva para la toma de decisiones, ayudando a los gerentes a obtener un conjunto de soluciones casi óptimas con un presupuesto limitado e incluirlas en el plan de atención a los riesgos de la empresa. Esto, permitirá anticiparse a las amenazas, fortalecer las vulnerabilidades y evolucionar sanamente evitando la quiebra o el fracaso empresarial

Dentro de las principales aportaciones están:

- Adaptación y desarrollo de la heurística computacional Recocido Simulado para el tratamiento óptimo de los riesgos empresariales.
- Desarrollo de un marco metodológico para optimizar la aplicación de salvaguardas para atender los factores de riesgo.

La organización del resto del documento es la siguiente: En el capítulo 2 se explican los conceptos generales de los elementos de esta investigación en el marco teórico. En el capítulo 3 se realiza una descripción de la propuesta para solucionar el problema del tratamiento óptimo de los riesgos empresariales usando la heurística RS. El capítulo 4 está dedicado al caso de estudio con la aplicación de la metodología. En el capítulo 5 se presentan los resultados experimentales del caso de estudio y del desempeño computacional del RS

con respecto a la solución en CPLEX. En el capítulo 6 se presenta la discusión de los resultados y el capítulo 7 contiene las conclusiones y trabajos futuros.

## **Capítulo 2. Marco Teórico**

En este capítulo se presenta el soporte conceptual y teórico en los que se cimienta la metodología planteada en este trabajo. Así también, se explica brevemente, en qué consisten las herramientas utilizadas para resolver el problema de realizar una selección óptima en la respuesta a cada uno de los factores de riesgo empresarial.

### **2.1 Riesgo Empresarial.**

La palabra riesgo, desde sus orígenes, se relaciona con peligro, con hechos adversos que pueden suceder, es incertidumbre ante una posible ocurrencia de un evento que puede generar pérdidas. Riesgo empresarial es toda aquella amenaza o vulnerabilidad que pudiera afectar la estabilidad o el cumplimiento de los objetivos estratégicos de una empresa. En la actualidad, se ha incrementado considerablemente el número y la potencialidad de los riesgos que inciden sobre la empresa, haciendo más imperiosa la necesidad de protección e impulsando a las empresas a emprender estrategias que permitan optimizar sus recursos y minimizar los riesgos. Inicialmente la disciplina de la administración de riesgos se dedicó al estudio de los riesgos financieros y con el tiempo amplió su cubrimiento a diversos tipos de riesgos que impactan el negocio [54-57]. Así que la evolución de la administración de riesgos ha permitido que se establezcan acciones, ya no aisladas, sino de manera estructurada e integral, para identificar, calificar, evaluar y monitorear todo tipo de riesgos que puedan afectar el cumplimiento de los objetivos de las organizaciones, con el propósito de responder con medidas efectivas para su manejo.

#### **2.1.1 ISO 31000**

Con respecto a la normatividad sobre riesgos empresariales, la norma ISO 31000 es un importante referente internacional que establece para la administración de éstos, principios que fundamentan la gestión de riesgos, un

marco de referencia que delimita y direcciona la misma y un proceso para la gestión de riesgo que facilita su ejecución. Define el riesgo en términos del efecto de la incertidumbre de los objetivos de la empresa y no está diseñada para un sistema de gestión en particular, sino que provee una estructura de mejores prácticas y guía, para todas las operaciones relacionadas con la gestión de riesgos. Dentro de los principios de esta norma está:

- Crear y proteger el valor
- Estar integrada a todos los procesos de la organización
- Ser parte de la toma de decisiones
- Tratar explícitamente la incertidumbre
- Ser sistemática, estructurada y oportuna
- Buscarse en la mejor información disponible
- Alinearse al contexto y al perfil del riesgo de la organización
- Tener en cuenta los factores humanos y culturales
- Ser transparente e inclusiva
- Ser dinámica, iterativa y sensible al cambio
- Facilitar la mejora continua

### 2.1.2 Modelo de Altman

El modelo de Altman se centra en un modelo financiero, donde la empresa es vista como un ente dentro de la actividad económica en la que los diferentes agentes actúan motivado por la obtención de un valor económico que esté de acuerdo con sus expectativas.

El modelo Z-Score de Altman ayuda a las empresas e inversionistas a identificar si están tomando las decisiones acertadas para continuar operando en el mercado. El modelo ZScore de Edward I. Altman [58,59] es uno de los primeros modelos que predice con anterioridad la quiebra a empresas de diferentes sectores. Mediante variables y ratios financieros, evalúa si una



empresa se está administrando de forma adecuada financieramente o si existe el riesgo de entrar en quiebra.

Los cinco indicadores financieros que fueron seleccionados por Altman para realizar el modelo, los retomó por ser los más significativos dentro de una derivación estadística, clasificados en: indicador de liquidez, indicador de rentabilidad, indicador de apalancamiento, indicador de solvencia e indicador de actividad.

Con estos indicadores se trabaja en la formulación para hallar el resultado final, este resultado tiene como nombre Z donde se usa la siguiente ecuación para empresas en el sector comercial y de servicios:

$$Z = 6.56X_1 + 3.267X_2 + 6.72X_3 + 1.05242X_4.$$

Y para empresas de manufactura

$$Z = 0.717X_1 + 0.847X_2 + 3.107X_3 + 0.42X_4 + 0.998X_5.$$

Donde

$X_1 = (\text{Capital de trabajo/Activos totales})$

$X_2 = (\text{Utilidad/Activo total})$

$X_3 = (\text{U.A.I./Activo total})$

$X_4 = (\text{Valor contable del patrimonio/Pasivo total})$

$X_5 = (\text{Ventas/Activos totales})$

$X_1 = (\text{Capital de trabajo/Activos totales})$  - para hallar este índice se utiliza el capital de trabajo neto, que es la diferencia entre los activos corrientes versus los pasivos corrientes, estos activos incluyen el disponible de la compañía y los bienes que rápidamente pueden volverse efectivo. Los pasivos corrientes son las obligaciones contraídas con terceros, los cuales deben estar pagados en un plazo de un año o menor.

$X_2 = (\text{Utilidad/Activo total})$  - para hallar este índice se toma la utilidad obtenida durante el ejercicio vigente (año gravable), tomado del Estado de Resultados.

$X_3 = (\text{U.A.I.I./Activo total})$  - el U.A.I.I., son las utilidades antes de impuestos e intereses y se obtienen del Estado de Resultados de la empresa, antes de calcular y provisionar los impuestos. El activo total es la suma de los activos corrientes y no corrientes.

$X_4 = (\text{Valor contable del patrimonio/Pasivo total})$  - el valor contable del patrimonio es tomado del balance general, se toma el valor registrado en libros, que es diferente al valor comercial de la empresa, comparado con el pasivo total que son las obligaciones financieras que tiene la empresa.

$X_5 = (\text{Ventas/Activos totales})$  - las ventas corresponden a los ingresos operacionales de la actividad de la empresa, es un indicador muy significativo, porque muestra la evolución de facturación de la empresa a través de los años.

La figura 2.1 y la tabla de resultados 2.1, presentan los puntajes y la gráfica donde un valor alto indica cuando la situación de la empresa es saludable, es decir, que ésta no tendrá problemas de solvencia en el corto plazo. Un resultado intermedio indica incertidumbre o alerta, es decir, que la empresa está en observación o prueba, un resultado intermedio inclinado hacia un resultado bajo, implica decir que la empresa está en una zona muy alerta por los próximos dos años y un resultado bajo nos indica que la empresa probablemente entrará en quiebra y es muy posible que no pueda recuperarse financieramente.



Figura 2.1: Gráfica de Zonas Riesgo, Modelo Altman.

Dentro de las ventajas de este modelo es que permite determinar de forma sencilla la situación general que presenta una empresa en un determinado momento, con información clave de la empresa, además puede analizarse la tendencia de su situación para tomar medidas correctivas en forma oportuna.

Tabla 2.1: Tabla de Resultados

Puntaje	Probabilidad de quiebra
1.1 o menos	Muy alta (zona de insolvencia)
De 1.2 a 2.6	Posible (zona de precaución)
De 2.7 a mayor	Baja (zona segura)

Entre las desventajas en la economía de países latinoamericanos, la calificación que se obtiene del modelo está basada en el desempeño operativo y financiero, pero, ante la incertidumbre, la operación de la empresa está sujeta a cambios más bruscos en aspectos como el riesgo cambiario, las diferencias en los registros contables y la intervención gubernamental.

### 2.1.3 Riesgo de Quiebra o Insolvencia.

Sobre los componentes o áreas a evaluar sobre el desempeño de la empresa para prevenir el riesgo del fracaso empresarial. El modelo presentado en [10], considera datos de los factores de riesgo cualitativos mediante decisiones reales de los expertos recopiladas de la base de datos de gestión de préstamos de un Banco de Corea.

Se realiza un análisis con un benchmark del riesgo más básico que es la insolvencia que deriva en la quiebra empresarial. En esta metodología, se utiliza una muestra recolectada durante 2001–2002 de uno de los bancos comerciales más grandes de Corea. Esta muestra consta de 250 compañías

de manufactura y servicios, 143 están clasificadas en el grupo de quiebra y 107 en el grupo de No-quiebra. Los expertos que calificaron estas compañías son oficiales de crédito experimentados que trabajan para el banco comercial. Cuentan con 9 años en promedio de experiencia en esta área. Realizaron una evaluación de seis dimensiones o áreas de riesgo cualitativos y asignaron un nivel apropiado a cada una de las seis áreas de riesgo cualitativos: riesgo de la industria, riesgo administrativo, flexibilidad financiera, credibilidad de la empresa, competitividad y riesgo operativo.

Así, se presenta un modelo cualitativo que considera las diferentes dimensiones que intervienen y que son parte importante en el nivel de riesgo de permanencia de la empresa, éstas son:

- Riesgo de la Industria
- Riesgo Administrativo
- Flexibilidad Financiera
- Competitividad
- Cumplimiento
- Riesgo Operativo

Cada una de estas dimensiones contiene elementos vulnerables que pueden ser sus procesos, la maquinaria y/o herramientas tecnológicas, las personas que realizan las funciones, etc., éstos son llamados *factores de riesgo*, en la Tabla 2.2 se muestran estos factores agrupados por las dimensiones del modelo.

Tabla 2.2 Factores de Riesgo por Dimensión del Modelo.

Dimensiones	Factores de Riesgo
Riesgo de la Industria	Políticas gubernamentales y acuerdos internacionales
	Ciclicidad. Las ganancias fluctúan ampliamente (ej. por temporada)
	Grado de competencia

	Estabilidad en el mercado de suministros y sus costos
	Tamaño y crecimiento del mercado (demanda de su producto/servicio)
	Sensibilidad a los cambios en los factores macroeconómicos
	Poder competitivo nacional e internacional
	Ciclo de vida de producto
Riesgo Administrativo	Capacidad y competencia en la gestión (administración)
	Estabilidad en el área administrativa
	Relación entre la gerencia / propietario
	Administración de recursos humanos
	Crecimiento / desempeño comercial
	Planificación de negocios a corto y largo plazo
	Logros y viabilidad
Flexibilidad Financiera	Financiamiento Directo
	Financiamiento Indirecto
	Otras fuentes de financiamiento
Cumplimiento	Historial Crediticio
	Información confiable
	Relación con instituciones financieras
Competitividad	Posición en el mercado
	Nivel de Capacidad en la actividad principal (core business)
	Estrategia diferenciada
Riesgo Operacional	Estabilidad y diversidad de las adquisiciones (compras)
	Estabilidad de las transacciones
	Eficiencia en la producción
	Perspectivas de la demanda de productos y servicios
	Diversificación de ventas
	Precio de venta y condición de liquidación
	Cartera vencida (clientes y/o proveedores)
	Eficacia de la red de venta

## 2.2. Respuesta a los factores de riesgo

El tratamiento de riesgos es el proceso que consiste en seleccionar y aplicar medidas para modificar los factores de riesgo. El tratamiento de riesgos incluye, como principal elemento, el control o mitigación del riesgo, pero también se extiende más allá, por ejemplo, a evitar los riesgos, o a su transferencia.

Cualquier sistema de tratamiento de riesgos debe proporcionar como mínimo:

- Un funcionamiento efectivo y eficiente de la organización.
- Controles internos efectivos.
- Conformidad con las leyes y reglamentos.

Una vez identificados aquellos factores que desencadenan riesgos, la dirección, deberá priorizar acciones de control de riesgos en función de su potencial para beneficiar a la empresa. La efectividad del control interno constituye el grado en que el riesgo será eliminado o reducido mediante las medidas de control propuestas. La rentabilidad de los controles internos está relacionada con el coste del control, comparado con los beneficios esperados de la reducción de riesgos. Los controles propuestos tienen que medirse según el posible efecto económico en caso de que no se tome ninguna acción, comparado con el coste de la(s) acción(es) propuesta(s) y necesariamente requieren más información detallada e hipótesis que las que están disponibles inmediatamente.

En primer lugar, hay que fijar el coste de la puesta en práctica. Este debe calcularse con precisión, ya que se convertirá rápidamente en la referencia con la que se medirá la rentabilidad. También hay que calcular la pérdida que se debe prever si no se toma ninguna medida, y, al comparar los resultados, la dirección puede decidir si poner o no en práctica las medidas de control de riesgos. Ocasionalmente se dispone de cierta flexibilidad cuando el coste de reducción de un riesgo es totalmente desproporcionado con relación a ese riesgo.

### 2.2.1 Tipos de salvaguardas (preventivas, correctivas, de control)

El término 'Salvaguarda' lo encontramos comúnmente en el lenguaje de las instituciones financieras [60-62] (como el del Banco Mundial), el cual se refiere a 'salvaguardas' como las "medidas para anticipar, minimizar, mitigar o tratar de otro modo los impactos adversos asociados a una actividad dada."

Por lo tanto, las salvaguardas tienen un 'enfoque basado en los riesgos', que supone implementar acciones para 'gestionar los riesgos' de una manera económicamente eficiente.

### 2.3. Escalas de Valoración del Riesgo

Establecer criterios de medición del riesgo consiste en definir los criterios que permitan conocer la postura de la organización en cuanto a su propensión a los riesgos. Se trata de la base para la evaluación, ya que sin esta actividad no se puede medir el grado en el que la organización se ve afectada cuando se materializa una amenaza.

#### 2.3.1. Metodología PMI

Project Management Institute, PMI, de la Universidad Carnegie Mellon, ha diseñado una metodología reconocida internacionalmente como PMBoK (Project Management Body of Knowledge) la cual contempla herramientas y técnicas para el análisis de riesgos que son presentados en la tabla 2.3.

Tabla 2.3 Técnicas para análisis de riesgo del PMBoK.

Evaluación de Probabilidad e Impacto de los Riesgos	Investiga la probabilidad de ocurrencia de cada riesgo específico, investiga el posible efecto sobre un objetivo determinado, como tiempo, coste, alcance o calidad, incluidos tanto los efectos negativos por las amenazas que implican, como los efectos positivos por las oportunidades que generan.
---	---

	<p>El nivel de probabilidad de cada riesgo y su impacto sobre cada objetivo se evalúa durante la entrevista o reunión. Los detalles explicativos, incluidos los motivos que justifican los niveles asignados, también se registran. Las probabilidades y los impactos de los riesgos se califican de acuerdo con las definiciones que se dan en el plan de gestión de riesgos</p>
Matriz de probabilidad e impacto	<p>La evaluación de la importancia de cada riesgo y, por consiguiente, de su prioridad, generalmente se realiza usando una tabla de búsqueda o una matriz de probabilidad e impacto. Esta matriz especifica combinaciones de probabilidad e impacto que llevan a la calificación de los riesgos como de prioridad baja, moderada o alta. Pueden usarse términos descriptivos o valores numéricos, dependiendo de la preferencia de la organización</p>
Evaluación de la Calidad de los Datos sobre Riesgos	<p>Un análisis cualitativo de riesgos requiere datos exactos y sin sesgos para que sea creíble. El análisis de la calidad de los datos sobre riesgos es una técnica para evaluar el grado de utilidad de los datos sobre los riesgos para su gestión. Implica examinar el grado de entendimiento del riesgo, y la exactitud, calidad, fiabilidad e integridad de los datos sobre el riesgo.</p>
Categorización de Riesgos	<p>Los riesgos pueden categorizarse por fuentes de riesgo, área afectada u otra categoría útil para determinar las áreas que están más expuestas a los efectos de la incertidumbre. Agrupar los riesgos por causas comunes puede contribuir a desarrollar respuestas efectivas a los riesgos.</p>



Las estrategias para el tratamiento de los riesgos negativos o amenazas se presentan en cuatro acciones que son: evitar, transferir, mitigar y aceptar.

Evitar: Es una estrategia de respuesta a los riesgos según la cual se actúa para eliminar la amenaza o para proteger del impacto. Por lo general implica realizar modificaciones, a fin de eliminar por completo la amenaza. Evitar cualquier contacto con la causa que pueda generar la amenaza.

Transferir: Trasladar el impacto negativo del riesgo hacia un tercero. Se asigna cuando se cuenta con un tercero que asume las tareas correctivas como consecuencia del daño ocasionado por la amenaza.

Mitigar: Disminuir la probabilidad de ocurrencia y/o el impacto. Son atenuados y son aquellos que establecen criterios y medidas de seguridad preventivas y correctivas para reducir el daño causado ante la exposición del riesgo.

Aceptar: No realizar cambios. Una aceptación activa consiste en dejar establecida una política de cómo actuar en caso de que ocurra el evento negativo. Los aceptados se dan cuando el costo de la solución es más costoso que el daño y la exposición al riesgo mismo.

### 2.3.2. COSO

Metodología del COSO (Committee of Sponsoring Organizations of the Treadway) desarrollada en principios y lineamientos de ISO31000. Surge en 1985 con el objeto de ayudar a las organizaciones a evaluar y mejorar sus sistemas de control interno. La definición del Riesgo empresarial

*“... es un proceso, efectuado por la junta de directores, gerentes y el resto del personal, que impactará a la estrategia de toda la empresa, diseñado para identificar eventos potenciales que pueden afectar a la entidad, y gestionar el riesgo dentro de un margen que se está dispuesto a aceptar, para garantizar el logro de los objetivos de la empresa”. COSO*

Presenta técnicas y principios de la administración del riesgo, que pueden ser usadas dentro de la organización a varios niveles. Comenta que es útil para aquellos que buscan como aplicar las técnicas de la administración del riesgo empresarial. Las componentes de dicha gestión se muestran en la figura 2.2, en cada rectángulo de la gráfica se presentan los elementos claves de ese componente.

Iniciando con el ambiente interno con la filosofía de la administración del riesgo, el tono del riesgo, como se mostró en la figura 1.2, los valores, el compromiso, la estructura de la organización, asignación de las responsabilidades y los estándares de los recursos humanos. En la siguiente componente, se consideran el conjunto de objetivos dentro de los cuales están los estratégicos, los relacionados entre las diversas áreas de la empresa, los seleccionados a corto y mediano plazo, la tolerancia ante el riesgo y el apetito del mismo, elementos que impactarán posteriormente en la asignación de presupuestos para atender los factores de riesgo. Sobre la identificación de eventos se contempla determinar la influencia de los factores, la interdependencia de los eventos, realizar su categorización y distinguir claramente los riesgos entre las oportunidades. Posteriormente se ingresa a las acciones para valorar el riesgo, determinando riesgos inherentes y residuales, estableciendo impactos y probabilidades, ir a las fuentes de datos y aplicar las técnicas de identificación y la relación de los eventos. En la componente de la respuesta al riesgo, los elementos son evaluar las respuestas al riesgo, seleccionarlas y presentarlas como un portafolio. Las actividades de control, información - comunicación y monitoreo son acciones de seguimiento encaminadas a la mejora continua.



Figura 2.2 Elementos Claves de cada Componente sobre la gestión del riesgo. Fuente COSO.

Así, esta metodología determina que el valor de la empresa es una función del riesgo y el rendimiento ya que cada decisión aumenta, preserva o erosiona el valor. Dado que el riesgo es parte integral de la búsqueda del valor, las empresas con mentalidad estratégica no se esfuerzan por eliminar el riesgo o incluso por minimizarlo, por el contrario, buscan gestionar las exposiciones al riesgo en todas las partes de sus organizaciones para que, en cualquier momento, incurran en los tipos de riesgo adecuados, ni más ni menos, para perseguir eficazmente objetivos estratégicos. Este es el “punto óptimo” de toma de riesgo, a la que hace referencia la figura 2.3, ya que para incrementar el valor empresarial esperado, la zona ideal representa un cambio crítico de la visión tradicional del riesgo como algo a evitar.

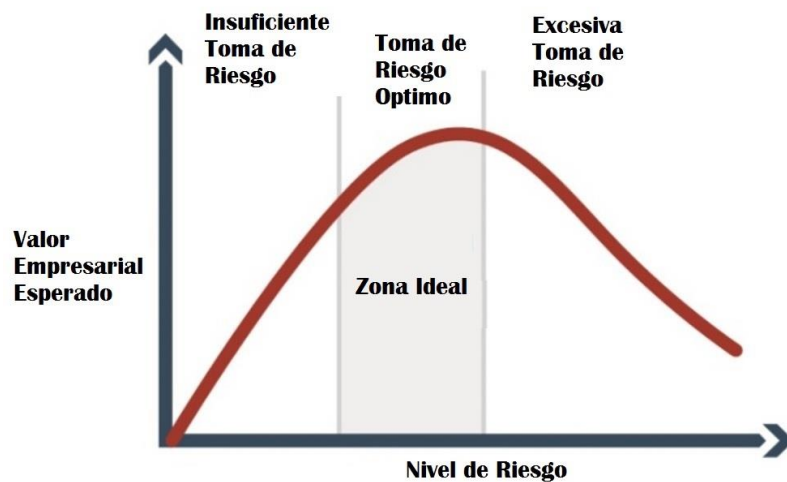


Figura 2.3 Óptima Toma de Riesgo. Fuente COSO.

Dentro de las estrategias y técnicas del COSO para la identificación y tratamiento del riesgo empresarial, el riesgo se considera: 1) por forma de Gobierno: la manera en que se marca los lineamientos de control de operaciones, código ético, etc.; 2) por Operaciones e Infraestructura; 3) por Finanzas: como el manejo de portafolio de inversión, acciones, etc.; 4) de Cumplimiento: regulaciones fiscales, de medio ambiente, etc. y 5) por Estrategias: con clientes, proveedores, empleados, etc.

De manera que, todo lo que amenace el valor para la empresa, que según el COSO es: lo que impida el crecimiento de ganancias, lo que limite el margen operativo y lo que interfiera en la eficiencia de activos, todo esto es lo que genera los principales riesgos empresariales.

### 2.3.3. Método William Fine. Riesgo Matemático

Para facilitar el control expedito del riesgo para la prevención de accidentes, se han reconocido dos grandes necesidades [17]. La primera es (1) un método para determinar la seriedad relativa de todos los peligros para contar con una guía en la asignación de prioridades para el esfuerzo preventivo; y (2) un método para dar una determinación definitiva sobre si el costo estimado de la acción correctiva contemplada para eliminar el peligro está justificado.

Para satisfacer estas necesidades, se ha diseñado una fórmula, “*GP*”, que sopesa los factores de control y “calcula el riesgo” de una situación peligrosa, dando una evaluación numérica a la urgencia de una atención remediadora del peligro. Los puntajes de riesgo calculados se usan luego para establecer prioridades para el esfuerzo correctivo. Una fórmula adicional, “*J*”, pondera el costo estimado y la efectividad de cualquier acción correctiva contemplada en relación con el puntaje de riesgo y proporciona una determinación en cuanto a si el costo está justificado.

#### **Se recomienda el siguiente procedimiento:**

Con las tablas de escalas (ajustadas del riesgo matemático de William Fine, tablas 2.4 y 2.5), se determina la calificación del grado de riesgo y la justificación de las salvaguardas. Una vez que se ha reconocido una amenaza o factor de riesgo, se recomienda el siguiente procedimiento:

- a. Exponga el factor de riesgo brevemente.
- b. Decida las consecuencias más probables debido al riesgo en la dimensión correspondiente.

- c. Revise todos los elementos cuidadosamente, en la escena. Enumere la secuencia de eventos paso a paso real que probablemente resulte en las consecuencias elegidas. Debe ser específico.
- d. Decida sobre la acción correctiva más apropiada y obtener o hacer una estimación aproximada de su costo.
- e. Considere cuidadosamente el efecto de la acción correctiva propuesta y el riesgo, y calcule de forma aproximada el grado en que se mitigará la situación peligrosa.
- f. Si son posibles medidas correctivas alternativas, repita los pasos (d) y (e) para ellas.
- g. Seleccione el evento de riesgo (la primera ocurrencia no deseada que pudiera iniciar la secuencia).
- h. Considere cuidadosamente la situación existente para determinar la frecuencia de la ocurrencia del evento de riesgo, mediante la observación de la escena y luego decida sobre la Clasificación de exposición, tabla 2.4. Si tiene dudas entre dos clasificaciones, interpole.
- i. Para la Calificación de Probabilidad, considere la probabilidad de que ocurra cada evento de la secuencia del problema, incluyendo la pérdida y/o daño resultante, y forme una opinión basada en las palabras descriptivas, tabla 2.4. Por ejemplo, si se requieren dos coincidencias “inusuales”, esto podría considerarse “raramente”; dos ocurrencias “raramente” podrían ser “ocasionalmente”, etc. Si tiene dudas entre dos clasificaciones, puede realizar interpolación. Considere la ocurrencia de solo las mismas consecuencias que se decidieron en el paso (b) anterior. Por ejemplo, si decide las consecuencias de un gran daño, entonces en este paso solo se puede considerar la probabilidad de una fatalidad. Si también desea considerar consecuencias menores, se debe realizar un cálculo separado y adicional, ya que las evaluaciones de Consecuencias y Probabilidad serían diferentes. Se deben agregar puntuaciones.

j. Ahora se han obtenido calificaciones para todos los elementos de la fórmula “*GP*”. Sustituir en la fórmula, tabla 2.5, y calcular el puntaje de justificación, “*J*”.

k. Si se consideran medidas correctivas alternativas para disminuir el riesgo, calcule también sus puntajes de justificación.

l. Si hay medidas correctivas alternativas que tienen puntajes de justificación aceptables, la más deseable desde el punto de vista de la seguridad es la que haría la mayor reducción en el puntaje de riesgo. Por lo tanto, para cada alternativa, suponga que las medidas correctivas están en vigencia y vuelva a calcular el puntaje de riesgo. Por supuesto, esta selección también puede verse afectada por consideraciones externas (Inseguridad) tales como el tamaño de la inversión requerida, los efectos relativos sobre la moral, la estética, la eficiencia, la comodidad, la facilidad de implementación, etc.

**Excepciones en la justificación de la salvaguarda.** Puede existir una situación de alto riesgo de la que no se pueda idear ninguna acción correctiva que dé un puntaje de justificación aceptable. Obviamente, en tal caso, cualquier acción correctiva necesaria para reducir el puntaje de riesgo debe tomarse independientemente del puntaje de justificación.

Tabla 2.4 Escala del Método W. Fine para valorar el grado de riesgo.

<b>Consecuencias (C)</b>	
<b>Valor</b>	<b>Descripción</b>
25	Muerte o daño superior a 5 nóminas mensuales
15	Lesiones incapacitantes permanentes y/o daño entre 1 y 5 nóminas mensuales
5	Lesiones con incapacidades no permanentes y/o daños entre 10 y 100% de la nómina mensual
1	Lesiones con heridas leves, contusiones, golpes y/o daños menores del 10% de la nómina mensual
<b>Probabilidad (P)</b>	
<b>Valor</b>	<b>Descripción</b>
10	Es el resultado más probable y esperado si la situación de riesgo tiene lugar
6	Es completamente posible, nada extraño. Probabilidad de ocurrencia del 50%
3	Sería una coincidencia rara. Probabilidad de ocurrencia del 20%
1	Nunca ha sucedido en muchos años de exposición al riesgo, pero es concebible. Probabilidad del 5%
<b>Exposición (E)</b>	
<b>Valor</b>	<b>Descripción</b>
10	La situación de riesgo ocurre continuamente o muchas veces al día
6	Frecuentemente o una vez al día
3	Ocasionalmente o una vez por semana
1	Remotamente posible



Tabla 2.5 Escala del W. Fine para cálculo del beneficio de la salvaguarda.

<b>Justificación de la Salvaguarda (J)</b>		
<b>Descripción</b>	<b>Nivel</b>	<b>Valor</b>
GC Grado de Corrección. Grado en el que será reducido el riesgo.	Eliminado 100%	1
	Reducido al menos 75%	2
	Reducido del 50% al 75%	3
	Reducido del 25% al 50%	4
	Ligero efecto sobre el riesgo (menos del 25%)	6
FC Factor de Coste. Coste estimado de la salvaguarda propuesta	Superior a X nóminas mensuales	10
	Entre 1 y X nóminas mensuales	6
	Entre 50% al 100% de la nómina mensual	4
	Entre del 10% al 50% de la nómina mensual	2
	Menos del 10% de la nómina mensual	1

El valor del Riesgo “GP” se calcula con la fórmula  $GP=C \times E \times P$  donde  $C$  es valor de la escala de la Consecuencia,  $E$  es el valor de la Exposición al riesgo, y  $P$  es el valor de la probabilidad de ocurrencia. Para las salvaguardas, el valor de la justificación, “J” del beneficio se calcula de  $J=GP/(GC \times FC)$ .

## 2.4. Modelo de Optimización

### 2.4.1. Optimización

Los problemas de optimización tienen gran relevancia práctica en muchas áreas donde se requiere de un método que garantice un resultado casi inmejorable. La optimización se presenta en dos clases. La primera es la continua y la segunda es la combinatoria o discreta. Los problemas combinatorios se caracterizan por tener un número muy grande de posibles

soluciones y, en general, la mayoría de este tipo de problemas encontrados en la vida real están considerados como muy difíciles de resolver, es decir, son del tipo NP [63,64]. En principio, es posible hacer una enumeración de todas las combinaciones que integran al espacio de búsqueda para encontrar la solución que cumpla con las restricciones del problema. Sin embargo, el número de posibles combinaciones, aunque finito, es muy grande y no es práctico debido al tiempo que se puede tardar.

Por otra parte, en cuanto a las técnicas de la IA (inteligencia artificial), el paradigma principal para la resolución de problemas de optimización está en la búsqueda heurística [34,35,65,66]. Este es un método muy general que se puede aplicar a muchos tipos de problemas difíciles con el fin de encontrar una solución cercana a la óptima global en un tiempo prudente, pero sin garantizar que ésta sea la solución óptima (global). Su objetivo es que casi siempre se encuentre una buena solución (óptimo local).

De manera simple, un algoritmo es un programa que resuelve un problema, está identificado generalmente con un programa computacional. Dependiendo de la dificultad inherente del problema, la solución requerirá de una cierta cantidad de recursos computacionales, así los problemas catalogados como NP son tratados con técnicas combinatorias y generando soluciones aproximadas. Todos los algoritmos exactos conocidos para problemas NP-completos utilizan tiempo exponencial con respecto al tamaño de la entrada. Ahora, para encontrar soluciones cercanas a la óptima de un problema NP-completo de tamaño arbitrario, se utiliza uno de los enfoques de la tabla 2.6.

Tabla 2.6 Enfoques de solución de problemas computacionales.

Aproximación:	Un algoritmo que rápidamente encuentra una solución no necesariamente óptima, pero dentro de un cierto rango de error. No todos los problemas NP-completos tienen algoritmos de aproximación.
Probabilístico:	Mediante aleatoriedad se obtiene en promedio una buena solución al problema planteado con una pequeña probabilidad de fallar, para una distribución de los datos de entrada dada.
Restricciones:	Se encuentra un algoritmo más rápido al restringir la estructura de las entradas. Los casos particulares se presentan en problemas en los que existen soluciones rápidas solo para algunos casos.
Algoritmo inspirado en la naturaleza:	Algoritmo que mejora las posibles soluciones hasta encontrar una que posiblemente esté cerca del óptimo. Utiliza analogías de sistemas derivados de la física, biología y comportamiento social.
Heurísticas:	Un algoritmo que trabaja razonablemente bien en muchos casos. En general son rápidos, pero no existe medida de la calidad de la respuesta.

#### 2.4.2. Formulación Matemática del Modelo de Optimización

La representación de la selección de salvaguardas genera un problema de optimización binario (0/1) [67] con un vector  $V_a$  de factores de riesgo y una matriz  $M_{a \times s}$  salvaguardas, donde  $a$  es el número de factores y  $s$  es el número de salvaguardas posibles para cada factor. Cada salvaguarda tiene un valor  $J$  el cual indica la justificación de la reducción del riesgo de la amenaza  $i$  usando la salvaguarda tipo  $k$ , con un costo de implementación  $CS$ . El valor  $CS$  y  $J$ , son obtenidos al aplicar las tablas 2.4 y 2.5 de William Fin

### Modelo de Optimización del problema

$$\min f(x) = \min \left[ \sum_{i=1}^a \sum_{k=1}^s \left( \frac{CS_{i,k}}{J_{i,k}} \right) \times x_{i,k} \right] \quad (1)$$

$$\text{sujeto a} \quad \sum_{i=1}^a \sum_{k=1}^s CS_{i,k} x_{i,k} \leq P \quad (2)$$

$$\sum_{k=1}^s x_{i,k} \leq 1 \text{ para } i = \{1, \dots, a\} \quad (3)$$

$$\text{donde} \quad x_{i,k} = \begin{cases} 1 & \text{si es seleccionada } S_{i,k} \\ 0 & \text{de otra manera} \end{cases} \quad (4)$$

La función objetivo,  $FO$ , en (1) representa el total del cociente costo-beneficio por la selección de las salvaguardas, la restricción en (2) es de no exceder el presupuesto  $P$  y el conjunto de restricciones en (3) son restricciones de que solo una salvaguarda puede ser seleccionada para cada factor de riesgo. En (4), la restricción es que la variable  $x_{i,k}$  es un indicador de la selección de la salvaguarda  $k$  que atiende al factor de riesgo  $i$ , esto es si  $x_{i,k}=1$ . Si  $x_{i,k}=0$  indica que no ha sido seleccionada esa salvaguarda. Se asume que  $J_{i,j} > 0$  y  $CS \geq 0$  para  $i=1, \dots, a$  y  $j=1, \dots, s$ .

Integrando la algorítmica y los procesos de optimización, Goldberg [68] menciona que un aspecto importante para determinar la eficiencia y la precisión de un algoritmo de optimización, es el balance entre exploración y explotación. La exploración es la capacidad de un algoritmo de búsqueda para examinar diferentes regiones del espacio de búsqueda con el fin de encontrar una buena solución, un óptimo, mientras que la explotación es la capacidad de un algoritmo para concentrar la búsqueda en una región prometedora con el fin de refinar una solución posible. Un buen algoritmo para problemas de optimización debe de conjugar, de la mejor manera, estos dos objetivos [69].

De manera que las heurísticas computacionales y los procesos de optimización proporcionan a las herramientas tecnológicas, una gran capacidad para identificar buenas soluciones, en nuestro caso, sobre el tratamiento de los riesgos empresariales, ofreciendo así una solución de gran valor para la elaboración de planes de protección específicos.

## 2.5. Heurísticas

El concepto de heurística es difícil de comprender. Newell, Shaw y Simon [70,71] en 1963 dieron la siguiente definición: “Un proceso que puede resolver un problema dado, pero que no ofrece ninguna garantía de que lo hará, se llama una heurística para ese problema”

### 2.5.1 Heurísticas de Búsqueda Local

Los métodos de búsqueda heurística disponen de alguna información sobre la proximidad de cada estado a un estado objetivo, lo que permite explorar en primer lugar los caminos más prometedores.

Son características de los métodos heurísticos:

- No garantizan que se encuentre una solución, aunque existan soluciones.
- Si encuentran una solución, no se asegura que ésta tenga las mejores propiedades (que sea de longitud mínima o de coste óptimo).
- En algunas ocasiones (que, en general, no se podrán determinar a priori), encontrarán una solución (aceptablemente buena) en un tiempo razonable.

### 2.5.2. Heurística Recocido Simulado

El recocido simulado, RS es un algoritmo de búsqueda local. Debido a su sencilla implementación, las propiedades de convergencia y la capacidad de escape de óptimos locales, se ha convertido en una metaheurística popular en las últimas décadas. Es llamado así por la analogía en el proceso físico del recocido de sólidos, en el cual la estructura cristalina sólida es calentada y luego se lleva a un proceso lento de enfriamiento hasta alcanzar el estado estable, es decir, a un estado mínimo de energía de manera que queda libre de defectos en la cristalización. RS se inspira en este tipo de comportamiento termodinámico en la búsqueda del óptimo para problemas de optimización discreta [65,66,72].

En cada iteración de la metaheurística RS aplicada al problema de optimización, la función objetivo se valora con dos soluciones, la solución actual y una solución nueva generada en la vecindad, las cuales son comparadas. Las mejores soluciones siempre se aceptan, mientras que solo una fracción de soluciones “no buenas” son aceptadas con el propósito de escapar de óptimos locales para continuar la búsqueda de mejores soluciones. La probabilidad de aceptar soluciones “no buenas” depende del parámetro T (temperatura) el cual decrece en cada iteración del algoritmo.

Un elemento clave del RS es que provee una manera de escape del óptimo local al permitir movimientos malos (i.e. se mueve a soluciones que generan un peor valor en la función objetivo). Al decrecer la temperatura hasta cero, los peores movimientos ocurren con menor frecuencia, y la distribución de la solución asociada con la cadena de Markov no homogénea que modela el comportamiento del algoritmo, converge a una distribución en la que toda la probabilidad se concentra en el conjunto de soluciones óptimas, lo que significa que el algoritmo es asintóticamente convergente.

Es necesario especificar un conjunto de parámetros que determinen la convergencia del algoritmo [73, 74].

$\omega_0$	Solución inicial del problema
T	Parámetro de control del RS
$T_0$	Valor inicial del parámetro de control $T_0=2 \times \sigma$
$\alpha$	Coefficiente de control. $0 < \alpha < 1$
M	Tamaño de la Cadena de Markov
f	Función de costo
$\Omega$	Estructura de Vecindad
$\beta$	Coefficiente de Criterio de Paro. Mínimo valor de $T = \beta \times T_0$

Para describir formalmente el algoritmo de recocido simulado, se necesitan algunas definiciones. Sea  $\Omega$  el espacio solución; se define  $\eta(\omega)$  como la función de vecindad para  $\omega \in \Omega$ . Sea una solución inicial  $\omega \in \Omega$ . Una solución vecina  $\omega' \in \eta(\omega)$  se genera al azar en la mayoría de los casos. El recocido simulado se basa en el criterio de aceptación de Metrópolis, que modela cómo el sistema se mueve de la solución actual  $\omega \in \Omega$  a una solución candidata  $\omega' \in \eta(\omega)$ , en el que la energía está siendo minimizada. La solución candidata,  $\omega'$  es aceptada como la solución actual basada en la probabilidad de aceptación

$$P\{\omega'\} = \begin{cases} \exp((f(\omega') - f(\omega))/t_k) & \text{si } f(\omega') - f(\omega) < 0 \\ 1 & \text{si } f(\omega') - f(\omega) \geq 0 \end{cases}$$

$t_k$  se define como la temperatura en la iteración  $k$  tal que  $t_k > 0$  para todo  $k$  y  $\lim_{k \rightarrow \infty} t_k = 0$

El algoritmo de recocido simulado se describe a continuación en pseudocódigo:

1. Seleccionar una solución inicial  $\omega = (x_1, \dots, x_n) \in \Omega$ ; una temperatura inicial  $t = t_0$ ; un parámetro de control  $\alpha$ ; y la temperatura final  $e$ ; un programa

de repetición,  $M$  que define el número de iteraciones ejecutadas a cada temperatura

2. solución  $\leftarrow f(\omega)$
3. Repite
4. Inicializar contador  $m = 0$ ;
5. Repite
6. Selecciona un entero aleatorio  $i$  del conjunto  $\{1, 2, \dots, n\}$ ;
7. Si  $x_i = 0$ , selecciona elemento  $i$ , i.e.  $x_i = 1$ , obtener una nueva solución  $\omega'$ ,
8. Mientras solución  $\omega'$  es no-factible haz
9. Selecciona aleatoriamente otro elemento; nueva solución  $\omega'$ ;
10. Calcula  $\Delta = f(\omega_0) - f(\omega)$ ;
11. Mientras  $\Delta \geq 0$  o  $\text{Random}(0, 1) < e^{-\Delta t}$ , haz  $\omega \leftarrow \omega'$ ;
12. otro
13. Selecciona elemento  $i$ , y otro elemento aleatoriamente, obtener  $\omega'$ ;
14. Calcula  $\Delta = f(\omega_0) - f(\omega)$ ;
15. Mientras  $\Delta \geq 0$  o  $\text{Random}(0, 1) < e^{-\Delta t}$ , haz  $\omega \leftarrow \omega'$ ;
16. termina si
17. si solución  $< f(\omega)$  entonces solución  $\leftarrow f(\omega)$ ;
18.  $m = m + 1$ ;
19. Hasta  $m = M$
20. Calcula  $t = a * t$ ;
21. Hasta  $t < e$ .

## 2.6. Aprendizaje Automático (Machine Learning)

Las técnicas de Machine Learning consisten básicamente en automatizar, mediante distintos algoritmos, la identificación de patrones o tendencias que se “escondan” en los datos. Por ello, la elección del algoritmo no es fácil, es muy importante no sólo la elección del algoritmo más adecuado (y su posterior parametrización para cada problemática concreta), sino también el hecho de disponer de un gran volumen de datos de suficiente calidad [75–79].



**Aprendizaje Automático.** El aprendizaje automático o aprendizaje automatizado o aprendizaje de máquinas (del inglés, machine learning) es el subcampo de las ciencias de la computación y una rama de la inteligencia artificial, cuyo objetivo es desarrollar técnicas que permitan que un agente (sistema, dispositivo, programa, etc.) aprende cuando su desempeño mejora con la experiencia [80–82]. En muchas ocasiones el campo de actuación del aprendizaje automático se solapa con el de la estadística inferencial, ya que las dos disciplinas se basan en el análisis de datos.

**Aprendizaje no supervisado.** Tiene lugar cuando no se dispone de datos “etiquetados” para el entrenamiento. Sólo conocemos los datos de entrada, pero no existen datos de salida que correspondan a un determinado *input*. Por tanto, sólo podemos describir la estructura de los datos, para intentar encontrar algún tipo de organización que simplifique el análisis. Por ello, tienen un carácter exploratorio. El problema de agrupación de objetos de acuerdo a sus atributos es ampliamente utilizado debido a sus aplicaciones en áreas como el aprendizaje de máquina, minería de datos y descubrimiento de conocimiento, reconocimiento de patrones. El objetivo de la clasificación es particionar el conjunto de objetos de manera que los patrones en cada grupo sean similares en los atributos que son vectores multidimensionales asociados.

#### 2.6.1. K-means

Es un algoritmo de aprendizaje no supervisado que particiona el conjunto de objetos en un determinado número de grupos de acuerdo con un criterio de optimización [83-85]. Su implementación es relativamente simple, cabe mencionar que la convergencia a una agrupación óptima no está garantizada y para problemas grandes, se requiere de un mayor número de iteraciones. Para determinar el número de grupos o clústeres, se utiliza el algoritmo de Elbow [86].

**Descripción del algoritmo K-means estándar.** De acuerdo con la literatura especializada [87–90] se pueden identificar cuatro pasos en el algoritmo:

Paso 1. Inicialización: Se definen un conjunto de objetos a particionar, el número de grupos y un centroide por cada grupo. Algunas implementaciones del algoritmo estándar determinan los centroides iniciales de forma aleatoria; mientras que algunos otros procesan los datos y determinan los centroides mediante cálculos.

Paso 2. Clasificación: Para cada objeto de la base de datos, se calcula su distancia a cada centroide, se determina el centroide más cercano, y el objeto es incorporado al grupo relacionado con ese centroide.

Paso 3. Cálculo de centroides: Para cada grupo generado en el paso anterior se vuelve a calcular su centroide.

Paso 4. Condición de convergencia: Se han usado varias condiciones de convergencia, de las cuales las más utilizadas son las siguientes: converger cuando alcanza un número de iteraciones dado, converger cuando no existe un intercambio de objetos entre los grupos, o converger cuando la diferencia entre los centroides de dos iteraciones consecutivas es más pequeña que un umbral dado. Si la condición de convergencia no se satisface, se repiten los pasos dos, tres y cuatro del algoritmo.

Se ha identificado como una de las clasificaciones más populares [91]. Supongamos que el número de grupos sea  $K=3$ , la formulación matemática es un conjunto de datos  $n$  o instancias en el vector  $d$ -dimensional  $D=\{x_1, \dots, x_n\} \in R^d$ , donde el problema es determinar un conjunto de 3 centroides  $C=\{c_1, c_2, c_3\} \in R^d$  donde se minimice la siguiente función del error.

$$E(C) = \sum_{x \in D} \min_{k=1,2,3} \|x - C_k\|^2$$

Consiste en encontrar la partición de las  $n$  instancias en 3 grupos cuya distancia al centro sea la mínima. La estadística  $F$  se calcula como el cociente de medias de cuadrados de la suma de cuadrados entre los grupos, dividida por la suma de sumas de cuadrados dentro del grupo,  $F > 1$  es cuando los grupos están suficientemente diferenciados entre sí.

**Aprendizaje supervisado:** En el aprendizaje supervisado, los algoritmos trabajan con datos “etiquetados” (labeled data), intentando encontrar una función que, dadas las variables de entrada (input data), les asigne la etiqueta de salida adecuada. El algoritmo se entrena con un “histórico” de datos y así “aprende” a asignar la etiqueta de salida adecuada a un nuevo valor, es decir, predice el valor de salida.

### 2.6.2. Árboles de decisión

El aprendizaje supervisado de árboles de decisión es una de las técnicas de inferencia inductiva más usadas [92,93]. Se trata de un método para aproximar funciones de valores discretos que sea capaz de expresar hipótesis disyuntivas y que presenten capacidad en la discriminación del ruido de los datos en los ejemplos de entrenamiento.

Cada nodo del árbol está conformado por un atributo y puede verse como la pregunta: ¿Qué valor tiene este atributo en el caso a clasificar? Las ramas que salen de los nodos corresponden a los posibles valores del atributo correspondiente. Un árbol de decisión clasifica a un caso, filtrándolo de manera descendente, hasta encontrar una hoja, que corresponde a la clasificación buscada.

De manera simple, los árboles de decisión permiten representar en forma gráfica una serie de reglas sobre la decisión que debe tomarse en la asignación de un valor de salida a un registro dado [94-98], se usan comúnmente en el análisis para toma de decisión, ayudando a identificar los elementos relevantes para alcanzar una meta. Al crear un árbol completo que ha clasificado todo el conjunto de datos, una medida de "bondad", es decir, qué tan bien los atributos se discriminan entre las clases, se empleó el índice de GINI, [99], esta función mide la impureza "de un atributo con respecto a las clases". Dadas las probabilidades para cada clase  $p_i$ , la función general de GINI, o medida de impureza, es

$$\sum \sum p_i p_j = (\sum p_i)^2 = \sum p_i^2 = 1 - \sum p_i^2$$

## 2.7. Complejidad computacional.

De acuerdo con la complejidad computacional, los problemas se pueden clasificar según el nivel de dificultad que muestran, esto puede ser en el uso de recursos necesarios para resolver el problema implementando un algoritmo [100]. En general, la complejidad computacional se enfoca en los posibles algoritmos que podrían resolver el mismo problema. Realiza una clasificación de problemas que podrían o no ser resueltos con recursos restringidos. En [63] se realiza una explicación completa de los problemas de clasificación de NP.

Computacionalmente, un problema de decisión es un problema lineal entero, donde el número de soluciones factibles con  $n$  variables podría ser  $2^n$ , este es un problema NP Completo [64,101].

Dentro de las herramientas para resolver modelos matemáticos de optimización, se encuentra el CPLEX. Debido a la naturaleza combinatoria de

los programas enteros, para un buen rendimiento con este tipo de problemas se presentan mayores retos que con los problemas de programación lineal o cuadráticos. CPLEX tiene muchos parámetros que permiten a los usuarios personalizar la forma en que opera el algoritmo Branch and Bound de CPLEX [102,103], el cual es de tipo heurístico-determinista y exacto por lo que encuentra soluciones óptimas, pero presenta un comportamiento exponencial en tiempo y/o memoria. Si bien esta variedad de parámetros proporciona muchas formas diferentes de mejorar el rendimiento, un usuario no puede experimentar de manera realista con todas las combinaciones posibles de configuraciones de parámetros. En [32] se hace una recomendación que, antes de tratar de mejorar el rendimiento, primero es necesario localizar el cuello de botella actual del rendimiento. CPLEX proporciona un registro de nodo que muestra el progreso del algoritmo enlazado en un MIP. Por otro lado, la ventaja de las heurísticas no deterministas radica en su simplicidad, esto es lo que las convierte en un método bastante rápido y con un menor uso de memoria, funcionan en un método de dos fases; la primera fase es la construcción, mientras, que la segunda es la búsqueda de mejores soluciones [104].

En algunas ocasiones, con un algoritmo exacto como CPLEX, se puede obtener la solución global óptima de ciertas instancias de problemas. Esto permite realizar un contraste con el algoritmo RS, que encuentra soluciones casi óptimas que dependen de la solución inicial factible. En [105] se prueba que un algoritmo heurístico puede devolver una solución muy buena, casi óptima y sin exceder las capacidades de la computadora.

## 2.8. Errores en la Toma de decisiones

Muchas organizaciones esperan consistencia de sus gerentes o mandos directivos.

Sin embargo, el juicio humano a menudo está influenciado por factores tan irrelevantes como el clima y el último caso visto. Además, las decisiones a menudo varían de empleado a empleado. La variabilidad casual de los juicios se llama ruido, este es sorprendentemente costosa para las empresas.

Decidir es un proceso que implica evaluar distintas variables. La información previa con la que contamos a la hora de decidir nos ayuda a dar una solución en una situación específica, pero el que decide no puede tener la certeza absoluta de que fue la mejor decisión, ya que la información puede ser distorsionada dependiendo del estado de ánimo o de los últimos acontecimientos presentados, como se mencionó en el párrafo anterior.

Aunque existen herramientas formales que nos ayudan a decidir, estas pueden verse afectadas por los “sesgos” de nuestra mente. Son distorsiones inevitables en la manera de evaluar las opciones y realizar la elección. En la toma de decisiones, estos complican una óptima selección de las acciones para atender los riesgos empresariales, Kahneman [106-111].

El sesgo y el ruido son distintos tipos de error. Cada uno viene en diferentes variantes y requiere diferentes acciones correctivas. Kahneman indica que, la solución más radical a problemas de este tipo es reemplazar el juicio humano con algoritmos. Los algoritmos no son difíciles de construir, pero requieren amplia experiencia para diseñar y codificar las políticas y procedimientos. De ahí la importancia de que las empresas establezcan procedimientos automatizados para ayudar a los profesionales a lograr una mayor coherencia.

### **Capítulo 3. Metodología para la Optimización del Tratamiento de Riesgos Empresariales**

En este capítulo se estructura la metodología, describiendo cada una de las etapas que la constituyen, planteando cuáles son las restricciones que las limitan y que además sirven como guía para llegar a diversas soluciones cercanas al óptimo en el tratamiento de los factores de riesgo empresarial.

En este proyecto se propone un modelo cuantitativo con la cual es posible estimar el nivel de riesgo en el que se encuentra una empresa y revelar áreas prioritarias para atender el riesgo [112]. Los elementos del modelo son obtenidos de los resultados publicados en [10] y son la base para la elaboración del cuestionario para la autoevaluación de los elementos vulnerables con la herramienta FODA [113].

El diagrama de flujo de la figura 3.1, muestra la metodología propuesta con la realización de las etapas.

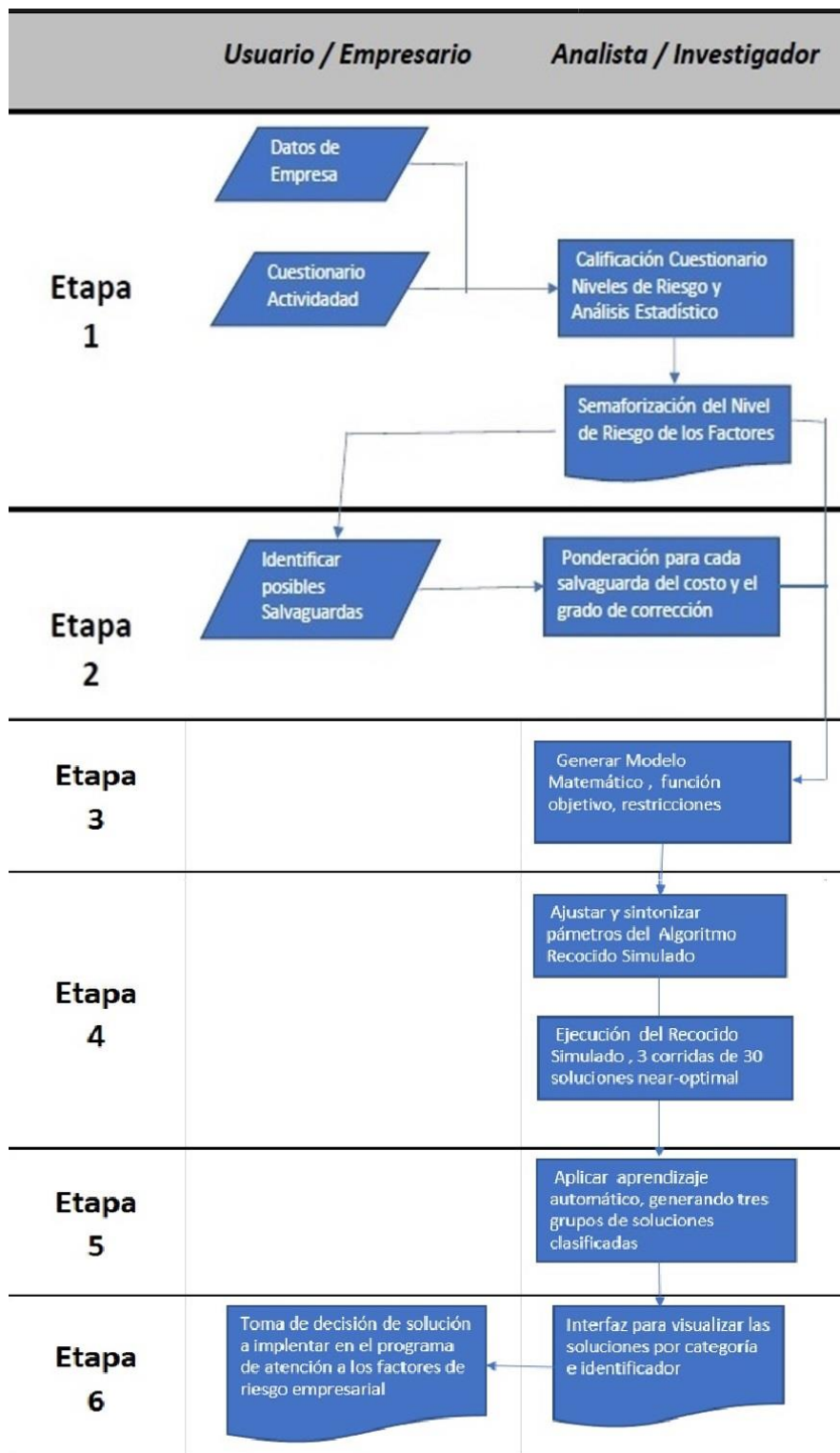


Figura 3.1 Diagrama de la metodología para el Tratamiento de los Factores de Riesgo.



Etapa I. Recopilación de datos. Los datos comerciales principales son necesarios para la identificación de los factores de riesgo en todas las áreas empresariales. El reconocimiento de las fortalezas y debilidades internas, así como las oportunidades y amenazas externas [113], se realiza sobre la base de un cuestionario cualitativo de análisis FODA (ver Apéndice A). Este cuestionario está basado en [10] para descubrir los factores de riesgo que pueden llevar a la quiebra a las empresas conforme a las decisiones cualitativas que realizan los expertos. Está conformado por 32 factores categorizados en seis áreas de riesgo establecidas por uno de los mayores bancos comerciales coreanos, ver figura 3.2 Clasificación de los factores de riesgos en seis categorías.

Las seis áreas son: i. Riesgo de la industria (IR), ii. Riesgo de gestión administrativa (MR), iii. Flexibilidad Financiera (FF), iv. Credibilidad (CR), v. Competitividad (CO) y vi. Riesgo operativo (OP). La siguiente figura muestra los 32 factores en estas seis áreas.

i. Riesgo Industria	ii. Riesgo Admitivo.	iii. Flexibilidad Financiera	iv. Credibilidad	v. Competitividad	vi. Riesgo Operativo
Grado de competencia					Eficacia de la red de venta
Sensibilidad a cambios factor macroeconómico	Administración de recursos humanos				Estabilidad de las transacciones
Políticas gubernamentales y acuerdos internacionales	Planificación de negocios a corto y largo plazo				Precio de venta y condición de liquidación
Ciclicidad. El negocio fluctúa por temporada	Capacidad y competencia en la gestión (administración)				Estabilidad y diversidad de las adquisiciones (compras)
Estabilidad en el mercado de suministros y costos de estos	Estabilidad en el área administrativa				Eficiencia en la producción
Tamaño y crecimiento mercado	Relación entre la gerencia / propietario	Financiamiento Directo	Información confiable	Posición en el mercado	Perspectivas de la demanda de productos y servicios
Poder competitivo nacional e internacional	Crecimiento / desempeño comercial	Financiamiento Indirecto.	Historial Crediticio	Nivel de Capacidad en la actividad principal (core business)	Diversificación de ventas
Ciclo de vida de producto	Logros y viabilidad	Otras fuentes de financiamiento	Relación con instituciones financieras	Estrategia diferenciada	Cartera vencida (clientes y/o proveedores)

Figura 3.2 Factores en seis áreas empresariales.

Los resultados del cuestionario sobre actividad y riesgo empresarial, cumple en parte con las normas establecidas en el ISO31000, con los resultados se determina una ponderación del grado de peligrosidad para cada uno de los elementos determinados como factores de amenaza y debilidades con las tablas de William Fine, tabla 2.4.

Etapa II. Selección de contramedidas o salvaguardas. En esta etapa, y con la información obtenida acerca de los riesgos presentes en la empresa, se seleccionan las salvaguardas (acciones o medidas preventivas o correctivas de cada factor) que minimizan el riesgo en cada una de las amenazas detectadas. Se analizan e identifican posibles soluciones para cada uno de los factores de riesgo detectados. Lo anterior con el objeto de clasificarlas por tipo de respuesta (aceptar, mitigar, transferir o evitar). Se determina el grado de corrección alcanzado y el coste de la implementación de esa salvaguarda en la empresa, mediante tablas ajustadas del método de riesgo matemático de William Fine, tabla 2.5.

Etapa III. Desarrollo del modelo matemático. Ajuste de la función objetivo, planteamiento de las restricciones para la selección óptima de las salvaguardas a implementar en los negocios y el presupuesto límite requerido. El modelo es un problema lineal entero mixto (MIP), con variables binarias (Restricciones 1-4, pág. 47)

Etapa IV. Determinación de soluciones óptimas locales. Resolver el modelo MIP mediante la heurística RS realizando varias ejecuciones para obtener un conjunto de ellas. En esta etapa y con la información obtenida, se genera la sintonización entre los parámetros que utiliza el RS. El formato de la solución con las salvaguardas seleccionadas para cada factor de riesgo es

$$S=\{JT, CT, (X_{1,1}, X_{1,2}, \dots, X_{1,s}), (X_{2,1}, X_{2,2}, \dots, X_{2,s}), \dots, (X_{a,1}, X_{a,2}, \dots, X_{a,s})\}$$

donde  $JT$  es el valor total de la justificación del beneficio al implementar el conjunto de salvaguardas,  $CT$  es el costo total por la implementación,  $X_{i,k} \in \{0,1\}$ , es 1 si la respuesta  $k$  es seleccionada para el factor  $i$ , o 0 si no fue seleccionada.

La figura 3.3 muestra la matriz  $ES_{axk}$  con la estructura de los valores de las salvaguardas para cada uno de los factores de riesgo

Fac \ Safg	$S_{i1}$	$S_{i2}$	...	$S_{is}$
$A_1$	$(J_{1,1}, CS_{1,1})$	$(J_{1,2}, CS_{1,2})$	...	$(J_{1,s}, CS_{1,s})$
$A_2$	$(J_{2,1}, CS_{2,1})$	$(J_{2,2}, CS_{2,2})$	...	$(J_{2,s}, CS_{2,s})$
$A_3$	$(J_{3,1}, CS_{3,1})$	$(J_{3,2}, CS_{3,2})$	...	$(J_{3,s}, CS_{3,s})$
...	...	...	...	...
$A_a$	$(J_{a,1}, CS_{a,1})$	$(J_{a,2}, CS_{a,2})$	...	$(J_{a,s}, CS_{a,s})$

Figura 3.3 Matriz  $ES_{axk}$ , estructura de los valores de las Salvaguardas.

Para generar la solución inicial,  $\omega_0$ , este otro formato es usado  $\omega = \{JT, CT, (k_1, k_2, k_3, k_4 \dots k_{11})\}$ , donde  $k_i$  es el tipo de respuesta de la salvaguarda implementada al correspondiente factor de riesgo  $i$ . Se genera asignando aleatoriamente un tipo de respuesta  $k = \{1,2,3,4\}$  para cada fila  $i$  de la matriz  $M_{axs}$ . Es decir, sean  $nA1, nA2$  dos números generados aleatoriamente,  $nA1 = rand(0, a)$  y  $nA2 = rand(0, s)$ , el primero es para seleccionar una fila de la matriz  $M_{axs}$  y el segundo para el tipo de salvaguarda, que es la columna en la que se asigna el número 1, es necesario utilizar una bandera que se levante para esa fila, esto es para no volver a seleccionarla, este procedimiento continuará hasta que todas las filas tengan un tipo de salvaguarda asignado, o si el costo total exceda el presupuesto  $P$ . En la tabla 4.3, se encuentra un ejemplo de la solución inicial  $\omega_0$

La adaptación de la heurística RS para este trabajo, se muestra en el algoritmo 1, donde el valor de cada uno de los parámetros se define de la siguiente manera: en primer lugar, la desviación estándar de un conjunto de soluciones

generadas aleatoriamente se obtiene como en [73], el valor  $T_0$  es igual a dos veces la desviación estándar, el parámetro de control definido  $T$ , disminuye por el coeficiente  $\alpha \times T$ , donde  $\alpha=0.98$ , la longitud de la cadena de Markov está en función del tamaño del problema,  $M=(a \times s)(a \times s-1)$ , donde  $a$  es el número de factores de riesgo y  $s$  es el tipo de salvaguardas, este valor determina el número de iteraciones en el ciclo de Metrópolis, que comienza en línea 5 a la 19. Se repite el ciclo, línea 3 a 21, hasta que se alcanza el equilibrio con un valor mínimo del parámetro de control  $T$ , que disminuye en cada iteración de acuerdo con el coeficiente  $\alpha$ , como se ve en la Sección 2.5.2, con el algoritmo general RS. El enunciado de la línea 6 genera nuevas soluciones, se calcula la diferencia de energía en la línea 7, si es negativa, se selecciona la nueva solución, de lo contrario, las líneas 11 a 14 asignan una nueva solución de acuerdo con la probabilidad determinada por la función de aceptación  $e^{\Delta/T}$ . En las líneas 16-18, se actualiza la mejor solución encontrada durante la ejecución completa del algoritmo.

Algoritmo 1. Recocido Simulado para tratamiento factores de riesgo.

```

1  Calcular una solución inicial factible  $\omega_0$ 
2   $\omega = \omega_0, \omega_{opt} = \omega_0$ 
3  repeat
4      Inicializar contador  $m = 0$ 
5      for  $m=0$  to  $M$  do
6          Generar nueva solución  $\omega' = \Omega(\omega)$ 
7          Calcular  $\Delta = f(\omega') - f(\omega)$ 
8          if  $\Delta < 0$  then
9               $\omega = \omega'$ 
10         else
11             Generar random  $\rho(0,1)$ 
12             if  $\rho \leq e^{\Delta/T}$  the
13                  $\omega = \omega'$ 
14             end if
15         end if
16         if  $\omega_{opt} > \omega$  then
17              $\omega_{opt} = \omega$ 

```

```

18     end if
19     end for
20     T =  $\alpha \times T$ 
21 until T  $\leq \beta \times T_0$ 

```

Algoritmo 2. Estructura de Vecindad  $\Omega$  del RS.

```

1 function  $\Omega(\omega)$ 
2 Suma=0
3 SX = {SX1 = 0, ..., SXa = 0}
4 bandera= 0; contador = 0
5 while bandera = 0 and contador  $\leq a$  do
6     contador = 1
7      $\omega T = \omega$ 
8     Generar nA1 random(1,a)
9     Generar nA2 random(1,s)
10    while SX(nA2) = 1 do
11        Generar nA1 random(1,a)
12    end while
13    SX(nA2)=1
14    for j=1 to s do
15        if j <> nA2 then
16             $\omega T(nA1,j) = 0$ 
17        else
18             $\omega T(nA1,nA2) = 1$ 
19        end if
20    end for
21    Calcular Suma
22    if Suma  $\leq P$  then
23        bandera = 1
24    else
25        Solucion no factible
26        Incrementar contador
27    end if
28 end while
29  $\omega' = \omega T$ 
30 return  $\omega'$ 

```

El intercambio consiste en generar dos números aleatorios,  $nA1=rand(0,a)$  y  $nA2=rand(0,s)$  para que desde la solución  $\omega$  enviada a la función  $\Omega$ , la línea  $nA1$  se modifica y se le asigna 1 al elemento  $\omega T(nA1,nA2)=1$  y  $\omega T(nA1,j)=0$  para  $j \neq nA2$ . El costo total de  $\omega T$ ,  $CT(\omega T) = \sum_{i=1}^a \sum_{j=1}^s CT(i,j)\omega T(i,j)$  se calcula y se verifica que no excede el presupuesto  $P$ , en caso de excederlo, se rechaza  $\omega T$  y se calcula otra solución vecina, si  $\omega T$  es una solución factible, entonces  $\omega_0=\omega T$  y la función devuelve  $\omega_0$ . Así,  $\omega_0$  es una solución factible generada aplicando el intercambio (swap) que no excede el presupuesto asignado  $P$ .

Etapa V. Clasificar las soluciones. Mediante la técnica de aprendizaje automático, k-means, las soluciones cercanas al óptimo se clasifican en tres grupos, de acuerdo con el grado de justificación que incluye el grado de corrección de riesgos, y de acuerdo al presupuesto asignado. Con el lenguaje de software R [30] se usaron las bibliotecas /commands para análisis de clúster [114], las funciones k-means están disponibles en el paquete clúster, las instrucciones en el script se encuentran en el algoritmo 3.

Algoritmo 3. R Script para clasificación de soluciones RS con k-means.

```

1  library(cluster)
2  k<- 3 # numero de grupos
3  DatRD = read.csv("c:\\data\\SRSol.csv")
4  NLong=length(DatRD$Num)
5  resultado <- kmeans(DatRD,k)
6  clusplot(DatRD, resultado)
7  DatRD=cbind(DatRD,Cluster=resultado$cluster)
8  write.csv(DatRD, file = "c:\\data\\DtC.csv",row.names=FALSE)

```

Los datos de entrada están en uno de los archivos generados por el algoritmo RS (SRSol.csv), la línea 6 obtiene el gráfico de los principales componentes del clúster.

Con este conjunto de datos que mediante la función k-means, ha sido clasificado y contiene el número de clúster o grupo al que pertenece, se determinaron los elementos relevantes de cada grupo con árboles de decisión [115]. El paquete utilizado en el software R era rpart y rpart. plot [116], el script está en el algoritmo 4.

Algoritmo 4. R Script para factores relevantes de grupos de soluciones mediante Árbol de Decisión.

```
1 set.seed(4355) # utilizar una semilla
2 library(rpart)
3 library(rpart.plot)
4 DatClus= read.csv("c:\\data\\DtC.csv")
5 ArbolK <- rpart(Cluster~.,method="class",data=DatClus)
6 print(ArbolK)
7 rpart.plot(ArbolK)
```

La línea 6 imprime la estructura del árbol y la línea 7 imprime el árbol de decisión con los nodos raíz y las hojas finales.

Etapa VI. Visualización de soluciones para la toma de decisión. La presentación final de las soluciones se muestra en un archivo MS Excel. Mediante una interfaz de usuario con tablas dinámicas [117], se integran las soluciones con la categoría asignada en la etapa anterior, para que, con el estudio realizado y los niveles de beneficio y costo total, se seleccione la solución más viable, apoyando en la toma de decisión de la solución a implementar por la empresa. Esto facilita la toma de decisión para la elaboración de un programa específico y óptimo para la atención a los factores de riesgo empresarial.

## **Capítulo 4. Aplicación de la metodología (caso de estudio)**

En este capítulo, para establecer claridad en el procedimiento de aplicación de la metodología propuesta en este proyecto, para explicitar su sistematización y para garantizar su validez y su eficacia, se aplica sobre un problema representativo real. La información generada será utilizada para evaluar su eficacia y eficiencia.

### **4.1 Caso de Estudio**

La empresa participante se encuentra en el sector de hidrocarburos en la ciudad de Cuernavaca en México. La compañía cuenta con amplia experiencia y se dedica a dar asesoría en protección civil y en brindar consultoría para el cumplimiento de las disposiciones administrativas de carácter general y normatividad aplicable de la NOM-005-ASEA-2016. La empresa desarrolla y asesora en la elaboración de análisis de riesgo para el sector hidrocarburos (ASAH) y cuenta con más de treinta clientes de estaciones de servicio de hidrocarburos. Dentro de sus líneas estratégicas, está ampliar sus servicios a otros sectores industriales. Inicialmente determinan que tienen una oportunidad de mejora que atender para evitar el riesgo de insolvencia y este consiste en aumentar el crecimiento de ganancias (más clientes o proyectos); tener un margen operativo mayor (cash Flow, disminuir cartera vencida) y mejorar el uso de los activos (RH, herramientas tecnológicas, entre otros).

La siguiente figura 4.1, muestra la necesidad de la empresa para continuar operando, es decir, presenta la situación de atender las fallas que podrían generarse y el esquema de qué hacer para detectar, prevenir o superar esas fallas mediante una serie de acciones denominadas salvaguardas en una solución cercana a la óptima para evitar el fracaso empresarial.



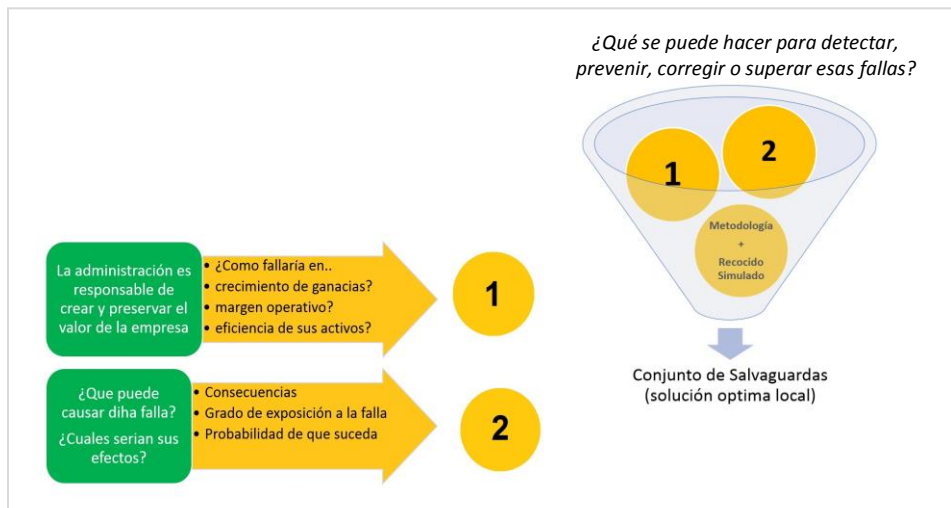


Figura 4.1: Necesidades de la empresa para continuar operando.

#### 4.2. Aplicación de la Metodología.

Considerando la información proporcionada por la empresa y sus necesidades, se aplica la metodología:

##### Etapa 1. Recopilación de datos

En esta etapa inicial, el administrador con experiencia académica y profesional adecuada para responder preguntas sobre la gestión de riesgos empresariales fue la persona que respondió el cuestionario ([Apéndice A](#)).

Se identificaron y cuantificaron once factores de riesgo,  $a=11$ . Para la identificación de posibles salvaguardas en los cuatro tipos de respuestas,  $s=4$ , (capítulo 3 etapa II), se realizó una lluvia de ideas y se determinó en función de las necesidades y posibilidades de la empresa. El gráfico de semáforo que muestra el nivel de riesgo para todos los factores se encuentra en la figura 4.2.

i. Riesgo Industria	ii. Riesgo Admitivo.	iii. Flexibilidad Financiera	iv. Credibilidad	v. Competitividad	vi. Riesgo Operativo
Grado de competencia					Eficacia de la red de venta
Sensibilidad a cambios factor macroeconómico	Administración de recursos humanos				Estabilidad de las transacciones
Políticas gubernamentales y acuerdos internacionales	Planificación de negocios a corto y largo plazo				Precio de venta y condición de liquidación
Ciclicidad. El negocio fluctúa por temporadas	Capacidad y competencia en la gestión (administración)				Estabilidad y diversidad de las adquisiciones (compras)
Estabilidad en el mercado de suministros y costos de estos	Estabilidad en el área administrativa				Eficiencia en la producción
Tamaño y crecimiento mercado	Relación entre la gerencia / propietario	Financiamiento Directo	Información confiable	Posición en el mercado	Perspectivas de la demanda de productos y servicios
Poder competitivo nacional e internacional	Crecimiento / desempeño comercial	Financiamiento Indirecto.	Historial Crediticio	Nivel de Capacidad en la actividad principal (core)	Diversificación de ventas
Ciclo de vida de producto	Logros y viabilidad	Otras fuentes de financiamiento	Relación con instituciones financieras	Estrategia diferenciada	Carta vencida (clientes y/o proveedores)

Figura 4.2 Nivel de riesgo en los 32 factores por dimensión.  
Riesgo en Rojo, SIN Riesgo en Verde.

Con las respuestas obtenidas, se calificó y realizó la ponderación de las amenazas, etiquetando cada una con valores de Grado Peligrosidad, GP, conforme al método de William Fine. La figura 4.3, muestra las imágenes de la interfaz para la valoración de los factores de riesgo que fueron automatizados en una hoja de cálculo, la figura del inciso a, presenta los factores de riesgo por renglón, el usuario selecciona la celda que presenta la consecuencia que resultaría por dicho factor, el inciso b presenta las opciones de selección con respecto a la exposición indicando la frecuencia con que ocurre o se presenta dicho factor, y en el inciso c, se debe seleccionar la probabilidad de que se presente el factor indicado. Una vez marcadas las celdas correspondientes, se ejecuta la macro que esta codificada en el botón de la esquina superior izquierda calculando el grado de peligrosidad e imprimiéndolo en el renglón correspondiente y en la celda de la columna C.

a)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
id	Vulnerabilidad	GP	1.-Consecuencias (C) (Resultado más probable del evento potencial)				2.-Exposición (E) (Frecuencia con que ocurre la situación de riesgo) La situación de riesgo ocurre:				3.-Probabilidad (P) (Probabilidad de que la secuencia del evento se complete)			
			25	15	5	1	10	6	3	1	10	6	3	1
			a	b	c	d	a	b	c	d	a	b	c	d
1	12 Administración de recursos humanos	75	x											
2	17 Financiamiento Indirecto.	900	x											
3	32 Eficacia de la red de venta	75												
4	3 Grado de competencia que tiene su personal	75												
5	6 Sensibilidad de afectación a su producción	75	x											
6	14 Planificación de negocios a corto y largo plazo	900	x											

b)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
id	Vulnerabilidad	GP	1.-Consecuencias (C) (Resultado más probable del evento potencial)				2.-Exposición (E) (Frecuencia con que ocurre la situación de riesgo) La situación de riesgo ocurre:				3.-Probabilidad (P) (Probabilidad de que la secuencia del evento se complete)			
			25	15	5	1	10	6	3	1	10	6	3	1
			a	b	c	d	a	b	c	d	a	b	c	d
1	12 Administración de recursos humanos	75	x											
2	17 Financiamiento Indirecto.	900	x											
3	32 Eficacia de la red de venta	75												
4	3 Grado de competencia que tiene su personal	75												
5	6 Sensibilidad de afectación a su producción	75	x											
6	14 Planificación de negocios a corto y largo plazo	900	x											
7	16 Financiamiento Directo	750	x											
8	18 Otras fuentes de financiamiento	45												
9	26 Estabilidad de las transacciones (ausencia de impago)	15												
10	30 Precio de venta y condición de liquidación	25	x											

c)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
id	Vulnerabilidad	GP	1.-Consecuencias (C) (Resultado más probable del evento potencial)				2.-Exposición (E) (Frecuencia con que ocurre la situación de riesgo) La situación de riesgo ocurre:				3.-Probabilidad (P) (Probabilidad de que la secuencia del evento se complete)			
			25	15	5	1	10	6	3	1	10	6	3	1
			a	b	c	d	a	b	c	d	a	b	c	d
1	12 Administración de recursos humanos	9												
2	17 Financiamiento Indirecto.	270	x											
3	32 Eficacia de la red de venta	2500	x											
4	3 Grado de competencia que tiene su personal	270												
5	6 Sensibilidad de afectación a su producción	75	x											
6	14 Planificación de negocios a corto y largo plazo	900	x											
7	16 Financiamiento Directo	750	x											
8	18 Otras fuentes de financiamiento	45												
9	26 Estabilidad de las transacciones (ausencia de impago)	15												
10	30 Precio de venta y condición de liquidación	25	x											
11	20 Información confiable para la toma de decisiones	270												
12														

Figura 4.3 Tablas de William Fine en interfaz electrónica para ponderar los factores. a) valores para selección de las Consecuencias; b) valores de la Exposición y c) valores de la Probabilidad.

Se observan en la gráfica 4.4 a los porcentajes y en la tabla 4.1 con la categoría/dimensión de riesgo empresarial al que pertenece.

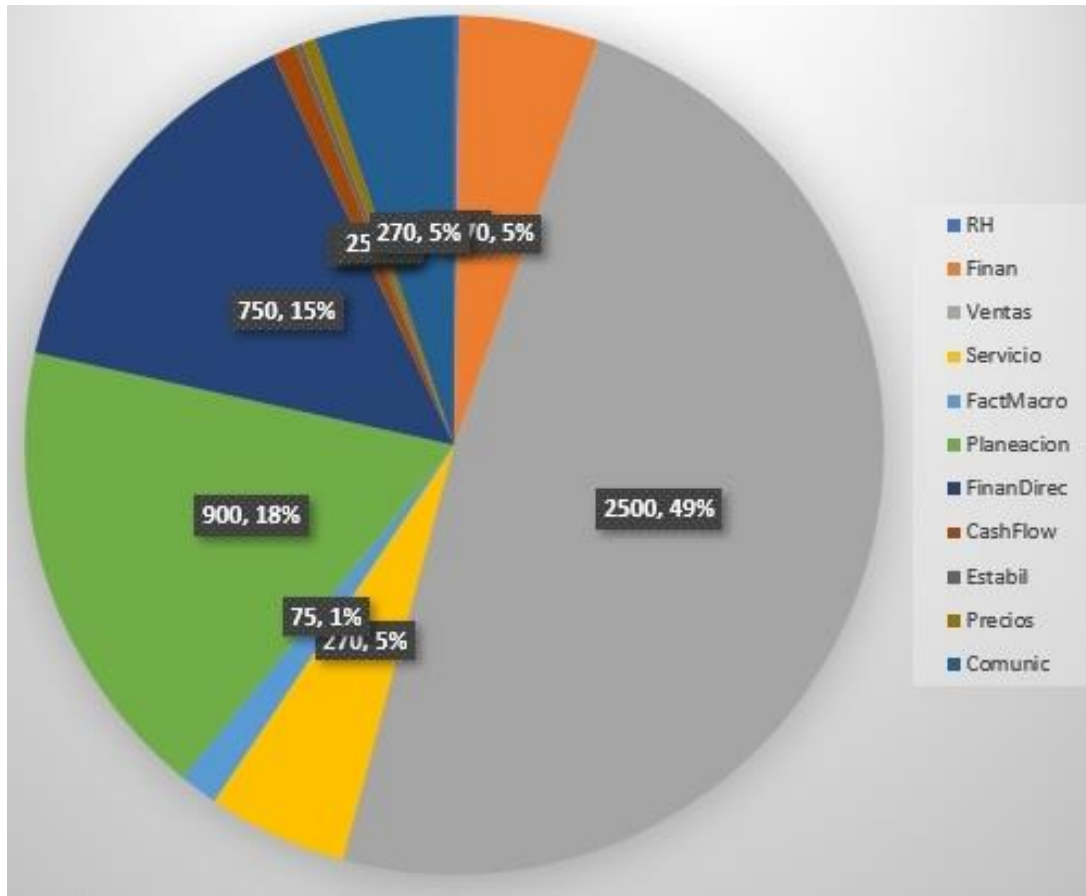


Figura 4.4: Porcentajes del nivel de riesgo

Tabla 4.1: Categoría de los Factores de Riesgo.

ID CATEG	CATEGORÍA	FACTOR DE RIESGO
2	Riesgo Administrativo	Administración de recursos humanos (plan carrera, plan de comunicación, evaluaciones de desempeño, etc.)
3	Flexibilidad Financiera	Financiamiento Indirecto.
6	Riesgo Operativo	Eficacia de la red de venta
1	Riesgo de la Industria/ Sector	Grado de competencia que tiene su producto/servicio
1	Riesgo de la Industria/ Sector	Sensibilidad de afectación a su producto/servicio a los cambios en los factores macroeconómicos (inflación, desempleo, inversión extranjera, etc.)
2	Riesgo Administrativo	Planificación de negocios a corto y largo plazo
3	Flexibilidad Financiera	Financiamiento Directo
3	Flexibilidad Financiera	Otras fuentes de financiamiento
6	Riesgo Operativo	Estabilidad de las transacciones (ausencia de crisis importantes)
6	Riesgo Operativo	Precio de venta y condición de liquidación
5	Cumplimiento	Información confiable para la toma de decisiones

Es importante resaltar que el mayor factor de riesgo está en “Eficiencia de la red de ventas”. Esto se concluye, por la ponderación asignada al responder el cuestionario, de la baja eficacia en la red de ventas además de calificar la consecuencia de pérdida entre un 10 y 100 % del parámetro de Valor Nómina Mensual, con una exposición semanal y probabilidad de ocurrencia de que se presenta una falta de eficacia del 20 %.

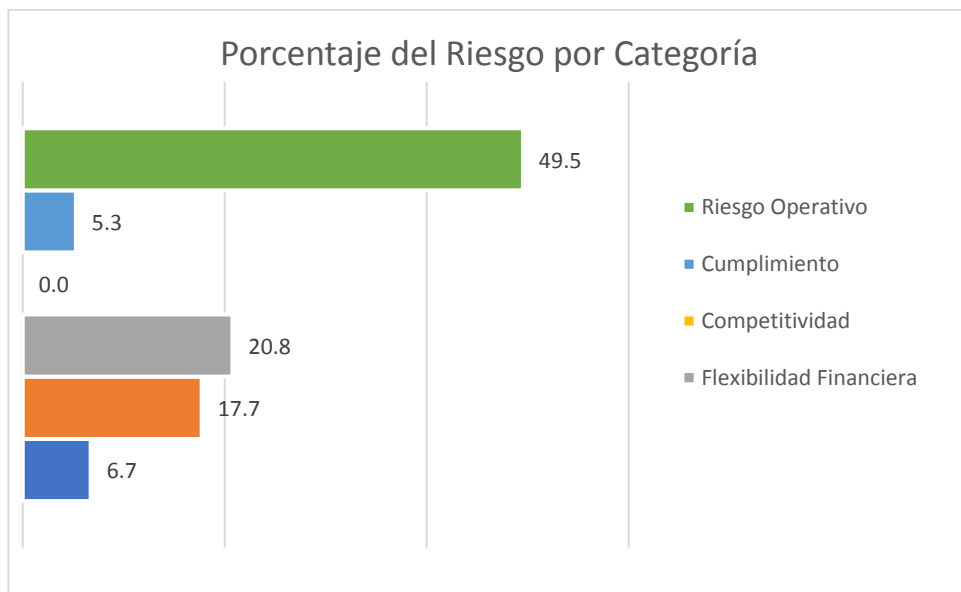


Figura 4.5 Porcentaje del Riesgo por Categoría.

Haciendo un análisis con respecto a los riesgos por Categoría, en la figura 4.5 se muestra los porcentajes del total del riesgo empresarial. Como puede observarse, el mayor riesgo, con un valor de casi del 50% está en el riesgo operativo. Se puede notar que no se observa la existencia de un riesgo por competitividad, esto se debe a que la empresa se ha posicionado en el mercado con una estrategia diferenciada y un alto nivel de capacidad en la actividad que desarrolla. Se puede concluir, sin embargo, que el riesgo está siendo en la operación por la poca eficacia en la red de ventas y poca estabilidad de las transacciones.

#### Etapa II. Análisis de contramedidas o salvaguardas

En función de las condiciones de riesgo detectadas, las necesidades y posibilidades de la empresa, se determinaron las salvaguardas para dar cuatro tipos de respuestas. La figura 4.6 presenta las 44 salvaguardas ordenadas por los tipos de respuesta proporcionada por el personal administrativo responsable, para cada uno de los 11 factores de riesgo detectados en la empresa.



<b>Id</b>	<b>1 Aceptar</b>	<b>2 Mitigar</b>	<b>3 Transferir</b>	<b>4 Evitar</b>
V1	Área de mesa de ayuda	Capacitación servicio al cliente y Elaboración de plan de carrera para el personal	Capacitación de acuerdo a necesidades detectadas. Programa de capacitación	Contratar expertos en RH
V2	Se utiliza solo el recurso actual	Tener bienes de garantía	Buscar asociación con capitalista	Ahorro de utilidades para emergencias
V3	Realizar un análisis de mercado y detectar necesidades	Elaborar plan publicidad, preparación equipo de ventas directas y plataforma tecnológica	Contratar un servicio marketing	Aceptar clientes nuevos solo por recomendación
V4	Realizar alianzas comerciales	Elaborar plan ampliar el mercado, nuevos tipos servicio, contacto clientes antiguos	Contratación asesores para plan de creatividad e innovación	Consultoría para plan de crecimiento
V5	Tener una partida para multas	Realidad un monitoreo diario oficial periódicamente	Contratación despacho para monitoreo	Vincularse con un agente del giro adecuado
V6	Registrar entradas y salidas de recursos	Elaborar plan estratégico, proyección anual, presupuesto	Contratar outsourcing para elaboración plan estratégico	Registro de incidencias
V7	Aumentar monto caja chica	Habilitar depto. de cobranzas	Contratar servicio cobranza	Cobro pólizas por adelantado a los clientes
V8	Aumentar salarios al personal actual	Solicitar de 3 a 5 jóvenes "construyendo el futuro"	Contratar personal técnico en outsourcing	Limitarse a la atención de clientes con la plantilla actual
V9	Aumentar monto gastos por doble visita	Implementar un plan de mejora de la comunicación con clientes para evitar doble visita	Servicio de implementación a clientes de terceros (GASAF)	Implementar servicio en línea
V10	Tener actualizado el costo de liquidación	Seguir la planeación, y vigilar la tendencia para innovar en estrategias	Generar venta Accionistas	Adquirir crédito
V11	Elaborar y difundir información	Contar con un buen plan de comunicación interna y externa en plataforma	Contratar servicios de atención a redes sociales y página web	Solicitar informes semanales

Figura 4.6: Salvaguardas ordenadas por los tipos de respuesta proporcionada. V1 - Administración de recursos humanos; V2 - Financiamiento Indirecto; V3 - Eficacia de la red de venta; V4 -Grado de competencia que tiene su producto/servicio; V5 -Sensibilidad de afectación a producto/servicio a cambios económicos; V6 - Planificación de negocios a corto y largo plazo; V7 - Financiamiento Directo; V8 - Otras fuentes de financiamiento; V9 - Estabilidad de las transacciones (ausencia de crisis importantes); V10 - Precio de venta y condición de liquidación; V11 - Información confiable para la toma de decisiones.

Para determinar la ponderación sobre el grado de corrección que aporta la salvaguarda sobre el factor de riesgo, los valores de las tablas se cargaron a una hoja de cálculo, figura 4.7, de manera que se marca con una **X** la casilla del porcentaje de corrección, con los siguientes rangos que se fijaron en función del método de William Fine, tabla 2.5, corrección del 100%, entre 75 al 99%, entre 50 al 74% y, entre un 25 al 49%. El factor de coste se basa en el monto de la nómina mensual establecido por acuerdo de la empresa y se determina el valor del parámetro definido como el número de nóminas tope (máximo fijado por la empresa) que se está dispuesto a ejercer para la atención a los riesgos detectados con la implementación de las salvaguardas. El valor de este parámetro  $X$ , para este estudio fue de 3, ( $X=3$ ). Con el valor GP obtenido en la etapa anterior, se aplica la fórmula del Método de William Fine para obtener la Justificación,  $J=GP/(GC*FC)$ , esta justificación  $J_{i,k}$  es la que determina el beneficio de la implementación de esa salvaguarda para responder al factor de riesgo  $i$ .

En la figura 4.7 se muestran los valores de la tabla 2.4 ingresados en la interfaz y en la figura 4.8, se muestra las elecciones realizadas para cada uno de los once factores de riesgo en cada uno de sus cuatro tipos de respuesta, aceptar, mitigar, transferir y evitar.

Grado de corrección (GC)				Factor de coste(FC)				
1	2	3	4	10	6	4	2	1
a	b	c	d	a	b	c	d	e
Riesgo completamente eliminado 100%.	Riesgo reducido al menos el 75%.	Riesgo reducido del 50% al 75%.	Riesgo reducido del 25% al 50%.	Superior a X nóminas mensuales	Entre 1 y X nóminas mensuales	Entre el 50% al 100% de la nómina mensual	Entre el 10% al 50% de la nómina mensual	Menos del 10% de la nómina mensual

Figura 4.7 Interfaz automatizada con los valores de Fine para valorar el grado de riesgo.



o Con	D <sub>1</sub> Aceptar										D <sub>2</sub> Mitigar									
	Grado de corrección (GC)					Factor de coste(FC)					Grado de corrección (GC)					Factor de coste(FC)				
	1	2	3	4	10	6	4	2	1	1	2	3	4	10	6	4	2	1		
GP	a	b	c	d	a	b	c	d	e	a	b	c	d	a	b	c	d	e		
9				x					x				x					x		
270				x					x				x					x		
2500			x					x					x					x		
270	x							x					x					x		
75				x				x					x					x		
900				x				x					x					x		
750				x				x					x					x		
45			x					x					x					x		
15			x					x					x					x		
25				x				x					x					x		
270				x				x					x					x		

Figura 4.8 Interfaz automatizada para realizar la elección de corrección y coste.

**Etapa III. Definir el modelo de optimización.**

Una vez conocidos los valores  $J_{i,k}$  y  $CS_{i,k}$ , se genera la matriz de Justificación y de Costes,  $M_{axs}$ . Esta matriz muestra para cada factor de riesgo, los valores del beneficio y costo en caso de seleccionar la salvaguarda de la columna especificada, es decir, la salvaguarda del tipo de respuesta al factor de riesgo del renglón.

Tabla 4.2 Matriz de Justificación y de Costo de Implementación.

V <sub>i</sub> /S <sub>j</sub>	J - Justificación				CS – Coste de implementación (miles de pesos)			
	S <sub>i1</sub>	S <sub>i2</sub>	S <sub>i3</sub>	S <sub>i4</sub>	S <sub>i1</sub>	S <sub>i2</sub>	S <sub>i3</sub>	S <sub>i4</sub>
	Aceptar	Mitigar	Transferir	Evitar	Aceptar	Mitigar	Transferir	Evitar
1	2.25	3.00	2.25	1.50	1.767	2.799	47.386	77.374
2	33.75	33.75	45.00	135.00	7.812	53.242	60.841	1.436
3	208.33	625.00	250.00	138.89	54.435	10.254	908.163	84.842
4	22.50	67.50	22.50	33.75	111.757	10.743	82.412	58.859
5	3.13	75.00	9.38	6.25	112.287	5.663	58.487	40.921
6	56.25	900.00	150.00	112.50	45.746	1.397	6.321	9.555
7	46.88	375.00	62.50	46.88	44.061	4.211	97.362	49.435
8	2.50	7.50	3.75	2.50	75.828	7.676	54.894	109.47
9	1.25	2.50	3.75	1.25	47.675	11.917	57.329	73.612
10	1.56	6.25	2.08	1.56	50.853	11.880	74.636	46.016
11	33.75	45.00	67.50	67.50	6.638	119.965	10.057	5.906

Con la información de la tabla 4.2, se define la función objetivo y se establecen las restricciones del modelo, restricciones 1-4. En el modelo  $a=11$ ,  $s=4$ , la

ecuación 1 es la función objetivo, FO, la suma del cociente de costo-beneficio de la selección de salvaguardas, la restricción en 2 es la de no exceder el presupuesto asignado P, la restricción en 3, es donde se limita a que solo se puede seleccionar una salvaguarda para cada factor de riesgo y la restricción en 4 es sobre la variable binaria  $x_{i,k}$ , que indica que la salvaguarda tipo  $k$  del factor  $i$  es seleccionada, sólo si  $x_{i,k}=1$  o 0 si ésta no es seleccionada. Como se puede observar, en la restricción 3 únicamente se puede seleccionar una o ninguna salvaguarda, para cada factor de riesgo. En función de que la justificación y el costo existen, se asume que  $J_{i,k} > 0$  y  $CS_{i,k} \geq 0$  para  $i=1, \dots, 11$  y  $k=1, \dots, 4$ .

#### Etapa IV. Aplicación de heurística Recocido Simulado (RS).

La interpretación del problema con el RS resulta al especificar la relación entre los parámetros y las variables de salida, para el problema de optimización de la respuesta a los factores de riesgo empresarial. El proceso inicia, con una solución inicial factible, donde la selección de las salvaguardas, para cada factor  $i$ , se realiza en forma aleatoria. El valor de la Justificación total de la solución se determina calculando la suma de la variable binaria  $x_{i,k}$  con la matriz  $J$ , (valores de la justificación de cada tipo de salvaguarda para cada amenaza) y también con la suma de la variable binaria  $x_{i,k}$  con la matriz de costo  $CS$ . La función energía, definida como el cociente del costo entre la justificación que se obtiene de las salvaguardas elegidas en esa solución, *Costo/Justificación*, a medida que se van generando las iteraciones del RS con soluciones factibles cercanas a la anterior, el valor del cociente disminuye. El resultado final se da cuando el cociente converge al mínimo costo de la implementación de las salvaguardas seleccionadas y con la mayor justificación posible de que las salvaguardas seleccionadas disminuyen los factores de riesgo empresarial.

Como se puede observar, el espacio solución se presenta en la instancia es de  $4^{11} = 4'194,304$ , lo anterior se debe a que se determinaron 11 factores de

riesgo, los cuales pueden ser atendidos cada uno de 4 maneras diferentes, esto es  $4^{11}$ . La solución inicial factible se construye de manera aleatoria. Se genera un número aleatorio que indica el renglón en el cual se selecciona una salvaguarda del tipo  $k$ ,  $k=\{1,2,3,4\}$ , que también se genera aleatoriamente.

El algoritmo construye una estructura de vecindad  $\eta(\omega)$ , donde  $\Omega$  se define como el conjunto de soluciones  $\omega$  factibles que pueden ser generadas de las soluciones vecinas  $\omega$ ,  $\omega=\{J,CS,(k_1,k_2,k_3,k_4\dots k_{11})\}$ , donde  $k_j$  es el tipo de respuesta de la salvaguarda implementada a la amenaza  $i$  correspondiente. Cada solución  $\omega' \in \eta(\omega)$  puede ser obtenida al seleccionar aleatoriamente alguna de las amenazas y generar una pequeña perturbación  $Y$ . La vecindad queda definida con la función  $\eta(\omega) = \{\omega' \in W \mid \omega \rightarrow \omega'\}$  donde  $\Omega$  es el conjunto de soluciones factibles para el problema y  $\eta(\omega)$  define la vecindad de  $\omega$ , de manera que las soluciones  $\omega'$  son muy similares a  $\omega$ , la perturbación  $Y$  que puede ser un swap (intercambio), una inserción de selección si la amenaza no tiene seleccionada una salvaguarda, o la eliminación de la salvaguarda para esa amenaza, posteriormente se evalúa si esta solución no excede el presupuesto asignado  $P$ , en caso de exceder se eliminará una selección de salvaguarda de manera aleatoria obteniendo así otra solución  $\omega$ .

La siguiente tabla muestra las soluciones vecinas generadas a partir de la solución inicial, que, como se mencionó anteriormente, es generada aleatoriamente,  $\omega =\{1653.3, 1256, (2,3,3,4,4,2,2,4,2,1,1)\}$ . Se puede observar que la primera solución  $\omega'$  se obtuvo mediante un intercambio (swap) en V2,  $\omega'=\{1642.06, 1203.04, (2,1,3,4,4,2,2,4,2,1,1)\}$ .

Tabla 4.3 Conjunto de Soluciones Vecinas, cada renglón es una solución con el tipo de respuesta a cada elemento del vector de factores de riesgo V.

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
(1)2	(2)3	(3)3	(4)4	(5)4	(6)2	(7)2	(8)4	(9)2	(10)1	(11)1
2	1	3	4	4	2	2	4	2	1	1
2	1	3	4	4	2	2	4	2	2	1
2	1	1	4	4	2	2	4	2	2	1

2	1	1	4	3	2	2	4	2	2	1
2	1	2	4	3	2	2	4	2	2	1
2	1	2	4	2	2	2	4	2	2	1

Para alcanzar la convergencia del algoritmo, es necesario especificar un conjunto de parámetros, éstos son: la temperatura inicial  $t_0$ , el parámetro de control de temperatura  $\alpha$ , la temperatura final con el criterio de paro, y longitud de cadena de Markov [118].

Considerando lo recomendado en [119] y [73] se estableció que la temperatura inicial es igual al doble de la desviación estándar de las soluciones generadas al azar. Posteriormente, a partir de los valores de las soluciones generadas al azar, se calcularon los promedios, desviación estándar, el valor máximo y mínimo de las variables del modelo que son: Justificación, Costo y el cociente  $f(x) = \text{Costo} / \text{Justificación}$ , tabla 4.4.

Tabla 4.4 Variables RS.

	Justificación	Costo	f(x)
PROMEDIO	1257.7161	372.063245	31.7084369
DESV ESTANDAR	549.9868	321.806131	23.7618965
MAXIMO	2263.5	1256.07593	95.521835
MINIMO	466.763885	72.848808	8.838078

Tabla 4.5 Parámetros RS.

$T_0$	Valor inicial del parámetro de control	$T_0 = 2 * \text{DesvEstandar} = 47.52$
$\alpha$	Coficiente de control	$\alpha = 0.98; T_k = T_{k-1} * \text{alfa}$
M	Tamaño Vecindad, Cadena de Markov	$M = (a*s)*(a*s-1) = 1892$
$\beta$	Coficiente del Criterio	$\beta = 0.001; T_k < \text{beta} * T_0$

En la figura 4.9 se muestra el diagrama de flujo del RS, con los valores de la tabla 4.5. El ciclo interno del RS conocido como Ciclo de Metrópolis, mostrado en la figura 4.10, haciendo referencia al proceso de la función vecindad, la función de aceptación, al esquema de enfriamiento y finalmente, a la condición de paro. Este ciclo se realiza en una corrida del RS.

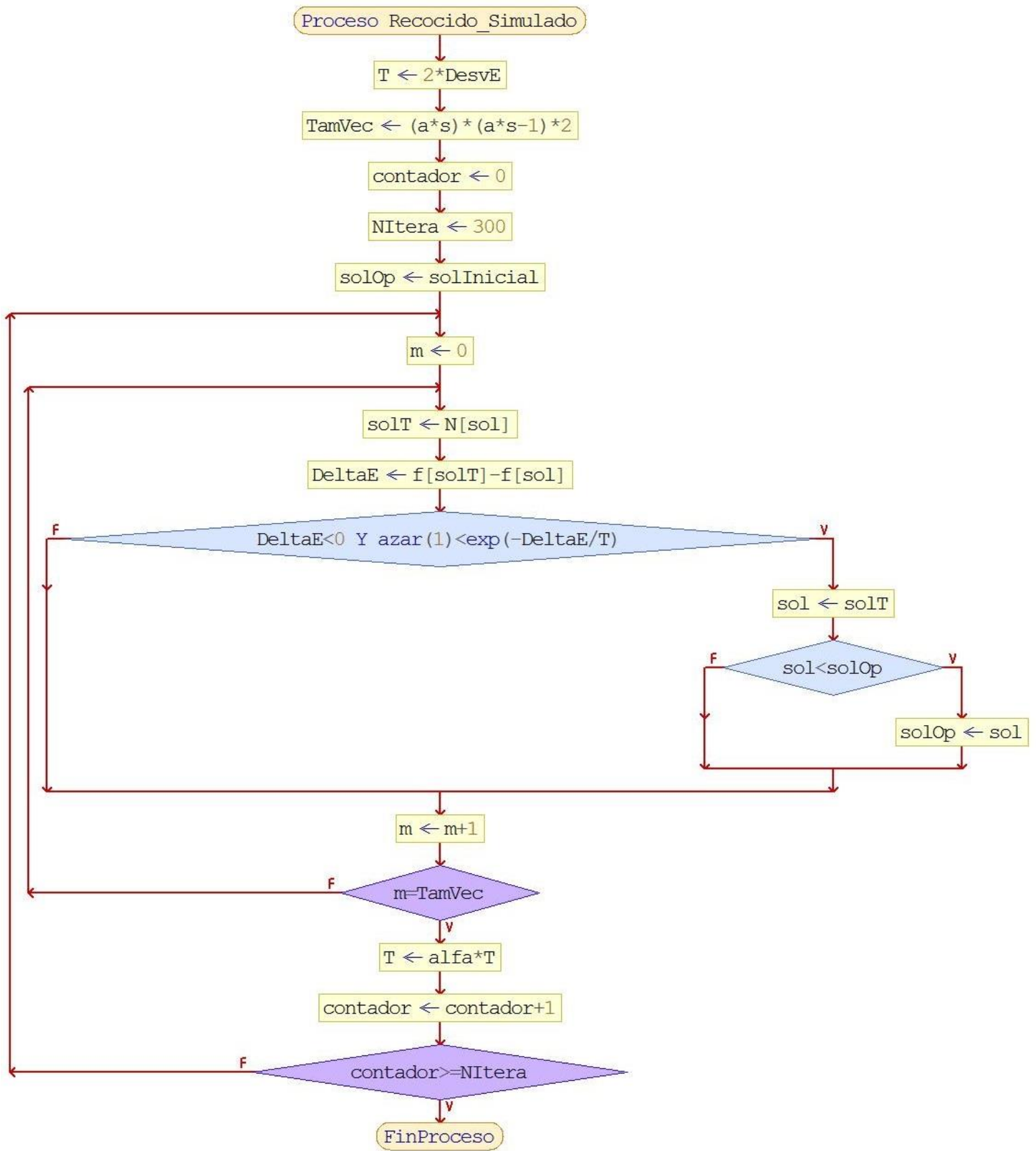


Figura 4.9 Diagrama de Flujo del RS adaptado.

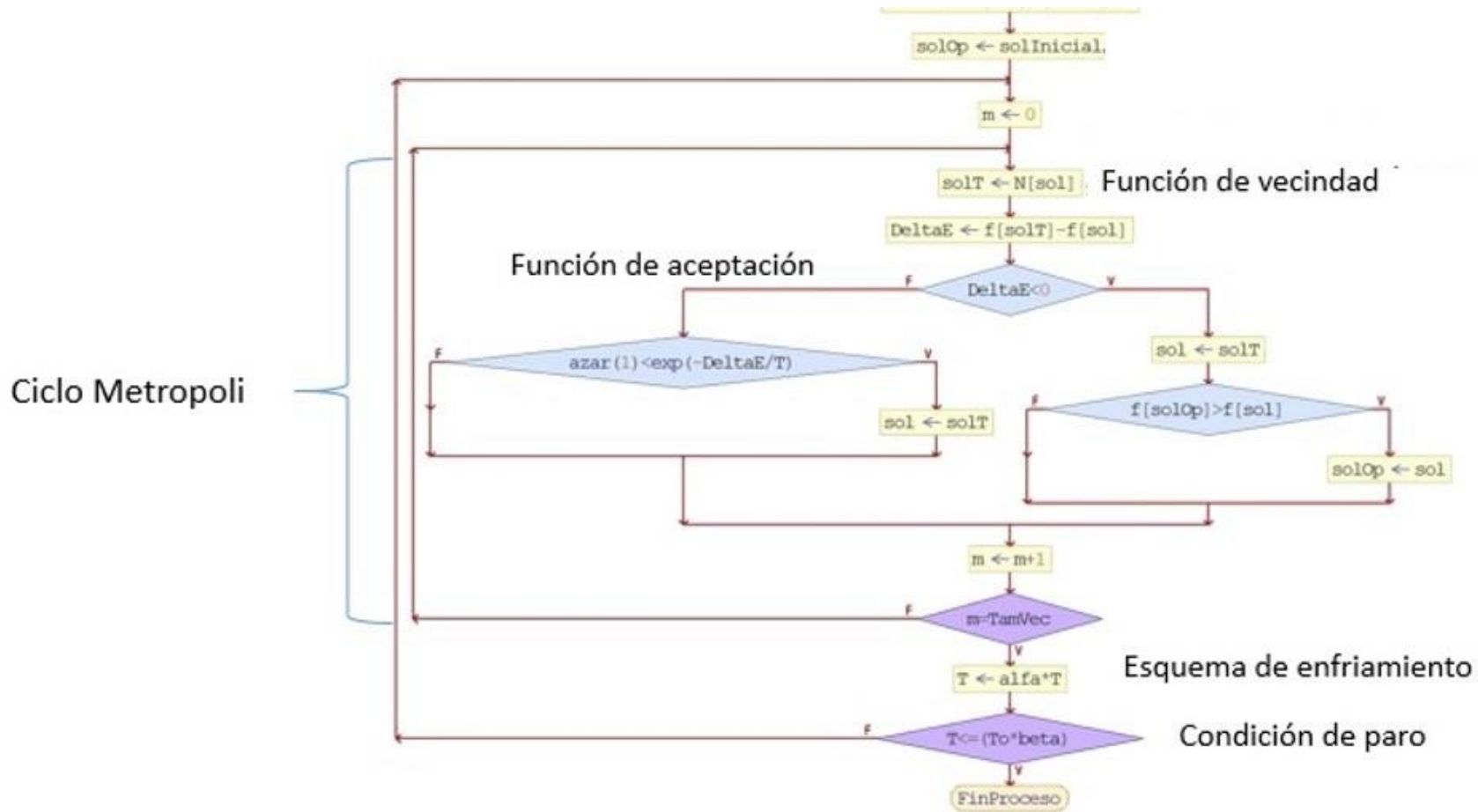


Figura 4.10 Ciclo Metr poli del RS.

**Encontrar soluciones óptimas locales.** En la siguiente gráfica se observa la convergencia de la heurística a la solución óptima local, se puede apreciar que conforme avanza el algoritmo, se encuentra una mejor solución minimizando la función  $f(x)$ .

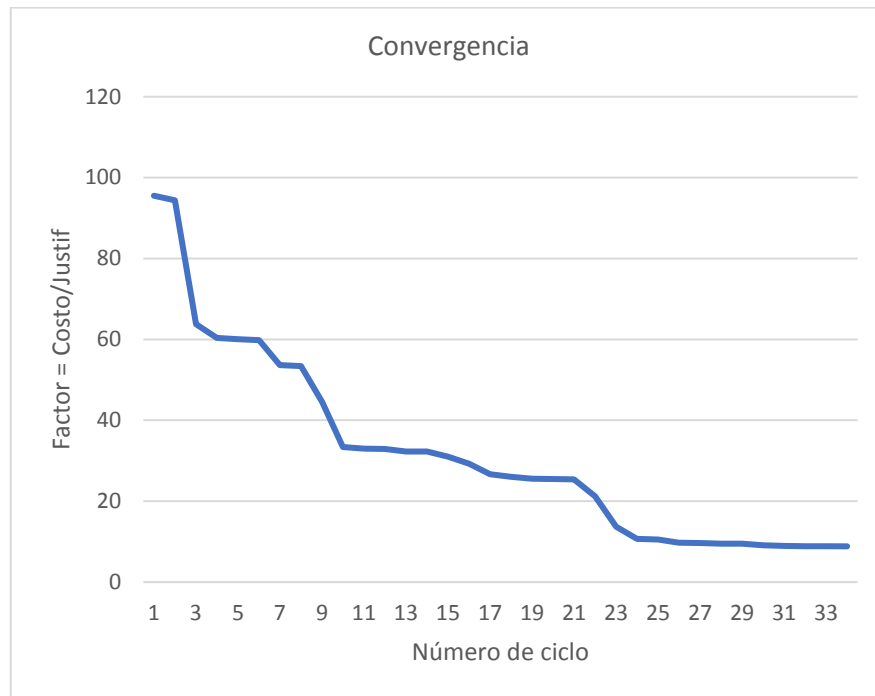


Figura 4.11 Convergencia del RS.

Se generan varias corridas que serán clasificadas y calificadas como en mejores, promedio y peores, para que, con los resultados obtenidos, se presenten varias opciones. Se aplicó el RS con un presupuesto amplio de 25 veces la nómina mensual de la empresa. Se realizaron 30 corridas y el algoritmo converge en un 58.2% a la misma solución en un tiempo promedio de 1.2 segundos. Dado que la empresa tiene una nómina de 60 mil pesos mensuales, el presupuesto autorizado máximo que determinó la empresa, fue de 1.10 veces la nómina, de manera que para la implementación de las salvaguardas el presupuesto es de \$66,000.

Entre las diversas soluciones cercanas al óptimo que arrojó el RS, se presenta la siguiente en el formato descrito en el capítulo 3, etapa IV, pág. 60 como

$$S=\{2261, 60.92, (1\ 0\ 0\ 0), (0\ 0\ 0\ 1), (0\ 1\ 0\ 0), (0\ 1\ 0\ 0), (0\ 1\ 0\ 0), (0\ 1\ 0\ 0), (0\ 1\ 0\ 0), (0\ 1\ 0\ 0), (0\ 0\ 0\ 0), (0\ 1\ 0\ 0), (0\ 0\ 0\ 1)\}$$

y en el formato  $\omega$  utilizado para la generación de soluciones vecinas del RS..

$$\omega=\{2261, 60.92, (1,4,2,2,2,2,2,2,0,2,4)\}$$

Se realizaron tres bloques obteniendo 90 soluciones cercanas al óptimo, en la figura 4.12 se observa que el valor de  $f(x)$  es menor que 10 unidades, por lo que todas las soluciones son cercanas al óptimo.

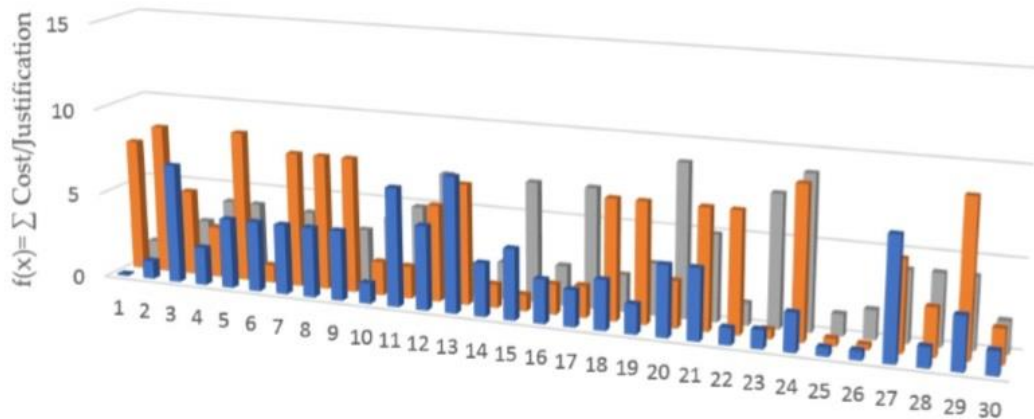


Figura 4.12: Tres pruebas de 30 soluciones casi óptimas de RS.

#### Etapas IV. Categorizar las soluciones.

Se procede ahora a categorizar las soluciones cercanas al óptimo generadas por el algoritmo RS, mediante el software R, aplicando la técnica de aprendizaje automático (machine learning) con k-means. La clasificación se realizó en 3 grupos, los promedios, desviación estándar, mínimo y máximo se muestran en la tabla 4.6.



Tabla 4.6 Promedios, desviación estándar, mínimo y máximo por Clúster.

	JT				CT				FACTOR			
	Prom	DS	Máy	Mín	Prom	DS	Máy	Mín	Prom	DS	Máy	Mín
<b>Clúster 1</b>	2249.1	37.1	2261.0	2061.0	61.5	2.0	65.5	54.3	5.8	1.8	8.7	3.2
<b>Clúster 2</b>	1331.6	281.0	1729.8	919.8	36.0	17.4	65.9	5.6	3.5	2.9	8.8	0.5
<b>Clúster 3</b>	300.5	231.2	767.5	33.8	47.8	19.7	65.5	5.9	2.2	1.4	5.2	0.1

Se observa en la gráfica de la figura 4.13 los grupos bien definidos, sin embargo, al observar los costos entre los grupos o clústeres, se encuentran algunas soluciones con costo mayor en el grupo 2 y con una justificación menor que el grupo 1.

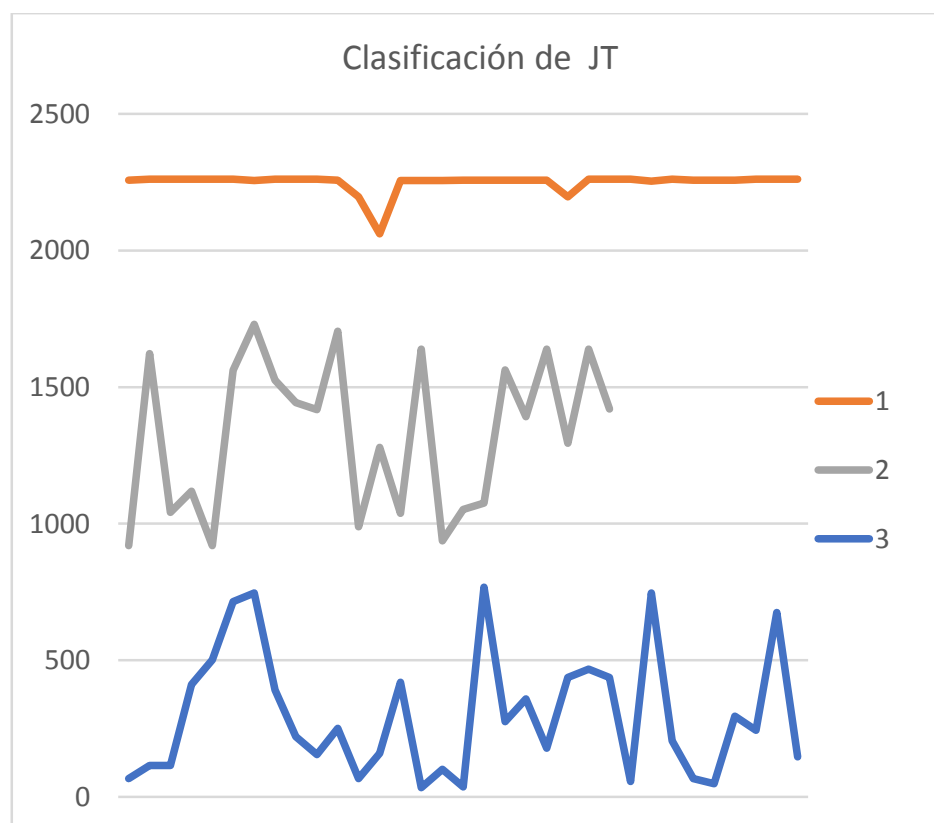


Figura 4.13 Grupos clasificados por la Justificación Total.

El estadístico F es calculado como el cociente de los cuadrados de las medias, como se indicó en la sección 2.6.1. El total de valor *within-cluter* es la suma de

cuadrados de las distancias euclidianas entre los elementos y su respectivo centroide, en R, la función es  $\$withinss$ , es un vector de las sumas de cuadrados *within-cluster*, una componente del clúster,  $E(C)=Sum(withinss)$  es la suma para todos los clústeres. La suma de cuadrados *between-cluster* en R es  $\$betweenss$ . Calculando  $F=17.4694$ , es un valor apropiado, ya que un buen modelo tiene un valor esperado de  $F>1$ . Esto indica que los clústeres están suficientemente diferenciados entre ellos.

Aplicando después árboles de decisión, se obtiene el punto de corte para la Justificación.

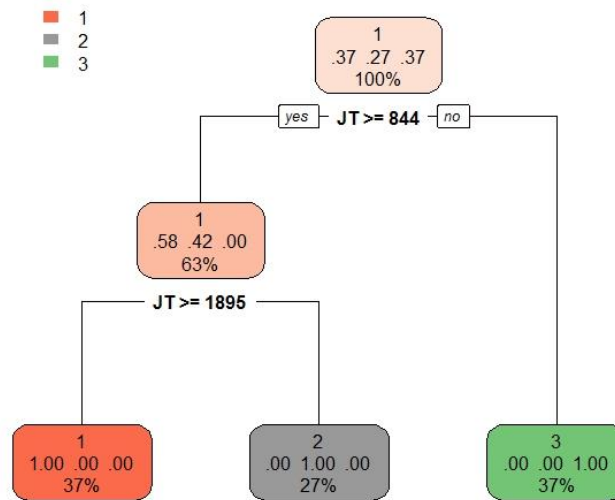


Figura 4.14 Árbol con la clasificación sobre el Total Justificación por Clúster.

Se interpreta de la siguiente manera: Si La Justificación Total,  $JT$ , de la solución de las salvaguardas seleccionadas es mayor que 1895 entonces las soluciones seleccionadas son del Clúster 1, en otro caso, si la Justificación esta entre 844 y 1895, las soluciones son del Clúster 2, por último, si la Justificación por la selección realizada es menor que 844, es del Clúster 3.

Finalmente, con el propósito de obtener los factores de riesgo relevantes para los grupos, se aplicó la heurística de árboles de decisión de nuevo, pero ahora sobre los factores y el clúster al que pertenece la solución. La entrada son los tipos de salvaguarda en texto para cada uno de los once factores de riesgo, la tabla 4.7, muestra algunos registros.

Tabla 4.7 Datos de entrada para obtener los tipos de respuesta más relevantes para cada factor. Ac- Aceptar; Mi - Mitigar; Tr - Transferir; AV - Evitar

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	Clúster
Null	Null	Null	Null	Null	Null	Null	Null	Null	Null	AV	C3
Ac	Null	Mi	Null	Mi	Tr	Null	Null	Null	Null	Tr	C2
Ac	AV	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Null	AV	C1

El árbol de decisión se muestra en la siguiente figura 4.15.

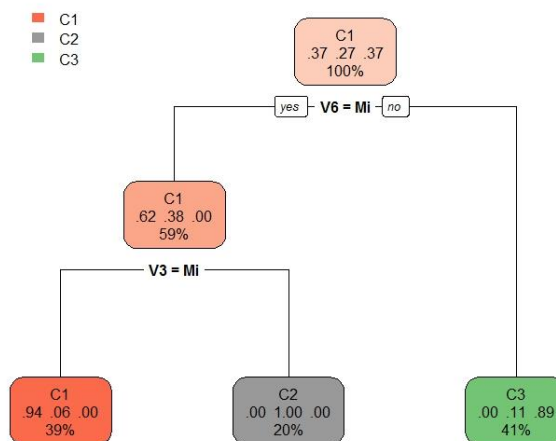


Figura 4.15 Respuestas Relevantes a los factores de riesgo por clúster.

Se interpreta así, si el factor de riesgo 6 se mitiga y el factor de riesgo 3 también se mitiga, entonces la solución está en el Clúster 1. Pero, si el factor 6 se mitiga y el factor 3 se acepta, transfiere o evita, la solución está en el Clúster

2, finalmente si el factor de riesgo 6 se acepta, transfiere o evita, la solución está en el Clúster 3. Con esta información, el gerente puede considerar las respuestas a los principales factores de riesgo.

#### Etapa V. Presentación de soluciones en la interfaz para el usuario.

Para visualizar las mejores soluciones cercanas al óptimo obtenidas, se elaboró una interfaz con tablas dinámicas en hojas de cálculo. La siguiente imagen, figura 4.16, muestra la carga a la interfaz, de tres pruebas de 30 ejecuciones de la heurística RS, las tablas dinámicas permiten escoger una solución y ver las salvaguardas seleccionadas y el tipo de respuesta para cada riesgo, muestra también el valor de la justificación total y el costo de implementación en miles de pesos.

1	Test	Run #	JT	CT	f(x)	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	Cluster
32	2	1	988.5	53.631264	7.670808	Ac	Ev	Mi	Null	Null	Tr	Null	Null	Mi	Mi	Tr	C2
33	2	2	2196	62.105927	8.678924	Ac	Ev	Mi	Null	Mi	Mi	Mi	Mi	Mi	Mi	Ev	C1
34	2	3	220	28.980467	4.967917	Null	Null	Null	Mi	Null	Tr	Null	Null	Mi	Null	Null	C3
35	2	4	154.75	29.839687	2.967629	Ac	Null	Null	Null	Null	Ev	Null	Null	Null	Mi	Ac	C3
36	2	5	2061	65.50737	8.73995	Ac	Null	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Null	C1
37	2	6	250.5	13.789864	1.028718	Mi	Ev	Null	Null	Null	Ev	Null	Null	Null	Null	Null	C3
38	2	7	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1
39	2	8	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1
40	2	9	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1

Figura 4.16 Interfaz con Tablas Dinámicas.

En la figura 4.17, hay una imagen de la primera hoja de cálculo en la que el gerente puede seleccionar una solución de cada grupo y ver las salvaguardas y el tipo de respuesta para cada factor de riesgo.



Figura 4.17. Interfaz de usuario para seleccionar y ver una solución.

Al seleccionar una solución del grupo 1, es decir, que ejercen casi todo el presupuesto con la mayor justificación, la solución local óptima tiene una justificación total de 2261, con un costo de 60.93 unidades monetarias y es la que acepta el factor de riesgo 1, evita 2 y 11 y mitiga el factor 3 a 8 y el factor de riesgo 10. La salvaguarda elegida para cada riesgo y el tipo de salvaguarda se encuentran en la tabla 4.8 con la que se puede desarrollar una versión inicial del plan de atención a los factores de riesgo.

Tabla 4.8 Solución elegida a partir de la interfaz figura 4.17.

<b>Id Factor</b>	<b>Factor de Riesgo</b>	<b>Salvaguarda</b>	<b>Tipo de Respuesta</b>
1	Administración de recursos humanos	Área de mesa de ayuda	Aceptar
2	Financiamiento Indirecto	Ahorro de utilidades para emergencias	Evitar
3	Eficacia de la red de venta	Elaborar plan publicidad, preparación equipo de ventas directas y plataforma tecnológica	Mitigar
4	Grado de competencia que tiene su producto/servicio	Elaborar plan de ampliar el mercado, nuevos tipos servicio, tener contacto clientes antiguos	Mitigar
5	Sensibilidad de afectación a su producto/servicio a los cambios en los factores macroeconómicos (inflación, desempleo, inversión extranjera, etc.)	Realizar un monitoreo diario oficial periódicamente	Mitigar
6	Planificación de negocios a corto y largo plazo	Elaborar plan estratégico, proyección anual, presupuesto	Mitigar
7	Financiamiento Directo	Habilitar depto. de cobranzas	Mitigar
8	Otras fuentes de financiamiento	Solicitar de 3 a 5 jóvenes "construyendo el futuro"	Mitigar
9	Estabilidad de las transacciones	Sin selección	Sin respuesta

10	Precio de venta y condición de liquidación	Implementar un plan de mejora de la comunicación con clientes para evitar doble visita	Mitigar
11	Información confiable para la toma de decisiones	Tener actualizado el costo de liquidación.	Evitar

Las soluciones cercanas al óptimo generadas por RS brindan a los gerentes más de una opción ajustada al presupuesto asignado. Se observa que algunas soluciones se repiten con la posibilidad de ser la más cercana a la óptima y podría ser la mejor decisión.

Los factores relevantes de las "mejores soluciones" las genera el árbol de decisión y es responder con la salvaguarda de mitigación para los factores seis y tres, es decir, preparar el plan estratégico con proyección presupuestal anual y preparar el plan de ventas y al equipo e incorporarlo en una plataforma tecnológica.



## Capítulo 5. Resultados Experimentales.

En este capítulo, para muestra de la eficiencia y robustez de la heurística RS, se realizan dos tipos de análisis: el primero del caso real de optimización de las respuestas a los factores de riesgo de la empresa y el segundo generando diez instancias de prueba de diversos tamaños, es decir, con 20, 32, 50, 100, 200, 300 y 400 factores de riesgo y con diferentes tipos de respuestas en las salvaguardas, 4, 6 y 8. Se contrasta la heurística con un algoritmo exacto (CPLEX) mostrando que la solución óptima, se obtiene en las soluciones generadas por RS y este utiliza menos recursos de memoria RAM. La computadora utilizada para estas pruebas es una i3 Intel de 4 Gb RAM con S.O. Windows 8.1 y Visual C++.

### 5.1 Caso de Estudio.

Para mostrar la eficiencia y robustez de la heurística RS, se realizó una verificación de su efectividad con un algoritmo exacto (CPLEX) en el software R. Con el presupuesto asignado por la empresa,  $P = 66$  unidades monetarias, el CPLEX no da una solución, ya que el óptimo global se encuentra con un presupuesto mínimo de 72.85 unidades monetarias y el conjunto de salvaguardas es el mismo que en la tabla 4.8, pero incluye la respuesta mitigar al factor V9,  $\omega = \{2264, 72.85, (1, 4, 2, 2, 2, 2, 2, 2, 2, 4)\}$ , el valor está en la tabla 5.1, donde se muestran los óptimos local generado por RS, todos estos son los clasificados por k-means y árbol de decisión, son del clúster 1.

Tabla 5.1 Soluciones Clúster 1.

<i>JT</i>	<i>CT</i>	<i>f(x)</i>	<i>Solución</i> <i>ω</i>	<i>Repeticiones</i> <i>Clúster 1</i>	<i>%</i> <i>(de 90 Sol)</i>
<b>2261</b>	<b>60.93</b>	<b>4.071</b>	{2261, 60.92, (1,4,2,2,2,2,2,0,2,4)}	<b>15</b>	<b>16.67%</b>
<b>2257</b>	<b>60.97</b>	<b>6.937</b>	{2257, 60.97, (1,4,2,2,2,2,2,0,4)}	<b>10</b>	<b>11.1%</b>
<b>2256</b>	<b>65.17</b>	<b>7.815</b>	{2256, 65.17, (1,4,2,2,2,2,0,2,2,4)}	<b>4</b>	<b>4.44%</b>
<b>2254</b>	<b>54.29</b>	<b>3.196</b>	{2254, 54.29, (2,4,2,2,2,2,0,0,2,4)}	<b>1</b>	<b>1.11%</b>
<b>2196</b>	<b>62.11</b>	<b>8.868</b>	{2196, 62.11, (1,4,2,0,2,2,2,2,2,4)}	<b>2</b>	<b>2.22%</b>
<b>2061</b>	<b>65.50</b>	<b>8.739</b>	{2061, 65.50, (1,0,2,2,2,2,2,2,2,0)}	<b>1</b>	<b>1.11%</b>

Uno de los principales objetivos de esta investigación fue implementar la heurística RS para optimizar las respuestas a los factores de riesgo, dando importancia en ofrecer múltiples soluciones para la mejor toma de decisiones con un presupuesto limitado y con los máximos beneficios para la empresa. Se demostró que la metodología es adecuada al realizar el análisis de los resultados en el caso de estudio, con las restricciones planteadas y la aplicación de la heurística RS. El conjunto de soluciones locales óptimas, se categoriza y simplifica con algoritmos de aprendizaje automático.

## 5.2 Desempeño Computacional del Recocido Simulado.

Las pruebas para heurísticas con datos artificiales se utilizan para evaluar la eficacia de éstas. Dentro de estos análisis, la comparación se puede realizar con un algoritmo exacto. En este caso, se continúa utilizando el solver CPLEX en R.

Con los datos reales se generan datos aleatorios artificiales, utilizando parámetros semejantes obtenidos del problema a resolver, esto con el fin de evaluar el comportamiento del algoritmo [120]. Los valores para los parámetros ingresados en el algoritmo RS se ajustaron con la desviación estándar, de acuerdo con el número de factor de riesgo,  $\alpha$ , y el número de tipos de respuesta

a cada factor,  $s$ . Los valores en  $T_0 = \sigma T$  y  $M = (a \times s)(a \times s - 1)$ , los parámetros  $\alpha$ ,  $\beta$ , son los mismos para todas las instancias,  $\alpha = 0.98$ ,  $\beta = 0.001$ .

Para estas diez instancias, la solución óptima se obtiene del software R con el solver CPLEX del paquete Rglpk [31] y slam [121]. Los datos de entrada de las instancias de prueba deben estar con formato del solver CPLEX, estos fueron generados con un algoritmo en C++ con la matriz de entrada  $M_{a \times s}$  utilizada por RS, capítulo 3, etapa IV de la metodología.

Tabla 5.2 Valor de la FO obtenida por las 10 instancias.

RFa -TRs	CPLEX			SA			
	Time sec	Memory Mb	O. F.	Time sec	Memory Mb	O. F.	RE
RF20-TR4	0.22	30.20	15.941530	1.67	0.40	15.941500	0.0000
RF20-TR6	0.23	30.40	14.588370	2.66	0.40	14.588370	0.0000
RF32-TR4	0.23	30.60	18.721720	5.52	0.40	18.721725	0.0000
RF32-TR6	0.23	30.50	16.473470	7.19	0.40	16.473473	0.0000
RF32-TR8	0.23	30.60	14.000400	11.38	0.40	14.000403	0.0000
RF50-TR8	0.24	30.90	11.165620	13.10	0.40	11.169275	0.0003
RF100-TR8	0.25	31.40	33.067110	56.11	0.50	33.083324	0.0005
RF200-TR8	0.35	32.00	38.589580	142.13	0.60	38.605324	0.0004
RF300-TR8	0.40	32.60	44.299110	277.36	0.80	44.331455	0.0007
RF400-TR8	0.48	33.10	55.076540	496.41	1.10	55.146198	0.0013

Se observa que, para las diez instancias, la heurística RS siempre encuentra una solución muy cercana al óptimo. El promedio de la prueba para cada instancia se realizó enfocado en el clúster 1 que es el que ofrece mayor beneficio ajustado al presupuesto  $P$  asignado, donde RS repite la solución óptima, tabla 5.1. Los errores relativos están por debajo de 0.04%. El mejor resultado obtenido fue para el modelo 20x6 y el peor resultado fue para el modelo 400x8.

Para medir el consumo de memoria (uso en RAM), la figura 5.1 muestra que el algoritmo CPLEX está en función del tamaño de la instancia de prueba, por otro lado, el uso de RS es un recurso constante que es realmente muy bajo.

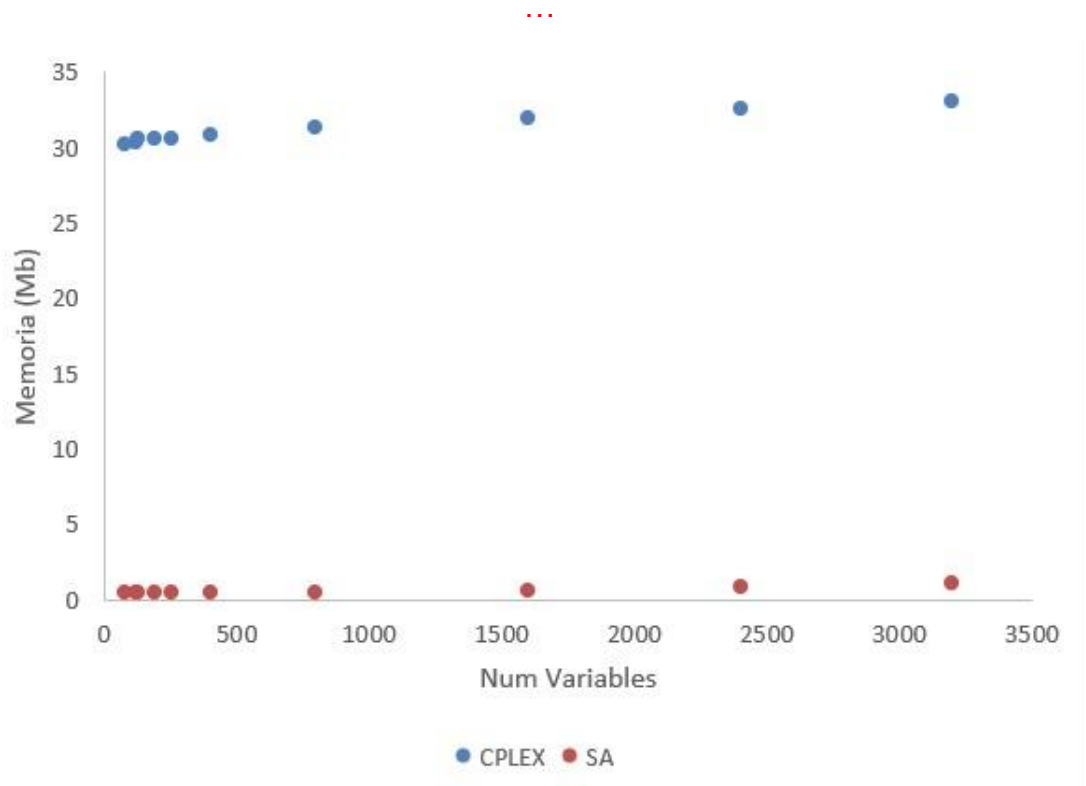


Figura 5.1 Consumo de memoria utilizado (Mb) por RS y CPLEX solver.

Para realizar el análisis estadístico y ver diferencias de comportamiento entre los métodos CPLEX y RS, primero se tiene que evaluar si existe normalidad de los resultados obtenidos por ambos métodos. Tomando como datos el valor de la *FO* obtenida por los 10 problemas, Tabla 5.2.

Se realiza el análisis de la normalidad mediante el gráfico de normalidad de los datos con la función *qqnorm* de R Software.

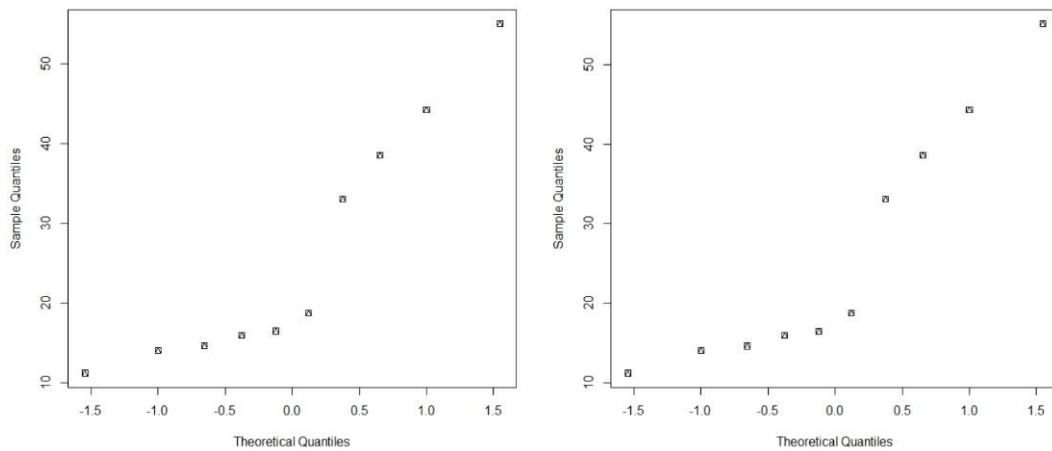


Figura 5.2 Prueba de Normalidad para los datos de CPLEX y RS.

La figura 5.2, muestra que no existe normalidad de los datos en los dos métodos ya que los puntos no se sitúan en la diagonal de los gráficos. Además, el comportamiento que se observa en ambas gráficas prácticamente es el mismo por ser los datos muy parecidos.

Se realiza el análisis de homocedasticidad con el gráfico de cajas y bigotes con la función `boxplot` de R Software.

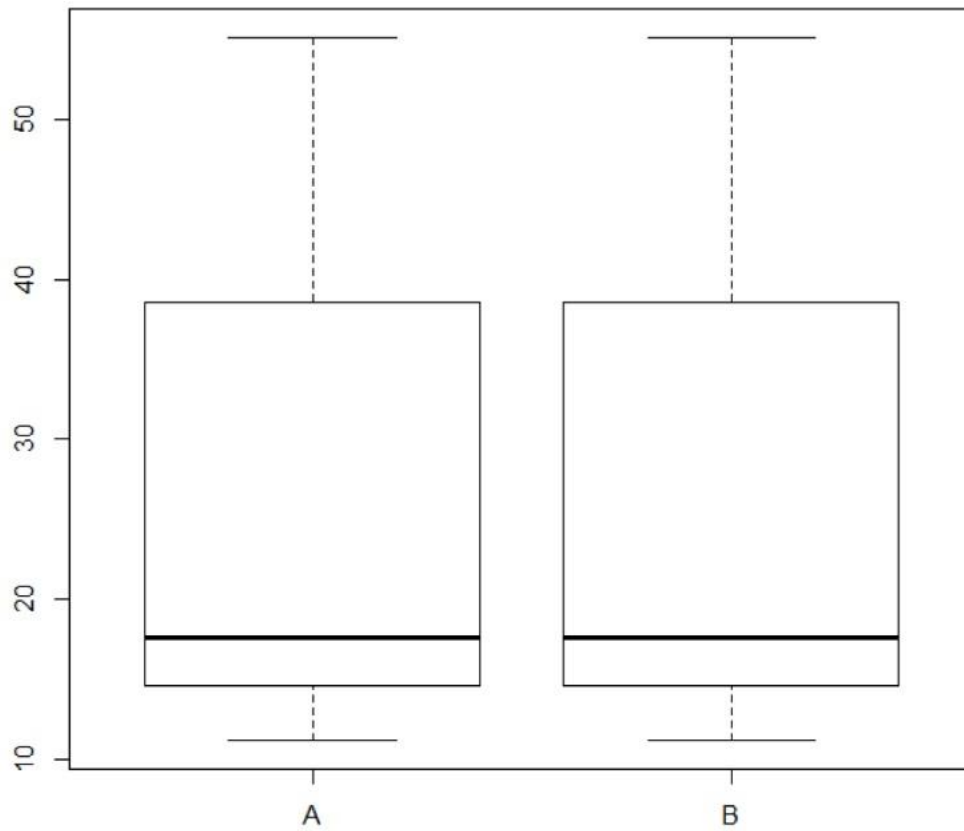


Figura 5.3 Análisis de Homocedasticidad de los datos de CPLEX y RS.

La figura 5.3, se observa que las cajas para cada método son iguales por lo que se puede admitir una varianza, esto implica que si existe homocedasticidad. Como no existe uno de los dos supuestos, en este caso normalidad de los datos (homocedasticidad si existe) para poder aplicar una prueba paramétrica como ANOVA clásico, entonces se puede aplicar ANOVA robusto con Welch y Box [122-124], este método también es paramétrico. Pero dado el comportamiento muy parecido de los datos obtenidos por los métodos CPLEX y RS, si se realiza el análisis estadístico de estos dos métodos con ANOVA robusto, el resultado será que el comportamiento de ambos métodos CPLEX y RS, son iguales ya que RS obtiene en sus mejores resultados para cada problema RFa-TRs, prácticamente el valor óptimo global, lo mismo que CPLEX, donde el peor valor de RS en RE es de 0.0013% (Ver Tabla 5.2). Este comportamiento de los resultados de la FO para problemas RFa -TRs con el

método de RS es muy importante, debido a que CPLEX es un método exacto que siempre obtiene la solución óptima global, siempre y cuando se cuente con los recursos computacionales necesarios, como se mostró en el caso del uso de la memoria RAM del equipo computacional y que para el caso de RS el comportamiento en el uso de memoria es muy diferente (Figura 5.1).

De esta manera, conociendo el valor de la solución óptima, podemos evaluar con precisión qué tan cerca se encuentran las soluciones generadas por la heurística para instancias de cualquier tamaño.

Con los resultados experimentales de las diez instancias y el procedimiento realizado en esta investigación, se observan los beneficios de la clasificación realizada por las técnicas de machine learning y nos enfocamos en el mejor grupo, obteniendo un conjunto de soluciones muy buenas y cercanas al óptimo con las restricciones del presupuesto, como nuestro caso de estudio, donde el presupuesto era inferior al presupuesto del valor óptimo, y donde el algoritmo exacto no obtiene una solución factible.

Con respecto al uso de memoria, se pueden observar los beneficios de las heurísticas no-deterministas, como RS. CPLEX utiliza algoritmos de Branch and Cut que realizan una enumeración sistemática de soluciones candidatas que se encuentran en nodo raíz del árbol [102,103], el algoritmo explora ramas y consume más memoria que la búsqueda local realizada por la heurística RS.

## Capítulo 6. Discusión de los Resultados

El desarrollo de esta investigación se enfocó en optimizar las respuestas de los factores de riesgo. Su importancia radica en apoyar el proceso de toma de decisión que implica la evaluación de diversas variables con información previa disponible. El uso de metodologías como la presentada, permite la sistematización de procesos evitando algunas veces que el juicio humano genere sesgos en la toma de decisiones. El caso de estudio se realizó para evaluar en qué medida la metodología propuesta es funcional para la detección, análisis, búsqueda y encontrar una serie de soluciones óptimas locales a los riesgos detectados. Cabe mencionar que la aplicación de la metodología y la solución encontrada es propia de la empresa, pues es función de: a) la veracidad de la información, b) las características de operación de la empresa, c) las necesidades expuestas por la empresa y d) las capacidades de la empresa. Los resultados experimentales muestran consistencia en la solución, lo que al momento implica que la metodología es eficiente.

1) Se desarrolla una nueva metodología considerando las seis áreas de riesgo más importantes en las empresas. Esta metodología se aplica en un caso de estudio real que lleva a cabo la identificación y evaluación de factores de riesgo. Se crea una matriz de salvaguardas aplicando el método William Fine. Al ejecutar el algoritmo RS, se obtienen múltiples soluciones cercanas al óptimo. Estas soluciones se clasifican utilizando un algoritmo de agrupación por los niveles de uso de recursos y sus beneficios. En la empresa del caso de estudio, se identifican los factores de riesgo  $V_6$  (planificación comercial a corto y largo plazo) y  $V_3$  (eficiencia de la red de ventas). Se tratan utilizando las salvaguardas  $S_{6,2}$  (Preparar plan estratégico, proyección anual) y  $S_{3,2}$  (Preparar el plan de ventas e incorporarlo en una plataforma tecnológica), respectivamente. A través de una interfaz de usuario para la toma de



decisiones, los gerentes del negocio pueden usar una herramienta informática que contribuya a la generación de un plan de atención a los factores de riesgo. La correcta administración de éste, generará beneficios que eviten el fracaso empresarial.

2) El modelo de optimización se resuelve utilizando el algoritmo RS que muestra ser eficaz al proporcionar soluciones casi óptimas con un presupuesto más bajo que el de la solución óptima alcanzada por el solver CPLEX.

En <https://drive.google.com/drive/folders/1V-cl21LrOoCOOQf9a9Cac1UPLhDSKHVf?usp=sharing> se encuentran archivos del caso de estudio con las corridas del CPLEX y el RS.

3) En las diez instancias generadas con datos artificiales, la eficacia y la efectividad del algoritmo RS se evalúan en contraste con la solución óptima generada por el solver CPLEX. RS obtiene el óptimo global en todos los casos en un tiempo razonable y consume menos recursos de memoria, lo cual se demuestra estadísticamente.

## Capítulo 7. Conclusiones y Trabajos Futuros

### 7.1 Conclusiones

Como resultado de este proyecto, se ha desarrollado una metodología para la optimización del tratamiento de riesgos y la búsqueda de soluciones cercanas al óptimo a través de la aplicación de la heurística RS y algoritmos de aprendizaje automático, la cual, fue comparada con otras técnicas computacionales (algoritmos exactos vs heurísticas) y evaluada con análisis estadístico, llegando a las conclusiones presentadas en esta sección.

La metodología desarrollada fue evaluada mediante su aplicación a un problema real de una empresa del sector de hidrocarburos y se comprobó su eficacia al identificar y reducir los efectos clasificados como perjudiciales en la empresa. Lo anterior, mediante la selección de las mejores soluciones a los factores de riesgo detectados arrojados por las técnicas utilizadas.

La metodología con el algoritmo de RS que minimiza el cociente del costo entre el beneficio a la empresa es útil para la generación del plan de atención a los riesgos empresariales de fracaso o quiebra, en las dimensiones de riesgo de la industria, riesgo operativo, riesgo administrativo, flexibilidad financiera, competitividad y cumplimiento. En la validación de la metodología en una empresa, se comprobó la eficacia del algoritmo con un algoritmo exacto contenido en CPLEX, los resultados obtenidos, tanto por el algoritmo exacto como por el RS, para minimizar el cociente del costo entre la justificación o beneficio, coinciden en la solución óptima.

El algoritmo RS proporciona un conjunto de soluciones cercanas al óptimo siendo esto una de las principales contribuciones y originalidad de esta investigación, donde se observó que, en caso de tener un presupuesto limitado, la solución óptima con el algoritmo CPLEX en R, no proporcionó

ninguna solución factible, sin embargo, con la metodología propuesta es posible obtener una solución muy cercana al óptimo. En el caso de estudio, esta solución contiene solo una diferencia en la salvaguarda asignada al factor  $V_9$ , la cual consiste en que no se le asigna ninguna acción, mientras que en la solución óptima generada por CPLEX, al asignar al factor de riesgo  $V_9$  una salvaguarda, se excede al presupuesto designado por la empresa (tablas 4.8 y 5.1).

El modelo minimiza la relación entre el costo y la justificación del beneficio al implementar la solución elegida, este cociente está de acuerdo con el principio ALARP, que indica que el riesgo debe ser minimizado, y ser tan bajo como sea razonablemente posible.

Aplicando la clasificación con la heurística k-medias, el análisis de RS se realiza con las soluciones del mejor grupo, clúster 1.

Sobre los objetivos: de esta investigación, la metodología ha sido evaluada con datos reales de las condiciones de riesgo en una empresa. También se desarrolló una interfaz gráfica para facilitar la toma de decisiones. En esta interfaz se carga el conjunto de soluciones clasificadas en tres grupos:

- Soluciones que consumen casi todo el presupuesto con la mayor justificación.
- Soluciones de bajo costo
- Soluciones con menor costo, pero con un nivel básico de justificación (seguridad mínima)

Con respecto al comportamiento del problema de tratamiento de factores de riesgo empresarial, no se presentan características de un problema tipo NP-Completo, sin embargo, la pendiente con respecto al consumo de recursos

computacionales, como lo es la memoria RAM, es menos pronunciada que el que presenta el algoritmo exacto, CPLEX.

## 7.2 Trabajos Futuros

Para trabajos futuros, la asignación de costos y beneficios a las salvaguardas podrán ser considerados dependientes entre ellas, esto es, que el grado de corrección de un factor de riesgo puede variar dependiendo del tratamiento que se ha seleccionado a los factores de riesgo previos. Esta modificación podría ser tratada con lógica difusa y un modelo de optimización bi-objetivo.

Esta metodología está diseñada para ajustar el número de respuestas a los elementos de riesgo. En el caso real de estudio, se usaron cuatro (aceptar, mitigar, transferir o evitar), que son los que se implementan en las mejores prácticas sugeridas por el PMI para el tratamiento de riesgo, pero puede expandirse o reducirse conforme a los requerimientos del problema. Esto permite que esta metodología pueda ser utilizada en diversos entornos donde la solución es estructuralmente equivalente. Además, se podrá ampliar el uso de esta metodología, en trabajos futuros, al integrar los factores que no presentan riesgo a la empresa pero que pueden promover el desarrollo de la misma al atenderlos con estrategias que impulsen las fortalezas y oportunidades.

## Referencias

- [1] Dorothy Gjerdrum y Mary Peter. The new international standard on the practice of risk management –a comparison of iso 31000: 2009 and the coso erm framework. *Risk management*, 31(21):8–12, 2011.
- [2] María Luisa Saavedra García y Blanca Tapia Sánchez. Mejores prácticas y factores de competitividad en las micro, pequeñas y medianas empresas mexicanas. *Economía*, (32):11–36, 2011.
- [3] Alfonso Eduardo Márquez Chamorro, Manuel Resinas, y Antonio Ruiz-Cortés. Predictive monitoring of business processes: a survey. *IEEE Transactions on Services Computing*, 11(6):962–977, 2017.
- [4] Enrique Ávila Heredia et al. Las pymes en México: desarrollo y competitividad. *Observatorio de la Economía Latinoamericana*, 201, 2014.
- [5] John R Graham, Hyunseob Kim, Si Li, y Jiaping Qiu. Human capital loss in corporate bankruptcy. *US Census Bureau Center for Economic Studies Paper No. CES-WP-13-37*, 2013.
- [6] Fred R David y Forest R David. *Strategic management: Concepts and cases: A competitive advantage approach*. Pearson, 2013.
- [7] María E Gómez-Gamero et al. Las habilidades blandas competencias para el nuevo milenio. *DIVULGARE Boletín Científico de la Escuela Superior de Actopan*, 6(11), 2019.
- [8] Sergio Javier Jasso Villazul. Empresas, violencia y competitividad en México: un creciente proceso de destrucción productiva y moral. In *XVII Congreso Internacional de Contaduría, Administración e Informática*, páginas 1–22, 2012.
- [9] 2014 COSO. *Enterprise Risk Management-Integrated Framework*. Committee of Sponsoring Organizations of the Treadway Commission, 2004.

- [10] Myoung-Jong Kim and Ingoo Han. The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms. *Expert Systems with Applications*, 25(4):637–646, 2003.
- [11] Rosalba Martínez Hernández, Pastor Pérez, y María del Pilar. Interrelación entre riesgo e innovación: percepción del riesgo por gestores de proyectos. *Journal of technology management & innovation*, 13(2):94–103, 2018.
- [12] Kenneth H Rose. A guide to the project management body of knowledge (pmbok®guide)fifth edition. *Project management journal*, 44(3):e1–e1,2013.
- [13] Damon Jones, David Molitor, y Julian Reif. What do workplace wellness programs do? evidence from the illinois workplace wellness study. Technical report, National Bureau of Economic Research, 2018.
- [14] Sara Marcelino-Sádaba, Amaya Pérez-Ezcurdia, Angel M Echeverría Lazcano, y Pedro Villanueva. Project risk management methodology for small firms. *International journal of project management*, 32(2):327–340,2014.
- [15] Grant Purdy. Iso 31000: 2009 setting a new standard for risk management. *Risk Analysis: An International Journal*, 30(6):881–886, 2010.
- [16] Robert R Moeller. *COSO enterprise risk management: understanding the new integrated ERM framework*. John Wiley & Sons, 2007.
- [17] William T Fine. Mathematical evaluations for controlling hazards. Technical report, NAVAL ORDNANCE LAB WHITE OAK MD, 1971.
- [18] Nataly Ruiz Quiñones, Oscar Castro Aguilar, y Martha Lily Ocampo Guerrero. Estimation of the risk associated to marketing of swine meat contaminated with salmonella spp., employing the William T. Fine method. *Procedia Food Science*, 7:137–140, 2016.

- [19] M Meknatjoo y M Omidvari. Safety risk assessment by using william–fine method with compilation fuzzy dematel in machining process. *Iran Occupational Health*, 12(5):31–42, 2015.
- [20] M Ebrahimzadih, GH Halvani, E Darvishi, y F Froghinasab. Application of job safety analysis and william fine methods to identify and control hazards in a uranium mine in central area of iran. *Journal of Health*, 6(3):313–324, 2015.
- [21] Gholamhosein Halvani, Mohammad Hasan Ehrampoush, Mohammdd Taghi Ghaneian, Ali Dehghani, y Mohsen Hesami Arani. Applying job hazard analysis and William Fine methods on risks identification and assessment of jobs in hot rolling steel, iran. *Journal of Mazandaran University of Medical Sciences*, 26(145):293–303, 2017.
- [22] Derrick Nanda Tchiehe y Francois Gauthier. Classification of risk acceptability and risk tolerability factors in occupational health and safety. *Safety science*, 92:138–147, 2017.
- [23] Richard Maguire. *Safety cases and safety reports: meaning, motivation and management*. CRC Press, 2017.
- [24] MH Ehrampoush, GH Halvani, MT Ghaneian, A Dehghani, y M Hesami Arani. Usage of williamfine, pha methods for environmental risks management and estimation of control costs in the kavir steel complex. *Tolooebehdasht*, 15(1):10–25, 2016.
- [25] Thomas L Norman et al. *Risk analysis and security countermeasure selection*. CRC press, 2016.
- [26] Pierre Bonami y A Tramontani. Recent improvement to misocp in cplex, 2017.
- [27] Stefan Eggenschwiler. Rs: Parameter tuning for the cplex. *Bachelor Thesis*, 2016.

- [28] Michael Barry, Hubert Abgottspon, y René Schumann. Solver tuning and model configuration. In *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*, páginas 141–154. Springer, 2018.
- [29] Stefan Eggenschwiler, Michael Barry, y René Schumann. Exploring effects of parameter configurations on runtime, using mathematical solvers. In *Proceedings of the 31. Workshop "Planen, Scheduling und Konfigurieren, Entwerfen" (PuK 2018), n°CONFERENCE*. 24 September 2018, 2018.
- [30] R Core Team et al. R: A language and environment for statistical computing. 2013.
- [31] Kurt Hornik, Stefan Theussl, y Maintainer Stefan Theussl. Package 'rglpk'.2012.
- [32] Ricardo Lima y EWO Seminar. Ibm ilog cplex-what is inside of the box? In *Proc. 2010 EWO Seminar*, páginas 1–72, 2010.
- [33] Alessandro Berti. Statistical sampling in process mining discovery. In *The 9th International Conference on Information, Process, and Knowledge Management*, páginas 41–43, 2017.
- [34] Jia Hui Liang, Vijay Ganesh, Pascal Poupart, y Krzysztof Czarnecki. Learning rate based branching heuristic for sat solvers. In *International Conference on Theory and Applications of Satisfiability Testing*, páginas 123–140. Springer, 2016.
- [35] Stelios H Zanakis y James R Evans. Heuristic "optimization": Why, when, and how to use it. *Interfaces*, 11(5):84–91, 1981.
- [36] Jan Kronqvist, David E Bernal, Andreas Lundell, y Ignacio E Grossmann. A review and comparison of solvers for convex minlp. *Optimization and Engineering*, 20(2):397–455, 2019.



- [37] Rimmi Anand, Divya Aggarwal, y Vijay Kumar. A comparative analysis of optimization solvers. *Journal of Statistics and Management Systems*, 20(4):623–635, 2017.
- [38] Rui He, Xinhong Li, Guoming Chen, Guoxing Chen, y Yiwei Liu. Generative adversarial network-based semi-supervised learning for real-time risk warning of process industries. *Expert Systems with Applications*, 150:113244, 2020.
- [39] Valentina Viduto, Carsten Maple, Wei Huang, y David López-Peréz. A novel risk assessment and optimisation model for a multi-objective network security countermeasure selection problem. *Decision Support Systems*, 53(3):599–610, 2012.
- [40] Faisal Aqlan y Sarah S Lam. Supply chain risk modelling and mitigation. *International Journal of Production Research*, 53(18):5640–5656, 2015.
- [41] K Vergidis. Business process optimisation using an evolutionary multi-objective framework. 2008.
- [42] Kostas Vergidis, Christopher Turner, Alex Alechnovic, y Ashutosh Tiwari. An automated optimisation framework for the development of reconfigurable business processes: a web services approach. *International Journal of Computer Integrated Manufacturing*, 28(1):41–58, 2015.
- [43] Kostas Georgoulakos, Kostas Vergidis, George Tsakalidis, Nikolaos Samaras. Evolutionary multi-objective optimization of business process designs with pre-processing. In *2017 IEEE Congress on Evolutionary Computation (CEC)*, páginas 897–904. IEEE, 2017.
- [44] Jiri Tupa, Jan Simota, y Frantisek Steiner. Aspects of risk management implementation for industry 4.0. *Procedia Manufacturing*, 11:1223–1230, 2017.

- [45] Antoni-Lluís Mesquida y Antonia Mas. A project management improvement program according to iso/iec 29110 and pmbok®. *Journal of Software: Evolution and Process*, 26(9):846–854, 2014.
- [46] Marika Arena, Michela Arnaboldi, y Giovanni Azzone. The organizational dynamics of enterprise risk management. *Accounting, Organizations and Society*, 35(7):659–675, 2010.
- [47] Olivier C Martin y Steve W Otto. Combining simulated annealing with local search heuristics. *Annals of Operations Research*, 63(1):57–75, 1996.
- [48] Tim Furche, Georg Gottlob, Leonid Libkin, Giorgio Orsi, y Norman WPaton. Data wrangling for big data: Challenges and opportunities. In *EDBT*, tomo 16, páginas 473–478, 2016.
- [49] Mahfoud Bala, Omar Boussaid, y Zaia Alimazighi. A fine-grained distribution approach for etl processes in big data environments. *Data & Knowledge Engineering*, 111:114–136, 2017.
- [50] John Meehan, Cansu Aslantas, Stan Zdonik, Nesime Tatbul, y Jiang Du. Data ingestion for the connected world. In *CIDR*, 2017.
- [51] Varun Grover, Roger HL Chiang, Ting-Peng Liang, y Dongsong Zhang. Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, 35(2):388–423,2018.
- [52] Aleš Popovič, Ray Hackney, Rana Tassabehji, y Mauro Castelli. The impact of big data analytics on firms' high value business performance. *Information Systems Frontiers*, 20(2):209–222, 2018.
- [53] David Hillson y Ruth Murray-Webster. *Understanding and managing risk attitude*. Routledge, 2017.

- [54] Dian Agustia, Nur Pratama Abdi Muhammad, y Yani Permatasari. Earnings management, business strategy, and bankruptcy risk: evidence from indonesia. *Heliyon*, 6(2): e03317, 2020.
- [55] Shu Ling Chiang y Ming Shann Tsai. The valuation of deposit insurance allowing for the interest rate spread and early-bankruptcy risk. *The Quarterly Review of Economics and Finance*, 2019.
- [56] Mürüvvet Büyükboyacı, Mehmet Y Gürdal, Arzu Kıbrıs, y Özgür Kıbrıs. An experimental study of the investment implications of bankruptcy laws. *Journal of Economic Behavior & Organization*, 158:607–629, 2019.
- [57] She-I Chang, Albert Huang, Li-Min Chang, y Jhan-Cyun Liao. Risk factors of enterprise internal control: Governance refers to internet of things (iot) environment. In *Pacific Asia Conference On Information Systems (PACIS)*. Association For Information System, 2016.
- [58] Edward I Altman y Edith Hotchkiss. Corporate financial distress and bankruptcy. 1993.
- [59] Edward I Altman, Małgorzata Iwanicz-Drozdowska, Erkki K Laitinen, y Arto Suvas. Financial distress prediction in an international context: A review and empirical analysis of altman's z-score model. *Journal of International Financial Management & Accounting*, 28(2):131–171, 2017.
- [60] Ricarda B Bouncken y Viktor Fredrich. Good fences make good neighbors? directions and safeguards in alliances on business model innovation. *Journal of Business Research*, 69(11):5196–5202, 2016.
- [61] Rachida Parks, Heng Xu, Chao-Hsien Chu, y Paul Benjamin Lowry. Examining the intended and unintended consequences of organisational privacy safeguards. *European Journal of Information Systems*, 26(1):37–65, 2017.
- [62] Céline Bak. Generating growth from innovation for the low-carbon economy: Exploring safeguards in finance and regulation. 2017.

- [63] Michael R Garey y David S Johnson. “strong”np-completeness results: Motivation, examples, and implications. *Journal of the ACM ( JACM)*, 25(3):499–508, 1978.
- [64] Christos H Papadimitriou. *Computational complexity*. John Wiley andSons Ltd., 2003.
- [65] Zixiang Li, Ibrahim Kucukkoc, y J Mukund Nilakantan. Comprehensive review and evaluation of heuristics and meta-heuristics for two-sided assembly line balancing problem. *Computers & Operations Research*, 84:146–161, 2017.
- [66] Saied Asghari y Nima Jafari Navimipour. Review and comparison of meta-heuristic algorithms for service composition in cloud computing. *Majlesi Journal of Multimedia Processing*, 4(4):28–34, 2016.
- [67] Singiresu S Rao. *Engineering optimization: theory and practice*. John Wiley & Sons, 2019.
- [68] David E Goldberg. *Genetic algorithms*. Pearson Education India, 2006.
- [69] Anupriya Shukla, Hari Mohan Pandey, y Deepti Mehrotra. Comparative review of selection techniques in genetic algorithm. In *2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)*, páginas 515–519. IEEE, 2015.
- [70] Allen Newell, J Clifford Shaw, and Herbert A Simon. The processes of creative thinking. In *Contemporary Approaches to Creative Thinking, 1958, University of Colorado, CO, US; This paper was presented at the aforementioned symposium*. Atherton Press, 1962.
- [71] Allen Newell, John Calman Shaw, y Herbert A Simon. Elements of a theory of human problem solving. *Psychological review*, 65(3):151, 1958.

- [72] LM Rasdi Rere, Mohamad Ivan Fanany, y Aniati Murni Arymurthy. Simulated annealing algorithm for deep learning. *Procedia Computer Science*, 72(1):137–144, 2015.
- [73] Martín G Martínez-Rangel, Marco Antonio Cruz-Chávez, J Crispín Zavala-Díaz, David Juárez-Romero, y Ocotlán Díaz-Parra. Analysis of the simulated annealing convergence in function of the standard deviation and the boltzmann quotient for scheduling problems. *Research in Computing Science, IPN, ISSN*, 4069:282–293, 2007.
- [74] Mark Jerrum y Alistair Sinclair. The Markov chain Monte Carlo method: an approach to approximate counting and integration, approximation algorithms for np-hard problems, 1996.
- [75] Nicola Paltrinieri, Louise Comfort, y Genserik Reniers. Learning about risk: Machine learning for risk assessment. *Safety science*, 118:475–486, 2019.
- [76] Mark Stephenson, Saman Amarasinghe, Martin Martin, and Una May O'Reilly. Meta optimization: improving compiler heuristics with machine learning. In *ACM SIGPLAN Notices*, tomo 38, páginas 77–90. ACM, 2003.
- [77] Felix Agakov, Edwin Bonilla, John Cavazos, Björn Franke, Grigori Fursin, Michael FP O'Boyle, John Thomson, Marc Toussaint, y Christopher KI Williams. Using machine learning to focus iterative optimization. In *Proceedings of the international symposium on code generation and optimization*, páginas 295–305. IEEE Computer Society, 2006.
- [78] Andika Rachman y RM Chandima Ratnayake. Machine learning approach for risk-based inspection screening assessment. *Reliability Engineering & System Safety*, 185:518–532, 2019.
- [79] Farrokh Farrokhi, Quinlan D Buchlak, Matt Sikora, Nazanin Esmaili, Maria Marsans, Pamela McLeod, Jamie Mark, Emily Cox, Christine

- Bennett, y Jonathan Carlson. Investigating risk factors and predicting complications in deep brain stimulation surgery with machine learning algorithms. *World Neurosurgery*, 134: e325–e338, 2020.
- [80] Horace B Barlow. Unsupervised learning. *Neural computation*, 1(3):295–311, 1989.
- [81] Erik Brynjolfsson y ANDREW McAfee. The business of artificial intelligence. *Harvard Business Review*, páginas 1–20, 2017.
- [82] Saumya Chaturvedi, Vimal Mishra, y Nitin Mishra. Sentiment analysis using machine learning for business intelligence. In *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, páginas 2162–2166. IEEE, 2017.
- [83] Anil K Jain. Data clustering: 50 years beyond k-means. *Pattern recognition letters*, 31(8):651–666, 2010.
- [84] Marco Capó, Aritz Pérez, y Jose A Lozano. An efficient approximation to the k-means clustering for massive data. *Knowledge-Based Systems*, 117:56–69, 2017.
- [85] Bo Yang, Xiao Fu, Nicholas D Sidiropoulos, y Mingyi Hong. Towards k-means-friendly spaces: Simultaneous deep learning and clustering. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, páginas 3861–3870. JMLR. org, 2017.
- [86] Purnima Bholowalia y Arvind Kumar. Ebk-means: A clustering technique based on elbow method and k-means in wsn. *International Journal of Computer Applications*, 105(9), 2014.
- [87] Leon Bottou y Yoshua Bengio. Convergence properties of the k-means algorithms. In *Advances in neural information processing systems*, páginas 585–592, 1995.

- [88] Greg Hamerly y Charles Elkan. Alternatives to the k-means algorithm that find better clusterings. In *Proceedings of the eleventh international conference on Information and knowledge management*, páginas 600–607, 2002.
- [89] Tapas Kanungo, David M Mount, Nathan S Netanyahu, Christine DPiatko, Ruth Silverman, and Angela Y Wu. A local search approximation algorithm for k-means clustering. In *Proceedings of the eighteenth annual symposium on Computational geometry*, páginas 10–18, 2002.
- [90] Dan Pelleg, Andrew W Moore, et al. X-means: Extending k-means with efficient estimation of the number of clusters. In *icml*, tomo 1, páginas 727–734, 2000.
- [91] Xindong Wu, Vipin Kumar, J Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J McLachlan, Angus Ng, Bing Liu, S Yu Philip, et al. Top 10 algorithms in data mining. *Knowledge and information systems*, 14(1):1–37, 2008.
- [92] Wang Bao, Ning Lianju, and Kong Yue. Integration of unsupervised and supervised machine learning algorithms for credit risk assessment. *Expert Systems with Applications*, 128:301–315, 2019.
- [93] Shengqiang Chi, Xinhang Li, Yu Tian, Jun Li, Xiangxing Kong, Kefeng Ding, Chunhua Weng, and Jingsong Li. Semi-supervised learning to improve generalizability of risk prediction models. *Journal of biomedical informatics*, 92:103117, 2019.
- [94] Joseph F McCarthy and Wendy G Lehnert. Using decision trees for coreference resolution. *arXiv preprint cmp-lg/9505043*, 1995.
- [95] Ameer Almomani, Paula Saavedra, and Eduardo Sánchez. Ensembles of decision trees for recommending touristic items. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, páginas 510–519. Springer, 2017.

- [96] Kazeem Adesina Dauda, Biswabrata Pradhan, B Uma Shankar, and Sushmita Mitra. Decision tree for modeling survival data with competing risks. *Biocybernetics and Biomedical Engineering*, 39(3):697–708, 2019.
- [97] Yilong Cao y Peter I Rockett. The use of vicinal-risk minimization for training decision trees. *Applied Soft Computing*, 31:185–195, 2015.
- [98] BN Lakshmi, TS Indumathi, y Nandini Ravi. A study on c. 5 decision tree classification algorithm for risk predictions during pregnancy. *Procedia Technology*, 24:1542–1549, 2016.
- [99] Leo Breiman. *Classification and regression trees*. Routledge, 2017.
- [100] Harry R Lewis. Michael R. Garey and David S. Johnson. computers and intractability. a guide to the theory of np-completeness. wh freeman and company, san francisco 1979, x+ 338 pp. *The Journal of Symbolic Logic*, 48(2):498–500, 1983.
- [101] Herbert S Wilf. *Algorithms and complexity*. AK Peters/CRC Press, 2002.
- [102] Tobias Achterberg, Thorsten Koch, y Alexander Martin. Branching rules revisited. *Operations Research Letters*, 33(1):42–54, 2005.
- [103] John E Mitchell. Branch-and-cut algorithms for combinatorial optimization problems. *Handbook of applied optimization*, 1:65–77, 2002.
- [104] Mauricio GC Resende y Celso C Ribeiro. Greedy randomized adaptive search procedures. In *Handbook of metaheuristics*, páginas 219–249. Springer, 2003.
- [105] Marco Antonio Cruz-Chávez, Fredy Juárez-Pérez, Erika Yesenia Ávila-Melgar, y Alina Martínez-Oropeza. Simulated annealing algorithm for the weighted unrelated parallel machines problem. In *2009 Electronics, Robotics and Automotive Mechanics Conference (CERMA)*, páginas 94–99. IEEE, 2009.



- [106] Dan Lovallo y Daniel Kahneman. Delusions of success. *Harvard business review*, 81(7):56–63, 2003.
- [107] Richard H Thaler. The psychology and economics conference handbook: Comments on Simon, on Einhorn and Hogarth, and on Tversky and Kahneman. *The Journal of Business*, 59(4): S279–S284, 1986.
- [108] Daniel Kahneman, Dan Lovallo, and Olivier Sibony. Before you make that big decision. *Harvard business review*, 89(6):50–60, 2011.
- [109] Daniel Kahneman, DAN LOVALLO, and Olivier Sibony. Un enfoque estructurado para las decisiones estratégicas. *Harvard Deusto business review*, (293):6–16, 2019.
- [110] Daniel Kahneman, Dan Lovallo, Olivier Sibony, Andrew Torrance, and Christiana von Hippel. *A structured approach to strategic decisions*. MIT Sloan Management Review, 2019.
- [111] Hugo Lindborg. Improving decision making in the digital era: Human and machine. 2018.
- [112] Marta Lilia Eraña-Díaz, Marco Antonio Cruz-Chávez, Rafael Rivera-López, Beatriz Martínez-Bahena, Erika Yesenia Ávila-Melgar, Martín H. Cruz-Rosales (en prensa). Optimization for Risk Decision-Making through Simulated Annealing, in *IEEE Access*, 2020.
- [113] Ghislain Houben, Kwan Lenie, y Koen Vanhoof. A knowledge-based swot-analysis system as an instrument for strategic planning in small and medium sized enterprises. *Decision support systems*, 26(2):125–135, 1999.
- [114] Martin Maechler, Peter Rousseeuw, Anja Struyf, Mia Hubert, y Kurt Hornik. cluster: *Cluster Analysis Basics and Extensions*, 2019. R package version 2.1.0 — For new features, see the 'Changelog' file (in the package source).

- [115] Terry Therneau, Beth Atkinson, Brian Ripley, and Maintainer Brian Ripley. Package 'rpart'. Available online: [cran.r-project.org/web/packages/rpart/rpart.pdf](http://cran.r-project.org/web/packages/rpart/rpart.pdf) (accessed on 20 April 2016), 2015.
- [116] Stephen Milborrow. Plotting rpart trees with the rpart. plot package, 2016.
- [117] Felicidad Marqués. *Microsoft EXCEL 2010. En profundidad*. RC Libros, 2010.
- [118] Scott Kirkpatrick, C Daniel Gelatt, y Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.
- [119] Giorgio Ausiello, Pierluigi Crescenzi, Giorgio Gambosi, Viggo Kann, Alberto Marchetti-Spaccamela, y Marco Protasi. *Complexity and approximation: Combinatorial optimization problems and their approximability properties*. Springer Science & Business Media, 2012.
- [120] Ronald L Rardin y Reha Uzsoy. Experimental evaluation of heuristic optimization algorithms: A tutorial. *Journal of Heuristics*, 7(3):261–304,2001.
- [121] Margaret E Roberts, Brandon M Stewart, Dustin Tingley, et al. stm: R package for structural topic models. *Journal of Statistical Software*, 10(2):1–40, 2014
- [122] Hangcheng Liu. Comparing welch's anova, a kruskal-wallis test and traditional anova in case of heterogeneity of variance. 2015.
- [123] Rand R Wilcox. *Introduction to robust estimation and hypothesis testing*. Academic press, 2011.
- [124] Patrick Mair y Rand Wilcox. Robust statistical methods in r using the wrs2 package. *Behavior research methods*, páginas 1–25, 2019

## Apéndice A. Cuestionario sobre Factores de Riesgo Empresarial



Cuestionario num. \_\_\_\_\_

Fecha \_\_\_\_\_

### CUESTIONARIO SOBRE FACTORES DE RIESGO EMPRESARIAL

Este cuestionario es parte de un estudio sobre los factores externos e internos que enfrentan las empresas. El uso de la información proporcionada y sus resultados serán exclusivamente para fines académicos. Los datos proporcionados son confidenciales.

Por favor conteste todas las preguntas.

Tome su tiempo. No hay límite de tiempo para responder a este instrumento.

Seleccione solo una respuesta, es un formato FODA.

Marque una "X" si considera que el factor son fortalezas, debilidades, oportunidades o amenazas.

		Interno / Externo	F Strengths	D Weaknesses	O Opportunities	A Threats
<b>i. Riesgo de la Industria (IR) :</b>						
1	Políticas gubernamentales y acuerdos internacionales	E				
2	Ciclicidad. Es cuando en el negocio las ganancias fluctúan ampliamente según las variaciones en la economía o el ciclo de las estaciones (por temporada)	E				
3	Grado de competencia	E				
4	Estabilidad en el mercado de suministros y costos de estos	E				
5	Tamaño y crecimiento del mercado (demanda de su producto/servicio)	E				
6	Sensibilidad a los cambios en los factores macroeconómicos	E				
7	Poder competitivo nacional e internacional	E				
8	Ciclo de vida de producto	E				
<b>ii. Riesgo Administrativo (MR):</b>						
9	Capacidad y competencia en la gestión (administración)					
10	Estabilidad en el área administrativa	I				
11	Relación entre la gerencia / propietario	I				
12	Administración de recursos humanos	I				
13	Crecimiento / desempeño comercial	I				
14	Planificación de negocios a corto y largo plazo	I				
15	Logros y viabilidad	I				

	S	W	O	T
Intern al/Ext ernal	Stren gths	Wea kness es	Oppo rtunit ies	Threa ts

iii.Flexibilidad Financiera (FF):					
16	Financiamiento Directo	I			
17	Financiamiento Indirecto.	E			
18	Otras fuentes de financiamiento	E			

iv.Credibilidad (CR):					
19	Historial Crediticio	I			
20	Información confiable	E			
21	Relación con instituciones financieras	E			
22	Posición en el mercado				
23	Nivel de Capacidad (core business)	I			
24	Estrategia diferenciada	I			
24	Differentiated strategy	I			

vi.Riesgo Operativo (OP):					
25	Estabilidad y diversidad de las adquisiciones (compras)	I			
26	Estabilidad de las transacciones	I			
27	Eficiencia en la producción	I			
28	Perspectivas de la demanda de productos y servicios	I			
29	Diversificación de ventas	I			
30	Precio de venta y condición de liquidación	I			
31	Cartera vencida (clientes y/o proveedores)	I			
32	Eficacia de la red de venta	I			

Gracias por su participación.

Apéndice B. Código fuente de la metaheurística Recocido Simulado.

```
/* Programa realizado para resolver la FO del Modelo de Tratamiento de
Factores Riesgo
Marta Lilia Eraña Díaz
*/
#define TPre 0.01
#define alfa 0.98
#define beta 0.001
#define SDesv 23.75

#define C 12
#define NItera 10
#define Presup 66.0
#define PresMin 60.0
#define amenaz 11
#define salvag 4

#include <stdio.h>
#include <stdlib.h>
#include <conio.h>
#include <time.h>
#include <math.h>

typedef struct
{
    float JT;
    float CT;
    float Fc;
    int x[amenaz+1][salvag+1];
```

```

}solucion;

solucion Sol, SolT, SolOp;

void coef_matriz(int *, int *, float(*)[C]);
void sec_NPre2(int,int, int(*)[salvag+1], int(*)[salvag+1], float(*)[C]);
void sec_Inicial(int, int , int(*)[salvag+1], float(*)[C]);

FILE *fpG;
FILE *fInter;
FILE *fpG2;
int ireng, jsalv, opc;
void main()
{
/* Programa Recocido Simulado minimizando ratio de costo / beneficio, ALARP
riesgo*/
long start;
float tOPTi;
float ACoef[C][C];
int XSel[amenaz+1][salvag+1];
int XSeN[amenaz+1][salvag+1];
int m,corridas;
int Nsolucion,i,j, Nveci,NumAc;
float T0, DeltaEne, nAlea, T, TamVeci, EnerIni, EnergAc;
int a , s ;
int xaleat, posi2,bande,acepta,acepOP;
float sumaJT, sumaCT, sFac;
srand(time(NULL));
fpG=fopen("SRDV.txt","w");
fpG2=fopen("SRSol.txt","w");

```

```

fInter=fopen("SRIn.txt","w");
start = clock();
coef_matriz(&a,&s,ACoef);
TamVeci=(a*s)* (a*s-1);
printf("\nTamaño Vecindad %6.0f",TamVeci);
printf("\nRestriccion de Vulnerabilidad atendida\n1.Si \n2.No\n");
scanf("%d",&opc);
if (opc==1)
{ printf("\nNumero Vulnerabilidad\t");
scanf("%d",&i reng);
printf("\nTipo de Respuesta:(1 al 4)\t");
scanf("%d",&j salv);
}
fprintf(fInter, "\nCorrida\t#Aceptad\tJT\tCT\tRatio\tTime Opt\tTime\n");
corridas=1;
acepOP=0;
do {
T0=2*SDesv;
T=T0;
sec_Inicial(a, s, XSel,ACoef);
Nsolucion =0;
SolOp=Sol;
fprintf(fpG, "\nInicio: %f", ((float)clock() - start) / CLOCKS_PER_SEC);
fprintf(fpG, "\nNum \tJustf \t\tCosto \t\tFactor\t");
for (i=1;i<=a;i++)
fprintf(fpG, "\tV%d", i);
fprintf(fpG, "\nSolucion inicial %d\n\t%f \t %f \t%f",
Nsolucion,SolOp.JT,SolOp.CT,SolOp.Fc);
for(i=1;i<=a;i++) {
bande=0;

```

```

for(j=1;j<=s;j++) {
    if (SolOp.x[i][j]==1)
        { bande=1;
          fprintf(fpG, "\t(%d)%d", i, j);
        }
    }
if (bande==0)
    fprintf(fpG, "\t(%d)0",i);
}

while (T>beta * T0) // INICIO CICLO METROPOLI
{
    Nsolucion ++;
    m=0;
    acepta=0;
    while (m<=TamVeci)
    {
        for(i=1;i<=a;i++)
            for(j=1;j<=s;j++)
                XSel[i][j]=Sol.x[i][j];
        sec_NPre2(a, s, XSel, XSeN,ACoef);
        DeltaEne= SolT.Fc - Sol.Fc;
        nAlea= (float)rand()/(float)RAND_MAX;
        if (DeltaEne<= 0 || nAlea < exp(-1*DeltaEne/T))
            {
                acepta++;
                Sol=SolT;
                for(i=1;i<=a;i++)
                    for(j=1;j<=s;j++)
                        XSel[i][j]=Sol.x[i][j];
                sec_NPre2(a, s, XSel, XSeN,ACoef);
            }
        if (Sol.Fc < SolOp.Fc && Sol.Fc>0.0 )

```



```

        {   acepOP= acepOP + 1;
            SolOp=Sol;
            printf(" .");
            tOPTi=(float)((float)clock() - start) / CLOCKS_PER_SEC;
        }
    m++;
}
T=alfa*T;
}
for(i=1;i<=a; i++) {
    sumaJT=0.0;
    sumaCT=0.0;
    fprintf(fpG2, "\n%d\tv%1d", corridas, i);
    for(j=1; j<=s; j++) {
        fprintf(fpG2, "\t%3d", SolOp.x[i][j]);
        sumaJT=sumaJT+ SolOp.x[i][j]*ACoef[i][j];
        sumaCT=sumaCT+SolOp.x[i][j]*ACoef[i][j+s];    }
    if (sumaJT>0)
        sFac=sumaCT / sumaJT;
    else
        sFac=0.0;
    fprintf(fpG2, "\t%4.1f\t%4.1f\t%8.4f ", sumaJT, sumaCT, sFac);
}
fprintf(flInter, "\n%2d\t%4d\t%d\t%4.2f\t%4.2f\t
%4.2f", corridas, acepta, acepOP, SolOp.JT, SolOp.CT, SolOp.Fc);
fprintf(flInter, " \t %f \t %f", ((float)clock() - start) / CLOCKS_PER_SEC, tOPTi);
fprintf(fpG, "\n%d\t%f \t %f \t%f ", corridas, SolOp.JT, SolOp.CT, SolOp.Fc);
for(i=1; i<=a; i++) {
    bande=0;
    for(j=1; j<=s; j++) {

```

```

        if (SolOp.x[i][j]==1)
            { bande=1;
              fprintf(fpG, "\t%1d", j);
            }
    }
    if (bande==0)
        fprintf(fpG, "\t0");
    }
    corridas++;
} while (corridas<=NItera);
fprintf(flInter, "\nFINAL\n %d\t%4.2f\t%4.2f\t%4.2f\t%4.2f\t
%4.2f", corridas, T, DeltaEne, SolOp.JT, SolOp.CT, SolOp.Fc);
fprintf(flInter, "\t%f", ((float)clock() - start) / CLOCKS_PER_SEC);
} // FIN MAIN

```

```

void coef_matriz(int *n, int* nm, float(*amat)[C])
{
    FILE *fLec;
    int temene, tempeme, i, j ;
    float Atemp;
    float fdigito;
    int bandera=0;
//
    tempeme= amenaz;
    temene= salvag;
    printf("\nValor de m (numero de amenazas): %d", tempeme);
    printf("\nValor de n (numero de salvaguardas por amenaza): %d", temene);
    printf("\n Cargar de Matriz MatrizJC.txt ");
    fLec = fopen("MatrizJC.txt", "rt" );

```

```

if (fLec==NULL)
printf("\n Error Archivo"); //Realizamos el ingreso por archivo
else
{
    if (tempeme>0 && temene>0)
    {
        for(i=1;i<=tempeme;i++)
        {
            for(j=1;j<=temene*2;j++)
            {
                fscanf(fLec,"%f",&digito);
                // amat[i][j]=float(digito);
                amat[i][j]=digito;
            }
        }
        printf("\nMatriz Cargada");
    }
    fclose(fLec);
}
printf("\n Imprimir Elementos ");
getch();
if (tempeme>0 && temene>0)
{
    printf("\n");
    for(i=1;i<=tempeme;i++)
    {
        for(j=1;j<=temene*2;j++)
        {
            printf("\t %5.1f", amat[i][j]);
        }
    }
}

```

```

        printf("\n");
    }
    getch();
}
*nm=temene;
*n=tempeme;
}

void sec_Inicial(int m, int g, int(*xT)[salvag+1], float(*amat)[C])
{
    int xaleat,xaleat2,i,j,k,tope;
    int yaS,contador;
    int iireng, jsalv,iopc;
    int siYa[amenaz+1];
    float suma, suma2,sumaPre,sumaJus,suma3,sumaFc,Presupu,PreMinimo;
    Presupu=Presup;
    PreMinimo=PresMin;
    iireng=iireng;
    jsalv=jsalv;
    iopc=opc;
    yaS=0;
    sumaPre=0.0;
    for(i=1;i<=m; i++)
    for(j=1;j<=g; j++)
        { siYa[i]=0;
          xT[i][j]=0;
        }
    tope=m;
    if (iopc==1)

```

```

    { xT[iireng][jialv]=1;
      siYa[iireng]=1;
      sumaPre= amat[iireng][g + jialv];
      tope=m-1;
    }
contador=1;
while (contador <= tope)
{
xaleat = rand()%(g) + 1;
xaleat2 = rand()%(m) + 1;
while (siYa[xaleat2]==1)
{
    xaleat2 = rand()%(m) + 1;
}
siYa[xaleat2]=1;
xT[xaleat2][xaleat] = 1;
suma=0.0;
for(k=g+1;k<=2*g;k++)
    suma=suma + amat[xaleat2][k] * xT[xaleat2][k-g] ;
sumaPre=sumaPre+suma;
if ( sumaPre>Presupu)
{
    xT[xaleat2][xaleat] = 0;
    sumaPre=sumaPre-suma;
}
contador++;
} // del while, llenar los renglones
printf("\n***");
sumaJus=0.0;
sumaFc=0.0;

```

```

for(i=1;i<=m; i++) {
    suma=0.0;
    suma2=0.0;
    suma3=0.0;
    for(k=g+1;k<=2*g;k++)
        suma=suma + amat[i][k] * xT[i][k-g] ;
    for(k=1;k<=g;k++)
        suma2=suma2 + amat[i][k] * xT[i][k] ;
    sumaJus=sumaJus+suma2;
    if (suma2 != 0)
        suma3=suma3 + suma/suma2;
    sumaFc=sumaFc+suma3 ;
    for(j=1;j<=g;j++)
        Sol.x[i][j]= xT[i][j];
    }
    Sol.JT=sumaJus;
    Sol.CT=sumaPre;
    Sol.Fc=sumaFc;
}

void sec_NPre2(int m, int g, int(*x)[salvag+1],int(*xT)[salvag+1],
float(*amat)[C])
{
int i,j,jj,k,ii,ji;
int siYa[amenaz+1];
int xaleat,xaleat2, posi1, posi2,tempo;
int yaS,contador,tope;
int iireng, jjsalv;
float suma, suma2,sumaPre,sumaJus,suma3,sumaFc,Presupu,PresuMin;

```

```

int xTem[amenaz+1][salvag+1];
Presupu=Presup;
PresuMin=PresMin;
iireng=iireng;
jjsalv=jsalv;
do {
sumaPre=0.0;
sumaJus=0.0;
sumaFc=0.0;
for(i=1;i<=m; i++)
    siYa[i]=0;
tope=m;
if (opc==1)
    { xT[iireng][jjsalv]=1;
      siYa[iireng]=1;
      sumaPre= amat[iireng][jjsalv];
      tope=m-1;
    }
yaS=0;
contador=1;
while ( yaS==0 && contador<=tope)
    {
sumaJus=0.0;
sumaFc=0.0;
for(ii=1;ii<=m; ii++)
    for(j=1;j<=g; j++)
        xTem[ii][j]= x[ii][j];
xaleat = rand()%(g) + 1;
xaleat2 = rand()%(m) + 1;
while ( siYa[xaleat2]==1)

```

```

        xaleat2 = rand()%(m) + 1;
    siYa[xaleat2]=1;
    for (j=1; j<=g; j++)
        xTem[xaleat2][j] = 0;
        xTem[xaleat2][xaleat] = 1;
suma=0.0;
    for(ii=1;ii<=m; ii++)
        for(k=g+1;k<=2*g;k++)
            suma=suma + amat[ii][k] * xTem[ii][k-g] ;
    sumaPre=sumaPre+suma;
    if ( sumaPre>Presupu)
        contador++;
    else {
yaS=1;

    for(ii=1;ii<=m; ii++) {
        suma=0.0;
        suma2=0.0;
        suma3=0.0;
        for(k=g+1;k<=2*g;k++)
            suma=suma + amat[ii][k] * xTem[ii][k-g] ;
        for(k=1;k<=g;k++)
            suma2=suma2 + amat[ii][k] * xTem[ii][k] ;
        sumaJus=sumaJus+suma2;
        if (suma2 != 0)
            suma3=suma3 + suma/suma2;
        sumaFc=sumaFc+suma3 ;
    }
    SolT.JT=sumaJus;
    SolT.CT=sumaPre;

```



```
        SolT.Fc=sumaFc;
    for(i=1;i<=m;i++) {
        for(j=1;j<=g;j++) {
            xT[i][j]= xTem[i][j];
            SolT.x[i][j]=xTem[i][j];
        }
    }
} // del else
}
} while(sumaPre<PresuMin);
}
```

Apéndice C. Productos Generados  
Publicación de Artículo.

ScholarOne Manuscripts™

**IEEE** **IEEE Access**  
Multidisciplinary | Rapid Review | Open Access Journal

[Home](#)
[Author](#)
[Review](#)

Author Dashboard

MARTA LILIA ERAÑA DIAZ ▾

[Instructions & Forms](#)
[Help](#)
[Log Out](#)

## Manuscripts I Have Co-Authored

**Author Dashboard**

- 2 Manuscripts I Have Co-Authored
- Start New Submission
- Legacy Instructions
- 5 Most Recent E-mails

STATUS	ID	TITLE	CREATED	SUBMITTED
EIC: Abbott, Derek ADM: Kukreti, Mansi	Access-2020-30152	Optimization for Risk Decision-Making through Simulated Annealing Submitting Author: CRUZ CHAVEZ, MARCO ANTONIO	14-Jun-2020	14-Jun-2020
<ul style="list-style-type: none"> <li>Accept (22-Jun-2020)</li> <li>Accept for Final Submission</li> </ul>				
EIC: Abbott, Derek ADM: Mongia, Garima	Access-2020-23335	Optimization for Risk Decision-Making through Simulated Annealing <a href="#">View Submission</a> Submitting Author: CRUIZ CHAVEZ MARCO ANTONIO	04-May-2020	04-May-2020

# Optimization for Risk Decision-Making through Simulated Annealing

MARTA LILIA ERAÑA-DÍAZ<sup>1</sup>, MARCO ANTONIO CRUZ-CHÁVEZ<sup>1</sup>,  
RAFAEL RIVERA-LÓPEZ<sup>2</sup>, BEATRIZ MARTÍNEZ-BAHENA<sup>1</sup>, ERIKA  
YESENIA ÁVILA-MELGAR<sup>1</sup>, MARTÍN H. CRUZ-ROSALES<sup>3</sup>

<sup>1</sup>Research Center in Engineering and Applied Sciences, Autonomous University of Morelos State (UAEM), Av. Universidad 1001, Colonia Chamilpa, 62209, Cuernavaca, Morelos, México; merana@uaem.com, mcruz@uaem.mx, bmartinezb@uaem.mx, erikay@uaem.mx

<sup>2</sup>Computation and Systems Department, National Technological Institute of Mexico/Veracruz Technological Institute, Calzada Miguel Ángel de Quevedo 2779, 91860, Veracruz, México; [rivera@itver.edu.mx](mailto:rivera@itver.edu.mx)

<sup>3</sup>Faculty of Accounting, Administration Informatics, UAEM, Av. Universidad 1001 Colonia Chamilpa, 62209, Cuernavaca, Morelos, México; [mcr@uaem.mx](mailto:mcr@uaem.mx)

Corresponding author: Cruz-Chávez M.A. (e-mail: mcruz@uaem.mx).

This research was funded by CONACYT grant number INNOVAPYME-21105

**ABSTRACT** In this paper, a computational methodology combining the simulated annealing algorithm with two machine learning techniques to select a near-optimal safeguard set for business risk response is presented. First, a mathematical model with four types of risk factor responses (avoid, mitigate, transfer, and accept) is constructed. Then, the simulated annealing algorithm is applied to find a set of near-optimal solutions to the model. Next, these solutions are processed by the k-means clustering algorithm for identifying three categories, and with a decision tree classifier, the most relevant elements of each one are obtained. Finally, the categorized solutions are shown to the decision-makers through a user interface. These stages are designed with the aim of the users can take an appropriate safeguard set and develop one specific and optimal program to respond to business risk factors. The results generated by the proposed approach are reached in a reasonable time using less computational resources than those used by other procedures. Furthermore, the best results obtained by the simulated annealing algorithm use a lower business budget, and they have a relative-error less than 0.0013% of the optimal solution given by a deterministic method.

**INDEX TERMS** Risk factor to bankruptcy; metaheuristic, machine learning; k-means; decision trees

## I. INTRODUCTION

In the present day, any business organization works in a changing environment exposing it to external and internal risk factors. Business risk factors are all those actions or elements that could diminish the business profit or lead it to failure [1]–[4]. They are commonly organized in five main categories: governance, strategic, financial, operational, and compliance [1]. Risks factors require an adequate response to the achievement of the strategic business objectives. Their attention must begin with early identifying the causes that trigger them [5]. In its treatment, decision-making about all the applicable actions to modify the detected risk factors is implied [6], [7]. It is essential to generate optimal plans for providing adequate risk factors attention, and the existence of a suitable communication of the actions and factors to be monitored. This promotes a favorable organizational environment, which is the foundation of a productive and competitive business [8]–[11].

The selection of safeguards (preventive or corrective measures attending risk factors) represents an organizational challenge since it is often needed to reduce threats in a short time with a limited budget. The ALARP (As Low As Reasonably Practicable) principle standing that the best cost-benefit analysis must be reflect in the weight the decision-making gives to several attributes [12], [13]. This is since both the

uncertainty and the possible consequences are linked to various alternatives. Cost alone should never be the justification, since besides direct financial loss, indirect economic impacts such as more extended business interruption and the reputation impact, can cause more substantial concern due to their possible overall financial scope [14].

There are multiple methodologies for risk assessment [1]–[4]. Many of them apply risk matrices depending entirely on the skills, experience, and critical thinking of the emulators. The methodology described in this paper agrees with the decision-maker skills but proposes the realization and weighting of more than one response to each factor, that is, designate many types of responses. However, it is essential to use critical thinking and experience to identify various solutions to the risk factors. Subsequently, the safeguard justification is assessed with the William Fine method [15]–[19], computing the cost factors, and the degree of correction or the improvement benefits in the organizational environment. In particular, cost factors should include the economic aspect and the degree of effort required to implement the safeguard (training time, resistance to change, and so on). Economic and competitive risks have been classified as natural, operational, physical, and electronic risks [20], and the Project Management Institute (PMI) [2], [21] divides the risk responses into four categories: avoid, transfer, mitigate, and accept. In both cases, a matrix is first generated using a set of possible safeguards, such as portfolio choices, and the optimal response for each risk factor is obtained by solving a combinatorial problem.

On the other hand, CPLEX is a high-performance mathematical programming solver for linear, mixed-integer, and quadratic programming [22]–[25], and R is one of the most significant tools for computational statistics, perception, and data science [26]. R is enriched with many packages and libraries for diverse applications. In particular, the Rglpk package is a high-level interface to connect R with CPLEX [27]. Then, among the multiple options available to solve a mathematical model, it can be observed that, regardless of the programming language used, it is necessary to consider three points for the use of any software: (1) build the model, (2) set the solver parameters, and (3) compute and extract the solution [28]. However, it is known that several combinatorial optimization problems are NP-complete [29], and mathematical programming solvers can be spending excessive time finding the best solution. Other strategies, such as metaheuristics (MHs), can be applied to reach near-optimal solutions in a reasonable time. They are reliable and straightforward strategies for solving complex problems [30][32]. MHs, try to simulate both intelligent processes and behaviors observed in nature and other disciplines. They are characterized by combining the search space exploration to identify promising areas and exploit these areas to improve the known solution or solutions.

In particular, concerning the treatment of business risks through MHs, several frameworks to solve decision problems in many business areas are described in the existing literature. In Reference [33], a review of the use of evolutionary algorithms (EAs) to solve diverse financial problems such as fraud and bankruptcy detection, credit portfolio, credit scoring, and forecasting, among others, is provided. Authors include several EAs such as genetic algorithms (GAs), genetic programming (GP), multi-objective EAs (MOEAs), among others. Furthermore, in Reference [34], a GA is applied to solve a financial optimization problem by maximizing the profits and minimizing the default probability. In Reference [35], the simulated annealing (SA) algorithm, GA, and particle swarm optimization (PSO) are combined in a hybrid decision-making method to solve portfolio optimization problems with different risk measures. Reference [36] describes a data-driven threshold accepting-heuristic optimizing financial risks with operational data.

Reference [37] describes the use of several MOEAs such as the Non-Dominated Sorting GA-2 (NSGA2), and the Strength Pareto EA-2 (SPEA2) to solve the Business Process Optimization (BPO) problem. The authors provide a set of BPO alternative solutions for several experimental and real-life scenarios and produce satisfactory results, since this approach generates diverse designs, and selects those with optimal objective values for business processes in less time. Furthermore, in Reference [38], a GA-based procedure is applied in a computer security decision-support system seeking a near-optimal combination of threats costs, safeguards, and the impact on assets. In Reference [39], a multi-objective tabu-search method to minimize the risk of network vulnerabilities is developed. In Reference [40], a GP-based approach to generate fuzzy association rules used in one intrusion detection system is described.

Finally, machine learning techniques such as classifiers, clustering algorithms, and feature selection methods have been used to treat business risks. Reference [41] uses a Support Vector Machine (SVM) classifier and the Dynamically Growing Self-Organizing Tree (DGSOT) clustering algorithm to implement an intrusion detection system. Reference [42] describes another intrusion detection system, including four stages using machine learning techniques: preprocessing, classification, feature reduction, and feature selection. Reference [43] proposes an adaptive method with logical procedures for detection of anomalies and cyber-attacks, based on the coverage matrices of features and the use of elementary classifiers.

In the previous paragraphs, the business risk treatment is approached over particular types of risks such as operational risks, financial risks, and those related to informatic security. The use of an integral approach attending all business risk factors, including the industrial risks and its competitiveness, compliance with regulations, and all financial and operational aspects, is crucial to achieving the business objectives. Furthermore, although MHs such as GA, GP, and PSO have been used to treat business risks, the use of the SA algorithm is reported in only one study to solve portfolio optimization problems.

In this paper, a new methodology to address business risk factors with a mathematical optimization model is presented. This methodology combines the SA algorithm with two machine learning techniques to select a near-optimal safeguard set to implement a robust business risk response program. The main contributions of this methodology are the follows:

- 1) The definition of a mathematical model for the treatment of the risk factors in all business areas, based on four types of risk factor responses.
- 2) The use of an efficient MH such as the SA algorithm to reach near-optimal solutions to the mathematical model, which uses less computational resources than those used by the CPLEX solver.
- 3) The combined use of two machine learning methods to categorize the near-optimal solutions, and to identify the most relevant elements to be included in an optimal response plan to risk-decision making.
- 4) The implementation of a user-interface using dynamic tables to facilitate the decision-makers the creation of an optimal plan for the attention of business risk factors.

The rest of the paper is structured as follows: Section 2 outlines the theoretical framework to build this computational methodology. The mathematical formulation of the optimization model, the simulated annealing algorithm, and the two machine learning techniques used in this proposal are described in Section 3. In section 4, the six steps comprising the methodology proposed in this paper, are detailed. The experimental study, including the results of a real case study and ten random instances, is discussed in Section 5. Finally, Section 6 holds the conclusions and the future work of this proposal.

## II. THEORETICAL FRAMEWORK

In this research, the theoretical framework adopted is based on two pillars. The first one consists of a conceptual model to identify the risk factors and the risk level maintained by the business. A set of safeguards for four types of responses (avoid, mitigate, transfer, and accept) is implemented using a framework suggesting the best practices for attending business risks. The other pillar is the use of a mathematical formulation of the problem and the application of the SA algorithm to find a set of near-optimal solutions. These solutions are processed by two machine learning algorithms to identify three categories consuming the budget in different ways and give different benefits to business security.

The risk factors are those that can difficult the fulfillment of the company's mission, vision, and strategic objectives. They can be derived from the nature of its activities or the company's external conditions [44]–[46]. On the other hand, safeguards are actions taken to anticipate, minimize, mitigate, or otherwise treat the adverse impacts associated with vulnerable activity [47].

### A. SAFEGUARDS (COUNTERMEASURES) MODEL

Once the risk factors have been identified, a safeguards model can be proposed [21]. The model is a set of selected actions allowing for anticipating, minimizing, mitigating, or otherwise treating the adverse impacts triggered by those risk factors. Risk factors must be identified based on (1) a valid and efficient operation, (2) the adequate internal controls, and (3) following the laws and regulations. Furthermore, the safeguards

model should consider the implementation costs and possible losses if no actions to resolve any risk problem are taken.

**B. RISK ASSESMENT SCALES**

In this paper, two methodologies are used to risk quantification: the PMI framework [2], [20], [47-49], and the William Fine method [15].

PMI is a framework based on the exact information obtained from the company. First, the risk factors are grouped based on the possible effect (positive or negative) on a given objective (time, cost, scope, or quality). Next, four strategies are defined to deal with the risks: avoid, transfer, mitigate, and accept.

On the other hand, the William Fine method is used to facilitate the expedited control of the risks (accidents and job losses). This method first assigns a priority level for each risk, and one estimated cost for the implementation of corrective actions contemplated to eliminate them. Next, it computes the justification value of the safeguard benefit.

The risk score *R* is determined as follows:

$$R = C \times E \times P \tag{1}$$

where *C* is the consequence rating value, *E* is the exposure value, and *P* is the probability value, which values are obtained applying the Table 1.

**TABLE 1 Values for William Fine's Process.**

<b>C</b>	<b>CONSEQUENCES (MOST PROBABLE RESULT OF POTENTIAL RISK)</b>
25	Mortal or superior damage to 5 monthly payrolls
15	Discapacity injury or damage between 1 to 5 monthly payrolls
5	Injury not discapacity or damage between 10% to 100% of the monthly payroll
1	Minor injuries or damage 10% less monthly payroll
<b>P</b>	<b>PROBABILITY (LIKELIHOOD THAT THE RISK SEQUENCE WILL FOLLOW TO COMPLETION)</b>
10	Is the most likely and expected result if the event takes place
6	Is quite possible, not unusual, has an even 50-50 chance.
3	Would be an unusual sequence or coincidence, 20% chance.
1	Has never happened after many years of exposure, but is 5% conceivably possible
<b>E</b>	<b>EXPOSURE (FREQUENCY OF OCCURRENCE OF THE RISK EVENT)</b>
10	Continuously (or many times daily)
6	Frequently (about once daily)
3	Occasionally (once per week to once per month)
1	Remotely possible (it has been known to occur)

The justification value of the safeguard benefit is computed as follows:

$$J = R / (CF \times DC) \tag{2}$$

where *CF* is the cost factor, and *DC* is the degree of correction, which values are taken from Table 2.

**TABLE 2. Values for William Fine's Cost Justification.**

<b>CF</b>	<b>COST FACTOR (ESTIMATED COST IN MONTHLY PAYROLL IF THE SAFEGUARD IS ELECTED)</b>
10	More than 5 monthly payrolls
6	Between 1 to 5 monthly payrolls
4	Between 50% to 100% of the monthly payroll
2	Between 10% to 50% of the monthly payroll

1	10% less monthly payroll
<b>DC</b>	<b>DEGREE OF CORRECTION (DEGREE TO WHICH RISK WILL BE REDUCED)</b>
1	Risk almost eliminated at 100%
2	Reduce at least 75% 3 Reduce by 50% - 75%
4	Reduce by 25% - 50%
6	Slight effect on risk, less than 25%

### III. MATHEMATICAL FORMULATION OF THE OPTIMIZATION MODEL

The safeguards selection can be modeled as a binary integer linear programming (ILP) optimization problem [29], [50]. This model uses a set of risk factors and a safeguards matrix  $M_{a \times s}$ , where  $a$  is the number of risk factors, and  $s$  is the number of safeguard types could be implemented for each risk factor. Each safeguard value  $J_{i,k}$ ,  $i=\{1, \dots, a\}$ ,  $k=\{1, \dots, s\}$ , represents the justification to reduce the  $i$ -th risk factor using the  $k$ -th type of safeguard, with an implementing cost  $CS_{i,k}$ . The values  $CS_{i,k}$  and  $J_{i,k}$  are computed using the William Fine method.

Then, the optimization model proposed in this paper is as follows:

$$\min (f) = \min \left[ \sum_{i=1}^a \sum_{k=1}^s \frac{CS_{i,k}}{J_{i,k}} x_{i,k} \right] \quad (1)$$

subject to

$$\sum_{i=1}^a \sum_{k=1}^s CS_{i,k} x_{i,k} \leq P \quad (2)$$

$$\sum_{k=1}^s x_{i,k} \leq 1 \quad i=\{1, \dots, a\} \quad (3)$$

$$x_{i,k} = \begin{cases} 1 & \text{if } S_{i,k} \text{ is selected} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The objective function (1) represents the total benefit-cost ratio for the selection of safeguards. Constraints in (2) avoid exceeding the budget  $P$ . Constraints in (3) ensure that one safeguard can be selected for each risk factor. Constraints in (4) indicate, with the binary variable  $x_{i,k}$ , that the  $k$ -th type of safeguard for the  $i$ -th risk factor is or not selected. It is assumed that  $J_{i,k} > 0$  and  $CS_{i,k} > 0$ , for each  $i=\{1, \dots, a\}$  and  $k=\{1, \dots, s\}$ .

#### A. THE SIMULATED ANNEALING ALGORITHM

SA is an algorithm inspired in a type of thermodynamic behavior [51]. It has been widely successful in solving complex optimization problems due to its ability to escape from local points by accepting not good solutions [52]–[54]. SA is a simple algorithm whose competitive results are achieved in a short time due to its asymptotic convergence.

The SA process begins with an initial feasible solution  $\omega_0$ , and with the initialization of the control parameter  $T$ . With an iterative structure, SA disturbs one candidate solution until  $T$  reaches a value less than the stop criterion. During this process, in each iteration a new candidate solution is selected from neighborhood of the current solution. These solutions are compared, and the best one is chosen as the new current solution. In some cases, one not improving candidate solution is accepted to escape a local optimum and to continue searching for better solutions. The probability of taking these solutions depends on the  $T$  parameter, which decreases in each algorithm iteration using a control coefficient  $\alpha$ . A number of parameters need to be adjusted to ensure the algorithm convergence [55], [56]. They are shown in Table 3.



TABLE 3. The SA parameters

PARAMETER	DESCRIPTION
$\omega_0$	Initial candidate solution
$T$	External cycle control parameter
$T_0$	Initial value of the control parameter
$\alpha$	Control coefficient ( $0 < \alpha < 1$ )
$L$	Markov chain size
$f$	Cost function
$\Omega$	Neighborhood structure
$\beta$	Stop criterion coefficient
$T_f$	Final value of the control parameter

It is clear that when SA perturbs only one solution in each stage of its iterative process, it promotes the solution space exploitation. Furthermore, by the use of its acceptance criteria, the solution space exploration is encouraged. Although other MHs such as GA, GP, and PSO have demonstrated to reach near-optimal solutions for diverse optimization problems, they consume more computational resources than that used by SA. These MHs disturb a set of candidate solutions in each step of their iterative process, and this implies the evaluation of several solutions, unlike SA, evaluating only one candidate solution.

However, one correct setting of the control parameter  $T$  and the definition of one useful stopping criteria are crucial to the algorithm convergence since the use of inadequate values can affect the SA performance. Some straightforward methods to parameter setting can be found in the existing literature [57].

The SA-based algorithm used for this work is described in Section 4, where a candidate solution  $\omega$  is defined by the set of selected safeguards addressing the risk factors. The cost function  $f(\omega)$  is the benefit-cost ratio of a set of chosen safeguards. The neighborhood of a current solution  $\Omega(\omega)$  is defined as the set of feasible solutions generated by disturbing the response type of the selected safeguard, for a given risk factor.

## B. MACHINE LEARNING

Machine learning procedures aim to allow an agent (system, device, or program) to learn when their performance improves with experience [58]–[62]. In this work, two machine learning techniques are used to generate optimal plans for adequate attention to business risk factors: The K-means clustering algorithm and a decision tree classifier.

K-means is an unsupervised learning algorithm dividing a dataset instances into a certain number of groups according to an optimization criterion [63]–[65]. It has been identified as one of the most popular clustering algorithms [66]–[68]. If the number of clusters is  $k$ , the algorithm determines a set of  $k$  centroids  $C = \{c_1, \dots, c_k\} \in \mathbb{R}^d$  for a set of  $n$  instances on one  $d$ -dimensional vector  $D = \{x_1, \dots, x_d\} \in \mathbb{R}^d$ , where the following error function is minimized:

$$E(C) = \sum_{x \in D} \min_{i \in \{1, \dots, k\}} |x - c_i|^2 \quad (7)$$

On the other hands, decision tree induction is a supervised learning procedure to build a hierarchical classifier [69], [70]. Decision trees represent in a graphical form a set of decision rules to determine the class membership of unclassified instances [71], [72]. They are commonly used in decision analysis, helping identify the relevant items to reach a goal. Each tree node usually contains one attribute, and the branches leaving it corresponding to the possible attribute-values. In order to classify a new instance, the classifier first evaluates the root node, and the instance is filtered downwards until a leaf is found, which corresponds to the instance class [73]–[77].

The decision tree induction process requires using a criterion to select the attribute evaluated in each new tree node. This criterion measures the *goodness of split*, i.e., how well the instances are discriminated between classes. In this work, the Gini index is used, since it is applied in one of the most efficient induction classifiers: the CART method [78]. Gini index measures the *impurity* of an attribute concerning the classes. Given the probabilities for each class  $p_i$ , the general Gini function is defined as follows:

$$\sum_{j \neq i} p_i p_j = 1 - \sum p_i^2 \quad (8)$$



### C. COMPUTATIONAL COMPLEXITY

Computational complexity is a computer science area focus on to classify and compare the computational problems based on the difficulty level to solve them, i.e., the number of resources required to find a solution to a problem using some algorithm [79]. In particular, ILP problems can be classified as P or NP problems [80]. For those ILP problems considered NP problems, a near-optimal solution can be found in polynomial time using non-deterministic computational methods. On the other hand, since the CPLEX solver uses the branch-and-cut algorithm [81], [82] to solve an ILP problem, it always finds an optimal solution to the problem. However, the computational resources utilized increasing at an exponential rate according to the problem data.

In the present work, a binary ILP optimization model to business risk response is developed. The theoretical study to know if this particular model belongs to P or NP classes has not been performed, which is part of one future work. The importance that when there is no knowledge of the theoretical complexity of a problem, it always has a near-optimal solution through some MH. In this work, the proposed optimization model is solved using the SA algorithm, which has the advantage of its simplicity and of being a rather fast method with less memory usage [83].

Since of the combinatorial nature of ILP optimization problems, CPLEX users may have difficulties getting good performance with them. CPLEX has many parameters allowing customize the way the algorithm operates. While this variety provides many alternatives to improve the solver performance, a user cannot realistically experiment with all the possible combinations of parameter settings. Reference [28] indicates that, before trying to improve the algorithm performance, it is needed previously to locate the current bottleneck.

### IV. METHODOLOGY

In this research, a quantitative model to estimate the risk levels is developed. The response strategies are applied by optimizing the set of safeguards that are implemented with a limited budget. This methodology is described in six stages, as is shown in a flowchart in Fig. 1. In the following paragraphs, these stages are detailed.

**Step One (Data Collection):** The main business-data are required to identify the risk factors in all business areas. The internal strengths and weaknesses, as well as the external opportunities and threats, are identified using one SWOT-analysis qualitative questionnaire [84] (see Appendix A). This questionnaire is applied to discover the bankruptcy decision rules applied by the experts [85].

There are 32 factors used by one of the largest Korean commercial banks. They are categorized into six risk areas: industry risk (IR), management risk (MR), financial flexibility (FF), credibility (CR), competitiveness (CO), and operating risk (OP). Fig. 2 shows the 32 factors in these six areas. The risk score for weaknesses and threats is computed using the William Fine tables.

**Step Two:** A selection of safeguards for each risk factor is conducted in this step. For each risk factor, a possible action for each response type (accept, mitigate, transfer, or avoid) is needed. The correction-levels reached, and the implementation costs are then defined using the William Fine tables.

**Step Three:** This step concerns the development of the mathematical model, including adjusting the objective function and the constraints for the optimal selection of safeguards. This is a binary ILP model

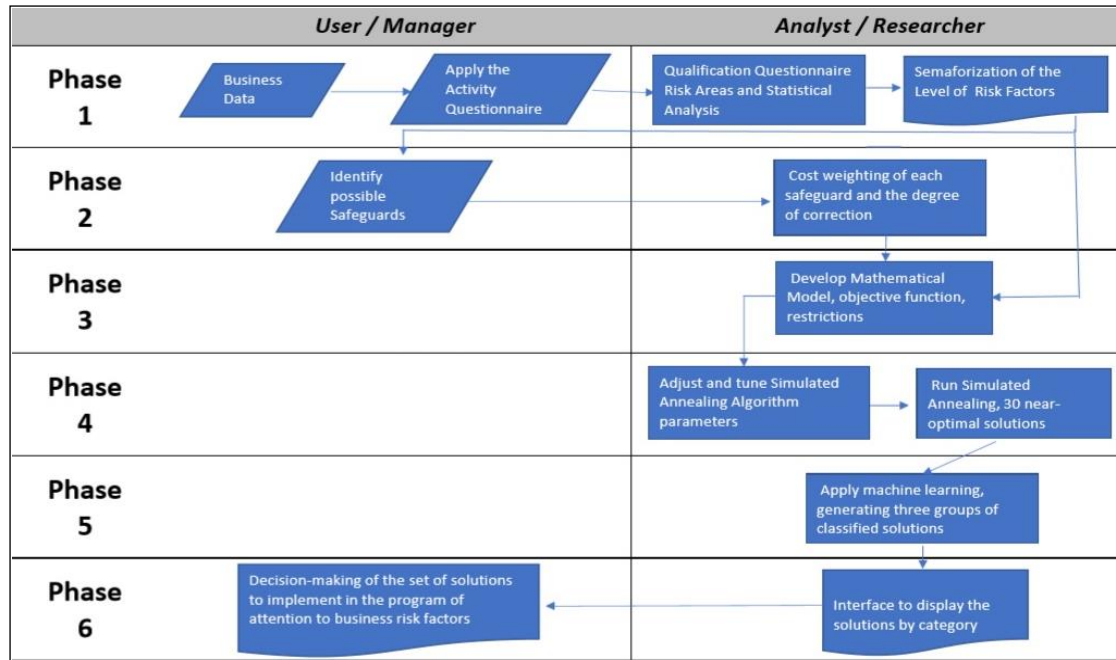


FIGURE 1. Flowchart for the methodology for Treatment of Business Risk Factors.

described by (3)-(6).

**Step Four:** In this stage, the binary ILP model is solved using the SA algorithm. SA is executed repeatedly to obtain several near-optimal solutions. The heuristic adaptation, solution structure, benefit-cost matrix, and the feasible solutions within the neighborhood are presented below.

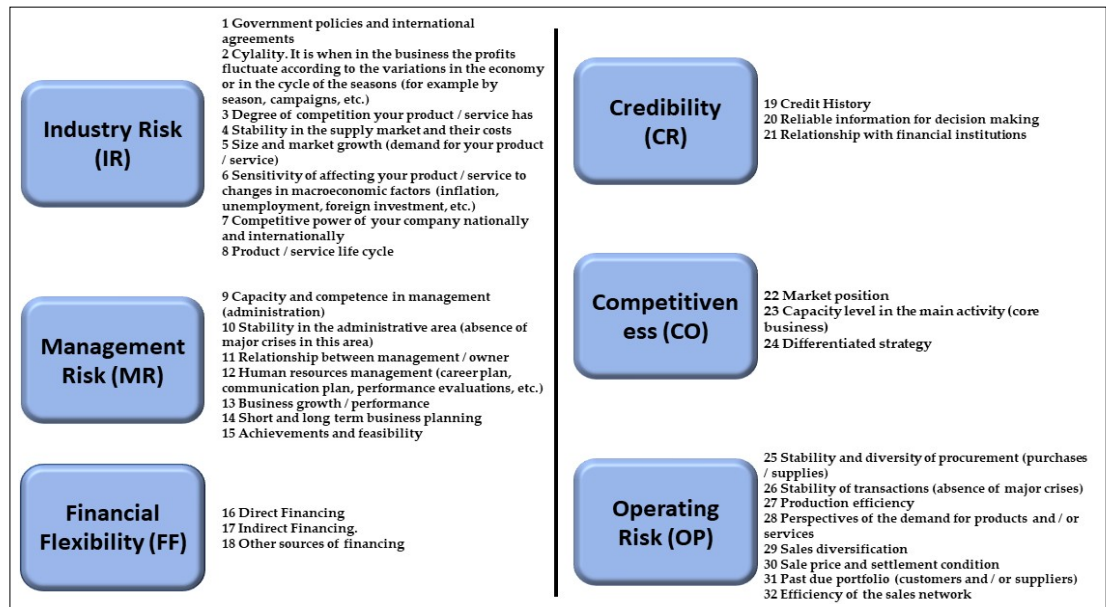


FIGURE 2. Factors in six business areas.

The format for the safeguards for the risk factors is as follows:

$$S = \{JT, CT, (x_{1,1}, \dots, x_{1,s}), (x_{2,1}, \dots, x_{2,s}), \dots, (x_{a,1}, \dots, x_{a,s})\} \quad (9)$$

where  $JT$  is the total justification of the benefit for implementing the set of safeguards selected,  $CT$  is the total cost of the implementation,  $x_{i,k} \in \{0, 1\}$ , is 1 if the response  $k$  is taken to factor  $i$ , or 0 if it is not selected. Table 4 shows the structure of values for each safeguard for each risk factor.

The solution structure  $\omega$  is described as follows:

$$\omega = \{JT, CT, (k_1, \dots, k_a)\} \quad (10)$$

where  $k_i$  is the type of response to the safeguard implemented for each  $i$ -th risk factor,  $i \in \{1, \dots, a\}$ .

TABLE 4. Structure of the safeguards values.

FAC \ SAFG	S <sub>1,1</sub>	S <sub>1,2</sub>	...	S <sub>1,s</sub>
V <sub>1</sub>	(J <sub>1,1</sub> , CS <sub>1,1</sub> )	(J <sub>1,2</sub> , CS <sub>1,2</sub> )	...	(J <sub>1,s</sub> , CS <sub>1,s</sub> )
V <sub>2</sub>	(J <sub>2,1</sub> , CS <sub>2,1</sub> )	(J <sub>2,2</sub> , CS <sub>2,2</sub> )	...	(J <sub>2,s</sub> , CS <sub>2,s</sub> )
V <sub>3</sub>	(J <sub>3,1</sub> , CS <sub>3,1</sub> )	(J <sub>3,2</sub> , CS <sub>3,2</sub> )	...	(J <sub>3,s</sub> , CS <sub>3,s</sub> )
...	...	...	...	...
V <sub>a</sub>	(J <sub>a,1</sub> , CS <sub>a,1</sub> )	(J <sub>a,2</sub> , CS <sub>a,2</sub> )	...	(J <sub>a,s</sub> , CS <sub>a,s</sub> )

The algorithm 1 shows the SA-based approach used in this methodology. The parameter values are defined as follows: First,  $T_0$  is set to  $2\rho$ , where the standard deviation  $\rho$  is obtained using a set of candidate solutions selected at random [57]. The Markov chain length  $L$  is defined as follows:

$$L = a \times s \times (a \times s - 1) \quad (11)$$

where  $a$  is the number of risk factors, and  $s$  is the number of safeguard types. This value determines the number of iterations in the Metropolis cycle, which is described in lines 5-19 of Algorithm 1. Furthermore, the SA external cycle is showed in lines 4-21. In particular, line 6 describes the creation of a new solution, line 7 computes the difference in the energy cost, and the acceptance criteria is showed in lines 8-15. The selection of the best solution of the SA algorithm is described in lines 16-18. Finally, the neighbor  $\omega'$  is created by randomly select the elements described in (10), and consistent with all problem constraints.

---

**Algorithm 1.** SA-based algorithm.

---

```

1   $\omega_0 \leftarrow$  Compute the initial solution
2   $T_0 \leftarrow$  Initial value of the control parameter
3   $\omega \leftarrow \omega_0, \omega^* \leftarrow \omega_0, T \leftarrow T_0, T_f \leftarrow \beta \times T_0$ 
4  while  $T \geq T_f$  do
5    for each  $i \in \{1, \dots, L\}$  do
6       $\omega' \leftarrow \Omega(\omega)$ 
7       $\Delta \leftarrow f(\omega') - f(\omega)$ 
8      if  $\Delta < 0$  then
9         $\omega \leftarrow \omega'$ 
10     else
11        $\rho \leftarrow$  random uniform value in  $[0, 1]$ 
12       if  $\rho \leq e^{-\Delta/T}$  then
13          $\omega \leftarrow \omega'$ 
14       end if
15     end if
16     if  $\omega^* > \omega$  then
17        $\omega^* \leftarrow \omega$ 
18     endif
19   end for
20    $T \leftarrow \alpha \times T$ 
21 end while

```

---

Reference [86] describes the computational complexity of the SA algorithm as follows:

$$O(\tau S \ln|R|) \quad (12)$$

Where  $\tau$  is the maximum number of steps required to generate and evaluate a solution,  $S$  is the neighborhood size, and  $R$  is the size of the solution space.

Based on the input algorithm parameters,  $a$  and  $s$ , it is possible to evaluate the complexity of Algorithm 1. If  $\tau = as^2$ ,  $S = as(as - 1)$ , and  $R = s^a$ . Then, (12) can be represented as follows:

$$O(as^2as(as - 1) \ln|s^a|) \quad (13)$$

And, simplifying (13), the computational complexity of the SA algorithm used to solve the Risk Decision-Making optimization problem, is as follows:

$$O\left(\left((as)^4\left(1 - \frac{1}{as}\right)\right) \ln(s)\right) \quad (14)$$

**Step Five (Categorize the solutions):** The near-optimal solutions generated by the SA algorithm are categorized into three clusters through the K-means clustering algorithm. This procedure is implemented using the cluster R package [87], as is shown in Algorithm 2.

---

**Algorithm 2.** R script for clustering using the k-means method.

---

```
1 library(cluster)
2 k <- 3 # number of groups
3 dataset <- A set of SA near-optimal solutions
4 results <- kmeans(dataset,k)
5 clusplot(dataset,results$cluster,color=TRUE,
6   shade=TRUE,labels=2,lines=0)
7 # Add the cluster number to each solution
8 dataset <- cbind(dataset,cluster=results$cluster)
9 # Write the dataset in an external file
```

---

The categorized solutions are processed to identify each group's relevant elements utilizing decision trees [88]. This procedure is implemented using the rpart and rpart.plot R packages [89], as is shown in Algorithm 3.

---

**Algorithm 3.** R Script for inducing decision trees

---

```
1 library(rpart)
2 library(rpart.plot)
3 dataset <- Clustered solutions
4 dt <- rpart(cluster~.,method="class",data=dataset)
5 print(dt)
6 rpart.plot(dt)
```

---

**Step Six (Visualization of solutions):** The solutions are shown to the decision-maker in a spreadsheet, in a user interface with dynamic tables [90]. All possible solutions with the assigned category are displayed. This presentation helps the decision-makers develop a specific and optimal program for the attention to business risk factors.

## V. EXPERIMENTAL STUDY

In this section, the experimental study adopted to analyze and compare the methodology performance is presented. First, the experimental methodology applied in this proposal, and the definition of the algorithm parameters is detailed. Then, the experimental results and the statistical tests carried out to evaluate these results are outlined. Finally, a discussion about the performance of the proposed methodology is provided.

### A. EXPERIMENTAL METHODOLOGY

Two experiments are conducted: first, the proposed scheme is applied in a Mexican company from the private sector to analyze the pertinence of the provided solution, and then, a benchmark of ten datasets randomly created is used to conduct an statistical analysis and to evaluate the methodology performance. Table 5 shows the values of the ten datasets created at random.

TABLE 5. BENCHMARK WITH 10 DATASETS

INSTANCE	RISK FACTORS	RESPONSE TYPES
1	20	4
2	20	6
3	32	4
4	32	6
5	32	8
6	50	8
7	100	8
8	200	8
9	300	8
10	400	8

The Mexican company belongs to the hydrocarbon sector and is located in Cuernavaca, Mexico. It has extensive experience in civil protection consulting, providing advice for compliance with general administrative provisions and applicable regulations of international hydrocarbon standards. The questionnaire to identify risk factors is applied to the company manager, having the adequate academic and professional experience to answer questions about risk management. Eleven risk factors are identified and quantified. A brainstorming is carried out to identify possible safeguards for the four types of responses, selecting 44 of them according to the company's needs and possibilities. Fig. 3 shows a semaphoring graph with the risk level for all factors, where red boxes identify the risk factors.

Furthermore, Fig. 4 shows the 44 safeguards ordered by type of response, for each risk factor detected in the company. The safeguard values and the implementing costs are computed using the William Fine method, which values are shown in Table 6 and Table 7, respectively. Furthermore, the maximum budget defined by the company is 66 monetary units.

This study is carried out on a computer with Intel Core i3 CPUs and 4 GB RAM. The results of the SA algorithm are compared with those obtained by the CPLEX solver through the Rglpk package. Table 8 shows the SA parameters. 90 independent runs of the SA algorithm are conducted, to get several near-optimal solutions and to apply the machine learning methods included in the proposed methodology.

TABLE 6. JUSTIFICATION MATRIX

$V_i/S_{ii}$	$S_{i1}$	$S_{i2}$	$S_{i3}$	$S_{i4}$
	ACCEPT	MITIGATE	TRANSFER	AVOID
1	2.25	3.00	2.25	1.50
2	33.75	33.75	45.00	135
3	208.33	625.00	250.00	138.89
4	22.50	67.50	22.50	33.75
5	3.13	75.00	9.38	6.25
6	56.25	900.00	150.00	112.50
7	46.88	375.00	62.50	46.88
8	2.50	7.50	3.75	2.50
9	1.25	2.50	3.75	1.25
10	1.56	6.25	2.08	1.56
11	33.75	45.00	67.50	67.50

TABLE 7. COST MATRIX

$V_i/S_{ij}$	$S_{i1}$ ACCEPT	$S_{i2}$ MITIGATE	$S_{i3}$ TRANSFER	$S_{i4}$ AVOID
1	1.767	2.799	47.386	77.374
2	7.812	53.242	60.841	1.436
3	54.435	10.254	908.163	84.842
4	111.757	10.743	82.412	58.859
5	112.287	5.663	58.487	40.921
6	45.746	1.397	6.321	9.555
7	44.061	4.211	97.362	49.435
8	75.828	7.676	54.894	109.470
9	47.675	11.917	57.329	73.612
10	50.853	11.880	74.636	46.016
11	6.638	119.965	10.057	5.906

TABLE 8. SA PARAMETERS

PARAMETER	VALUE
$T_0$	47.520
$\alpha$	0.980
$L$	1,892.000
$\beta$	0.001

## B. RESULTS

1) **Details for the Mexican company results:** Fig. 5 shows the best solution reached for each of the 90 runs of the SA algorithm. Based on the total justification of one solution, they are used to build three categories through the k-means clustering algorithm. The averages, standard deviations, minimum and maximum values for each cluster are shown in Table 9.

TABLE 9. AVERAGES, STANDARD DEVIATION, AND MINIMUM AND MAXIMUM VALUES FOR EACH CREATED CLUSTER.

	TOTAL JUSTIFICATION			TOTAL COST			OBJECTIVE FUNCTION VALUE		
	C1	C2	C3	C1	C2	C3	C1	C2	C3
	$\mu$	2249	1332	301	61.5	36	47.8	5.8	3.5
$\sigma$	37.1	281	231	2	17.4	19.7	1.8	2.9	1.4
Max	2261	1730	768	65.5	65.9	65.5	8.7	8.8	5.2
Min	2061	920	33.8	54.3	5.6	5.9	3.2	0.5	0.1

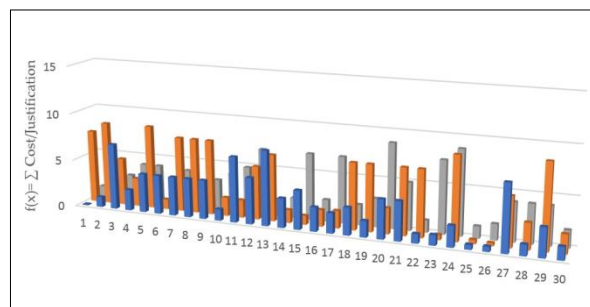
i. Industry risk (IR)	ii. Management risk (MR)	iii. Financial Flexibility (FF)	iv. Credibility (CR)	v. Competitiveness (CO)	vi. Operating Risk (OP)
3 Degree of competition product/service					26 Stability of transactions
6 Sensitivity macroeconomic factors affecting product	12 Human resources management				30 Sale price and settlement condition
4 Stability in the supply market and their costs	14 Short and long term business planning				32 Efficiency of the sales network
5 Size and market growth (demand for your product / service)	9 Capacity and competence in management (administration)				25 Stability and diversity of procurement (purchases)
7 Competitive power nationally and internationally	10 Stability in the administrative area				27 Production efficiency
8 Product / service life cycle	11 Relationship between management / owner	16 Direct Financing	20 Reliable information for decision making	22 Market position	28 Perspectives of the demand for products and / or services
1 Government policies and international agreements	13 Business growth / performance	17 Indirect Financing.	19 Credit History	23 Capacity level in the main activity (core business)	29 Sales diversification
2 Cylality. Fluctuating to the variations in the economy	15 Achievements and feasibility	18 Other sources of financing	21 Relationship with financial institutions	24 Differentiated strategy	31 Past due portfolio (customers and / or suppliers)

FIGURE 3. Risk levels of the 32 factors identified in the real case study (Risk in red, no risk in green).



V <sub>i</sub> Risk Factor	S <sub>i,1</sub> Accept	S <sub>i,2</sub> Mitigate	S <sub>i,3</sub> Transfer	S <sub>i,4</sub> Avoid
V1	Help desk area	Customer service training and staff career plan development	Training according needs detected. Training program	Hire HR experts
V2	Only the current resource is used	Have collateral assets	Search partnership with capitalist	Utility savings for emergencies
V3	Perform a market analysis and detect needs	Prepare advertising plan, direct sales team preparation and technology platform	Hire a marketing service	Accept new customers only by recommendation
V4	Make business alliances	Develop plan to expand the market, new types of service, contact with old customers	Hiring consultants for creativity and innovation plan	Growth plan consulting
V5	Have a ticket for fines	Reality an official daily monitoring periodically	Hiring dispatch for monitoring	Link with a suitable turn agent
V6	Register entries and exits of resources	Prepare strategic plan, annual projection, budget	Hire outsourcing for strategic plan development	Event log
V7	Increase small cash amount	Enable dept. of collections	Contract collection service	I charge policies in advance to customers
V8	Increase salaries for current staff	Request 3 to 5 young people "building the future"	Hire outsourcing technical staff	Limited to customer service with the current template
V9	Increase the amount of expenses per double visit	Implement a communication improvement plan with clients to avoid double visits	Implementation service to third-party clients (GASAF)	Implement online service
V10	Have the settlement cost updated	Follow the planning, and monitor the trend to innovate in strategies	Generate sale Shareholders	Acquire Credit
V11	Prepare and disseminate information	Have a good internal and external platform communication plan	Hire social network services and website	Request weekly reports

**FIGURE 4.** Possible safeguards V1 (Human resources management), V2 (Indirect financing), V3 (Efficiency of the sales network), V4 (Degree competition product/service), V5 (Sensitivity of affecting your product/service to changes in macroeconomic factors), V6 (Short and long term business planning), V7 (Direct financing), V8 (Other sources of financing), V9 (Stability of transactions, absence of major crises), V10 (Sale price and settlement condition), and V11 (Reliable information for decision making).



**FIGURE 5.** Best solutions of the 90 algorithm runs.

Three well-differentiated groups are shown in Fig. 6. However, Table 9 shows that some solutions have higher implementation costs in cluster 2 but with lower justification degrees than those in cluster 1. This information must be provided for an adequate decision of the decision-makers.



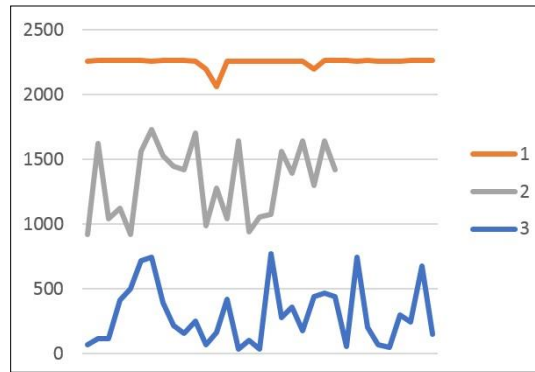


FIGURE 6. Groups classified by total justification.

Fig. 7 shows a decision tree induced using the clustering algorithm values: JT, CT, objective function value, and the number of assigned cluster. In this tree can be observed that when the JT value is not less than 1895, the solution of the selected safeguards belongs in Cluster 1. Furthermore, if the JT value is between 844 and 1895, the solution is in Cluster 2, and when the JT value is lower than 844, the solution is in Cluster 3.

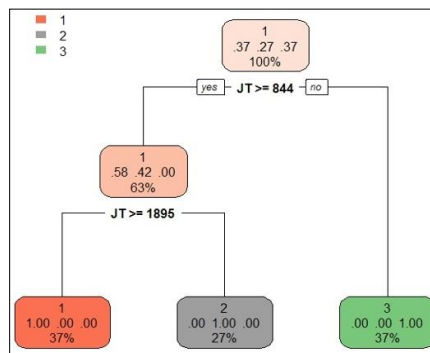


FIGURE 7. Decision tree induced using the cluster information.

Finally, to obtain the most relevant risk factor of each cluster, a decision tree is induced using the responses types suggested in each solution, as shown in Fig. 8. In this tree can be observed that when the risk factors 3 and 6 are mitigated, the solution belongs in Cluster 1. If factor 6 is mitigated, and factor 3 is accepted, transferred, or avoided, the solution is in Cluster 2. Finally, if the risk factor 6 is accepted, transferred, or avoided, the solution is in Cluster 3. With this information, the decision-maker can consider the responses type to treat the most relevant risk factor.

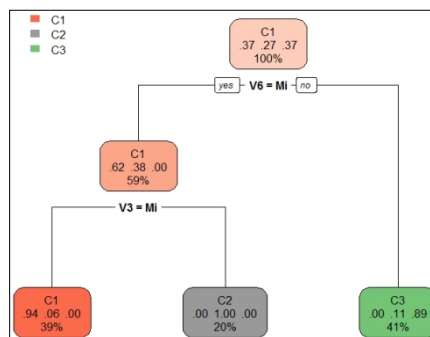


FIGURE 8. Decision tree using the response type for risk factors.

The last step of the proposed methodology is to visualize the near-optimal solutions in a user interface, as shown in Fig. 9. Dynamic tables in the spreadsheet allow the decision-maker to choose a specific solution and to see the selected safeguards, the type of response for each factor risk, as well as the JT and CT values. Furthermore, Fig. 10 shows a spreadsheet where the decision-maker can select one solution of each cluster and see the safeguards, and the type of response for each factor risk. A detailed description of this interface is provided in Appendix B.

As an example, a near-optimal solution selected by the decision-maker is the one with a total justification of 2,261, and one implementation cost of 60.93 monetary units. Table 9 shows the safeguards and the type of response for each risk factor defined in this selected solution.

TABLE 10. AN EXAMPLE OF SELECTION CHOSEN FOR THE DECISION-MAKER.

ID	FACTOR	SAFEGUARD	RESPONSE TYPE
1	V1	Help desk area	Accept
2	V2	Savings of utilities for emergencies	Avoid
3	V3	Prepare advertising plan, direct sales team and technology platform	Mitigate
4	V4	Develop a plan to expand the market	Mitigate
5	V5	Perform an official daily monitoring periodically	Mitigate
6	V6	Prepare strategic plan, annual projection, budget	Mitigate
7	V7	Enable department of collections	Mitigate
8	V8	Request 3 to 5 young people "Building the future" project.	Mitigate
9	V9	No selected	No response
10	V10	Follow the planning, and monitoring for innovated strategies	Mitigate
11	V11	Request weekly reports	Avoid

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Test	Run #	JT	CT	f(x)	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	Cluster
32	2	1	988.5	53.631264	7.670808	Ac	Ev	Mi	Null	Null	Tr	Null	Null	Mi	Mi	Tr	C2
33	2	2	2196	62.105927	8.678924	Ac	Ev	Mi	Null	Mi	Mi	Mi	Mi	Mi	Mi	Ev	C1
34	2	3	220	28.980467	4.967917	Null	Null	Null	Mi	Null	Tr	Null	Null	Mi	Null	Null	C3
35	2	4	154.75	29.839687	2.967629	Ac	Null	Null	Null	Null	Ev	Null	Null	Null	Mi	Ac	C3
36	2	5	2061	65.50737	8.73995	Ac	Null	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Null	C1
37	2	6	250.5	13.789864	1.028718	Mi	Ev	Null	Null	Null	Ev	Null	Null	Null	Null	Null	C3
38	2	7	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1
39	2	8	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1
40	2	9	2256	65.172752	7.814604	Ac	Ev	Mi	Mi	Mi	Mi	Mi	Null	Mi	Mi	Ev	C1

FIGURE 9. User interface to select a near-optimal solution.



FIGURE 10. User interface to select a near-optimal solution.

2) CPLEX and SA results for the Mexican company: Table 11 shows all different near-optimal solutions obtained by the SA algorithm and belonging in cluster 1, as well as the solution found using CPLEX through the Rglpk package.

TABLE 11. SA AND CPLEX SOLUTIONS FOR THE MEXICAN COMPANY.

JT	CT	OBJECTIVE FUNCTION VALUE	SOLUTION
<i>SA solutions in Cluster 1:</i>			
2261	60.93	4.071	{2261, 60.92, (1,4,2,2,2,2,2,2,0,2,4)}
2257	60.97	6.937	{2257, 60.97, (1,4,2,2,2,2,2,2,0,4)}
2256	65.17	7.815	{2256, 65.17, (1,4,2,2,2,2,2,0,2,2,4)}
2254	54.29	3.196	{2254, 54.29, (2,4,2,2,2,2,2,0,0,2,4)}
2196	62.11	8.868	{2196, 62.11, (1,4,2,0,2,2,2,2,2,2,4)}
2061	65.50	8.739	{2061, 65.50, (1,0,2,2,2,2,2,2,2,2,0)}
<i>CPLEX solution:</i>			
2262	72.85	8.838	{2262, 72.85, (1,4,2,2,2,2,2,2,2,2,4)}

In this Table can be observed that the CPLEX solver obtains a better solution than those found by the SA algorithm, concerning the total justification of implementing the safeguards set, but using a total cost higher than the maximum budget defined by the company.

3) Random instances: Table 12 shows the results obtained by the CPLEX solver and the SA algorithm with the ten random instances. This Table shows the time (in seconds), the memory used (in MB), and the value of the objective function of the mathematical model introduced in this paper. The last column in Table 12 shows the relative error of the near-optimal solution obtained by SA with reference to the solution reached by the CPLEX solver.

TABLE 12. SA AND CPLEX SOLUTIONS FOR RANDOM INSTANCES.

INSTANCE	CPLEX			SA			RELATIVE ERROR
	TIME	MEMORY	OBJECTIVE	TIME	MEMORY	OBJECTIVE	
	SEC	MB	FUNCTION VALUE	SEC	MB	FUNCTION VALUE	
1	0.22	30.2	15.9415	1.67	0.4	15.9415	0.0000
2	0.23	30.4	14.5884	2.66	0.4	14.5884	0.0000
3	0.23	30.6	18.7217	5.52	0.4	18.7217	0.0000
4	0.23	30.5	16.4735	7.19	0.4	16.47347	0.0000
5	0.23	30.6	14.0004	11.38	0.4	14.00040	0.0000
6	0.24	30.9	11.1656	13.1	0.4	11.16927	0.0003
7	0.25	31.4	33.0671	56.11	0.5	33.08332	0.0005
8	0.35	32.0	38.5896	142.1	0.6	38.60532	0.0004
9	0.40	32.6	44.2991	277.3	0.8	44.33145	0.0007
10	0.48	33.1	55.0765	496.4	1.1	55.14620	0.0013

It is observed that the SA algorithm always reach near-optimal solutions, using less memory than that used by the CPLEX solver, as is shown in Fig. 11. On the other hand, the runtime of the SA algorithm increases as the problem size rises.

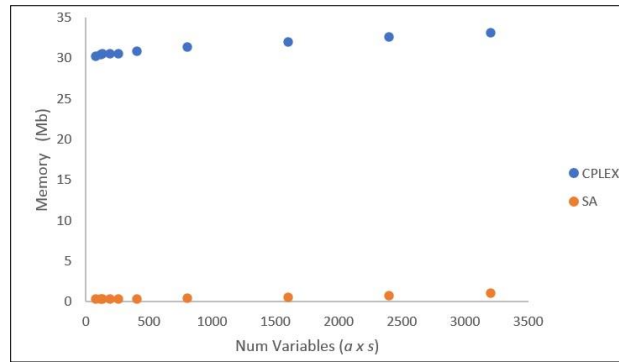


FIGURE 11. Memory used by the compared procedures.

### C. STATISTICAL ANALYSIS

Before performing the statistical analysis of the results generated by the compared methods, one study of the three conditions to apply a parametric test is conducted: independence, normality, and homoscedasticity [91]. The independence condition is evident since they are independent runs of the compared algorithms. Normality is verified through one Q-Q plot analysis to compare the quartiles from the data observed. Finally, a homoscedasticity analysis to check equality of variance is performed using the boxplot graph and the Bartlett test [92]. These conditions are evaluated for three compared measures: solutions quality, memory usage, and the runtime of the algorithms.

**1) Solution quality:** The Q-Q plot in Fig. 12 shows that there is no normality of the data in the two methods since the points do not lie on the diagonal of the graphs. Furthermore, the behavior observed in both graphs is practically the same as the data is very similar.

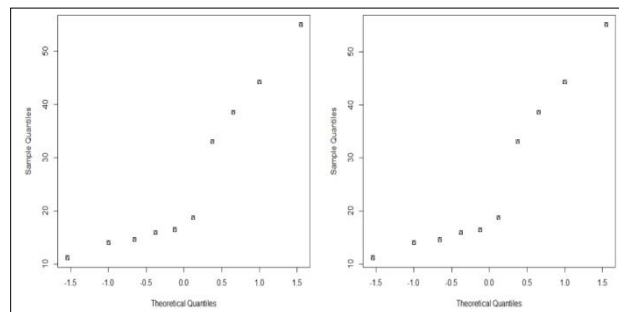


FIGURE 12. Q-Q plots for solutions quality.

Fig. 13 shows that the boxes for each method are the same, implying that data homoscedasticity exists.

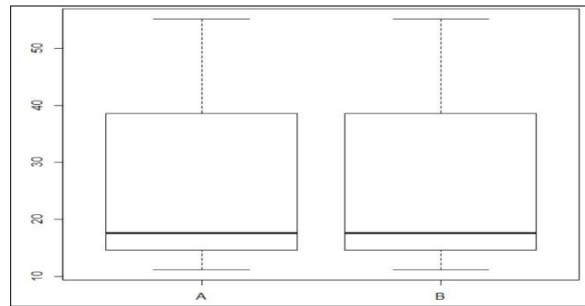


FIGURE 13. Box plots for solutions quality.

Since one of the three assumptions to apply a parametric test such as ANOVA is not fulfilled, a robust test such as the Welch and Box procedures [93]-[95] should be used. On the other hand, the near-optimal results of the SA algorithm are very similar to those reached by the CPLEX solver, with a relative error of 0.0013% in the worst case. If a statistical analysis using robust ANOVA is performed, the null hypothesis cannot be rejected, implying the compared algorithms have the same behavior. The SA behavior with the random instances is significant since the CPLEX solver always uses an exact method to obtain the optimal global solution, as long as the necessary computational resources are available, as in the case of the memory usage.

**2) Memory usage and running time:** Figure 16 presents the Q-Q plots to compare memory usage and running time of the compared procedures. In this Figure can be observed that normality condition is not accomplished both memory usage and running time.

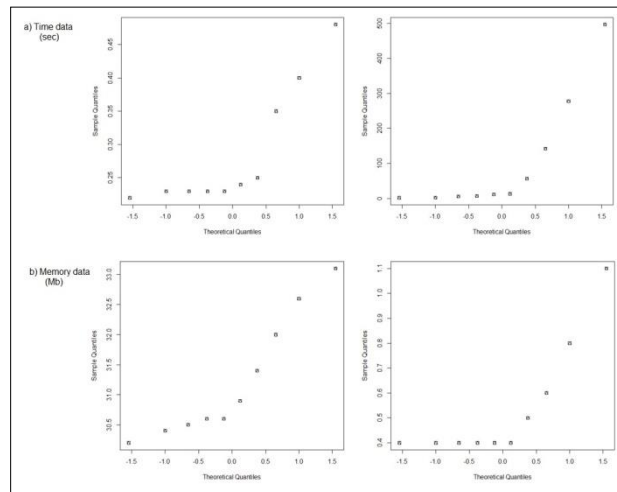


FIGURE 12. Q-Q plots for (a) running time, and (b) memory usage.

The homoscedasticity condition for the running time and memory usage are analyzed using the Bartlett test. The resulting p-values are  $2.2 \times 10^{-16}$  and  $4.978 \times 10^{-16}$ , respectively. Since both are less than a significance level of 5%, it is concluded that the results have different variances, and data homoscedasticity does not exist. In both cases, the Welch test is applied to verify if statistical differences exist in both variables. The p-values obtained are 0.1985711 and  $2.055245 \times 10^{-12}$ , respectively. These values indicate there is no statistical difference in running time, but it does exist for memory usage.

#### D. DISCUSSION

The development of this research refers to optimizing the responses of risk factors. Its importance lies in supporting the risk decision-making process using a limited budget, focus on improving the company's maximum benefits. The implementation of the methodology and the advantages of the solution found in the real case study depend on several elements provided by the company, such as (a) the information veracity, (b) the operative characteristics, c) the needs exposed, and d) the capabilities. It is observed that the SA algorithm generates near-optimal solutions for the optimization model constructed. These solutions are refined using machine learning procedures to be categorized in a group, providing to the decision-maker more than one possible solution, and, since the best group has repeated solutions, it is an indication that it may be the best solution. The more relevant risk factors to be treated in the solutions grouped in cluster 1 are identified using a decision tree induction procedure, showing that factors three and six should be mitigated.

With the experimental results of the ten instances created at random and the procedure implemented in this paper, the benefits of the solution categorization are perceived: The clustering algorithm identifies the best near-optimal solutions consistent with the budget limitation, and the CPLEX solver is unable to get a feasible solution. Regarding memory usage, the benefits of a non-deterministic heuristics, such as the SA algorithm, can be observed. The CPLEX solver uses the branch-and-cut algorithm implementing a systematic enumeration of solutions in a rooted tree, and consuming more memory resources than those used by the neighborhood structure used by the SA algorithm.

In light of the results obtained in this experimental study:

- 1) A new methodology considering the six more important risk areas in companies is developed. This methodology is applied in a real case study carrying out the identification and evaluation of risk factors. A matrix of safeguards is created by applying the William Fine method. When executing the SA algorithm, multiple near-optimal solutions are obtained. These solutions are categorized using a clustering algorithm by the resource use levels and their benefits. In the Mexican company, the risk factors  $V_6$  (Short and long term business planning) and  $V_3$  (Efficiency of the sales network) are identified. They are treated using the  $S_{6,2}$  (Prepare strategic plan, annual projection, budge) and  $S_{3,2}$  (Prepare advertising plan, direct sales team preparation and technology plataform) safeguards, respectively. Through a user-interface for decision-making, the business managers can use an informatics tool contributing to the generation of an attention plan to business risk factors. Being well managed, it would avoid business failure.
- 2) The optimization model is solved using the SA algorithm. The parameter tuning is performed based on the standard deviation of the ten dataset instances created at random. SA is effective in providing near-optimal solutions with a lower budget than that of the optimal solution reached by the CPLEX solver.
- 3) In the ten instances generated with simulated data, the SA algorithm efficiency and effectiveness are evaluated by contrast with the optimal solution generated by the CPLEX solver. SA obtains the global optimum in all cases in a reasonable time and consumes fewer memory resources, statistically proving that there is a means difference in the memory usage.

#### VI. CONCLUSIONS AND FUTURE WORK

The methodology presented in this work has shown to be effective for decision-making, helping managers obtain a set of near-optimal solutions with a limited budget. The SA algorithm provides a set of near-optimal solutions, being one of the research contributions. The optimal solution reached by the CPLEX solver, by using a limited budget, did not provide any solution, but the proposed methodology generates it. The proposed model minimizes the justification-cost ratio of implementing a solution, according to the ALARP principle. The results show that the maximum justification (the benefits for the company) is near the maximum possible value for this instance, and the cost is less than the budget assigned by the manager.

Concerning the research objectives, the methodology has been evaluated with real data of the company risk conditions. A user-interface for decision making is also developed. In this interface, the set of solutions classified in three groups is loaded: solutions consuming almost the entire budget with the most significant justification, low-cost solutions, they are solutions with lower cost but with a basic justification level (minimum security). Statistical analysis to evaluate significant differences between the compared methods shows that neither presents specific gaps in their efficacy and temporal behavior. On the other hand,

significant differences are found in the memory usage. This behavior indicates that the algorithmic proposal implemented with the SA algorithm for the optimization of risk decision-making can compete efficiently with the CPLEX solver.

This methodology is designed to adjust the number of responses to risk elements. Four response types (accept, mitigate, transfer, or avoid) are used in the real case study. These are the ones that are implemented in the best practices suggested by the PMI when analyzing risk treatment but can be expanded or reduced as appropriate, allowing to be used in various environments where the solution is structurally equivalent.

As a future work, the allocation of costs and benefits to the safeguards may be considered dependent on each other, that is, the degree of correction of a risk factor may vary depending on its selected treatment. This modification could be treated with a fuzzy-logic-based approach and a bi-objective optimization model. In addition, the use of this methodology may be expanded by integrating factors that do not present a risk to the company but that can promote its development providing with strategies increasing business strengths and opportunities.

### APPENDIX A: BUSINESS FACTORS QUESTIONNAIRE

The questionnaire is online in <https://goo.gl/forms/3R2LaqWDKsp8HvV33>. A paper-based questionnaire is shown in Fig. 13 and Fig. 14.

		S	W	O	T
		Strengths	Weaknesses	Opportunities	Threats
		Internal/External			
<b>i. Industry risk (IR):</b>					
1	Government policies and International agreements	€			
2	Cyclical	€			
3	Degree of competition	€			
4	The price and stability of market supply	€			
5	The size and growth of market demand	€			
6	The sensitivity to changes in macroeconomic factors	€			
7	Domestic and international competitive power	€			
8	Product Life Cycle	€			
<b>ii. Management risk (MR):</b>					
9	Ability and competence of management				
10	Stability of management	I			
11	The relationship between management/owner	I			
12	Human resources management	I			
13	Growth process/business performance	I			
14	Short and long term business planning	I			
15	Achievement and feasibility	I			
<b>iii. Financial Flexibility (FF):</b>					
16	Direct financing	I			
17	Indirect financing	€			
18	Other financing	€			

FIGURE 13. Paper-based questionnaire.



		S	W	O	T
		Strengths	Weaknesses	Opportunities	Threats
<b>iv. Credibility (CR):</b>					
19	Credit history	I			
20	Reliability of information	E			
21	The relationship with financial institutes	E			
<b>v. Competitiveness (CO):</b>					
22	Market position	I			
23	The level of core capacities	I			
24	Differentiated strategy	I			
<b>vi. Operating Risk (OP):</b>					
25	The stability and diversity of procurement	I			
26	The stability of transaction	I			
27	The efficiency of production	I			
28	The prospects for demand for product and service	I			
29	Sales diversification	I			
30	Sales price and settlement condition	I			
31	Collection of A/R	I			
32	Effectiveness of sale network	I			

Thank you for your participation.

FIGURE 14. Paper-based questionnaire.

## APPENDIX B: DECISION-MAKING USER INTERFACE

Fig. 15 shows a description of the user-interface. The elements of this interface are the follows:

- Select a test between the three SA test. Each one has 30 executions or solutions.
- Select the number of the solution, the cluster to which it belongs will appear in position c.
- Indicates the cluster to which the chosen solution belongs.
- The color of the bar indicates the section where the safeguards of the chosen solution for each risk factor  $V_i$  appear in a text, the type of each.
- Bars that indicate what type of response each safeguard has for the chosen solution.
- Values of the chosen solution. Justification is the maximum value close to the optimum of the benefit. Total Cost is the value in monetary units that are exercised by the said solution.
- It is the cost-benefit function that has been optimized.

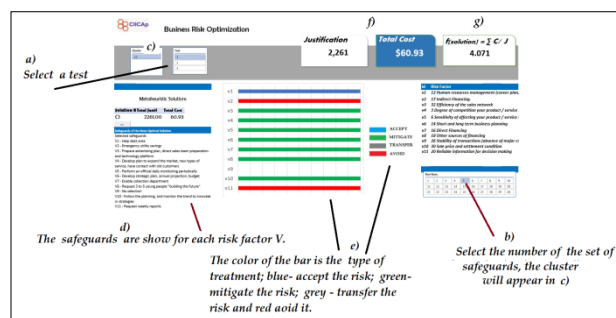


FIGURE 15. User-interface.

## REFERENCES

- [1] R.R. Moeller, *COSO enterprise risk management: understanding the new integrated ERM framework*. New Jersey, USA: Jhon Wiley & Sons, 2007.
- [2] S. Marcelino-Sádaba, A. Pérez-Ezcurdia, A.M.E. Lazcano, P. Villanueva, Project risk management methodology for small firms. *Int. Journal of Project Management*, vol. 32, no. 2, pp. 327–340, 2014.
- [3] A.L. Mesquida, A. Mas, A., A project management improvement program according to ISO/IEC 29110 and PMBOK. *Journal of Software: Evolution and Process*, vo. 26, no. 9, pp. 846–854, 2014.



- [4] M. Arena, M. Arnaboldi, G. Azzone, The organizational dynamics of enterprise risk management. *Accounting, Organizations and Society*, vol. 35, no. 7, pp. 659–675, 2010.
- [5] R. A. Bettis, Organizationally intractable decision problems and the intellectual virtues of heuristics. *Journal of Management*, vol. 43, no. 8, pp. 2620–2637, 2017.
- [6] J. Fu, X. S. Luo, A. M. Luo, J. X. Liu, Enterprise-level business component identification in business architecture integration. *Frontiers of Information Technology & Electronic Engineering*, vol. 18, no. 9, pp. 1320–1335, 2017.
- [7] R. Martínez-Hernández, M. del P. Pastor-Pérez, Interrelación entre riesgo e innovación: percepción del riesgo por gestores de proyectos. *Journal of Technology Management & Innovation*, vol. 13, n. 2, pp. 94–103, 2018.
- [8] J.S. Zamora-Salas, Salud Laboral, variable promotora de la productividad. *Revista de la Sociedad Española de Salud Laboral en la Administración Pública*, vol. 2, no. 10, pp. 4–9, 2010.
- [9] F.C. Roa-Herrera, *Bienestar laboral*, B.S. thesis, Fundación Universidad de América, 2019.
- [10] M. J. Pereira, B. K. Coombes, T. A. Comans, V. Johnston, The impact of onsite workplace health-enhancing physical activity interventions on worker productivity: a systematic review. *Occupational & Environmental Medicine*, vol. 72, no. 6, pp. 401–412, 2015.
- [11] D. Jones, D. Molitor, J. Reif, What do workplace wellness programs do? Evidence from the Illinois workplace wellness study, *The Quarterly Journal of Economics*, vol. 134, no. 4, pp. 1747–1791, 2019.
- [12] D.N. Tchiche, F. Gauthier, Classification of risk acceptability and risk tolerability factors in occupational health and safety, *Safety Science*, vol. 92, pp. 138–147, 2017.
- [13] R. Maguire, *Safety cases and safety reports: meaning, motivation and management*, CRC Press, 2017.
- [14] M. H. Ehrampoush, G. H. Halvani, M. T. Ghaneian, A. Dehghani, M. Hesami Arani, Usage of William Fine, PHA Methods For Environmental risks management and estimation of control costs in the Kavir Steel Complex, *Tolooebehdasth*, vol. 15, no. 1, pp. 10–25, 2016.
- [15] W.T. Fine, *Mathematical evaluations for controlling hazards*, Tech. Report, Naval Ordnance Lab, Maryland, USA, 1971.
- [16] N.R. Quiñones, O. C. Aguilar, M. L. O. Guerrero, Estimation of the risk associated to marketing of swine meat contaminated with Salmonella spp., employing the William T. Fine method. *Procedia Food Science*, vol. 7, pp. 137–140, 2016.
- [17] M. Meknatjoo, M. Omidvari, Safety risk assessment by using William–Fine method with compilation fuzzy DEMATEL in machining process, *Iran Occupational Health*, vol. 12, no. 5, pp. 31–42, 2015.
- [18] M. Ebrahimzadih, G. Halvani, E. Darvishi, F. Froghinasab, Application of Job Safety Analysis and William Fine Methods to Identify and Control Hazards in a Uranium Mine in Central Area of Iran, *Journal of Health*, vol. 6, no. 3, pp. 313–324, 2015.
- [19] G. Halvani, M. H. Ehrampoush, M. T. Ghaneian, A. Dehghani, M. H. Arani, Applying job hazard analysis and William Fine methods on risks identification and assessment of jobs in hot rolling steel Iran. *Journal of Mazandaran University of Medical Sciences*, vol. 26, no. 145, pp. 293–303, 2017.
- [20] K. H. Rose, A Guide to the Project Management Body of Knowledge (PMBOK R Guide), *Project Management Journal*, vol. 44, no. 3, pp. e1–e1, 2013.
- [21] T. L. Norman, *Risk analysis and security countermeasure selection*, CRC press, Florida, USA, 2016.
- [22] P. Bonami, A. Tramontani, Recent improvement to MISOCP in CPLEX, *INFORMS 2015*, Philadelphia, USA, 2015.
- [23] S. Eggenschwiler, M. Barry, R. Schumann, R. Exploring effects of parameter configurations on runtime, using mathematical solvers, *PuK 2018*, pp. 1–14, Berlin, Germany, 2018.
- [24] M. Barry, H. Abgottspon, R. Schumann, Solver tuning and model configuration, in: F. Trollmann, et al. (eds), *KI 2018*, LNCS, vol. 11117. Springer, Berlin, Germany, 2018.
- [25] S. Eggenschwiler, *RS: Parameter tuning for the CPLEX*, B.S. Thesis, University of Applied Sciences and Arts Western Switzerland, 2016.
- [26] R Core Team, R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria, pp. 1–16, 2019.
- [27] S. Theussl, K. Hornik, C. Buchta, F. Schwendinger, H. Schuchardt, M. S. Theussl, *Package ‘Rglpk’*, 2019.
- [28] R. Lima, E. Seminar, IBM ILOG CPLEX-What is inside of the box? *Proc. of the 2010 EWO Seminar*, pp. 1–72. Pittsburgh, USA, 2010.
- [29] C. H. Papadimitriou, K. Steiglitz, *Combinatorial optimization, Algorithms and Complexity*, Dover Publications, NY, USA, 1982.
- [30] K. Du, M. Swamy, *Search and Optimization by Metaheuristics*, Springer, 2016 (2016).
- [31] E. G. Talbi, *Metaheuristics: From Design to Implementation*, Wiley, 2006.
- [32] A. Nayyar, D. N. Le, N. G. Nguyen, *Advances in swarm intelligence for optimizing problems in computer science*. CRC Press, 2018.
- [33] R. Aguilar-Rivera, M. Valenzuela-Rendón, J. Rodríguez-Ortiz, Genetic algorithms and darwinian approaches in financial applications: A survey, *Expert Systems with Applications*, vol. 42, no. 21, pp. 7684–7697, 2015.
- [34] N. Metawa, M. K. Hassan, M. Elhoseny, Genetic algorithm based model for optimizing bank lending decisions, *Expert Systems with Applications*, vol. 80, pp. 75–82, 2017.
- [35] I. B. Salehpoor, S. Molla-Alizadeh-Zavardehi, A constrained portfolio selection model at considering risk-adjusted measure by using hybrid meta-heuristic algorithms, *Applied Soft Computing*, vol. 75, pp. 233–253, 2019.
- [36] M. Gilli, E. Kellezi, H. Hysi, *A data-driven optimization heuristic for downside risk minimization*, Swiss Finance Institute Research Paper no. 06-2, 2006.
- [37] K. Georgoulakos, K. Vergidis, G. Tsakalidis, N. Samaras, Evolutionary multi-objective optimization of business process designs with pre-processing, *CEC-2017*, IEEE, San Sebastian, Spain, pp. 897–904, 2017.
- [38] L. P. Rees, J. K. Deane, T. R. Rakes, W. H. Baker, Decision support for cybersecurity risk planning, *Decision Support Systems*, vol. 51, no. 3, pp. 493–505, 2011.
- [39] V. Viduto, C. Maple, W. Huang, D. López-Peréz, A novel risk assessment and optimisation model for a multi-objective network security countermeasure selection problem. *Decision Support Systems*, vol. 53, 599–610, 2012.
- [40] G. G. R. Guntur, Modeling an intrusion detection system using data mining and genetic algorithms based on fuzzy logic, *Int. Journal of Computer Science and Network Security*, vol. 8, no. 7, p. 319–325, 2008.
- [41] L. Khan, M. Awad, B. Thuraisingham, A new intrusion detection system using support vector machines and hierarchical clustering, *The VLDB journal*, vol. 16, no. 4, pp. 507–521, 2007.
- [42] A. I. Madbouly, A. M. Gody, T. M. Barakat, Relevant feature selection model using data mining for intrusion detection system, *Int. Journal of Engineering Trends and Technology*, vol. 9, no. 10, pp. 501–512, 2014.

- [43] V. Lakhno, S. Kazmirchuk, Y. Kovalenko, L. Myrutenko, T. Zhmurko, Design of adaptive system of detection of cyber-attacks, based on the model of logical procedures and the coverage matrices of features, *East European Journal of Advanced Technology*, vol. 3, no. 9, pp. 30–38, 2016.
- [44] D. Agustia, N. P. A. Muhammad, Y. Permatasari, Earnings management, business strategy, and bankruptcy risk: evidence from Indonesia, *Heliyon*, vol. 6, no. 2, pp. 1–9, 2020.
- [45] S. L. Chiang, M. S. Tsai, The valuation of deposit insurance allowing for the interest rate spread and early-bankruptcy risk. *The Quarterly Review of Economics and Finance*, vol. 76, pp. 345–356, 2019.
- [46] M. Büyükboyacı, M. Y. Gürdal, A. Kıbrıs, Ö. Kıbrıs, An experimental study of the investment implications of bankruptcy laws, *Journal of Economic Behavior & Organization*, vol. 158, pp. 607–629, 2019.
- [47] S.I. Chang, A. Huang, L. M. Chang, J. C. Liao, Risk factors of enterprise internal control: Governance refers to internet of things (iot) environment, *PACIS 2016*, pp. 1–19, 2016.
- [48] T. Seymour, S. Hussein, The history of project management. *Int. Journal of Management & Information Systems*, vol. 18, no. 4, pp. 233–240, 2014.
- [49] G. L. Richardson, *Project Management. Theory and Practice*, CRC Press, Florida, USA, 2014.
- [50] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, M. Protasi, *Complexity and approximation: Combinatorial optimization problems and their approximability properties*, Springer Berlin, Germany, 2012.
- [51] S. Kirkpatrick, C.D. Gelatt, M. P. Vecchi, Optimization by Simulated Annealing, *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [52] L. R. Rere, M. I. Fanany, A. M. Arymurthy, Simulated annealing algorithm for deep learning, *Procedia Computer Science*, vol. 72, pp. 137–144, 2015.
- [53] Z. Li, I. Kucukkoc, J. M. Nilakantan, Comprehensive review and evaluation of heuristics and meta-heuristics for two-sided assembly line balancing problem. *Computers & Operations Research*, vol. 84, pp. 146–161, 2017.
- [54] S. Asghari, N.J. Navimipour, Review and comparison of meta-heuristic algorithms for service composition in cloud computing, *Majlesi Journal of Multimedia Processing*, vol. 4, no. 4, pp. 28–34, 2015.
- [55] M. G. Martínez-Rangel, M. A. Cruz-Chávez, J. C. Zavala-Díaz, D. Juárez-Romero, O. Díaz-Parra, Analysis of the simulated annealing convergence in function of the standard deviation and the Boltzmann quotient for scheduling problems, *Research in Computing Science*, 4069, 282–293.
- [56] M. Jerrum, A. Sinclair, The Markov chain MonteCarlo method: An approach to approximate counting and integration. *Approximation algorithms for NP-hard problems*, pp. 482–520, 1996.
- [57] H. Sanvicente-Sánchez, J. Frausto-Solís, A Method to Establish the Cooling Scheme in Simulated Annealing Like Algorithms. In: A. Laganá et al. (eds) ICCSA 2004, LNCS, vol. 3045. Springer, Berlin, pp.755–763, Springer, 2004.
- [58] N. Paltrinieri, L. Comfort, G. Reniers, Learning about risk: Machine learning for risk assessment, *Safety Science*, vol. 118, pp. 475–486, 2019.
- [59] M. Stephenson, S. Amarasinghe, M. Martin, U. M. O’Reilly, Meta optimization: improving compiler heuristics with machine learning. *ACM SIGPLAN Notices*, vol. 38, no. 5, pp. 77–90, 2003.
- [60] F. Agakov, E. Bonilla, J. Cavazos, B. Franke, G. Fursin, M. F. O’Boyle, J. Thomson, M. Toussaint, C. K. Williams, Using machine learning to focus iterative optimization, *CGO’06*, New York, USA, pp. 305–305, 2006.
- [61] A. Rachman, R. C. Ratnayake, Machine learning approach for risk-based inspection screening assessment, *Reliability Engineering & System Safety*, vol. 185, pp. 518–532, 2019.
- [62] S. Chaturvedi, V. Mishra, N. Mishra, Sentiment analysis using machine learning for business intelligence, *ICPCSI, IEEE*, Chennai, India, pp. 2162–2166, 2017.
- [63] A. K. Jain, Data clustering: 50 years beyond K-means, *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.
- [64] M. Capó, A. Pérez, J. A. Lozano, An efficient approximation to the K-means clustering for massive data, *Knowledge-Based Systems*, vol. 117, pp. 56–69, 2017.
- [65] B. Yang, X. Fu, N. D. Sidiropoulos, M. Hong, Towards k-means-friendly spaces: Simultaneous deep learning and clustering, D. Precup et al. (eds), *ICML’17*, Sydney, Australia, vol. 70, pp. 3861–3870, 2017.
- [66] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. McLachlan, A. Ng, B. Liu, S. Y. Philip, Top 10 algorithms in data mining, *Knowledge and Information Systems*, vol. 14, pp. 1–37, 2008.
- [67] H. Xu, C. Ma, J. Lian, K.Xu, E. Chaima, Urban flooding risk assessment based on an integrated k-means cluster algorithm and improved entropy weight method in the region of Haikou, China, *Journal of Hydrology*, vol. 563, pp. 975–986, 2018.
- [68] C. Li, L. Sun, J. Jia, Y. Cai, X. Wang, Risk assessment of water pollution sources based on an integrated k-means clustering and set pair analysis method in the region of Shiyan, China, *Science of the Total Environment*, vol. 557, pp. 307–316, 2016.
- [69] W. Bao, N. Lianju, K. Yue, Integration of unsupervised and supervised machine learning algorithms for credit risk assessment, *Expert Systems with Applications*, vol. 128, pp. 301–315, 2019.
- [70] S. Chi, X. Li, Y. Tian, J. Li, X. Kong, K. Ding, C. Weng, J. Li, Semi-supervised learning to improve generalizability of risk prediction models, *Journal of Biomedical Informatics*, vol. 92, pp. 1–11, 2019.
- [71] J. F. McCarthy, W. G. Lehnert, Using decision trees for coreference resolution, R. Kohavi (ed), *IJCAI’95*, Montréal, Canada, 1995.
- [72] A. Almomani, P. Saavedra, E. Sánchez, Ensembles of Decision Trees for Recommending Touristic Items. J. Ferrández Vicente et al. (eds), *IWINAC 2017*, LNCS, vol. 10338, Springer, La Coruña, Spain, pp. 510–519, 2017.
- [73] L. Milani, S. Grumi, E. Camisasca, S. Miragoli, D. Traficante, P. Di Blasio, Familial risk and protective factors affecting CPS professionals’ child removal decision: A decision tree analysis study, *Children and Youth Services Review*, vol. 109, 104687, 2020.
- [74] M. P. Romero, Y. M. Chang, L. A. Brunton, J. Parry, A. Prosser, P. Upton, E. Rees, O. Tearne, M. Arnold, L. Stevens, Decision tree machine learning applied to bovine tuberculosis risk factors to aid disease control decision making, *Preventive Veterinary Medicine*, vol. 175, 104860, 2020.
- [75] K. A. Dauda, B. Pradhan, B. U. Shankar, S. Mitra, Decision tree for modeling survival data with competing risks, *Biocybernetics and Biomedical Engineering*, vol. 39, pp. 697–708, 2019.
- [76] Y. Cao, P. I. Rockett, The use of vicinal-risk minimization for training decision trees, *Applied Soft Computing*, vol. 31, pp. 185–195, 2015.
- [77] B. Lakshmi, T. Indumathi, N. Ravi, A Study on C.5 decision tree classification algorithm for risk predictions during pregnancy, *Procedia Technology*, vol. 24, pp. 1542–1549, 1016.
- [78] L. Breiman, J. Friedman, C. J. Stone, R. A. Olshen, *Classification and regression trees*, CRC Press, Florida, USA, 1984.
- [79] H. R. Lewis, M. R. Garey, D. S. Johnson, *Computers and intractability. A guide to the theory of NP-completeness*, W. H. Freeman & Co., USA, 1990.

- [80] M. T. Garey, D. S. Johnson, “Strong” NP-Completeness Results: Motivation, Examples, and Implications, *Journal of the ACM*, vol. 25, no. 3, pp. 499-508, 1978.
- [81] T. Achterberg, T. Koch, A. Martin, Branching rules revisited, *Operations Research Letters*, vol. 33, pp. 42–54, 2005.
- [82] J. E. Mitchell, Branch-and-cut algorithms for combinatorial optimization problems, *Handbook of Applied Optimization*, vol. 1, pp. 65–77, 2002.
- [83] T. A. Feo, M. G. C. Resende, Greedy Randomized Adaptive Search Procedures, *Journal of Global Optimization*, vol. 6, pp. 109–133, 1995.
- [94] G. Houben, K. Lenie, K. Vanhoof, A knowledge-based SWOT-analysis system as an instrument for strategic planning in small and medium sized enterprises, *Decision Support Systems*, vol. 26, pp. 125–135, 1999.
- [85] M.J. Kim, I. Han, The discovery of experts’ decision rules from qualitative bankruptcy data using genetic algorithms, *Expert Systems with Applications*, vol. 25, pp. 637–646, 2003.
- [86] E.H. Aarts, P.J. Van Laarhoven, Statistical cooling: A general approach to combinatorial optimization problems, *Philips Journal Research*, vol. 40, no. 4, pp. 193–226, 1985.
- [87] M. Maechler, P. Rousseeuw, A. Struyf, M. Hubert, K. Hornik, cluster: Cluster Analysis Basics and Extensions, *R package version 2.1.0*, 2012.
- [88] T. Therneau, B. Atkinson, B. Ripley, M B. Ripley, Package ‘rpart’, 2019.
- [89] S. Milborrow, Plotting rpart trees with the rpart. plot package, 2016.
- [90] F. Marqués, *Microsoft EXCEL 2010. En profundidad*, RC Libros, Madrid, España, 2010.
- [91] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, CRC Press, Florida, USA, 2003.
- [92] M. S. Bartlett, Properties of sufficiency and statistical tests, *Proceedings of the Royal Society, A*, vol. 160, pp. 318-33, 1937.
- [93] H. Liu, *Comparing Welch’s ANOVA, a Kruskal-Wallis test and traditional ANOVA in case of Heterogeneity of Variance*, M.S. Thesis, Virginia Commonwealth University, 2015.
- [94] R. R. Wilcox, *Introduction to robust estimation and hypothesis testing*, Academic Press, USA, 2011.
- [95] P. Mair, R. Wilcox, Robust statistical methods in R using the WRS2 package, *Behavior Research Methods*, vol. 52, pp. 464–488, 2020.



**MARTA L. ERAÑA-DÍAZ** was born in Tamaulipas, México. She received the B.S. in mathematics applied to computer science at the University Autonomous Metropolitan, Mexico City in 1986. In 2016, the M.S. in cognitive sciences and an upcoming Ph.D. degree in engineering and applied sciences, both from the University Autonomous of Morelos State in Cuernavaca, México. Since 1997, she has been a Professor with the Industrial Engineering Department, University Autonomous of Morelos State. Her research interest includes the design and development of technological tools for information management and data analysis.

Mrs. Eraña is the founder and partner of an information technology company, has been a technical manager and was part of the jury in transdisciplinary R&D projects, and is a member of the liaison committee of the National Center for Technological Research and Development.



**MARCO ANTONIO CRUZ CHÁVEZ** received the Ph.D. degree in computer sciences from the Tec de Monterrey, in 2004. Since 2004, he has been a Research Professor with the Research Center in Engineering and Applied Sciences (CIICAP), Autonomous University of Morelos State (UAEM). He is also a Leader of Grid Morelos Project, High Performance of Computing Laboratory. He is the Manager of High Performance Grid Morelos. He is also the Leader of Research Group Optimization and Software. He has 35 international publications and 30 national publications. Since 2005, he has been a Reviewer of international journals. He

is a National Researcher of Mexico (SNI II).



**RAFAEL RIVERA-LÓPEZ** was born in Poza Rica, Veracruz, México in 1965. He received the B.S. degree in computer systems engineering from the Instituto Tecnológico de Veracruz, in Veracruz, México in 1989, the M.S. degree in computer sciences from the Instituto Tecnológico y de Estudios Superiores de Monterrey, in Cuernavaca, Morelos, México, in 2000, and the Ph.D. degree in computer sciences from the Universidad Juárez Autónoma de Tabasco, in Cunduacán, Tabasco, México, in 2017. From 1991 to 1998, he was a systems analyst and a software developer in several companies in Mexico. Since 1992, he has been a research professor with the Computer Systems Department in the Instituto Tecnológico de Veracruz, in México. His research interests include the study and application of metaheuristics for solving complex problems and the implementation of object-oriented models in machine learning procedures. He holds the status of National Researcher (SNI C), in Mexico.



**BEATRIZ MARTÍNEZ BAHENA** was born in Taxco, Guerrero, Mexico in 1987. She was graduated as computer Engineering from Polytechnic University of Morelos, Mexico, in 2009. She received the M.Sc. and Ph.D. degrees in Engineering and Applied Sciences from Center for Applied Research in Engineering and Applied Sciences (CIICAp-UAEM), Mexico (2011 and 2016, respectively). From 2014 to 2019 she was professor in Polytechnic University of Morelos. She works, since 2014, at Autonomous University of Morelos, Mexico, where she is currently professor. She has participated in national and international conferences. She is an associate editor of the journal Mathematical Programming and Software. She has participated as an organizing committee in the International Congress on Computer Optimization and Software (CICOS). Her main research interest includes Theoretical and Applied Optimization, Parallel and Distributed Programming, Grid Computing and Network Design. She holds the status of National Researcher (SNI C), in Mexico.



**ERIKA YESENIA ÁVILA-MELGAR** was born in Coatlan del Rio, Morelos, Mexico, in 1976. She received the B.S. degree, in engineering, in computing systems, from Instituto Tecnológico de Zacatepec, in 2003. She received the M.S. degree, in computing sciences, from Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), in Cuernavaca, Morelos, Mexico, in 2008. She received the Ph.D degree, in Engineering, in Applied Sciences, in optimization and software area, from Universidad Autónoma del Estado de Morelos, in 2015. She has been titular professor, in the Chemistry and Engineering Faculty, for the Distance Education System, since 2011. She is a Research Assistant with the Optimization and Software Laboratory, in the Research Center of Engineering and Applied Sciences. She is author of one article and co-author of ten articles. Her research interests include the design of algorithms to solve combinatorial optimization problems, by applying distributed parallel programming in C language, with MPI library. She is also interested in solving mathematical models and water distribution networks designs problems, by designing and codifying high performance computing programs. She holds the status of National Researcher (SIN C), in Mexico.



**MARTÍN H. CRUZ ROSALES** received the Ph.D. degree from the Autonomous University of Morelos State. He was Research Professor with the Simulation and Nuclear Energy Department, Electrical Research Institute, Morelos, Mexico. He then became a Researcher and a Lecturer with the Science Faculty, Autonomous University of Morelos State, Morelos. He is currently a Researcher and a Lecturer with the Faculty of Accounting, Administration and Computing, Autonomous University of Morelos State, Morelos. He has several publications and articles on scheduling algorithms, and he has also given several lectures and seminars. His research interest includes the area of combinatorial optimization. He holds the status of National Researcher (SNI C), in Mexico.