

**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

**FACULTAD DE ECONOMÍA Y EMPRESAS
CARRERA NEGOCIOS INTERNACIONALES**

TEMA:

Análisis de factores predominantes en la rotación temprana de empleados de grupo Difare de la región Sierra aplicando Machine Learning.

AUTORAS:

Arroyo Becerra, Doris Nayeli

Ríos Arambulo, Sara María

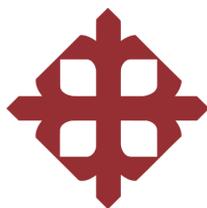
**Trabajo de titulación previo a la obtención del título de LICENCIADO EN
NEGOCIOS INTERNACIONALES**

TUTOR:

Ing. Carrera Buri, Félix Miguel, Mgs.

Guayaquil, Ecuador

15 de febrero del 2024



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA

CARRERA NEGOCIOS INTERNACIONALES

CERTIFICACIÓN

Certificamos que el presente trabajo de titulación fue realizado en su totalidad por **Arroyo Becerra Doris Nayeli** y **Ríos Arambulo Sara María**, como requerimiento para la obtención del título de **Licenciado en Negocios Internacionales**.

Guayaquil, a los 15 días del mes de febrero del año 2024

TUTOR

Ing. Carrera Buri, Félix Miguel Mgs.

DIRECTOR DE LA CARRERA

Ing. Hurtado Cevallos, Gabriela Elizabeth Mgs.



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA

CARRERA NEGOCIOS INTERNACIONALES

DECLARACIÓN DE RESPONSABILIDAD

Nosotras, **Arroyo Becerra Doris Nayeli** y **Ríos Arambulo Sara María**

DECLARAMOS QUE:

El Trabajo de Titulación: **Análisis de factores predominantes en la rotación temprana de empleados de grupo Difare de la región Sierra aplicando Machine Learning**, previo a la obtención del título de **Licenciadas en negocios internacionales** ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de nuestra total autoría. En virtud de esta declaración, nos responsabilizamos del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guayaquil, a los 15 días del mes de febrero del año 2024

LAS AUTORAS:

Doris Arroyo

Arroyo Becerra, Doris Nayeli

Sara Ríos.

Ríos Arambulo, Sara María



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA

CARRERA NEGOCIOS INTERNACIONALES

AUTORIZACIÓN

Nosotras, **Arroyo Becerra Doris Nayeli** y **Ríos Arambulo Sara María**

Autorizamos a la Universidad Católica de Santiago de Guayaquil a la **publicación** en la biblioteca de la institución del Trabajo de Titulación: **Análisis de factores predominantes en la rotación temprana de empleados de grupo Difare de la región Sierra aplicando Machine Learning**, cuyo contenido, ideas y criterios son de nuestra exclusiva responsabilidad y total autoría.

Guayaquil, a los 15 días del mes de febrero del año 2024

LAS AUTORAS:

Doris Arroyo

Arroyo Becerra, Doris Nayeli

Ríos Arambulo, Sara María



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA CARRERA NEGOCIOS

INTERNACIONALES

REPORTE COMPILATIO



CERTIFICADO DE ANÁLISIS
magister

Tesis Final Doris Arroyo Becerra - Sara Rios Arambulo

3%
Textos
sospechosos



3% Similitudes
< 1% similitudes
entre comillas
2% entre las
fuentes
mencionadas
< 1% Idioma no
reconocido
(ignorado)

Nombre del documento: Tesis Final Doris Arroyo Becerra - Sara Rios Arambulo.docx
ID del documento: 6e924556037d7f502b3ed814f8450128a9dc452d
Tamaño del documento original: 1,5 MB

Depositante: Felix Miguel Carrera Buri
Fecha de depósito: 26/1/2024
Tipo de carga: interface
fecha de fin de análisis: 26/1/2024

Número de palabras: 22.297
Número de caracteres: 151.162

Ubicación de las similitudes en el documento:



Fuentes principales detectadas

Nº	Descripciones	Similitudes	Ubicaciones	Datos adicionales
1	repositorio.ucsg.edu.ec http://repositorio.ucsg.edu.ec/bitstream/3317/18340/3/T-UCSG-PRE-CEAC-CNI-24.pdf.txt 37 fuentes similares	1%		Palabras idénticas: 1% (240 palabras)
2	repositorio.ucsg.edu.ec http://repositorio.ucsg.edu.ec/bitstream/3317/18340/1/T-UCSG-PRE-CEAC-CNI-24.pdf 33 fuentes similares	1%		Palabras idénticas: 1% (218 palabras)
3	Trabajo de titulación - 50% Carvache y Rodriguez (1).docx Trabajo de tit... #f8cd6b El documento proviene de mi grupo 29 fuentes similares	< 1%		Palabras idénticas: < 1% (141 palabras)

Ing. Félix Miguel, Carrera Buri, Mgs.

AGRADECIMIENTO

Quiero expresar mi agradecimiento a Dios por brindarme salud y vida para celebrar cada uno de mis logros. Además de guiarme en la consecución de mis metas y acompañarme en cada elección que he hecho.

Mi gratitud se dirige a mis padres por su apoyo constante e inagotable paciencia. Indudablemente, su dedicación ha sido el pilar que ha hecho viable todo este camino. Agradezco infinitamente su incansable labor de guiarme en la construcción de mi futuro. Son unos padres excepcionales, y me enorgullece ser su hija. Un agradecimiento especial a mí misma por la perseverancia y esfuerzo desplegados a lo largo de los años. Aunque nada ha sido sencillo, he sabido enfrentar cada obstáculo con valentía, siempre recordando que la clave del éxito radica en la confianza que deposite en mis habilidades, conocimientos y determinación.

También, agradezco a las grandes amistades que he cosechado en estos 4 años de estudio, a mis compañeros y jefes de trabajo por su respaldo incondicional durante este proceso, ya que cada uno ha desempeñado un papel fundamental en este logro, y estoy agradecida de contar con todos ustedes.

A todos, gracias por formar parte de este trayecto académico.

Arroyo Becerra, Doris Nayeli

DEDICATORIA

Este trabajo de investigación está dedicado a mis padres, José Arroyo y Doris Becerra, cuyo amor y apoyo ilimitado han sido la fuerza motriz detrás de cada logro no solo en mi carrera académica, sino que también deportiva en su momento. Son la inspiración que me ha impulsado a alcanzar el éxito a lo largo de mi vida.

A mis mejores amigos Lexy, Naomy, Pablo, Mirkala y Maoli, gracias por su comprensión y apoyo incomparable durante este desafiante proceso. Sus palabras y consejos han sido cruciales para mantenerme valiente frente a diversas adversidades. Gracias por siempre alentarme y ayudarme a entender que soy capaz de lograr todo aquello que tengo en mente incluso cuando ni yo misma lo creo. Me siento afortunada de contar con ustedes.

Al concluir, destino este proyecto a mis hermanos Kevin, Dylan y José Ivan con la esperanza de ser su guía y respaldarlos en cada etapa, tal como lo han hecho conmigo. Recuerden que nada es imposible siempre que se trabaje duro por ello.

Arroyo Becerra, Doris Nayeli

AGRADECIMIENTO

Primordialmente, quiero expresar mi agradecimiento a Dios por guiar mis pasos a lo largo de este arduo pero gratificante camino académico. De igual manera, a mis dos pilares fundamentales en la vida: mis padres, Sara Arámbulo y Carlos Ríos. A lo largo de los años, han sido la guía y el ejemplo que han forjado mi desarrollo personal y profesional. Agradezco el esfuerzo y sacrificio diario que han dedicado para brindarme grandes oportunidades y enseñanzas. Sin su dedicación incansable, nada de esto habría sido posible.

Asimismo, quiero expresar mi agradecimiento al apoyo incondicional de mis hermanas, Michelle Ríos y Carla Ríos. Ustedes han sido mi inspiración y ejemplo para seguir a lo largo de mi vida.

Indudablemente, agradezco a las amistades que he cultivado a lo largo de mis 20 años. A mis incondicionales del colegio, a las nuevas amistades de la universidad, y a cada compañero de trabajo que ha creído en mis capacidades. Su apoyo constante ha sido un pilar fundamental que me ha impulsado a afrontar nuevos retos, permitiéndome crecer tanto a nivel personal como profesional.

Y por último, pero no menos importante, mi más sincero agradecimiento a los docentes, personal de trabajo de la Facultad de Ciencias Económicas, Administrativas y Empresariales de la universidad, y todo aquel profesor que me brindó herramientas claves para cosechar el éxito que lo auguro desde el presente.

Ríos Arambulo, Sara María

DEDICATORIA

El presente trabajo está dedicado a mis padres, Sara Arámbulo y Carlos Ríos. Sin su apoyo incondicional, este logro no habría sido posible, gracias por brindarme las herramientas necesarias para cosechar los éxitos. Ustedes son mi claro ejemplo de superación tras un trabajo arduo.

A mis tíos, tías y demás familiares, en especial a ti, abuela Marina Arámbulo, gracias por ser mi ejemplo de trabajo arduo, amor y humildad. Asimismo, dedico este logro a mis dos grandes motivaciones personales, mis sobrinos Joaquín y Emilio. Espero ser siempre el ejemplo que desean seguir, y más aún, superar. Que este trabajo de titulación sea un testimonio tangible de que, con esfuerzo y dedicación, podemos alcanzar y superar nuestros objetivos.

Finalmente, me dedico este logro a mí misma. Gracias, Sara Ríos, por tu valentía, persistencia y autoconfianza. Aprecio la fuerza interior que has demostrado en cada desafío, la resiliencia ante las adversidades y el constante esfuerzo por crecer. Este logro es testimonio de tu dedicación y autoaprecio. Que esta auto-dedicación sea un recordatorio constante de tu capacidad para alcanzar tus metas.

Ríos Arambulo, Sara María



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES
TRIBUNAL DE SUSTENTACIÓN**

Ing. Felix Miguel Carrera Buri, Mgs
TUTOR

Ing. Gabriela Hurtado Cevallos, Gabriela Elizabeth, Mgs
DIRECTOR DE CARRERA

COORDINADOR DE ÁREA



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES**

CALIFICACIÓN

Doris Arroyo

Arroyo Becerra, Doris Nayeli

Sara Arambulo

Ríos Arambulo, Sara María

ÍNDICE GENERAL

Capítulo I	6
Problemática	6
Justificación	9
Alcance	11
Objetivos	12
Objetivo General	12
Objetivos Específicos	12
Preguntas de Investigación	13
Marco Teórico	14
Inteligencia Artificial	14
Machine Learning	15
Tipos de Aprendizaje	16
Técnicas de Machine Learning	18
Deep Learning	23
Marco Referencial	24
Marco conceptual	35
Rotación de Personal	35
Intención de Rotación de Personal	36
Desempeño Laboral	36
Factores Intrínsecos	36
Factores Extrínsecos	36
Inteligencia Artificial	37
Machine Learning	38
Matriz de Confusión	39
Outliers	39
Business Intelligence	40
Marco Legal	41
Normas Internacionales	41
Normas Nacionales	42
Metodología	43

Procedimiento en RStudio	45
Resultados	51
Análisis de Resultados	58
Conclusiones	61
Anexos	73
Anexo 1	73
Base de datos (Neta)	73
Anexo 2	75
Extracción de la base consolidada de Probabilidades en Rstudio.....	75

ÍNDICE DE TABLAS

Tabla 1	31
Tabla 2	51
Tabla 3	57
Tabla 4	57

ÍNDICE DE FIGURAS

Figura 1	16
Figura 2	17
Figura 3	18
Figura 4	19
Figura 5	21
Figura 6	22
Figura 7	44
Figura 8	52
Figura 9	52
Figura 10	53
Figura 11	53
Figura 12	54
Figura 13	54
Figura 14	55
Figura 15	55
Figura 16	56
Figura 17	56

RESUMEN

La Inteligencia Artificial (IA) se centra en el desarrollo de sistemas que puedan realizar tareas que normalmente requieren inteligencia humana, incluidas actividades como el aprendizaje, la resolución de problemas, el reconocimiento de patrones y la toma de decisiones. La Inteligencia Artificial tiene una variedad de aplicaciones, desde asistentes virtuales y chatbots hasta algoritmos de recomendación. Su continuo desarrollo ha tenido un impacto significativo en áreas como la salud, la industria y la investigación, cambiando la forma en que interactuamos con la tecnología y creando nuevas oportunidades en una variedad de campos.

El presente trabajo de investigación tuvo como objetivo principal elaborar un modelo de Machine Learning para clasificar los factores predominantes que influyen en la rotación temprana de empleados de Grupo Difare, de la región Andina del Ecuador, lo que a su vez permita tomar decisiones estratégicas acerca del pronóstico del comportamiento a futuro. Dicho análisis fue llevado mediante la implementación del algoritmo Random Forest o Bosque Aleatorio, en el software estadístico RStudio. Lo que permitió evaluar las causas principales de las renuncias voluntarias del personal de talento de la entidad comercial, y por consiguiente, brindar recomendaciones para prevenir la tendencia creciente del fenómeno de rotación de personal, el mismo que tiene un impacto en la productividad y dinámica organizacional.

Palabras claves: Inteligencia Artificial, Machine Learning, Random Forest, Rotacion Temprana de Personal.

ABSTRACT

Artificial Intelligence (AI) focuses on the development of systems that can perform tasks that normally require human intelligence, including activities such as learning, problem solving, pattern recognition and decision making. Artificial Intelligence has a variety of applications, from virtual assistants and chatbots to recommendation algorithms. Its continuous development has had a significant impact in areas such as healthcare, industry and research, changing the way we interact with technology and creating new opportunities in a variety of fields.

The main objective of this research work was to develop a Machine Learning model to classify the predominant factors that influence the early turnover of employees of Difare Group, in the Andean region of Ecuador, which in turn allows making strategic decisions about the forecast of future behavior. This analysis was carried out through the implementation of the Random Forest algorithm in the statistical software RStudio. This allowed us to evaluate the main causes of voluntary resignations of talented personnel of the commercial entity, and consequently, to provide recommendations to prevent the growing trend of the phenomenon of staff turnover, which has an impact on productivity and organizational dynamics.

Keywords: Artificial Intelligence, Machine Learning, Random Forest, Early Staff Rotation.

RÉSUMÉ

L'intelligence artificielle (IA) se concentre sur le développement de systèmes capables d'effectuer des tâches qui requièrent normalement l'intelligence humaine, y compris des activités telles que l'apprentissage, la résolution de problèmes, la reconnaissance de formes et la prise de décision. L'intelligence artificielle a de nombreuses applications, depuis les assistants virtuels et les chatbots jusqu'aux algorithmes de recommandation. Son développement continu a eu un impact significatif sur des domaines tels que les soins de santé, l'industrie et la recherche, modifiant la façon dont nous interagissons avec la technologie et créant de nouvelles opportunités dans une variété de domaines.

L'objectif principal de ce travail de recherche était de développer un modèle d'apprentissage automatique pour classer les facteurs prédominants qui influencent la rotation précoce des employés de Difare S.A., dans la région andine de l'Équateur, ce qui permet à son tour de prendre des décisions stratégiques sur la prévision du comportement futur. Cette analyse a été réalisée en implémentant l'algorithme Random Forest dans le logiciel statistique RStudio. Cela a permis d'évaluer les principales causes des démissions volontaires du personnel talentueux de l'entité commerciale et, par conséquent, de fournir des recommandations pour prévenir la tendance croissante de la rotation du personnel, qui a un impact sur la productivité et la dynamique organisationnelle.

Mots-clés : Intelligence Artificielle, Apprentissage Automatique, Forêt Aléatoire, Rotation Précoce du Personnel.

Introducción

En la actualidad, en lo que concierne al ámbito laboral, se evidencia una alta tasa de desempleo no solo en Ecuador, como país en vía de desarrollo, sino que también a nivel mundial. En adición, esto está ligado a la rotación del personal existente en las compañías, mismas que concentran una gran cantidad de horas para el análisis enfocado en detectar las causas que están generando aquella rotación, para de esta manera, aplicar estrategias empresariales que permitan la disminución de esta.

Las organizaciones se enfrentan regularmente a cuestiones como la rotación de empleados y los requisitos de formación. En una economía como la que tenemos actualmente, que se caracteriza por un entorno de constante transformación, dinamismo y complejidad, la integración de trabajadores experimentados y plenamente capacitados a los procesos productivos es fundamental. Una función esencial de la gestión empresarial es retener a quienes obtienen información y luego convierten esa información en conocimiento que pueda distribuirse por toda la organización, lo que también permite que el negocio mejore y sea más productivo. Una pérdida de conocimiento resulta de la partida de cualquier contribuyente importante, así como una disminución en términos económicos. De igual manera, las empresas realizan un análisis de la dotación presente en la misma, así como también de la rotación laboral a la cual se enfrentan en periodos de corto plazo.

Por este motivo es indispensable para las compañías contar con una transformación digital en sus procesos que les permita competir en el mercado y adherirse a un entorno económico más estable. El uso de la inteligencia artificial ha tenido un gran impacto dentro del sector farmacéutico ya que con una recolección de datos eficaz y un software conveniente para el estudio se puede lograr tener una gran precisión de los posibles resultados que una empresa puede llegar a presenciar en sus procesos. De ahí que, la importancia del uso de nuevas tecnologías las cuales les permiten a las compañías mantener un mayor control en sus diferentes procesos como la investigación de mercados, desarrollos logísticos, reclutación de personal, entre otros. Esto con el propósito de aumentar la competitividad de estas dentro del sector farmacéutico local e internacional.

Por otro lado, la rotación es un fenómeno que afecta a las organizaciones a nivel mundial. Además, existen varios elementos que pueden influir en este movimiento y parten desde la economía global y las oportunidades laborales presente en cada país hasta la satisfacción laboral

y su cultura organizacional. Alvares (2020) explicaron que “la rotación de personal es la separación laboral unilateral por parte de un trabajador de una organización que deja a la misma con un desajuste en su fuerza de trabajo” (p.14), lo que a su vez implica un coste en la nómina.

Mientras que, (Aranibar et al., 2018) afirmaron que “la rotación de personal no es más que el movimiento o migración de los empleados o trabajadores que se da de una empresa u organización a otra en un determinado periodo de tiempo y por causas diversas” (p.26).

La comprensión de la rotación de personal se ha vuelto evidente en varias organizaciones mundiales debido a que estas buscan adaptarse a las cambiantes dinámicas laborales que les permitan la retención de recursos necesarios para prosperar en un entorno empresarial cada vez más desafiante y diverso.

En vista de ello, en el sector farmacéutico se muestra que la rotación es un desafío constante debido a la alta demanda de talento especializado y a la naturaleza global de la industria, a pesar de que, en el año 2020 a nivel mundial el mercado farmacéutico “alcanzó los 974.000 millones de euros, aproximadamente un 3,7% más que el año anterior. En términos de ubicaciones, más de las tres cuartas partes de las ventas totales del mercado farmacéutico provienen de América del Norte, Europa y Japón” (Instituto Español de Comercio Exterior, 2022, p. 9).

Por lo consiguiente, en España, según Del Río (2023), el sector farmacéutico estaba compuesto por 42.600 empleados de trabajo directo y 170.000 de forma indirecta, en total 212.600 trabajadores en la industria. Donde se puede destacar que el 94% de los contratos laborales tuvieron una duración indefinida y más del 62% de los colaboradores tenían un título universitario. Por lo que, se puede evidenciar que la pandemia generó un impacto positivo en el ámbito laboral del sector, por la alta demanda de productos para la salud. Lo que a su vez generó que los farmacéuticos tuvieran la necesidad de adaptar sus farmacias, en pequeños dispensarios de atención a clientes, como también, adaptarse a la nueva ola de delivery y vender sus productos, con envío incluido a los destinos finales donde se encontraba el consumidor final (Carbonell,2022).

En adición, se debe tener en cuenta de que las empresas han tenido un margen de utilidad alto, sin embargo, existe un rubro considerable de gastos relacionados con las tasas de rotación voluntaria de las empresas a nivel mundial, creados por los costos relacionados con los ingresos, retiros y capacitaciones. Además, existen, costos intangibles relacionados a la adaptabilidad, calidad y servicio brindado a los clientes por los auxiliares y administradores de los diversos puntos de venta (Globalratings, 2022).

Así mismo, en el continente asiático, varias empresas han presentado una rotación elevada de personal por motivos de sus instalaciones, mismas que no son óptimas para el desarrollo de las actividades laborales. Lo que ha conllevado a que se originen renunciaciones voluntarias, ya que los colaboradores no sienten seguridad de realizar su jornada laboral bajo malas condiciones.

Por otra parte, es importante destacar que la rotación de personal en las entidades comerciales se encuentra presente desde la década de 1920 en Estados Unidos, según Macas et al., (2022) cuando las organizaciones identificaron que el accionar de despido de personal representaba una pérdida de recursos.

En paralelo, tras el surgimiento de la pandemia de Coronavirus (COVID-19), en el 2021 surge un nuevo fenómeno denominado la “Gran Dimisión”, donde los empleados buscaban alternativas de trabajo desde casa para lograr obtener un equilibrio laboral y personal tras el temor de contagio, y por consecuencia, se dio inicio a una desmesurada ola de renunciaciones que afectó al sector económico de la región (Portalo & Portalo, 2022).

De acuerdo con Adecco, 2022, en el continente de América Latina, el porcentaje de renunciaciones voluntarias en el año 2022 incrementó en un 12%, mientras que el 5% de colaboradores presentaron planes de renunciar en menos de 6 meses, el 8% en un periodo de 1 año, y en 2 años el 10%. Un dato para destacar es que Perú en el 2022 se convirtió en el país de Latinoamérica con más renunciaciones, al ascender a 20% (Arroyo, 2022).

Estas estadísticas se han visto sustentadas por el contexto de los últimos años en el continente. En primer lugar, la inestabilidad social derivada de conflictos políticos y la informalidad de América Latina. Asimismo, el contexto de trabajo híbrido ha generado un gran impacto en la mente de los trabajadores, haciendo hincapié a la Generación Z, puesto que las exigencias con respecto a beneficios laborales son cada vez más concretas y directas.

Según entrevistas realizadas por Deloitte Consulting (2021) a 157 entidades nacionales y multinacionales, el 91% de las empresas en Ecuador en el 2021, año post pandemia, aseguraron haber contado con un alto índice de rotación de personal por diversas razones, donde destacaron, el cambio en las políticas salariales, reestructuración de las entidades o por renunciaciones voluntarias.

Adicionalmente, a pesar del contexto de la pandemia donde en el año 2020 la economía ecuatoriana se vio afectada con una caída de su Producto Interno Bruto (PIB) de 7.8 puntos (Banco Central del Ecuador, 2023). Uno de los sectores económicos que logró mantener un desempeño

positivo de ventas fueron las empresas farmacéuticas, donde en el 2020, se incrementaron un 13%, y en el primer cuatrimestre del 2021 un 17%, valor significativo para la economía del país (Deloitte Consulting, 2021).

En síntesis, se puede evidenciar que, en los últimos dos años, las empresas farmacéuticas a nivel global y nacional han presentado una reducción de estructura post-pandemia, lo que a su vez ha implicado gastos de nómina, por motivos de la rotación presente. Lo que a su vez ha generado que las empresas tengan como prioridad mantener a sus colaboradores satisfechos, cumplir con sus obligaciones ante la ley, pero además priorizar la activación de incentivos económicos que le permitan mantener la fidelidad de estos.

Capítulo I.

Problemática

Distribuidora Farmacéutica Ecuatoriana Difare S.A., a lo largo de 39 años se ha caracterizado en el mercado farmacéutico ecuatoriano por ser una entidad comercial de capital 100% nacional, el mismo que tiene como objetivo mejorar la salud y el bienestar de las personas impulsando el crecimiento del sector a través del consumo masivo de sus productos, por medio de sus cadenas de retail en el territorio ecuatoriano, como lo son: Pharmacys, Farmacias Comunitarias, Cruz Azul y Dromayor. Sin embargo, la empresa desde el año 2021 hasta la actualidad, ha venido experimentando un margen significativo del 25% de rotación de personal, tanto a nivel corporativo como a nivel de retail.

La compañía ha evidenciado que el nivel de rotación ha estado en aumento en las diferentes partes del país en donde sus cadenas tienen alto nivel de presencia. En la Región Costa en el año 2022 el porcentaje de rotación tuvo un aumento del 7% versus al año anterior. Así mismo, en la Región Sierra el porcentaje de aumento en el año 2022 fue de 18% especialmente en la cadena Cruz Azul, donde los colaboradores presentaron la renuncia voluntaria antes de los 3 meses, es decir durante su periodo de 90 días de prueba.

Siendo así, que esta problemática refleja un gasto representativo en la nómina de la compañía ya que se invierten recursos cuantitativos y cualitativos en la contratación de cada uno de los colaboradores. Y estos, al retirarse de la entidad de forma voluntaria o por falta de cumplimiento a requisitos esenciales del puesto, generan un retraso en el flujo de la misma. Debido a que la posición queda libre pero la demanda del sector farmacéutico desde el 2021 año post pandemia va en crecimiento cada vez más. Por lo tanto, se deben suplir las necesidades de cada uno de los puntos de ventas que requieren de personal capacitado lo que como primera opción ha derivado a la contratación de personal backorder, mismo que suple las posiciones durante un periodo determinado hasta que se logre la captación de nuevo personal.

Además, es importante destacar que una característica que le da valor agregado al proceso de selección de Grupo Difare, es el enfoque en el cumplimiento de contratación CONADIS mismo que es del 4% a nivel nacional. (Código del trabajo, 2009). Sin embargo, muchos de estos colaboradores no han logrado adaptarse a las exigencias laborales y al entorno de esta, por ende,

se evidencia su renuncia voluntaria durante el periodo de prueba, lo que continúa afectando a la nómina de la compañía.

En la actualidad, la organización a nivel nacional mantiene un enfoque directo en la disminución del nivel de rotación presente en la región Sierra, a través de la implementación de KPI 's tales como: Nivel de Rotación, mismo que busca medir la variación existente entre un mes y otro. Y, por otro lado, el Cumplimiento de Dotación Aprobada, mediante el cual se busca controlar la cantidad de contrataciones que se hacen de forma trimestral. Siendo así, que con estos indicadores la empresa busca equilibrar y mejorar su problemática. Sin embargo, esta medición se está realizando de forma empírica, misma que no permite evidenciar la raíz del problema.

En otras palabras, la entidad comercial no cuenta con un análisis eficaz que permita la correcta identificación de factores determinantes de la rotación temprana de personal de puntos de ventas de retail, para ser más específicos de la región Sierra del Ecuador. Lo que conlleva un gran reto para el área de Talento Humano junto a Planeación Estratégica, puesto que a pesar de la implementación de los dos KPIs, como lo son cumplimiento de dotación y nivel de rotación, no se logra evidenciar las posibles razones de salidas de colaboradores en periodo de 3 meses (periodo de prueba). Por ende, no permite una correcta visualización de la data, que es valiosa. Por consiguiente, las tomas de decisiones de parte de los gerentes y miembros del directorio no han dado resultados fructíferos para reducir este margen de rotación.

De tal manera, que de ahí deriva la importancia de la implementación de ramificaciones de la inteligencia artificial, como lo es machine learning. La misma que se encarga de aprender las reglas estáticas codificadas en un programa por parte del área usuaria, para así, por consiguiente, poder identificar patrones de entrada y a través de algoritmos desarrollados, llevar a cabo sus funciones destinadas a lo largo del tiempo (Jiménez & Díaz, 2021).

En adición, dentro de este subcampo de la ciencia, se puede implementar un algoritmo de clasificación, denominado Random Forest. El mismo que se caracteriza por ser una técnica de aprendizaje supervisado, donde se consolida un conjunto de árboles de decisión para la clasificación, con el objetivo de emplear simultáneamente el aprendizaje de estos y poder otorgar una predicción eficiente (Lizares, 2019).

Por lo tanto, el aprendizaje automatizado en conjunto de esta técnica de clasificación les otorga a las organizaciones una mejor visibilidad empresarial para lograr una oportuna

identificación de patrones, lo que a su vez refleja predicciones correctas, y al contar con una considerable cantidad de datos, proveerá un resultado final con mayor eficiencia.

Por ende, a través de la implementación de machine learning, se busca que Ecuador, como país en vía de desarrollo, opte por una ciencia de gran utilidad y demanda en compañías del primer mundo. Siendo más específicos, para Grupo Difare, la misma que pueda obtener una visión clara, real y objetiva de los factores influyentes en este movimiento basada en procesos estadísticos. En conclusión, la implementación de esta ramificación de la inteligencia artificial permite la toma de decisiones basadas en información real, y no por inferencias. Por lo tanto, estas decisiones a futuro tienen consigo efectos positivos para la empresa, tanto en términos económicos como sociales.

Justificación

El presente trabajo de indagación tiene como enfoque principal analizar la rotación temprana de empleados de Grupo Difare, y a su vez lograr identificar los factores determinantes que influyen en dicho fenómeno, a través de la implementación de la inteligencia artificial en conjunto al método no paramétrico denominado Random Forest o también conocido como Bosque Aleatorio. Para, por consiguiente, poder tener una visión mejorada de la realidad de la entidad comercial, y a su vez a futuro, poder disminuir la rotación de la compañía en la región Sierra del Ecuador.

La Inteligencia Artificial (I.A.), al ser descrita como una disciplina dentro de la rama de la informática, la misma que se caracteriza por el desarrollo de dispositivos y sistemas que se encargan de emitir funciones cognitivas, propias del ser humano; permite realizar tareas de procesamiento y análisis de datos (Ahumada, 2016). Dentro de ella, se encuentran dos campos de gran relevancia, como son el aprendizaje automatizado y el aprendizaje profundo. El aprendizaje automático (machine learning) a través de la minería de datos permite extraer información por medio de análisis predictivos para identificar tendencias, y establecer patrones de comportamiento, sin ser programados explícitamente (López et al., 2019). Mientras que el aprendizaje profundo (deep learning), a través de un conjunto de códigos con diferentes capas de procesamiento, haciendo una imitación de redes neuronales, se encarga de aprender, procesar datos, y, por lo tanto, coopera en la identificación de la estructura intrincada en conjuntos de datos de gran dimensión (LeCun, Bengio, & Hilton, 2015). Por ende, se puede hacer hincapié, que el método de aprendizaje profundo describe una subárea del aprendizaje automático.

Por lo tanto, el uso de la Inteligencia Artificial (I.A) le permite a las compañías una mejor toma de decisiones mediante la automatización de tareas. Debido a que la I.A utiliza algoritmos y modelos matemáticos, los cuales otorgan una óptima clasificación de los datos. Esto se debe a que, mediante la aplicación de inteligencia artificial se puede identificar variables relevantes para análisis estadísticos, lo que a su vez simplifica y mejora la precisión de los modelos.

Siendo así que, una de las razones principales que impulsan esta investigación, es que en la actualidad se puede evidenciar las aplicaciones de la I.A en las empresas, las cuales persiguen integrar distintas estimaciones, tratamiento de datos y modelos para conseguir una aproximación

coherente y total en análisis de los mismos, para llevar a cabo una correcta interpretación de los resultados obtenidos, lo que a su vez permite el desarrollo competitivo de las entidades comerciales del sector farmacéutico.

Por otro lado, es importante destacar, que la empresa durante el 2022 llegó a tener una rotación laboral en la región Sierra del Ecuador del 24,61%, a diferencia de la rotación de personal en la Costa del territorio, que llegó a ser del 18,57%. Es decir, la región Andina tuvo una rotación laboral superior a 6,04% con respecto a la Costa, dato significativo para el análisis de la problemática. Por lo cual, se evidencia un desbalance en los colaboradores de los centros farmacéuticos, puesto que los mismos abandonan sus posiciones de trabajo en periodos cortos, inferior a los 90 días de prueba.

De igual forma, Difare S.A., está concentrando una gran cantidad de horas de análisis en encontrar las causas que están generando la rotación, ya que esto representa gastos significativos en la nómina, misma que se ve afectada tras la pérdida de un recurso. Además de que conlleva a un mayor esfuerzo por parte de los empleados de un punto de venta que pierden un miembro de su fuerza de trabajo, debido a que deben cubrir de forma momentánea la posición abandonada hasta el ingreso de un nuevo colaborador.

En síntesis, la importancia de esta investigación se basa en que al aplicar machine learning dentro de los análisis de rotación de personal, la organización podrá determinar los factores que inciden en el retiro de los colaboradores en cortos plazos mediante la correcta clasificación de datos, misma que permite una mejor interpretación de los resultados y de esta forma aplicar estrategias empresariales para la disminución del índice de renunciadas evidenciadas especialmente en la región Andina del Ecuador.

Alcance

El siguiente trabajo de titulación está dirigido hacia aquellas personas; tales como estudiantes de Estadística, Business Intelligence y a Gerencias de Talento Humano. Debido a que hoy en día, es imprescindible, que los estudiantes y nuevos profesionales se encuentren inmersos en nuevas ramificaciones, como lo es la Inteligencia Artificial, la misma que permite la automatización de procesos, a través de algoritmos, lo que simplifica largos procedimientos realizados de manera empírica en tiempos anteriores. Puesto que, la IA junto a la estadística es una necesidad en el mundo laboral evidente en la actualidad, por lo que, para las compañías es importante la contratación de nuevos talentos que aporten a la mejora continua de las mismas, mediante la aplicación o uso de nuevas tecnologías.

Siendo así que, para la alta gerencia estratégica de entidades comerciales, la presente investigación será de gran utilidad debido a que creará oportunidades para un mejor uso de la estadística con respecto a la toma de decisiones empresariales. Debido a que, siendo más específicos, los procesos de análisis que actualmente utiliza la empresa Distribuidora Farmaceutica ecuatoriana Difare S.A., se consideran obsoletos, por lo que se espera que el uso del modelo de clasificación facilite la toma de decisiones en el futuro.

En síntesis, la implementación de conceptos de machine learning mismos que trabajan bajo un modelo de Inteligencia Artificial otorgan un enfoque con mayor claridad respecto a la rotación temprana de los colaboradores de la compañía, lo que a su vez permite la proposición de mejores soluciones dentro del análisis.

Objetivos

Objetivo General

Elaborar un modelo de machine learning para clasificar los factores predominantes que influyen en la rotación temprana de empleados de Grupo Difare, de la región Andina del Ecuador, lo que permite tomar decisiones estratégicas acerca del pronóstico del comportamiento a futuro.

Objetivos Específicos

1. Analizar el marco teórico y conceptual a través de la revisión literaria acerca del tema de investigación.
2. Aplicar la metodología de la revisión literaria por medio de un modelo de análisis supervisado.
3. Analizar los resultados obtenidos e identificar los factores predominantes de la rotación de empleados de Grupo Difare, de la región Andina del Ecuador.

Preguntas de Investigación

1. ¿Cuál es la relación entre los factores predominantes y la rotación temprana de Grupo Difare, de la región Andina del Ecuador?
2. ¿Cómo afectaría la implementación de machine learning a la eficiencia del proceso de reducción de rotación de personal?
3. ¿En qué medida beneficia a Grupo Difare, la implementación de un modelo de machine learning para la identificación de los factores predominantes en la rotación temprana?

Capítulo II.

Marco Teórico

Inteligencia Artificial

Dentro de la ciencia de la informática, desde mediados del siglo XX se ha desarrollado un campo denominado inteligencia artificial (IA). La misma, que tuvo su origen a través de McCulloch & Pitts (1943), quienes crearon un modelo de redes neuronales artificiales, el mismo que tuvo como finalidad, demostrar que aquellas redes tenían la capacidad de aprender y lograr resolver ecuaciones lógicas proposicionales a partir de patrones de comportamientos desarrollados previamente. Siendo así, que se define a la IA, como aquella suficiencia propia de un sistema computacional, con aptitudes características de la conducta del cerebro humano, capaz de recibir y aprender patrones externos en función a la información otorgada, y, por consiguiente, a partir de entrenamientos, y aprendizaje, lograr llevar a cabo los objetivos para los cuales fueron entrenados (Brasil & Jorge, 2001; Incio et al., 2021).

Sin embargo, no fue hasta 1950 cuando Alan Turing difundió su artículo titulado *Computing Machinery and Intelligence*, donde sustentó un ensayo denominado la Prueba de Turing (García, 2012). De acuerdo con (Shah & Warwick, 2017) si un artificio puede ejercer como lo haría un humano, aquella máquina es inteligente. De esta suposición, partió su prueba, y permitió definir si un dispositivo era inteligente o no. A pesar del aporte de Turing; en 1958 Jon McCarthy llevó a cabo la creación del lenguaje de procesamiento de listas (LISP), convirtiéndose en el lenguaje de la Inteligencia Artificial. Simultáneamente junto a Minsky, Claude, Shannon y Newell iniciaron la época moderna de esta disciplina, por medio de la conferencia de verano en el Dartmouth College donde propusieron un proyecto de investigación de inteligencia artificial con el objetivo de estudiar los computadores, sus programas y teorías de automatización. (Barrera 2012).

Por lo tanto, es evidente que hasta la actualidad el campo de la Inteligencia Artificial (IA) ha ido evolucionado, logrando combinar análisis de datos, minería de datos, representación de datos, herramientas y fundamentos de los mismos, otorgando a varias áreas como la empresarial, ventajas competitivas que impulsan su crecimiento y a su vez les permite predecir situaciones en

las cuales puedan tomar decisiones prácticas y proactivas para el bienestar de todos los usuarios (Jiménez & Diaz, 2021).

Machine Learning

Dentro de los conceptos que comprende la Inteligencia Artificial, una de las ramas que ha marcado significancia en la era digital, es el machine learning (ML) o también conocido como aprendizaje automático. De acuerdo a Arthur (1959), fundador del término ML, define al aprendizaje automático como aquella ciencia que se caracteriza por la implementación de la tecnología junto a técnicas estadísticas y algoritmos informáticos, con el objetivo de otorgar a los ordenadores la capacidad de aprender, gestionar, y por ende, brindar resultados satisfactorios, tras llevar a cabo actividades específicas, sin la necesidad de instrucciones explícitas externas a las otorgadas por el programador.

De tal manera, que el aprendizaje de máquinas se ha diferenciado desde su origen hasta la actualidad, por aprender patrones a través del análisis de datos, y por medio de algoritmos genéricos lograr la clasificación automatizada (Bobadilla, 2020).

A mediados de 1979, en la universidad de Stanford un grupo de estudiantes llevaron a cabo un proyecto que consistía en la creación de un robot, al cual lo denominaron Stanford Car. Aquel robot estaba configurado con un algoritmo Nearest Neighbor, capaz de identificar patrones, desplazándose sin obstáculos (Cárdenas, 2018). De tal forma, que marcó su origen a la primera implementación de aprendizaje automático, y hoy en día, es evidente su crecimiento en distintas industrias.

Para ser más específicos, el machine learning está presente en la cotidianidad del ser humano, desde las recomendaciones en las páginas webs por medio de robots, hasta algoritmos en las redes sociales (Ramírez, 2018). En 2008, Microsoft decide apostar por el aprendizaje automático, y lanza la versión beta de Azure Machine learning, la cual permite al usuario subir a la nube aplicaciones y su sistema lo ejecuta automáticamente, además que otorga mayor confiabilidad que otros sistemas convencionales (Faggela, 2018).

Hoy en día, en el ámbito financiero es evidente su implementación, como, por ejemplo, en aplicaciones como TurboTax, el software recopila información del área usuaria, y gestiona los trámites y transacciones sin necesidad que el usuario intervenga en todo el proceso, más solo que introduzca información descrita como patrón de conducta (Ramírez, 2018).

Tipos de Aprendizaje

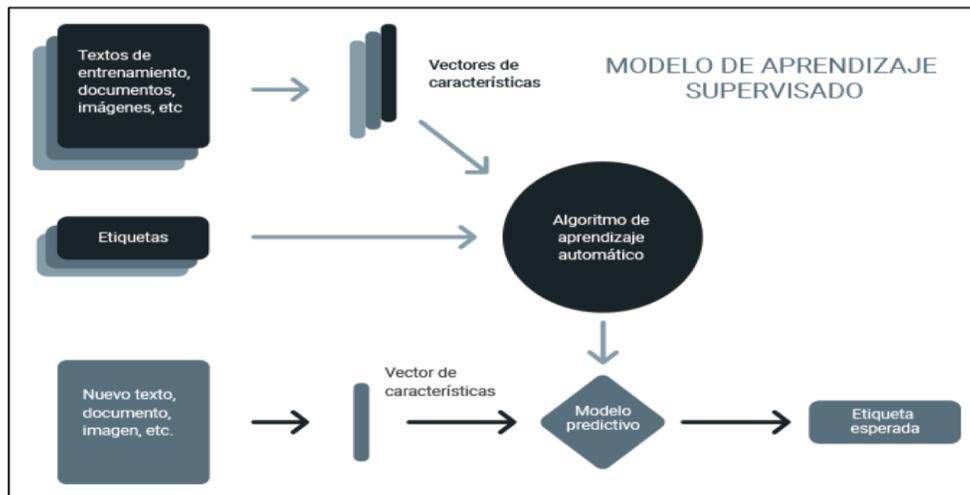
Aprendizaje Supervisado. Consiste en aquel entrenamiento brindado al dispositivo, a través de la proporción de características o entradas, conocidas como inputs o features, las cuales darán un resultado conocido como outputs o etiquetas (Álvarez et.al, 2020). De tal forma, que en un futuro dichos algoritmos puedan otorgar predicciones precisas en base a las entradas registradas.

Es decir, este aprendizaje se encarga de enseñar a dicho algoritmo la forma de cómo debe desempeñar su labor (Rojas, 2018), utilizando un conjunto de datos clasificados y etiquetados bajo ciertos patrones previamente definidos, para ser aplicados posteriormente en un análisis, y, por consecuencia, generar una salida, la cual ya es conocida (Mueller et al., 2016).

De acuerdo con Chaviano (2015), el aprendizaje supervisado representa a dicho algoritmo de aprendizaje, el cual se basa en ejemplos donde el nuevo conocimiento es impulsado en función a una serie de ejemplos o contraejemplos.

Figura 1

Modelo de Aprendizaje Supervisado



Nota. Tomado de Ruiz (2019).

Como se puede observar la Figura 1 representa el proceso que conlleva el aprendizaje supervisado. Este tipo de aprendizaje busca llevar a cabo un mapeo entre un conjunto de entrada X y un conjunto de salida Y. En situaciones, donde el conjunto de salida es discreto, esto hace hincapié a algoritmos de clasificación, mientras que si el conjunto de salidas es continuo hace hincapié a algoritmos de regresión.

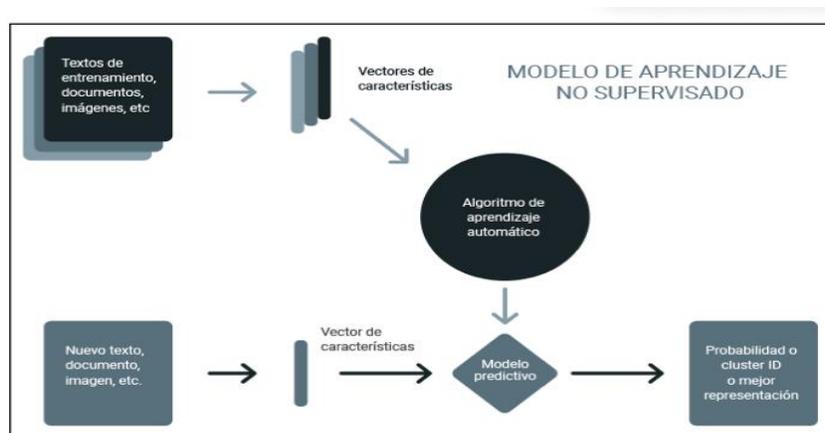
Algoritmo de Clasificación. El algoritmo se encarga de encontrar patrones en los datos a estudiar, y los clasifica en grupos. Posteriormente, estudia cada nuevo dato y lo ubica en el grupo correspondiente a esa clasificación. El conjunto de datos se caracteriza por ser finito de posibles resultados de la predicción, mientras que las variables a predecir son de carácter discreto o categórico, las mismas que pueden ser binarias (0 o 1) o de multi-clases (múltiples categorías) (Sandoval, 2018). Dentro de los algoritmos de clasificación, podemos encontrar árbol de decisión, clasificación de Naive Bayes, redes neuronales entre otros (Maday et al., 2022).

Algoritmo de Regresión. Este algoritmo hace referencia al subcampo del machine learning supervisado, el cual tiene como finalidad relacionar un cierto número de características con una variable continua. Dentro de los algoritmos de regresión, encontramos a la regresión lineal, regresión logística, entre otros (Sandoval, 2018).

Aprendizaje No Supervisado. Consiste en aquel modelo predictivo bajo el mismo entrenamiento que el aprendizaje supervisado, sin embargo, su diferencia recae en que los datos de entrada son no clasificados o etiquetados, puesto que no es necesario realizar dicho entrenamiento. Debido a que, puede descubrir patrones de ejemplos semejantes entre grupo de datos (Bishop, 2007; Rojas, 2018). Dentro del aprendizaje no supervisado podemos encontrar algoritmos como K- means (K-media) (Ruiz, 2019).

Figura 2

Modelo de Aprendizaje No Supervisado



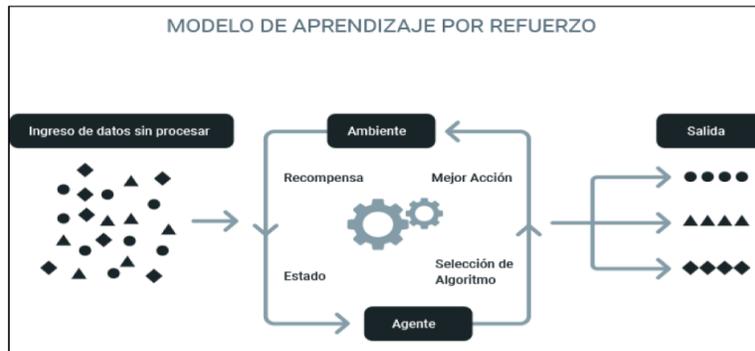
Nota. Tomado de Ruiz (2019).

Como se puede observar en la Figura 2 representa el proceso que conlleva el aprendizaje no supervisado, también se puede evidenciar como las etiquetas de salida no son proporcionadas al algoritmo, sino que son presentadas como se muestra en el gráfico.

Aprendizaje Reforzado. Consiste en aquel modelo de resolución de problemas de toma de decisiones progresivas, donde no existen datos de entrenamiento como entrada, por ende, no se tiene información sobre la posible salida. Sin embargo, se debe entregar una retroalimentación escalar, en función al rendimiento general del sistema

Figura 3

Modelo de Aprendizaje por Refuerzo



Nota. Tomado de Ruiz (2019).

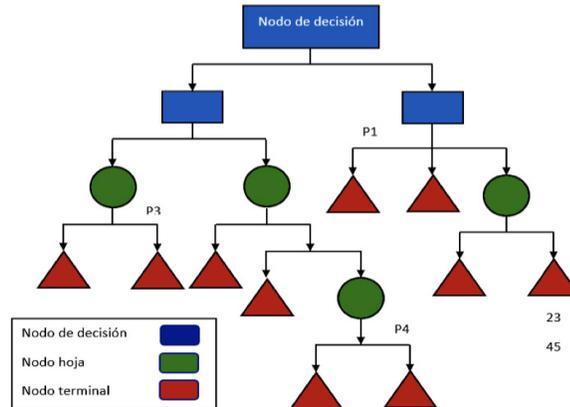
Como se puede observar en la Figura 3 representa el proceso que conlleva el aprendizaje reforzado, aquí se entregan datos sin procesar y el sistema a través del ambiente los clasifica y entrega las respuestas, en función a la mejor recompensa.

Técnicas de Machine Learning

Árbol de Decisión. De acuerdo con Quinlan (1986) & Breiman (1984), un árbol de decisión es una técnica de aprendizaje supervisado de clasificación y regresión, la cual trabaja con variables binarias (0 o 1), representando predicciones por medio de un diagrama de árbol, donde cada nodo interno refleja una prueba sobre un rasgo de entrada y cada rama representa una etiqueta o un valor de salida (Rockach & Maimon, 2008).

Figura 4

Estructura de un Árbol de Decisión



Nota. Tomado de Rivero (2022).

Como se puede observar en la Figura el nodo de decisión contiene al Nodo Raíz el cual es el Nodo inicial o rasgo de entrada, el nodo de decisión es aquel nodo o subnodo que se divide en sub-nodos adicionales, representando una decisión. El nodo de probabilidad u hoja es aquel nodo u hoja que representa diversos resultados inciertos. El nodo terminal es aquel nodo que representa la decisión final, resultado del estudio de nodos anteriores.

Las ramas son los caminos o subsección del árbol que encamina la decisión. Los random forest o también conocido como bosques aleatorios, consiste en la combinación de árboles de decisión para la clasificación, los cuales son entrenados simultáneamente para finalmente utilizar el aprendizaje de todos en conjunto y brindar una predicción mejorada (Amit & Geman, 2001).

De acuerdo con Breiman (2001), un bosque aleatorio es aquel clasificador formado a partir de un consolidado de clasificadores estructurados en forma de árboles, donde cada uno emite una etiqueta unitaria para la clase más concurrente en la entrada x

$$\{h(x, O_t), t=1, \dots\}$$

donde;

- $\{O_t\}$: aquellos vectores aleatorios independientes que genera al t-ésimo árbol.
- x : es el vector de cualidades en R

Según Hastie & Friedman (2001), random forest es una modificación del Bagging, el mismo que trabaja con un consolidado de árboles incorrelacionados y los pondera.

Error de Generalización. Hace referencia a aquella proporción de variables mal clasificadas por el bosque aleatorio, pero al ser una probabilidad su valor está dentro del intervalo de [0,1].

$$PE^* = P_{x,y} (mg(X,Y) < 0)$$

donde;

- X, Y: Aquella distribución de los datos de entrenamiento.
- mg (X, Y): Representa aquella función de margen.

De acuerdo a Liaw & Wiener (2002), la expresión matemática de bosque aleatorio es randomForest(), tanto para la clasificación como para la regresión. La cual se detalla a continuación;

$$\text{randomForest}(\text{formula}, \text{data}, \text{ntree}, \text{mtry}, \text{importance}, \dots)$$

donde;

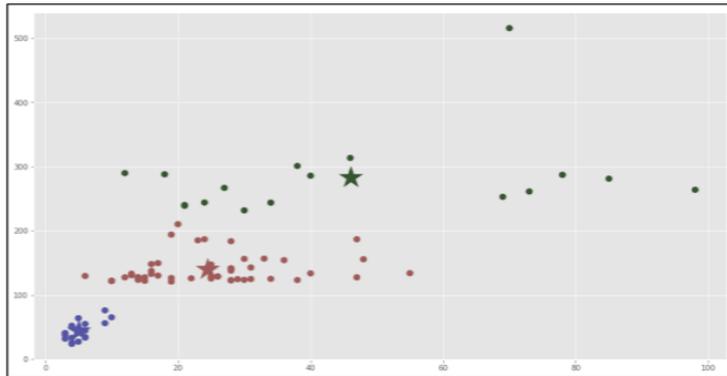
- Fórmula: Aquella fórmula que detalla las variables predictoras y la variable de salida.
- Data: Conjunto de datos que contiene las variables en estudio.
- Ntree: Número de árboles en el bosque aleatorio.
- Mtry: Número de variables predictoras elegidas de forma aleatoria.
- Importance: Indicador lógico, que detalla si se debe calcular la importancia de las variables predictoras.

K-means. De acuerdo a Ruiz (2019), k- media es un algoritmo de segmentación no supervisado, el cual desarrolla una clasificación de las variables en estudio en grupos, en función a sus características de entrada y las distancias entre cada uno de los datos, utilizando algoritmos de cluster o agrupamiento.

La recopilación consiste en llevar a cabo una minimización de la suma de distancias entre cada una de las variables y el clúster de su grupo.

Figura 5

Descripción del Modelo de Agrupación por K-means



Nota. Tomado de Ruiz (2019).

Como se observa en la Figura 5 esto representa el agrupamiento por k-means de un conjunto de datos, también puede observar cómo hay 3 grupos, que se diferencian entre ellos, y la estrella representa el centro de todas las observaciones.

Regresión Logística. Según Hosmer & Lemeshow (2000), el modelo de regresión logística se define como un modelo de clasificación donde la variable réplica (Y) es de tipo dicotómica (1 o 0), con probabilidad $\chi^2(i)$ para $Y(i)=1$ y probabilidad $1 - \chi^2(i)$ para $Y_i=0$. La regresión logística busca estimar la probabilidad de éxito de un suceso, a través de la probabilidad de una variable binaria, en función de los valores de las variables independientes predictoras p.

Dicha probabilidad de éxito de que el suceso $Y=1$ ocurra está dada bajo la siguiente expresión matemática;

$$\pi = \pi(x) = P(Y = 1|x) = \frac{e^{x'\beta}}{1 + e^{x'\beta}}$$

donde;

- $Y=1$ es el suceso de éxito
- $x' = (x_1, x_2, x_p)$; un consolidado de p variables independientes (predictoras). (Lizares, 2017).

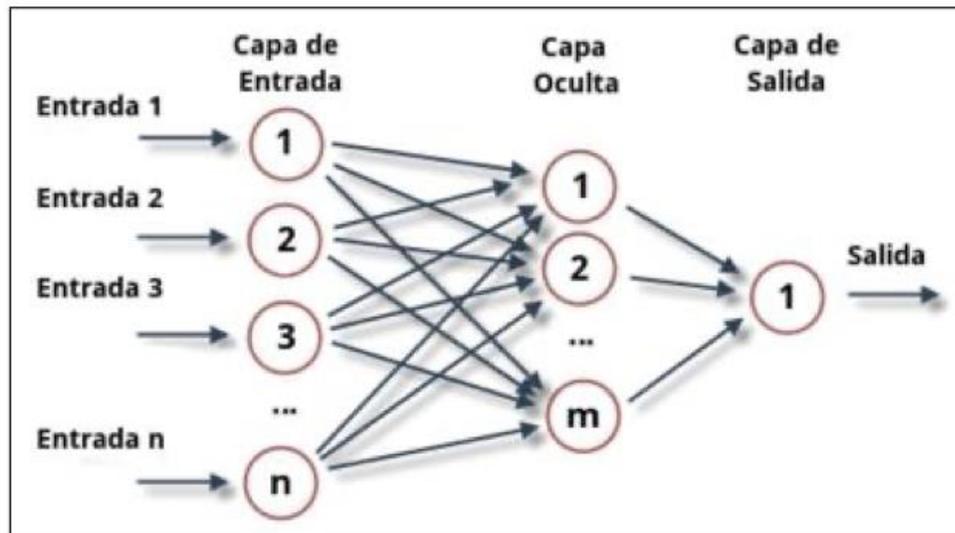
Redes Bayesianas (RB). También conocidas como redes probabilísticas, tiene como finalidad representar un modelo base en función de una ilustración gráfica de las variables y sus dependencias e independencias, al combinarlo con la potencia del teorema de Bayes (Menacho, 2017).

El formato de una Red Bayesiana Naive Bayes, está sintetizado por un nodo raíz (variable dependiente Y_g) y un consolidado de variables independientes $X = (X_1, X_2, X_3, \dots, X_p)$. Donde el teorema de Bayes define que Y_g , es el responsable de todas las variables X , y al mismo tiempo, este conjunto de variables es independientes entre ellas dada la variante clase (Ruiz, 2019).

Redes Neuronales Artificiales (RNA). De acuerdo con McCulloch (1943), una red neuronal artificial es un diseño informático, el cual fue creado en función a las neuronas humanas, con el objetivo de desempeñar tareas que la programación sistematizada no lo podía cumplir sin novedad. Siendo así, que se puede definir a una RNA como aquel modelo matemático e informático, formado a partir de un conjunto de neuronas interrelacionadas entre sí por medio de arcos dirigidos que definen la relación entre los nodos (Higiano, 2017).

Figura 6

Estructura de una Red Neuronal Artificial (RNA)



Nota. Tomado de Ruiz (2019).

Como se puede observar en la Figura 6 esto representa los componentes que forman una red neuronal, también se puede observar que la estructura de una red neuronal está dada por entradas, capaz de salida y capas ocultas. Sus conexiones son de tipo sinápticas, donde llega y sale neuronas, cada una con un peso sináptico correspondiente.

Matriz de Confusión. Según Higiano (2017), una matriz de confusión es una técnica utilizada para la evaluación de clasificadores, la cual consiste en una tabla de contingencia que refleja la repartición de la distribución observada (real) y la predicha (clasificador).

Deep Learning

De acuerdo con Hinton (2006), fundador del término Deep Learning o Aprendizaje Profundo, lo define como aquel conjunto particular de algoritmos del Machine Learning, el cual implementa configuraciones profundas de redes neuronales para poder detectar patrones en los datos (Giletta et.al, 2020).

Sus inicios se dieron a partir del uso de redes neuronales, y la necesidad de explicar sus nuevas arquitecturas, con la capacidad de experimentar de una mejor manera los modelos más planos. En la actualidad, es evidente los avances que han surgido gracias a esta área de la IA, como, por ejemplo, los modelos predictivos o reconocimientos faciales con imágenes, que podemos encontrar en instituciones corporativas (Giletta et.al, 2020).

Marco Referencial

El marco referencial de esta investigación abarca la revisión crítica de teorías, modelos y estudios relacionados con la rotación laboral a través de esto se pretende comprender las dinámicas y factores que influyen en la rotación temprana de empleados. La revisión explorará investigaciones previas que han aplicado técnicas de machine learning en problemas similares para entender cómo estas metodologías pueden aportar al análisis de la rotación laboral. El marco referencial proporcionará una base teórica sólida para contextualizar la investigación, identificar lagunas en el conocimiento existente y fundamentar la elección de enfoques metodológicos, contribuyendo así a la comprensión integral de los factores que afectan la rotación temprana de empleados en el Grupo DIFARE de la región Sierra.

Según el estudio realizado por Parrales & Rosales (2022), mencionó que los hallazgos más sobresalientes del estudio incluyen dos aspectos fundamentales. En primer lugar, se llega a la conclusión de que la implementación de herramientas de inteligencia artificial, especialmente algoritmos de Machine Learning, tiene un efecto positivo en el proceso de selección de personal. Según lo mencionado por los autores, esto implica que la utilización de la inteligencia artificial puede tener un impacto significativo en la mejora de la eficiencia y efectividad en la selección de candidatos. Lo mencionado por los autores también resalta que, en segundo lugar, se determina que el área de Recursos Humanos está preparada para modernizarse mediante la adopción de herramientas basadas en algoritmos de Machine Learning, independientemente del tamaño o la naturaleza de la entidad. Esto resalta la factibilidad de incorporar estas tecnologías en microempresas y otras organizaciones, lo que podría contribuir a una gestión de recursos humanos más eficaz y eficiente.

Además de estos hallazgos, el estudio también presenta una propuesta detallada para la implementación de técnicas de Machine Learning en el proceso de selección de personal en microempresas, junto con las tecnologías requeridas para llevar a cabo esta implementación. Lo mencionado por los autores también señala ejemplos de empresas que ya están haciendo uso de herramientas de inteligencia artificial en el reclutamiento, lo que destaca la importancia y la tendencia en este campo. En resumen, estos resultados subrayan el potencial impacto positivo de la inteligencia artificial, particularmente el Machine Learning, en la gestión de recursos humanos

y el reclutamiento, ofreciendo una valiosa perspectiva para la mejora de los procesos de selección de personal en microempresas y otras organizaciones (Parrales & Rosales, 2022).

Asimismo, el estudio realizado por Barbosa (2021), abordó diversas técnicas de Machine Learning, incluyendo Aprendizaje Supervisado, Semi Supervisado, No Supervisado y por Refuerzo, junto con el uso de Deep Learning y Cognitive Computing. Lo mencionado por el autor destaca el algoritmo K-NN (k-Nearest Neighbors), Algoritmos Genéticos y el Modelo Random Forest como enfoques de aprendizaje conjunto. Además, se discuten medidas de desempeño como el AUC y la curva ROC para evaluar modelos de Machine Learning en la clasificación de datos. Estas técnicas se aplican en diversos campos para resolver problemas y optimizar soluciones.

En la creación de un modelo predictivo para la rotación de empleados, el estudio enfatiza la importancia de una selección cuidadosa de variables que puedan influir en las decisiones de los trabajadores de abandonar la organización. Lo mencionado por el autor también subraya la necesidad de considerar el contexto organizacional y la combinación adecuada de características en un marco temporal específico para evitar desafíos como el sobreajuste o el desajuste durante el entrenamiento del modelo. En este proceso, se recomienda explorar diversas herramientas y enfoques de modelado y, basándose en la lógica empresarial o en las métricas de desempeño de los modelos, elegir la solución que mejor se adapte a las necesidades de la organización. Estos hallazgos resaltan la importancia de una estrategia de modelado sólida y adaptada a las circunstancias específicas de la rotación de empleados en una organización (Barbosa, 2021).

De la misma manera Zaballa et al. (2021) mencionó en su investigación que el procedimiento propuesto destaca por su enfoque detallado y flexible en el análisis de la rotación laboral en las organizaciones. La segunda fase se centra en revisar documentos y realizar entrevistas con especialistas del área de Capital Humano para obtener una comprensión profunda de la organización bajo estudio. La recopilación de datos se realiza mediante la revisión de documentos, registros y entrevistas con expertos y empleados de la organización. El procedimiento se caracteriza por su enfoque en la caracterización del sistema productivo, análisis histórico de bajas, identificación de causas y formulación de propuestas de solución, así como la medición y control de los costos asociados a la rotación laboral. Destaca la importancia de predecir la rotación

para tomar medidas proactivas, proporcionando a las empresas una guía integral para la toma de decisiones informadas y la gestión efectiva del personal.

En conjunto, este enfoque proporciona a las empresas una guía integral para tomar decisiones informadas y abordar eficazmente la rotación de personal. Resalta la necesidad de medir y controlar los costos asociados a la fluctuación laboral, así como la importancia de prevenirla para implementar medidas proactivas. El procedimiento no solo se centra en la identificación de causas, sino que también se enfoca en la formulación de propuestas de solución, ofreciendo una perspectiva completa para abordar los desafíos relacionados con la rotación de empleados. Este enfoque detallado y flexible se presenta como una herramienta valiosa para las organizaciones que buscan comprender y gestionar eficazmente la rotación del personal, proporcionando una base sólida para la toma de decisiones estratégicas y la implementación de políticas laborales efectivas (Zaballa et al., 2021).

Así pues, Zaballa (2020) se centró en analizar la rotación del personal en la Empresa de Construcción y Montaje de Matanzas. Para ello, se recopilaron datos de los Registros de Personal de la organización y se llevaron a cabo entrevistas con la Dirección de Capital Humano. En el proceso de análisis, se aplicaron enfoques teóricos y empíricos, incluyendo análisis documental, entrevistas y técnicas como la tormenta de ideas y el diagrama de causa y efecto. Los resultados revelaron que existen Unidades Empresariales de Base (UEB) con altas tasas de bajas de empleados, especialmente en oficios como albañil, aprendiz de oficios básicos y ayudante.

Además, se identificó que las fuentes de empleo de la empresa no satisfacen completamente la demanda de recursos humanos, y se han adoptado medidas como la contratación de personas de otras provincias y la colaboración con reclusos. Aunque el índice de rotación se encuentra por debajo del umbral máximo permitido, se destaca la importancia de un seguimiento constante y se recomienda implementar soluciones identificadas, analizar las causas subyacentes de las deficiencias en los indicadores relacionados con la rotación del personal, y extender el estudio a otras UEB de la empresa para evaluar el comportamiento de la rotación en diferentes áreas.

Estos hallazgos proporcionan una valiosa visión sobre la situación de la rotación de personal en la Empresa de Construcción y Montaje de Matanzas, destacando la necesidad de abordar las tasas de bajas en ciertos oficios y mejorar la satisfacción de las fuentes de empleo. Resaltan la importancia

de la gestión de recursos humanos en el sector de la construcción y sugieren la implementación de medidas específicas para mantener un equilibrio adecuado en el personal y prevenir la fluctuación laboral en el futuro (Zaballa, 2020).

En consecuencia, Henao (2021) en su estudio, se enfocó en la construcción de un modelo de pronóstico para determinar la propensión a la renuncia voluntaria de los empleados en la Caja de Compensación Comfenalco Antioquia S.A. Los datos utilizados para este trabajo se obtuvieron de las bases de datos de la empresa y abarcan el período de enero de 2010 a abril de 2021, sin incurrir en costos adicionales para la empresa o la universidad, y estuvieron sujetos a un acuerdo de confidencialidad. Se evaluaron cinco modelos diferentes, con el modelo XGBoost destacándose por su rendimiento sobresaliente en la predicción de la renuncia voluntaria, alcanzando una precisión del 87.3% cuando se emplearon técnicas como StratifiedKFold y GridSearchCV para optimizar sus parámetros.

Estos hallazgos subrayan el cumplimiento exitoso de los objetivos del estudio, que incluyen la recopilación, limpieza y depuración de datos, el desarrollo de modelos de clasificación y la creación de un prototipo de herramienta de datos. Además, resaltan que el modelo XGBoost se destaca como una herramienta efectiva para pronosticar la renuncia voluntaria en la organización estudiada, lo que proporciona información valiosa para la toma de decisiones y la gestión de la rotación de personal. La viabilidad de poner este modelo en producción se confirma, y se sugiere la posibilidad de recalibrar con la incorporación de más datos en el futuro, lo que podría contribuir significativamente a la estrategia de recursos humanos de la empresa (Henao, 2021).

Del mismo modo Sierra (2022), se enfocó en una población de 1497 trabajadores de una empresa de software en Medellín. De estos, 900 estaban activos en la empresa y el resto habían abandonado sus puestos de trabajo. El análisis incluyó 19 características de los empleados, como salario, edad, tiempo de permanencia en la empresa y nivel educativo. El modelo más efectivo para predecir la deserción laboral se identificó como un árbol de decisión con 14 capas, destacando la importancia de variables específicas como el salario total y la duración del empleo en la probabilidad de que los empleados abandonen sus puestos de trabajo. Además, se observó que los empleados más jóvenes mostraban una probabilidad más alta de deserción laboral. El estudio enfatiza que los esfuerzos por reducir la deserción no han tenido éxito en la empresa analizada, lo

que genera costos financieros significativos debido a las liquidaciones de empleados que optan por dejar la organización. Se sugiere ampliar el estudio a diversas empresas y regiones para obtener resultados más generalizables, explorar otros algoritmos de aprendizaje y considerar la inclusión de variables adicionales, con el objetivo de mejorar la comprensión y gestión de la deserción laboral en un contexto empresarial más amplio.

Estos hallazgos resaltan la relevancia de abordar la deserción laboral en la industria de software colombiana y proporcionan información valiosa sobre las variables clave que influyen en la decisión de los empleados de abandonar su empleo. El modelo de árbol de decisión identificado puede ser una herramienta efectiva para predecir y gestionar la deserción laboral en esta organización y posiblemente en otros contextos similares. Las recomendaciones de ampliar el estudio a nivel regional y sectorial sugieren un enfoque más generalizable y brindan pautas para futuras investigaciones en el campo de la gestión de recursos humanos (Sierra, 2022).

Por consiguiente, Palacio (2023), se basó en datos recopilados de la red social pública LinkedIn, abarcando una población diversa de 3369 empleados de diferentes países de América Latina, como Argentina, Chile, Colombia, México, Panamá y Perú. El análisis de datos incluyó variables como la cantidad de roles iniciados por año y la duración de los roles, con una limpieza de datos para garantizar la calidad. Además, se incorporaron variables externas, como la tasa de desempleo y el crecimiento del producto interno bruto, para comprender el contexto macroeconómico de los países de origen de los empleados. Se observó que modelos de predicción como regresión logística, random forest, gradient tree boosting y LSTM superaron a un modelo de referencia simple en la predicción de la rotación de empleados, destacando el gradient tree boosting con un rendimiento superior del 17%. A pesar de limitaciones como la falta de información salarial y de satisfacción, se identificaron variables clave, como el tiempo en la empresa y la tasa de desempleo, y se exploró el uso de LSTM para captar historias laborales.

Los hallazgos indican que estos modelos de predicción son aplicables a diversas empresas en América Latina y que la asignación de costos por clasificación errónea mejora su desempeño. Aunque el enfoque de LSTM no superó a otros modelos en este estudio, se sugiere que futuras investigaciones podrían obtener resultados mejores con más datos y variables que cambien con el tiempo. En resumen, este estudio ofrece resultados alentadores que abren la puerta a futuras

investigaciones centradas en empleados con trabajos similares, proporcionando valiosos conocimientos sobre la predicción de la rotación de empleados en el contexto empresarial de América Latina (Palacio, 2023).

Consecutivamente Cahuasqui & Zapata (2023), se centraron en una revisión sistemática de la literatura que involucró la selección y análisis de 2,531 documentos iniciales, con exclusiones basadas en criterios como la fecha de publicación y la naturaleza de los artículos. La investigación se enfocó en estrategias para prevenir la rotación de personal en microempresas y en los beneficios y desafíos de aplicar la inteligencia artificial en la gestión de recursos humanos. Estos pasos fundamentales permitieron recopilar datos y analizar tendencias en la implementación de la IA en la gestión de talento humano en microempresas.

Los hallazgos principales de la investigación indican que la implementación de mantenimiento predictivo respaldado por inteligencia artificial es una herramienta valiosa para reducir la rotación de personal en las microempresas. Este enfoque permite anticipar patrones de rotación y tomar medidas preventivas, lo que se traduce en la reducción de costos, mejora de la productividad y la creación de un ambiente de trabajo más favorable para los colaboradores. La gestión de talento humano puede tomar decisiones informadas y abordar desafíos mediante la utilización de la IA en este contexto. Aunque la investigación reconoce limitaciones como la necesidad de datos precisos y posibles sesgos, resalta que la aplicación del mantenimiento predictivo con IA en microempresas tiene un impacto significativo en la prevención de la rotación de personal y contribuye positivamente a la eficiencia operativa y la sostenibilidad empresarial (Cahuasqui & Zapata, 2023).

Finalmente, Navarro (2022) en su investigación enfrentó desafíos para obtener datos reales de empleados debido a restricciones de protección de datos en algunas empresas. Como solución, se optó por generar datos simulados que se asemejaran lo más posible a la realidad, utilizando una base de datos propuesta por Huebner y Patalano como referencia.

La investigación propone una metodología innovadora basada en Machine Learning para agrupar a los empleados según su desempeño y potencial. Los resultados indican que es factible clasificar a los empleados en grupos homogéneos en términos de variables como retribución, formación, compromiso, satisfacción, rotación, entre otras. Estos hallazgos ofrecen la posibilidad

de diseñar políticas específicas de recursos humanos para cada grupo de empleados, lo que puede ser valioso para los departamentos de RR.HH. a la hora de aplicar estrategias adaptadas a las necesidades de cada grupo o clúster (Navarro, 2022).

Tabla 1*Matriz referencial*

Título de investigación	Autores	Año	País	Población	Hallazgos
Predicción de rotación de empleados usando modelos de aprendizaje automático	Luis Javier Palacio Mesa	2023	Colombia	Los datos se adquirieron de una red social pública conocida como LinkedIn, y que el conjunto de datos consta de 3369 empleados que residían en diferentes países de América Latina, a saber, Argentina, Chile, Colombia, México, Panamá y Perú	Este estudio resalta la relevancia de la rotación de empleados en las empresas y subraya los costos asociados a la pérdida de empleados talentosos. Además, enfatiza la utilidad de los modelos de aprendizaje automático, como las redes neuronales LSTM (Long Short-Term Memory - Memoria a Largo Plazo de Corto Plazo), para predecir la rotación de empleados y cómo estos modelos pueden ayudar a las empresas a tomar medidas preventivas para retener a su personal. Se destaca la importancia de considerar las implicaciones éticas de la recopilación y el uso de datos de empleados para predecir la rotación. Asimismo, se mencionan los desafíos potenciales en la implementación de soluciones basadas en modelos de aprendizaje automático, como la necesidad de datos de alta calidad y la importancia de involucrar a los empleados en el proceso. El estudio también subraya la utilidad de las técnicas de modelado de secuencias de tiempo para predecir la rotación de empleados, especialmente dada la dimensión temporal de los datos de empleados.
Utilización de Machine Learning para el proceso de selección de personal en una microempresa	María Carolina Rosales Roldán & Estefanía María Parrales Toledo	2022	Ecuador	52 personas con rango de edades desde los 18 a 46 años, en su mayoría de las ciudades Guayaquil, Quito y Cuenca.	Entre los hallazgos que se destacan es que la aplicación de herramientas de inteligencia artificial, particularmente algoritmos de Machine Learning, tiene un impacto positivo en el proceso de selección de personal. En segundo lugar, se determina que el área de Recursos Humanos está preparada para modernizarse en la selección de personal a través de la incorporación de herramientas basadas en algoritmos de Machine Learning, independientemente del tamaño o la naturaleza de la entidad. Además de estos hallazgos, el trabajo presenta una propuesta detallada para la implementación de técnicas de Machine Learning en el proceso de selección de personal en una microempresa, junto con las tecnologías necesarias para llevar a cabo esta implementación. También se mencionan ejemplos de empresas que ya emplean herramientas de inteligencia artificial en el reclutamiento, como L'Oreal, Adecco, Hays y Deloitte, y se describe una herramienta específica llamada MYA Systems que utiliza chatbots con IA conversacional para agilizar el proceso de selección. Estos resultados resaltan el potencial impacto positivo de la inteligencia artificial en la gestión de recursos humanos y el reclutamiento.
Modelos Predictivos para la Rotación del Talento Humano	José Luis Barbosa Fontecha	2020	Colombia	Personal de la empresa Colsubsidio	En el trabajo de investigación, se destaca la importancia de una cuidadosa selección de variables que puedan influir en las decisiones de los trabajadores de dejar la organización. Además, se enfatiza la necesidad de considerar el contexto

					organizacional y la combinación apropiada de características en un marco temporal específico para evitar desafíos como el sobreajuste o el subajuste durante el entrenamiento del modelo. En este proceso, se recomienda explorar diversas herramientas y enfoques de modelado, y luego, basándose en la lógica empresarial o en las métricas de desempeño de los modelos, elegir la solución que mejor se adapte a las necesidades de la organización.
Procedimiento para el análisis de la rotación del personal	Pablo Enrique Zaballa Gomariz, Yusef El Assafiri Ojeda, Yuly Esther Medina Nogueira, Dianelys Nogueira Rivera & Alberto Medina León	2021	Cuba	No aplica	Entre sus hallazgos más relevantes esta el procedimiento propuesto para analizar la fluctuación laboral en las organizaciones se destaca por su enfoque detallado y flexible. Se enfoca en la caracterización del sistema productivo y las fuentes de empleo, así como en el análisis histórico de las bajas en la organización, proporcionando una base sólida para comprender el contexto. Luego, se centra en el análisis de la fluctuación real y potencial, identificando causas y formulando propuestas de solución. También resalta la importancia de medir y controlar los costos asociados a la fluctuación laboral y la necesidad de predecir para tomar medidas proactivas. En conjunto, este enfoque proporciona a las empresas una guía integral para tomar decisiones informadas y abordar eficazmente la rotación de personal.
Desarrollo de un procedimiento para el análisis de la rotación del personal en la Empresa de Construcción y Montaje de Matanzas	Pablo Enrique Zaballa Gomariz	2020	Cuba	Empleados de la Empresa de Construcción y Montaje de Matanzas	Los hallazgos encontrados revelaron resultados significativos, entre los cuales se destacan la identificación de Unidades Empresariales de Base (UEB) con altas tasas de bajas de empleados, resaltando oficios como albañil, aprendiz de oficios básicos y ayudante como los más afectados. También se constató que las fuentes de empleo de la empresa no satisfacen completamente la demanda de recursos humanos, y los largos periodos de preparación han llevado a recurrir a medidas como la contratación de personas de otras provincias y la colaboración con reclusos. A pesar de que el índice de fluctuación se encuentra por debajo del umbral máximo admitido, se subraya la importancia de un seguimiento constante. Se recomienda implementar las soluciones identificadas, analizar las causas subyacentes de las deficiencias en los indicadores relacionados con la fluctuación laboral y extender el estudio a otras UEB de la empresa para evaluar el comportamiento de la fluctuación en diferentes áreas.
Modelo de medición de la rotación de personal como variable de decisión estratégica	Carolina Henao Ríos	2021	Colombia	Colaboradores de la Caja de Compensación Comfenalco Antioquia S.A.	Entre los hallazgos más importantes están la construcción de un modelo de pronóstico para determinar la propensión a la renuncia voluntaria de los empleados de una Caja de Compensación. Se destaca que el modelo XGBoost demostró una precisión del 87.3% en la predicción de la renuncia voluntaria, utilizando técnicas como StratifiedKfold y GridSearchCV para optimizar sus parámetros. Además, se enfatiza el cumplimiento de los objetivos específicos del

Análisis y predicción de la deserción de empleados: un caso de estudio en la industria de software colombiana	Eliana Marcela Sierra Buriticá	2022	Colombia	La población en este estudio fue de 1497 trabajadores de una compañía de software.	<p>estudio, que incluyen la recolección, limpieza y depuración de datos, el desarrollo de modelos de clasificación y la creación de un prototipo de herramienta de datos. Estos hallazgos proporcionan información valiosa para la toma de decisiones y la gestión de la rotación de personal en la organización estudiada.</p> <p>Entre sus hallazgos más importantes se establece que el modelo más efectivo para predecir la deserción laboral se identifica como un árbol de decisión con 14 capas, demostrando un rendimiento superior en comparación con otras alternativas propuestas. Se enfatiza la relevancia de variables específicas, como el salario total y la duración del empleo en la probabilidad de que los empleados abandonen sus puestos de trabajo. Se destaca que los empleados más jóvenes muestran una probabilidad más alta de deserción laboral. En el contexto de la empresa analizada, se señala que los esfuerzos por reducir la deserción no han tenido éxito, lo que se traduce en costos financieros considerables debido a las liquidaciones de empleados que optan por dejar la empresa. En cuanto a las recomendaciones, se sugiere ampliar el tamaño de la base de datos y realizar un muestreo aleatorio estratificado en diversas empresas y regiones para lograr resultados más generalizables. Se recomienda explorar otros algoritmos de aprendizaje y considerar la inclusión de variables adicionales. Además, se propone extender este tipo de estudio a empresas de diferentes sectores y regiones para obtener conclusiones más amplias y aplicables en diversos contextos organizacionales.</p>
Mantenimiento Predictivo de rotación de personal en Microempresas con IA	Erika Priscila Cahuasqui Molina & Shirley Viviana Zapata Romero	2023	Ecuador	2,531 documentos relevantes. Estos fueron esenciales para la recopilación de datos y el análisis de tendencias en las publicaciones sobre la implementación de la IA en la gestión de talento humano en microempresas.	<p>Entre sus hallazgos la investigación revela que la implementación de mantenimiento predictivo respaldado por inteligencia artificial (IA) es una valiosa herramienta para reducir la rotación de personal en microempresas. Este enfoque permite anticipar patrones de rotación y tomar medidas preventivas, lo que conlleva a la reducción de costos, mejora de la productividad y creación de un entorno laboral más favorable para los colaboradores. La gestión de talento humano puede tomar decisiones informadas y abordar desafíos mediante el uso de la IA. La incorporación de la IA facilita la prevención de la rotación de empleados al identificar factores determinantes y diseñar un marco metodológico para prever y contrarrestar rotaciones innecesarias, representando una revolución en el ámbito empresarial y, en particular, en el área de talento humano. Sin embargo, la investigación también destaca algunas limitaciones, como la necesidad de contar con datos precisos y la posibilidad de sesgos en la incorporación de la IA. Además, se reconoce que la inversión monetaria puede ser un obstáculo en microempresas con recursos limitados. A pesar de estas limitaciones, se demuestra que la aplicación del mantenimiento predictivo con IA en microempresas aporta de manera significativa a la prevención de la</p>

Análítica de RR. Machine Learning	HH:	Guillermo Navarro Meseguer	2022	España	<p>Datos generados mediante una simulación, partiendo de una investigación inicial para crear unos datos que se acercaran lo máximo posible a la realidad.</p>	<p>rotación de personal, generando un impacto positivo en la eficiencia operacional y la sostenibilidad empresarial.</p> <p>El trabajo de investigación propone una metodología innovadora para agrupar a los empleados en función de su desempeño y potencial, utilizando técnicas de Machine Learning. Los resultados obtenidos muestran que es posible agrupar a los empleados en distintos grupos homogéneos entre sí y heterogéneos entre grupos, utilizando variables independientes como la retribución, formación, compromiso, satisfacción, rotación, etc. Además, se presentan propuestas de actuación para cada grupo de empleados, lo que puede ayudar a los departamentos de RR.HH. a aplicar políticas concretas en función del grupo o clúster asignado a cada empleado.</p>
-----------------------------------	-----	----------------------------	------	--------	----------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Marco conceptual

En el marco conceptual de esta investigación se pretende proporcionar una base teórica sólida para el análisis de los factores que influyen en la rotación temprana de empleados en el Grupo Difare de la Región Sierra, empleando herramientas de Machine Learning. Se abordarán diversas dimensiones, tales como el contexto laboral e inteligencia artificial, con el objetivo de comprender la complejidad de este fenómeno a través de sus diferentes conceptos.

Rotación de Personal

Es la partida de un empleado de la organización, ya sea de manera voluntaria o involuntaria, marca el cierre de su relación laboral y se puede describir como el porcentaje de individuos que dejan la organización en un período específico (Macario, 2018).

Del mismo modo Mobley (1982) proporcionó una definición de la rotación de personal en la que describe este fenómeno como la separación de un empleado afiliado a una organización que recibe una remuneración de esta.

Por otro lado, Chiavenato (1990), resaltó que uno de los aspectos de suma importancia para las organizaciones, que debe ser constantemente considerado para lograr una efectiva gestión empresarial, es la rotación de personal. Este factor desempeña un papel fundamental en la dinámica de la organización, ya que hace referencia a los cambios en el número de empleados, tanto internos como externos, que forman parte de la empresa.

Facilita la evaluación del período durante el cual los empleados prestan sus servicios en una posición específica, departamento o en la organización en general, junto con la frecuencia con la que deben ser sustituidos. Cuanto más frecuentemente se vacía un puesto dentro de la empresa, mayor será la tasa de rotación. Además, esta métrica suele estar relacionada con la ausencia de liderazgo y orientación eficaces en una empresa en lo que respecta a su personal (Martín et al., 2013).

Rotación de Personal Interna. Según Chiavenato (2005) definió la rotación de personal interna como el movimiento de los empleados dentro de la misma organización. Además, en su

investigación, sugiere que una empresa que fomenta la rotación interna puede ser considerada como una organización conscientemente productiva.

Rotación de Personal Externa. Chiavenato (2005), describió la rotación de personal externa como el flujo de empleados que entran y salen de la compañía, es decir, las contrataciones y despidos de colaboradores. Además, señala que la repetición de este fenómeno puede tener un impacto negativo en la empresa.

Intención de Rotación de Personal

La intención de rotación se describe como el proceso de reflexión previo a la finalización voluntaria del contrato, en el cual el empleado está considerando dejar la organización (Lodhi et al., 2020).

Desempeño Laboral

El desempeño laboral abarca la evaluación tanto del logro de metas por parte del empleado como de los recursos empleados para alcanzarlas. Puede ser evaluado en términos de eficacia y ofrece una medida de cuán exitosa ha sido una organización en la consecución de sus objetivos (Alarcón, 2022).

Factores Intrínsecos

Los factores intrínsecos hacen referencia a las características personales y de salud que influyen en los empleados, y, en consecuencia, en la organización. A menudo, estos factores pueden estar relacionados con las responsabilidades laborales y las oportunidades de progresar, desarrollarse, ejercer autonomía, alcanzar metas y obtener reconocimiento en sus roles (Charaja & Mamani, 2013).

Factores Extrínsecos

Los factores extrínsecos engloban las influencias del entorno que impactan en los trabajadores, como su grado de satisfacción con aspectos como el salario, la participación en la toma de decisiones, la supervisión, la formación, los beneficios y otros elementos proporcionados por la organización a sus empleados (Carreón, 2016).

Inteligencia Artificial

Según Leyva (2018), la inteligencia artificial se refiere a la creación de agentes racionales que perciben su entorno y toman decisiones con el fin de maximizar las posibilidades de alcanzar metas específicas. En este sentido, la inteligencia artificial se enfoca en emular el pensamiento humano de manera muy cercana, de modo que, al aplicar modelos o algoritmos, pueda generar resultados que beneficien a negocios, empresas y otros ámbitos.

En la actualidad, la inteligencia artificial desempeña un papel crucial en el avance tecnológico de diversos sectores dentro de un país, la inteligencia artificial utiliza una variedad de técnicas para abordar una amplia gama de tareas, lo que implica que desde sus inicios, que incluyen el uso de estadísticas para el análisis de datos, hasta avances significativos como el reconocimiento de voz y el texto predictivo en dispositivos móviles, la inteligencia artificial ha experimentado una notable evolución (Boden, 2018).

Es importante destacar que la inteligencia es un término subjetivo que no se limita a una sola dimensión, sino que abarca un amplio conjunto de capacidades para el procesamiento y manejo de información que se utilizará en diversas situaciones. En este contexto, la IA emplea una amplia variedad de técnicas para abordar diversas tareas, siendo el enfoque principal el aprendizaje automático o *machine learning*, que se encuentra a la vanguardia del estudio y desarrollo en el campo de la IA (Boden, 2018).

Desde sus inicios, se han atribuido varias definiciones a la Inteligencia Artificial (IA). La primera mención conocida fue hecha por el Dr. Marvin Minsky, quien la describió como la ciencia de hacer que las máquinas realicen tareas que requerirían inteligencia si fueran realizadas por seres humanos (Minsky, 1956, citado en Alfonso et al., 2003, p.4).

Otra definición más específica, aportada por Pérez (2018), consideró que la IA se manifiesta cuando una máquina puede imitar las funciones cognitivas propias de la mente humana, tales como la creatividad, la sensibilidad, el aprendizaje, la comprensión, la percepción del entorno y el uso del lenguaje. Sin embargo, no existe una definición única debido a la amplia gama de aplicaciones y usos que abarca la IA, así como a la continua evolución y descubrimiento de nuevas metodologías.

Una definición más completa que engloba las anteriores sería que la IA es la capacidad de las máquinas para utilizar algoritmos, aprender a partir de los datos y aplicar ese aprendizaje en la toma de decisiones, de manera similar a como lo haría un ser humano (Rouhiainen, 2018). En resumen, la IA busca que las máquinas emulen acciones humanas mientras adquieren y desarrollan conocimientos para mejorar sus habilidades, es decir, para desarrollar inteligencia.

Machine Learning

Según Mahesh (2018) el aprendizaje automático (*Machine Learning, ML*) se emplea para instruir a las máquinas en la gestión más efectiva de datos. Esto se vuelve especialmente relevante cuando se lidia con volúmenes significativos de datos, lo que puede hacer que su interpretación resulte complicada. La finalidad del aprendizaje automático consiste en adquirir conocimientos a partir de los datos. En este sentido, muchos expertos en matemáticas y programación utilizan diversos enfoques para abordar este desafío asociado a conjuntos de datos extensos.

Asimismo, Hurwitz & Kirsch (2018) mencionaron que el aprendizaje automático puede definirse como una forma de IA que permite que un sistema adquiera conocimientos a partir de los datos en lugar de depender de una programación explícita. En otras palabras, el *machine learning* permite que las máquinas identifiquen patrones de manera autónoma, sin necesidad de una programación específica, a través de la aplicación de algoritmos en relación con los datos para aprender de manera continua. Este enfoque se puede desglosar en tres tipos principales de algoritmos: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo (Rouhiainen, 2018).

Con el paso de los años, el *machine learning* ha experimentado un crecimiento significativo, dando lugar al desarrollo de nuevas metodologías para aprovechar al máximo su potencial dentro del campo de la IA. Esto incluye el surgimiento de las redes neuronales, cuyo funcionamiento aún no se comprende por completo, pero que han logrado emular, hasta cierto punto, el funcionamiento del cerebro humano. Esto ha permitido el procesamiento eficiente de grandes conjuntos de datos y ha dado lugar al surgimiento del aprendizaje profundo, que representa un hito significativo en el campo de la inteligencia artificial y el *machine learning* (IBM Cloud Education, 2021).

Matriz de Confusión

También conocida como matriz de error o tabla de contingencia, la matriz de confusión desempeña un papel esencial en el ámbito del aprendizaje automático cuando se trata de algoritmos de clasificación. Su propósito principal es evaluar la precisión de las clasificaciones realizadas. La interpretación de esta matriz se efectúa de manera horizontal, donde los datos reales se representan en las columnas y las clasificaciones en las filas. Aquellos elementos que se encuentren fuera de la diagonal principal corresponden a las clasificaciones incorrectas, y a partir de estos, es posible calcular el error en forma de porcentaje. Existen dos tipos de errores: el error de omisión, que ocurre cuando elementos de una categoría se asignan a otra incorrecta, y el error de comisión, que ocurre cuando elementos de una categoría incorrecta se asignan a una determinada clase (Sánchez 2015).

Outliers

Gosh y Vogt (2012) describieron a los valores atípicos (*Outliers*) como datos que se apartan significativamente de la mayoría o del conjunto de observaciones. Por lo general, estos valores atípicos tienen un impacto en el modelo y sus resultados. La presencia de varios valores atípicos puede afectar negativamente la calidad de los resultados, lo que requiere que el científico de datos realice una limpieza de datos. Estos valores o datos atípicos suelen ser, en su mayoría, errores. Su origen a menudo se encuentra en el proceso de recopilación de datos, como posibles errores de medición debido a la falta de precisión en la respuesta del instrumento. También pueden surgir por errores administrativos, como la introducción incorrecta de los valores de la muestra. Finalmente, los errores de muestreo, que se refieren a la posibilidad de que los datos no se hayan registrado o recolectado de manera adecuada en el informe o en la recopilación de datos.

En el análisis de datos, la mayoría de las bases de datos utilizadas para llevar a cabo estudios e inferencias contienen datos que tienen un cierto grado de influencia en los resultados, y estos datos se denominan *Outliers* o valores atípicos. En términos estadísticos, los *Outliers* se definen como observaciones que se distancian significativamente numéricamente del conjunto de datos y, en consecuencia, no representan adecuadamente la totalidad de los datos. Cuando estos valores atípicos alcanzan un cierto porcentaje, es necesario depurarlos. Una forma eficaz de identificarlos

es mediante el uso de diagramas de caja. Además, los valores atípicos también incluyen datos faltantes o *missing* (Tejada, 2021).

En general, los analistas deben ejercer su juicio para manejar estos datos, ya que no se pueden categorizar de manera definitiva como beneficiosos o problemáticos; en su lugar, deben ser considerados en el contexto del análisis y se debe evaluar el tipo de información que pueden proporcionar (Ocaña, 2018).

Business Intelligence

Choi et al. (2021) describieron la inteligencia de negocios como un concepto que implica la adquisición, interpretación, recopilación, análisis y obtención de información relacionada con empresas o emprendimientos en diversos entornos. En una sociedad moderna que depende en gran medida de la tecnología, esta herramienta resulta sumamente beneficiosa. Desde una perspectiva empresarial, la utilidad principal de los datos obtenidos a partir de la información recopilada en sistemas o programas de otras empresas, así como encuestas, radica en comprender la situación o contexto actual de un sector de producción o mercado con el fin de tomar decisiones adecuadas.

Marco Legal

Normas Internacionales

ISO 26000. De acuerdo con la Organización Internacional de Empleadores OIE (2010), el ISO 26000 representa aquella norma espontánea diseñada con la finalidad de mejorar el desempeño de una entidad comercial con sus grupos de interés tanto externos como internos a la organización. Dicha gestión se basa en las siguientes vertientes de la Responsabilidad Social, Gobernanza de la Entidad Comercial, Derechos Humanos, Prácticas Laborales, Prácticas Equitativas de la Gestión, Desarrollo de Proveedores, Participación y Desarrollo de la Comunidad y Medio ambiente.

De las cuales, para la empresa Distribuidora Farmacéutica ecuatoriana Difare S.A, es de suma relevancia las siguientes;

Gobernanza de la Entidad Comercial. Una buena administración corporativa guía el accionar diario y vela por el imprescindible balance tanto en lo interno como externo de la compañía, precautelando la sustentabilidad y bienestar de todos a largo plazo (Vives & Vara, 2011).

Derechos Humanos. Según la Organización Internacional del Trabajo (OIT) y la Declaración Universal de Derechos Humanos de las Naciones Unidas, describe que los derechos humanos están regidos bajo 4 principios;

- Respeto por la libertad de alianza e identificación efectiva del derecho de negociación colectivo.
- Prohibición completa de cualquier forma de trabajo forzado.
- Prohibición de la labor infantil.
- Cualquier tipo de discriminación en relación con el puesto de trabajo.

Prácticas Laborales. Las entidades comerciales no solo deben velar por brindarles a los empleados un salario justo, los beneficios según la ley y condiciones adecuadas para trabajar, ya que es de igual de relevante brindar un excelente clima laboral, que vaya de la mano de una estabilidad dentro de la compañía (Ibarra, 2015).

Normas Nacionales

De acuerdo con la Constitución de la República del Ecuador, se establece como estabilidad mínima y excepciones que el tiempo mínimo de duración de un colaborador de todo contrato por tiempo fijo o indefinido debe ser de un año, y cuando la actividad sea de naturaleza estable o permanente, se modifique dicho contrato a largo plazo (Código del Trabajo, 2012, Artículo 14).

Sin embargo, se detallan ciertas restricciones;

- a) Contratos no habituales en la actividad de la compañía
- b) Contratos eventuales, o por un periodo ocasional.
- c) Servicio doméstico
- d) Por aprendizaje
- e) Por artesanos y sus operarios
- f) Contratos prueba (Código del Trabajo, 2012, Artículo 14).

Según la Constitución de la República del Ecuador, se describe a un contrato a prueba, como todo aquel contrato que se puede llegar a celebrar sólo una vez, con un periodo máximo de duración de 90 días. Una vez vencido este periodo, la entidad comercial debe formalizar el contrato, o darlo por terminado (Código del Trabajo, 2012, Artículo 15).

De igual forma, existen los contratos eventuales, por un periodo ocasional o por temporadas, los cuales son todos aquellos que se llevan a cabo para satisfacer requerimientos momentáneos del empleador, donde se debe puntualizar los requisitos circunstanciales del contrato, como el nombre de la persona a reemplazar y reemplazar, como tiempo de duración (Código del Trabajo, 2012, Artículo 17).

Capítulo III.

Metodología

Para este trabajo de indagación se implementará el algoritmo de clasificación denominado Random Forest o Bosque Aleatorio en conjunto a la matriz de confusión. El Bosque Aleatorio es una técnica de aprendizaje supervisado existente tanto para la clasificación como para la regresión. La misma que consiste en un conjunto de árboles de decisión, donde sus predicciones son correlacionadas y promediadas por la clasificación (Liaw & Wiener, 2002). Para ser más específicas, cada árbol de decisión se entrena en un subconjunto en específico, y en paralelo, genera su respuesta binaria (“Si” o “No”). Y, por consiguiente, se combinan todos estos resultados y se obtiene una respuesta concluyente, la cual permite identificar el éxito o fracaso de la predicción.

El Random Forest es una variante mejorada de la técnica “bagging”, la cual hace referencia a diversos clasificadores débiles, los cuales han sido tratados de manera independiente. (un modelo en cada subconjunto) (Breiman, 2001).

Siendo así, que de manera estadística en Random Forest, el conjunto de árboles débiles se resume en;

$$\{h(x, O_k), k=1 \dots\}$$

donde el $\{O_k\}$, son aquellos segmentos aleatorios distribuidos de igual manera, donde cada clasificador emite una etiqueta unitaria para la clase más concurrente en la entrada x (Breiman, 2001).

Cuando los árboles de decisión suelen ser correlacionados, lo que empeora la predicción concluyente, se implementa la técnica “dropout”. La cual permite reducir la correlación promedio entre dichas predicciones ejecutadas, y por consiguiente, mejorar la tendencia de la predicción general a diferencia del “bagging” estándar.

En relación con esto, otra de las herramientas a utilizar es la matriz de confusión, o matriz de error, la misma que es empleada para evaluar el desempeño de un modelo de clasificación. Donde se resumen las predicciones correctas e incorrectas mediante conteos y se desglosan por

cada clase. Suele estar compuesta por cuatro elementos: verdaderos positivos (VP), falsos positivos (FP), verdaderos negativos (VN) y falsos negativos (FN) (Gil y Seguro, 2022).

- Verdaderos Positivos (VP): Instancias correctamente clasificadas como positivas.
- Falsos Positivos (FP): Instancias incorrectamente clasificadas como positivas.
- Verdaderos Negativos (VN): Instancias correctamente clasificadas como negativas.
- Falsos Negativos (FN): Instancias incorrectamente clasificadas como negativas.

La matriz de confusión es un arreglo de números que expresa el número de unidades de píxeles asignados a una categoría particular con respecto a otra de prueba. En las columnas se encuentran los datos verificados en el campo y sobre los que se tiene certeza del tipo de cubierta que representan, que en este caso sería la clasificación supervisada (Muñoz, 2016).

Figura 7

Matriz de confusión

Valores Actuales	Valores Predichos	
	No cumple	Cumple
	No cumple	(VN) Verdaderos Negativos
Cumple	(FN) Falsos Negativos	(VP) Verdaderos Positivos

Nota. Adaptado de Análisis de Calidad Cartográfica mediante el estudio de la Matriz de Confusión (2016)

Procedimiento en RStudio

Para el desarrollo del modelo de clasificación se implementará la herramienta RStudio, la misma que permite la visualización de datos, manejo de paquetes e integración de gráficos estadísticos.

En primer lugar, se procede a llamar a la base de datos, por lo cual se utiliza la función (*read.csv*) o (*read.csv2*), de acuerdo con el tipo de datos que se vaya a estudiar. En este caso, implementamos (*read.csv2*), puesto que son caracteres.

```
Difare <- read.csv2("../Data/Base Doris y Sara TESIS (1).csv")
```

Una vez se tuvo la base de datos lista en el software, se procedió a empezar a tratar las variables, convirtiéndolas de datos cuantitativos a categóricos.

Conversión de Datos Cuantitativos a Categóricos

```
Difare$DIVISION <- factor(Difare$DIVISION, levels= c(1,2,3),
                        labels = c("Cruz Azul", "Pharmacys", "Dromayor"))
Difare$AREA <- factor(Difare$AREA, levels= c(1,2,3),
                    labels = c("Ventas Farmacias Propias", "Contac
                              Center", "Franquicia"))
Difare$PUESTO <- factor(Difare$PUESTO, levels = c(1,2),
                      labels = c("Auxiliar", "Administrador"))
Difare$GENERO <- factor(Difare$GENERO, levels = c(1,2),
                      labels = c("Masculino", "Femenino"))
Difare$DISCAPACIDAD <- factor(Difare$DISCAPACIDAD, levels =
                             c(0,1), labels = c("No", "Si"))
Difare$ROTACION.TEMPRANA <-
    factor(Difare$ROTACION.TEMPRANA, levels=
           c(0,1), labels= c("No", "Si"))
Difare$CAUSAS_DE_SALIDA <-
    factor(Difare$CAUSAS_DE_SALIDA, levels =
           c(1,2,3), labels = c("Clima laboral",
                              "Despido", "Oportunidad de crecimiento"))
Difare$MOTIVO <- factor(Difare$MOTIVO, levels = c(1,2,3), labels =
```

```
c("RV","RN","DI"))
```

Para dar inicio al análisis descriptivo de la problemática, se procedió a realizar una tabla de frecuencia y un gráfico de barras con la variable de estudio identificada como Motivo, donde las 3 respuestas son: RV (Renuncia Voluntaria), RN (Renuncia Negociada) y DI (Despidos Intempestivos). Donde, a la base Difare la tratamos con *as.data.frame*, la cual es una función que permite cohesionar a data.frame por la variable en estudio y nos otorga su frecuencia relativa.

Posterior a esto, por medio de *transform* se trata la frecuencia estudiada (FreDifare), calculando sus 3 Tipos (Relativa, Acumulada y Relativa Acumulada), para con ello, a través de su Frecuencia Relativa poder obtener una mejor visualización mediante un gráfico de barras, mismo que al utilizar *ggplot* en conjunto de *aes* (), se genera una relación de las variables y adiciona la leyenda de los datos de acuerdo a formas, tamaño, identificación y color, los cuales, para Difare fueron detalladas como; en el eje de las **X** la variable y en el eje de las **Y** su frecuencia relativa.

#Ejecución Tabla de Frecuencia y Gráfico de Barras Variable Motivo

```
FreDifare <- as.data.frame(table(Difare$MOTIVO))
FreDifare
TabDifare <- transform(FreDifare,
                        FreRel= round(prop.table(FreDifare$Freq),3),
                        FreAcu = cumsum(FreDifare$Freq),
                        FreRelAcu= round(cumsum(prop.table(FreDifare$Freq)),3))
TabDifare
GraDifare <- ggplot(TabDifare,aes(x=TabDifare$Var1,y=TabDifare$Freq))+
  geom_bar(stat = "identity",fill= "yellow",colour= "black")
GraDifare
```

Por consiguiente, se llegó a elaborar distintos gráficos estadísticos (Gráficos Combinados y Evolutivos). Los gráficos combinados (*Geom_bar*) son implementados para poder contrarrestar dos variables con varios subconjuntos durante un periodo de tiempo establecido.

#Graficos Combinados (Estadistica Descriptiva)

```
GCANO <- ggplot(Difare,aes(Difare$ANO))+
  geom_bar(aes(fill=Difare$GENERO))
GCANO
GCAREA <- ggplot(Difare,aes(Difare$AREA))+
```

```

    geom_bar(aes(fill=Difare$MOTIVO))
GCAREA
GCPUESTO <- ggplot(Difare,aes(Difare$PUESTO))+
    geom_bar(aes(fill=Difare$MOTIVO))
GCPUESTO
GCAREAYCAUSA <- ggplot(Difare,aes(Difare$AREA))+
    geom_bar(aes(fill=Difare$CAUSAS_DE_SALIDA))
GCAREAYCAUSA
GCPUESTOYCAUSA <- ggplot(Difare,aes(Difare$PUESTO))+
    geom_bar(aes(fill=Difare$CAUSAS_DE_SALIDA))
GCPUESTOYCAUSA
GCGENEROYCAUSA <- ggplot(Difare,aes(Difare$GENERO))+
    geom_bar(aes(fill=Difare$CAUSAS_DE_SALIDA))
GCGENEROYCAUSA

```

Mientras que, por otro lado se desarrollaron gráficos de caja o boxplot (*Geom_boxplot*), los cuales muestran la distribución estadística de un conjunto de datos, mediante la proporción de información de la mediana, cuartiles y presencia de posibles valores atípicos.

```

#Graficos Evolutivos ( Estadística Descriptiva)
GECGENERO <- ggplot(Difare,aes(x=Difare$AREA,
    y=Difare$SueldosYComisiones))+
    geom_boxplot(fill="yellow",colour="black")
GECGENERO
GECPUUESTO <- ggplot(Difare,aes(x=Difare$PUESTO,
    y=Difare$SueldosYComisiones))+
    geom_boxplot(fill="yellow",colour="black")
GECPUUESTO

```

Una vez, se obtuvo el análisis descriptivo, se procedió a descargar y llamar a un conjunto de paquetes, los cual son útiles para el desarrollo del modelo de clasificación:

```

#Instalar y Llamar Librerías
install.packages("caret")

```

```
install.packages("rpart")
install.packages("rpart.plot")
install.packages("randomForest")
library(caret)
library(rpart)
library(rpart.plot)
library(randomForest)
```

La forma de llamar estas librerías es mediante la descarga de sus paquetes con la función (*install.packages*) y luego utilizando la función (*library*) tal como se muestra anteriormente.

El primero, se llama *Caret* (Clasificación y Entrenamiento de Regresión), el cual nos brinda una variedad de funciones que permiten implementar diversos métodos elaborados de clasificación y regresión. Además, a través de un código unificado posibilita la ejecución de distintas reglas de clasificación en distintos paquetes, y al mismo tiempo, permite realizar particiones entre los datos tratados.

Mientras que otro paquete utilizado es *Rpart*, el cual engloba todo lo necesario para la debida implementación de árboles de clasificación, y a partir del conjunto de datos tratados obtener una predicción confiable.

Adicional, se utilizó *Rpart.plot*, el cual permite graficar los resultados arrojados por la función *Rpart*. Por último, se realizó el uso de *RandomForest*, mismo que facilita la construcción, ajuste y evaluación de modelos.

En consecuencia, para el desarrollo del modelo de clasificación, Bosque Aleatorio, se procedió a consolidar una nueva base de datos identificada como *Difare2*, la cual está constituida por la eliminación de las variables cuantitativas, y de esta manera, se conservó las variables convertidas en factor.

#Eliminación de Variables Cuantitativas

```
Difare2 <- Difare [, c (-4, -5, -12, -13)]
```

Una vez se hayan descargado todos los paquetes, llamado a través de sus librerías respectivas, y al tener las variables categóricas listas para ser tratadas, se procedió a llamar a la semilla denominada, *Set.seed (123)*, la cual nos permite obtener una mejor clasificación de datos. Dentro de R, sembrar una semilla implica activar un generador de números pseudoaleatorios, dado que la mayoría de los métodos de simulación se apoyan en la presencia de este generador.

#Plantar Semilla

```
set.seed(123)
```

El siguiente paso consistió en llevar a cabo el entrenamiento de los datos. El entrenamiento se da a partir de la función *createDataPartition*, la cual realiza una serie de particiones de los datos de la variable analizada, en este caso la variable *Motivo* de la base de datos *Difare2*, donde *p* explica el porcentaje de datos a ser entrenados (0.90), mientras que *list* es un argumento lógico que indica como los datos serán presentados, para *Difare2*, en una lista *false*.

#Entrenamiento de Datos

```
dataentrenamiento <- createDataPartition(Difare2$MOTIVO, p=0.90, list = FALSE)
```

De forma consecuente, se procedió a crear el Bosque Aleatorio, el cual se lo denominó como *RandomTreeModel*, mismo que hace referencia a la función *Random Forest* que estudia la data en función a; por el eje de las X desde la primera variable a la décima (*División, Área, Puesto, Género, Discapacidad, Edad, Rotacion.Temprana, Meses_Ant, Causas_De_Salida; Sueldos Y Comisiones*), mientras que por el eje de las Y estudia a la variable onceava (*Motivo*). Para un árbol (*ntree*) de 3000 ramificaciones, y bajo el parámetro *keep.forest*, el cual está definido como *True*,

que establece que para el test el valor predeterminado será verdadero y el bosque se mantendrá en el objeto de salida.

#Ejecución de Bosque Aleatorio (Random Forest)

```
RandomTreeModel <- randomForest(x=Difare2[dataentrenamiento,  
1:10],y=Difare2[dataentrenamiento,11],ntree= 3000,  
keep.forest = TRUE)
```

RandomTreeModel

Por tanto, se llevaron a cabo las pruebas del modelo a través de la predicción. La función implementada se denomina *predict*, la cual se utiliza para predecir los valores en función de los datos de entrada. Donde el objeto es *RandomTreeModel*, y los datos de entrada para pronosticar los valores hacen referencia a la nueva data denominada *Difare2[-data entrenamiento,]*.

#Ejecución de Predicción

```
Prediccion = predict (RandomTreeModel,Difare2[-dataentrenamiento,])  
Prediccion
```

A partir de esto, se pudo desarrollar la matriz de confusión, la cual nos permitió evaluar la fiabilidad de los datos obtenidos de la predicción. Esta vez la identificó con el nombre *Matriz*, donde la tabla, bajo el nombre *table*, agrupa los datos de entrada, la variable de estudio Motivo y Predicción, y bajo el vector *dnn=c()* se personaliza la tabla en resultados identificados como Actual y Predicho.

#Ejecución de la matriz de confusión

```
Matriz =table(Difare2 [-dataentrenamiento,"MOTIVO"], Prediccion,  
dnn = c ("Actual", "Predicho"))
```

Matriz

Y, por último, se llevó a cabo la clasificación mediante la probabilidad de éxito. La misma que está establecida como *Difare2\$Probabilidad* y a través de la función *predict*, se estudia el Bosque aleatorio, los argumentos de la base de datos *Difare2*, bajo un tipo de probabilidad “*prob*”.

#Clasificación mediante la probabilidad de éxito

```
Difare2$Probabilidad = predict(RandomTreeModel,Difare2,type = "prob")
```

Probabilidad

Resultados

A través de los modelos empleados en la investigación, se logró realizar un análisis detallado, el cual permitió la debida identificación de los factores determinantes de las renuncias tempranas de la Distribuidora Farmaceutica ecuatoriana Difare S.A.

Para empezar, se realizó una tabla de Frecuencias (Tabla 2) en conjunto a un gráfico de barras (Figura 8), sobre la variable en estudio Motivo. A partir de lo cual se logró identificar que el 98% del personal de trabajo de Grupo Difare ha renunciado de manera voluntaria. Mientras que, el 10% ha sido por un despido intempestivo, y apenas el 6% lo ha hecho por términos negociados.

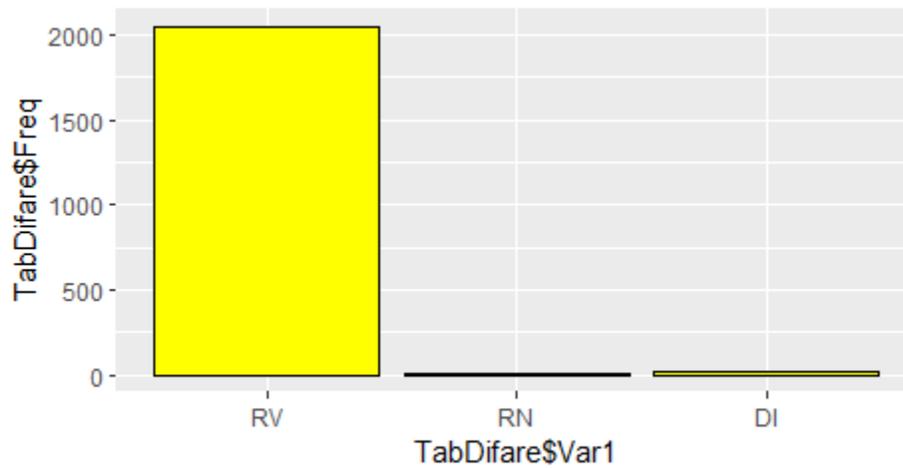
Tabla 2

Tabla de Frecuencias Variable Motivo en Rstudio

	Var1	Freq	FreRel	FreAcu	FreRelAcu
1	RV	2054	0.985	2054	0.985
2	RN	12	0.006	2066	0.99
3	DI	20	0.01	2086	1

Figura 8

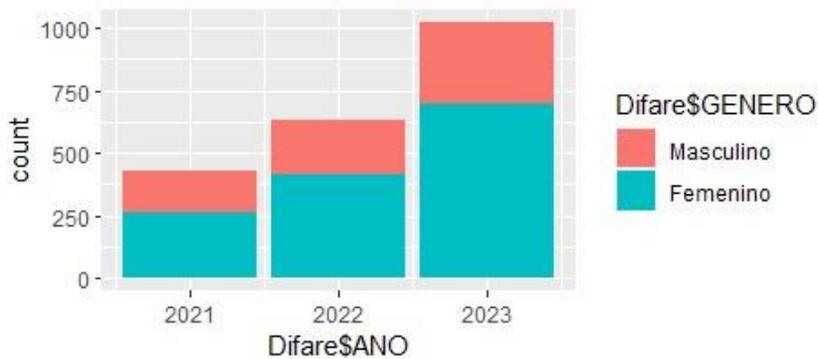
Gráfico de Barras de Variable Motivo en RStudio



Posteriormente, se realizaron los siguientes gráficos combinados con el objetivo de tener una visión clara de la problemática de la indagación.

Figura 9

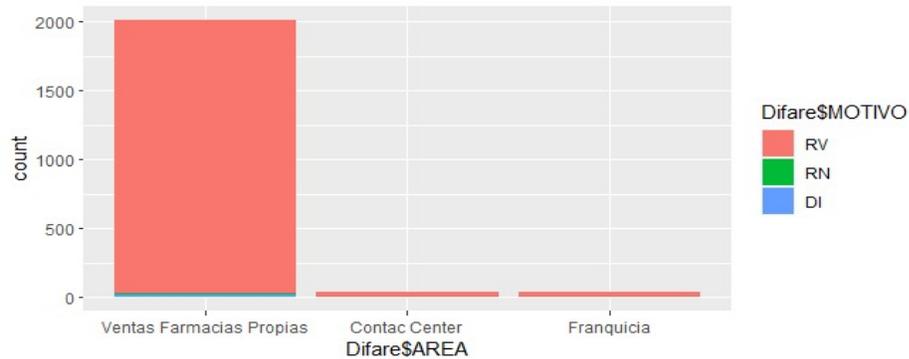
Gráfico Combinado Año vs Género en Rstudio



Mediante el gráfico Combinado de la Figura 9, se puede evidenciar que existió mayor rotación del género femenino en los 3 años de análisis con un crecimiento exponencial a diferencia del género masculino.

Figura 10

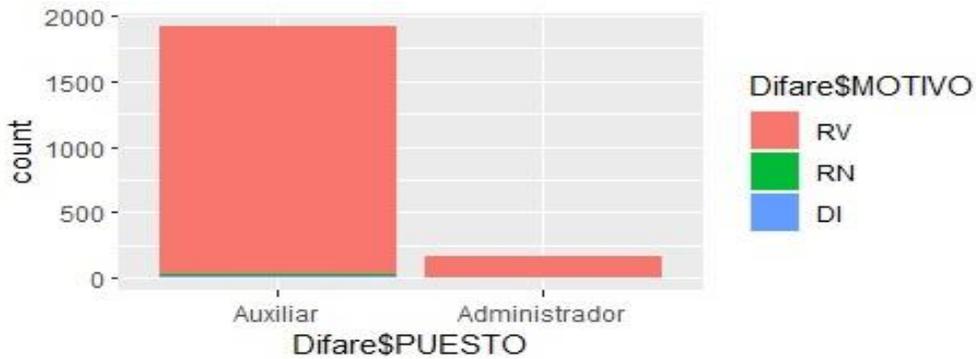
Gráfico Combinado Área vs Motivo en RStudio



Por otro lado, en la Figura 10, se evidencia que el área de farmacias propias tiene una mayor cantidad de renuncias voluntarias frente a las otras áreas como contact center y franquicia.

Figura 11

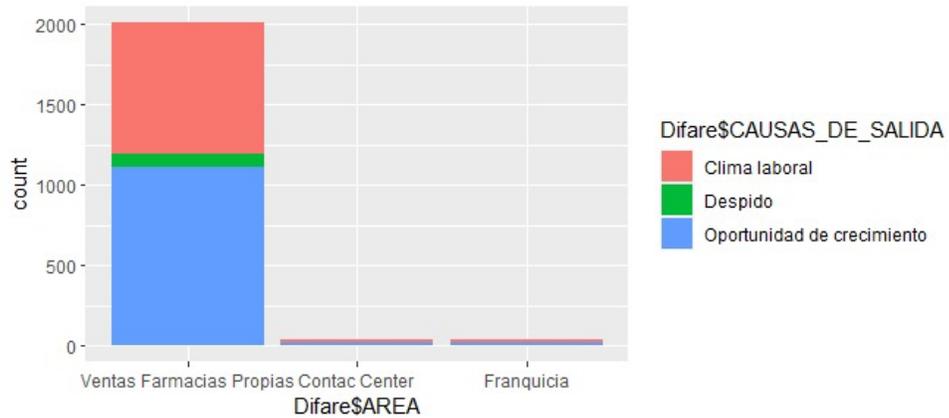
Gráfico Combinado Puesto vs Motivo en Rstudio



En la Figura 11 se muestra que los auxiliares representan un mayor flujo de rotación de manera voluntaria, mientras que los despidos intempestivos no han sido una opción para el personal de farmacias de Difare.

Figura 12

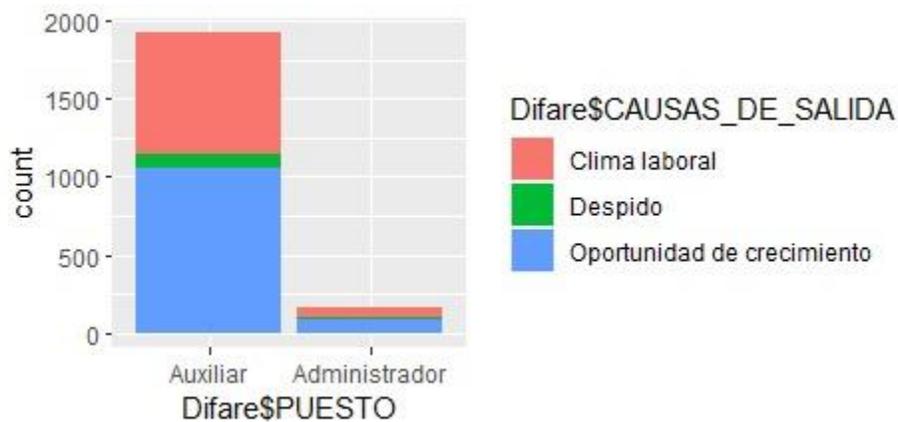
Gráfico Combinado Área vs Causa de Salida en RStudio



Por otra parte, en la Figura 12, se refleja que la mayor causa de rotación en las farmacias propias han sido la oportunidad de crecimiento como también el clima laboral.

Figura 13

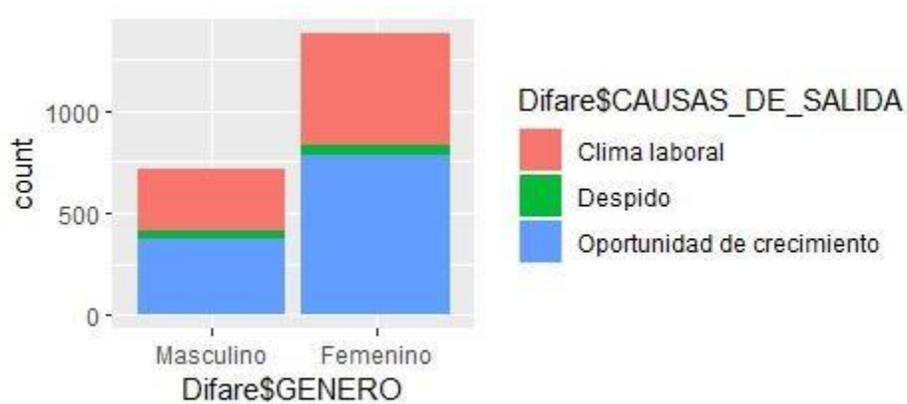
Gráfico Combinado Puesto vs Causa de Salida en Rstudio



En la Figura 13 se confirma que las causas de salida de los auxiliares también fueron el clima laboral y oportunidad de crecimiento.

Figura 14

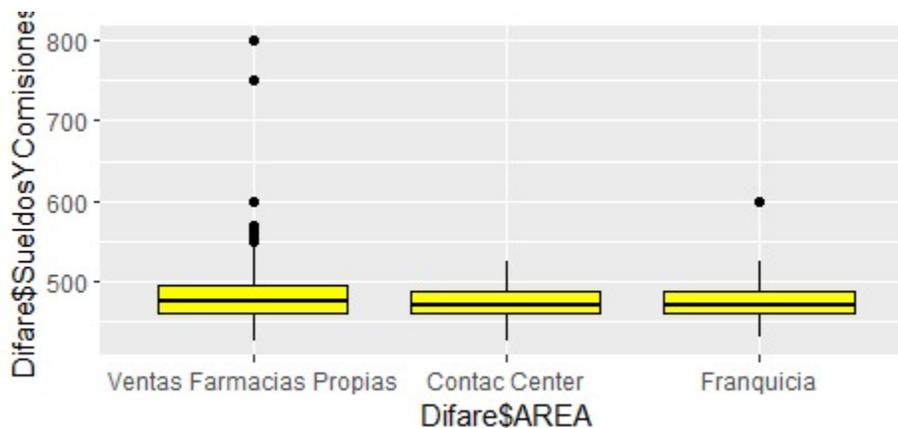
Gráfico Combinado Género vs Causa de Salida en Restudio



En en la Figura 14 se explica que el género femenino presenta una mayor rotación por el clima laboral y oportunidad de crecimiento.

Figura 15

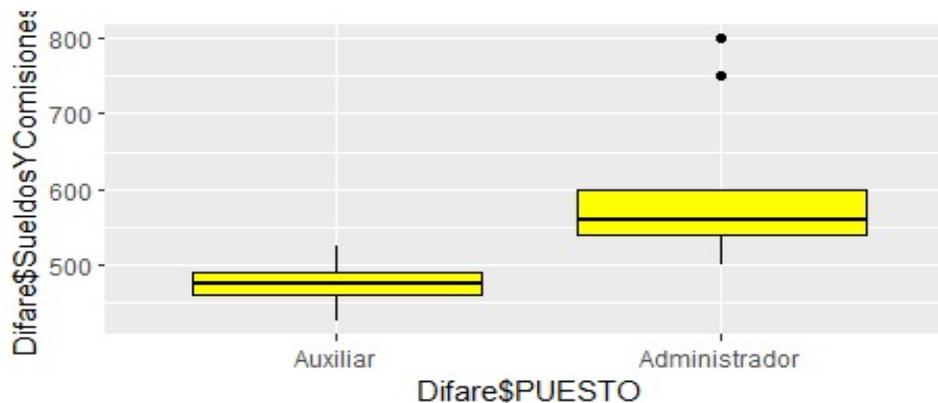
Gráfico Evolutivo Área vs Sueldos y Comisiones en Rstudio



Según el Figura 15 se puede evidenciar que el área de contact center tiene mayor simetría debido a que ambas semi-cajas tienen la misma longitud, además de no poseer valores atípicos.

Figura 16

Gráfico Evolutivo Puestos vs Sueldos y Comisiones en Rstudio



Por último, la Figura 16 indica que la distribución de los auxiliares ha sido aproximadamente simétrica al tener una semicaja con la misma longitud, mientras que los administradores han tenido un sesgo negativo, debido a que la semi-caja es más larga del lado posterior de la mediana que el lado inferior.

Luego de realizar las figuras descriptivas, mismos que permitieron identificar la problemática, se desarrolló el modelo de clasificación Random Forest como se puede evidenciar en la figura a continuación:

Figura 17

Modelo Random Forest en RStudio

Call:

```
randomForest (x = Difare2 [ dataentrenamiento, 1:10] , y = Difare2 [dataent
```

```
11] , ntree = 3000 , keep.forest = TRUE
```

```
Type of random Forest: classification
```

```
Number of trees: 3000
```

```
No. of variables tried at each split: 3
```

```
OOB estimate of error rate: 1.6%
```

Tabla 3*Matriz de Confusión*

	RV	RN	DI	class.error
RV	1848	0	1	0.0005408329
RN	11	0	0	1.000000
DI	18	0	0	1.000000

La Tabla 3 mediante el OOB estimate of error rates demuestra que el bosque aleatorio tiene una eficiencia del 98.4% lo que nos indica que la predicción ha sido confiable.

Tabla 4*Matriz de Confusión en Rstudio*

	Predicho		
<i>Actual</i>	<i>RV</i>	<i>RN</i>	<i>DI</i>
<i>RV</i>	204	0	1
<i>RN</i>	1	0	0
<i>DI</i>	2	0	0

Como último punto la Tabla 4 muestra que a través de la matriz de confusión se comprueba que los datos son confiables, puesto a que el class.error coincide con lo predicho por el random forest al ser menor a 1%. Lo que indica que la mayoría de las renuncias tempranas se han dado de forma voluntaria, por dos razones evidentes; clima laboral y oportunidad de crecimiento.

Análisis de Resultados

Entre los principales hallazgos que ha revelado el análisis descriptivo en conjunto al modelo de clasificación, se pudo identificar que el género con mayor rotación de Grupo Difare, ha sido el femenino en el periodo 2020 - 2023, donde la mayor parte de estas renunciaciones han sido de forma voluntaria y se evidencia que pertenecen al área de farmacias propias, dando como resultado que la principal causa de salida sea el clima laboral. Además, se refleja que el cargo con mayor rotación es el de “auxiliar de farmacia propias”, los cuales constantemente se encuentran en búsqueda de oportunidad de crecimiento.

Es importante mencionar que, aunque es inevitable que exista rotación de personal en las compañías, esta puede afectar la productividad y la cohesión de los equipos. A través de los gráficos evolutivos, se pudo evidenciar que los sueldos y comisiones no están siendo la causa principal de salida de los colaboradores. Sin embargo, los auxiliares de contact center son los que demuestran ser remunerados de una forma más equitativa, a diferencia de farmacias propias y franquicias.

Por consiguiente, mediante el algoritmo Random Forest se comprueba lo anteriormente descrito. Puesto que, el análisis realizado en RStudio, nos arrojó un OOB estimate of error rate del 1.60%, lo que representa fiabilidad de la predicción. Lo mismo, que es justificado a través de la matriz de confusión, al arrojar un class.error de menos del 1%. Siendo así, que se ha logrado identificar que el trabajo en la Distribuidora Farmacéutica Ecuatoriana Difare S.A. es un empleo seguro, pero no poseen un ambiente laboral ideal para conservar a su equipo.

Dado que, de manera consecuente se visualiza que las renunciaciones de forma voluntaria tienen una probabilidad de ocurrencia del 100%, y esto se relaciona con lo identificado por el análisis descriptivo, de que los colaboradores rechazan el clima laboral, lo que a su vez conlleva a la búsqueda de otras oportunidades laborales.

Por otro lado, un estudio de investigación centralizado en analizar el comportamiento de deserción de los clientes de una entidad comercial enfocada en la administración de fondos en Ecuador decide de igual forma que Grupo Difare, implementar el modelo de predicción, bosque aleatorio. Pero por su parte, con el objetivo de gestionar mejoras que les permita alcanzar la fidelización de sus clientes. Donde la predicción de los clientes de tipo desertores a partir de Random Forest, arrojo que las cualidades que definen a estos 150000 usuarios entrevistados son; el saldo de fondo, tipo de sistema de aporte (pago por rol, tarjeta de débito o crédito) y el tipo de dependencia del colaborador. Adicionalmente, Rstudio arrojo un OOB estimate of error rate del 7%, lo que resumió que las predicciones obtenidas son correctas, con un grado de confiabilidad del 93% (Bohórquez et al., 2020).

Por tal motivo, que la empresa administradora de fondos en Ecuador logró identificar cuáles son los atributos que definen a los desertores de su compañía, y según la perfilación obtenida tras la investigación, se pudo identificar que 30577 entrevistados afiliados presentaron características de abandono. Por lo cual, la compañía está a tiempo, de tomar acciones correctivas que permitan obtener la fidelización de esos usuarios y atraer a más que no se hayan involucrado en el negocio brevemente gracias a la predicción de la técnica de machine learning (Bohórquez et al., 2020).

En síntesis, el modelo de clasificación Bosque Aleatorio, permite evaluar e identificar el modelo de predicción óptimo para cada estudio realizado, y de tal forma que conserve un balance entre el error de su predicción y su varianza. Donde para Grupo Difare, las probabilidades de renuncia voluntaria demostraron que será complicado mantener a los recursos si el clima laboral no presenta mejoría y si no se ofrece oportunidades de crecimiento que les permita a los colaboradores desarrollar sus habilidades e incrementar su motivación de continuidad en la compañía.

En los anexos se podrá visualizar la base de datos con los resultados de las probabilidades y predicción del modelo.

Conclusiones

La Inteligencia Artificial, identificada como una disciplina dentro de la rama de la informática, en conjunto de una de sus ramificaciones como lo es el Machine Learning han logrado en la actualidad, ser relevantes para las entidades comerciales, al posibilitar el análisis de extensas cantidades de datos, identificando patrones y tendencias que mejoran la toma de decisiones. Siendo así, que el papel desempeñado por la IA dentro de las organizaciones ha sido evolutivo al facilitar la automatización de tareas y procesos, contribuyendo al aumento de la eficiencia operativa.

Siendo así, mediante la utilización de Random Forest en análisis predictivos sobresale por su habilidad para gestionar una base de datos complejos, tal como la usada en el desarrollo de este proyecto para mitigar el sobreajuste y generar predicciones precisas. Este algoritmo versátil proporciona resultados sólidos en diversas aplicaciones desde clasificación hasta regresión, así misma mejora los árboles de decisión creándose así un modelo más fiable.

Por ende, a través del modelo de clasificación aplicado, y el algoritmo Random Forest o Bosque Aleatorio, se pudo identificar que las renuncias dentro de Distribuidora Farmaceutica Ecuatoriana Difare S.A, se dan en mayor frecuencia por motivo voluntario, y por razones como; el clima laboral y por consiguiente, una búsqueda de alternativas de crecimiento. Lo que llega a demostrar que el alto ejecutivo de la compañía no es la razón de las renuncias, sin embargo, se debe dar una atención estricta a los incentivos que se brinda a su personal, para poder conservar al talento de la entidad comercial.

Recomendaciones

Dentro de las sugerencias para la Distribuidora Farmacéutica Ecuatoriana Difare S.A., se incluye la evaluación del entorno laboral de la empresa con el fin de detectar los factores que contribuyen a la rotación. La atmósfera laboral en los puntos de venta cobra gran relevancia para los vendedores, quienes pasan la mayor parte de su jornada laboral en estos lugares. Es esencial involucrar a los jefes de zona y administradores que poseen fuertes cualidades de liderazgo positivas para abordar este problema.

Se sugiere examinar la estructura de beneficios y compensaciones para garantizar su competitividad, utilizando KPIs como Ventas por Punto de Venta y Storecheck. Al implementar el indicador de Ventas al 35%, se beneficiaría a los empleados con el 50% de las ventas mensuales de 17 productos específicos, preferiblemente aquellos con mayor inventario. Esta medida no solo aumentaría la rentabilidad al agotar stock, sino que también agregaría un monto adicional a las remuneraciones mensuales de los colaboradores.

Asimismo, con el indicador de Storecheck al 20% se busca beneficiar a los colaboradores por la exhibición y marketing del punto de venta. Con esto se puede verificar si el producto está llegando al anaquel y se encuentra disponible para el cliente.

Siendo así, que a través de programas de reconocimiento y motivación se puede lograr que los colaboradores sientan oportunidades de desarrollo.

Referencias

- Ahumada, (2016). Aplicaciones de inteligencia artificial en procesos de cadenas de suministros: una revisión sistemática. *Ingeniare. Revista chilena de ingeniería*, 24(4), 663-679.
<https://doi.org/10.4067/s0718-33052016000400011>
- Alarcón, K. (2022). *El teletrabajo y su incidencia en el desempeño de los colaboradores de una empresa del sector privado*. [Tesis de Posgrado, Universidad Católica de Santiago de Guayaquil]. <http://repositorio.ucsg.edu.ec/handle/3317/20158>
- Alfonso, M., Cazorla, M., Colomina, O., Escolano, F. y Lozano, M. (2003). Inteligencia artificial: Modelos, técnicas y áreas de aplicación [Digital]. En Modelos, técnicas y áreas de aplicación (1.a ed.). *Paraninfo*.
https://books.google.com.ec/books?id=_spC6S7UfZgC&dq=inteligencia+artificial+pdf&lr=&hl=es&source=gbs_navlinks_s
- Alvares (2020). Rotación de personal ¿Qué es y cómo combatirla? *ResearchGate*.
https://www.researchgate.net/publication/353023868_ROTACION_DE_PERSONAL_Qu_e_es_y_como_combatirla
- Álvarez, M., Quirós, L y Cortés, M. (2020). Inteligencia artificial y aprendizaje automático en medicina. *Revista Médica Sinergia*, 5(8).
<https://www.revistamedicasinergia.com/index.php/rms/article/view/557/923>
- Amit, Y. & Geman, D. (1997). Shape quantization and recognition with randomized trees. *Neural Computation*, 9, 1545–1588.
- Aranibar, M., Melendres, V., Ramírez, M., & García, B. (2018). Los factores de la rotación de personal en las maquilladoras de exportación en ensena, B.C. *Revista Global de Negocios*, 6(2), 25-40. <https://www.theibfr2.com/RePEc/ibf/rgnego/rgn-v6n2-2018/RGN-V6N2-2018-3.pdf>
- Arroyo, E. (2022). Tendencias laborales 2022. *Adecco Ecuador*.
<https://adecco.com.ec/tendencias-laborales-2022/>
- Arthur, L. (1959). Some Studies in Machine Learning Using the Game of Checkers, *IBM Journal*.

- Banco Central del Ecuador (2023). *La pandemia incidió en el crecimiento 2020: La economía ecuatoriana decreció 7,8%*. <https://www.bce.fin.ec/index.php/boletines-de-prensa-archivo/item/1421-la-pandemia-incidio-en-el-crecimiento-2020-la-economia-ecuatoriana-decrecio-7-8>.
- Barbosa, J. (2021). Modelos Predictivos para la Rotación del Talento Humano. *Tecnología Investigación y Academia*, 8(1), 54–71.
<http://revistas.udistrital.edu.co:8080/index.php/tia/article/view/14728>
- Barrera, L. (2012). Fundamentos históricos y filosóficos de la inteligencia artificial UCV-HACER. *Revista de Investigación y Cultura*, 1(1), 87-92.
<https://www.redalyc.org/pdf/5217/521752338014.pdf>
- Bishop, C. (2007). *Pattern Recognition and Machine Learning*. New York, NY. Springer.
- Bobadilla, J. (2020). *Machine Learning y Deep Learning Usando Python, Scikit y Keras*. Ra-Ma Editorial
- Bohórquez, M., Torys, J., y Paredes, M. (2020). Modelos de Predicción de Deserción de Clientes para una Administradora de Fondos Ecuatoriana. *Revista Compendium. Vol.7 (1),1-11*.
<http://www.revistas.espol.edu.ec/index.php/compendium/article/view/777/702>
- Boden, M. (2022). *Inteligencia artificial*. Edición Turner de Madrid, España, 192.
<https://www.marcialpons.es/libros/inteligencia-artificial/9788418895357/>
- Brazdil, P., y Jorge, A. (2001). Progress in Artificial Intelligence: Knowledge Extraction, Multi-agent Systems, Logic Programming, and Constraint Solving. *Heidelberg: Springer Berlin Heidelberg*. <https://doi.org/10.1007/3-540-45329-6>
- Breiman, L. (1984). Classification and regression trees. *CRC press*. doi: 10.1201/9780429259260
- Breiman, L. (2001). Random Forest. *Machine Learning*, 45, 5–32.
<https://link.springer.com/article/10.1023/a:1010933404324>
- Cahuasqui, E., y Zapata, S. (2023). Mantenimiento Predictivo de rotación Cahuasqui, E., & Zapata, S. (2023). Mantenimiento Predictivo de rotación de personal en Microempresas con IA. *Journal of Science and Research*, 8(4), 131–151.
<https://revistas.utb.edu.ec/index.php/sr/article/view/2952>

- Carbonell, L. (2022). *Análisis financiero y evaluación del crecimiento de las principales empresas mundiales de los sectores de salud y farmacia*. [Trabajo de Fin de Grado, Universidad Politécnica de Valencia]. Repositorio Universidad Politécnica de Valencia. <https://riunet.upv.es/bitstream/handle/10251/192810/Carbonell%20-%20ANALISIS%20FINANCIERO%20Y%20EVALUACION%20DEL%20CRECIMIENTO%20DE%20LAS%20PRINCIPALES%20EMPRESAS%20MUNDIALE....pdf?sequence=1&isAllowed=y>
- Cárdenas, J. (2018). *El machine learning a través de los tiempos, y los aportes a la humanidad*. [Trabajo de Fin de Grado, Universidad Libre Seccional Pereira]. Universidad Libre Seccional Pereira. <https://repository.unilibre.edu.co/bitstream/handle/10901/17289/EL%20MACHINE%20LEARNING.pdf?sequence=1&isAllowed=y>
- Carreón, L. (2016). Estudio de caso de satisfacción laboral en trabajadores de una institución pública de salud. *Revista de Psicología Procesos Psicológicos y Sociales*, 1-51. <https://www.uv.mx/psicologia/files/2016/10/liliana.pdf>
- Charaja, Y., y Manami, J. (2013). Satisfacción laboral y motivación de los trabajadores de la dirección regional de comercio exterior y turismo - Puno – Perú. *Comuni@cción*, 5(1), 05-13. http://www.scielo.org.pe/scielo.php?pid=S2219-71682014000100001&script=sci_abstract
- Chaviano, H. (2015). Técnicas de aprendizaje supervisado y no supervisado para el aprendizaje automatizado de computadoras, *Dialnet*, 549-564. <https://dialnet.unirioja.es/servlet/articulo?codigo=7192675>
- Chiavenato. (1990). *Administración de Recursos Humanos*. México: 5.
- Chiavenato. (2005). *Administración de Recursos Humanos*. México: 5
- Choi, J., Robb, C., Mifli, M., & Zainuddin, Z. (2021). University students' perception to online class delivery methods during the COVID-19 pandemic: A focus on hospitality education in Korea and Malaysia. *Journal of Hospitality, Leisure, Sport & Tourism Education*. <https://www.sciencedirect.com/science/article/pii/S147383762100037X>

- Constitución de la República del Ecuador. (2008). Código de Trabajo de 2012.
- Del Río, R. (2023). La industria farmacéutica se convierte en la tercera fuerza exportadora en España en 2022 - FarmaIndustria. *FarmaIndustria*.
<https://www.farmaindustria.es/web/prensa/notas-de-prensa/2023/02/20/la-industria-farmacautica-se-convierte-en-la-tercera-fuerza-exportadora-en-espana-en-2022/>
- Deloitte Consulting. (2021). *Doing Business Ecuador*
<https://www2.deloitte.com/ec/es/pages/about-deloitte/articles/doing-business-ecuador-2021.html>
- Faggella, D. (2018). 7 Applications of Machine Learning in Pharma and Medicine. *TechEmergence*. <https://www.techemergence.com/machinelearning-in-pharma-medicine/>
- García, A. (2012). *Inteligencia Artificial. Fundamentos, práctica y aplicaciones*. Libros RC.
- Ghosh, D., y Vogt, A. (2012). *Outliers: An Evaluation of Methodologies. Proceedings of the Survey Research Methods Section, American Statistical Association*.
http://www.amstat.org/sections/srms/Proceedings/y2012/Files/304068_72402.pdf
- Gil, V., y Seguro, C. (2022). Machine Learning aplicado al análisis del rendimiento de desarrollos de software. *Revista politecnica*, 18(35), 128-139.
<https://doi.org/10.33571/rpolitec.v18n35a9>
- Giletta, M., Giordano, A., Mercau, N., Orden, P., y Villareal, V. (2020). Inteligencia Artificial: Definiciones en disputa. *Revista Sociales Investiga. Escritos Académicos, de Extensión y Docencia*, 9, 20-33.
<https://socialesinvestiga.unvm.edu.ar/ojs/index.php/socialesinvestiga/article/view/320/287>
- Globalratings (2022). *Calificadora de Riesgos S.A. – GlobalRatings Calificadora de Riesgos S.A. WEB Publica y corporativa*. <https://www.globalratings.com.ec/site1/>
- Henao, C. (2021). *Modelo de medición de la rotación de personal como variable de decisión estratégica* [Tesis de pregrado, Universidad Nacional de Colombia].
<https://repositorio.unal.edu.co/handle/unal/81925>
- Higiano, C. (2017). Predicción del rendimiento académico aplicando técnicas de minería de datos. *Revista Dialnet*, 78(1). <https://dialnet.unirioja.es/servlet/articulo?codigo=6171237>

- Hosmer, W., y Lemeshow, S. Applied Logistic Regression. Second Edition. A Wiley-Interscience Publication. John Wiley & Sons, INC. 2000.
- Hurwitz, J., y Kirsch, D. (2018). Machine Learning: For Dummies (Vol. 1). John Wiley & Sons, Inc. <https://www.ibm.com/downloads/cas/GB8ZMQZ3>
- Ibarra, M. (2015). *Programas y Estrategias asociadas a la Responsabilidad Social en la gestión de una Distribuidora Farmacéutica: Estudio de Caso DIFARE S.A.* [Tesis de Bachelor, Universidad Casa Grande del Ecuador]. Repositorio Institucional. <http://dspace.casagrande.edu.ec:8080/handle/ucasagrande/734>
- IBM Cloud Education. (2021). Redes neuronales. IBM. <https://www.ibm.com/cloud/es/cloud/learn/neural-networks>
- Incio, F., Capuñay, D., Estela, R., Valles, M., Vergara, S. y Elera, D. (2021). Inteligencia artificial en educación: una revisión de la literatura en revistas científicas internacionales. *Revista de Investigación Apuntes Universitarios*, 12(1). <https://apuntesuniversitarios.upeu.edu.pe/index.php/revapuntes/article/view/974/866>
- Instituto Español de Comercio Exterior (2022). *ICEX España Exportación e Inversiones*. <https://www.icex.es/>
- Jiménez, A. y Díaz, J (2021). Técnicas de aprendizaje automático (machine learning). *Revista sistemática de Literatura*, (13), 113-121. <https://ojs.tdea.edu.co/index.php/cuadernoactiva/article/view/849/1366>
- Jiménez, A., y Díaz, J. (2021). Revisión sistemática de literatura: Técnicas de aprendizaje automático (machine learning). *Revisión sistemática de literatura: Técnicas de aprendizaje automático (machine learning)*. *Cuaderno Activa*, 13, 113-121. <https://ojs.tdea.edu.co/index.php/cuadernoactiva/article/view/849/1366>
- Kenzo, L. (2018). Aplicaciones del aprendizaje reforzado en robótica móvil [Universidad de Chile]. <https://repositorio.uchile.cl/bitstream/handle/2250/152970/Aplicaciones-del-aprendizaje-reforzado-en-robotica-movil.pdf?sequence=1&isAllowed=y>
- LeCun, Y., Bengio, Y., y Hinton, G. E. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>

- Leyva, M., y Smarandache, F. (2018). Inteligencia Artificial: retos, perspectivas y papel de la Neutrosología. *Revista Dilemas Contemporáneos: Educación, Política y Valores*, 5(1).
<https://dilemascontemporaneoseducacionpoliticayvalores.com/index.php/dilemas/article/view/645/1166>
- Liaw, A., y Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22. <https://doi.org/10.1177/004912419002900307>
- Lizares, M. (2017). *Comparación de modelos de clasificación: regresión logística y árboles de clasificación para evaluar el rendimiento académico*. [Tesis de Bachelor, Universidad Nacional Mayor de San Marcos]. Repositorio Universidad del Perú.
https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/7122/Lizares_cm.pdf?sequence=3&isAllowed=y
- Lodhi, H., Masood, S., Shahbaz, S. y Siddiqui, G. (2020). Effect of Leadership Styles on Organizational Citizenship Behavior and Employee Turnover Intention. *Journal of Accounting and Finance in Emerging Economies*, 6(2), 487-495.
<https://publishing.globalcsrc.org/ojs/index.php/jafee/article/view/1200>
- López, L., Acosta, N., y Gago, A. (2019). Detección de anomalías basada en aprendizaje profundo: revisión. *Universidad de las Ciencias Informáticas*.
<https://www.redalyc.org/journal/3783/378365913008/html/>
- Macario, F. (2018). *Rotación de personal y clima organizacional* [Tesis de Pregrado, Universidad Rafael Landívar]. <http://recursosbiblio.url.edu.gt/tesisjrzd/2018/05/43/Macario-Fredy.pdf>
- Macas, S., Fernández, M., y Espinoza, N. (2022). Análisis del índice de rotación de personal y sus posibles causas en farmacia CIA. Ltda. *Ciencia latina*, 6(4), 3630-3648.
https://doi.org/10.37811/cl_rcm.v6i4.2866
- Maday, Y., Machado, M., Machado, N., y Velastegui, E. (2022). Técnicas de aprendizaje supervisado para la detección y clasificación de enfermedades y defectos en imágenes de frutas: revisión. *Revista de Investigación e Innovación*, 7(1).
<https://revistas.utb.edu.ec/index.php/magazine/article/view/2330/1983>

- Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*, 9, 381-386.
- Martín, J., Rodrigo, M., y Rodríguez, E. (2013). Las Competencias en Contextos de Riesgo Psicosocial. Colegio Oficial de Psicólogos de Madrid, España
- McCulloch, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4),115–133. <https://doi.org/10.1007/BF02478259>
- McCulloch, W. S., y Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4),115–133. <https://doi.org/10.1007/BF02478259>
- Menacho, C. (2017). Predicción del rendimiento académico aplicando técnicas de minería de datos. *Dialnet*, 78(1), 26-33. <https://dialnet.unirioja.es/servlet/articulo?codigo=6171237>
- Mobley. (1982). La satisfacción laboral y la rotación de personal en una mediana empresa del sector transportista
- Mueller, J., y Massaron, L. (2016). Machine learning for dummies. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Muñoz. (2016). *Análisis de calidad cartográfica mediante el estudio de la matriz de confusión*. Dialnet. <https://dialnet.unirioja.es/servlet/articulo?codigo=5998855>
- Navarro, G. (2022). *Agrupación de empleados en función de su desempeño y potencial. Propuestas de actuación frente a dichos grupos* [Tesis de pregrado, Universidad CUNEF]. https://www.researchgate.net/profile/Guillermo-Navarro-10/publication/361364194_Analitica_de_RRHH_Machine_Learning/links/62ac39ac23f3283e3ace884e/Analitica-de-RRHH-Machine-Learning.pdf
- Ocaña, F. (2018). Tratamiento estadístico de outliers y datos faltantes. *Técnicas estadísticas en Nutrición y Salud*. <https://www.ugr.es/~fmocan/MATERIALES%20DOCTORADO/Tratamiento%20de%20outliers%20y%20missing.pdf>

- Palacio, L. (2023). *Predicción de rotación de empleados usando modelos de aprendizaje automático* [Tesis de postgrado, Universidad EAFIT]. Escuela de Ingenierías, Maestría en Ciencias de los datos y la Analítica, Medellín.
https://repository.eafit.edu.co/bitstream/handle/10784/32615/LuisJavier_PalacioMesa_2023.pdf?sequence=4&isAllowed=y
- Parrales, E., y Rosales, M. (2023). *Utilización de Machine Learning para el proceso de selección de personal en una microempresa* [Tesis de pregrado, Universidad Politécnica Salesiana].
<https://dspace.ups.edu.ec/handle/123456789/23335>
- Pérez, B. (2018). Inteligencia artificial. INCyTU, 012, 1.
https://www.foroconsultivo.org.mx/INCYTU/documentos/Completa/INCYTU_18-012.pdf
- Portalo, A., y Portalo, A. (2022). La gran Renuncia (I): ¿Por qué la pandemia ha sacudido el mercado laboral? *Forbes España*. <https://forbes.es/economia/152055/la-gran-renuncia-eeuu/>
- Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.
<https://doi.org/10.1007/BF00116251>
- Ramírez, J. (2018). El machine learning a través de los tiempos, y los aportes a la humanidad. [Tesis de pregrado, Universidad Libre Seccional Pereira]. Universidad Libre Seccional Pereira.
<https://repository.unilibre.edu.co/bitstream/handle/10901/17289/EL%20MACHINE%20LEARNING.pdf?sequence=1&isAllowed=y>
- Rivero, F. (2022). Árbol de Decisión en Aprendizaje Automático. *Revista Varianza 19* (39-46).
http://revistasbolivianas.umsa.bo/pdf/rv/n19/n19_a05.pdf
- Rojas, E. (2018). *Glosario de los seis términos básicos del Machine Learning*.
<https://www.muycomputerpro.com/2018/02/07/glosario-terminos-basicos-machine-learning>.
- Rokach, L., & Maimon, O. (2008). Data mining with decision trees: theory and applications. *World Scientific*.

- Rouhiainen, L. (2018). *Inteligencia artificial: 101 cosas que debes saber hoy sobre nuestro futuro*. Alianza Editorial.
- Ruiz, G. (2019). *Modelo de análisis de datos utilizando técnicas de aprendizaje supervisado y no supervisado, para identificar patrones en la información generada por los pacientes, sometidos a juegos diseñados como un instrumento de apoyo terapéutico* [Trabajo de Grado, Universidad Jorge Tadeo Lozano]. Repositorio Universidad Jorge Tadeo Lozano. <https://expeditiorepositorio.utadeo.edu.co/bitstream/handle/20.500.12010/8502/Trabajo%20de%20grado.pdf?sequence=1&isAllowed=y>
- Sánchez, J. (2015). Análisis de Calidad Cartográfica mediante el estudio de la Matriz de Confusión. *Pensamiento Matemático*, 6(2), 9- 26.
- Sandoval, L. (2018). Algoritmos de Aprendizaje Automático para Análisis y Predicción de Datos. *Revista Tecnológica*, 11.
http://redicces.org.sv/jspui/bitstream/10972/3626/1/Art6_RT2018.pdf
- Shah, H., y Warwick, K. (2017). El futuro de la comunicación humano-máquina: el test de Turing. *OpenMind*. <https://www.bbvaopenmind.com/articulos/el-futuro-de-la-comunicacion-humano-maquina-el-test-de-turing/>
- Sierra, E. (2022). *Análisis y predicción de la deserción de empleados: un caso de estudio en la industria de software colombiana* [Tesis de postgrado]. Universidad EAFIT, Escuela de Ciencias, Maestría en Ciencias de los datos y la Analítica, Medellín.
<https://repository.eafit.edu.co/handle/10784/32155>
- Tejada, P. (2021). ¿Cómo identificar y filtrar outliers (valores atípicos) en Tableau?. *The Information Lab*. <https://theinformationlab.es/como-identificar-y-filtrar-outliers-valoresatipicos-en-tableau/>
- Vives, A., & Vara, E. (2011). *La Responsabilidad Social de la Empresa en América Latina*. Washington D.C.: Banco Interamericano de Desarrollo.

Zaballa, P. (2020). Desarrollo de un procedimiento para el análisis de la rotación del personal en la Empresa de Construcción y Montaje de Matanzas [Tesis de pregrado]. Universidad de Matanzas, Título de Ingeniero Industrial. <https://rein.umcc.cu/handle/123456789/1282>

Zaballa, P., Assafiri, Y., Medina, Y., Nogueira, D., & Medina, A. (2021). Procedimiento para el análisis de la rotación del personal. *Academo (Asunción)*, 8(1), 29-41. <https://doi.org/10.30545/academo.2021.ene-jun.3>

Anexos

Anexo 1.

Base de datos (Neta)

DIVISION	AREA	PUESTO	AÑO	MESES	GENERO	DISCAPACIDAD	EDAD	ROTACION TEMPRANA	MESES_ANT	CAUSAS_DE_SALIDA	MOTIVO	SUELDO FIJO	COMISIONES
1	1	1	2022	3	2	0	35	1	1	1	1	425	10
1	1	1	2022	2	2	0	36	0	0	1	1	425	30
1	1	1	2022	7	2	0	21	1	0	3	1	425	25
1	1	1	2022	5	2	0	27	0	4	3	1	425	35
1	1	1	2023	7	2	0	30	0	0	3	1	425	45
1	1	1	2023	5	1	0	28	0	0	1	1	425	75
1	1	1	2022	3	2	0	27	0	3	1	1	425	10
1	1	1	2023	9	2	0	22	0	7	3	1	425	100
1	1	1	2023	9	2	0	31	1	2	3	1	425	20
1	1	1	2023	4	2	0	31	0	1	3	1	425	30
1	1	1	2023	5	2	0	29	1	2	3	1	425	25
1	1	1	2022	1	2	0	24	1	1	2	1	425	35
2	1	1	2022	4	1	0	25	1	2	3	1	425	45
1	1	1	2023	9	2	0	28	0	46	1	1	425	75
1	1	1	2021	12	2	0	34	0	41	2	1	425	10
1	1	1	2023	8	1	0	23	0	1	3	1	425	100
1	1	1	2023	2	2	0	22	0	1	3	1	425	20
2	1	2	2022	8	2	0	40	0	145	3	1	500	300
2	1	1	2021	12	2	0	28	0	7	2	1	425	25
1	1	1	2023	9	1	0	26	1	1	3	1	425	30
1	1	2	2023	9	1	0	36	0	155	3	1	500	300

1	1	1	2022	12	2	0	27	1	0	3	1	425	40
1	1	1	2023	1	2	0	25	0	2	3	1	425	55
1	1	1	2023	7	2	0	22	1	1	1	1	425	75
2	1	1	2022	8	2	0	34	1	2	3	1	425	100
1	1	1	2023	9	2	0	32	0	53	3	1	425	20
1	1	1	2023	9	1	0	23	0	25	3	1	425	25
1	1	1	2023	9	1	0	44	1	1	1	1	425	30
2	1	1	2023	6	1	0	21	1	2	3	1	425	40
2	1	1	2021	2	1	0	21	0	5	3	1	425	55
1	1	1	2022	11	2	0	26	0	41	1	1	425	75
1	1	1	2021	9	2	0	28	0	9	1	1	425	100
1	1	1	2022	11	2	0	29	0	10	3	1	425	20
1	1	1	2022	11	2	0	31	0	49	3	1	425	25
2	2	1	2022	5	1	0	35	0	3	3	1	425	30
1	1	1	2022	3	2	0	27	0	55	1	1	425	40
2	1	1	2022	2	2	0	37	0	78	1	1	425	55
1	1	1	2021	1	1	0	22	0	15	1	1	425	75
1	1	2	2023	9	2	0	51	0	11	3	1	500	300
2	1	1	2021	11	1	0	34	0	56	2	1	425	100
1	1	1	2023	7	2	0	22	0	8	3	1	425	20
2	1	1	2023	6	2	0	35	0	3	3	1	425	25
1	1	1	2023	5	2	0	23	0	0	3	1	425	30
1	1	1	2021	10	2	0	22	1	2	1	1	425	40
1	1	1	2023	7	2	0	25	0	15	1	1	425	55
2	3	1	2021	8	2	0	26	1	0	1	1	425	75
2	1	1	2022	11	2	0	34	0	37	3	1	425	30
2	1	1	2023	9	1	0	40	1	2	3	1	425	25
1	1	1	2023	7	2	0	22	0	1	1	1	425	15
2	1	1	2021	5	2	0	29	1	2	1	1	425	20

Anexo 2.

Extracción de la base consolidada de Probabilidades en Rstudio

DIVISION	AREA	PUESTO	GENERO	DISCPACIDAD	EDAD	ROTACION.TE.MPRANA	MESES_ANT	CAUSAS_DE_SALIDA	SUELDOS Y COMISIONES	MOTIVO	Probabilidad.RV	Probabilidad.RN	Probabilidad.DI	Prediccion
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	35	Si	1	Clima laboral	435	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	36	No	0	Clima laboral	455	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	21	Si	0	Oportunidad de crecimiento	450	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	27	No	4	Oportunidad de crecimiento	460	RV	1.00	0.00	0.00	RV

Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	30	No	0	Oportunidad de crecimiento	470	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Masculino	No	28	No	0	Climalaboral	500	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	27	No	3	Climalaboral	435	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	22	No	7	Oportunidad de crecimiento	525	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	31	Si	2	Oportunidad de crecimiento	445	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	31	No	1	Oportunidad de crecimiento	455	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	29	Si	2	Oportunidad de crecimiento	450	RV	1.00	0.00	0.00	RV

Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	24	Si	1	Despido	460	RV	0.93	0.00	0.07	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Masculino	No	25	Si	2	Oportunidad de crecimiento	470	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	28	No	46	Clima laboral	500	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	34	No	41	Despido	435	RV	0.93	0.00	0.07	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Masculino	No	23	No	1	Oportunidad de crecimiento	525	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	22	No	1	Oportunidad de crecimiento	445	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Administrador	Femenino	No	40	No	145	Oportunidad de crecimiento	800	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Femenino	No	28	No	7	Despido	450	RV	0.92	0.00	0.08	RV

Cruz Azul	Ventas Farmacias Propias	Auxiliar	Masculino	No	26	Si	1	Oportunidad de crecimiento	455	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Administrador	Masculino	No	36	No	155	Oportunidad de crecimiento	800	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	27	Si	0	Oportunidad de crecimiento	465	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	25	No	2	Oportunidad de crecimiento	480	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	22	Si	1	Climat laboral	500	RV	1.00	0.01	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Femenino	No	34	Si	2	Oportunidad de crecimiento	525	RV	1.00	0.00	0.00	RV

Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	32	No	53	Oportunidad de crecimiento	445	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Masculino	No	23	No	25	Oportunidad de crecimiento	450	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Masculino	No	44	Si	1	Climas laboral	455	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Masculino	No	21	Si	2	Oportunidad de crecimiento	465	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Masculino	No	21	No	5	Oportunidad de crecimiento	480	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	26	No	41	Climas laboral	500	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	28	No	9	Climas laboral	525	RV	1.00	0.00	0.00	RV

Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	29	No	10	Oportunidad de crecimiento	445	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	31	No	49	Oportunidad de crecimiento	450	RV	1.00	0.00	0.00	RV
Pharmacys	Contac Center	Auxiliar	Masculino	No	35	No	3	Oportunidad de crecimiento	455	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	27	No	55	Clim laboral	465	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Femenino	No	37	No	78	Clim laboral	480	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Masculino	No	22	No	15	Clim laboral	500	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Administrador	Femenino	No	51	No	11	Oportunidad de crecimiento	800	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Masculino	No	34	No	56	Despido	525	RV	0.94	0.00	0.06	RV

Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	22	No	8	Oportunidad de crecimiento	445	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Femenino	No	35	No	3	Oportunidad de crecimiento	450	RV	0.99	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	23	No	0	Oportunidad de crecimiento	455	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	22	Si	2	Climat laboral	465	RV	0.83	0.17	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	25	No	15	Climat laboral	480	RV	1.00	0.00	0.00	RV
Pharmacys	Franquicia	Auxiliar	Femenino	No	26	Si	0	Climat laboral	500	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Femenino	No	34	No	37	Oportunidad de crecimiento	455	RV	1.00	0.00	0.00	RV

Pharmacys	Ventas Farmacias Propias	Auxiliar	Masculino	No	40	Si	2	Oportunidad de crecimiento	450	RV	1.00	0.00	0.00	RV
Cruz Azul	Ventas Farmacias Propias	Auxiliar	Femenino	No	22	No	1	Climat laboral	440	RV	1.00	0.00	0.00	RV
Pharmacys	Ventas Farmacias Propias	Auxiliar	Femenino	No	29	Si	2	Climat laboral	445	RV	1.00	0.00	0.00	RV



**Presidencia
de la República
del Ecuador**



**Plan Nacional
de Ciencia, Tecnología,
Innovación y Saberes**



DECLARACIÓN Y AUTORIZACIÓN

Nosotras, **Arroyo Becerra, Doris Nayeli**, con C.C: # **0954708244** y **Ríos Arambulo, Sara María**, con C.C: # **0951747856** autoras del trabajo de titulación: **Análisis de factores predominantes en la rotación temprana de empleados de grupo Difare de la región Sierra aplicando Machine Learning**, previo a la obtención del título de **Licenciada en Negocios Internacionales** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, 15 de febrero del 2024

Doris Arroyo

Arroyo Becerra, Doris Nayeli
0954708244

Sara Ríos

Ríos Arambulo, Sara María
0951747856

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TEMA Y SUBTEMA:	Análisis de factores predominantes en la rotación temprana de empleados de grupo Difare de la región Sierra aplicando Machine Learning.		
AUTOR(ES)	Arroyo Becerra, Doris Nayeli y Rios Arambulo, Sara Maria		
REVISOR(ES)/TUTOR(ES)	Ing. Félix Miguel, Carrera Buri, Mgs.		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Facultad de Economía y Empresas		
CARRERA:	Negocios Internacionales		
TÍTULO OBTENIDO:	Licenciadas en Negocios Internacionales		
FECHA DE PUBLICACIÓN:	15 de febrero del 2024	No. DE PÁGINAS:	81
ÁREAS TEMÁTICAS:	Análisis de datos y modelado predictivo, Psicología Organizacional y Comportamiento del Empleado, Gestión de Recursos Humanos y Estrategias de Retención.		
PALABRAS CLAVES/KEYWORDS:	Palabras claves: Inteligencia Artificial, Machine Learning, Random Forest, Rotacion Temprana de Personal. / Keywords: Artificial Intelligence, Machine Learning, Random Forest, Early Staff Rotation.		
RESUMEN/ABSTRACT (150-250 palabras):			
<p>Resumen: La Inteligencia Artificial (IA) se centra en el desarrollo de sistemas que puedan realizar tareas que normalmente requieren inteligencia humana, incluidas actividades como el aprendizaje, la resolución de problemas, el reconocimiento de patrones y la toma de decisiones. La Inteligencia Artificial tiene una variedad de aplicaciones, desde asistentes virtuales y chatbots hasta algoritmos de recomendación. Su continuo desarrollo ha tenido un impacto significativo en áreas como la salud, la industria y la investigación, cambiando la forma en que interactuamos con la tecnología y creando nuevas oportunidades en una variedad de campos.</p> <p>El presente trabajo de investigación tuvo como objetivo principal elaborar un modelo de Machine Learning para clasificar los factores predominantes que influyen en la rotación temprana de empleados de Grupo Difare, de la región Andina del Ecuador, lo que a su vez permita tomar decisiones estratégicas acerca del pronóstico del comportamiento a futuro. Dicho análisis fue llevado mediante la implementación del algoritmo Random Forest o Bosque Aleatorio, en el software estadístico RStudio. Lo que permitió evaluar las causas principales de las renuncias voluntarias del personal de talento de la entidad comercial, y por consiguiente, brindar recomendaciones para prevenir la tendencia creciente del fenómeno de rotación de personal, el mismo que tiene un impacto en la productividad y dinámica organizacional.</p> <p>Abstract: Artificial Intelligence (AI) focuses on the development of systems that can perform tasks that normally require human intelligence, including activities such as learning, problem solving, pattern recognition and decision making. Artificial Intelligence has a variety of applications, from virtual assistants and chatbots to recommendation algorithms. Its continuous development has had a significant impact in areas such as healthcare, industry and research, changing the way we interact with technology and creating new opportunities in a variety of fields. The main objective of this research work was to develop a Machine Learning model to classify the predominant factors that influence the early turnover of employees of Difare Group, in the Andean region of Ecuador, which in turn allows making strategic decisions about the forecast of future behavior. This analysis was carried out through the implementation of the Random Forest algorithm in the statistical software RStudio. This allowed us to evaluate the main causes of voluntary resignations of talented personnel of the commercial entity, and consequently, to provide recommendations to prevent the growing trend of the phenomenon of staff turnover, which has an impact on productivity and organizational dynamics.</p>			
ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	
CONTACTO CON AUTOR/ES:	Teléfono: +593 99 014 0456	E-mail: doris.arroyo1213@gmail.com	
CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UTE)::	Nombre: Freire Quintero, Cesar Enrique		
	Teléfono: +593-4-90090702		
	E-mail: cesar.freire@cu.ucsg.edu.ec		
SECCIÓN PARA USO DE BIBLIOTECA			
Nº. DE REGISTRO (en base a datos):			
Nº. DE CLASIFICACIÓN:			
DIRECCIÓN URL (tesis en la web):			