



**FACULTAD DE POSTGRADO
TESIS DE POSTGRADO**

**IMPLEMENTACIÓN DE SISTEMA DE DATOS PARA ANÁLISIS
PREDICTIVO DE ROTACIÓN DE PERSONAL EN LOS PINOS
APPAREL S.A.**

SUSTENTADO POR:

**KAREN YAMILETH CRUZ MATA
STEPHANY GABRIELA OSORTO FLORES**

PREVIA INVESTIDURA AL TÍTULO DE

**MÁSTER EN
ANALÍTICA DE NEGOCIOS**

TEGUCIGALPA, FRANCISCO MORAZAN, HONDURAS, C.A.

FEBRERO, 2024

**UNIVERSIDAD TECNOLÓGICA CENTROAMERICANA
UNITEC**

FACULTAD DE POSTGRADO

AUTORIDADES UNIVERSITARIAS

PRESIDENTE EJECUTIVO /

RECTORA

ROSALPINA RODRÍGUEZ

SECRETARIO GENERAL /

PRORRECTOR

ROGER MARTÍNEZ MIRALDA

VICERRECTOR ACADÉMICO NACIONAL

JAVIER ABRAHAM SALGADO LEZAMA

DIRECTORA NACIONAL DE POSTGRADO

ANA DEL CARMEN RETTALLY

**IMPLEMENTACIÓN DE SISTEMA DE DATOS PARA
ANÁLISIS PREDICTIVO DE ROTACIÓN DE PERSONAL
EN LOS PINOS APPAREL S.A.**

**TRABAJO PRESENTADO EN CUMPLIMIENTO DE LOS
REQUISITOS EXIGIDOS PARA OPTAR AL TÍTULO DE**

MÁSTER EN

ANALÍTICA DE NEGOCIOS

ASESORA METODOLOGICA

ALBA GABRIELA GARAY ROMERO

MIEMBROS DE LA TERNA:

**KEVIN EDUARDO FUNEZ FUNEZ
DANIEL ANTONIO LUNA RODRIGUEZ
ERICK LEONARDO IZAGUIRRE ARDON**

DERECHOS DE AUTOR

© Copyright 2024
Karen Yamileth Cruz Mata
Stephany Gabriela Osorto Flores

Todos los derechos son reservados.



FACULTAD DE POSTGRADO

IMPLEMENTACIÓN DE SISTEMA DE DATOS PARA ANÁLISIS PREDICTIVO DE ROTACIÓN DE PERSONAL EN LOS PINOS APPAREL S.A.

**Karen Yamileth Cruz Mata
Stephany Gabriela Osorto
Flores**

Resumen

En cualquier organización, el activo máspreciado indudablemente es su capital humano, ya que una fuerza laboral bien capacitada se convierte en el pilar esencial para alcanzar los objetivos anuales. Entre los diversos indicadores críticos para el departamento de recursos humanos, la tasa de rotación destaca como un elemento crucial debido a su impacto directo en la producción de la empresa. El propósito fundamental de la presente investigación es analizar en detalle el comportamiento de la rotación de personal en la empresa, con el objetivo de desarrollar perfiles con características similares que permitan prever si un colaborador(a) tiene la probabilidad de abandonar la organización y así poder diseñar estrategias que prolonguen su estadía en la organización o que permitan rediseñar estrategias de reclutamiento. Este estudio se llevó a cabo utilizando la base de datos proporcionada generosamente por la empresa Los Pinos Apparel. El proceso metodológico incluyó una exhaustiva limpieza de datos y un análisis exploratorio detallado para garantizar la integridad y fiabilidad de la información. Posteriormente, se crearon cuatro modelos de predicción, siendo la regresión logística el modelo con mayor exactitud, alcanzando un impresionante 97.10%

Palabras claves: (Modelos, Predicción, Recurso Humano, Rotación)



GRADUATE SCHOOL

IMPLEMENTATION OF A DATA SYSTEM FOR PREDICTIVE ANALYSIS OF EMPLOYEE TURNOVER AT LOS PINOS APPAREL S.A.

**Karen Yamileth Cruz Mata
Stephany Gabriela Osorto
Flores**

Abstract

In any organization, the most valuable asset is undoubtedly its human capital, since a well-trained workforce becomes the essential pillar to achieve annual objectives. Among the various critical indicators for the human resources department, the turnover rate stands out as a crucial element due to its direct impact on the company's production. The fundamental of this research is to analyze in detail the purpose of personnel rotation in the company, with the objective of developing profiles with similar characteristics that allow us to predict whether an employee is likely to leave the organization and thus be able to. design strategies that prolong their stay in the organization or that allow redesigning recruitment strategies. This study was carried out using the database generously provided by the company Los Pinos Apparel. The methodological process included extensive data cleaning and detailed exploratory analysis to ensure the integrity and reliability of the information. Subsequently, four prediction models were created, with logistic regression being the model with the highest accuracy, reaching an impressive 97.10%.

Keywords: (Models, Human Resource, Prediction, Rotation)

DEDICATORIA

Amado señor Jesús, gracias por acompañarme en estos dos años de post-grado, por estar conmigo en las buenas y malas situaciones, este triunfo es para ti y tu madre la virgen de Suyapa que nunca me desampara. También lo dedico a mis amados padres, gracias por guiar mis pasos y nunca soltar mi mano, a mi familia por su apoyo incondicional y también a mis ángeles en el cielo que se desde arriba están viendo este gran logro.

“Gracias a todos, por siempre estar”

STEPHANY GABRIELA OSORTO FLORES

A mi amado Padre celestial que nunca me suelta de su mano y siempre me bendice grandemente. También dedico este logro profesional a mi familia, especialmente a mi madre que me heredó su carácter persistente para nunca darme por vencida y a mi hermana Ana que me enseñó desde pequeña, la pasión de cultivar mi mente con la lectura.

“Gracias a todos por su paciencia y amor”

KAREN YAMILETH CRUZ MATA

AGRADECIMIENTO

Queremos expresar nuestro agradecimiento inicialmente a Dios por permitirnos alcanzar este logro académico, una meta que hemos alcanzado con dedicación y esfuerzo sostenido. También extendemos nuestro agradecimiento a nuestras familias, quienes desde que iniciamos este largo camino nos apoyaron y estuvieron ahí hasta el final.

Extendemos nuestro agradecimiento a todos los profesores y a la Universidad Tecnológica Centroamericana que compartieron su conocimiento con nosotros, contribuyendo así a enriquecer nuestra base de saberes.

Y a cada una de las personas que nos brindó apoyo a lo largo de esta etapa, les expresamos nuestro agradecimiento y pedimos a Dios que les bendiga abundantemente, permitiéndonos celebrar juntos este importante logro.

ÍNDICE DE CONTENIDO

DEDICATORIA	vii
AGRADECIMIENTO	viii
ÍNDICE DE CONTENIDO	ix
ÍNDICE DE TABLAS	xi
ÍNDICE DE FIGURAS.....	xii
CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN	1
1.1 INTRODUCCIÓN	1
1.2 ANTECEDENTES DEL PROBLEMA	1
1.3 DEFINICIÓN DEL PROBLEMA	3
1.3.1 Enunciado del Problema.....	3
1.3.2 Formulación del Problema	3
1.3.3 Preguntas de Investigación.....	3
1.4 OBJETIVOS DEL PROYECTO.....	4
1.4.1 Objetivo General	4
1.4.2 Objetivos Específicos	4
1.5 JUSTIFICACIÓN.....	4
CAPÍTULO II. MARCO TEÓRICO	6
2.1 ANÁLISIS DE LA SITUACIÓN ACTUAL.	6
2.2 CONCEPTUALIZACIÓN.....	11
2.3 TEORÍAS DE SUSTENTO	13
2.3.1 BASES TEÓRICAS	13
2.3.2 METODOLOGÍAS DESARROLLADAS.....	18
2.3.3 INSTRUMENTOS UTILIZADOS.....	20
2.4 MARCO LEGAL	22
CAPÍTULO III. METODOLOGÍA	29
3.1 CONGRUENCIA METODOLÓGICA.....	29
3.1.1 MATRIZ METODOLÓGICA.....	29
3.1.2 ESQUEMA DE VARIABLES DE ESTUDIO.....	31
3.1.3 OPERACIONALIZACIÓN DE LAS VARIABLES	31
3.1.4 HIPÓTESIS	35

3.2	ENFOQUE Y MÉTODOS	35
3.2.1	ENFOQUE.....	35
3.2.2	DISEÑO.....	35
3.2.3	ALCANCE	36
3.3	DISEÑO DE LA INVESTIGACIÓN	36
3.3.1	POBLACIÓN	37
3.3.2	MUESTRA.....	37
3.3.3	TÉCNICAS DE MUESTREO.....	37
3.4	TÉCNICAS, INSTRUMENTOS Y PROCEDIMIENTOS APLICADOS	37
3.4.1	TÉCNICAS.....	37
3.4.2	INSTRUMENTOS	38
3.5	FUENTES DE INFORMACIÓN.....	38
3.5.1	FUENTES PRIMARIAS	38
3.5.2	FUENTES SECUNDARIAS.....	38
CAPÍTULO IV. RESULTADOS Y ANÁLISIS		39
4.1	INFORME DE PROCESO DE RECOLECCIÓN DE DATOS	39
4.1.1	CRITERIOS DE EXCLUSION DE VARIABLES.....	39
4.2	RESULTADOS Y ANÁLISIS DE LAS TÉCNICAS APLICADAS	42
4.2.1	ANÁLISIS CUANTITATIVO.....	42
4.2.2	ANÁLISIS CUALITATIVO	50
4.3	ANÁLISIS INFERENCIAL Y MODELOS APLICADOS.....	52
CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES.....		60
5.1	CONCLUSIONES	60
5.2	RECOMENDACIONES	62
CAPÍTULO VI. APLICABILIDAD.....		64
6.1	NOMBRE DE LA PROPUESTA	64
6.2	JUSTIFICACIÓN DE LA PROPUESTA	64
6.3	ALCANCE DE LA PROPUESTA	64
6.3.1	OBJETIVO GENERAL	64
6.3.2	OBJETIVOS ESPECIFICOS	65
6.4	DESCRIPCIÓN Y DESARROLLO	65

6.4.1	DESCRIPCIÓN	65
6.4.2	DESARROLLO.....	66
6.4.2.1	HERRAMIENTAS	66
6.5	MEDIDAS DE CONTROL	73
6.6	CRONOGRAMA DE IMPLEMENTACIÓN Y PRESUPUESTO	75
6.6.1	CRONOGRAMA DE ACTIVIDADES.....	75
6.6.2	PRESUPUESTO	76
6.7	CONCORDANCIA DE LOS SEGMENTOS DE LA TESIS CON LA PROPUESTA	77
	REFERENCIAS BIBLIOGRÁFICAS.....	82
	ANEXOS	84
	Anexo 1	84

ÍNDICE DE TABLAS

Tabla 1	Matriz Metodológica.....	30
Tabla 2	Operacionalización de las Variables.....	32
Tabla 3	Tabla de Variables de Base de Datos.....	40
Tabla 4	Resultados de Modelos de Predicción	59
Tabla 5	Nodos Limpieza de Datos.....	69
Tabla 6	Nodos Análisis Exploratorio de Datos.....	70
Tabla 7	Nodos de Entrenamiento Modelos de Predicción.....	73
Tabla 8	Cronograma de Actividades.....	75
Tabla 9	Capacitación de Knime impartidas por IQuartil	76
Tabla 10	Capacitación de Udemty Machine Learning con Knime	77
Tabla 11	Presupuesto de Proyecto	77
Tabla 12	Concordancia de los segmentos de la Tesis con la Propuesta Parte I.....	78
Tabla 13	Concordancia de los segmentos de la Tesis con la Propuesta Parte II	79

ÍNDICE DE FIGURAS

Ilustración 1 Funcionamiento de Machine Learning	16
Ilustración 2 Esquema de Variables de Estudio.....	31
Ilustración 3 Enfoque Metodológico	36
Ilustración 4 Total de Colaboradores	42
Ilustración 5 Colaboradores por Rango de Edad	43
Ilustración 6 Colaboradores por Estado Civil.....	44
Ilustración 7 Colaboradores por Cantidad de Hijos	45
Ilustración 8 Colaboradores por Antigüedad	46
Ilustración 9 Colaboradores por Sección	47
Ilustración 10 Comportamiento de Salidas	48
Ilustración 11 Salidas por Sección.....	49
Ilustración 12 Salidas por Rango de Hijos.....	49
Ilustración 13 Salidas por Antigüedad.....	50
Ilustración 14 Regresión Logística – Variables Categóricas	54
Ilustración 15 Matriz de Regresión Logística – Variables Categóricas.....	54
Ilustración 16 Regresión Logística – Variables Binarias.....	54
Ilustración 17 Matriz de Regresión Logística – Variables Binarias	55
Ilustración 18 Naive Bayes – Variables Categóricas.....	55
Ilustración 19 Matriz de Naive Bayes – Variables Categóricas	55
Ilustración 20 Árbol de Decisión – Variables Categóricas.....	56
Ilustración 21 Matriz de Árbol de Decisión – Variables Categóricas	56
Ilustración 22 Árbol de Decisión – Variables Binarias	56
Ilustración 23 Matriz de Árbol de Decisión – Variables Binarias	57
Ilustración 24 Random Forest – Variables Categóricas.....	57
Ilustración 25 Matriz de Random Forest – Variables Categóricas	57
Ilustración 26 Random Forest – Variables Binarias	58
Ilustración 27 Matriz de Random Forest – Variables Binarias.....	58
Ilustración 28 Limpieza de Datos	69
Ilustración 29 Análisis Exploratorio de los Datos	70
Ilustración 30 Flujo Modelo de Predicción Regresión Logística.....	71

Ilustración 31 Flujo Modelo de Predicción Naive Bayes	71
Ilustración 32 Flujo Modelo de Predicción Árbol de Decisión	72
Ilustración 33 Flujo Modelo de Predicción Random Forest	72
Ilustración 34 Matriz de Confusión	74
Ilustración 35 Adquisición de Laptop.....	76

CAPÍTULO I. PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1 INTRODUCCIÓN

El presente trabajo se enfoca en el desafío de la alta rotación de empleados, un problema que influye significativamente en la eficiencia económica y productiva de las compañías. Las organizaciones deben afrontar los gastos relacionados con la contratación y capacitación de nuevos trabajadores, y a medida que la rotación de personal aumenta, estos costos se elevan. Además, la alta rotación de empleados perturba la continuidad de los diversos procesos empresariales, lo que resulta en una reducción de la productividad.

El área de Recursos humanos en Los Pinos Apparel, S.A. tiene el propósito de sistematizar la organización de documentos y datos del personal activo que trabaja en la empresa, así como generar la visibilidad de la información y los datos en tiempo y forma para facilitar la comprensión y la generación de conocimiento para la toma de decisiones y generación de nuevas estrategias que permitan cumplir con los objetivos corporativos. Al digitalizar los datos, estos se utilizarán para un modelo de predicción el cual nos ayudara a identificar la rotación y sus posibles factores.

1.2 ANTECEDENTES DEL PROBLEMA

En la Honduras actual, las fuentes de empleo más importantes están en el rubro de la manufactura, ya que permite emplear masivamente, mano de obra semi calificada para procesamiento de diferentes tipos de materias primas. Al manipular materias primas uno de los indicadores más importantes es la eficiencia de la mano de obra. En la actualidad las empresas tienen como objetivo principal el de reducir los costos innecesarios y aumentar la competitividad.

La eficiencia productiva es una métrica clave para todas las empresas, ya que permite medir el rendimiento de los procesos y, en consecuencia, la capacidad de generar ingresos.

“La eficiencia productiva en recursos humanos se refiere a la capacidad de una organización para gestionar y desarrollar su fuerza laboral de manera efectiva, maximizando la productividad y contribuyendo al logro de los objetivos empresariales.”. (Dessler & Varela, 2011)

Considerando el concepto anterior se podría decir que el recurso más importante en Los Pinos Apparel S.A, es el recurso humano. En una ciudad poco industrializadas como Siguatepeque, la mano de obra no calificada abunda, pero requiere un nivel de inversión en horas de capacitación para que se logre llegar a un nivel óptimo de producción y cumplir con el objetivo establecido de

eficiencia. Sin embargo, pese a la escasez de oportunidades de empleo en la zona se ha observado un fenómeno de abandono y rotación laboral.

- Según Chiavenato, la rotación de personal consiste en:

“La expresión rotación de personal se utiliza para definir la fluctuación de personal entre una organización y su ambiente; en otras palabras, el intercambio de personas entre la organización y el ambiente está definido por el volumen de personas que ingresan y que salen de la organización” (Chiavenato, Subsistema de integración de recursos humanos, 2007).

El índice de rotación de personal en condiciones controladas, es saludable sin embargo cuando es por iniciativa del empleado y se pierde mano de obra entrenada y con experiencia, puede provocar un descenso en la eficiencia en los procesos productivos, el personal nuevo que ingresa debe ser entrenado y esperar que se cumpla la curva de aprendizaje de no menos de dos meses, para lograr llevar al nuevo empleado al nivel de habilidades y competencias laborales apropiadas para cubrir la baja causada por la rotación.

- Chiavenato opina con respecto a la importancia de retener al personal para cumplir con los objetivos estratégicos:

Los desafíos estratégicos de las organizaciones deben ocupar un lugar importante y fundamental en la administración de personal. Hoy, una condición básica para que los gerentes de línea puedan tener éxito en sus actividades es que las personas, su desempeño y sus competencias coincidan con las estrategias del negocio y las metas de la organización. En ese contexto, las relaciones entre la organización y las personas que trabajan en ella adquieren una importancia vital (Chiavenato, Relaciones con los Empleados, 2008)

Un hecho importante en las empresas hoy en día es contar con la visibilidad de los datos en tiempo y forma, que permitan tener una perspectiva más clara y definida para la toma de decisiones y generación de conocimiento a partir de la comprensión de los datos históricos de la empresa. Actualmente, se puede tener acceso a diferentes algoritmos inteligentes de aprendizaje automático supervisados o no y que permiten cuantificar los datos de la empresa, aplicar métodos predictivos y generar conocimiento en base a la información de los datos de la empresa.

1.3 DEFINICIÓN DEL PROBLEMA

En toda empresa los altos índices de rotación afectan la productividad e ingresos de la empresa, lo que se busca en toda empresa es que este porcentaje de rotación sea lo más bajo posible, con esto se asegura una mano de obra calificada, a continuación, se detalla el enunciado del problema, así como la pregunta que se responderá con esta investigación.

1.3.1 Enunciado del Problema

Los Pinos Apparel S.A se dedica a fabricar uniformes para exportación y uso de oficiales de policía y bomberos. Tiene una planilla activa de 559 empleados directos del área de producción. Actualmente reporta una rotación de personal en su sección de manufactura, esto genera que la mano de obra pierda antigüedad y cada vez sean más nuevas las personas que se dedican a fabricación de uniformes, si esto sigue así, podría salirse de control y dificultar el cumplimiento de la producción y fechas de entrega, generar un incremento de horas extras e incremento en los costos financieros de selección, contratación y de entrenamiento. Actualmente Los Pinos Apparel, maneja sus datos de empleados en expedientes impresos, y calcula sus planillas en Excel. Lo anterior implica que no se tiene consolidada la información de los empleados para generar estadísticas o indicadores, por lo tanto, datos como la rotación del personal son impredecibles.

1.3.2 Formulación del Problema

Dado que esta rotación encarece la mano de obra y compromete la productividad y sobre todo los limita para mantenerse competitivos en un mercado demandante. Este enunciado tiene como finalidad orientar las respuestas que se buscan en toda la investigación, dando origen a la siguiente pregunta

¿Es posible entrenar un modelo de predicción para estimar la probabilidad de que un colaborador renuncie, utilizando datos históricos sobre el personal y su historial de rotación almacenados en las bases de datos de la empresa?

1.3.3 Preguntas de Investigación

A continuación, detallamos las siguientes preguntas de investigación que se responderán con esta investigación:

¿Cuáles son las causas que generan la rotación en los Pinos Apparel S.A.?

¿Existirán características o variables comunes en los perfiles de los empleados que renuncian a la empresa?

¿Existen técnicas de machine learning y analítica de negocios que pueden implementarse para la generación de un perfil del personal candidato a dejar la organización en un corto plazo?

1.4 OBJETIVOS DEL PROYECTO

Una vez planteadas las preguntas de investigación, se ha logrado plantear el objetivo principal de la investigación, así como los objetivos específicos los cuales se detallan a continuación.

1.4.1 Objetivo General

A continuación, se detalla el objetivo general de la investigación:

Analizar si posible entrenar un modelo de predicción para estimar la probabilidad de que un colaborador renuncie, utilizando datos históricos sobre el personal y su historial de rotación almacenados en las bases de datos de la empresa.

1.4.2 Objetivos Específicos

A continuación, se detallan los objetivos específicos de la investigación:

Identificar las causas que generan la rotación en los Pinos Apparel S.A.

Definir cuáles son las características o variables comunes en los perfiles de los empleados que renuncian a la empresa.

Analizar si existen técnicas de machine learning y analítica de negocios que pueden implementarse para la generación de un perfil del personal candidato a dejar la organización en un corto plazo.

1.5 JUSTIFICACIÓN

La rotación de personal implica costos significativos, como los relacionados con la contratación y capacitación de nuevos empleados. Comprender y reducir esta rotación puede resultar en ahorros económicos sustanciales para la empresa. También permitirá conocer los factores que impulsan a los empleados a quedarse o a irse, lo que es crucial para mantener a los profesionales valiosos en la organización.

Comprender la rotación del personal también ayuda a adaptarse a las tendencias cambiantes del mercado laboral, lo que es esencial para competir con éxito por el talento y mantener una ventaja competitiva.

En resumen, investigar la rotación del personal es esencial para abordar los desafíos económicos, operativos y de gestión de talento que las empresas enfrentan en un entorno empresarial altamente competitivo. Además, puede ayudar a mejorar la retención de empleados, fortalecer la cultura organizacional y mantenerse al día con las tendencias laborales, lo que es fundamental para el éxito a largo plazo de la empresa.

CAPÍTULO II. MARCO TEÓRICO

2.1 ANÁLISIS DE LA SITUACIÓN ACTUAL.

2.1.1 LA ROTACION DE PERSONAL EN PERSPECTIVA CONCEPTUAL

Cuando hablamos de rotación nos referimos a la salida voluntaria de un colaborador de la empresa, esta rotación puede ser salir de la empresa o realizar un cambio voluntario de una maquila a otra.

En un acercamiento preliminar, se puede afirmar 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. Es decir, la salida y la entrada laboral por parte del personal de una empresa a otra. Partiendo de la premisa que la mayoría de los trabajadores en cualquier parte del mundo intentarán marcharse de sus trabajos y buscarán nuevas oportunidades de empleo, porque consideran que un nuevo trabajo les dará mayores beneficios, mayor seguridad laboral, nuevas experiencias o simplemente les permitirá evitar los conflictos en el actual empleo, la rotación de personal es inevitable. Sin embargo, desde la empresa, una alta tasa de rotación de personal es indeseable por la carencia de personal habilidoso o entrenado, por la pérdida de productividad, por la pérdida de negocios para competir o los altos costos que se tienen que pagar en el reclutamiento y el entrenamiento. Estos efectos negativos y la importancia del tema han permitido el desarrollo y la evolución de la literatura, donde encontramos teorías, tipos y modelos de rotación de personal en las empresas. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

La rotación de personal como objeto de estudio se ha abordado desde diferentes disciplinas ya sea desde la psicología, la economía o la sociología. La explicación de este fenómeno obtiene diversas respuestas dependiendo de su perspectiva. Desde el punto de vista de la economía se ha subrayado, por ejemplo, su costo, su relación con la oferta y la demanda del empleo, la inflación, el crecimiento de la economía. Desde la psicología se busca la explicación de la rotación de personal voluntaria en la falta de motivación y en la insatisfacción del personal, se subrayan necesidades, actitudes, o percepciones, dentro de la organización; desde el punto de vista social, se trata de identificar el perfil sociodemográfico y los problemas de carácter social que influyen en la rotación de personal. Estos enfoques han subrayado diversos aspectos de la rotación de personal y la importancia de disminuir su impacto negativo por lo que se pueden encontrar diversas

definiciones, todas ellas señalan el compromiso laboral con una empresa y el inicio en otra tratando de identificar sus causas. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

Cálculo del índice de rotación

Para poder medir la rotación de personal y que al área de recursos humanos pueda realizar una planeación sobre las estrategias a utilizar, suelen trabajar con índices mensuales o anuales, permitiéndoles comparar y analizar los resultados para que puedan tomar todas las acciones correspondientes en el desarrollo de estrategias que puedan mitigar este indicador.

Para términos de esta investigación y a sugerencia de la empresa manufacturera es utilizada la siguiente formula:

$$\left(\text{Índice de rotación de Empleados} = \frac{D \text{ (Empleados Separados en el Periodo)}}{\text{Promedio de Empleados en el Periodo}} \times 100 \right)$$

Siendo D = separaciones de personal (voluntaria e involuntaria) en el periodo considerado

PE = Personal empleado promedio en el periodo

Esta fórmula se utiliza “cuando se analizan las pérdidas de personas y sus causas” (Chiavenato, 2007, p. 137).

Todas las organizaciones hoy en día presentan porcentajes de rotación de personal, las que pueden resultar positivas si de estas salen las personas correctas ya que como lo menciona (Robbins, 1996, p. 27), la rotación puede crear oportunidad de sustituir a un individuo con un desempeño pobre por alguien que tenga mayores habilidades o motivación, abrir mayores oportunidades de ascenso, e incorporar ideas nuevas y recientes a la organización.

2.1.2 FACTORES QUE INFLUYEN O DETERMINAN LA ROTACION DEL PERSONAL

2.1.2.1 FACTORES DESDE LA PERSPECTIVA ECONOMISTA

El primer modelo completo de la rotación de personal fue presentado por Slichter en 1919 quien afirmaba que la rotación es el producto de muchas variables tanto a nivel de economía, de empresa y factores a nivel empleado. Son ocho las causas que identifica en la rotación de personal:

1) Reducción en la cantidad de trabajo que puede ser por la depresión industrial hasta la discontinuidad en el negocio o la quiebra;

2) Causas atribuibles al trabajo como puede ser salarios, turnos o falta de oportunidad para desarrollarse;

3) Métodos para dirigir al personal donde destaca pagos de salarios en base a méritos, promociones sistemáticas basadas en méritos, método adecuado para el manejo de quejas;

4) Causas de relaciones entre los trabajadores como relaciones desagradables entre los trabajadores o renuncia para acompañar una amistad que deja la empresa;

5) Causas relacionadas con el trabajador como puede ser deseo de vacacionar o de cambiar de empleo, problemas de salud o casamiento;

6) Oportunidades externas más atractivas que pueden ser común en épocas de prosperidad o intento para mejorar la condición del trabajador;

7) Las condiciones de vida en la localidad por ejemplo que no les gusta el clima;

8) Cambios debido a la familia del trabajador como seria enfermedad de la familia o que no les gusta la ciudad. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

2.1.2.2 FACTORES DESDE LA PERSPECTIVA PSICOLOGICA Y SOCIOLOGICA

Inicialmente los modelos económicos eran predominantes y llegaron a tener un enfoque demasiado estrecho que no permitían explicar adecuadamente el fenómeno de la rotación; lo que permite la aseveración que los factores económicos son necesarios, pero no suficientes para explicar la rotación de personal. Partiendo de esta insuficiencia en la explicación de la rotación de personal se construyeron otros modelos que incluyeron factores no monetarios, estos modelos fueron propuestos por psicólogos y sociólogos. En estos modelos se ha sugerido que la satisfacción en el trabajo juega un papel importante en el proceso de rotación lo mismo que el compromiso con la organización sobre todo en la intención de rotar. Por ejemplo, Luu & Hatstrup en 2010 establecieron la correlación significativa de la satisfacción en el trabajo con la intención de rotar, en los cuatro países que estudiaron; los resultados fueron que la correlación fue más fuerte en Francia, seguida por Estados Unidos, Japón y Filipina respectivamente. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

Otro elemento que ha sido estudiado como explicación de la rotación de personal, desde la perspectiva de la psicología, es la del arraigo. Cada vez es más importante para las organizaciones reconocer el arraigo de los empleados como un verdadero recurso a largo plazo para las

organizaciones. Considerando que el arraigo puede ayudar a reducir la rotación voluntaria y retener a los empleados valiosos que hacen que la organización tenga un alto rendimiento. Por su parte, Price en 2004 elaboró su primer modelo causal sobre la rotación que se publicó en 1975, “este modelo inicial tenía cuatro variables exógenas: pago, grupo primario, comunicación y centralización. El pago es un determinante significativo cuando se cumple alguna de las siguientes condiciones: Primero, si el pago es importante para los empleados y segundo si se percibe el pago como alto. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

El segundo determinante se refiere a la participación del empleado en grupos primarios con los que tiene afinidad. El tercer determinante se refiere a la comunicación como la transmisión de información entre los miembros de la organización. El cuarto determinante es la centralización y se refiere a la distribución del poder dentro de la organización. La rotación de una organización, según su hipótesis, es baja si el pago, el grupo primario y la comunicación son altos, mientras que un alto nivel de centralización resultaba en una alta rotación. El modelo tiene dos variables que intervienen, la satisfacción en el empleo y la oportunidad. Las variables demográficas (nivel educativo, rango militar, edad y antigüedad dentro de las fuerzas armadas) se incluyeron dentro de los análisis, pero no fueron incluidas dentro del modelo. Price en 2004 consideró que, si las variables demográficas arrojaban resultados significativos, eso indicaría que el modelo no funcionaba bien. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

Debido a que la rotación es alta en los nuevos empleados en una organización, se ha sugerido, por ejemplo, y se han desarrollado tácticas de socialización por parte de la organización para los nuevos empleados que puedan influir en la inserción del nuevo trabajador y evitar la rotación. Las tácticas de socialización tratan de ayudar a los nuevos trabajadores a una pronta adaptación y a reducir la incertidumbre y la ansiedad propia, que se genera al unirse a una nueva organización y al mismo tiempo adquirir las actitudes, los comportamientos y los conocimientos necesarios. Por su parte, Chen et al en 2011 considera que a pesar de que las investigaciones sobre la rotación de personal afirman que la satisfacción en el trabajo juega un papel importante, -incluso, aunque varias teorías sugieren relaciones longitudinales entre la satisfacción laboral y las intenciones de rotación, no han explicado suficientemente por qué, cómo y cuándo los cambios en la satisfacción del trabajo ejercen influencias únicas sobre las decisiones de rotar. Asimismo, considera que la mayoría de las investigaciones han considerado la satisfacción en el trabajo y las intenciones de rotación, de manera estática y no han considerado su aspecto dinámico. Por lo que

el hallazgo principal de la investigación de Chen et al en 2011 es resaltar la visión consistente que se obtiene al centrarse en la dinámica del cambio de satisfacción en el trabajo y no en los niveles estáticos de satisfacción en el trabajo. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

2.1.3 COSTO DE LA ROTACION DEL PERSONAL

Uno de los temas más complicados de cuantificar para las áreas de Recursos Humanos en las empresas es sin duda el costo que implica reemplazar a un colaborador que ha sido desvinculado o que renunció voluntariamente. Los argumentos que se han utilizado han tenido relación con criterios de índole subjetiva más que objetiva. Esta quizás es la crítica que más se escucha hacia muchos departamentos de Recursos Humanos en las empresas: la incapacidad para demostrar el impacto que tiene la rotación de personal en términos económicos. (Estrada, 2018)

Los mecanismos para retener a los mejores colaboradores se vuelven cada vez más creativos y los líderes de Recursos Humanos invierten tiempo y recursos para reducir su rotación de personal en puestos clave. Sólo las áreas de reclutamiento en las organizaciones conocen lo complejo que puede ser encontrar un reemplazo de un colaborador considerado de alto potencial. (Estrada, 2018)

Existen dos rubros que impactan directamente en los costos directos al momento de reemplazar a un colaborador: la selección y el entrenamiento (inducción). Aquí se debe considerar que no será el mismo costo involucrado, si estamos hablando de reemplazar a un nivel ejecutivo, que a un “entry level” en la empresa. La complejidad y tiempo invertido es directamente proporcional mientras más alto es el nivel jerárquico. Los valores estándares de mercado si se hace una selección a nivel gerencial con un headhunter, pueden llegar a ser de hasta 2 remuneraciones del ejecutivo. A eso se debe agregar el tiempo requerido para su inducción, capacitación en los temas propios de su área, conocimiento de clientes, proveedores, colaboradores, etc. Normalmente se estima entre 6 meses y 1 año que se alcanza a cubrir satisfactoriamente todos esos frentes. (Estrada, 2018)

La salida de un colaborador también trae una disminución del costo laboral, considerando el ingreso que la persona percibía en la empresa y el número de días que el cargo estará vacante. Según estudios realizados por Deloitte, el tiempo promedio de reponer un cargo gerencial, es de aproximadamente 45 días. Este es un valor referencial que puede variar según la especialización del cargo. (Estrada, 2018)

(Chiavenato, 2000) afirma: La rotación de personal es un problema que cada organización debe evaluar según sus propios cálculos y sus intereses básicos, la que implica costos primarios, secundarios y terciarios, los cuales se explican a continuación:

Costos primarios de la rotación de personal: Se relacionan directamente con el retiro de cada empleado y su reemplazo por otro. Incluyen: costos de reclutamiento y selección, costos de registro y documentación, costos de ingreso, costos de desvinculación.

Costos secundarios de la rotación de personal: Abarcan aspectos intangibles difíciles de evaluar en forma numérica porque sus características son cualitativas en su mayor parte, se refieren a los efectos colaterales inmediatos de la rotación. Los costos secundarios de la rotación de personal incluyen: efectos en la producción, efectos en la actitud del personal, costos extralaborales, costos operacionales.

Costos terciarios de la rotación de personal: Se relacionan con los efectos colaterales mediatos de la rotación, que se manifiestan a mediano y largo plazo. En tanto los costos primarios son cuantificables y los secundarios son cualitativos, los costos terciarios son sólo estimables. Entre dichos costos se cuentan: costos de inversión extra, pérdidas en los negocios. (p. 198)

2.2 CONCEPTUALIZACIÓN

Algoritmos de predicción: son conjuntos de reglas y procedimientos matemáticos que utilizan datos históricos para hacer pronósticos sobre eventos futuros. Estos algoritmos se aplican en una variedad de campos, desde finanzas hasta pronósticos meteorológicos y análisis de datos. (James, Witten, Hastie, & Tibshirani, 2013)

Arraigo: El arraigo se refiere a la sensación de conexión profunda y arraigada que una persona experimenta con un lugar, una comunidad, una cultura o una identidad. Se manifiesta como un fuerte sentido de pertenencia y compromiso con ciertas ubicaciones geográficas o grupos sociales. El arraigo puede influir en las decisiones, el comportamiento y las identidades individuales, y a menudo se asocia con la sensación de arraigo cultural o arraigo territorial. (Relph, 1976)

Costo Laboral: es el costo total de emplear a un trabajador en una organización. Esto abarca no solo el salario o sueldo bruto pagado al empleado, sino también todos los gastos

adicionales asociados con la relación laboral, como contribuciones a la seguridad social, beneficios de salud, seguros de compensación laboral, impuestos sobre la nómina, gastos de reclutamiento y selección, capacitación, desarrollo profesional y otros gastos relacionados con el personal. (Newman, Gerhart, & Milkovich, 2021)

KNIME: es una plataforma de código abierto utilizada para el análisis de datos, la minería de datos y la creación de flujos de trabajo de procesamiento de datos. Permite a los usuarios diseñar, implementar y ejecutar flujos de trabajo que involucran tareas como la limpieza de datos, la transformación, la integración de datos y el análisis predictivo. (Berthold, y otros, 2008)

Machine Learning: es una disciplina de la inteligencia artificial que se enfoca en desarrollar algoritmos y modelos que permiten a las computadoras aprender y mejorar su rendimiento en tareas específicas a través de la experiencia con datos, sin una programación explícita. (Hastie, Tibshirani, & Friedman, 2009)

Perfil Sociodemográfico: es una representación detallada de una población o grupo de personas que incluye información sobre sus características demográficas (como edad y género) y sociodemográficas (como nivel educativo, ingresos, ocupación, estado civil, ubicación geográfica, entre otros). Este tipo de perfil proporciona una comprensión más completa de la composición de una población y es fundamental para tomar decisiones informadas en una variedad de campos. (Borjas, 2014)

Premisa: Una premisa es una declaración o proposición que se acepta como verdadera y que se utiliza como punto de partida en un razonamiento lógico o un argumento. Las premisas son fundamentales en la construcción de argumentos válidos y sólidos, ya que proporcionan el respaldo lógico necesario para llegar a una conclusión. (Hurley, 2011)

Rotación de Personal: La rotación de personal es un indicador clave en la gestión de recursos humanos que mide la movilidad de los empleados dentro de una organización. Puede ser voluntaria, cuando los empleados renuncian por su propia elección, o involuntaria, cuando la organización decide terminar el empleo de un trabajador. La rotación de personal tiene implicaciones significativas en la cultura laboral, la estabilidad de la fuerza laboral y la rentabilidad de una empresa. (Bluedorn, 1982)

Salida Voluntaria: también conocida como renuncia, se refiere a la decisión de un

empleado de abandonar su puesto de trabajo de manera deliberada y por elección propia. Este tipo de salida no implica presiones o decisiones forzadas por parte de la organización. En la mayoría de los casos, un empleado decide renunciar debido a motivos personales, profesionales o una combinación de ambos. Se produce cuando un empleado toma la decisión de dejar su empleo de manera autónoma, sin ser influenciado por factores externos, como despidos o presiones de la organización. Esta decisión puede estar motivada por una variedad de razones, que van desde oportunidades de carrera, insatisfacción en el trabajo, conflictos personales o profesionales, hasta razones de estilo de vida. (Steel & Ovalle, 1984)

Variables Demográficas: son atributos o características específicas de una población que se utilizan para describir y analizar sus características y composición. Estas variables incluyen información sobre la edad, el género, la etnia, la ubicación geográfica, el nivel educativo, el estado civil, la ocupación, el tamaño de la familia, la religión, entre otros. El análisis de variables demográficas es esencial para comprender las dinámicas de la población y tomar decisiones informadas en diversos campos. (Hauser & Duncan, 2001)

Variables Exógenas: son aquellas que se consideran independientes en un modelo o sistema y que no son afectadas por otras variables dentro del modelo. Estas variables se utilizan para representar factores externos o condiciones que influyen en el sistema, pero que no están determinados por las relaciones internas del sistema en sí. Son fundamentales en el análisis de causalidad y en la modelización de procesos complejos. (Greene, 2012)

2.3 TEORÍAS DE SUSTENTO

2.3.1 BASES TEÓRICAS

2.3.1.1 TEORIA DEL EQUILIBRIO ORGANIZACIONAL

El equilibrio organizacional es un factor clave para el crecimiento sostenible de cualquier empresa. Esto se debe a que una organización bien equilibrada puede maximizar su eficiencia y minimizar sus costos, lo que a su vez aumenta su rentabilidad. El equilibrio organizacional implica la adecuada distribución de responsabilidades y poder entre los colaboradores y los departamentos, lo que permite una fluidez en los procesos y evita la sobrecarga de trabajo. Además, el equilibrio organizacional permite la creación de una cultura empresarial sólida y coherente, lo que contribuye a la lealtad y compromiso de los colaboradores y, por ende, a la retención de talentos. (Maldonado,

n.d.)

Otro aspecto importante del equilibrio organizacional es la alineación de los objetivos y estrategias empresariales con los recursos y capacidades disponibles. Es decir, se trata de contar con una planificación estratégica sólida que considere la situación actual y las necesidades futuras de la empresa, para luego diseñar una estructura organizacional que permita el logro de dichos objetivos de manera eficiente y efectiva. La innovación y la adaptación a los cambios también son cruciales para el equilibrio organizacional, ya que permiten mantener la competitividad en un mercado en constante evolución. (Maldonado, n.d.)

El equilibrio organizacional es fundamental para lograr un crecimiento sostenible en el largo plazo. Para lograrlo, es necesario considerar aspectos como la distribución adecuada de responsabilidades y poder, la creación de una cultura empresarial sólida, la alineación de objetivos y estrategias, la innovación y adaptación a cambios, entre otros. Todo esto permitirá a la empresa tener una base sólida para desarrollarse de manera sostenible y competitiva en el mercado. (Maldonado, n.d.)

Es importante reconocer que cada empresa es única y que la forma en que se aplica la teoría del equilibrio organizacional puede variar. Sin embargo, existen algunas estrategias clave que pueden ayudar a lograr este equilibrio. Aquí te presentamos algunas de ellas: (Maldonado, n.d.)

1. Evalúa la situación actual: Para aplicar la teoría del equilibrio organizacional en tu empresa, es necesario tomar una fotografía de la situación actual en la que se encuentra tu organización. Esto te permitirá identificar las áreas que necesitan mejorar y las que están funcionando bien. (Maldonado, n.d.)

2. Establece objetivos claros: Una vez que hayas evaluado la situación actual, es importante establecer objetivos claros y alcanzables para tu empresa. Estos objetivos deben ser medibles y específicos para que puedas evaluar el progreso a lo largo del tiempo. (Maldonado, n.d.)

3. Desarrolla una cultura organizacional sólida: La cultura organizacional es uno de los elementos más importantes para lograr el equilibrio organizacional. Para desarrollar una cultura sólida, es necesario fomentar valores y comportamientos enfocados en el trabajo en equipo, la colaboración y el respeto mutuo. (Maldonado, n.d.)

4. Invierte en tus empleados: Los empleados son la fuerza laboral detrás de cualquier

empresa exitosa. Por lo tanto, es importante invertir en ellos a través de oportunidades de capacitación y desarrollo, beneficios y recompensas por desempeño. (Maldonado, n.d.)

5. Utiliza los recursos adecuadamente: Finalmente, es importante utilizar adecuadamente los recursos de tu empresa, incluyendo el presupuesto, el tiempo y los equipos. Identifica cuáles son los recursos críticos para la operación de tu empresa y asegúrate de utilizarlos de manera efectiva. (Maldonado, n.d.)

La teoría del equilibrio organizacional es un enfoque estratégico que busca equilibrar las necesidades y objetivos de la organización con las de sus empleados y otros interesados clave. Esta teoría se enfoca en la creación de un ambiente de trabajo saludable y productivo donde se apoya a los empleados para que logren sus metas y así puedan contribuir al éxito de la organización. Para implementar esta teoría, se deben ser consciente de las dinámicas internas y externas que afectan a la organización y adaptarse a ellas de manera efectiva. Al crear un ambiente laboral equilibrado, las organizaciones pueden tener un equipo motivado y comprometido que puede trabajar en conjunto para alcanzar los objetivos de la empresa. si se implementa correctamente, la teoría del equilibrio organizacional puede ayudar a las organizaciones a lograr un rendimiento óptimo mientras se asegura que sus empleados estén felices y se sientan valorados en el trabajo. (Maldonado, n.d.)

2.3.1.2 MODELO DESDOBLANTE

En el modelo desdoblante (unfolding model) de rotación, hay cuatro rutas diferentes para un empleado que decide salir de su trabajo: insatisfacción; mejores alternativas; seguir un plan y partir sin un plan. Cualquiera que sea la ruta que se siga, habrá algo de choque; estos pueden ser esperados o no esperados, relacionados o no con el trabajo y, en general, positivos o negativos. He aquí la explicación de cada uno: (Melendres, Ramirez, Garcia, & Aranibar, 2017)

La insatisfacción: es el tipo tradicional de rotación, en el que la causa se vincula con las actitudes que llevan finalmente a la salida de la empresa. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

Mejores alternativas: cuando un empleado parte por una oportunidad más atractiva, independientemente de si hay satisfacción o no en su actual puesto de trabajo. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

Seguir un plan: este tipo de rotación se da en las personas que previamente han planeado salir. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

Partir sin un plan: este tipo de rotación puede ser resultado de una acción o reacción impulsiva, y por lo general responde a un shock negativo. (Melendres, Ramirez, Garcia, & Aranibar, 2017)

2.3.1.2 MACHINE LEARNING

El machine learning es un subconjunto de la inteligencia artificial (IA). Se enfoca en enseñar a las computadoras para que aprendan de los datos y mejoren con la experiencia –en lugar de ser explícitamente programadas para hacerlo–. En el machine learning, los algoritmos están entrenados para encontrar patrones y correlaciones en grandes data sets, y para lograr las mejores decisiones y proyecciones en base a ese análisis. Las aplicaciones de machine learning mejoran con el uso y se vuelven más precisas a medida que tienen acceso a más datos. (SAP, n.d.)

El machine learning se compone de diferentes tipos de modelos de aprendizaje automático, y usa varias técnicas algorítmicas. Dependiendo de la naturaleza de los datos y el resultado deseado, se puede usar uno de los cuatro modelos de aprendizaje: supervisado, no supervisado, semisupervisado, o de refuerzo. Dentro de cada uno de esos modelos, se pueden aplicar una o más técnicas algorítmicas –relacionadas con los data sets en uso y los resultados buscados–. Los algoritmos de machine learning están diseñados básicamente para clasificar cosas, encontrar patrones, proyectar resultados, y tomar decisiones fundamentadas. Los algoritmos pueden usarse uno a la vez o combinarse para lograr la mayor precisión posible cuando se trata de datos complejos y más impredecibles. (SAP, n.d.)

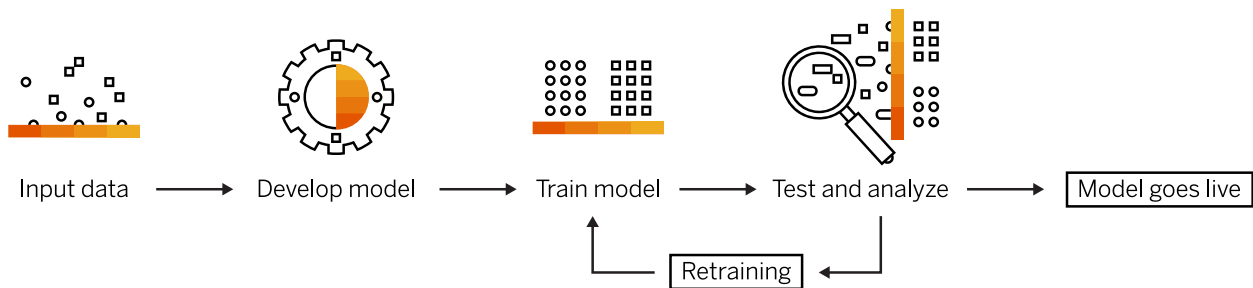


Ilustración 1 Funcionamiento de Machine Learning

Fuente: (SAP, n.d.)

2.3.1.2.1 MODELOS DE MACHINE LEARNING

Modelo de aprendizaje supervisado: Es el primero de los cuatro modelos de machine learning. En sus algoritmos, a la máquina se la entrena mediante ejemplos. Los modelos de supervised learning consisten en pares de datos de "entrada" y "salida", donde el de salida se etiqueta con el valor deseado. Por ejemplo, supongamos que el objetivo es que la máquina diga la diferencia entre margaritas y pensamientos. Uno de los pares de datos de entrada binarios incluye las imágenes de una margarita y de un pensamiento. El resultado deseado para ese par en particular es escoger la margarita, así que se la identificará previamente como el resultado correcto. (SAP, n.d.)

Mediante un algoritmo, el sistema compila todos esos datos del entrenamiento a lo largo del tiempo, y comienza a determinar similitudes correlativas, diferencias, y otros puntos de lógica –hasta que puede prever por sí mismo las respuestas a la pregunta "margarita o pensamiento"–. Es el equivalente de darle a un niño un conjunto de problemas con una clave de respuestas, y luego pedirle que muestre su trabajo y explique su lógica. Los modelos de supervised learning se usan en muchas de las herramientas con las que interactuamos todos los días, como los motores de recomendación de productos y las aplicaciones de análisis de tránsito como Waze, que prevén la ruta más rápida en diferentes horarios. (SAP, n.d.)

Modelo de aprendizaje no supervisado: es el segundo de los cuatro modelos de machine learning. En este, no existe una clave de respuestas. La máquina estudia los datos de entrada – muchos de los cuales no están etiquetados ni estructurados– y comienza a identificar patrones y correlaciones, usando todos los datos relevantes y accesibles. En muchos sentidos, el unsupervised learning sigue el modelo humano de observación del mundo. Usamos la intuición y experiencia para agrupar las cosas juntas. A medida que experimentamos más ejemplos de algo, nuestra capacidad para categorizarlo e identificarlo se vuelve cada vez más precisa. Para las máquinas, la "experiencia" se define por la cantidad de datos que se introducen y ponen a disposición. Ejemplos comunes para aplicar unsupervised learning incluyen reconocimiento facial, análisis de secuencias genéticas, investigación de mercado, y ciberseguridad. (SAP, n.d.)

Modelo de aprendizaje semi-supervisado: es el tercero de los cuatro modelos de machine learning. En un mundo perfecto, todos los datos se estructurarían y etiquetarían antes de ser ingresados a un sistema. Pero como obviamente esto no es factible, el semi-supervised learning se

convierte en una solución viable cuando hay grandes cantidades de datos crudos y no estructurados. Este modelo consiste en ingresar pequeñas cantidades de datos etiquetados para aumentar los data sets no etiquetados. Esencialmente, los primeros actúan para dar un inicio de funcionamiento al sistema, y pueden mejorar considerablemente la velocidad y precisión del aprendizaje. Un algoritmo de semi-supervised learning instruye a la máquina para que analice los datos etiquetados según propiedades correlativas que podrían aplicarse a los no etiquetados. (SAP, n.d.)

Modelo de aprendizaje reforzado: es el cuarto modelo de machine learning. En el supervised learning, la máquina recibe la respuesta de referencia y aprende encontrando correlaciones entre todos los resultados correctos. El modelo de reinforcement learning no incluye una respuesta de referencia, sino que más bien ingresa un conjunto de acciones permitidas, reglas, y potenciales estados finales. Cuando el objetivo deseado del algoritmo es fijo o binario, las máquinas pueden aprender mediante el ejemplo. Pero en los casos donde el resultado deseado es mutable, el sistema debe aprender por experiencia y recompensa. En los modelos de reinforcement learning, la "recompensa" es numérica, y se programa dentro del algoritmo como algo que el sistema busca recopilar. (SAP, n.d.)

2.3.2 METODOLOGÍAS DESARROLLADAS

2.3.2.1 RECOLECCION DE DATOS

La recolección de datos es considerada como la medición es una precondición para obtener el conocimiento científico. El instrumento de recolección de datos está orientado a crear las condiciones para la medición. Los datos son conceptos que expresan una abstracción del mundo real, de lo sensorial, susceptible de ser percibido por los sentidos de manera directa o indirecta, donde todo lo empírico es medible. (Hernandez Mendoza & Duana Avila, 2020)

Por lo que las técnicas de recolección de datos comprenden procedimientos y actividades que le permiten al investigador obtener información necesaria para dar respuesta a su pregunta de investigación. Existen múltiples y diferentes instrumentos útiles para la recolección de datos y para ser usados en todo tipo de investigaciones ya sean cuantitativas, cualitativas o mixtas. En la actualidad en investigación científica hay gran variedad de técnicas e instrumentos

para la recolección de información, según Muñoz Giraldo et al. (2001) la investigación cuantitativa utiliza generalmente la encuesta, entrevista, observación sistemática, análisis de contenidos, fichas de cotejo etc. Todo instrumento utilizado en la recolección de datos en una investigación científica debe ser confiable, objetivo y que tenga validez, si alguno de estos elementos no se cumple el instrumento no será útil y los resultados obtenidos no serán legítimos. (Hernandez Mendoza & Duana Avila, 2020)

Para obtener una adecuada recolección de datos el investigador debe seguir un planeamiento detallado de lo que se hará en la recolección de datos como:

Autorización: condición necesaria previa a la recolección de datos.

Tiempo: es la estimación del tiempo que el investigador para requerir para la realización de la investigación (recolección, tabulación y análisis de datos).

Recursos: humanos, económicos y físicos.

Capacitación: referente a los objetivos, selección de muestra, instrumentos y procedimientos para la recolección de datos.

Supervisión y coordinación: manera en que se distribuirá el recurso humano para cubrir los componentes de información del estudio.

Por lo anterior las técnicas de recolección de datos aluden a procedimientos de actuación concreta y particular de recogida de información relacionada con el método de investigación que se esté utilizando, el uso de unas técnicas u otras van a depender del marco de investigación a realizar. (Hernandez Mendoza & Duana Avila, 2020)

2.3.2.2 ALGORITMOS DE PREDICCIÓN

Los algoritmos de predicción se pueden agrupar en tres modelos:

1) Modelos lineales

Estos tratan de encontrar una línea que se “ajuste” bien a la nube de puntos que se disponen. Aquí destacan desde modelos muy conocidos y usados como la regresión lineal (también conocida como la regresión de mínimos cuadrados), la logística (adaptación de la lineal a problemas de clasificación -cuando son variables discretas o categóricas-). Estos dos modelos tienen el problema del “overfit”, esto significa que se ajustan “demasiado” a los datos disponibles, con el riesgo que

esto tiene para nuevos datos que pudieran llegar. Al ser modelos relativamente simples, no ofrecen resultados muy buenos para comportamientos más complicados. (Sandoval, 2018)

2) Modelos de árbol

Son modelos precisos, estables y más sencillos de interpretar básicamente porque construyen unas reglas de decisión que se pueden representar como un árbol. A diferencia de los modelos lineales, pueden representar relaciones no lineales para resolver problemas. En estos modelos, destacan los árboles de decisión y los random forest (una media de árboles de decisión). Al ser más precisos y elaborados, obviamente ganamos en capacidad predictiva, pero perdemos en rendimiento. (Sandoval, 2018)

2.3.3 INSTRUMENTOS UTILIZADOS

2.3.3.1 ALGORITMOS DE PREDICCIÓN

2.3.3.1.1 REGRESIÓN LINEAL

La regresión lineal es un enfoque estadístico que se utiliza para modelar la relación entre una variable dependiente (la variable que se quiere predecir) y una o más variables independientes (predictoras) mediante una ecuación lineal. En una regresión lineal simple, se ajusta una línea recta a los datos para predecir la variable dependiente en función de una única variable independiente. En una regresión lineal múltiple, se utilizan múltiples variables independientes para predecir la variable dependiente. (Montgomery, Peck, & Vining, 2012)

2.3.3.1.2 REGRESIÓN LOGÍSTICA

La regresión logística es un enfoque estadístico que se utiliza para modelar la relación entre una variable dependiente categórica (por ejemplo, sí/no, éxito/fallo, categorías multinomiales) y una o más variables independientes. Esta técnica utiliza una función logística para modelar la probabilidad de que la variable dependiente pertenezca a una categoría en función de las variables independientes. La regresión logística es ampliamente utilizada en campos como la epidemiología, la investigación médica, el análisis de datos de encuestas y la clasificación en aprendizaje automático. (Hosmer, Lemeshow, & Sturdivant, 2013)

2.3.3.1.3 NAIVE BAYES

El algoritmo de Naive Bayes es un método de clasificación que se basa en el teorema de

Bayes para estimar la probabilidad de que un elemento pertenezca a una clase particular. Asume independencia condicional entre las características utilizadas para la clasificación, lo que significa que se considera que las características son independientes entre sí, lo cual es una simplificación que permite realizar cálculos más eficientes. El clasificador de Naive Bayes es ampliamente utilizado en la clasificación de texto, minería de datos y muchas otras aplicaciones de aprendizaje automático. (Russell & Norvig, 2009)

2.3.3.1.4 RANDOM FOREST

Random Forest es un algoritmo de aprendizaje automático que se basa en la construcción de múltiples árboles de decisión. Cada árbol se entrena en una submuestra aleatoria de datos y, a menudo, con un subconjunto aleatorio de características. Luego, para la clasificación, se toma la decisión por mayoría de votos de los árboles (votación en clasificación) o se promedian las predicciones (promedio en regresión) para obtener un resultado final. Random Forest es conocido por su capacidad para manejar grandes conjuntos de datos, manejar variables categóricas y numéricas, y reducir el sobreajuste, lo que lo convierte en una elección popular en aplicaciones de aprendizaje automático. (Breiman, Random forests., 2001)

2.3.3.1.5 ARBOL DE DECISION

Un árbol de decisión es una estructura jerárquica que se utiliza para tomar decisiones basadas en datos. Cada nodo del árbol representa un atributo o característica, y las ramas salientes de cada nodo representan los posibles valores o resultados de esa característica. Al seguir un camino desde el nodo raíz hasta un nodo hoja, se toma una serie de decisiones basadas en los valores de las características, lo que lleva a una decisión final o clasificación. Los árboles de decisión son ampliamente utilizados en clasificación, regresión, y toma de decisiones automatizadas. (Breiman, Friedman, Olshen, & Stone, 1984)

2.3.3.2 KNIME

KNIME pertenece a una nueva generación de herramientas dominadas como Plataformas de Data Science y Machine Learning por Gartner. Estas herramientas permiten a científicos de datos expertos, analistas o usuarios de negocio interactuar con sus datos y crear, desplegar y gestionar sus modelos de analítica avanzada. Las herramientas integran las funcionalidades principales para realizar proyectos de minería de datos: importación de datos, preparación de datos,

exploración de datos, modelado, evaluación y despliegue. Dentro de estas herramientas KNIME se encuentra en el grupo de líderes del último diagrama de Gartner conocido también como la “navaja suiza” del mercado. (Lis Data Solutions, n.d.)

Como puntos fuertes podemos destacar:

Facilidad de uso: a través de interfaz visual la programación de aplicaciones en KNIME es altamente intuitiva. Conectando visualmente nodos que encapsulan distintas funciones e integrando módulos automatizados de Machine Learning y modelos preprogramados, facilita la analítica avanzada a los usuarios de negocio sin experiencia en este ámbito. Al mismo tiempo brinda un ecosistema óptimo para desarrolladores avanzados con la posibilidad de integrar programación en Python o R. (Lis Data Solutions, n.d.)

Extensas funcionalidades: KNIME ofrece el ciclo completo de Data Mining. Con la posibilidad de conectarse a múltiples y heterogéneas fuentes de datos podremos unificar datos provenientes de distintas BBDD, archivos y servicios web diversos como Azure, etc. con muy poco esfuerzo. Con una gran variedad de nodos para el preprocesamiento la herramienta ofrece las condiciones óptimas para la generación de procesos de ETL automatizada. Finalmente ofrece los principales algoritmos y métodos de evaluación para la generación de modelos potentes. Adicionalmente dispone de múltiples extensiones (Text Processing, Big Data con Spark y Hadoop, Deep Learning con TensorFlow y Keras y muchas más) que empoderan la herramienta aún más. (Lis Data Solutions, n.d.)

Bajo costo de adquisición e implantación: KNIME cuenta con una versión gratuita “KNIME Analytics Platform” para el uso personal, así como una versión de pago “KNIME Server” para el uso en organizaciones que quieren llevar sus actividades de data mining a un nuevo nivel. (Lis Data Solutions, n.d.)

2.4 MARCO LEGAL

En Honduras no existe una ley actualizada para normar los temas laborales, sin embargo, tenemos un Código Laboral, vigente desde 1977, que dicta las normas y limitantes en cuanto la gestión laboral de empleados que son contratados o que son dados de baja en una empresa. A continuación, se describen algunos de los artículos del código del trabajo que se asocian con la

investigación:

CÓDIGO DEL TRABAJO

Artículo 95. Además de las contenidas en otros artículos de este Código, en sus reglamentos y en las leyes de previsión social, son obligaciones de los patronos:

1. Pagar la remuneración pactada en las condiciones, períodos y lugares convenidos en el contrato, o en los establecidos por las leyes y reglamentos de trabajo, o por los reglamentos internos o convenios colectivos, o en su defecto por la costumbre;

2. Pagar al trabajador el salario correspondiente al tiempo que dejare de trabajar por causas imputables al patrono;

3. Proporcionar oportunamente a los trabajadores los útiles, instrumentos y materiales necesarios para ejecutar el trabajo convenido, los cuales dará de buena calidad y repondrá tan pronto como dejen de ser eficientes, siempre que aquellos no se hayan comprometido a usar herramientas propias;

4. Proporcionar local seguro para la guarda de los instrumentos y útiles de trabajo pertenecientes al trabajador, siempre que aquellos deban permanecer en el lugar en que presten los servicios, sin que sea lícito al patrono retenerlos a título de indemnización, garantía o cualquier otro. El inventario de instrumentos o útiles de trabajo deberá hacerse siempre que cualquiera de las partes lo solicite;

5. Conceder licencia al trabajador para que pueda cumplir con las obligaciones de carácter público impuestas por la ley; en caso de grave calamidad doméstica debidamente comprobada; para desempeñar comisiones sindicales inherentes a la organización o para asistir al entierro de sus compañeros, siempre que avise con la debida oportunidad al patrono o a su representante y que, en los dos (2) últimos casos, el número de los que se ausenten no sea tal que perjudique el funcionamiento de la empresa; pero el patrono no está obligado a reconocer por estas causas más de dos (2) días con goce de salario en cada mes calendario, y en ningún caso más de quince (15) días en el mismo año. Cuando la comisión sea de carácter permanente o desempeñen cargos públicos de elección popular, el trabajador o trabajadores podrán volver al puesto que ocupaban, conservando todos los derechos derivados de sus respectivos contratos, siempre y cuando regresen a sus labores dentro del término de dos (2) años. Los sustitutos tendrán carácter de interinos. Cuando el trabajador desempeñe cargos de dirección sindical, las licencias durarán por el tiempo que permanezca en sus funciones. Se prohíbe al patrono reconocer salarios por esta causa. Dicha licencia será solicitada por la organización sindical respectiva;

6. Guardar a los trabajadores la debida consideración, absteniéndose de maltratos de palabra o de obra y de actos que pudieran afectar su dignidad;

7. Adoptar medidas adecuadas para crear y mantener en sus empresas las mejores

condiciones de higiene y seguridad en el trabajo;

8. Permitir y facilitar la inspección y vigilancia que las autoridades de trabajo, sanitarias y administrativas deban practicar en su empresa, establecimiento o negocio, y darles los informes que a ese efecto sean indispensables, cuando lo soliciten en cumplimiento de las disposiciones legales correspondientes;

9. Tomar las medidas indispensables y las que fijen las leyes para prevenir accidentes en el uso de maquinarias, instrumentos o material de trabajo, y mantener una provisión de medicinas y útiles indispensables para la atención inmediata de los accidentes que ocurran;

10. Cubrir las indemnizaciones por los accidentes que sufran los trabajadores con motivo del trabajo o a consecuencia de él, y por las enfermedades profesionales que los mismos contraigan en el trabajo que ejecuten, o en el ejercicio de la profesión que desempeñen;

11. Mantener a la disposición de empleados o dependientes en los almacenes, tiendas, farmacias, bazares, bodegas, depósitos de mercaderías y demás establecimientos análogos, el número suficiente de sillas;

12. Hacer las deducciones que por cuotas sindicales ordinarias o extraordinarias soliciten los sindicatos. Estos comprobarán que las cuotas cuyo descuento piden, son las que establecen sus estatutos. Asimismo, deberán hacer las deducciones que fija el Artículo 60-A de este Código a los trabajadores no sindicalizados que en él se especifican, y las pondrán a disposición del sindicato sin necesidad de solicitud ni requerimiento.¹²

13. Hacer las deducciones de cuotas ordinarias para la constitución y fomento de las cooperativas y cajas de ahorro formadas por los trabajadores sindicalizados. Unas y otras comprobarán que las cuotas cuyo descuento piden son las que establecen sus estatutos;

14. Reservar, cuando la población fija de un centro rural de trabajo excede de doscientos (200) habitantes, un espacio de terreno no menor de cinco mil (5,000) metros cuadrados para el establecimiento de mercados públicos, edificios para los servicios municipales y centros recreativos, siempre que dicho centro de trabajo esté a una distancia no menor de cinco (5) kilómetros de la población más próxima;

15. Suministrarle al trabajador habitación higiénica y alimentación sana y suficiente, en el caso de que, de acuerdo con el contrato, se haya obligado a hospedarle y alimentarle;

16. Derogado 13

17. Establecer y sostener escuelas de educación primaria en beneficio de los hijos de los trabajadores, cuando se trate de centros rurales y siempre que el número de niños en edad escolar sea mayor de veinte (20). La educación que se imparta en esos establecimientos se sujetará a los planes y programas de estudios de las escuelas oficiales. Los sueldos no serán menores que los retribuidos a los maestros en las escuelas que costee el Estado;

18. Los patronos que empleen más de doscientos (200) y menos de dos mil (2,000)

trabajadores, harán por su cuenta los gastos indispensables para sostener en forma decorosa los estudios técnicos, industriales o prácticos, en centros especiales, nacionales o extranjeros, de uno de sus trabajadores o de uno de los hijos de éstos, designado en atención a sus aptitudes, cualidades y dedicación, por los mismos trabajadores y el patrono. Cuando tenga a sus órdenes más de dos mil (2,000) trabajadores, deberá sostener, en las condiciones antes indicadas, tres (3) pensionados. El patrono sólo podrá cancelar la pensión cuando sea reprobado el pensionado en el curso de un (1) año o cuando observe mala conducta; pero en estos casos será sustituido por otro. Los pensionados que hayan terminado sus estudios deberán prestar sus servicios por lo menos durante dos (2) años, al patrono que los hubiere pensionado;

19. Llevar a cabo los reajustes de acuerdo con las estipulaciones del contrato colectivo. A falta de éstas, respetarán los derechos de antigüedad y, en igualdad de condiciones, preferirán a los elementos sindicalizados para que sigan trabajando;

20. En los lugares en donde existan enfermedades tropicales o endémicas, proporcionar a sus trabajadores los medicamentos profilácticos que determine la autoridad sanitaria del lugar;

21. En los cortes de piedra, cantera, minas de arena, hornos de calcinación, basalto y fábricas de cemento, observar los reglamentos de policía y seguridad expedidos por la Secretaría de Trabajo y Seguridad Social sobre trabajos mineros, fijando tales reglamentos en lugares visibles de las minas, cañones o niveles para conocimiento de los trabajadores;

22. Establecer un escalafón que rijan los ascensos y demás cambios en el personal, tomando en cuenta fundamentalmente la capacidad y eficiencia del trabajador y en igualdad de condiciones, su antigüedad dentro de la empresa. La capacidad y eficiencia de los trabajadores debe ser apreciada por organismos compuestos de trabajadores y patronos y, cuando no se lograse acuerdo, con la intervención de la Secretaría de Trabajo y Seguridad Social;

23. Cumplir las demás obligaciones que le impongan las leyes y reglamentos de trabajo.

Artículo 98. Se prohíbe a los trabajadores:

1. Faltar al trabajo, o abandonarlo en horas de labor, sin justa causa de impedimento o sin permiso del patrono;

2. Presentarse al trabajo en estado de embriaguez, o bajo la influencia de drogas estupefacientes, o en cualquier otra condición anormal análoga;

3. Portar armas de cualquier clase durante las horas de labor, excepto en los casos especiales autorizados debidamente por las leyes, o cuando se trate de instrumentos punzantes, cortantes o punzo cortantes que formen parte de las herramientas o útiles propios del trabajo;

4. Sustraer de la fábrica, taller o establecimiento, los útiles de trabajo y las materias primas o productos elaborados, sin permiso del patrono;

5. Disminuir intencionalmente el ritmo de ejecución del trabajo, suspender ilegalmente

labores, promover suspensiones intempestivas del trabajo, o excitar a su declaración y mantenimiento, sea que se participe o no en ellas;

6. Coartar la libertad para trabajar o no trabajar; o para afiliarse o no a un sindicato o permanecer en él o retirarse;

7. Usar los útiles o herramientas suministrados por el patrono para objeto distinto de aquél a que están normalmente destinados; y,

8. Hacer durante el trabajo, propaganda político electoral o contraria a las instituciones democráticas creadas por la Constitución, o ejecutar cualquier acto que signifique coacción de la libertad de conciencia que la misma establece, lo mismo que hacer colectas o suscripciones en las horas de trabajo.

Referente a lo relacionado con dar de baja a un empleado, el código laboral vigente norma como causales de terminación de contrato:

Artículo 111. Son causas de terminación de los contratos de trabajo:

1. Cualquiera de las estipuladas en ellos si no fueren contrarias a la ley;
2. El mutuo consentimiento de las partes;
3. Muerte del trabajador o incapacidad física o mental del mismo, que haga imposible el cumplimiento del contrato;
4. Enfermedad del trabajador en el caso previsto por el artículo 104;
5. Pérdida de la libertad del trabajador en el caso previsto en el artículo 106;
6. Caso fortuito o fuerza mayor;
7. Perder la confianza del patrono, el trabajador que desempeñe un cargo de dirección, fiscalización o vigilancia; tales como mayordomos, capataces, debiendo justificarse a juicio de la Dirección General del Trabajo, o sus representantes, los motivos de tal desconfianza; más si había sido promovido de un puesto de escalafón en las empresas en que éste existe, volverá a él, salvo que haya motivo justificado para su despido. Lo mismo se observará cuando el trabajador que desempeñe un puesto de confianza solicite volver a su antiguo empleo;
8. La suspensión de actividades por más de ciento veinte (120) días en los casos 1), 3), 4), 5), y 6) del artículo 100;
9. Liquidación o clausura definitiva en la empresa o establecimiento;
10. Ejercicio de las facultades que conceden a las partes los artículos 112 y 114;
11. Insolvencia o quiebra; y,
12. Resolución del contrato decretada por autoridad competente. En los casos previstos en los siete (7) primeros incisos de este artículo, la terminación del contrato no acarreará

responsabilidades para ninguna de las partes.

En los casos del inciso 8o. poco habrá responsabilidad para las partes a excepción del que se refiere a muerte o incapacidad del patrono, en que los trabajadores tendrán derecho al pago del preaviso.

En el caso del inciso 9), el patrono estará obligado a proceder en la misma forma que para la suspensión establecen los artículos 101 y 102; a menos que la causa haya sido la insolvencia o quiebra fraudulenta o culpable, declarada por autoridad competente, en cuyo caso estará obligado también al pago de las demás indemnizaciones y prestaciones a que tengan derecho los trabajadores.

En el caso del inciso 12) se procederá de acuerdo con lo que se disponga en la sentencia que ordene la resolución del contrato.

Artículo 112. Son causas justas que facultan al patrono para dar por terminado el contrato de trabajo, sin responsabilidad de su parte:

a) El engaño del trabajador o del sindicato que lo hubiere propuesto mediante la presentación de recomendaciones o certificados falsos sobre su aptitud. Esta causa dejará de tener efecto después de treinta (30) días de prestar sus servicios el trabajador;

b) Todo acto de violencia, injuria, malos tratamientos o grave indisciplina, en que incurra el trabajador durante sus labores, contra el patrono, los miembros de su familia, el personal directivo o los compañeros de trabajo;

c) Todo acto grave de violencia, injurias o malos tratamientos, fuera del servicio, en contra del patrono, de los miembros de su familia o de sus representantes y socios, o personal directivo, cuando los cometiere sin que hubiere precedido provocación inmediata y suficiente de la otra parte o que como consecuencia de ellos se hiciera imposible la convivencia o armonía para la realización del trabajo;

d) Todo daño material causado dolosamente a los edificios, obras, maquinaria o materias primas, instrumentos y demás objetos relacionados con el trabajo, y toda grave negligencia que ponga en peligro la seguridad de las personas o de las cosas;

e) Todo acto inmoral o delictuoso que el trabajador cometa en el taller, establecimiento o lugar de trabajo, cuando sea debidamente comprobado ante autoridad competente;

f) Revelar los secretos técnicos o comerciales o dar a conocer asuntos de carácter reservado en perjuicio de la empresa;

g) Haber sido condenado el trabajador a sufrir pena por crimen o simple delito, en sentencia ejecutoriada;

h) Cuando el trabajador deje de asistir al trabajo sin permiso del patrono o sin causa justificada durante dos (2) días completos y consecutivos o durante tres (3) días hábiles en el

término de un (1) mes;

i) La negativa manifiesta y reiterada del trabajador a adoptar las medidas preventivas o a seguir los procedimientos indicados para evitar accidentes o enfermedades; o el no acatar el trabajador, en igual forma y en perjuicio del patrono, las normas que éste o su representante en la dirección de los trabajos le indiquen con claridad, para obtener la mayor eficacia y rendimiento en las labores que se están ejecutando;

j) La inhabilidad o la ineficiencia manifiesta del trabajador que haga imposible el cumplimiento del contrato;

k) El descubrimiento de que el trabajador padece enfermedad infecciosa o mental incurable o la adquisición de enfermedad transmisible, de denuncia o aislamiento no obligatorio, cuando el trabajador se niegue al tratamiento y constituya peligro para terceros; y,

l) Cualquier violación grave de las obligaciones o prohibiciones especiales que incumben al trabajador, de acuerdo con los artículos 97 y 98, o cualquier falta grave calificada como tal en pacto o convenciones colectivas, fallos arbitrales, contratos individuales o reglamentos, siempre que el hecho esté debidamente comprobado y que en la aplicación de la sanción se observe el respectivo procedimiento reglamentario o convencional.

CAPÍTULO III. METODOLOGÍA

3.1 CONGRUENCIA METODOLÓGICA

En el próximo segmento, se proporcionará una descripción exhaustiva de la metodología de investigación empleada. Dentro de este apartado, se procederá a la definición precisa de las variables independientes y dependientes. Asimismo, se exhibirá un esquema detallado de dichas variables, acompañado de una clara definición conceptual para cada una de ellas.

Se abordará minuciosamente el alcance de la investigación, delineando los límites y la extensión de esta. Además, se presentarán los instrumentos que serán empleados a lo largo del estudio, dando inicio con la matriz metodológica. Esta matriz servirá como una representación gráfica que vincula las preguntas de investigación y los objetivos del proyecto con las variables pertinentes, proporcionando una visión integral de la estructura conceptual y la interrelación entre los elementos clave de la investigación.

3.1.1 MATRIZ METODOLÓGICA

En la matriz metodológica se presenta un resumen del objetivo de la investigación, así como las interrogantes que ayudaran a finalizar la misma.

Tabla 1 Matriz Metodológica

Problema	Objetivo General	Objetivos Específicos	Preguntas de Investigación	Variables Dependientes	Variable Independiente	Ítem
¿Es posible entrenar un modelo de predicción para estimar la probabilidad de que un colaborador renuncie, utilizando datos históricos sobre el personal y su historial de rotación almacenados en las bases de datos de la empresa?	Analizar si posible entrenar un modelo de predicción para estimar la probabilidad de que un colaborador renuncie, utilizando datos históricos sobre el personal y su historial de rotación almacenados en las bases de datos de la empresa.	Identificar las causas que generan la rotación en los Pinos Apparel S.A.	¿Cuáles son las causas que generan la rotación en los Pinos Apparel S.A.?	Colaboradores	Causas de Rotación	Salario Horas trabajadas Enfermedad Ubicación
		Definir cuáles son las características y variables comunes en los perfiles de los empleados que renunciaron a la empresa en el 2023.	¿Existirán características o variables comunes en los perfiles de los empleados que renuncian a la empresa?		Variables Comunes en perfiles	Genero Edad Estado Civil Número de Hijos Antigüedad Laboral Sección Clasificación Estatus Clasificación GK
		Analizar si existen técnicas de machine learning y analítica de negocios que pueden implementarse para la generación de un perfil del personal candidato a dejar la organización en un corto plazo	¿Existen técnicas de machine learning y analítica de negocios que pueden implementarse para la generación de un perfil del personal candidato a dejar la organización en un corto plazo?			

Fuente: Elaboración Propia

3.1.2 ESQUEMA DE VARIABLES DE ESTUDIO

A continuación, se presenta el esquema de variables de investigación

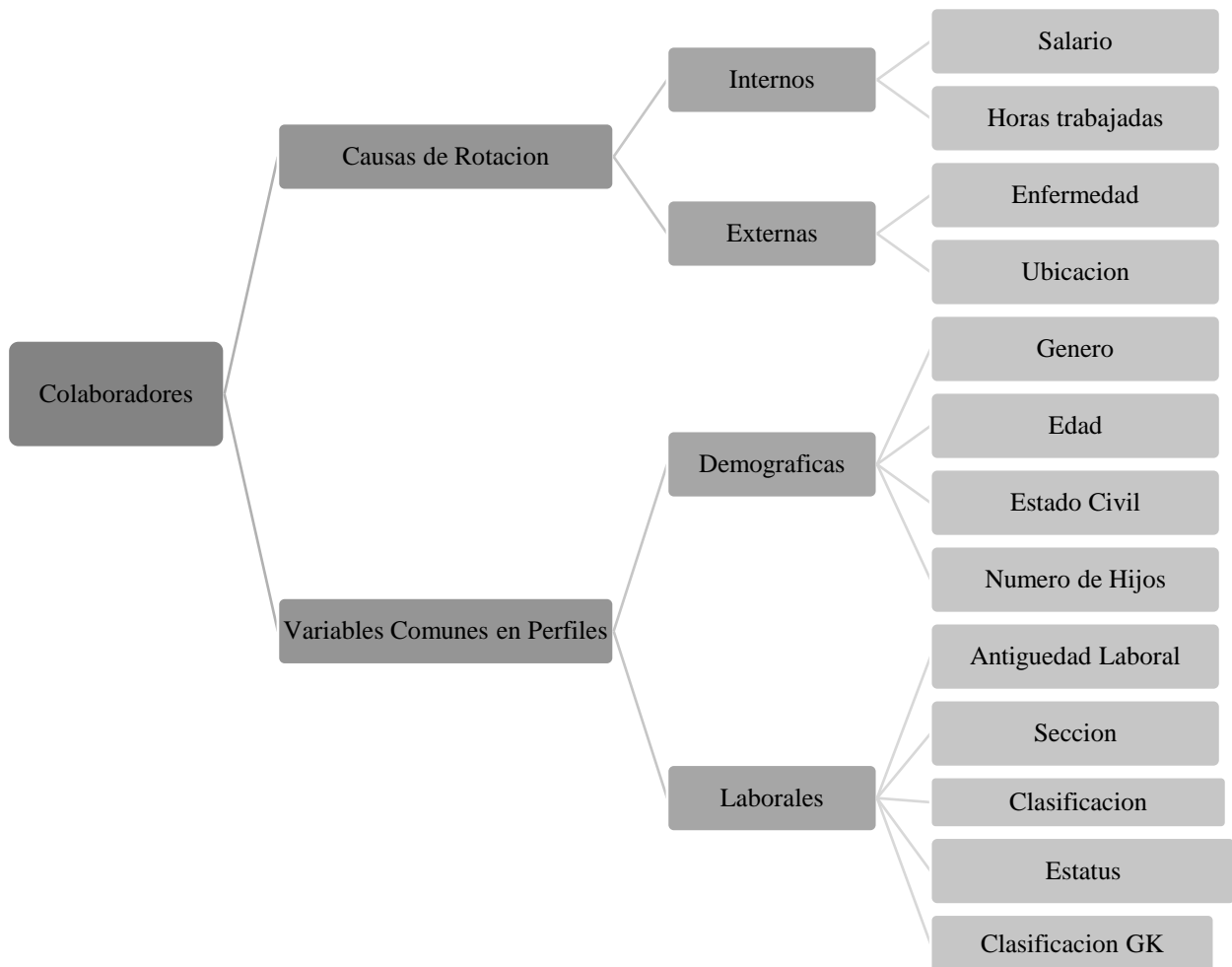


Ilustración 2 Esquema de Variables de Estudio

Fuente: Elaboración Propia

3.1.3 OPERACIONALIZACIÓN DE LAS VARIABLES

A continuación, se presenta la operacionalización de las variables de investigación

Tabla 2 Operacionalización de las Variables

Variables Dependientes	Variables Independientes	Definición Conceptual	Definición Operacional	Dimensiones	Ítem	Definición Ítems
Colaboradores candidatos(as) a dejar la organización	Causas de Rotación	Motivo por el cual los colaboradores de una organización dejan sus puestos de trabajo con cierto preaviso para que la empresa gestione su reemplazo	Describir de manera precisa los factores dentro de una organización que contribuyen a la rotación de colaboradores.	Internas	Salario	Remuneración económica que un empleado recibe a cambio de su trabajo en una organización.
					Horas trabajadas	Guarda las horas trabajadas por cada colaborador
			Describir de manera precisa los factores externos a la organización que contribuyen a la salida de los colaboradores de la organización.	Externas	Enfermedad	Alteración en el funcionamiento del cuerpo que causas síntomas malignos en la salud.
					Ubicación	Lugar donde está establecida una persona.

VARIABLES DEPENDIENTES	VARIABLES INDEPENDIENTES	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIONES	ÍTEM	DEFINICIÓN ÍTEMS
		Características o factores que son frecuentes en los colaboradores que tienden a rotar o cambiar de trabajo con mayor frecuencia.	Características estadísticas que tiene los colaboradores, con las que se pueden describir y generar un perfil.	Demográficas	Genero	Se refiere al género de una persona que puede ser masculino, femenino u otro.
					Edad	Edad de una persona.
	Estado Civil		Situación legal o social de una persona en relación con su matrimonio, como soltero, casado, divorciado o viudo.			
	Número de Hijos		Cantidad de personas dependientes que tiene un colaborador			
Variables Comunes en Perfiles			Aspectos que abarcan el entorno de trabajo, que	Laborales	Estatus	Afirma si el colaborador tiene un estatus activo o inactivo en la

VARIABLES DEPENDIENTES	VARIABLES INDEPENDIENTES	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIONES	ÍTEM	DEFINICIÓN ÍTEMS
			<p>incluyen condiciones físicas, emocionales y sociales, así como las relaciones entre los empleados y las características del empleo en sí</p>			planilla laboral
					Clasificación	Detalla si el colaborador trabaja directamente en los procesos de producción
					Código empleado	Es el número de ID de empleado con que se identifica cada colaborador
					Clasificación GK	Define si el colaborador trabaja administrativo o si pertenece al área operativa
					Sección	Nombre del área de trabajo dentro del proceso de producción y que es donde se encuentra asignado el colaborador
					Antigüedad Laboral	Años de antigüedad que tiene un colaborador donde ha formado conocimiento del trabajo.

Fuente: Elaboración Propia

3.1.4 HIPÓTESIS

Una hipótesis se puede definir como una proposición tentativa que busca explicar una relación entre variables en una investigación. Es una afirmación que se formula con base en conocimientos, para la siguiente investigación no se plantean hipótesis.

3.2 ENFOQUE Y MÉTODOS

A continuación, se presenta el enfoque y métodos utilizados en esta investigación, también se podrá encontrar el esquema de alcance, como también el diseño e instrumentos de la investigación.

3.2.1 ENFOQUE

El objetivo de la investigación es poder comprobar si se puede entrenar un modelo de predicción para evaluar el comportamiento de la rotación y verificar si existen variables y causas comunes para que esto surja por lo que nuestro enfoque será cuantitativo. En un enfoque cuantitativo se utiliza la recolección de datos para probar hipótesis con base en la medición numérica y el análisis estadístico, con el fin establecer pautas de comportamiento y probar teorías. (Sampieri, Fernandez Collado, & Baptista Lucio, 2014)

La investigación cuantitativa considera que el conocimiento debe ser objetivo, y que este se genera a partir de un proceso deductivo en el que, a través de la medición numérica y el análisis estadístico inferencial, se prueban hipótesis previamente formuladas. Este enfoque se comúnmente se asocia con prácticas y normas de las ciencias naturales y del positivismo. Este enfoque basa su investigación en casos “tipo”, con la intención de obtener resultados que permitan hacer generalizaciones. (AEL, n.d.)

3.2.2 DISEÑO

El diseño de la investigación es de tipo no experimental ya que este se realiza sin manipular deliberadamente las variables. Es decir, se trata de estudios en los que no hacemos variar en forma intencional las variables independientes para ver su efecto sobre otras variables. Lo que hacemos en la investigación no experimental es observar fenómenos tal como se dan en su contexto natural, para analizarlos. (Sampieri, Fernandez Collado, & Baptista Lucio, 2014)

En un estudio no experimental no se genera ninguna situación, sino que se observan situaciones ya existentes, no provocadas intencionalmente en la investigación por quien la realiza. En la investigación no experimental las variables independientes ocurren y no es posible manipularlas, no se tiene control directo sobre dichas variables ni se puede influir en ellas, porque ya sucedieron, al igual que sus efectos. (Sampieri, Fernandez Collado, & Baptista Lucio, 2014)

3.2.3 ALCANCE

El alcance de la investigación es correlacional. Un estudio correlacional tiene como finalidad conocer la relación o grado de asociación de exista entre dos o más variables en una muestra o contexto en particular. En ocasiones solo se analiza la relación entre dos variables, pero con frecuencia se ubican en el estudio vínculos entre, cuatro o más variables. (Sampieri, Fernandez Collado, & Baptista Lucio, 2014)

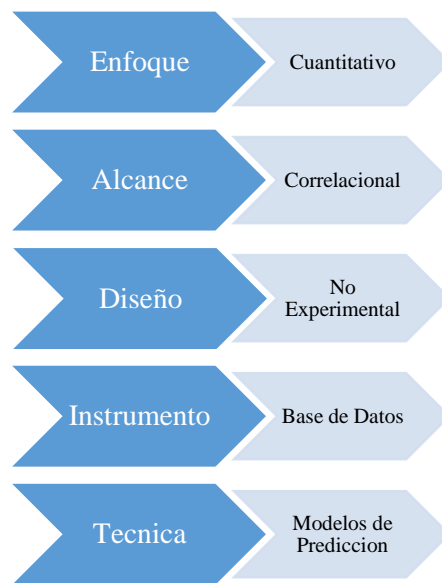


Ilustración 3 Enfoque Metodológico

Fuente: Elaboración Propia

3.3 DISEÑO DE LA INVESTIGACIÓN

A continuación, se detalla la población y muestra que se utilizarán para el proceso de la investigación. Como se mencionó en la sección anterior la investigación es de tipo no experimental y de alcance correlacional lo que significa que buscaremos qué relación tiene una o más variables.

Para nuestra investigación la población estará conformada por todos los colaboradores de la maquila de Siguatepeque en los Pinos.

3.3.1 POBLACIÓN

La población se refiere al conjunto completo de todos los elementos o individuos que comparten una característica común o que están sujetos a estudio en un contexto particular. La población representa el grupo completo que se desea investigar. (Trochim, 2006)

La población de la siguiente investigación comprende los 559 colaboradores que prestan servicios para la institución de los Pinos.

3.3.2 MUESTRA

La muestra es, en esencia, un subgrupo de la población. Digamos que es un subconjunto de elementos que pertenecen a ese conjunto definido en sus características al que llamamos población. (Sampieri, Fernández Collado, & Baptista Lucio, 2014)

Para la siguiente investigación no se realizó una muestra de población, ya que se utilizarán la información de los 559 colaboradores de institución, con el objetivo que los modelos puedan tomar todas las características necesarias.

3.3.3 TÉCNICAS DE MUESTREO

Ya que no se realizó una muestra de población, tampoco se realizarán técnicas de muestreo.

3.4 TÉCNICAS, INSTRUMENTOS Y PROCEDIMIENTOS APLICADOS

A continuación, se presentan los instrumentos y técnicas aplicadas en la investigación para la recolección de datos.

3.4.1 TÉCNICAS

Las técnicas de investigación son métodos o procedimientos utilizados para recolectar, analizar y obtener datos relevantes en un estudio de investigación. Estas técnicas se seleccionan y aplican de acuerdo con los objetivos de investigación, el tipo de datos requeridos y la naturaleza del fenómeno estudiado. (Romo, 1998)

En esta investigación se utilizarán como técnicas modelos de predicción, los modelos ayudarán a realizar un mejor análisis de cada una de variables y validar si existe alguna similitud

entre ellas. Con esto se llegará a la comprobación si cuando existe la rotación de personal se cuenta con variables comunes entre los colaboradores.

3.4.2 INSTRUMENTOS

Un instrumento de medición adecuado es aquel que registra datos observables que representan verdaderamente los conceptos o las variables que el investigador tiene en mente. En toda investigación cuantitativa aplicamos un instrumento para medir las variables contenidas en las hipótesis (y cuando no hay hipótesis simplemente para medir las variables de interés). (Sampieri, Fernandez Collado, & Baptista Lucio, 2014)

Para el instrumento de investigación se realizó una recolección de datos, tanto personales como de planillas históricos, con estos se conformó una base de datos con todas las variables de interés. Al contar con esta base de datos se crearán diferentes modelos de predicción para analizar si estas variables ayudaran a predecir la rotación antes de que suceda.

3.5 FUENTES DE INFORMACIÓN

Las fuentes de información son recursos que proporcionan datos, hechos, conocimientos o evidencias para respaldar la investigación o el estudio de un tema específico. Estas fuentes pueden ser de diferentes tipos y formatos, y se utilizan para obtener información relevante, actualizada y confiable. (Booth, Sutton, & Papaioannou, 2016)

Estas fuentes de información se pueden dividir en dos categorías:

3.5.1 FUENTES PRIMARIAS

Las fuentes primarias son los datos brutos o no analizados que se obtienen directamente de la realidad o de la fuente original de información. Son documentos o testimonios de primera mano que proporcionan información original, como entrevistas, cuestionarios, registros, observaciones directas, experimentos, entre otros. (Sampieri, Fernández Collado, & Baptista Lucio, 2014)

Para las fuentes primarias se cuenta con:

- Base de datos con variables demográficas y salariales de colaboradores.

3.5.2 FUENTES SECUNDARIAS

No se contó con fuentes secundarias en nuestro trabajo.

CAPÍTULO IV. RESULTADOS Y ANÁLISIS

4.1 INFORME DE PROCESO DE RECOLECCIÓN DE DATOS

El proceso de recolección de datos se llevó a cabo mediante la compilación de diversas bases de datos históricas, abarcando el periodo desde el 1 de abril de 2023 hasta el 31 de octubre de 2023. Esta compilación incorpora múltiples tablas de procesos internos, tales como la asistencia histórica, la compilación de planillas 2023, el histórico de atenciones médicas, el histórico de permisos acumulados y el registro de control de calidad. A partir de estas bases de datos, se generó una tabla más completa, detallando las características de cada empleado semana a semana, ya sea activo o inactivo.

Una vez presentada la información disponible, se procedió a explorar y verificar los datos, permitiendo así realizar un primer análisis para determinar la consistencia y la completitud de la información recopilada. En su forma inicial, la data consiste en 42 variables y 11,823 registros.

4.1.1 CRITERIOS DE EXCLUSION DE VARIABLES

De las 42 variables iniciales, se realizaron exclusiones estratégicas de 28 de ellas, dado que contenían información redundante. Estos elementos se consideraron poco relevantes para el enriquecimiento del entrenamiento del modelo predictivo. Esta decisión se fundamentó primeramente en la necesidad de simplificar el conjunto de datos para mejorar la eficiencia del modelo. Adicionalmente, la selección de las variables usadas en el algoritmo fue respaldada por la opinión de expertos en el área de Gestión Humana ya que según se lo manifiesta las investigaciones de (Melendres, Ramirez, Garcia, & Aranibar, 2017), las causas más comunes de rotación en las empresas se pueden categorizar por factores que responden a tres grandes perspectivas, la económica, Psicológicas y sociodemográficas, y dentro de las cuales se pueden mencionar factores, como salarios, turnos, falta de oportunidad para desarrollarse, pagos de salarios en base a méritos, cambios debido a la familia del trabajador, problemas de salud, cambios de estado civil, y variables de perfil sociodemográficas como nivel educativo, edad, antigüedad, entre otras. Tomando en consideración la opinión de los expertos se tomaron variables coherentes con estos conceptos mencionados para dar un respaldo teórico a nuestro modelo.

Tras este proceso de depuración, se logró definir un conjunto de variables finales que se consideraban más relevantes y específicas para el objetivo de entrenar el modelo de predicción de

rotación del personal.

Este enfoque selectivo en las variables promete mejorar la precisión y la capacidad predictiva del modelo al centrarse en factores más directamente relacionados con el fenómeno de la renuncia voluntaria en la organización.

Tabla 3 Tabla de Variables de Base de Datos

NO.	VARIABLE	TIPO DE DATOS	DESCRIPCIÓN
1	CODIGO EMPLEADO	String	Guarda los códigos de empleado de cada colaborador
2	NOMBRE EMPLEADO	String	Guarda los nombres completos de cada colaborador
3	SECCION	String	Nombre del área a donde se encuentra asignado el colaborador
4	GENERO	String	Genero de colaborador ya sea femenino o masculino
5	ESTATUS	String	Guarda si el colaborador tiene un estatus activo o inactivo
6	CLASIFICACION	String	Almacena si el colaborador trabaja directamente en los procesos de producción
7	CLASIFICACION GK	String	Almacena si el colaborador trabaja administrativo o si pertenece al área operativa
8	ESTADO CIVIL	String	Guarda el estado civil de cada uno de los colaboradores
9	ANTIGUEDAD	Double	Guarda los meses o años de antigüedad que tiene un colaborador en la empresa
10	EDAD	integer	Guarda la edad del colaborador
11	UBICACIÓN	String	Define el lugar de residencia del colaborador
12	HIJOS	integer	Guarda la cantidad de hijos que tiene el colaborador
13	HORAS ORDINARIAS	Double	Guarda las horas trabajadas por cada colaborador
14	SALARIO NETO A PAGAR	Double	Guarda el salario neto que se pago en planilla

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

Para las variables numéricas se recodificaron de la siguiente manera:

- Edad:
 - Sin Edad
 - Menor o igual a 20
 - Edad 21 a 30
 - Edad 31 a 40
 - Edad 41 a 50
 - Mayor o igual a 51
- Antigüedad:
 - De 0 a 3 Meses
 - De 4 a 6 Meses
 - De 7 a 9 Meses
 - De 10 a 12 Meses
 - Mayor a 12 Meses
- Hijos:
 - No tiene hijos
 - Rango de 0 a 2 hijos
 - Rango de 3 a 5 hijos
 - Rango de 6 a 8 hijos
 - Mayores a 8 hijos
- Horas Ordinarias:
 - De 0 a 22 horas
 - De 23 a 44 horas
 - De 45 a 66 horas
 - De 67 a 88 horas
- Salario Neto Por Pagar:
 - Salario Prom 0 a 3000 LPS
 - Salario Prom 3001 a 6000 LPS
 - Salario Prom 6001 a 9000 LPS

- Salario Prom Mayor 9000 LPS"

4.2 RESULTADOS Y ANÁLISIS DE LAS TÉCNICAS APLICADAS

Luego del proceso realizado para la recolección de datos, se prosiguió con el análisis de los datos obtenidos, esto con el objetivo de realizar un análisis exploratorio de los datos para comprender como se encuentra la empresa y con que variables se cuenta que puedan ser de interés para la investigación actual.

4.2.1 ANÁLISIS CUANTITATIVO

El análisis exploratorio de datos, realizado con la base de datos proporcionada, se inicia abordando una de las preguntas clave de esta investigación: la cantidad de colaboradores activos en la institución. Actualmente, se registra un total de 559 colaboradores, de los cuales 420 son del género femenino y 139 del género masculino. Es notable que la proporción entre ambos géneros revela una significativa disparidad, con el género femenino triplicando al masculino. Este fenómeno puede explicarse por la tendencia en el rubro de la maquila, donde la contratación de personal se inclina mayoritariamente hacia el género femenino. Este primer hallazgo sienta las bases para profundizar en el análisis y comprender las dinámicas de género dentro de la institución.

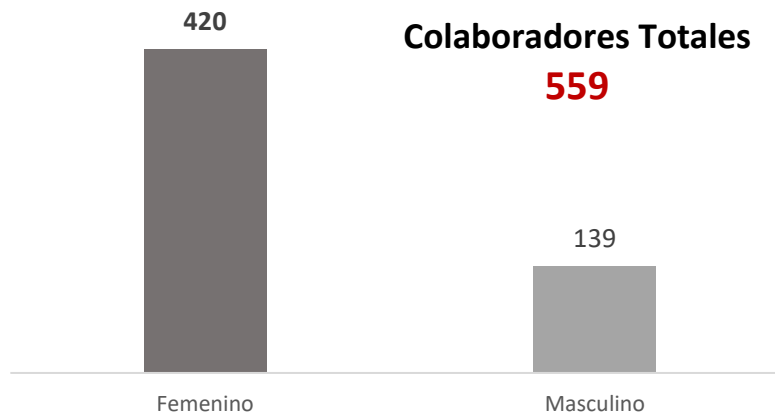


Ilustración 4 Total de Colaboradores

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

Una variable crucial que será empleada en los modelos de predicción es el rango de edad, proporcionando insights valiosos sobre la distribución demográfica de los colaboradores. Según la

gráfica adjunta, de los 559 colaboradores, la mayor concentración se encuentra en el rango de 21 a 30 años, con un total de 230 individuos. Le sigue el grupo de colaboradores con edades comprendidas entre 31 y 40 años, contabilizando 139 personas, y finalmente, 105 colaboradores se sitúan en la franja de edad entre 41 y 50 años. Este análisis detallado de la distribución por rangos de edad contribuirá significativamente a la construcción de modelos de predicción más precisos y contextualizados, permitiendo una comprensión más profunda de las dinámicas laborales en relación con la edad en nuestra institución.

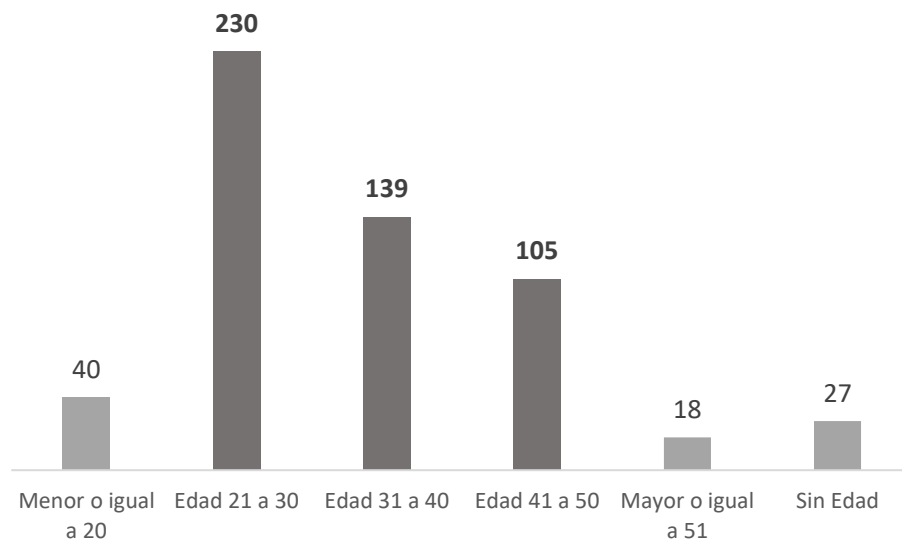


Ilustración 5 Colaboradores por Rango de Edad

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

La gráfica revela que 339 colaboradores tienen un estado civil de soltero. Sin embargo, es crucial destacar que el hecho de estar soltero no implica necesariamente la ausencia de familiares dependientes. Por otro lado, se observa que 142 colaboradores se encuentran en un estado civil de unión libre. Esta variable proporciona una valiosa herramienta para analizar los perfiles de los colaboradores y comprender mejor los factores que pueden influir en la rotación de personal. Se plantea la interrogante de si los colaboradores solteros encuentran más fácilmente motivos para abandonar la empresa, o si, por el contrario, son aquellos en uniones libres quienes muestran una mayor propensión a permanecer en la organización. Este análisis puede arrojar luz sobre las dinámicas de retención de personal y contribuir a estrategias más efectivas de gestión del talento.

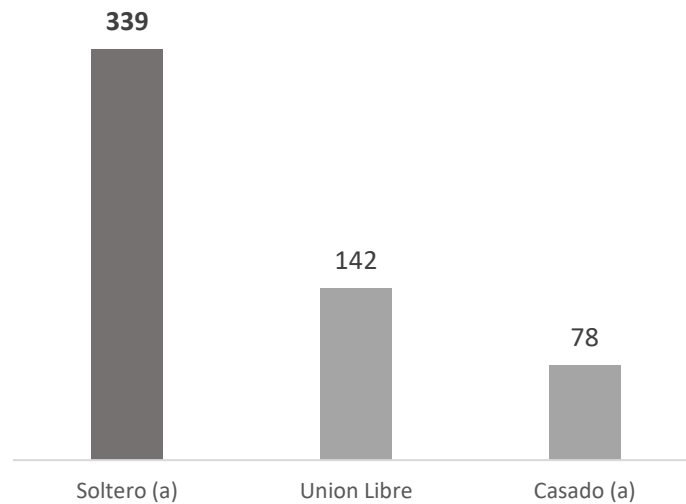


Ilustración 6 Colaboradores por Estado Civil

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

En el seno de nuestra institución, conformada por 559 colaboradores activos, se revela una interesante diversidad en la composición familiar. Entre este grupo, 264 colaboradores destacan al tener un rango de 1 a 2 hijos, mientras que 168 de ellos no cuentan con descendencia. Adicionalmente, se nota un segmento significativo de 123 colaboradores que disfrutan de familias más numerosas, con un rango de 3 a 5 hijos. Estos datos proporcionan una visión panorámica reveladora de la variabilidad en el tamaño de las familias dentro de nuestra organización, destacando la riqueza y complejidad de las experiencias familiares entre nuestros colaboradores activos.

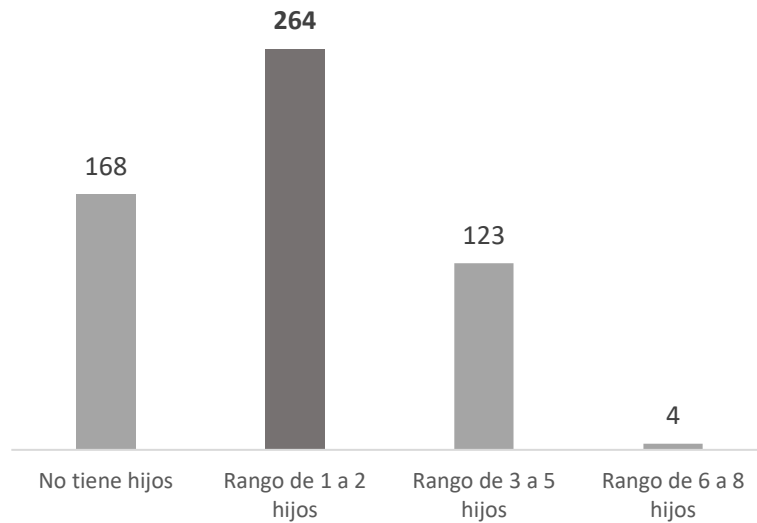


Ilustración 7 Colaboradores por Cantidad de Hijos

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

De los 559 colaboradores que conforman nuestra institución, se destaca que un significativo grupo de 395 posee una antigüedad en el rango de 7 a 9 meses. Este dato es indicativo de una fuerza laboral experimentada y calificada, lo cual repercute positivamente en la eficiencia operativa. La notable concentración de colaboradores con esta antigüedad sugiere que la mayoría ha alcanzado un nivel óptimo de aprendizaje en sus roles, lo que contribuye a la continuidad y fluidez de la producción. Esta experiencia acumulada no solo fortalece la estabilidad en el desempeño laboral, sino que también puede tener un impacto positivo en la calidad y eficiencia de los procesos, minimizando posibles interrupciones debido a la familiaridad y pericia de la mano de obra en la institución.

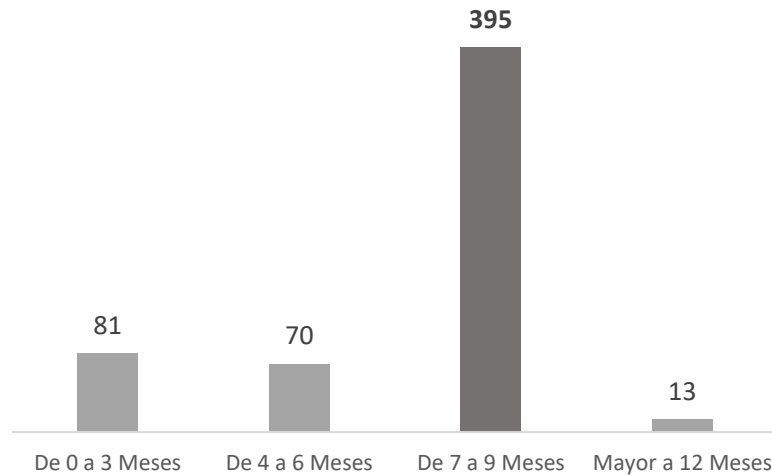


Ilustración 8 Colaboradores por Antigüedad

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

La gráfica siguiente revela claramente que las secciones con la mayor cantidad de colaboradores son la línea 6 y la línea 1. Este fenómeno se explica por el hecho de que estas secciones son fundamentales para la manufactura de los productos que nuestra institución realiza. La concentración de colaboradores en estas áreas específicas indica la importancia estratégica de las líneas 6 y 1 en el proceso productivo. Esta distribución probablemente refleje la necesidad de recursos humanos calificados en las etapas cruciales de la fabricación, subrayando la relevancia de estas secciones en la operación general de la institución. Este análisis puede orientar estrategias de gestión y asignación de recursos para optimizar la eficiencia en las áreas clave de producción.

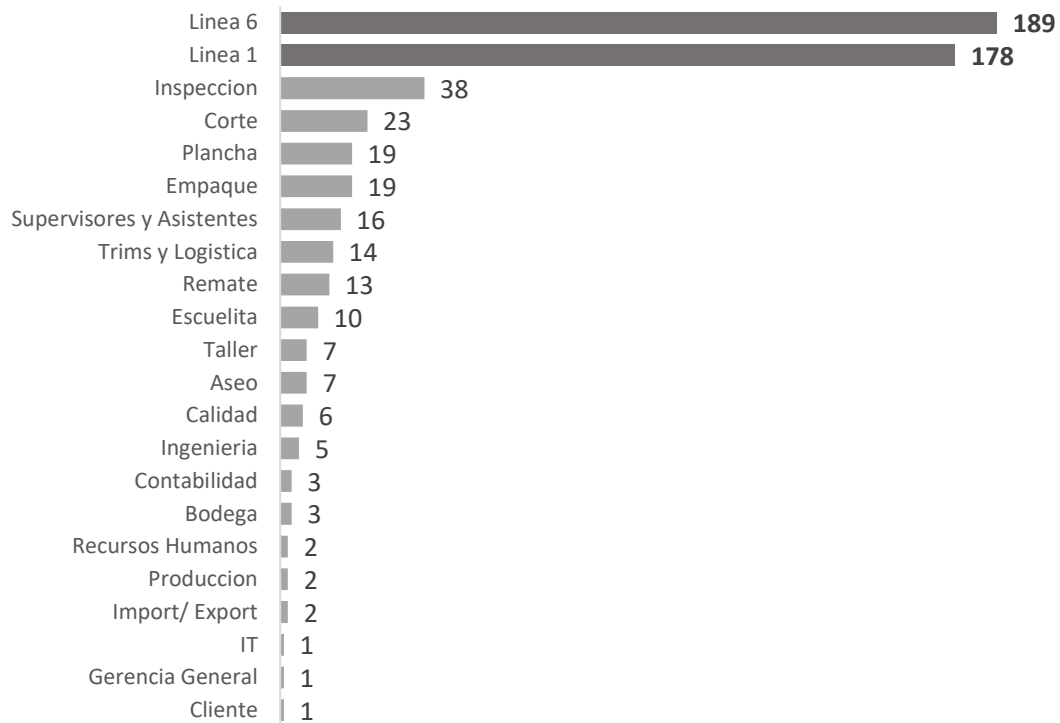


Ilustración 9 Colaboradores por Sección

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

La gráfica anterior proporciona una perspectiva detallada sobre el comportamiento de las salidas en nuestra institución desde diciembre de 2022 hasta la fecha actual. Se destaca que existen variaciones significativas en la cantidad de salidas a lo largo de este período, con meses en los que el número asciende más notablemente que en otros. Este patrón de fluctuación mensual en las salidas sugiere la existencia de factores estacionales, eventos específicos o tendencias que pueden influir en las decisiones de salida de los colaboradores. Analizar estos picos y valles en las salidas puede ofrecer valiosa información para comprender mejor las dinámicas laborales y, en última instancia, informar estrategias de retención y gestión del personal.

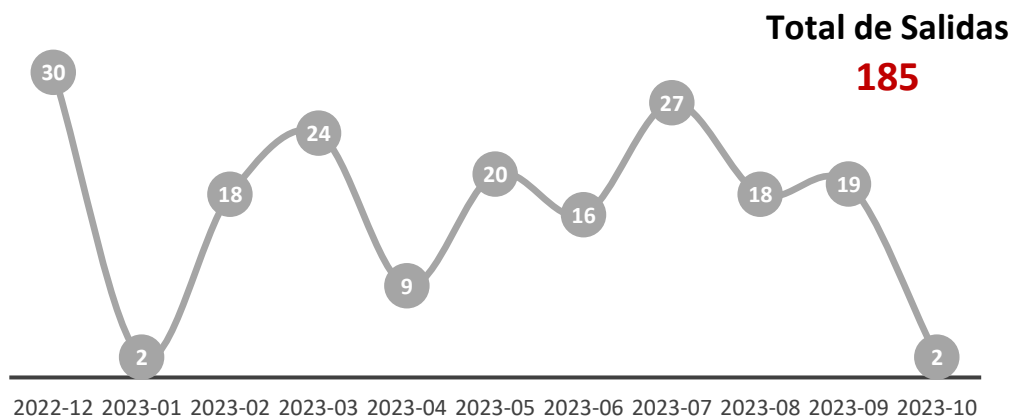


Ilustración 10 Comportamiento de Salidas

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

Como se ha evidenciado en gráficas previas, la preponderancia del género femenino en la plantilla de colaboradores se refleja también en las salidas de la institución. De un total de 185 salidas registradas, se destaca que 142 de estas corresponden al género femenino, mientras que las salidas por parte del género masculino ascienden a 43. Esta discrepancia en las cifras de salidas entre ambos géneros puede señalar patrones y dinámicas específicas que merecen un análisis más detenido. Comprender las razones detrás de estas salidas diferenciadas es crucial para desarrollar estrategias efectivas de retención y gestión del talento, con el objetivo de fomentar un entorno laboral equitativo y sostenible.

De las 185 salidas registradas en la institución, se observa una coincidencia notable con las secciones de mayor cantidad de colaboradores, específicamente la línea 6 y la línea 1. Estas secciones, que albergan la manufactura de los productos de la institución, no solo cuentan con la mayor cantidad de personal, como se ha visto en gráficas anteriores, sino que también experimentan el mayor número de salidas. Esta asociación entre la concentración de colaboradores y las salidas podría indicar la existencia de desafíos específicos o factores influenciadores dentro de estas áreas de trabajo. Analizar más a fondo esta relación puede proporcionar insights valiosos para implementar estrategias de retención personalizadas y mejorar las condiciones laborales en estas secciones críticas para la operación de la institución.

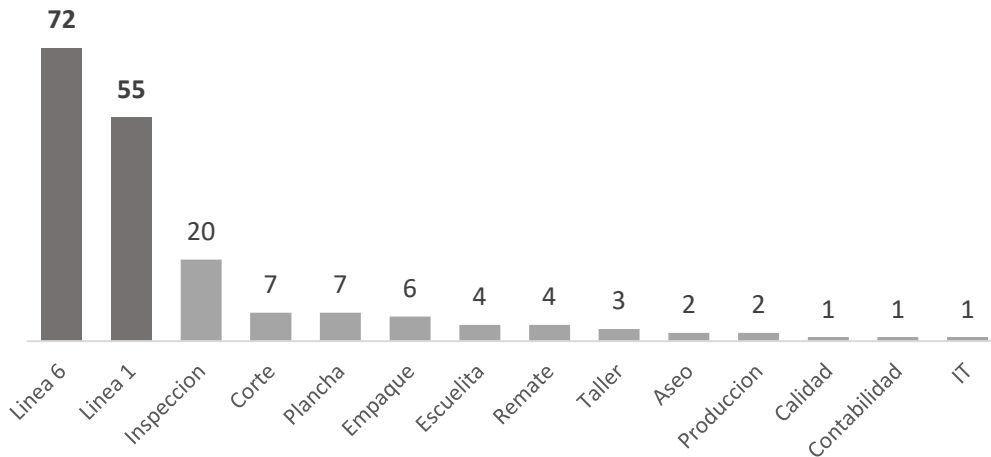


Ilustración 11 Salidas por Sección

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

De las 185 salidas registradas en la institución, se destaca que la mayoría de los colaboradores que dejaron la organización no tenían hijos. Específicamente, de este total, se observa que 170 salidas corresponden a colaboradores que no contaban con dependientes. Este dato sugiere que la ausencia de hijos podría ser un factor relevante en las decisiones de salida de los colaboradores. Analizar esta información aporta una perspectiva valiosa para comprender las dinámicas laborales y los posibles factores que pueden influir en las decisiones de los empleados de abandonar la institución.

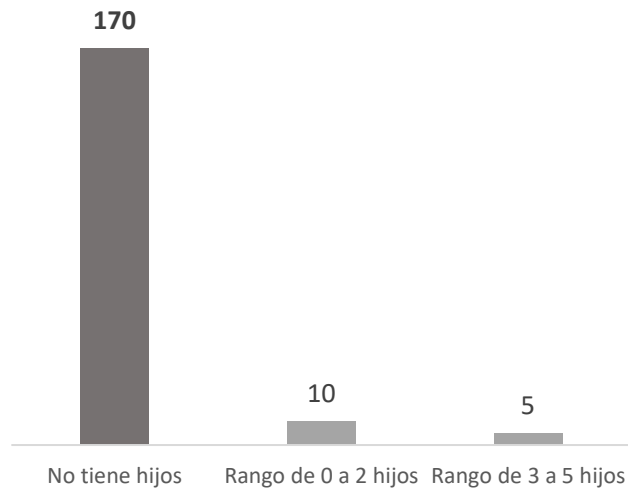


Ilustración 12 Salidas por Rango de Hijos

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

En la siguiente gráfica, se destaca que 94 de los excolaboradores tenían una antigüedad promedio de 0 a 3 meses. Este patrón sugiere que un número significativo de salidas se produjo durante el período de prueba, indicando que estos colaboradores dejaron la institución antes de completar un tiempo sustancial de servicio. Además, se observa que 49 excolaboradores tenían una antigüedad en el rango de 7 a 9 meses. Este grupo, aunque con una mayor permanencia en comparación con los de menor antigüedad, también experimentó una cantidad considerable de salidas. Analizar estos datos de antigüedad proporciona perspectivas valiosas sobre los momentos críticos en la relación laboral en los que se producen las salidas

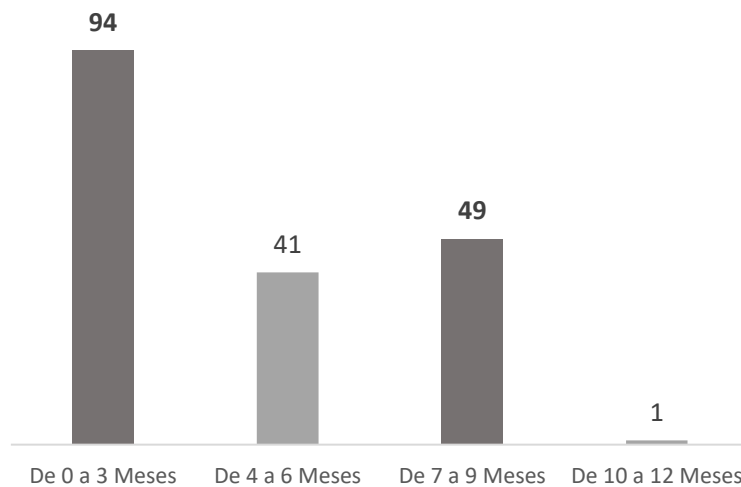


Ilustración 13 Salidas por Antigüedad

Fuente: Elaboración Propia. Base de Datos Recursos Humanos Los Pinos Apparel

4.2.2 ANÁLISIS CUALITATIVO

Después de realizar un análisis exploratorio de los datos se encontraron ciertas características, estas fueron obtenidas por ciertos patrones similares en las salidas que ha tenido en la institución, a continuación, se enlistan los hallazgos encontrados:

1. La mayoría de las salidas generadas han sido por el género femenino, esto en buena parte es porque en la institución la mayoría de los colaboradores es de género femenino debido al rubro y a lo que la institución genera como producto.
2. Las salidas se generan con mayor frecuencia en las áreas de manufactura, las líneas 6 y línea 1 que es donde se prepara y se fabrica el producto.

3. También se observó que la mayoría de las salidas no sobrepasaron los seis meses de antigüedad.
4. Se pudo agrupar las razones de salida en:
 - a. Renuncia voluntaria.
 - b. Salud o problemas familiares.
 - c. Bajo rendimiento o mala conducta.
 - d. Otros como viajes u otras ofertas de trabajo.
5. Otra de las características es que el estado civil que mayor predomina son los solteros y que no cuentan con dependientes, es decir que no tiene hijos.
6. En cuanto a las horas que trabajaron antes de salir de la institución, se observó que si se cumplía con las horas de trabajo semanales
7. También se observó que, para las renuncias con motivos de salud o personales, antes de que la salida ocurriera, el colaborador pedía permisos o se incapacitaba y con esto no cumplía con su objetivo semanal.

También se aplicó en la institución una pequeña encuesta donde se preguntaba a cada colaborador si estará en el siguiente año, esta encuesta es aplicada anualmente con el objetivo de saber con qué porcentaje de mano de obra se cuenta para el siguiente año y ver la antigüedad que tiene. También para escuchar las necesidades de los colaboradores de la institución. De esta pequeña encuesta se logró observar que las características más comunes de los que contestaron que si saldrán de la compañía son las siguientes:

1. Todas son colaboradoras de género femenino que actualmente se encuentran en la sección de línea 6 y 1, como lo mencionamos anteriormente esta sección es la encargada de la manufactura del producto.
2. Tienen un promedio de antigüedad de 9 meses, porque lo que podemos clasificarlas como una mano de obra con experiencia.
3. El promedio de hijos de las colaboradoras se encuentre entre 1 a 4 hijos y se encuentran solteras o en unión libre.

4. Tienen en un promedio de edad entre los 20 y 48 años.
5. En los últimos meses han cumplido con sus horas de trabajo laborales.
6. En algunos de los casos se han solicitado permisos y las razones más comunes han sido por enfermedad, por cuidado de algún pariente con mayor relevancia hijo o por algún otro problema personal.

4.3 ANÁLISIS INFERENCIAL Y MODELOS APLICADOS

Como se mencionó en el siguiente informe, la siguiente investigación busca poder validar si existe un modelo que pueda predecir la propensión de los colaboradores a renunciar. De acuerdo con esto, se consideran los siguientes modelos de predicción:

1. Regresión Logística
2. Naive Bayes
3. Árbol de Decisión
4. Random Forest

Para la creación de los modelos se utilizó la herramienta KNIME, en donde se cargó la base de datos, se realizó la limpieza de los datos y se crearon los cuatro modelos mencionados anteriormente, para estos modelos se utilizó el 80% de los registros para el entrenamiento y el 20% para evaluar los modelos.

En el capítulo 3, se seleccionaron diversas variables para ser utilizadas en el análisis, muchas de las cuales son categóricas. Entre ellas se encuentran la sección, género, estado civil, cantidad de hijos, rango de horas trabajadas, rango de salario promedio, rango de antigüedad, y la variable target que indica si el colaborador se encuentra activo o inactivo.

En el proceso de preparación de datos para los modelos, se exploraron dos enfoques distintos. El primero consiste en mantener las variables categóricas en su forma original, mientras que el segundo implica transformar estas variables categóricas en variables dummy. Esta última metodología, que implica asignar valores binarios a las distintas categorías, permite una representación numérica más adecuada para su integración en modelos de machine learning. Ambos enfoques fueron considerados para evaluar cómo la elección del método de codificación

afecta el rendimiento y la interpretación de los modelos utilizados en el análisis.

Para la siguiente investigación se realizaron los cuatro modelos de predicción mencionados anteriormente, estos se entrenaron de dos formas distintas:

Modelo con Variables Categóricas

Para desarrollar el modelo con variables categóricas, se emplearon nodos de Excel Reader para importar la base de datos proporcionada. Posteriormente, se llevaron a cabo conversiones de variables con el objetivo de optimizar al máximo la capacidad predictiva del modelo. Una vez completada la preparación de la información, se procedió a la partición del conjunto de datos para entrenar el modelo.

El proceso de entrenamiento y predicción de los modelos se llevó a cabo mediante nodos especializados para cada uno de ellos. Este análisis culminó con la generación de una matriz de confusión que presenta de manera clara los resultados obtenidos por el modelo.

Modelo con Variables Binarias

Para construir el siguiente modelo, se emplearon nodos de Excel Reader con el propósito de importar la base de datos suministrada. Posteriormente, se realizaron conversiones específicas de variables con el objetivo de potenciar al máximo la capacidad predictiva del modelo. Una vez que la información estaba preparada, se llevó a cabo una transformación de las variables categóricas, representándolas en formato binario.

El proceso continuó con la partición del conjunto de datos para el entrenamiento del modelo. Después de esta fase, se implementaron nodos especializados de entrenamiento y predicción. Como resultado final, se generó una matriz de confusión que presenta de manera detallada los resultados obtenidos por el modelo.

A continuación, se presentan cada uno de los modelos con sus resultados.

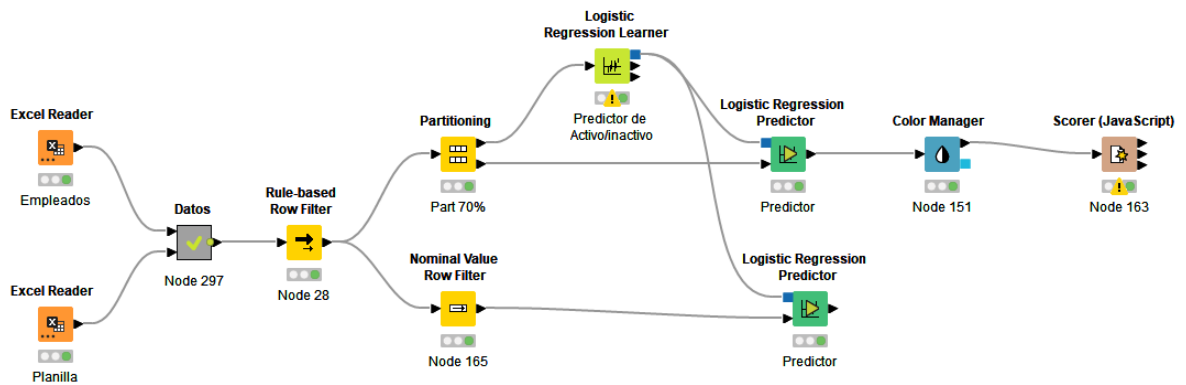


Ilustración 14 Regresión Logística – Variables Categóricas

Fuente: Elaboración Propia

Matriz de confusion

Confusion Matrix

Rows Number : 131	ACTIVO (Predicted)	INACTIVO (Predicted)	
ACTIVO (Actual)	109	1	99.09%
INACTIVO (Actual)	1	20	95.24%
	99.09%	95.24%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
ACTIVO	109	1	20	1	99.09%	99.09%	99.09%	95.24%	99.09%
INACTIVO	20	1	109	1	95.24%	95.24%	95.24%	99.09%	95.24%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
98.47%	1.53%	0.943	129	2

Ilustración 15 Matriz de Regresión Logística – Variables Categóricas

Fuente: Elaboración Propia

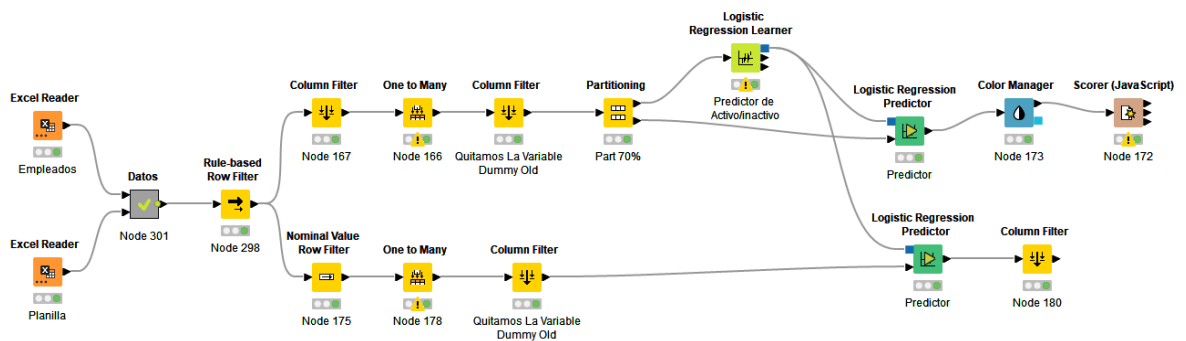


Ilustración 16 Regresión Logística – Variables Binarias

Fuente: Elaboración Propia

Matriz de confusion

Confusion Matrix			
Rows Number : 131	ACTIVO (Predicted)	INACTIVO (Predicted)	
ACTIVO (Actual)	109	1	99.09%
INACTIVO (Actual)	3	18	85.71%
	97.32%	94.74%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
ACTIVO	109	3	18	1	99.09%	97.32%	99.09%	85.71%	96.20%
INACTIVO	18	1	109	3	85.71%	94.74%	85.71%	99.09%	90.00%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
96.95%	3.05%	0.882	127	4

Ilustración 17 Matriz de Regresión Logística – Variables Binarias

Fuente: Elaboración Propia

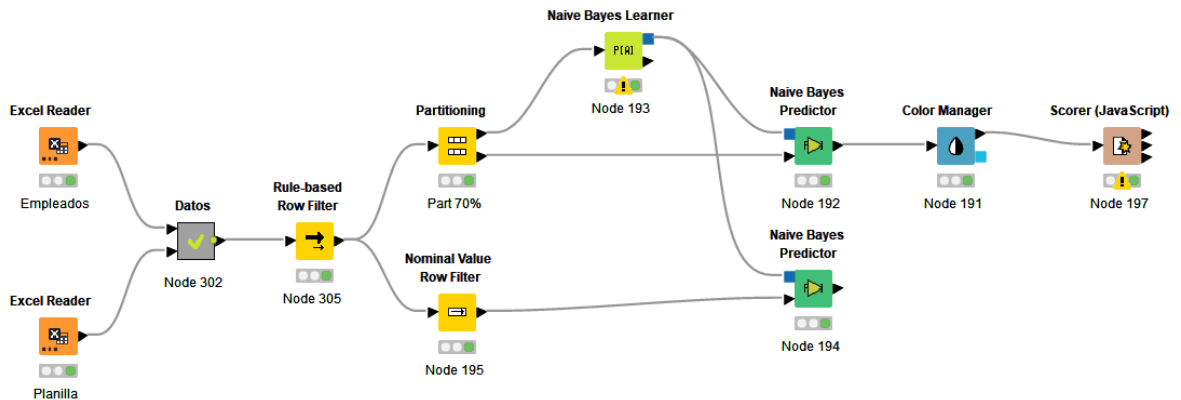


Ilustración 18 Naive Bayes – Variables Categóricas

Fuente: Elaboración Propia

Matriz de confusion

Confusion Matrix			
Rows Number : 131	ACTIVO (Predicted)	INACTIVO (Predicted)	
ACTIVO (Actual)	108	2	98.18%
INACTIVO (Actual)	1	20	95.24%
	99.08%	90.91%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
ACTIVO	108	1	20	2	98.18%	99.08%	98.18%	95.24%	98.63%
INACTIVO	20	2	108	1	95.24%	90.91%	95.24%	98.18%	93.02%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
97.71%	2.29%	0.917	128	3

Ilustración 19 Matriz de Naive Bayes – Variables Categóricas

Fuente: Elaboración Propia

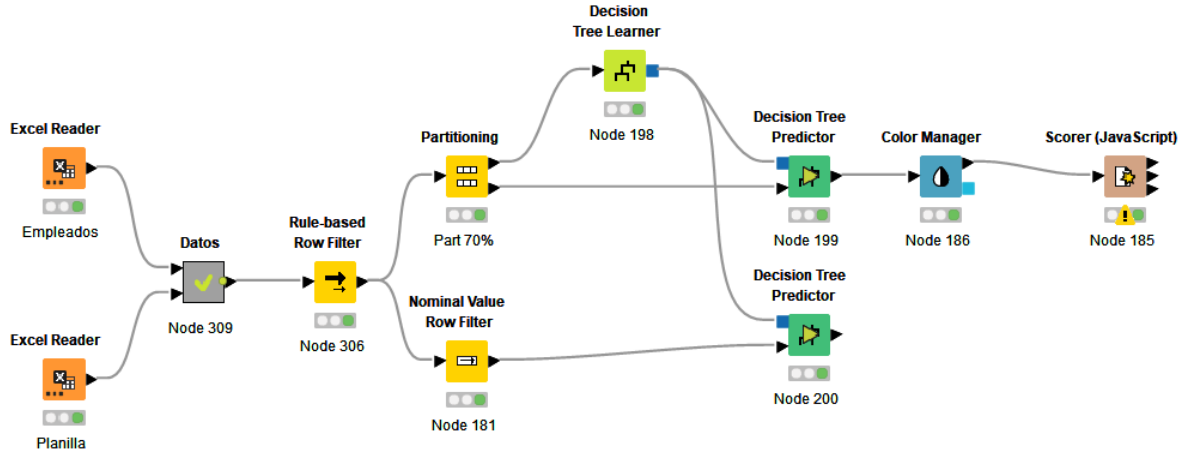


Ilustración 20 Árbol de Decisión – Variables Categóricas

Fuente: Elaboración Propia

Matriz de confusion

Confusion Matrix

Rows Number : 130	ACTIVO (Predicted)	INACTIVO (Predicted)	
ACTIVO (Actual)	109	1	99.09%
INACTIVO (Actual)	2	18	90.00%
	98.20%	94.74%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
ACTIVO	109	2	18	1	99.09%	98.20%	99.09%	90.00%	98.64%
INACTIVO	18	1	109	2	90.00%	94.74%	90.00%	99.09%	92.31%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
97.69%	2.31%	0.910	127	3

Ilustración 21 Matriz de Árbol de Decisión – Variables Categóricas

Fuente: Elaboración Propia

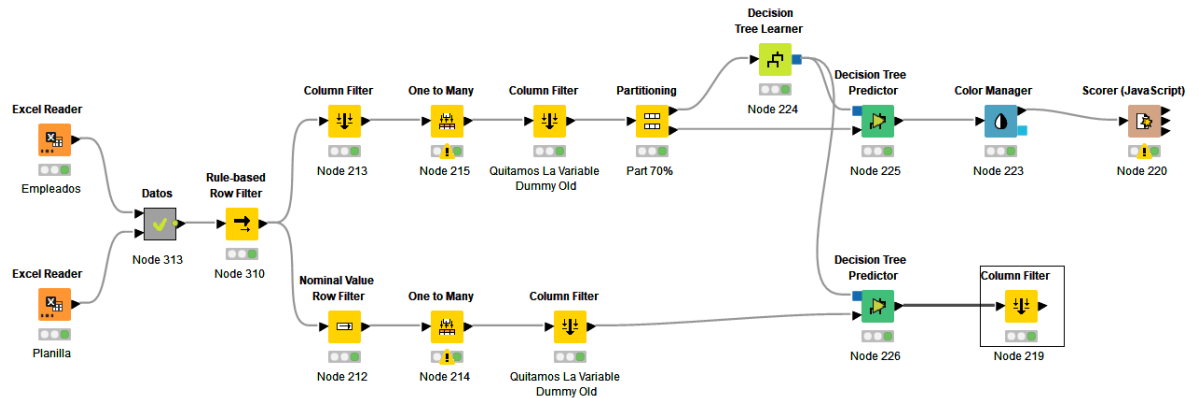


Ilustración 22 Árbol de Decisión – Variables Binarias

Fuente: Elaboración Propia

Matriz de confusión

Confusion Matrix

Rows Number : 131	ACTIVO (Predicted)	INACTIVO (Predicted)	
ACTIVO (Actual)	109	1	99.09%
INACTIVO (Actual)	3	18	85.71%
	97.32%	94.74%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
ACTIVO	109	3	18	1	99.09%	97.32%	99.09%	85.71%	98.20%
INACTIVO	18	1	109	3	85.71%	94.74%	85.71%	99.09%	90.00%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
96.95%	3.05%	0.882	127	4

Ilustración 23 Matriz de Árbol de Decisión – Variables Binarias

Fuente: Elaboración Propia

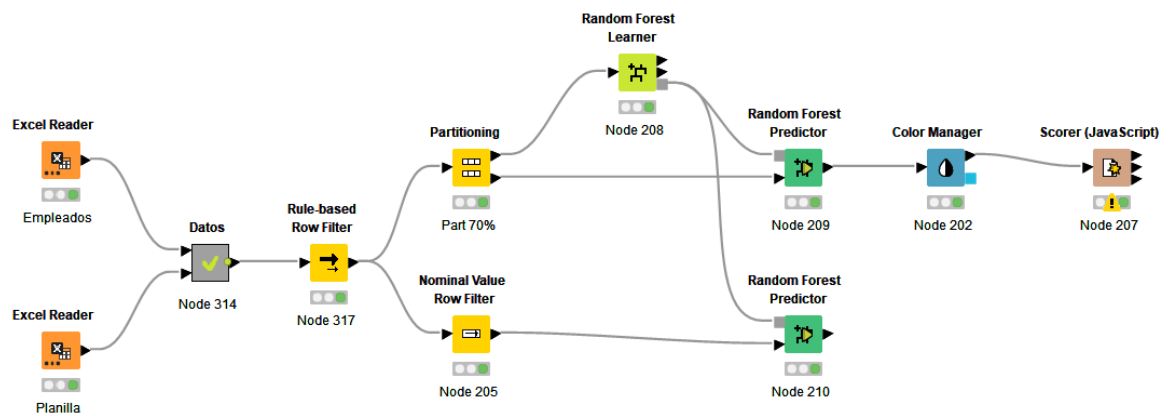


Ilustración 24 Random Forest – Variables Categóricas

Fuente: Elaboración Propia

Matriz de confusión

Confusion Matrix

Rows Number : 131	ACTIVO (Predicted)	INACTIVO (Predicted)	
ACTIVO (Actual)	109	1	99.09%
INACTIVO (Actual)	2	19	90.48%
	98.20%	95.00%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
ACTIVO	109	2	19	1	99.09%	98.20%	99.09%	90.48%	98.64%
INACTIVO	19	1	109	2	90.48%	95.00%	90.48%	99.09%	92.68%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
97.71%	2.29%	0.913	128	3

Ilustración 25 Matriz de Random Forest – Variables Categóricas

Fuente: Elaboración Propia

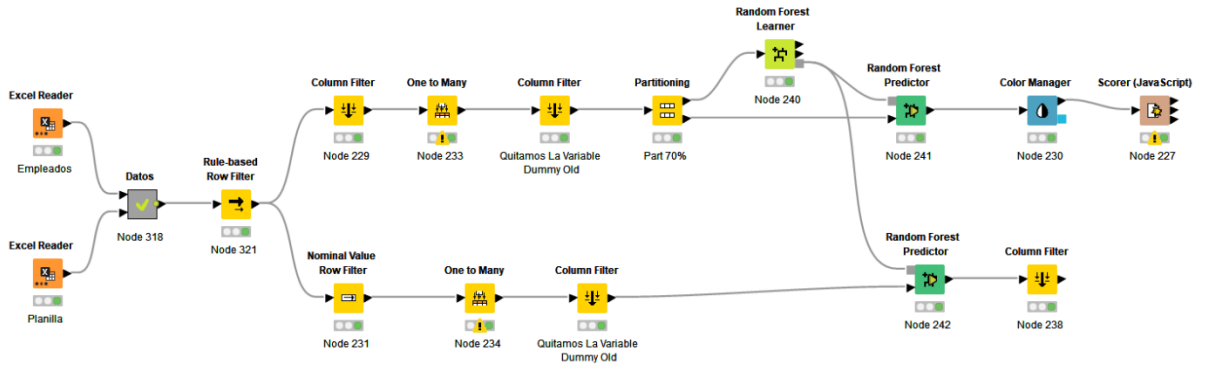


Ilustración 26 Random Forest – Variables Binarias

Fuente: Elaboración Propia

Matriz de confusion

Confusion Matrix

Rows Number : 131	ACTIVO (Predicted)	INACTIVO (Predicted)	
ACTIVO (Actual)	110	0	100.00%
INACTIVO (Actual)	3	18	85.71%
	97.35%	100.00%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
ACTIVO	110	3	18	0	100.00%	97.35%	100.00%	85.71%	98.65%
INACTIVO	18	0	110	3	85.71%	100.00%	85.71%	100.00%	92.31%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's Kappa (κ)	Correctly Classified	Incorrectly Classified
97.71%	2.29%	0.910	128	3

Ilustración 27 Matriz de Random Forest – Variables Binarias

Fuente: Elaboración Propia

Después de realizar el entrenamiento de cada uno de los modelos de predicción se resumen cada uno de los resultados en la siguiente tabla:

Tabla 4 Resultados de Modelos de Predicción

Tipo de Variables	Estadísticas	Regresión Logística	Naive Bayes	Árbol de Decisión	Random Forest
Variables Categóricas	Recall	95.24%	95.24%	90.00%	90.48%
	Precision	95.24%	90.91%	94.74%	95.00%
	Sensitivity	95.24%	95.24%	90.00%	90.48%
	Specificity	99.09%	98.18%	99.09%	99.09%
	F-Measure	95.24%	93.02%	92.31%	92.68%
Variables Dummies	Recall	85.71%		85.71%	85.71%
	Precision	94.74%		94.74%	100.00%
	Sensitivity	85.71%		85.71%	85.71%
	Specificity	99.09%		99.09%	100.00%
	F-Measure	90.00%		90.00%	92.31%

Fuente: Elaboración Propia

Se destaca que los tres modelos con la precisión más destacada en predicciones son la Regresión Logística con variables categóricas, alcanzando un impresionante 95.24%. Le sigue de cerca el Random Forest con variables categóricas, con un sólido 95.00%, y el Árbol de Decisión con variables categóricas, que se mantiene fuerte con un 94.74%. En contraste, el modelo con la menor exactitud fue el Naive Bayes con variables categóricas, registrando un 90.91%. También es notable que los cuatro modelos mencionados anteriormente logran un recall superior al 90%, lo que indica que clasifican correctamente la gran mayoría de los valores positivos.

A lo largo de la investigación, se planteó la interrogante sobre si las variables demográficas y laborales podrían influir y revelar patrones de similitud que permitieran generar un perfil para identificar posibles colaboradores propensos a rotar, ya sea dentro o fuera de la institución.

En base a los resultados, se puede concluir que tanto el modelo de Regresión Logística como el de Random Forest ofrecen herramientas eficaces para identificar colaboradores con perfiles similares, lo que podría ayudar en la predicción de quienes podrían salir de la institución. Estos modelos se presentan como valiosas herramientas para la gestión del talento y la retención de personal al permitir anticipar posibles movimientos dentro de la organización.

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

- Se concluye que es viable lograr una precisión del 80% en un modelo de predicción de clasificación y que es factible para estimar la probabilidad de rotación del personal. Esta precisión en la predicción de los algoritmos se fundamenta en el peso que aporta las diversas variables extraídas del perfil de colaboradores y que estuvieron presentes en los colaboradores que experimentaron rotación en el año 2023, así como en el análisis del comportamiento histórico de los colaboradores activos en la institución. Dentro de la variedad de modelos sometidos a evaluación, destaca la regresión logística y Random Forest, ambos alcanzando una destacada precisión del 97%. La implementación exitosa del algoritmo de predicción ha demostrado ser una herramienta sumamente valiosa como complemento a la gestión del departamento de recursos humanos, proporcionando una base sólida para la toma de decisiones informada y estratégica.
- La generación de perfiles ha permitido identificar de manera clara la variedad de características internas o externas y económicas, psicológicas, demográficas subyacentes a la realidad de los empleados que rotaron en 2023, y es notorio que las colaboradoras involucradas mayormente ostentan el estado civil de solteras. Así como las causas expresadas en su entrevista de salida donde en primer lugar, la presión laboral se erige como un factor significativo, ejerciendo una fuerza que conduce a la renuncia de algunos colaboradores. Problemas de salud y complicaciones familiares, especialmente prominentes entre el personal femenino, se perfilan como motivos recurrentes, atribuibles en gran medida a las responsabilidades familiares. Este análisis exhaustivo proporciona una comprensión más profunda de los factores que influyen en las salidas específicas, ofreciendo a la empresa una base sólida para implementar estrategias orientadas a mejorar la retención en estas áreas críticas.
- La elaboración de perfiles ha permitido identificar de manera sistemática patrones recurrentes asociados a las causas laborales que provocaron salidas de

colaboradoras de la empresa. Un análisis detallado revela que las renunciaciones están predominantemente concentradas en la Sección de Línea 6-camisas en primera instancia y, en segundo lugar, en la Línea 1-Pantalón. Estas salidas se caracterizaron por colaboradores con perfiles de una antigüedad laboral inferior a 6 meses. Un hallazgo constante en estas situaciones es que, de manera consistente, estas colaboradoras no lograban cumplir con su cantidad de horas laboradas con respecto a la jornada semanal, lo que podría estar estrechamente vinculado con la intención de renunciar. La baja ejecución y conducta laboral deficiente se vinculan de manera directa con las salidas del personal, destacando la importancia de abordar estas cuestiones para mejorar la retención. A su vez, la atracción de ofertas laborales externas, aunque menos impactante en comparación, emerge como una razón adicional que contribuye a la rotación del personal. Este análisis detallado proporciona una visión integral de los desafíos asociados con la retención de empleados y establece una base sólida para implementar estrategias efectivas en el ámbito de recursos humanos.

- Se puede inferir que las técnicas de aprendizaje automático y análisis empresarial poseen un innegable potencial para discriminar con precisión los perfiles característicos de aquellos empleados que presentan una probabilidad alta o significativa de abandonar la organización en un futuro cercano. Este potencial se manifiesta gracias a la notable capacidad predictiva que estas herramientas ofrecen, permitiendo así a los gerentes de Los Pinos Apparel, anticiparse y abordar proactivamente a los colaboradores o controlar los factores que pueden contribuir a la rotación del personal. La combinación de estas tecnologías emergentes proporciona una valiosa perspectiva estratégica que puede ser fundamental para la gestión efectiva del talento y la implementación de medidas preventivas en el ámbito de recursos humanos. Este enfoque integrado no solo facilita la retención de empleados clave, sino que también fortalece la capacidad de la organización para adaptarse dinámicamente a los desafíos y cambios en el entorno laboral.

5.2 RECOMENDACIONES

En base a los datos observados en el perfil histórico de los colaboradores que rotaron en 2023 con respecto al perfil de los empleados que aún están activas, se puede recomendar lo siguiente:

- 1- Se propone enriquecer de manera significativa la información histórica disponible con el objetivo de potenciar la capacidad de entrenamiento de los modelos predictivos, lo que, a su vez, permitirá una generación más precisa de perfiles de rotación del personal. La sugerencia radica en la complementación de la base de datos analizada hasta el momento, incorporando datos diarios de rendimiento y eficiencia que se implementarán a partir del año 2024. Esta estrategia va más allá de simplemente contar con un perfil de comportamiento al incluir un perfil adicional de rendimiento semanal. La integración de estos datos adicionales no solo enriquecerá la calidad de las predicciones, sino que también proporcionará una comprensión más holística de los factores que contribuyen a la rotación del personal. Este enfoque estratégico, respaldado por una información más completa y actualizada, fortalecerá la capacidad de la organización para anticipar y gestionar de manera proactiva los desafíos relacionados con la retención de empleados.
- 2- Para llevar a cabo esta optimización, factores clave como el rango de edad y la composición familiar, incluyendo la cantidad y edades de los hijos, pueden ser cuidadosamente considerados. Este análisis más detallado permitirá una identificación más precisa de los candidatos que no solo se ajustan a las necesidades inmediatas de la empresa, sino que también tienen un perfil y composición familiar que sugiere una mayor probabilidad de permanencia a largo plazo. En resumen, esta estrategia integrada basada en la interpretación de perfiles de empleadas anteriores proporciona un enfoque proactivo y personalizado para la retención del talento, fortaleciendo la capacidad de la organización para fomentar un entorno laboral estable y de crecimiento sostenido.
- 3- Los perfiles detallados de las empleadas que han dejado la empresa ofrecen valiosas perspectivas que pueden orientar la implementación de diversas estrategias de retención. Una sugerencia estratégica consiste en la introducción de iniciativas sociales, como la creación de guarderías destinadas a las empleadas con hijos menores de 6 años,

- con el objetivo de controlar y reducir el índice de rotación. Al aprovechar la información de los perfiles de aquellas que se han ido, se puede optimizar de manera eficaz las pautas de reclutamiento. Este enfoque implica dirigir las estrategias de contratación hacia individuos con mayores probabilidades de mantener una estabilidad a largo plazo o con expectativas de crecimiento sostenido en la empresa.
- 4- Se sugiere la ampliación y enriquecimiento de las de variables presentes en los perfiles, incorporando datos sociodemográficos que podrían desempeñar un papel crucial en la caracterización de la rotación del personal. Actualmente, existe una zona gris en la que estos datos han sido pasados por alto, y su inclusión podría proporcionar una comprensión más completa y matizada de los factores que influyen en la decisión de los empleados de abandonar la empresa. Dentro de estos datos sociodemográficos, se podrían considerar aspectos como el nivel educativo, la trayectoria profesional previa, la ubicación geográfica, la diversidad cultural y otros elementos relevantes. Al analizar estos factores aparentemente periféricos, pero potencialmente influyentes, se abriría la puerta a una interpretación más precisa de los patrones asociados con la rotación del personal. Esta estrategia de expansión en la recopilación de datos no solo mejora la calidad de los perfiles, sino que también ofrece una base más sólida para la formulación de estrategias de retención más efectivas y personalizadas.
 - 5- Se sugiere la realización de un análisis exhaustivo para evaluar la viabilidad de llevar a cabo una investigación más profunda. El enfoque principal de este análisis debería centrarse en la identificación de características específicas que distingan a aquellos colaboradores con una probabilidad significativamente mayor de mantener una permanencia prolongada en la institución. Además, se plantea la posibilidad de evaluar la factibilidad de realizar este análisis utilizando puntos de corte adaptados al ciclo de vida que los empleados pueden tener dentro de la empresa. Dado que se ha observado que la mayoría de los colaboradores tiene una antigüedad de 6 a 9 meses, considerar estos segmentos temporales puede brindar una perspectiva más contextualizada sobre las tendencias de retención. Identificar patrones específicos en diferentes fases del ciclo de vida del empleado podría revelar insights valiosos sobre los factores que contribuyen a la permanencia o a la rotación en momentos particulares de la trayectoria laboral.

CAPÍTULO VI. APLICABILIDAD

6.1 NOMBRE DE LA PROPUESTA

“Implementación de Algoritmos de Predicción de Rotación de Personal como Indicador Clave en Decisiones Estratégicas.”

6.2 JUSTIFICACIÓN DE LA PROPUESTA

Tras realizar un exhaustivo análisis del sistema de predicción de rotación de personal, se emprenderá la creación de diversos modelos predictivos con el objetivo fundamental de identificar el algoritmo más eficiente para anticipar la rotación de empleados. Estos algoritmos se sustentarán en las características comunes extraídas de los datos proporcionados en una base de datos específica, lo que posibilitará el perfilamiento de colaboradores con propensión a rotar, ya sea dentro o fuera de la institución.

La implementación de técnicas de machine learning se presenta como un elemento crucial en este proceso. Este enfoque no solo facilitará la limpieza y reestructuración de los datos, sino que también desempeñará un papel fundamental en la construcción de modelos predictivos robustos y precisos. La aplicación de machine learning permitirá la identificación de patrones complejos en los datos, mejorando así la capacidad de predicción y proporcionando insights valiosos para abordar estratégicamente la gestión de la rotación de personal en la institución.

6.3 ALCANCE DE LA PROPUESTA

Se reconoce que la tasa de rotación de personal constituye uno de los indicadores más cruciales para cualquier institución. Todas ellas aspiran a contar con un equipo laboral capacitado. Por ello, mediante el uso de algoritmos predictivos, se busca obtener un perfil detallado de colaboradores propensos a rotar, ya sea dentro o fuera de la institución. Una vez identificados, el departamento de recursos humanos empleará estos resultados como insumo clave para desarrollar propuestas de valor destinadas a fortalecer la retención y satisfacción de los colaboradores.

6.3.1 OBJETIVO GENERAL

- Desplegar el modelo de predicción con el propósito de analizar los datos de los colaboradores suministrados por recursos humanos, el cual permitirá la definición

precisa del perfil de aquellos propensos a rotar.

6.3.2 OBJETIVOS ESPECIFICOS

- Identificar patrones recurrentes en los perfiles de colaboradores que los algoritmos predicen que abandonarán la institución.
- Identificar la causa principal que genera las salidas en la institución.
- Desarrollar un plan operativo que defina tanto el cronograma como los costos asociados a la ejecución del proyecto.

6.4 DESCRIPCIÓN Y DESARROLLO

6.4.1 DESCRIPCIÓN

A continuación, se detallan los pasos que guiarán la creación del proyecto:

1. **Recolección de Datos:** Se recopilará la información necesaria en una base de datos, la cual será proporcionada por el área de recursos humanos.
2. **Herramienta de Knime Analytics Platform:** Se empleará la herramienta Knime Analytics Platform para realizar la limpieza de datos y la creación de modelos de predicción.
3. **Creación de Modelos Predictivos:** Se desarrollarán cuatro modelos de predicción con el objetivo de determinar cuál de ellos presenta la mejor precisión para anticipar la rotación del personal.
4. **Algoritmos Para Implementar:** Se implementarán cuatro algoritmos:
 - a. Regresión Logística
 - b. Naive Bayes
 - c. Árbol de Decisión
 - d. Random Forest
5. **Entrenamiento de Algoritmos:** Los algoritmos serán entrenados utilizando el

80% de los registros disponibles.

6. Evaluación de Modelos: Se llevará a cabo la evaluación de los modelos mediante matrices de confusión, las cuales proporcionarán información sobre la precisión de cada algoritmo. El algoritmo con mejor precisión será seleccionado.
7. Conexión con Colaboradores Activos: Se establecerá la conexión entre el algoritmo seleccionado y la información de los colaboradores activos, permitiendo así prever e identificar a aquellos propensos a abandonar la institución.

6.4.2 DESARROLLO

A continuación, se detallan las herramientas, instrumentos y proceso que se utilizarán para llevar a cabo el proyecto.

6.4.2.1 HERRAMIENTAS

Para el siguiente proyecto se necesitarán las siguientes herramientas de trabajo:

1. Laptop con las siguientes especificaciones:
 - a. Memoria RAM de 16GB
 - b. Procesador Intel Core i7
 - c. Disco Duro de 512GB
2. Conexión a la red de preferencia de 50 MB de velocidad
3. Descarga de Knime Analytics Platform
4. Microsoft Office.

6.4.2.2 INSTRUMENTOS

El desarrollo del producto lo dividiremos en 3 fases:

1. Limpieza de datos: Aquí se analizará todas las variables que contiene la base de datos, se procederá a realizar un matriz de congruencia de variables para poder seleccionar las que se utilizaran en el modelo de predicción para esta fase se utilizar los nodos:
 - a. Excel Reader
 - b. Group By
 - c. Column Filter
 - d. Missing Value
 - e. Rule Based Row Filter
 - f. Column Renamer
 - g. Rule Engine
 - h. Double to Integer
 - i. Number To String
 - j. Joiner
 - k. Round Double

2. Análisis exploratorio de los Datos: una vez realizada la limpieza de los datos, se realizará el análisis exploratorio de los datos el cual nos dará un panorama de cómo se encuentra la institución, para esta fase se utilizarán los nodos:
 - a. Nominal Value Row Filter
 - b. Extract Date&Time Fields
 - c. Number to String

- d. String Manipulation
 - e. Rule Based Row Filter
 - f. Group by
3. Entrenamiento de modelos de predicción: para esta fase se conectarán los nodos de entrenamiento al metanodo creado con todos los nodos que realizan la limpieza de los datos, los nodos que se utilizarán para los algoritmos serán:
- a. Rule Based Row Filter
 - b. Partitioning
 - c. Nominal Value Row Filter
 - d. Logistic Regression Learner
 - e. Logistic Regression Predictor
 - f. Color Manager
 - g. Scorer
 - h. One to many
 - i. Column Filter
 - j. Naive Bayes Learner
 - k. Naive Bayes Predictor
 - l. Decision Tree Learner
 - m. Decision Tree Predictor
 - n. Random Forest Learner

o. Random Forest Predictor

6.4.2.3 PROCESOS

1. Limpieza de datos:

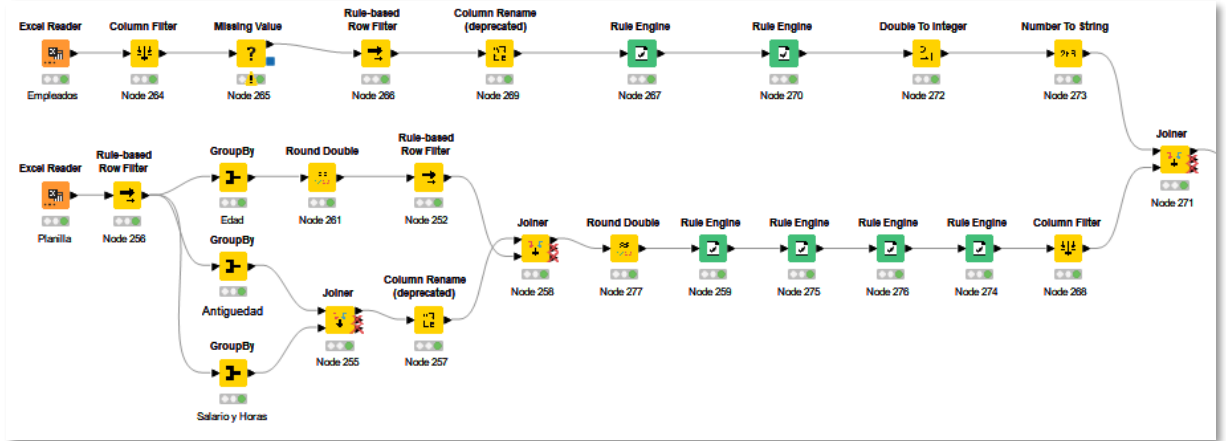


Ilustración 28 Limpieza de Datos

Fuente: Elaboración Propia

En la imagen anterior se observa el proceso de limpieza de datos, a continuación, se explica más a detalle el funcionamiento de los nodos:

Tabla 5 Nodos Limpieza de Datos

Excel Reader	Se conecta los Excel que contienen los datos de la planilla
Group By	Se realizan agrupaciones por edad, antigüedad y salario
Column Filter	Se seleccionan las columnas que contienen los datos se usaran para alimentar los algoritmos
Missing Value	Se limpia la data de valores faltantes y se reemplaza por ceros en caso numéricos
Rule Based Row Filter	Se asigna un valor de 1 en caso de que la línea de datos contenga valores mayores a 0
Column Renamer	Se renombran las columnas para una mejor identificación
Rule Engine	Se clasifica en segmentos la cantidad de hijos, antigüedad, edad, horas trabajadas y salario para un mejor agrupamiento
Double to Integer	Se pasa a números sin decimales la columna código
Number To String	Se convierte a variables string

Joiner	Todo esto se unifica con la data de planilla
Round Double	Se redondean las horas trabajadas

Fuente: Elaboración Propia

2. Análisis Exploratorio de los Datos:

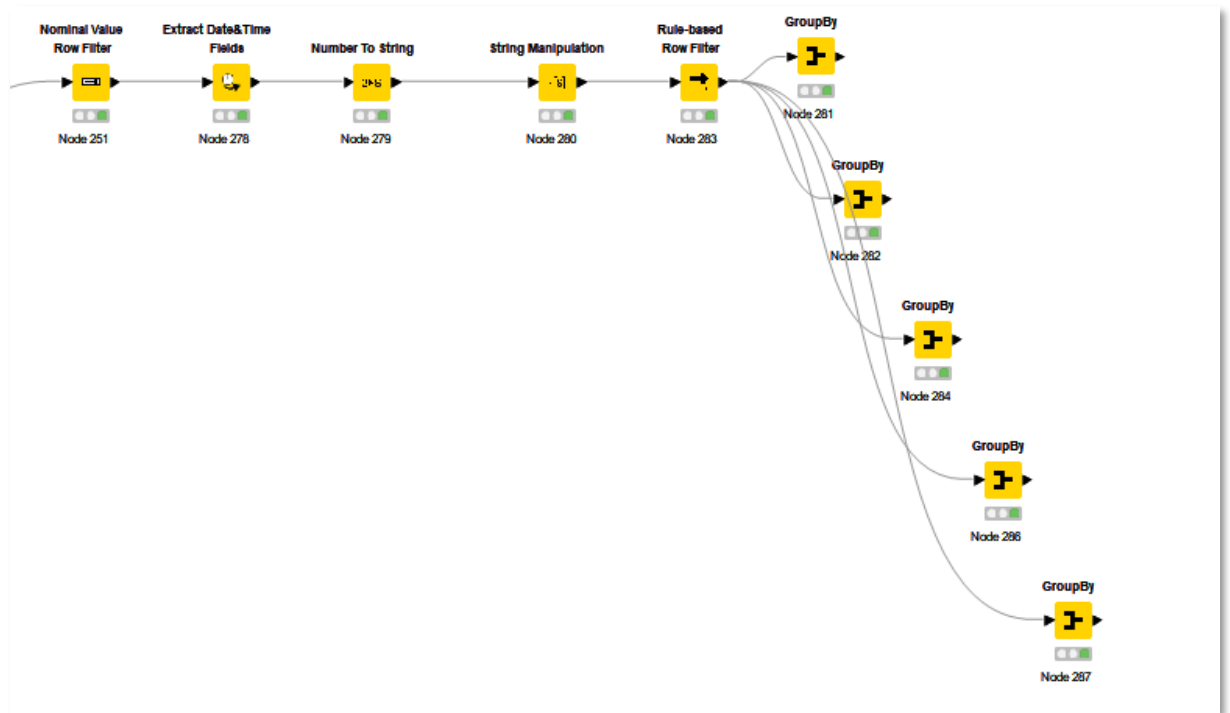


Ilustración 29 Análisis Exploratorio de los Datos

Fuente: Elaboración Propia

En la imagen anterior se observa el proceso de limpieza de datos, a continuación, se explica más a detalle el funcionamiento de los nodos:

Tabla 6 Nodos Análisis Exploratorio de Datos

Nominal Value Row Filter	Se filtra la data en base a la variable "activos"
Extract Date&Time Fields	Se extrae valores de fecha para agrupamiento
Number to String	Se convierten las variables año y mes de numéricas a string.
String Manipulation	Se concatena la variable, mes y año a un formato fecha para facilitar su comprensión.
Rule Based Row Filter	Se aplica una condición para filtrar algunas filas que contienen datos de diciembre

Group by

Se agrupan las variables considerando varias características como, fecha género, rango de antigüedad, etc.

Fuente: Elaboración Propia

3. Entrenamiento de modelos de predicción:

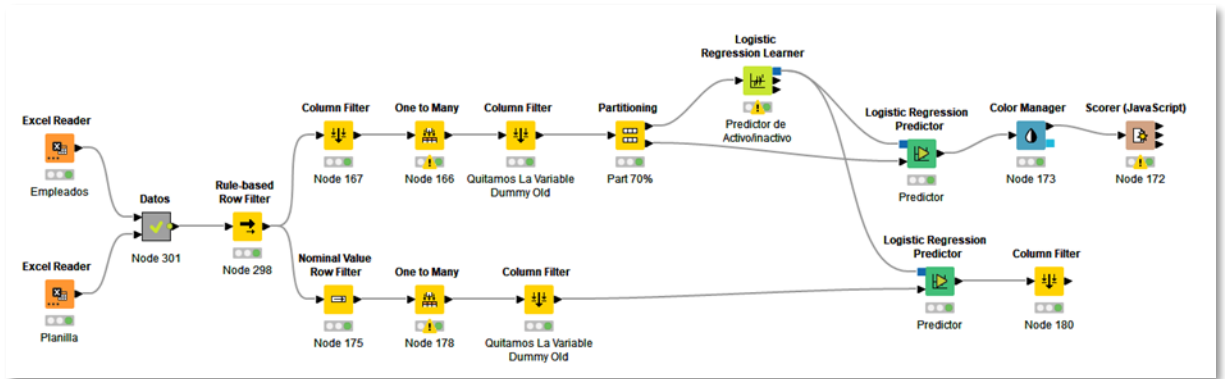


Ilustración 30 Flujo Modelo de Predicción Regresión Logística

Fuente: Elaboración Propia

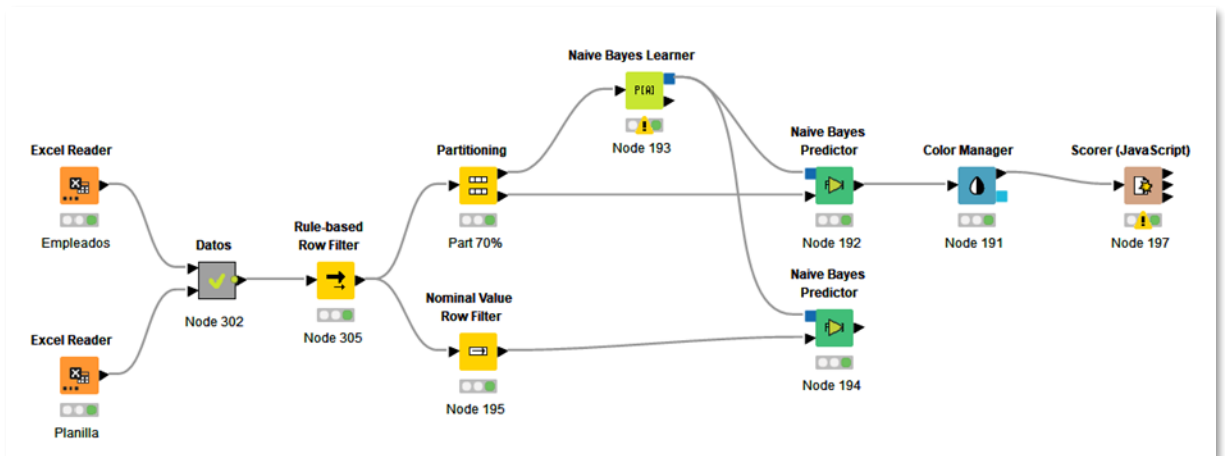


Ilustración 31 Flujo Modelo de Predicción Naive Bayes

Fuente: Elaboración Propia

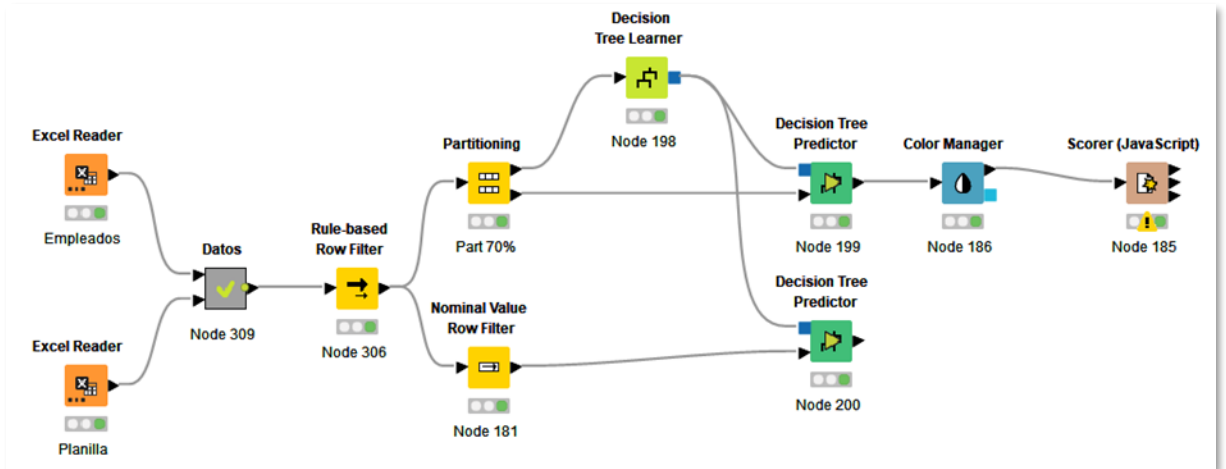


Ilustración 32 Flujo Modelo de Predicción Árbol de Decisión

Fuente: Elaboración Propia

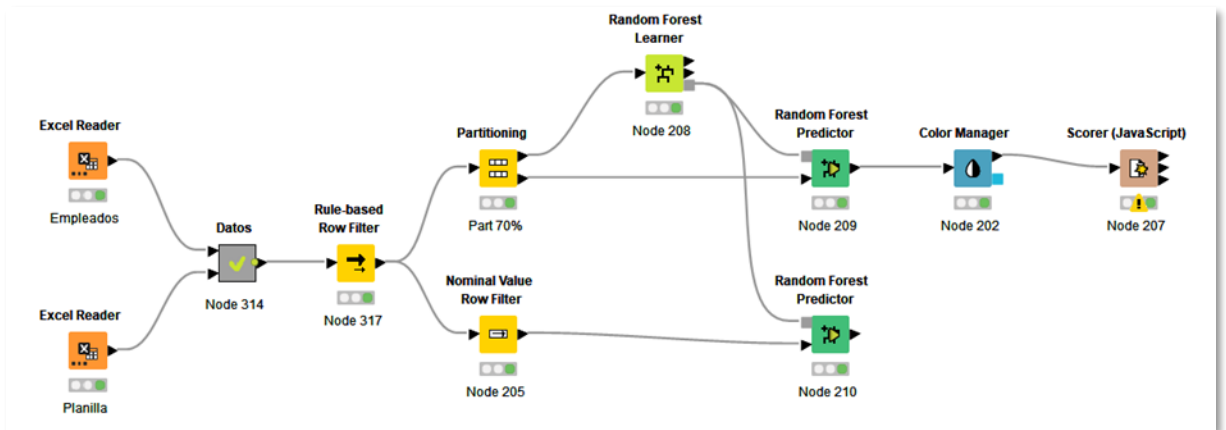


Ilustración 33 Flujo Modelo de Predicción Random Forest

Fuente: Elaboración Propia

En esta sección, se llevará a cabo el entrenamiento de los cuatro modelos de predicción, los cuales estarán conectados a la salida del proceso de limpieza de datos. Se llevará a cabo una partición de datos para realizar el entrenamiento de dichos modelos. Es importante destacar que cada modelo de predicción cuenta con sus propios nodos, los cuales se describen detalladamente a continuación.:

Tabla 7 Nodos de Entrenamiento Modelos de Predicción

Rule Based Row Filter	Se filtran las filas que no contienen edad definida.
Partitioning	Se particionan los datos al 80%, usando la variable estatus para el entrenamiento del algoritmo
Nominal Value Row Filter	Se filtran las filas que contienen datos de empleados activos de la data completa
Logistic Regression Learner	Se entrena el algoritmo de regresión logística usando las variables estatus, las filas de empleados inactivos
Logistic Regression Predictor	Se ejecuta la prueba de la data completa, aplicando los parámetros del algoritmo entrenado.
Color Manager	Se clasifica las variables activas o inactivas por colores, para su fácil apreciación
Scorer	Muestra los resultados en una matriz de confusión
One to many	Se convierte a variables dumies las variables nominales
Column Filter	Se filtran las variables dumies que no aportan al análisis predictivo.
Naive Bayes Learner	Se entrena el algoritmo de Navie Bayes usando la variable estatus.
Naive Bayes Predictor	Se ejecuta la prueba de la data completa, aplicando los parámetros del algoritmo entrenado.
Decision Tree Learner	Se entrena el algoritmo de Tree Learner usando las variables estatus, se define los parámetros no pruning para darle más amplitud al resultado
Decision Tree Predictor	Se ejecuta la prueba de la data completa, aplicando los parámetros del algoritmo entrenado.
Random Forest Learner	Se entrena el algoritmo Random Forest, usando las variables estatus, se define un máximo de 10 niveles del árbol.
Random Forest Predictor	Se ejecuta la prueba de la data completa, aplicando los parámetros del algoritmo entrenado.

Fuente: Elaboración Propia

6.5 MEDIDAS DE CONTROL

Para darle seguimiento a los modelos de predicción creados, las medidas de control que se deberán evaluar son aquellos indicadores que mostraran que modelo tiene mejor precisión y exactitud al momento de la predicción:

1. Exactitud: Este mostrara el porcentaje de clasificación correctas que realiza el modelo.

2. Precisión: Nos muestra que porcentaje de precisión tiene el modelo.
3. Error General: Porcentaje de error que tiene el modelo al momento de predecir.
4. Clasificados Correctamente: Registros clasificados de manera correcta en el modelo.
5. Clasificados Incorrectamente: Registros clasificados de manera incorrecta en el modelo

Matriz de confusion

Confusion Matrix

Rows Number : 131	ACTIVO (Predicted)	INACTIVO (Predicted)	
ACTIVO (Actual)	109	1	99.09%
INACTIVO (Actual)	1	20	95.24%
	99.09%	95.24%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
ACTIVO	109	1	20	1	99.09%	99.09%	99.09%	95.24%	99.09%
INACTIVO	20	1	109	1	95.24%	95.24%	95.24%	99.09%	95.24%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
98.47%	1.53%	0.943	129	2

Ilustración 34 Matriz de Confusión

Fuente: Elaboración Propia

Se aconseja llevar a cabo una minuciosa elaboración de una tabla de comparación que contemple detalladamente los indicadores previamente descritos. Esta práctica se revela esencial para obtener una perspectiva más exhaustiva acerca de la precisión de cada modelo, permitiendo un monitoreo meticuloso de los resultados obtenidos. Es crucial destacar que la elección del modelo con la mayor exactitud debe basarse en un análisis riguroso y en la consideración de factores adicionales, asegurando así una toma de decisiones fundamentada y respaldada por una evaluación integral de las capacidades predictivas de cada enfoque.

6.6 CRONOGRAMA DE IMPLEMENTACIÓN Y PRESUPUESTO

6.6.1 CRONOGRAMA DE ACTIVIDADES

A continuación, se detalla el cronograma de actividades que se realizarán en el proyecto.

Tabla 8 Cronograma de Actividades

Fase	Descripción	Semanas																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Fase 1	Adquisición de Laptop																					
Fase 2	Capacitación Introducción a Knime																					
	Capacitación Machine Learning con Knime																					
Fase 3	Recolección de Datos																					
	Limpieza de los Datos																					
	Análisis Exploratorio de los datos																					
Fase 4	Implementación de Algoritmos de Predicción																					
Fase 5	Análisis de Resultados																					
Fase 6	Presentación de Resultados a RRHH																					

Fuente: Elaboración Propia

6.6.2 PRESUPUESTO

A continuación, se detalla el presupuesto para llevar a cabo el proyecto:

1. Compra de Laptop con las siguientes especificaciones:
 - a. Memoria RAM de 16GB
 - b. Procesador Intel Core i7
 - c. Disco Duro de 512GB

The screenshot shows the Jetstereo website interface. At the top, there is a search bar containing 'laptop core i7'. Below the search bar is a navigation menu with options like 'Categorías', 'MARCAS', 'PROMOCIONES EXCLUSIVAS', 'CRÉDITO SOLVENZA', 'SERVICIO TÉCNICO', 'UBICANOS', and 'TECHLAB'. The main content area displays a product listing for a 'Laptop Dell Inspiron 5620 16"/ Intel Core i7/ 16GB RAM/ 512GB SSD/ Gris'. The price is listed as 'L 29,995.00' with a crossed-out original price of 'L 36,995.00' and a red 'Ahorra L 7,000.00' badge. A 'Ver más' button is visible below the product details.

Ilustración 35 Adquisición de Laptop

Fuente: (Jetstereo, 2023)

2. Capacitaciones

Tabla 9 Capacitación de Knime impartidas por IQuartil

Sesión	Objetivo	Duración	Valor Acumulado
Seminario de Introducción a Analítica con Knime	Contextualización de la Analítica y el Big Data y de KNIME como herramienta para la Ciencia de Datos Herramientas. La metodología CRISP-DM	3 semanas	\$ 488.00

Fuente: Elaboración Propia

Tabla 10 Capacitación de Udemey Machine Learning con Knime

Sesión	Objetivo	Duración	Valor Acumulado
Machine Learning con Knime para no programadores y Gerentes	Machine learning (aprendizaje automático) Utilizar modelos de regresión, clasificación, segmentación, árboles binarios y redes neuronales Cargar y transformar data (ETL) Visualización de datos	4 semanas	\$ 54.00

Fuente: Elaboración Propia

A continuación, se presenta el presupuesto para la aplicabilidad del proyecto

Tabla 11 Presupuesto de Proyecto

Producto	Costo	Unidades	Tasa de Cambio	Totales
Laptop Dell Intel core i7	\$. 1,217.33	1	L. 24.64	L. 29,995.00
Seminario de Introducción a Analítica con Knime	\$. 488.00	1	L. 24.64	L. 12,024.32
Machine Learning con Knime para no programadores y Gerentes	\$. 54.00	1	L. 24.64	L. 1,330.56
Pago de planilla para recurso humano encargado del proyecto	\$ 1,000.00	1	L. 24.64	L. 24,640.00
Total				L. 67,989.88

Fuente: Elaboración Propia

6.7 CONCORDANCIA DE LOS SEGMENTOS DE LA TESIS CON LA PROPUESTA

A continuación, se presenta la matriz de concordancia del trabajo de investigación con la aplicabilidad del proyecto:

Tabla 12 Concordancia de los segmentos de la Tesis con la Propuesta Parte I

Capítulo I			Capitulo II	Capitulo III		
Titulo Investigación	Objetivo General	Objetivos Específicos	Teorías/ Metodologías de sustento	Variables	Poblaciones	Técnicas
Implementación De Sistema De Datos Para Análisis Predictivo De Rotación De Personal En Los Pinos Apparel S.A.	Analizar si posible entrenar un modelo de predicción para estimar la probabilidad de que un colaborador renuncie, utilizando datos históricos sobre el personal y su historial de rotación almacenados en las bases de datos de la empresa.	Identificar las causas que generan la rotación en los Pinos Apparel S.A.	<p>Teorías:</p> <p>Teoría Del Equilibrio Organizacional</p> <p>Modelo Desdoblante</p> <p>Machine Learning</p> <p>Modelos De Machine Learning</p> <p>Metodologías:</p> <p>Recolección De Datos</p> <p>Algoritmos De Predicción</p>	<p>Colaboradores</p> <p>Causas de Rotación</p> <p>Variables Comunes en Perfiles</p>	<p>La población de la siguiente investigación comprende todos los colaboradores activos que prestan servicios para la institución de los Pinos Apparel</p>	<p>Algoritmos de Predicción:</p> <p>Regresión</p> <p>Logística</p> <p>Naive Bayes</p> <p>Árbol de Decisión</p> <p>Random Forest</p>
		Definir cuáles son las características y variables comunes en los perfiles de los empleados que renunciaron a la empresa en el 2023.				
		Analizar si existen técnicas de machine learning y analítica de negocios que pueden implementarse para la generación de un perfil del personal candidato a dejar la organización en un corto plazo.				

Tabla 13 Concordancia de los segmentos de la Tesis con la Propuesta Parte II

Capítulo V	Capítulo VI	
Conclusiones	Nombre de la Propuesta	Objetivos de la propuesta
<p>Se concluye que es viable lograr una precisión del 80% en un modelo de predicción de clasificación y que es factible para estimar la probabilidad de rotación del personal. Esta precisión en la predicción de los algoritmos se fundamenta en el peso que aporta las diversas variables extraídas del perfil de colaboradores y que estuvieron presentes en los colaboradores que experimentaron rotación en el año 2023, así como en el análisis del comportamiento histórico de los colaboradores activos en la institución. Dentro de la variedad de modelos sometidos a evaluación, destaca la regresión logística y Random Forest, ambos alcanzando una destacada precisión del 97%. La implementación exitosa del algoritmo de predicción ha demostrado ser una herramienta sumamente valiosa como complemento a la gestión del departamento de recursos humanos, proporcionando una base sólida para la toma de decisiones informada y estratégica.</p>	<p>Implementación de Algoritmos de Predicción de Rotación de Personal como</p>	<p>Desplegar el modelo de predicción con el propósito de analizar los datos de los colaboradores suministrados por recursos humanos, el cual permitirá la definición precisa del perfil de aquellos propensos a rotar.</p>
<p>La generación de perfiles ha permitido identificar de manera clara la variedad de características internas o externas y económicas, psicológicas, demográficas subyacentes a la realidad de los empleados que rotaron en 2023, y es notorio que las colaboradoras involucradas mayormente ostentan el estado civil de solteras. Así como las causas expresadas en su entrevista de salida donde en primer lugar, la presión laboral se erige como un factor significativo, ejerciendo una fuerza que conduce a la renuncia de algunos colaboradores. Problemas de salud y complicaciones familiares, especialmente prominentes entre el personal femenino, se perfilan como motivos recurrentes, atribuibles en gran medida a las responsabilidades familiares. Este análisis exhaustivo proporciona una comprensión más profunda de los factores que influyen en las salidas específicas, ofreciendo a la empresa una base sólida para implementar estrategias</p>	<p>Indicador Clave en Decisiones Estratégicas</p>	<p>Identificar patrones recurrentes en los perfiles de colaboradores que los algoritmos predicen que abandonarán la institución. Identificar la causa principal que genera las salidas en la institución.</p>

Capítulo V	Capítulo VI	
Conclusiones	Nombre de la Propuesta	Objetivos de la propuesta
orientadas a mejorar la retención en estas áreas críticas.		
<p>La elaboración de perfiles ha permitido identificar de manera sistemática patrones recurrentes asociados a las causas laborales que provocaron salidas de colaboradoras de la empresa. Un análisis detallado revela que las renunciaciones están predominantemente concentradas en la Sección de Línea 6-camisas en primera instancia y, en segundo lugar, en la Línea 1-Pantalón. Estas salidas se caracterizaron por colaboradoras con perfiles de una antigüedad laboral inferior a 6 meses. Un hallazgo constante en estas situaciones es que, de manera consistente, estas colaboradoras no lograban cumplir con su cantidad de horas laboradas con respecto a la jornada semanal, lo que podría estar estrechamente vinculado con la intención de renunciar. La baja ejecución y conducta laboral deficiente se vinculan de manera directa con las salidas del personal, destacando la importancia de abordar estas cuestiones para mejorar la retención. A su vez, la atracción de ofertas laborales externas, aunque menos impactante en comparación, emerge como una razón adicional que contribuye a la rotación del personal. Este análisis detallado proporciona una visión integral de los desafíos asociados con la retención de empleados y establece una base sólida para implementar estrategias efectivas en el ámbito de recursos humanos.</p>		<p>Desarrollar un plan operativo que defina tanto el cronograma como los costos asociados a la ejecución del proyecto.</p>

Capítulo V	Capítulo VI	
Conclusiones	Nombre de la Propuesta	Objetivos de la propuesta
<p>Se puede inferir que las técnicas de aprendizaje automático y análisis empresarial poseen un innegable potencial para discriminar con precisión los perfiles característicos de aquellos empleados que presentan una probabilidad alta o significativa de abandonar la organización en un futuro cercano. Este potencial se manifiesta gracias a la notable capacidad predictiva que estas herramientas ofrecen, permitiendo así a los gerentes de Los Pinos Apparel, anticiparse y abordar proactivamente a los colaboradores o controlar los factores que pueden contribuir a la rotación del personal. La combinación de estas tecnologías emergentes proporciona una valiosa perspectiva estratégica que puede ser fundamental para la gestión efectiva del talento y la implementación de medidas preventivas en el ámbito de recursos humanos. Este enfoque integrado no solo facilita la retención de empleados clave, sino que también fortalece la capacidad de la organización para adaptarse dinámicamente a los desafíos y cambios en el entorno laboral.</p>		

Fuente: Elaboración Propia

REFERENCIAS BIBLIOGRÁFICAS

- AEL. (s.f.). *El portal de la tesis, Universidad de Colima* . Obtenido de Universidad de Colima : <https://recursos.ucol.mx/tesis/investigacion.php>
- Berthold, M., Cebron, N., Dill, F., Gabriel, T., Kötter, T., Meinl, T., & Wiswedel, B. (2008). KNIME: The Konstanz Information Miner. In "Data Analysis, Machine Learning and Applications" . *Springer*, 319-326.
- Bluedorn, A. C. (1982). A Unified Model of Turnover from Organizations. *Human Relations*, 35.
- Borjas, G. J. (2014). *Labor Economics*. New York: McGraw-Hill Education.
- Breiman, L. (2001). Random forests. *Machine Learning*.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. CRC press.
- Chiavenato, I. (2007). Subsistema de integración de recursos humanos. En I. Chiavenato, *Administración de recursos Humanos* (pág. 135). Mexico,; Mc Graw-Hill.
- Chiavenato, I. (2008). Relaciones con los Empleados. En *Gestión del Talento Humano* (pág. 444). Mexico: Mc Graw Hill.
- Dessler, G., & Varela, R. (2011). *Administración de recursos humanos*. Mexico: Pearson.
- Estrada, R. (2018). El costo de la rotación de personal. *Deloitte*, 4.
- Greene, W. H. (2012). *Econometric analysis*. Pearson.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Hauser, R. M., & Duncan, G. J. (2001). Educational expansion and educational opportunity in the United States, 1947 to 1998. *Demography*.
- Hernandez Mendoza, S. L., & Duana Avila, D. (2020). Técnicas e instrumentos de recolección de datos. *UAEH*.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.
- Hurley, P. J. (2011). *A concise introduction to logic*. Cengage Learning.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer.
- Jetstereo. (2023). *Jetstereo*. Obtenido de Jetstereo: <https://www.jetstereo.com/>
- Lis Data Solutions. (s.f.). Obtenido de Lis Data Solutions: <https://www.lisdatasolutions.com/es/que-es-knime/#:~:text=KNIME%20pertenece%20a%20una%20nueva,sus%20modelos%20de%20analítica%20avanzada>.
- Maldonado, A. (s.f.). *Teoria Online*. Obtenido de Teoria Online: <https://teoriaonline.com/teoria-del-equilibrio-organizacional/>
- Melendres, V., Ramirez, M., Garcia, B. R., & Aranibar, M. F. (2017). LOS FACTORES DE LA ROTACIÓN DE PERSONAL EN LAS MAQUILADORAS DE EXPORTACIÓN DE ENSENADA, B.C. *ReSearchGet (Revista Global de Negocios)*, 17.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis*. Wiley.
- Morales, R. V. (1959). *Honduras Patente n° Código de Trabajo*.
- Newman, J. M., Gerhart, B. A., & Milkovich, G. T. (2021). *Compensation*. McGraw-Hill

Education.

Relph, E. (1976). *Place and placelessness*. Pion.

Romo, H. L. (1998). La metodología de Encuesta. Obtenido de https://biblioteca.marco.edu.mx/files/metodologia_encuestas.pdf

Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach*. Pearson.

Sampieri, R. H., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la Investigación*. Mc Grawhill Education.

Sandoval, L. (2018). ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PARA ANÁLISIS Y PREDICCIÓN DE DATOS. *Revista Tecnológica*.

SAP. (s.f.). Obtenido de SAP: <https://www.sap.com/latinamerica/products/artificial-intelligence/what-is-machine-learning.html#:~:text=learning%20en%20detalle,El%20machine%20learning%20es%20un%20subconjunto%20de%20la%20inteligencia%20artificial,ser%20explícitamente%20programadas%20par>

Steel, R. P., & Ovalle, N. K. (1984). A review and meta-analysis of research on the relationship between behavioral intentions and employee turnover. *Journal of Applied Psychology*.

Trochim, W. M. (2006). *Research Methods Knowledge Base*. Obtenido de Population: <https://conjointly.com/kb/population/>

ANEXOS

Anexo 1



NOTA DE AUTORIZACION

Yo, Marco Zavala, el suscrito Gerente General de la Empresa **LOS PINOS APPAREL S.A** por medio de la presente hace constar que autorizo a **Karen Yamileth Cruz Mata**, actual, Gerente de Recursos Humanos, para disponer de la información relacionada con planillas, códigos de empleado, asistencia, indicadores de calidad y eficiencia, etc. Para ser utilizados como data de investigación de su proyecto de tesis de la maestría de Analítica de Negocios.

Para los fines que el interesado estime conveniente, firmo la presente en la Ciudad de Siguatepeque, Departamento de Comayagua, a los 16 días del mes de diciembre del año 2023.

Atentamente

Marco Zavala
Gerente de General
Los Pinos Apparel S. A.

Barrió Zaragoza, 19.ave, 1 calle, Siguatepeque, Comayagua. Tel 2773-5353

CARTA DE AUTORIZACIÓN DE LA EMPRESA O INSTITUCIÓN

Sigüitepeque Comayagua 16/12/22
(Ciudad) (Departamento) (Día, mes y año)

Marco Zavala
(Nombre y apellidos del Director o Gerente)

Gerente General
(Puesto Laboral)

Los Pinos Apparel S.A.
(Empresa o Institución)

Ramón Ferreroza, Mave Torre
(Dirección principal de la empresa o institución)

Estimado Señor(a): Marco Zavala


Reciba un cordial y atento saludo. Por medio de la presente deseamos solicitar su apoyo, dado que somos alumnos de UNITEC y nos encontramos desarrollando el Trabajo de Tesis previo a obtener nuestro título de maestría en Análisis de negocios

Hemos seleccionado como tema Implementación de sistema de datos para análisis predictivo de rotación de personal, por lo que estaríamos muy agradecidos de contar con el apoyo de la empresa que usted representa para poder desarrollar nuestra investigación. En particular, dicha solicitud se circunscribe a petitionar que se nos autorice a realizar el uso de la base de datos de empleados y demás

(estadísticas, sueldos, etc.)

A la espera de su aprobación, me suscribo de Usted.

Atentamente,


Karen Cruz
Firma, nombre y apellidos
No. de cuenta: 21613114


Karen Cruz
Firma, nombre y apellidos
No. de cuenta: 2212044

Por este medio, Los Pinos Apparel S.A.
(empresa / institución)

Autoriza la realización dentro de sus instalaciones el proyecto de investigación de Tesis de Postgrado antes mencionado.

Marco Zavala
(Nombre y sello del Director / Gerente)


(Sello)