



UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO
DIVISIÓN ACADÉMICA DE
CIENCIAS Y TECNOLOGÍAS DE LA INFORMACIÓN



**IDENTIFICACIÓN DE PATRONES DE CONDUCTA, ASOCIADAS
CON VARIABLES DE INVESTIGADORES DE EDUCACIÓN
SUPERIOR, APLICANDO TÉCNICAS DE MINERÍA DE DATOS.**

Trabajo recepcional bajo la modalidad de Tesis
que para obtener el grado de:

**Maestro en Administración
de Tecnologías de la Información**

Presenta:

Angel Gustavo González Sánchez

Directores de Trabajo Recepcional:

Dra. Martha Patricia Silva Payró

Cuerpo Académico: Gestión de Tecnologías de la Información

Dr. Gerardo Arceo Moheno

Cuerpo Académico: Innovación en Organismos Sociales

Línea de Generación y Aplicación del Conocimiento de la Maestría:

**Administración, diseño e implementación de
Tecnologías de la Información.**

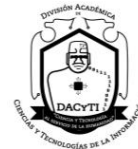
Cunduacán, Tabasco

Junio, 2023



UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

**DIVISIÓN ACADÉMICA DE
CIENCIAS Y TECNOLOGÍAS DE LA INFORMACIÓN**



**IDENTIFICACIÓN DE PATRONES DE CONDUCTA, ASOCIADAS
CON VARIABLES DE INVESTIGADORES DE EDUCACIÓN
SUPERIOR, APLICANDO TÉCNICAS DE MINERÍA DE DATOS.**

Trabajo recepcional bajo la modalidad de Tesis
que para obtener el grado de:

**Maestro en Administración
de Tecnologías de la Información**

Presenta:

Angel Gustavo González Sánchez

Directores de Trabajo Recepcional:

Dra. Martha Patricia Silva Payró

Cuerpo Académico: Gestión de Tecnologías de la Información

Dr. Gerardo Arceo Moheno

Cuerpo Académico: Innovación en Organismos Sociales

Jurado Revisor:

Dr. Guillermo de los Santos Torres

Dr. Pablo Payró Campos

Dr. Eric Ramos Méndez

MA. Rubicel Cruz Romero

MATI. Rafael Mena de la Rosa

Línea de Generación y Aplicación del Conocimiento de la Maestría:
**Administración, diseño e implementación de
Tecnologías de la Información**

Cunduacán, Tabasco

Junio, 2023



UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

**DIVISIÓN ACADÉMICA DE CIENCIAS
Y TECNOLOGÍAS DE LA INFORMACIÓN**



F8: Cesión de Derechos

Cunduacán, Tabasco, a 30 de mayo de 2023.

MTE. Óscar Alberto González González
Director de la División Académica de Ciencias y Tecnologías de la Información
Presente

El que suscribe la presente, declara que el trabajo de tesis titulado "IDENTIFICACIÓN DE PATRONES DE CONDUCTA, ASOCIADAS CON VARIABLES DE INVESTIGADORES DE EDUCACIÓN SUPERIOR, APLICANDO TÉCNICAS DE MINERÍA DE DATO" es de mi autoría intelectual y por lo tanto cedo todos los derechos sobre este proyecto a la Universidad Juárez Autónoma de Tabasco, a la cual relevo y al Director del mismo de cualquier sanción y asumo responder a cualquier reclamo de derecho de autor ante las autoridades competentes.

Atentamente

Angel Gustavo González Sánchez

Martha Patricia Silva Payró

Gerardo Arceo Moheno



c.c.p. Dr. Eddy Arquímedes García Alcocer. Encargada del despacho de la Coordinación de Posgrado.
Estudiante.

Cunduacán, Tabasco a 30 de mayo de 2023

Asunto: **Oficio de autorización**

A quien corresponda:

El que suscribe, autoriza por medio del presente escrito a la Universidad Juárez Autónoma de Tabasco para que utilice tanto física como digitalmente la tesis de grado denominada **“IDENTIFICACIÓN DE PATRONES DE CONDUCTA, ASOCIADAS CON VARIABLES DE INVESTIGADORES DE EDUCACIÓN SUPERIOR, APLICANDO TÉCNICAS DE MINERÍA DE DATOS”** de la cual soy autor y titular de los Derechos de Autor.

La finalidad del uso por parte de la Universidad Juárez Autónoma de Tabasco de la tesis antes mencionada, será única y exclusivamente para difusión, educación y sin fines de lucro; autorización que se hace de manera enunciativa mas no limitativa para subirla a la Red Abierta de Bibliotecas Digitales (RABID) y a cualquier otra red académica con las que la Universidad tenga relación institucional.

Por lo antes manifestado, libero a la Universidad Juárez Autónoma de Tabasco de cualquier reclamación legal que pudiera ejercer respecto al uso y manipulación de la tesis mencionada y para los fines estipulados en este documento.

Se firma la presente autorización en la ciudad de Villahermosa, Tabasco a los 29 días del mes de mayo del año 2023.

Atentamente



Angel Gustavo González Sánchez

202H19004

c.c.p. Dr. Eddy Arquimedes Garcia Alcocer. Encargado del despacho de la Coordinación de Posgrado.

Estudiante.



UNIVERSIDAD JUÁREZ
AUTÓNOMA DE TABASCO

"ESTUDIO EN LA DUDA. ACCIÓN EN LA FE"



DIVISIÓN ACADÉMICA DE
CIENCIAS Y TECNOLOGÍAS
DE LA INFORMACIÓN



"2021, Año de la Independencia de México"

Cunduacán, Tabasco a 18 de enero de 2021
Oficio No. 091/DACYTI/CP/2021

Asunto: Dirección de Tesis

Dra. Martha Patricia Silva Payro
Profesora Investigadora
PRESENTE

De conformidad con lo establecido en el Reglamento de Estudios de Posgrado Vigente, de la Universidad Juárez Autónoma de Tabasco, me permito informarle, que ha sido designada como Directora de la tesis titulada **"Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos"**, a realizar por el **C. Ángel Gustavo González Sánchez**, para obtener el grado de Maestro en Administración de Tecnologías de la Información.

Sin otro particular, aprovecho la ocasión para enviarle un afectuoso saludo.

Atentamente

MTE. Óscar Alberto González González
Director

UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO



DIVISIÓN ACADÉMICA DE
CIENCIAS Y TECNOLOGÍAS DE LA INFORMACIÓN

C.c.p: Coordinación de Posgrado de DACYTI
Archivo
Consecutivo

Carretera Cunduacán-Jalpa Km. 1, Colonia Esmeralda, C.P. 86690.
Cunduacán, Tabasco, México.
Tel: (993) 358 1500 ext. 6727; (914) 336 0616; Fax: (914) 336 0870
E-mail: direccion.dacyti@ujat.mx

www.ujat.mx



"2021, Año de la Independencia de México"

Cunduacán, Tabasco a 16 de junio de 2021
Oficio No. 706/DACYTI/CP/2021

Asunto: Dirección de Tesis

Dr. Gerardo Arceo Moheno
Profesor Investigador

De conformidad con lo establecido en el Reglamento de Estudios de Posgrado Vigente, de la Universidad Juárez Autónoma de Tabasco, me permito informarle, que ha sido designado como Director de la tesis titulada **"Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos"**, a realizar por el **C. Ángel Gustavo González Sánchez**, para obtener el grado de Maestro en Administración de Tecnologías de la Información.

Sin otro particular, aprovecho la ocasión para enviarle un afectuoso saludo.

Atentamente



MTE. Oscar Alberto González González
Director

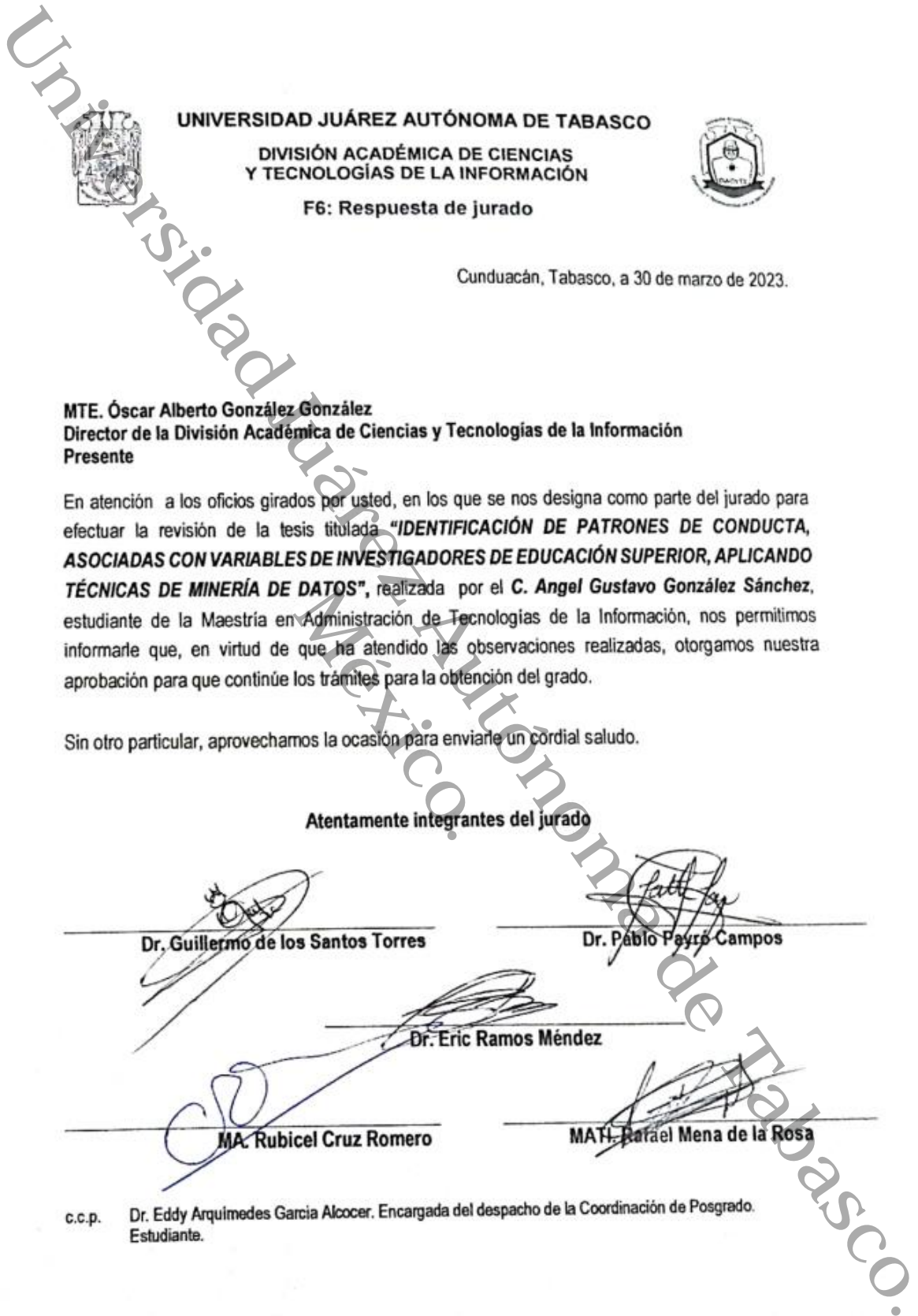


C.e.p. Coordinación de Posgrado
Alumno
Archivo

X

Carretera Cunduacán-Jalpa Km. 1, Colonia Esmeralda, C.P. 86690
Cunduacán, Tabasco, México.
Tel: (993) 358 1500 ext. 6727; (914) 336 0616; Fax: (914) 336 0877
E-mail: direccion.dacyti@ujat.mx

www.ujat.mx



UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

DIVISIÓN ACADÉMICA DE CIENCIAS Y TECNOLOGÍAS DE LA INFORMACIÓN



F6: Respuesta de jurado

Cunduacán, Tabasco, a 30 de marzo de 2023.

MTE. Óscar Alberto González González
Director de la División Académica de Ciencias y Tecnologías de la Información
Presente

En atención a los oficios girados por usted, en los que se nos designa como parte del jurado para efectuar la revisión de la tesis titulada **"IDENTIFICACIÓN DE PATRONES DE CONDUCTA, ASOCIADAS CON VARIABLES DE INVESTIGADORES DE EDUCACIÓN SUPERIOR, APLICANDO TÉCNICAS DE MINERÍA DE DATOS"**, realizada por el **C. Angel Gustavo González Sánchez**, estudiante de la Maestría en Administración de Tecnologías de la Información, nos permitimos informarle que, en virtud de que ha atendido las observaciones realizadas, otorgamos nuestra aprobación para que continúe los trámites para la obtención del grado.

Sin otro particular, aprovechamos la ocasión para enviarle un cordial saludo.

Atentamente integrantes del jurado



Dr. Guillermo de los Santos Torres



Dr. Pablo Payro Campos



Dr. Eric Ramos Méndez



MA. Rubicel Cruz Romero



MATH. Rafael Mena de la Rosa

c.c.p. **Dr. Eddy Arquimedes Garcia Alcocer**. Encargada del despacho de la Coordinación de Posgrado. Estudiante.



DIVISIÓN ACADÉMICA DE CIENCIAS Y TECNOLOGÍAS DE LA INFORMACIÓN



Cunduacán, Tabasco a 30 de marzo de 2023

Oficio No. 497/DACYTI/D/2023

Asunto: Autorización de impresión de Tesis

C. Angel Gustavo González Sánchez
Matricula: 202H19004

En virtud de que cumple satisfactoriamente los requisitos establecidos en el Reglamento General de Estudios de Posgrado vigente en la Universidad, informo a Usted que se autoriza la impresión del trabajo recepcional **"Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos"**, para presentar examen y obtener el Grado de Maestro en Administración de Tecnologías de la Información.

Sin otro particular, aprovecho la ocasión para enviarle un afectuoso saludo.

Atentamente

[Handwritten signature]
MTE. Oscar Alberto González Godínez
Director

UNIVERSIDAD JUÁREZ
AUTÓNOMA DE TABASCO



DIVISIÓN ACADÉMICA DE
CIENCIAS Y TECNOLOGÍAS
DE LA INFORMACIÓN

C.c.p. Dr. Eddy Arquímedes García Alcocer, - Encargado del Archivo de la Coordinación de Posgrado DACYTI
Archivo.
Consecutivo.
MTE OAGG/EAGA X

Carretera Cunduacán-Jalpa Km. 1, Colonia Esmeralda, CP. 86690
Cunduacán, Tabasco, México
Tel: (993) 358 1500 ext. 6727; (914) 336 0616; Fax: (914) 336 0870
E-mail: direccion.dacyti@ujat.mx

www.ujat.mx

Agradecimientos

Un agradecimiento especial a mis directores de tesis:

Dra. Martha Patricia Silva Payró y el Dr. Gerardo Arceo Moheno

Por haber depositado su confianza en mí para la realización y culminación de este proyecto; por haber dado las bases necesarias para lograr este objetivo, por cada mensaje de ánimo ante las diferentes situaciones presentadas, por tener siempre disponibilidad, apoyo incondicional, paciencia, constancia y amistad. De corazón ¡Gracias!

A los miembros del jurado revisor:

Dr. Guillermo de los Santos Torres, Dr. Pablo Payró Campos, Dr. Eric Ramos Méndez, MA. Rubicel Cruz Romero, MATI. Rafael Mena de la Rosa

Por su participación entusiasta, correcciones, comentarios y sugerencias que en su momento fueron pertinentes para el desarrollo de esta investigación, y por resolver cada una de las dudas presentadas a lo largo de la realización de este proyecto. En verdad ¡Gracias!

Dedicatoria

Dedico mi tesis a mi hijo Christopher por ser el motor que impulsa todos mis sueños, por darme la fortaleza que necesitaba en este camino tan largo, por ser mi motivo e inspiración para poder culminar con este objetivo fijado en mi vida; te amo hijo.

A mi mamá María del Rosario Sánchez Colomé y su esposo Abelzain Sánchez León, a mi papá José Guadalupe González Flores y su esposa Gloria Luciano Salvador, por no dejarme caer, por siempre apoyarme en cada paso y en cada decisión que he tomado hasta ahora, por darme la oportunidad de superarme, por acompañarme en mis logros y en mis fracasos y por haber estado conmigo de forma incondicional siempre.

A mis hermanos Miguel, Cristian y José por ser ejemplos de estudio y perseverancia, por enseñarme tanto, por compartir tristezas y alegrías juntos.

A ti Amor por siempre haberme apoyado, creer y siempre estar para mí en esta etapa de mi vida.

Pero, sobre todo, gracias a Dios por haberme permitido llegar a esta etapa de mi vida.

Ángel Gustavo González Sánchez

Resumen

El propósito de la investigación es identificar los patrones de conducta que afectan el desarrollo de la función de investigación en los docentes universitarios mediante el uso de diferentes técnicas de minería de datos. Lo anterior en virtud de que representa un área de sumo interés para el crecimiento y mejora de las Instituciones de Educación Superior. La técnica que se empleó para la recolección de datos son las encuestas aplicadas a investigadores de nivel superior, el análisis y comprensión de la información fue extraída con herramientas de minería de datos como WEKA, MOA, *Rapidminer*, por mencionar algunas, con la cuales se realizó el proceso KDD, así mismo se aplicaron algoritmos de minería de datos en estos la técnica de *Clustering* como método para la agrupación de los datos, también se aplicó la técnica de árbol de decisión para una representación analítica de los patrones, las cuales ayudaron a obtener los resultados que identificaron los patrones de conducta que más relevancia tienen los profesores investigadores. Cuidar el tiempo en la redacción, porque en algunas partes redactas en futuro y debe ser en pasado, porque es el resumen de lo que ya hiciste

Introducción

El presente trabajo que lleva por nombre "Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos" está dividido en 5 capítulos los cuales se mencionan a continuación.

En el capítulo uno llamado generalidades, se encuentran los antecedentes, el planteamiento de problema el cual está conformado por la definición del problema, delimitación de la investigación, el objetivo general, así como los objetivos específicos, también abarca la justificación y la metodología que se utilizó para el desarrollo de esta investigación.

En el capítulo dos, se aborda el sustento de la investigación ya que en él se ve los marcos de la investigación el cual se conforma por cinco:

Marco teórico, en el cual se analizan diferentes teorías de los factores que enfrentan los profesores investigadores durante su crecimiento profesional y de cómo el uso de la minería de datos ayudó a identificar patrones de conducta que afectan la productividad de los docentes.

Marco referencial, el cual contiene diferentes investigaciones en las cuales utilizaron la minería de datos para detectar patrones de conducta.

Marco conceptual, que permite conocer diferentes conceptos asociados a esta investigación como lo son: patrones de conducta, minería de datos, proceso KDD por mencionar algunos.

Marco tecnológico, que describe las diferentes tecnologías, herramientas y software, que pueden utilizarse en la minería de datos, así como los lenguajes de programación en los que están escritas dichas herramientas.

Marco legal, este último marco se encarga de mostrar las leyes, normas o reglamentos que le dan el sustento legal a la investigación.

En el capítulo tres, aplicación de la metodología y desarrollo, en primera instancia se encuentra la cuantificación y la depuración de la población de estudio con la cual se realizó esta investigación así mismo se encuentra la metodología del proceso KDD el cual cuenta con cinco etapas, selección de los datos, preprocesamiento y limpieza de la información, transformación, minería de datos en la cual se aplican los algoritmos e interpretación del conocimiento adquirido.

El capítulo cuatro resultados, muestra los resultados obtenidos con los algoritmos de árbol de decisión, agrupamiento y reglas de asociación que se implementaron con la herramienta de minería de datos RapidMiner.

Por último, en el capítulo cinco se presentan las conclusiones que brindan el cierre de la investigación, así como las recomendaciones y trabajos futuros, que puedan llegar a realizarse a partir de esta investigación.

Índice general

Índice de tablas	xx
Índice de figuras	xxi
Capítulo 1. Generalidades	1
1.1 Antecedentes	1
1.1.1 Factores que influyen en el desarrollo profesional del profesor	2
1.1.2 Minería de datos aplicados al rendimiento académico	4
1.2 Planteamiento del problema	5
1.2.1 Definición del problema	5
1.2.2 Delimitación de la investigación	10
1.2.3 Pregunta de investigación	12
1.3 Objetivos	12
1.3.1 Objetivo general	12
1.3.2 Objetivos específicos	12
1.4 Justificación	13
1.5 Metodología utilizada	15
1.6 Fuentes de investigación	15
1.7 Técnica de recolección de datos	16
1.8 Metodología para aplicar	17
1.9 Herramienta de minería datos	21
Capítulo 2. Marco de la investigación	22

2.1 Marco teórico 22

2.2 Marco referencial 26

 2.2.1 Afrontamiento, vida personal y familiar de profesores que pertenecen al Sistema Nacional de Investigadores (SNI) de una universidad del sureste de México. 26

 2.2.2 Las habilidades investigativas como eje transversal de la formación para la investigación..... 27

 2.2.3 Análisis del desempeño del profesorado universitario en el uso de MOODLE a través de técnicas de minería de datos: propuestas de necesidades formativas.28

 2.2.4 Minería de datos aplicada a la identificación de factores de deserción universitaria en programas de pregrado..... 30

 2.2.5 Factores que influyen en la productividad científica en una universidad pública. La perspectiva de los académicos. 31

 2.2.6 Utilización de técnicas de minería de datos para la identificación de rasgos de comportamiento en procesos de aprendizaje colaborativo en modelos de e-learning y b-learning. 32

 2.2.7 Modelo de minería de datos para identificación de patrones que influyen en el aprovechamiento académico..... 33

2.3 Marco conceptual..... 33

 2.3.1 Patrones de conducta. 33

 2.3.2 Minería de datos. 34

- 2.3.3 Barreras de datos..... 34
- 2.3.4 Variable..... 34
- 2.3.5 Proceso KDD. 35
- 2.3.6 Análisis de datos..... 35
- 2.3.7 Aprendizaje automático o *machine learning*. 35
- 2.4 Marco tecnológico..... 35
 - 2.4.1 Waikato Environment for Knowledge Analysis (WEKA). 35
 - 2.4.2 Konstanz Information Miner (KNIME)..... **¡Error! Marcador no definido.**
 - 2.4.3 Yet Another Learning Environment (YALE)..... 36
 - 2.4.4 RapidMiner..... 36
 - 2.4.5 Massive Online Analysis (MOA)..... 37
 - 2.4.6 Servicios en la nube para minería de datos. 37
 - 2.4.7 Just Another Vague Acronym (JAVA). 37
 - 2.4.8 Python..... 38
- 2.5 Marco legal 38
 - 2.5.1 Licencia Pública General (GNU). 38
 - 2.5.2 Ley Federal de Protección de Datos Personales en Posesión de los Particulares. 39
 - 2.5.3 Licencia de software propietario..... 39
 - 2.5.4 Licencia de software libre AGPL 40
- Capítulo 3. Aplicación de la metodología y desarrollo 41

3.1	Cuantificación de la población de estudio.....	41
3.2	Depuración de la población	42
3.3	Diseño de instrumentos	43
3.4	Aplicación del instrumento	44
3.5	Almacenamiento para el análisis de la información	44
3.6	Análisis para la selección de herramienta de minería de datos MOA, YALE y WEKA.	45
3.7.1	RapidMiner.....	46
3.7	Detección de las necesidades	52
3.8	Metodología KDD.....	53
3.8.1	Selección	54
3.8.2	Preprocesamiento/limpieza.....	58
3.8.3	Transformación de los datos.....	61
3.8.4	Minería de datos	62
3.8.5	Interpretación/Evaluación.....	75
Capítulo 4.	Resultados.....	77
4.1	Árbol de decisión	77
4.2	Clustering Algoritmo K-means	93
4.3	Reglas de asociación.....	99
Capítulo 5.	Conclusiones, recomendaciones y trabajos futuros	110
5.1	Recomendaciones	113

5.2 Trabajos futuros.....	116
Referencias.....	119
Glosario.....	133
Apéndice A. Cuestionario aplicado a profesores investigadores.....	135
Apéndice B. Relación de resultados de tamaños de clúster.....	146

Universidad Juárez Autónoma de Tabasco.
México.

Índice de tablas

Tabla 1. <i>Profesores investigadores participantes</i>	43
Tabla 2. <i>Comparación de herramientas de minería de datos</i>	51
Tabla 3. <i>Ventajas y desventajas en el uso de las herramientas de minería de datos</i> ...	52
Tabla 4. <i>Variables sociodemográficas</i>	54
Tabla 5. <i>Variables de carrera profesional y formación académica</i>	55
Tabla 6. <i>Variables de condiciones institucionales</i>	55
Tabla 7. <i>Variables de vida personal y familiar</i>	56
Tabla 8. <i>Variables de afrontamiento</i>	56
Tabla 9. <i>Variables de habilidades para la investigación</i>	56
Tabla 10. <i>Selección de algoritmo de minería de datos</i>	64

Juárez
México.

Autónoma de Tabasco.
XX

Índice de figuras

Figura 1.	Etapas del proceso KDD	18
Figura 2.	Población de estudio.	41
Figura 3.	Profesores investigadores encuestados.	42
Figura 4.	Interfaz de RapidMiner.	47
Figura 5.	<i>Interfaz gráfica de WEKA.</i>	49
Figura 6.	Interfaz gráfica de MOA.	50
Figura 7.	Implementación del proceso KDD.	53
Figura 8.	<i>Selección de los datos.</i>	58
Figura 9.	<i>Filtrado de datos para encontrar valores nulos.</i>	59
Figura 10.	Datos sin transformar.	60
Figura 11.	Uso del algoritmo Replace Missing Values.....	61
Figura 12.	Resultado del uso del algoritmo Replace Missing Values.	62
Figura 13.	<i>Importación del dataset</i>	65
Figura 14.	Almacenamiento del dataset.	66
Figura 15.	Lista de operadores.	67
Figura 16.	Proceso de creación de árbol de decisión.	67
Figura 17.	Árbol de decisión.	69
Figura 18.	<i>Algoritmo de agrupación k-means.</i>	70
Figura 19.	<i>Proceso para algoritmo K-means.</i>	71
Figura 20.	<i>Clústeres creados con RapidMiner.</i>	72
Figura 21.	Gráfico Plot.	73
Figura 22.	Proceso de creación de reglas de asociación.	74
Figura 23.	Reglas de asociación.....	75
Figura 24.	Matriz de confusión si el profesor cuenta con plaza definitiva	78
Figura 25.	Árbol de decisión trayectoria académica primera parte.	79
Figura 26.	Árbol de decisión trayectoria académica segunda parte	80
Figura 27.	Matriz de confusión condiciones institucionales	81
Figura 28.	Árbol de decisión condiciones institucionales primera parte.....	83
Figura 29.	Árbol de decisión condiciones institucionales segunda parte.....	84
Figura 30.	<i>Matriz de confusión vida personal y familiar.</i>	85
Figura 31.	<i>Árbol de decisión vida personal y familiar.</i>	86
Figura 32.	<i>Matriz de confusión afrontamiento.</i>	87
Figura 33.	<i>Árbol de decisión afrontamiento primera parte</i>	89
Figura 34.	<i>Árbol de decisión afrontamiento segunda parte</i>	90
Figura 35.	<i>Matriz de confusión habilidades para la investigación.</i>	91
Figura 36.	<i>Árbol de decisión habilidades para la investigación.</i>	92
Figura 37.	<i>Resultado de agrupamiento de trayectoria académica con RapidMiner.</i>	93
Figura 38.	<i>Resultados de agrupamiento de Condiciones institucionales con RapidMiner.</i>	94
Figura 39.	<i>Resultado de agrupamiento de vida personal y familiar con RapidMiner.</i> ...	95

Figura 40. *Resultado de agrupamiento de afrontamiento con RapidMiner.*..... 95

Figura 41. *Resultado de agrupamiento de habilidades para la investigación en RapidMiner.*..... 96

Figura 42. *Asignación de ítems en los clústeres.*..... 97

Figura 43. *Gráfica de agrupamiento en RapidMiner.*..... 98

Figura 44. *Resultados coincidencia de reglas de asociación.*..... 99

Figura 45. *Reglas de asociación con RapidMiner.*..... 100

Figura 46. *Descripción de las reglas de asociación con RapidMiner.*..... 101

Universidad Juárez Autónoma de Tabasco
México.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Capítulo 1. Generalidades

1.1 Antecedentes

En las universidades los factores que afectan a los profesores en su desarrollo profesional como investigador son muchos, estos factores no solo afectan a investigadores que laboran en México, sino también en otros países. Hoy en día se han desarrollado muchas estrategias para tratar de mitigar estos problemas a los que se enfrenta el profesor día con día.

Lo difícil es identificar cuáles son estos problemas, ya que cada profesor es diferente, con diferentes perspectivas y diferentes problemas que afrontan en su vida tanto personal como profesional, anteriormente para poder identificar estos problemas era muy tedioso y algo lento, ya que era mucha información que habría que tomar en cuenta, analizarla e identificar cuáles son los patrones que más se repiten entre los profesores.

Hoy en día existen muchas herramientas que permiten acortar y facilitar el análisis de estos datos, estas herramientas utilizan la minería de datos para la identificación de patrones, Espigares *et al.* (2009), mencionan que las técnicas estadísticas aplicadas en la minería de datos implican una generación de modelos estadísticos de carácter predictivo y correlacional con los datos almacenados en las plataformas.

En las instituciones educativas la minería de datos permite fundamentar mejor el uso de las Tecnologías de Información (TI) y les facilita el diseño de estrategias basadas

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

en la demanda y orientación de las plataformas virtuales por medio de las competencias y necesidades del entorno.

1.1.1 Factores que influyen en el desarrollo profesional del profesor

En el desempeño profesional del profesor investigador existen muchos factores que influyen en su desarrollo cotidiano, los cuales pueden afectar tanto su vida social, familiar y laboral. Estos factores en el entorno laboral pueden provocar problemas de salud, tales como estrés y cansancio entre otros.

Aguilar *et al.* (2015) mencionan que el ambiente educativo tiene características muy específicas tales como: más tiempo de clase, artículos, libros, capítulos de libros, reuniones, comités y participación en la actividad de la universidad, entre otras, lo que implica una serie de factores psicosociales que influyen en la satisfacción laboral, los tres más importantes son los problemas tradicionales, el trabajo que interactúa directamente con las personas y el trabajo irrestricto, cada uno de ellos es más propicio para un determinado tipo de trabajo. En cuanto a las actividades de los profesores investigadores, estas incluyen dos de los tres factores de riesgo: el trabajo consistente en interactuar con las personas y el trabajo irrestricto, porque el trabajo del profesor es el único trabajo sin horario, ya que éste realiza trabajo en casa.

Tardif (2004) menciona que un entorno de trabajo adecuado favorece el rendimiento y la autoestima. De hecho, una de las variables más importantes para crear este clima apropiado es la existencia de relaciones sociales adecuadas. Por lo tanto, hay

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

varios tipos de relaciones sociales en la enseñanza y estas relaciones pueden ser continuas, como las relaciones con los estudiantes, otros maestros u otros profesionales de otras instituciones educativas.

Es por esta situación que los factores que afectan a los profesores en su desarrollo profesional, pueden provenir de distintas fuentes (familia, trabajo, ambiente laboral, reglas y compañeros) estas mismas situaciones pueden provocar distintos problemas como estrés, cansancio, problemas de salud, entre otros.

Según Zuniga y Pizarro (2018) el estrés se relaciona con el trabajo, ya que es una respuesta emocional, física y conductual potencialmente patógena que puede afectar en las actividades laborales de los docentes y en sus resultados. Esto conduce a la pérdida de motivación, a volverse inadecuados y tienen sensación de fracaso, empeorando en términos de salud y relaciones interpersonales tanto dentro como fuera del lugar de trabajo.

Por otra parte, Fernández define el estrés como una "reacción psicofisiológica de tensión o esfuerzo que sufre una persona" (2014, p. 93). Lo que provoca un desgaste tanto mental, como físico, también menciona que dicho estrés puede provocar una respuesta positiva otorgando a la persona un desarrollo en sus capacidades, ya que bajo cierto estrés la persona aumenta su motivación y obtiene una respuesta favorable que lo hace más competitivo.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Gómez *et al.* (2016) mencionan que existen diferentes problemas de salud que enfrentan los profesores, y que estos problemas de salud están relacionados con los malos hábitos; mala alimentación, mucha inactividad, tabaquismo, entre otros, lo que puede provocar riesgos en la salud tales como: “trastornos psicológicos o psicosomáticos, que se manifiestan en fatiga, trastornos del sueño, ansiedad, estrés, depresión, trastornos del habla, enfermedades infecciosas y lesiones músculo-esqueléticas” (p. 1-2).

1.1.2 Minería de datos aplicados al rendimiento académico

La minería de datos educativos permite extraer conocimientos útiles y fáciles de entender de los datos académicos para resolver problemas relacionados con varios procesos de enseñanza. Una de las aplicaciones más populares de la minería de datos educativos es la predicción del rendimiento académico (Rico y Sánchez, 2017).

Peña (2014), menciona que la minería de datos en la educación está orientada al diseño, tarea, método y algoritmo, esto con la finalidad de analizar y comprender los datos en cualquier entorno educativo, también menciona que puede descubrir patrones y conocimientos de los datos.

Por otra parte, Martínez *et al.* (2015), sugiere que con el uso de la minería de datos se pueden distinguir características específicas y patrones que están relacionados directamente con el comportamiento del rendimiento académico, analizando distintos aspectos como sociales y económicos entre otros.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

En los profesores el rendimiento académico puede ser afectado por distintos factores, tales como problemas familiares, personales, laborales y de salud, por mencionar algunos. A través del uso de la minería de datos en la predicción de patrones es posible identificar las conductas provocadas por dichos factores; el uso de estas técnicas permitiría conocer cuál es factor que más afecta a los profesores.

1.2 Planteamiento del problema

1.2.1 Definición del problema

La importancia que tiene la investigación es muy considerable, ya que se ha convertido en la base del conocimiento y el proceso de aprendizaje. La investigación se refiere a un proceso sistemático y organizado, diseñado para responder preguntas. Esto permite incrementar el conocimiento y obtener información sobre lo desconocido. El propósito de la investigación es incrementar el conocimiento de un tema determinado.

Silva *et al.* (2020), mencionan que es necesario fortalecer los sistemas de investigación e innovación, el desarrollo de capacidades, el uso de tecnologías y redes científicas y la cooperación entre universidades y las industrias.

Por otra parte, Sandoval *et al.* (2013) menciona que las investigaciones que pueden realizar los profesores e investigadores se encuentran influenciadas por diversos esquemas de evaluación propuestos en el trabajo profesional, lo que ocasiona una desmotivación por parte del docente, ya que lo que busca el profesor es realizar investigaciones para obtener mejores ingresos económicos. Hoy en día un profesor

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

investigador en México que labore en alguna institución pública, llega a tener mejores incentivos económicos dependiendo de su desempeño laboral, existen instituciones y programas externos tales como: el Programa para el Desarrollo Profesional Docente (PRODEP), el Sistema Nacional de Investigadores (SNI) y el Consejo Nacional de Ciencia y Tecnología (CONACYT) que también ofrecen estos incentivos como ayuda al profesor investigador, para que pueda realizar sus investigaciones.

Silva *et al.* (2020) mencionan que, aunque se esperaría que una persona que se dedica a la investigación y a la ciencia, tuviera los mejores salarios, la realidad es diferente, ya que, aunque tenga una mejor preparación o el mayor grado de estudio, no hay una asociación directa a sus ingresos. Los autores también refieren que es de suma importancia valorar y reconocer el trabajo del personal que se dedica a la investigación.

Esto se debe a que el profesor investigador debe realizar investigaciones y al mismo tiempo dedicarse a dar clases, por lo tanto, debe de poner menos atención a una, entonces dependiendo de la carga de trabajo que tenga en ambas actividades, el profesor investigador se inclina ya sea por dar clases o investigar, lo que ocasiona, una baja productividad a la hora de realizar investigaciones y pueda provocar que no logre obtener el apoyo por parte de CONACYT, PRODEP o el SNI.

En la investigación de Silva *et al.* (2020), mencionan que en las universidades existen políticas desalentadoras que los limitan a desarrollar una carrera de investigación, como es la supresión de la modalidad de titulación por proyecto de investigación. Existe

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

la sobrecarga de actividades de docencia, tutorías, gestión e investigación lo cual hace que tengan poco tiempo para realizarlas todas satisfactoriamente.

Por lo tanto, la investigación hace que el profesor investigador lleve una carga laboral significativa, que puede afectar varios aspectos de su vida cotidiana, tanto en lo personal como en lo familiar. Actualmente el profesor investigador enfrenta muchas dificultades o barreras que le impiden poder realizar sus labores tanto académicas, como de investigación.

Diversos autores señalan barreras a las que se enfrentan los profesores al realizar su labor de investigación, Reyes y Hernández (2014) identifican cinco:

El tiempo es el principal obstáculo al que se enfrenta el profesor investigador para realizar investigaciones y publicaciones, ya que impartir clases, tutorías y otras funciones que realizan les consume el tiempo que tienen y no les alcanza para poder adelantar la investigación, es así que terminan usando su tiempo libre, lo cual puede llegar a afectarles en su entorno familiar.

En la formación no todos los investigadores están preparados, aunque cuenten con maestría no tienen la experiencia necesaria para investigar y publicar por sí solos, así que tienen que recurrir a otros investigadores para que los orienten o consultar artículos ya publicados para poder guiarse.

La definitividad indica que los profesores investigadores que no cuentan con una plaza no pueden ser acreedores de una beca de posgrado. Así que eso les impide

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

investigar y tener más publicaciones y esto a la vez le impide desarrollarse como profesor investigador.

El escaso acompañamiento indica que los maestros que no tienen experiencia y necesita del apoyo de algún profesor investigador con experiencia, terminan estancándose porque no saben cómo iniciar una investigación, ya que no cuentan con una asesoría por parte de los profesores más experimentados, puesto que ellos necesitan la interacción con otros investigadores para que con opiniones y críticas puedan construir conocimiento y realizar mejores investigaciones.

El escaso compromiso institucional indica que comúnmente las instituciones no brindan el espacio necesario para que los profesores investigadores puedan trabajar en un área que sea cómoda y privada, ya que comúnmente están en cubículos pequeños que son compartidos con otros profesores y están expuestos al ruido e interrupciones por parte de otros profesores y estudiantes lo que les dificulta concentrarse.

Ortega (2016), al igual que Reyes y Hernández (2014), considera que la falta de tiempo es el mayor obstáculo al que se enfrenta el profesor investigador. En su investigación clasifica las barreras en cuatro factores: características profesionales, características de la organización, calidad de la investigación y accesibilidad de la investigación, existen otros factores que se presentan en el profesor investigador un ejemplo es el estrés laboral, que puede ocasionar enfermedades físicas y emocionales.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Es debido que a la situación anterior, los profesores de la universidad están sobrecargados de trabajo, estrés, trastornos musculares y del sueño, ansiedad, fatiga, etc., todas estas situaciones se deben al trabajo y al uso prolongado de equipos informáticos, lo que conduce a un desgaste excesivo sonoro y visual, también se encontró que debido a la extensión de las jornadas laborales y al trabajo realizado en los días de descanso y feriados, se produjo un deterioro de la vida familiar, social y laboral de los profesores. La demanda laboral dentro de las instituciones ha provocado estrés, fatiga emocional y psicológica especialmente relacionada con la exigencia académica (Villarruel *et al.*, 2018).

En cuanto a la satisfacción laboral del profesorado influyen mucho las condiciones institucionales, en como las instituciones apoyan la labor del investigador respecto a la carga de trabajo que el docente maneja.

Por otra parte, en la vida personal y familiar del profesor investigador, se pueden observar variables como la paternidad y maternidad, actividades de casa, satisfacción personal, edad, género y el estado civil, que se encuentran relacionadas directamente con el estrés dentro de su entorno de trabajo (Esteras *et al.*, 2014), en consecuencia, la vida personal y familiar está ligado directamente con el afrontamiento y las estrategias que emplean los docentes para superar situaciones extraordinarias.

Sandoval *et al.* (2013) menciona que en término de género existen diferencias en cómo perciben la presión laboral, ya que los profesores investigadores hombres se

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

encuentran más estresados por demasiado trabajo o porque no tienen el tiempo para realizar su investigación.

Uno de los factores importantes no considerados es el número de profesores investigadores varones que tienen que lidiar con el trabajo académico y combinarlo con labores fuera de la institución, esto a su vez puede provocar conflictos debido a una constante presión en el trabajo académico. “En cuanto al estado civil, existe una gran diferencia entre estar casado y con familia, a estar soltero, ya que, si el profesor investigador se encuentra soltero, tienen menos libertad de dar clases” (p.19).

Una de las principales preocupaciones dentro del ambiente que viven los investigadores del nivel superior son las barreras a las que se enfrentan día con día que les impide desarrollar sus investigaciones y dentro de las universidades estas preocupaciones se convierten en un problema muy grande, actualmente no cuentan con alguna herramienta que permita determinar e identificar, cuáles son los patrones de conducta que adoptan al enfrentar estas barreras.

1.2.2 Delimitación de la investigación

La investigación se basó en el análisis de los factores a los cuales se enfrentan los profesores para su crecimiento profesional como investigadores, en donde los instrumentos de recolección de datos son las encuestas aplicadas a investigadores del nivel superior de una universidad pública de la república mexicana.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Alcances

- El estudio se enfocó en profesores investigadores del nivel superior de una universidad pública de México para conocer cuáles son las problemáticas familiares, laborales y personales a las cuales se enfrentan los profesores investigadores al realizar una investigación.
- Se conocieron los patrones de conducta que desarrollan el crecimiento en profesores investigadores de una institución de educación superior en México.
- En esta investigación se muestran los conceptos y las definiciones de los factores de conducta que sirven como base para realizar una investigación de tipo cuantitativa, la cual se realizó con información real; para recabar la información pertinente se empleó un cuestionario para conocer las características del perfil académico relacionadas con el desempeño del profesor investigador.

Limitaciones

- Se consideraron cinco factores de investigación: trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la investigación.
- Los datos del trabajo de campo son de la encuesta aplicada a profesores investigadores en una institución de educación superior de México.
- La información de los profesores se encuentra en un *dataset*.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

1.2.3 Pregunta de investigación

¿Cómo identificar cuáles son los patrones de conducta que afectan a los profesores investigadores, al realizar sus investigaciones?

1.3 Objetivos

1.3.1 Objetivo general

Identificar patrones de conducta, asociadas con variables de trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la investigación de profesores investigadores de educación superior, en una universidad pública de México, aplicando técnicas de minería de datos.

1.3.2 Objetivos específicos

- Aplicar la técnica de árbol de decisión para una representación analítica y gráfica de los patrones de trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades de los profesores investigadores.
- Emplear la técnica de agrupación para identificar las similitudes entre los patrones de conducta de trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la investigación de los profesores investigadores.
- Aplicar reglas de asociación para identificar las relaciones entre trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la investigación de los profesores investigadores.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

- Analizar los patrones de conducta obtenidos de trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la investigación para la toma de decisiones futuras.

1.4 Justificación

La minería de datos es el proceso de extraer conocimiento que puede ser de utilidad en la toma de decisiones, basándose en grandes volúmenes de datos, el continuo crecimiento en almacenamiento de información, proporciona nuevas posibilidades de analizar comportamientos.

Martínez y Hernández (2017) mencionan que el principal objetivo de la minería de datos es obtener información que existe entre la relación de las variables que se encuentran en una base de datos, esta información puede ser diversa. La minería de datos aprovecha la información para obtener datos importantes.

Es por eso que esta investigación utilizó herramientas de minería de datos para la identificación de patrones de conducta ya que, en las universidades tanto de México, como en universidades de todo el mundo existen barreras que impiden a los profesores realizar sus investigaciones. De acuerdo a esto se puede observar que existen factores que influyen en el desarrollo como investigadores, esto debido a que no se le da la importancia a la investigación.

Si bien se sabe que la sociedad mexicana sigue avanzando y buscando corregir las falencias, como por ejemplo recursos insuficientes para las instituciones, poco apoyo

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

económico a los profesores, entre otras, que han llevado al país a la situación que atraviesa actualmente en la educación formal de muchos mexicanos. Frente a estos cambios cotidianos en el entorno educativo, las IES han comenzado recientemente a realizar ajustes para mejorar la calidad de los servicios que brindan. Casi la mayoría de estas acciones se han centrado en la evolución del paradigma educativo, la certificación y recertificación de normas ISO, y diferentes procesos de certificación basadas en disciplinas a las que pertenece. Los principales participantes de estos avances son los profesores investigadores (Cárdenas *et al.*, 2013).

Se cree que los docentes, especialmente los universitarios, enfrentan exigencias constantemente, como lo es la toma de decisiones, introducir cambios e innovaciones efectivos, mantenerse al tanto de los avances tecnológicos y de áreas de conocimiento específicas. Por lo tanto, el estrés que existe en el desarrollo de las tareas profesionales es inherente a su desempeño profesional. (Sánchez y Maldonado, 2003).

En cuanto al SNI en su reglamento menciona que los docentes con título de doctorado, bajo contrato, están comprometidos con 20 horas de docencia y 20 horas de investigación, debiendo producir artículos de comunicación científica para obtener apoyo económico en función de su producción (CONACYT, 2020).

Del mismo modo, la situación relacionada con la percepción económica recibida por profesores de instituciones de educación superior no siempre tiene que ver con sus actividades escolares, las IES públicas dependen del presupuesto del gobierno, por lo

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

tanto, esto lleva a muchos profesores a realizar actividades distintas a la docencia para poder satisfacer sus necesidades.

Es por eso que se llevó a cabo esta investigación con la ayuda de técnicas de minería de datos, para obtener un indicador que permita analizar cuáles son los principales patrones de conducta y cuáles son las principales barreras que les impide a los profesores realizar su trabajo.

1.5 Metodología utilizada

Para el desarrollo de esta investigación se utilizó un diseño metodológico exploratorio, el cual permitió un acercamiento al problema de la investigación con un enfoque de investigación de carácter cuantitativo, mediante la aplicación de encuestas presenciales a profesores investigadores de una universidad pública en México. Martínez (2011) menciona que el método cuantitativo emplea la recopilación y el análisis de datos para dar respuesta a preguntas de investigación y comprobar hipótesis ya establecidas, las cuales están basadas en mediciones numéricas, conteo y el uso frecuente de datos estadísticos para establecer patrones más exactos de comportamiento de una población.

1.6 Fuentes de investigación

Como fuente de recolección de datos se utilizó un cuestionario, el cual fue aplicado a 250 profesores de nivel superior de una universidad pública de México, el propósito del cuestionario fue conocer las características del perfil académico relacionado con el desempeño como investigador, el cuestionario se dividió en secciones la primera sección

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

es la de datos personales, la cual incluye: edad, sexo y estado civil, en la segunda sección se encuentran los datos laborales en esta sección además de preguntas como ¿A qué división académica pertenece?, ¿Cuántos años lleva laborando dentro de la institución?, también se encuentran los factores de trayectoria académica, condiciones institucionales, vida personal y familiar y afrontamiento, estos 4 factores son importantes para conocer cómo afectan en el desarrollo del profesor investigador y de qué manera afrontan los distintos problemas que ellos enfrentan en su ambiente laboral, en la tercera y última sección del cuestionario, se aplica una autovaloración de sus condiciones como investigador, y se encuentra el factor de habilidades para la investigación, con el que se pretende conocer cuáles son las habilidades investigativas que más desarrollan los profesores. El cuestionario se consideró el método más adecuado para vincular datos con teoría, porque ayuda a describir, analizar e interpretar los hallazgos y compararlos con aspectos teóricos establecidos. En el apéndice A, se puede observar de manera detallada el cuestionario que fue aplicado a los profesores investigadores.

1.7 Técnica de recolección de datos

Es importante señalar que la información obtenida debe incluir aspectos relacionados con el trabajo de obtención, análisis y presentación de la información. Por lo tanto, se utilizará el cuestionario como técnica en la recolección de datos para el procesamiento y análisis de la información. Los datos recabados corresponden a profesores de una universidad pública en México, con la intención de identificar cuáles

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

son los patrones de conducta que afectan su desarrollo como investigadores, los datos fueron almacenados en un archivo *dataset* para su uso con la herramienta de minería de datos.

1.8 Metodología para aplicar

El poder obtener conocimiento a partir de conjuntos de datos ha sido considerado por diferentes investigadores como un proceso de investigación importante en los sistemas que utilizan bases de datos.

Timarán *et al.* (2016) mencionan que el hallazgo de información que brinda un análisis y comprensión de datos, dentro de un conjunto de datos en el proceso KDD es automático en el cual combina tanto la comprensión de los datos como el análisis de ellos. Este proceso trata específicamente en extraer los patrones en modo de reglas, para que le facilite al usuario el análisis de la información.

Es por eso que como parte de la metodología el proceso KDD se adecua a las necesidades de la investigación ya que brinda la facilidad de analizar e identificar los patrones de conducta de los profesores investigadores. En el proceso KDD que se muestra en la Figura 1 se observan las diferentes etapas que conlleva el proceso KDD las cuales se resumen en:

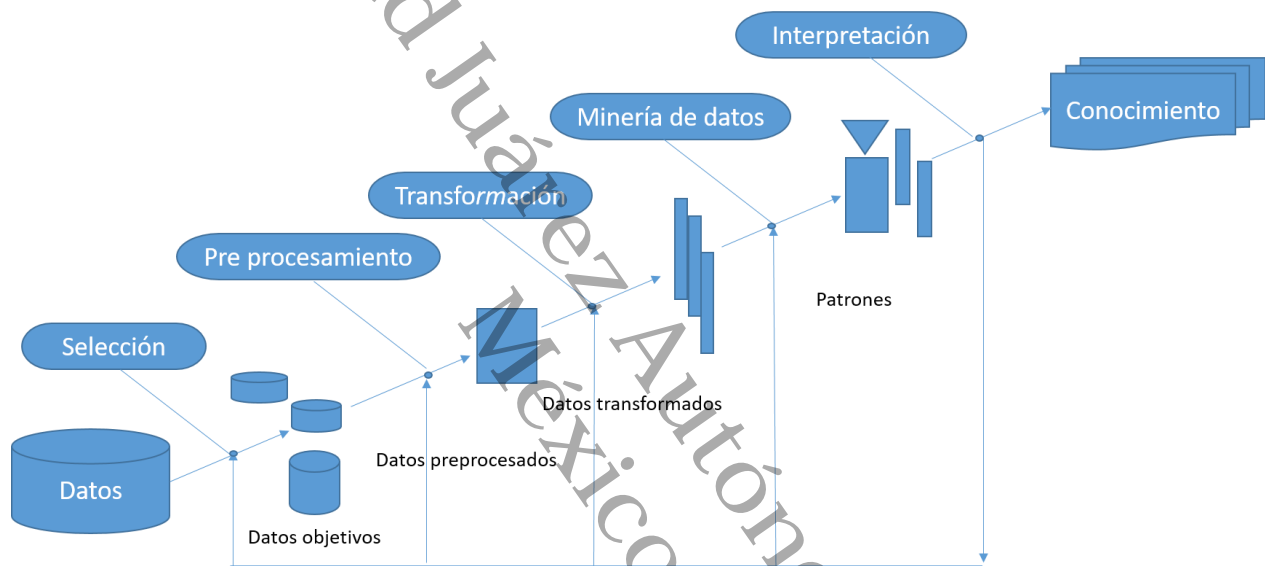
- Selección.
- Pre-procesamiento/limpieza.
- Transformación/reducción.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

- Minería de datos (*data mining*).
- Interpretación/evaluación.

Figura 1

Etapas del proceso KDD.



Nota: Timarán *et al.* (2016).

Dentro de las herramientas de minería de datos existen diferentes técnicas que permiten el análisis de la información, cada técnica se puede utilizar dependiendo el tipo de información que se quiere analizar, para esta investigación se optó por la tarea de *Clustering*, ya que esta es la agrupación de los datos en un conjunto particular de objetos en función de sus características, agrupándolos de acuerdo a sus similitudes, es una técnica de aprendizaje no supervisado y que pertenece directamente a las técnicas descriptivas de la minería de datos, puesto que esta permite el análisis y brinda una

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

descripción del comportamiento de los datos que se encuentren, en otras palabras, es adecuado para situaciones en las que la clasificación de los datos es desconocida.

El método de agrupación se obtiene directamente de los datos de entrada utilizando medidas similares, en cuanto a la seguridad de la aplicación está sujeta al tipo de datos y de la limpieza de la información obtenida Jiawei *et al.* (2012).

Con la tarea de *Clustering* se analizaron las diferentes variables asociadas a los factores que afectan a los profesores investigadores. Los resultados se obtuvieron a través del proceso de minería que se encuentra dentro de la herramienta *RapidMiner* en el cual se realizaron varias pruebas para la consolidación de los resultados. En dichas pruebas se visualizaron el comportamiento de los factores que afectan el desarrollo de los profesores como investigadores, de esta manera se realizaron conclusiones e interpretaciones a partir de la generación de grupos que se formaron.

Con la ayuda de la herramienta de minería de datos fue posible crear reglas de asociación con el algoritmo de Apriori, el cual permite conocer la relación que existe entre la información que se esté analizando dentro del *dataset* y aquellos patrones con rasgos de conducta entre las variables de los profesores investigadores. Sáenz *et al.* (2017), mencionan que estas reglas son utilizadas para conocer la veracidad de los parámetros en conjunto con los datos analizados. También menciona que el algoritmo Apriori crea varias reglas de asociación con pocos ítems dentro de la base de datos, de esta manera

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

es posible seleccionar las reglas de mayor validez en base a los parámetros de las mismas reglas.

Otras de las técnicas que se utilizaron en esta investigación, es la de árbol de decisión, esta técnica es un modelo de predicción, que permite al usuario obtener de manera gráfica datos que facilitan la toma de decisiones ya que crea un esquema de las diferentes posibilidades de una problemática y analiza las diferentes opciones que se pueden tomar.

Riascos y Molina (2016) mencionan que el uso de este método permite predecir diferentes escenarios posibles. Por otra parte, Damodaran (2014) menciona que los árboles de decisiones son muy prácticos en la toma de decisiones, ya que permiten valorar la mejor respuesta, asociando opciones y acciones de diferentes eventos, a través de los árboles de decisión las empresas pueden decidir cómo actuar ante un riesgo o cualquier circunstancia.

Es por eso que se optó por utilizar este método ya que con él se pudo crear una representación gráfica y analítica de los patrones de trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades de los profesores investigadores y permitió identificar cual es la variable, con la que los profesores más se identifican.

Es así que al cumplir con el proceso KDD e implementar los métodos y algoritmos de agrupamiento, reglas de asociación y árboles de decisión se encontraron los

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

resultados que brindaron los parámetros que facilitaron conocer cuáles son esos factores que más afectan a los docentes en su desarrollo como profesores investigadores.

1.9 Herramienta de minería datos

La herramienta de minería de datos que se utilizó para el análisis de los datos recabados fue *RapidMiner*, ya que es una herramienta muy completa que permite la implementación fácil y rápida de los algoritmos de agrupamiento, reglas de asociación y árboles de decisión.

Ruiz y Romero (2018) la definen como una herramienta libre, muy superior a WEKA, ya que el manejo de los resultados y las gráficas son de mayor calidad y con la característica de poder adjuntar diferentes extensiones, que aumentan el potencial del *software*.

Es así que el uso de esta herramienta facilitó de manera gráfica los resultados y con ella la identificación de los patrones de conducta de los profesores investigadores, para así poder cumplir con cada uno de los objetivos de esta investigación.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Capítulo 2. Marco de la investigación

2.1 Marco teórico

Esta investigación se basó en el análisis de los factores que enfrentan los profesores investigadores durante su crecimiento profesional y de cómo el uso de la minería de datos ayudo a identificar los patrones de conducta que afectan en la productividad de los docentes, por lo tanto, se revisaron las teorías de diferentes autores para así poder comprender los factores anteriormente mencionados.

Para ello es muy importante resaltar que existe una línea delgada entre los conceptos de minería de datos, análisis de datos y el aprendizaje automático, tanto la minería de datos como el aprendizaje automático, se fundamentan en la base de extraer información, pero con la diferencia de que la minería de datos se utiliza para identificar distintos patrones que permitan predecir tendencias futuras, en cambio el aprendizaje automático aunque incluye las bases de la minería de datos su función es crear relaciones automáticas y aprender de la información encontrada y así crear algoritmos que permitan un mejor funcionamiento de los sistemas, en cambio el análisis de datos se utiliza después de obtener la información, para comprender y examinar los datos encontrados, y así dar soluciones al problema en cuestión.

Este tipo de herramientas permite la extracción de datos, con lo cual es posible identificar patrones que sean similares, con estos patrones es posible mejorar el proceso en la toma de decisiones, en el caso de los profesores investigadores con la ayuda de

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

estas herramientas se pueden identificar cuáles son los obstáculos que dificultan su desarrollo como profesor investigador a lo largo de sus carreras.

Por otra parte, dentro de los factores se encuentra la trayectoria académica, Pérez (2016), menciona que las trayectorias laborales individuales pueden convertirse en una puerta de entrada para mirar lo institucional, también menciona el fuerte compromiso que los profesores mantienen con la institución, además del papel fundamental que desempeñan en momentos críticos y definitorios. De esta manera destaca la importancia de tres elementos: participación, posición y rol, ya que estos tres conceptos permiten observar los cambios que pueden experimentar los profesores en función de su antigüedad, formación, práctica profesional y redes de relación, en otros aspectos, los cuales se concretarán en posiciones, niveles de participación y roles específicos.

Tomando en cuenta esto se puede decir que la trayectoria en el docente puede pesar mucho al momento de adquirir un puesto más alto dentro de alguna institución, es por eso que como parte de su trayectoria tratan de entrar a programas como al del SNI para desarrollarse como profesor investigador.

El objetivo del SNI es contribuir en la importancia de la investigación científica, tecnológica y la innovación nacional, fomentando la formación de investigadores con un nivel alto en conocimiento científico y tecnológico para aumentar la educación, el rendimiento y el bien social (CONACYT, 2019).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Aunado a la formación académica también influyen las condiciones institucionales como parte del crecimiento académico de un profesor investigador, ya que, dadas las condiciones académicas, el docente puede o no trabajar cómodamente.

Santizo (2016) refiere que es necesario analizar más de cerca el proceso educativo, la forma en que opera la escuela y, además, considerar los factores que pueden afectar el comportamiento y la respuesta de los docentes a sus instituciones educativas y condiciones sociales de diferentes formas en que realizan sus tareas.

Rueda *et al.* (2014) dividen los factores y variables contextuales del desarrollo docente en tres niveles: macro, meso y micro.

El nivel macro corresponde al entorno de políticas nacionales e internacionales que orienta las metas, planes y programas generales de enseñanza de las IES; el nivel meso considera las condiciones institucionales para el desempeño de la labor docente, tales como planes y programas institucionales, condiciones laborales de los docentes, organización académica, las características de la labor docente e infraestructura y equipamiento institucional; el nivel micro se limita a las condiciones del aula que afectan la práctica docente en cuanto a infraestructura y equipamiento, características del docente y características del alumno. De esta forma, puede organizar diversas iniciativas para profesores y estudiantes.

Otro de los factores a los que se enfrentan los profesores a lo largo de su desarrollo académico es el afrontamiento o también llamado *coping*, Guerrero (2002), menciona que

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

las estrategias de afrontamiento utilizadas por los profesores varían mucho, incluso si son las mismas es posible utilizar diferentes estilos de afrontamiento teniendo un resultado exitoso, cabe resaltar que no existe la mejor manera de afrontar todas las situaciones.

Dentro de los afrontamientos que el profesor investigador enfrenta día a día en su desempeño académico existen diversos factores de riesgo y exigencias laborales. Entre los más importantes se encuentran los factores psicosociales que repercuten en la salud mental del trabajador debido al estrés que generan.

Lemos *et al.* (2019) mencionan que existe una relación muy estrecha entre los factores psicosociales y el estrés, esto debido a la carga laboral que el profesor lleva, de acuerdo con esto Villamar *et al.* (2019), en su investigación atribuyen que los factores psicosociales de mayor relevancia en profesores son: problemas con la institución, con alumnos y sociales.

Otro de los afrontamientos a los que se enfrentan los profesores investigadores de acuerdo a las funciones que desempeñan para llevar a cabo la investigación es el estrés. Esto derivado a que por lo regular tiende a llevar trabajo a casa, trabajar en sus horas libres, en vacaciones, días festivos, etc.

Para desarrollar el afrontamiento de estrés en el docente es necesario que maneje las demandas de las habilidades investigativas que se le exigen como profesor investigador, tales como saber administrar y planificar su tiempo, compromiso ético y social, entre otras.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

9 El último factor son las habilidades investigativas, de las cuales Moreno (2005), hace referencia a la expresión como un conjunto de habilidades de diversa naturaleza, que empiezan a desarrollarse desde antes de que el individuo tenga acceso a procesos sistemáticos de formación para la investigación, que en su mayoría no se desarrollan solo para posibilitar la realización de las tareas propias de la investigación, pero que han sido detectadas por los formadores como habilidades cuyo desarrollo en el investigador, en formación o en funciones, es una contribución fundamental para potenciar que este pueda realizar investigación de buena calidad.

En el trabajo, se considera que este concepto es el más íntimamente relacionado con el proceso de formación de pregrado, porque muestra las habilidades investigativas como un eje horizontal en el proceso sustantivo.

2.2 Marco referencial

10 2.2.1 Afrontamiento, vida personal y familiar de profesores que pertenecen al Sistema Nacional de Investigadores (SNI) de una universidad del sureste de México

10 En el artículo realizado por Ramón *et al.* (2020), se presenta una investigación de tipo cuantitativo transversal y de alcance descriptivo-exploratorio sobre el afrontamiento, vida personal y familiar de profesores que pertenecen al SNI de una universidad del sureste de México. Se realizó un cuestionario a 121 (96%) de los investigadores adscritos al SNI de los cuales 72 (59.5%) son hombres (M edad = 43.3, DE = 9.5 años) y 49 (40.5%)

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

mujeres (M edad = 44.2, DE = 8.4 años). El instrumento de recolección de la información fue adaptado y validado. Una vez establecida la validez de contenido mediante juicio de expertos hicieron los ajustes semánticos identificados.

Los resultados del estudio realizado por Ramón *et al.* (2020), de acuerdo a la técnica de recolección de datos dan como conclusión que los profesores investigadores de la Universidad Juárez Autónoma de Tabasco, nunca han tenido problemas con su familia por el trabajo, estos comprenden las demandas que este implica, casi siempre comparten actividades de recreación con su familia, siempre asisten a las actividades académicas de sus hijos, a veces asisten a eventos sociales no relacionados con la universidad, apoyan en las labores del hogar, siempre acostumbran a salir de vacaciones con su familia, a veces conviven con sus compañeros de trabajo fuera del horario laboral y a veces realizan actividades recreativas con sus amigos, se sienten satisfechos con su vida personal y su vida familiar. Casi siempre mantienen un buen estado de salud, regularmente saben manejar el estrés y casi siempre ha sabido sobrellevar el manejo de las presiones de tiempo en el trabajo.

2.2.2 Las habilidades investigativas como eje transversal de la formación para la investigación

En el artículo realizado por Martínez y Márquez (2014), se presenta una investigación de tipo cualitativa sobre las habilidades investigativas como eje transversal de la formación para la investigación. En esta investigación, partiendo de la lógica de la

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

formación de competencias investigadoras, la formación investigadora se considera insuficiente, ya que en el campo de la docencia y la investigación docente es necesario profundizar en la etapa de orientación consciente de estas conductas. Con el desarrollo de su investigación lograron identificar como parte de la formación y el desarrollo de las habilidades para la investigación lo siguiente:

Las habilidades de investigación representan la mayoría de la formación investigadora, por lo que los problemas teóricos se pueden resolver de manera que se comprendan conscientemente los métodos científicos y desarrollen progresivamente los métodos de acción, para solucionar problemas en la práctica del campo académico, laboral y de investigación.

La clasificación de las habilidades investigativas puede basarse en la lógica del método científico para determinar las constantes en una investigación, de esta manera al desarrollar estas clasificaciones y determinar las constantes de la investigación es posible resumir la función de las actividades investigativas científicas y el desarrollo de estrategias meta cognitivas.

2.2.3 **Análisis del desempeño del profesorado universitario en el uso de MOODLE a través de técnicas de minería de datos: propuestas de necesidades formativas**

En el artículo realizado por Camacho *et al.* (2018), se presenta un trabajo basado en análisis del desempeño del profesorado universitario en el uso de *Modular Object*

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Oriented Dynamic Learning Environment (MOODLE) en la Universidad Autónoma de Yucatán (UADY) a través de técnicas de minería de datos. Esta investigación se consideró no experimental debido a que los datos de las variables no se manipularon, la población estudiada fueron los profesores investigadores que utilizaron UADY Virtual.

En la investigación de Camacho *et al.* (2018), se analizó el desempeño del uso de los sistemas de gestión del aprendizaje denominado UADY Virtual, teniendo como muestra 484 profesores de nivel posgrado. UADY Virtual es una herramienta de uso para profesores como apoyo en sus clases presenciales y en línea. Teniendo como objetivo principal reconocer cuales son los patrones de comportamiento del profesor universitario con el uso de esta herramienta tecnológica.

En cuanto al análisis de la investigación, el estudio se realizó teniendo en cuenta un solo factor, para poder comparar el uso de la herramienta UADY virtual, con ciertas áreas de conocimiento de la institución. El estudio concluyó que las diferencias entre el uso del sistema y los diferentes campos de conocimiento no se encontraron diferencias. Por lo tanto, se concluye que los 484 docentes participantes en el estudio utilizan las actividades de tareas como herramienta principal. Mediante el uso de este sistema por parte de los docentes para apoyar sus necesidades de formación para las clases presenciales, se puede concluir que, aunque el hecho es que el desempeño general de la población es pobre, se pueden ver diferentes desempeños entre los docentes cuando se analiza por área de conocimiento.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

2.2.4 Minería de datos aplicada a la identificación de factores de deserción universitaria en programas de pregrado

4 En el artículo realizado por Choque (2019), se presenta una investigación bajo el enfoque cuantitativo y de carácter descriptivo sobre la minería de datos. En esta investigación se aplicó la técnica de *Clustering* con el algoritmo *K-means* para la identificación de factores de deserción universitaria en programas de pregrado y para analizar el estado de los datos históricos de la universidad, para así obtener una visión más amplia de la situación de deserción.

Se logró evaluar el modelo construido y se verificó su validez analizando los campos relacionados con los factores potenciales de deserción. Se aplicó la técnica *Clustering* con el algoritmo *K-means* pertenecientes a la minería de datos en cinco pruebas de acuerdo al número de centroides recomendados, a partir de los cuales se verificaron comportamientos de los distintos campos, identificando así factores de deserción.

4 Este artículo al ser una investigación basada en minería de datos de tipo descriptivo, examinó y explicó la información histórica de la Escuela Profesional de Ingeniería de Sistemas (EPIS) desde sus inicios con el comportamiento de patrones de deserción universitaria en el plan 1993 y los cambios hacia los planes 2005, 2013 y 2016. Lo que le proporcionó a la universidad un amplio conjunto de resultados que

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

permitirá a sus directores hacer un análisis retrospectivo, entendiendo lo acontecido en el pasado para tomar decisiones futuras basadas en el aprendizaje de sus antecedentes.

2.2.5 Factores que influyen en la productividad científica en una universidad pública. La perspectiva de los académicos

En el artículo realizado por Silva *et al.* (2020), se presenta una investigación de tipo cualitativa, con enfoque hermenéutico sobre los factores que influyen en la productividad científica en una universidad pública desde la perspectiva de los académicos. En la parte cualitativa de la investigación se realizaron entrevistas semiestructuradas a investigadores de una universidad pública del sureste de México, los cuales son profesores de tiempo completo del área de educación con grado de doctor, se refiere que estos no pertenecen al SNI. En total participaron trece académicos, siete mujeres y seis hombres. Como resultado de la investigación se obtuvieron los siguientes datos:

Como factores personales los investigadores entrevistados se refirieron a aspectos de formación académica y valoración del costo personal de la carrera de investigador como limitante de la productividad científica. En la formación para la investigación, los profesores admitieron que en su formación doctoral no adquirieron las habilidades que les permitan desarrollar proyectos de investigación científica, reconocieron que el dominio del inglés no lo han desarrollado de manera suficiente y que tienen limitaciones con la tecnología para llevar a cabo las actividades de investigación.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

2.2.6 Utilización de técnicas de minería de datos para la identificación de rasgos de comportamiento en procesos de aprendizaje colaborativo en modelos de e-learning y b-learning

De acuerdo a la investigación de Núñez (2016), donde presenta un estudio sobre el uso de técnicas de minería identificando patrones de comportamiento en el proceso de aprendizaje en las modalidades *e-learning* y *b-learning* mencionan que estas son cada vez más utilizadas ya que pueden agrupar volúmenes grandes de información importante para examinar y comprender el comportamiento de los estudiantes.

Para lograr los resultados en dicha investigación implementaron técnicas de minería de datos y la metodología de proceso de análisis jerárquico, para identificar las características del comportamiento en el aprendizaje colaborativo, el análisis diferentes actividades, a partir de información extraída de los registros de error que se obtienen de los servidores y de los reportes de actividad generados en las plataformas de gestión de cursos virtuales.

En la investigación se analizaron de igual manera, la interactividad en estos sistemas con el proceso de análisis jerárquico. Es por ello que utilizaron *Expert Choice*, *Weka*, *software* de minería de datos.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

2.2.7 Modelo de minería de datos para identificación de patrones que influyen en el aprovechamiento académico

6 En la tesis realizada por Hernández (2015), se presenta una investigación sobre el uso de técnicas de minería de datos para identificar patrones de comportamiento con la finalidad de predecir la causa del fracaso y la deserción escolar. En la investigación se identificó cuáles son las variables que intervienen en el aprovechamiento académico, indispensables para tomar decisiones y realizar acciones pertinentes.

13 En la investigación se utilizó la metodología *Cross Industry Standard Process for Data Mining* (CRISP-DM), que divide el proceso de minería de datos en seis fases que se relacionan entre sí de forma iterativa. También menciona que se aplicaron los modelos de redes neuronales, árboles de decisión y *Clustering K-means* para analizar el comportamiento. Los resultados de los modelos fueron calculados a partir del conjunto de datos de pruebas, los cuales indican los modelos predictivos que arrojaron resultados positivos.

2.3 Marco conceptual

2.3.1 Patrones de conducta

52 Corredor *et al.* (2009), mencionan que el patrón de conducta no solo tiene como característica el rasgo de personalidad o la reacción puntual a una situación diferente, sino que también tiene una reacción de una persona predispuesta biológicamente para una situación desafiante. Estas engloban la acción-emoción y las muestran en un

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

ambiente de demandas desafiantes, teniendo como conducta la impaciencia, la hostilidad. Por otra parte, Delgado (2017) define este concepto como la forma en que el ser humano piensa, siente, reacciona y actúa en alguna situación determinada.

2.3.2 Minería de datos

4 Para la ESIC *Business & Marketing School* (2018), la minería de datos es un proceso de identificación de información relevante extraída de grandes volúmenes de datos, con el objetivo de descubrir patrones y tendencias estructurando la información obtenida de un modo comprensible para su posterior utilización.

2.3.3 Barreras de datos

Fresneda *et al.* (2012), definen las barreras como aquellos factores que impiden la implementación de un cambio en la práctica profesional.

2.3.4 Variable

2 La variable es todo aquello que se mide, controla y se estudia en una investigación. Núñez (2007), lo define como un concepto clasificatorio pues toma valores diferentes los cuales pueden ser cuantitativos o cualitativos. También denomina variable a las propiedades, constructos o características que adquieren diversos valores. Es un elemento constitutivo de la estructura de la hipótesis, es decir del enunciado de la hipótesis que establece su relación.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

2.3.5 Proceso KDD

Timarán *et al.* (2016) mencionan el proceso KDD como aquel proceso automático en el cual el descubrimiento y el análisis se combinan. Este proceso implica extraer patrones en modo de reglas o funciones de los datos para el análisis del usuario.

2.3.6 Análisis de datos

Figueroa (2016), menciona que el análisis de datos se basa en identificar los principios elementales de la información y comprobar las diferentes cuestiones que se planteen, para así poder dar una respuesta más clara y concisa.

2.3.7 Aprendizaje automático o *machine learning*

Hinestroza (2018), menciona que es una herramienta diseñada que tiene el propósito de mejorar el análisis de datos, en funcionamiento de predicciones futuras, esto mediante el uso de diferentes algoritmos que se basan en los distintos tipos de información ya sea información antigua o actualizada, permitiendo una optimización de la implementación de nuevos sistemas.

2.4 Marco tecnológico

Para el desarrollo de esta investigación, es importante conocer qué herramientas y *software* se pueden utilizar para la minería de datos.

2.4.1 Waikato Environment for Knowledge Analysis (WEKA)

Se distribuye de forma gratuita con *General Public Licencie* (GPL), dentro de los *software* de minería de datos, WEKA destaca por la cantidad y la eficiencia de sus

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

algoritmos, fue desarrollada por la Universidad de *Waikato*, una de sus funcionalidades es permitir la programación en el lenguaje *Just Another Vague Acronym* (JAVA), para la creación de algoritmos más sofisticados, para realizar un análisis de datos y modelado predictivo, todo esto a través de una interfaz amigable con el usuario, WEKA es multiplataforma, así que puede ser utilizado en distintos sistemas operativos (Rodríguez y Díaz, 2009)

2.4.2 Yet Another Learning Environment (YALE)

Herramienta muy flexible para el uso en minería de datos, fue desarrollada en la universidad de Dortmund, el lenguaje de programación en el que está escrito YALE, es JAVA, este *software* trabaja bajo las licencias GPL, capaz de ofrecer representaciones de los datos de dispersión en 2D y 3D (Rodríguez y Díaz, 2009).

2.4.3 RapidMiner

Software de minería de datos, creado por la compañía *Rapidminer*, fue actualizada en el 2001 como una iniciativa de seguir con el proyecto de YALE. *RapidMiner* está disponible para los sistemas operativos *Windows*, *Mac OS* y *Linux* (Mariño, 2014).

lonos inc. (2018) menciona que *RapidMiner* está escrita en lenguaje JAVA y es capaz de importar tablas de *Excel*, archivos SPSS y conjunto de datos de diferentes bases de datos, también integra la minería de datos de WEKA y R., lo que resalta la versatilidad del *software*.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

2.4.4 Massive Online Analysis (MOA)

Frameworks gratuito de minería de datos el cual permite crear e implementar algoritmos de clasificación y *Clustering*, ya que es de código abierto, en los algoritmos de clasificación es posible asignar una categoría a ejemplos que aún no se hayan utilizado, en cambio en el algoritmo de *Clustering* se utiliza para dividir los datos en grupos que sean similares (Jaramillo *et al.*, 2015).

2.4.5 Servicios en la nube para minería de datos

Parra (2019) menciona que la computación en la nube o *cloud computing*, los lenguajes de programación son las herramientas básicas para la construcción de *software*, existen diversos tipos de lenguaje de programación, con los cuales se pueden construir herramientas para minería de datos.

2.4.6 Just Another Vague Acronym (JAVA)

JAVA es un lenguaje de programación que su principal característica es ser un lenguaje compilado e interpretado, el código que se genera en la compilación de JAVA se llama *bytecodes*, el cuál es interpretado en un entorno de ejecución llamado *Java Runtime Environment* (JRE), uno de los puntos fuertes de JAVA es su seguridad, ya que cuenta con distintos niveles de seguridad, desde la programación, hasta el poder restringir el acceso a los recursos del sistema desde su gestor de seguridad (Belmonte, 2004).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

2.4.7 Python

Fernández (2012), define a *Python* como un lenguaje de programación de alto nivel interpretado y multipropósito, que permite acortar el tiempo en el desarrollo de un programa, ya que no es necesario compilar, esto debido a que el intérprete se puede utilizar de modo interactivo, *Python* permite desarrollar programas compactos, legibles y compatibles.

2.5 Marco legal

Como parte de esta investigación es necesario revisar el aspecto legal que pueda afectar de manera directa o indirecta esta investigación. Cuyos aspectos están basados en leyes, licencias, políticas, manuales o lineamientos es por eso que en este apartado se muestran tales aspectos.

2.5.1 Licencia Pública General (GNU)

GNU o GPL por su acrónimo en inglés, es la licencia de *software* libre más conocida en todo el mundo, fue diseñado por Richard Stallman con la intención de que todo *software* que se creará en GNU fuera de uso libre y no privado. Este es el llamado *copyleft*, ya que utiliza las leyes de derechos de autor para otorgar libertad a los usuarios (Bain, 2009).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

2.5.2 Ley Federal de Protección de Datos Personales en Posesión de los Particulares

7 El 5 de julio del 2010 se expide la Ley Federal de Protección de Datos Personales en Posesión de los Particulares, ley que fue creada con el objetivo de dar protección a los datos personales.

En el capítulo uno: disposiciones generales se encuentra el artículo uno, el cuál menciona que la ley tiene como objeto la protección de los datos personales que esté en posesión de alguna institución u otra persona, esto con el propósito de regular y proteger el uso de estos datos y así garantizar la privacidad de las personas (Ley Federal de Protección de Datos Personales en Posesión de los Particulares, 2010).

Los datos recabados dentro del *dataset* proporcionados por los profesores investigadores de una universidad pública en México, fueron utilizados exclusivamente para uso académico de esta investigación, siendo así los datos personales recabados de los participantes para cumplimiento de esta ley no serán divulgados para salvar guardar la privacidad de los profesores investigadores.

2.5.3 Licencia de software propietario

Arriola *et al.* (2011), mencionan que el *software* de propietario es privado, mientras una persona tenga sus derechos de autor, esta tendrá exclusividad sobre él, no comparte el código fuente con otros usuarios, prohibiendo las opciones de copiar, modificar y estudiar el código fuente.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Rodríguez (2008), menciona que el *software* no puede ser utilizado en ningún otro hardware, tampoco puede ser modificado mientras que no se paguen los derechos de autor, para poder utilizar el *software* primero hay que pagar una licencia y así depender totalmente de quien lo desarrolla todo esto lleva un aspecto legal, ya que se firman contratos de mantenimientos los cuales los brindaran por año, si se acaba el año de la licencia hay que renovarla. El *software* de propietario se encuentra bajo *copyright* el cual le brinda los derechos de autor al desarrollador del *software*.

2.5.4 Licencia de software libre AGPL

La GNU Affero General Public License (AGPL), es una licencia *copyleft* libre, que se utiliza específicamente en *software* de servidores web. El código fuente de AGPL está disponible para el usuario, con el cuál él puede modificarlo de la manera que le convenga, esto después de que el operador de un servidor web proporcione el código fuente de la versión modificada para su uso público (*Free Software Foundation, Inc.*, 2007).

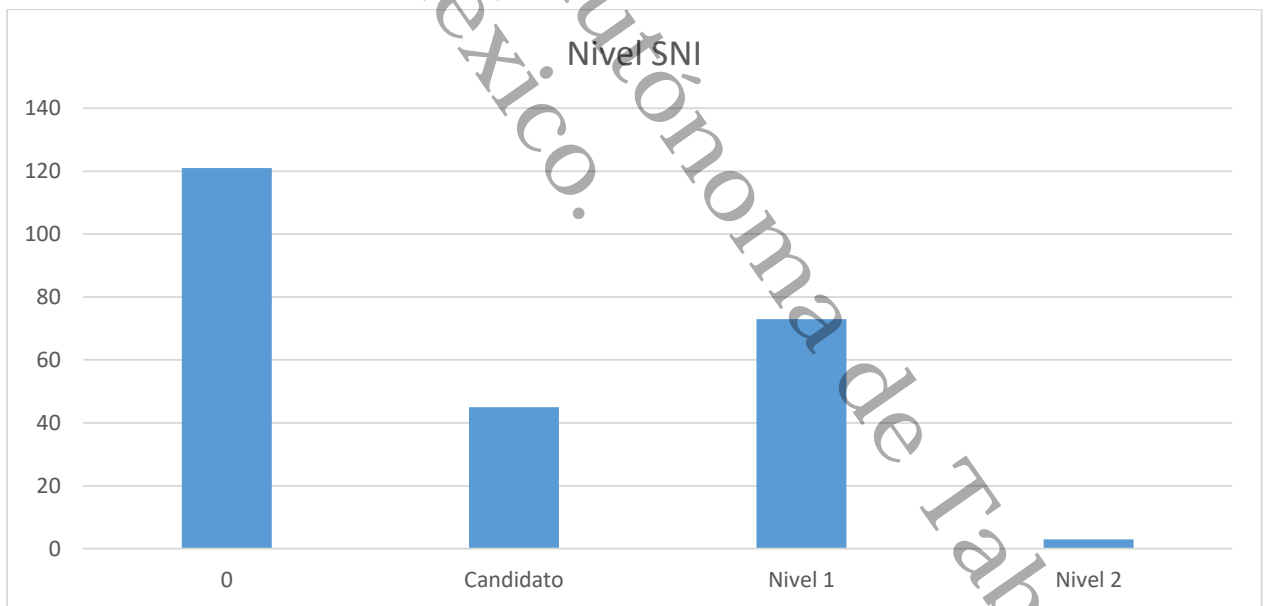
Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Capítulo 3. Aplicación de la metodología y desarrollo

3.1 Cuantificación de la población de estudio

Para determinar la población de estudio, se aplicaron 250 encuestas a profesores investigadores de una universidad pública en México, con base a las encuestas realizadas se determina que el número de encuestas que se respondieron satisfactoriamente son de 242 profesores, la cual resultó con 121 profesores que no cuentan con el SNI, mientras los otros grupos están divididos en 45 con nivel candidato, 73 con nivel 1 y por último tres con nivel 2, esto se observa en la Figura 2.

Figura 2
Población de estudio.



Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Del total de profesores encuestados el 39 % son del género femenino y 61 % son del género masculino, como se muestra en la Figura 3.

Figura 3

Profesores investigadores encuestados.



Nota: Elaboración propia

3.2 Depuración de la población

Para determinar la población de estudio se analizó el dataset para determinar el número de participantes, en esta investigación participaron 242 profesores, con edades entre 30 y 72 años pertenecientes a la Universidad Juárez Autónoma de Tabasco (UJAT), los cuales están distribuidos en sus diferentes divisiones como se muestra en la Tabla 1.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Tabla 1

Profesores investigadores participantes.

División Académica donde labora	Profesores
División Académica de Ciencias Agropecuarias	21
División Académica de Ciencias Sociales y Humanidades	21
División Académica de Ciencias Biológicas	43
División Académica de Ciencias Básicas	34
División Académica de Ingeniería y Arquitectura	27
División Académica de Ciencias Económico Administrativas	16
División Académica de Educación y Artes	20
División Académica de Ciencias y Tecnologías de la Información	14
División Académica de Ciencias de la Salud	16
Multidisciplinaria de Jalpa	11
Multidisciplinaria de Comalcalco	9
Multidisciplinaria de los Ríos	10

Nota: Elaboración propia

3.3 Diseño de instrumentos

Se elaboró un cuestionario para conocer la perspectiva de los profesores investigadores sobre su desempeño como investigador, y conocer las características de su perfil académico, que dio como resultado la construcción de un *dataset*, en el cual se almacenó cada una de las respuestas que los profesores contestaron. Para su elaboración se determinaron como criterio las variables de trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

investigación, la mayoría de las preguntas están basadas en la escala tipo Likert, para facilitarle al profesorado una respuesta más rápida.

Este tipo de escalas se utiliza para valorar las actitudes, lo que hace posible la creación de variables cuantitativas, las cuales se dividen en categóricas y numéricas, en las categóricas se encuentran las variables nominales, como el género y en las variables numéricas se encuentran las de intervalo, como cuánto es el salario que percibe un profesor (Obando, 2015).

La importancia de este instrumento permitió conocer el perfil y la postura en la que se encuentran los profesores con respecto a su desarrollo como investigadores, el dataset fue almacenado en un archivo de *Excel*, el cual permitió ser utilizado sin problemas por la herramienta de *RapidMiner*.

3.4 Aplicación del instrumento

El instrumento se aplicó de manera presencial mediante un cuestionario dirigido exclusivamente a profesores investigadores, posteriormente todas las respuestas fueron almacenadas en un archivo de *Excel*, dando como resultado el *dataset* el cual fue utilizado en esta investigación para la implementación de la metodología KDD.

3.5 Almacenamiento para el análisis de la información

Para el almacenamiento de la información se creó un *dataset*, obtenido de las encuestas aplicadas a los profesores investigadores en una universidad pública en México, dicho *dataset* está dividido en seis secciones con diferentes datos, tales como:

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

sociodemográficos, carrera profesional y formación académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la investigación, dicho dataset contiene 76 preguntas y 242 profesores encuestados.

3.6 Análisis para la selección de herramienta de minería de datos MOA, YALE y WEKA

Actualmente existen distintas herramientas que son utilizadas para la investigación con minería de datos estas herramientas pueden ser gratuitas o de paga, es por eso que para realizar la identificación de los patrones de conducta que afectan a los profesores en su desarrollo como investigadores es necesario un análisis de las herramientas de minería de datos, con los cuales se pretende obtener los resultados que ayudaron a identificar dichos patrones. Dentro de la investigación se habla de tres herramientas MOA, RapidMiner antes conocida como YALE y WEKA.

Es por ello que para elegir la herramienta que se utilizó en esta investigación se analizó que tuvieran los siguientes criterios:

- Que pueda aplicar el proceso KDD.
- Capaz de leer base de datos en diferentes formatos.
- Interfaz amigable y fácil de usar.
- Herramienta libre.
- Herramientas de visualización de resultados.
- Gráficos comprensibles para el usuario.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

- Multiplataforma.

3.7.1 RapidMiner

Antes conocida como *Yet Another Learning Environment* (YALE) es una de las herramientas de minería de datos más conocidas y más usadas, lo más destacado de esta herramienta es que es de acceso gratuito y de un fácil manejo, esta herramienta posee una gran selección de operadores que permiten la conexión de los datos.

Mariño (2014) menciona que es un software de minería de datos, creado por la compañía *RapidMiner*, y fue actualizada en el 2001 como una iniciativa de seguir con el proyecto de YALE. También menciona que esta herramienta está disponible para los sistemas operativos *Windows*, *Mac OS* y *Linux*.

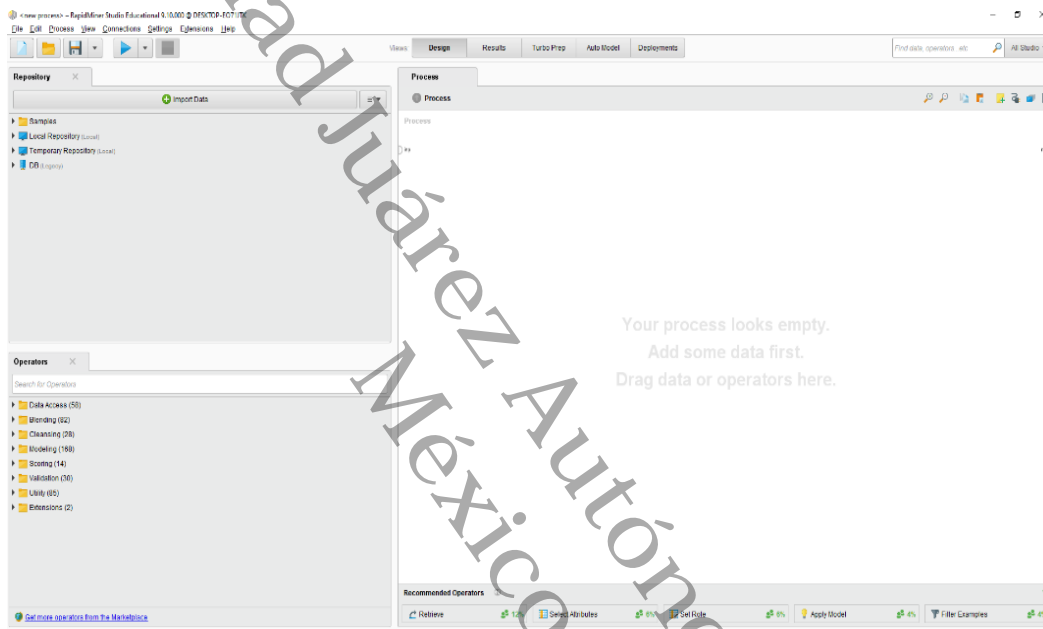
Ruiz y Romero (2018) la definen como una herramienta libre, muy superior a WEKA, ya que el manejo de los resultados y las gráficas son de mayor calidad y con la característica de poder adjuntar diferentes extensiones, que aumentan el potencial del software.

Ionos inc. (2018) menciona que es un software muy inclusivo con la minería de datos principalmente en la visualización de la información obtenida. La herramienta está formada por tres grandes módulos: *RapidMiner Studio*, *RapidMiner Server* y *RapidMiner Radoop*, cada uno encargado de una técnica diferente de minería de datos. Con *RapidMiner* es posible preparar los datos antes de realizar todo el proceso de análisis. El

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

punto más fuerte de esta herramienta es el análisis predictivo. En la Figura 4 se puede observar la interfaz de *RapidMiner*.

Figura 4
Interfaz de *RapidMiner*.



Nota: Elaboración propia

3.7.2 Waikato Environment for Knowledge Analysis (WEKA)

Es un software de código abierto creado en la universidad de Waikato a mediados de los años 90, está escrito en lenguaje JAVA lo que lo hace compatible con sistemas operativos Windows, macOS y Linux. WEKA tiene un gran número de funciones de aprendizaje automático, al igual gestiona tareas de minería de datos como clúster, correlación, regresión y también la clasificación de los datos, este último como uno de

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

sus puntos más fuertes ya que utiliza redes neuronales y árboles de decisión (Ionos inc., 2018).

Una de las desventajas de WEKA es que no es tan fuerte en el análisis de clúster, otra desventaja es el problema que presenta WEKA al momento de realizar el procesamiento cuando se trata de grandes volúmenes de datos, esto se debe a que intenta cargar la minería de datos desde el cache del software.

Chamba (2015) define a WEKA como un software que contiene un grupo de librerías de JAVA y está regida bajo una licencia GPL. También menciona que WEKA fue creado para realizar la extracción de conocimiento de base de datos con grandes volúmenes de datos.

Por otra parte, Brownlee (2020) menciona que WEKA, propone un aprendizaje automático relativamente fácil y eficiente, la define como una herramienta gráfica y de fácil uso, con la que es posible cargar grandes volúmenes de datos, ejecutar algoritmos y la creación de resultados estadísticamente completos.

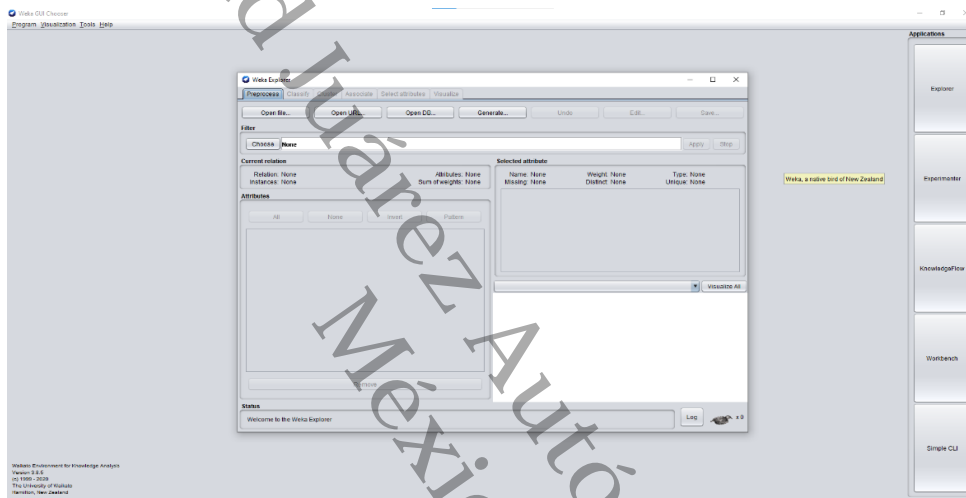
WEKA trabaja con un tipo de archivo llamado Attribute Relation File Format (ARFF), el cual se divide en tres partes, cabecera, declaración de atributos y selección de datos, también puede leer archivos de *Excel* y CSV (Navas, 2016). Estos tipos de archivos son los más comunes entre las herramientas y le facilita al usuario la aplicación de la minería de datos ya que *Excel* es de las herramientas más conocidas y al usuario le resulta más familiar trabajar con ellos. En la Figura 5 se observa la interfaz con la que

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

el usuario de WEKA puede interactuar, siendo una herramienta muy gráfica e intuitiva, aunque con poco diseño.

Figura 5

Interfaz gráfica de WEKA.



Nota: Elaboración propia

3.7.3 Massive Online Analysis (MOA)

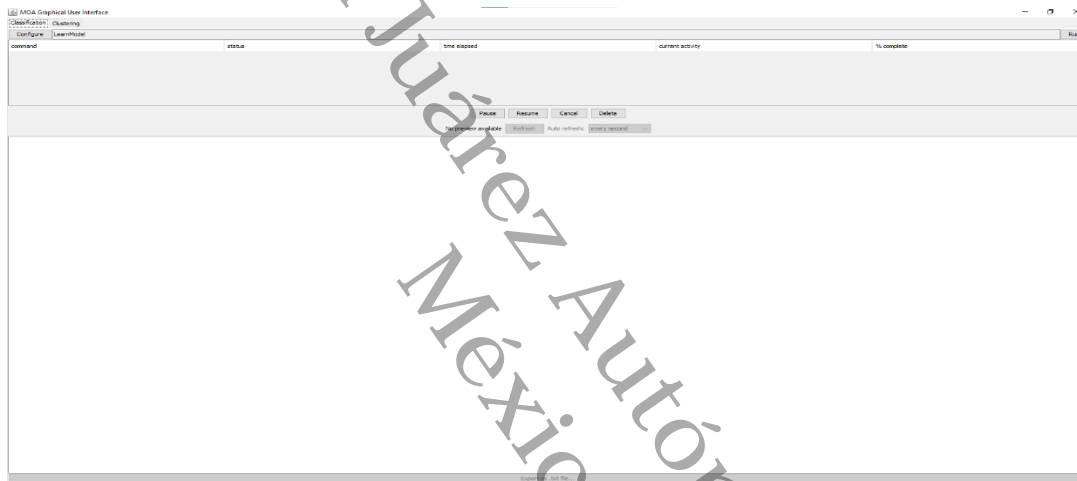
Al igual que WEKA y RapidMiner, MOA es una herramienta de minería de datos de uso libre que permite crear y ejecutar el aprendizaje automático. MOA tiene como objetivo la implementación de algoritmos más avanzados para convertir a tamaño real un conjunto de datos.

Contiene una serie de algoritmos online y offline para la clasificación, agrupamiento y minería de gráficos, así como herramientas para la evaluación de la información. MOA admite la interacción con WEKA. Además de proporcionar algoritmos

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

y medidas para evaluación y comparación, MOA es fácilmente ampliable con nuevas contribuciones. En la Figura 6 se puede observar la interfaz de MOA, la cual es un poco más simple que WEKA y RapidMiner.

Figura 6
Interfaz gráfica de MOA.



Nota: Elaboración propia

Bifet *et al.* (2010) menciona que MOA fue desarrollado en lenguaje JAVA, también menciona que la principal característica de la herramienta es su portabilidad, ya que se ejecuta en cualquier máquina virtual Java. El uso del lenguaje está muy extendido y características como la recolección automática de basura ayudan para reducir la carga y los errores del programador.

Como puede observarse en la Tabla 2, las tres herramientas son muy parecidas y cada una de ellas tiene su punto fuerte, que pueden ser utilizados para el análisis de los datos en esta investigación.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Tabla 2

Comparación de herramientas de minería de datos.

Herramienta	Características	Lenguaje	SO	Precio
RapidMiner (YALE)	Útil para todos los procesos de minería de datos, fácil uso y de interfaz amigable, su punto más fuerte es el análisis predictivo.	Java	Windows, MacOS, Linux	Software sin costo, con diferentes versiones de pago.
WEKA	Es una de las herramientas con más métodos de clasificación de la información, lo que la hace su punto más fuerte.	Java	Windows, MacOS, Linux	Software libre (GPL)
MOA	Es una herramienta poco conocida de software libre, muy parecida a WEKA, pero con funciones más limitadas, el punto fuerte de MOA, es la utilización de clúster, que permite una mejor agrupación de los resultados.	Java	Windows, MacOS, Linux	Software libre (GPL)

Nota: Elaboración propia.

Con el análisis de las herramientas y la comprensión de sus ventajas y desventajas como se observa en la Tabla 3, se puede decir que la herramienta de *RapidMiner* es quien cumple con los criterios de selección ya que tiene la interfaz más amigable lo que facilita su uso, y cuenta con las características suficientes para el desarrollo de esta investigación ya que también permite aplicar el proceso KDD y el uso de minería de datos,

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

así como la aplicación de clúster y diferentes algoritmos que facilitarán la identificación de los patrones de conducta de los profesores investigadores.

Tabla 3

Ventajas y desventajas en el uso de las herramientas de minería de datos.

Herramienta	Ventaja	Desventaja
MOA	No necesita instalación ya que es un ejecutable de java, Multiplataforma,	Solo reconoce archivos ARFF, No tiene una interfaz amigable, difícil de comprender.
RapidMiner	Fácil uso, Multiplataforma, Incluye gráficos y herramientas de visualización de resultados fácil de entender, lee varios tipos de archivos. Automáticamente elige el mejor algoritmo para los datos que se analizan	Ocupa mucho espacio de memoria, no es totalmente gratuito
WEKA	Fácil uso, interfaz amigable, multiplataforma, lee varios tipos de archivos	No incluye tantos gráficos, interpretación complicada, la visualización de resultados no es tan buena.

Nota: Elaboración propia

3.7 Detección de las necesidades

Un proceso KDD consiste en la extracción no trivial de conocimiento previamente desconocido y potencialmente útil a partir de un conjunto de datos, por lo tanto, en principio se establece la detección de las necesidades en función de tres elementos: muestra, instrumentos y proceso de aplicación.

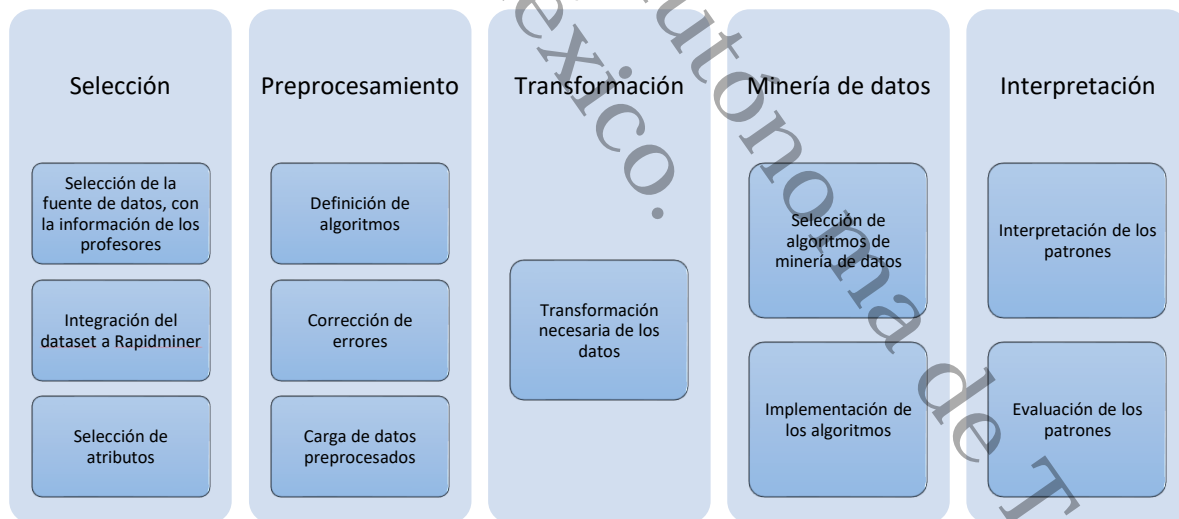
Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

3.8 Metodología KDD

Para el desarrollo de la metodología se utilizó el proceso KDD el cual cuenta con cinco etapas, selección, preprocesamiento o limpieza, transformación, minería de datos e interpretación, cada una de estas etapas fue fundamental para crear un modelo que permita el análisis de los datos, y obtener conocimiento aplicando la minería de datos a través de RapidMiner, en la Figura 7 se observa la implementación de este proceso.

Para el análisis de los datos se utilizaron los métodos de Clustering con el algoritmo *K-means*, reglas de asociación con el algoritmo Apriori y árbol de decisión.

Figura 7
Implementación del proceso KDD.



Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

3.8.1 Selección

Una vez identificado el conocimiento relevante y prioritario y definidas las metas del proceso KDD, se determina cuáles son los datos de mayor relevancia para así poder realizar el análisis de la información con la que se cuenta, como la trayectoria académica, problemas familiares o personales, afrontamiento. Para ello se deben eliminar el mayor número posible de datos erróneos o inconsistentes, utilizando métodos estadísticos. Las tareas de preparación o de limpieza de datos fueron realizadas repetidas veces.

En esta etapa se seleccionan los datos que se van analizar en este caso se selecciona el *dataset* donde se encuentran las respuestas de los 242 profesores encuestados, dentro del dataset se encuentran diferentes conjuntos de datos cada uno está dividida en diferentes variables tales como: variables socio demográficas (ver Tabla 4), carrera profesional y formación académica (ver Tabla 5), condiciones institucionales (ver Tabla 6), vida personal y familiar (ver Tabla 7), afrontamiento (ver Tabla 8) y habilidades para la investigación (ver Tabla 9), así como se observa en cada una de las tablas donde se muestra el número y el nombre de cada una de las variables utilizadas en esta investigación para la minería de datos.

Tabla 4

Variables sociodemográficas.

No.	Selección de variables
1	Edad
2	Sexo
3	Estado civil

Nota: elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Tabla 5

Variables de carrera profesional y formación académica.

No.	Selección de variables
1	División académica donde trabaja
2	Tipo de nombramiento
3	¿Cuenta con plaza definitiva?
4	¿Labora en otra institución?
5	¿Tiene puesto administrativo en la UJAT?
6	¿Cuántas horas de docencia imparte en licenciatura?
7	¿Cuántas horas de docencia imparte en posgrado?
8	Nombre del doctorado que estudió
9	Área del SNI
10	El Doctorado lo cursó a nivel Nacional o Internacional
12	¿Es miembro del SISTEMA NACIONAL DE INVESTIGADORES?
13	Qué nivel SNI

Nota: elaboración propia

Tabla 6

Variables de condiciones institucionales.

No.	Selección de variables
1	¿La institución promueve la investigación?
2	¿Se incentivan las actividades de vinculación con otras instituciones?
3	¿Se reconoce el trabajo que realizan los investigadores?
4	¿Se facilita el establecimiento de redes con otros investigadores?
5	¿Se apoya con recursos a los investigadores?
6	¿Existen reglamentos para regular la carga docente de los investigadores?
7	¿Las áreas administrativas apoyan efectivamente el trabajo de los investigadores?

Nota: elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Tabla 7

Variables de vida personal y familiar.

No.	Selección de variables
1	¿Ha tenido problemas familiares por su trabajo?
2	Su familia comprende las demandas de su trabajo
3	Comparte actividades de recreación con su familia
4	Asiste a las actividades académicas de sus hijos (si los tiene)
5	Asiste a eventos sociales no relacionados con la universidad
6	Apoya en las labores del hogar
7	Sale de vacaciones con su familia
8	Convive con compañeros de trabajo fuera del horario laboral
9	Convive con sus amigos en actividades recreativas
10	Se siente satisfecho con su vida personal
11	Se siente satisfecho con su vida familiar

Nota: elaboración propia

Tabla 8

Variables de afrontamiento.

No.	Selección de variables
1	Mantiene una salud buena
2	Puede manejar el estrés en su trabajo
3	Puede manejar las presiones de tiempo de su trabajo
4	Maneja efectivamente las demandas de nuevas habilidades que implica su trabajo
5	Se ejercita de manera constante

Nota: elaboración propia

Tabla 9

Variables de habilidades para la investigación.

No.	Selección de variables
1	Planificar el tiempo
2	Administrar el tiempo
3	Comunicarse de manera escrita
4	Comprensión de textos en un segundo idioma
5	Comunicarse de manera oral en un segundo idioma
6	Comunicarse de manera escrita en un segundo idioma
7	Uso de las Tecnologías de la Información y Comunicación

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

- 8 Crítica y autocrítica
 - 9 Actuar de manera creativa
 - 10 Actuar en nuevas situaciones
 - 11 Tomar decisiones
 - 12 Trabajar en contextos multidisciplinares
 - 13 Trabajar en contextos internacionales
 - 14 Compromiso ético y social
 - 15 Buscar información en bases de datos especializadas
 - 16 Desarrollar marcos teóricos de referencia
 - 17 Utilizar un sistema de referencias para dar crédito a las fuentes bibliográficas consultadas
 - 18 Elaborar fichas documentales y de trabajo
 - 19 Conocimientos acerca de los paradigmas de investigación
 - 20 Identificar problemas y/o necesidades de investigación o desarrollo tecnológico
 - 21 Formular problemas de investigación
 - 22 Redactar preguntas y objetivos de investigación
 - 23 Elaborar hipótesis de investigación
 - 24 Utilizar diseños experimentales
 - 25 Utilizar diseños no experimentales
 - 26 Diseñar instrumentos adecuados para la recolección de datos
 - 27 Validar instrumentos adecuados para la recolección de datos
 - 28 Analizar datos (cuantitativos)
 - 29 Analizar datos (cualitativos)
 - 30 Utilizar software para el análisis de datos
 - 31 Desarrollar prototipos de productos
 - 32 Elaborar informes técnicos
 - 33 Divulgar resultados en medios científicos (revistas, congresos, consejos técnicos entre otros)
 - 34 Divulgar resultados al público en general
 - 35 Conocimientos de las formas de patentar
 - 36 Conocimientos de los fondos públicos de apoyo a la investigación y desarrollo tecnológico
 - 37 Conocimiento de los fondos privados de apoyo a la investigación y desarrollo tecnológico
 - 38 Elaborar proyectos donde se gestionen fondos públicos
 - 39 Elaborar proyectos donde se gestionen fondos privados
 - 40 Desarrollar prototipos de procesos
 - 41 Implementar a nivel comercial de prototipos de productos y o procesos
 - 42 Conocimientos de las normas de propiedad intelectual
-

Nota: elaboración propia

El *dataset* está construido en *Excel* con formato (xlsx), el cual no tuvo problemas para ser integrado en la aplicación de *RapidMiner* ya que la herramienta permite que sea muy visual y de fácil uso. En la Figura 8 se observa, la selección de la base de datos con

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

la herramienta, esta nos presenta los datos, en una tabla e identifica el tipo de datos de cada columna, como *polynomial*, *integer*, reales y fecha.

Figura 8

Selección de los datos.

Import Data - Format your columns.

Format your columns.

Replace errors with missing values ⓘ

	Sociodemo...	B	C	D	E	F	Carrera pro...	H
	<i>polynomial</i>	<i>polynomial</i>	<i>integer</i>	<i>integer</i>	<i>polynomial</i>	<i>integer</i>	<i>polynomial</i>	<i>polynomial</i>
2	51	Femenino	1966	33	Soltero	0	DAEA	PTC A
3	52	Femenino	1965	50	Soltero	0	DAEA	PTC
4	40	Femenino	1977	38	Casado	1	DAEA	PTC
5	40	Femenino	1977	38	Casado	1	DAEA	PTC
6	54	Masculino	1963	48	Casado	3	DACEA	PTC
7	52	Femenino	1965	45	Casado	2	DAEA	PTC A
8	49	Femenino	1968	40	Casado	2	DAEA	PTC A
9	62	Femenino	1955	30	Casado	3	DAEA	PTC C
10	54	Femenino	1963	45	Soltero	1	DAEA	PTC
11	41	Masculino	1976	32	Soltero	0	DAEA	PTC A
12	41	Masculino	1976	39	Casado	2	DAIS	PTC
13	42	Masculino	1975	40	Casado	2	DAIS	PTC
14	63	Masculino	1954	57	Soltero	0	DAIS	PTC
15	42	Masculino	1975	41	Casado	0	DAIS	PTC
16	48	Masculino	1969	40	Casado	3	DAIS	PTC
17	47	Masculino	1970	45	Soltero	0	DAIS	PTC
18	52	Femenino	1965	50	Soltero	2	DAIS	PTC

Nota: Elaboración propia a través de la herramienta *RapidMiner*

3.8.2 Preprocesamiento/limpieza

En esta etapa se depura la información irrelevante utilizando técnicas como la eliminación de datos ruidosos, Timarán et al. (2016) menciona que, en esta etapa, se utilizan estrategias para el uso de datos desconocidos y se utilizan herramientas estadísticas para reemplazar datos nulos, desconocidos y duplicados.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

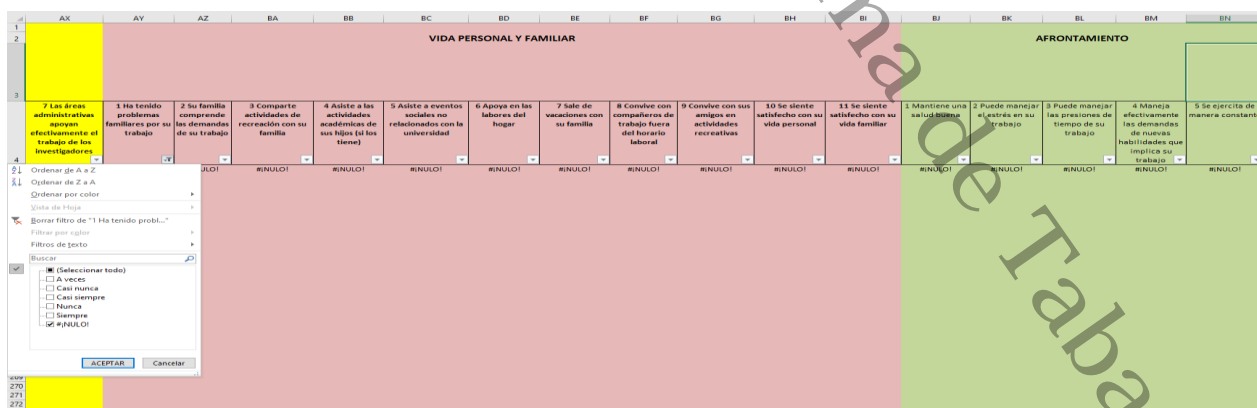
Los datos desconocidos también conocidos como *empty* son aquellos a los cuales no les corresponde un valor en el mundo real y los *missing* son aquellos que tienen un valor que no fue capturado.

Los datos nulos son datos desconocidos que son permitidos por los sistemas gestores de bases de datos relacionales. En el proceso de limpieza todos estos valores se ignoran, se reemplazan por un valor por omisión, o por el valor más cercano, es decir, se usan métricas de tipo estadístico como media, moda, mínimo y máximo para reemplazarlos.

Para la reducción de datos se utilizó un filtrado de datos en *Excel*, con el cual se pudo identificar los datos *missing* y *empty* algunos de los datos más comunes encontrados dentro del *dataset* fueron en la variable de afrontamiento y vida personal y familiar, en donde por lo menos hubo un valor nulo en cada uno como se observa en la Figura 9.

Figura 9

Filtrado de datos para encontrar valores nulos.



Nota: elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Cabe recalcar que la mayoría de los datos que se encuentran dentro de *dataset* fueron utilizados para la minería de datos, excluyendo solo algunos como los datos redundantes. Un ejemplo de estos es la edad y el año de nacimiento, ya que con ambos es posible obtener la misma información.

De esta manera se realizó el preprocesamiento de los datos en la herramienta de RapidMiner, para corroborar que los datos fueran leídos y visualizados de manera correcta.

En esta etapa RapidMiner puede rellenar los valores nulos, o donde faltaron respuestas con los valores de la media, utilizando el operador *replace missing values* el cual permite rellenar esos espacios vacíos con el valor promedio. En la Figura 10 se observan celdas con el signo de interrogación, los cuales son casillas que tienen un valor nulo y en la Figura 11 la construcción para utilizar el operador.

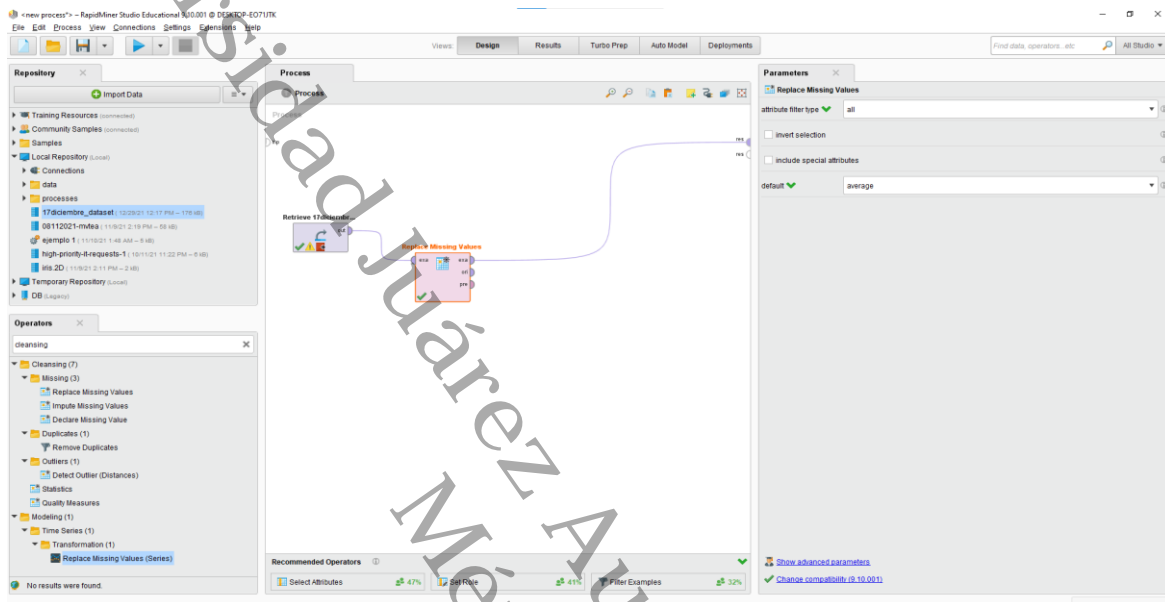
Figura 10
Datos sin transformar.

N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
Tiene puesto ...	Cuántos año...	Año de inicio ...	Año de fin del...	Cuántas hora...	Cuántas hora...	Nombre del d...	En qué año o...	Area del SNI	El Doctorado ...	Paisgrado	Cuántos año...	El grado de d...	Tuvo beca du...	Tuvo Beca C...	Tuvo beca P...	Tuvo otra bec...	D
Si	4	2000	2004	40	0	Doctorado	1999	IV Humanida...	Internacional	Francia	?	No	Si	No	No	No	Si
Si	4	2007	2010	22	0	Doctorado	2015	IV Humanida...	Nacional	México	?	No	No	No	No	No	?
No	0	0	0	20	0	Doctorado	2015	IV Humanida...	Nacional	México	?	Si	Si	Si	No	No	?
No	0	0	0	10	0	Ciencias de I...	2015	IV Humanida...	Nacional	México	3	Si	Si	Si	No	No	?
No	0	0	0	20	3	Finanzas Púb...	2011	V Ciencias S...	Nacional	México	4	Si	Si	No	Si	No	?
No	0	0	0	17	0	Doctorado	2010	IV Humanida...	Internacional	España	?	No	Si	No	No	Si	U
No	0	0	0	5	3	Doctorado	2008	IV Humanida...	Nacional	México	3	Si	Si	No	Si	No	?
Si	5	1990	1995	10	0	Doctorado en...	1985	IV Humanida...	Internacional	Estados Unid...	?	No	Si	Si	No	No	?
Si	2	2003	2005	5	5	Doctorado	2008	IV Humanida...	Internacional	Cuba	3	No	Si	No	No	Si	U
Si	3	1999	2001	4	6	Doctorado en...	2008	IV Humanida...	Internacional	Canadá	?	Si	Si	No	No	Si	U
No	0	0	0	20	0	Ingeniería en ...	2015	VII Ingenierías	Nacional	México	2	No	No	No	No	No	?
Si	2	0	0	18	0	Administración	2015	V Ciencias S...	Nacional	México	2	No	No	No	No	No	?
Si	5	0	0	16	0	Educación	2011	IV Humanida...	Nacional	México	2	No	No	No	No	Si	U
No	0	0	0	11	0	Ciencias de I...	2016	VII Ingenierías	Nacional	México	3	Si	Si	Si	No	No	?
Si	2	0	0	19	0	Administració...	2009	V Ciencias S...	Internacional	España	4	Si	Si	No	Si	No	?
No	0	0	0	20	0	Sistemas Co...	2015	VII Ingenierías	Nacional	México	2	No	No	No	No	No	?
Si	4	0	0	18	0	Educación	2015	IV Humanida...	Nacional	México	2	No	No	No	No	No	?
No	0	0	0	21	8	Dirección y A...	2012	V Ciencias S...	Internacional	España	5	Si	Si	No	Si	No	?
No	0	0	0	0	13	Tecnologías	2008	VII Ingenierías	Internacional	España	4	Si	Si	No	Si	No	?
No	0	0	0	10	10	Ciencias en I...	2012	VII Ingenierías	Nacional	México	4	Si	Si	No	Si	No	?
Si	3	2002	2005	0	10	Ciencias de I...	2010	I Física Mate...	Internacional	Estados Unid...	5	Si	Si	No	Si	No	Fu
No	0	0	0	9	0	Doctorado en ...	2002	I Física Mate...	Internacional	Francia	?	Si	Si	Si	No	No	?
No	0	0	0	5	8	Ciencias	2009	III Medicina y ...	Nacional	México	5	Si	Si	Si	No	No	?
No	0	0	0	10	0	Ciencias y m...	2010	VI Biotecnolo...	Nacional	México	6	Si	Si	Si	No	No	?

Nota: Elaboración propia a través de *RapidMiner*

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Figura 11
Uso del algoritmo Replace Missing Values



Nota: Elaboración propia a través de la herramienta de RapidMiner

3.8.3 Transformación de los datos

La etapa final de la preparación de los datos, que representa la vista final de ellos, lista para ser analizada por la minería de datos, esto con tal de facilitar la extracción de patrones dentro de los datos.

Timarán (2016) menciona que en esta etapa se pretende hallar las características más importantes y útiles que ayuden en la representación de los datos, en esta etapa se utilizan diferentes tipos de métodos de reducción y transformación de los datos, también menciona que se eliminan las tuplas que sean idénticas esto se hace una reducción horizontal, mientras que verticalmente se eliminan los atributos con menos importancia,

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

en el caso de esta investigación algunas de las columnas que fueron eliminadas fue la de fecha de nacimiento, ya que ya se contaba con la edad de cada uno de los profesores investigadores, en la Figura 12 se observan los datos transformados listos para su análisis.

Figura 12
Resultado del uso del algoritmo *Replace Missing Values*.

id	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AM	AI
1	5	Diccionario en.	2009	II Biotecnolog.	Nacional	México	8	Si	Si	No	Si	No	Institucional	3	No	0	0
1	10	Diccionario en.	2014	I Física Mate.	Nacional	México	8	Si	Si	No	Si	No	Institucional	3	Si	2017	Candidato
1	8	Diccionario en.	1999	II Humanida.	Internacional	Francia	7	No	No	No	No	No	Institucional	5	No	0	0
1	0	Diccionario en.	2014	V Ciencias S.	Internacional	España	7	No	Si	No	No	Si	Institucional	4	Si	2016	Candidato
1	0	Ciencias y m.	2010	VI Biotecnol.	Nacional	México	5	Si	Si	Si	No	No	Institucional	2	Si	2011	Candidato
1	12	Diccionario en.	2008	VIII Ingenierías	Nacional	México	6	Si	Si	No	No	No	Institucional	5	Si	2009	Nivel 1
1	0	Diccionario en.	2012	VIII Ingenierías	Nacional	México	6	Si	Si	No	No	No	Institucional	2	No	0	0
1	0	Diccionario en.	2012	VIII Ingenierías	Nacional	México	6	Si	Si	No	No	No	Institucional	2	No	0	0
1	0	Diccionario en.	2010	V Ciencias S.	Internacional	España	6	Si	No	No	No	No	Institucional	0	No	0	0
1	12	Diccionario en.	2009	I Física Mate.	Nacional	México	6	Si	No	No	No	No	Institucional	1	Si	2011	Nivel 1
1	4	Diccionario en.	2015	V Ciencias S.	Nacional	México	6	Si	Si	No	No	No	Institucional	2	Si	2017	Candidato
1	6	Diccionario en.	2011	I Física Mate.	Nacional	México	6	Si	Si	No	No	No	Institucional	1	Si	2017	Candidato
1	0	Diccionario en.	2016	II Biotecnolog.	Nacional	México	6	Si	Si	No	No	No	Institucional	1	No	0	0
1	0	Diccionario en.	2002	I Física Mate.	Nacional	México	6	Si	Si	No	No	No	Institucional	2	No	0	0
1	0	Diccionario en.	2013	II Medicina y	Nacional	México	6	Si	Si	No	No	No	COMECYT	2	Si	2015	Candidato
1	0	Diccionario en.	2009	II Medicina y	Nacional	México	6	Si	Si	No	No	No	Institucional	5	Si	2009	Nivel 1
1	5	Diccionario en.	1993	II Humanida.	Internacional	Estados Unid.	6	Si	Si	No	No	No	FORD	1	No	0	0
1	0	Diccionario en.	2012	VIII Ingenierías	Nacional	México	6	Si	Si	No	No	No	Institucional	3	No	0	0
1	0	Diccionario en.	1995	II Biotecnolog.	Nacional	México	6	No	No	No	No	No	Institucional	4	No	0	0
1	8	Dirección A.	2012	V Ciencias S.	Internacional	España	5	Si	No	Si	No	No	Institucional	3	No	0	0
1	10	Ciencias de L.	2010	I Física Mate.	Internacional	Estados Unid.	5	Si	No	Si	No	No	Futurget	3	Si	2016	Nivel 1
1	8	Ciencias	2009	II Medicina y	Nacional	México	5	Si	Si	No	No	No	Institucional	2	No	2011	0
1	0	Ciencias Bio.	2012	II Medicina y	Nacional	México	5	Si	Si	No	No	No	Institucional	4	Si	2015	Candidato
1	3	Diccionario en.	2011	VIII Ingenierías	Nacional	México	5	Si	Si	Si	No	No	Institucional	6	Si	2012	Nivel 1

Nota: Elaboración propia a través de la herramienta *RapidMiner*

3.8.4 Minería de datos

Para el diseño se emplea la técnica de agrupación para identificar las similitudes, aplicar reglas de asociación y establecer los patrones de conducta obtenidos. Inflow (2020) menciona que las técnicas de minería de datos crean modelos que son predictivos o descriptivos. Los modelos predictivos pretenden estimar valores futuros o desconocidos.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

En esta fase se buscan los siguientes cuatro tipos de relaciones (clases, clústeres, asociaciones y patrones secuenciales). Se aplicará la técnica del árbol de decisión para una representación analítica de los patrones de trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades de los profesores investigadores.

35 Todo el proceso de modelado se lleva a cabo sobre un conjunto de ejemplos formado tan sólo por entradas al sistema. El modelo KDD, como proceso, involucra una secuencia de etapas claramente definidas, cada una fundamental para la transición de los datos en conocimiento. Para Vieira et al. (2009), todos los sistemas de KDD mantienen la misma esencia, la minería de datos; su factor diferenciador radica en la implementación y la presentación. Todos transitan las mismas etapas: recolección, depuración y análisis de datos, de donde se obtiene como resultado un modelo descriptivo que puede ser convertido en un modelo predictivo.

15 En esta fase se realiza la explotación y uso de los resultados del proceso lo que, dependiendo de los requerimientos, puede ser tan sencillo como la generación de un informe o tan complejo como la realización repetida de un proceso cruzado de minería de datos. En la Tabla 10, se observa la selección de los algoritmos que se utilizaron para el desarrollo de la investigación.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Tabla 10

Selección de algoritmos de minería de datos.

Objetivo	Tarea	Algoritmo	Descripción
Identificar características comunes entre el ámbito personal y familiar y afrontamiento. Identificar entre las variables trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades investigativas, cual es la que más se relaciona con problemas ligados a las barreras que afectan al profesor.	Agrupar profesores	Clustering	Tiene como objetivo agrupar los datos que son similares dentro de un mismo grupo. Técnica utilizada para la toma de decisiones ya que se basa en el uso de resultados y probabilidades.
Identificar las relaciones entre trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la investigación de los profesores investigadores.	Toma de decisión	Algoritmo de árbol de decisión	
	Identificar patrones	Reglas de asociación	Permite encontrar asociaciones entre los datos en forma de relaciones

Nota: Elaboración propia

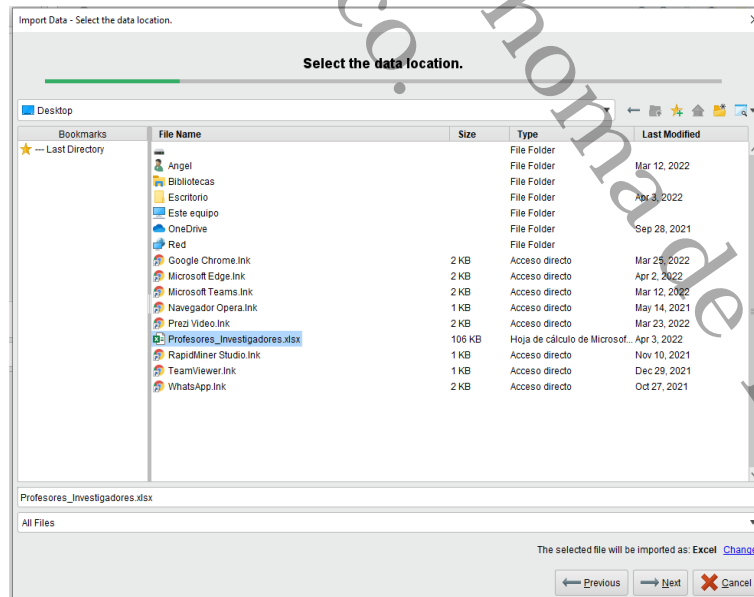
Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

RapidMiner

El primer paso de la minería de datos es la selección del *dataset* para ello se tiene que exportar el archivo como se muestra en la Figura 13, esto se hace dando *click* en el botón de *Import Data*, el cual abrirá un cuadro, que permite buscar la ubicación del *dataset* dentro del equipo de cómputo, se selecciona el *dataset*, el cual lleva el nombre de *Profesores_Investigadores.xlsx*, después de que se selecciona el *dataset*, se tiene que almacenar dentro del *software* de *RapidMiner* en el repositorio local, para que ya no sea necesario hacer el proceso de importación del *dataset* cada vez que se entre al *software* como se observa en la Figura 14.

Figura 13

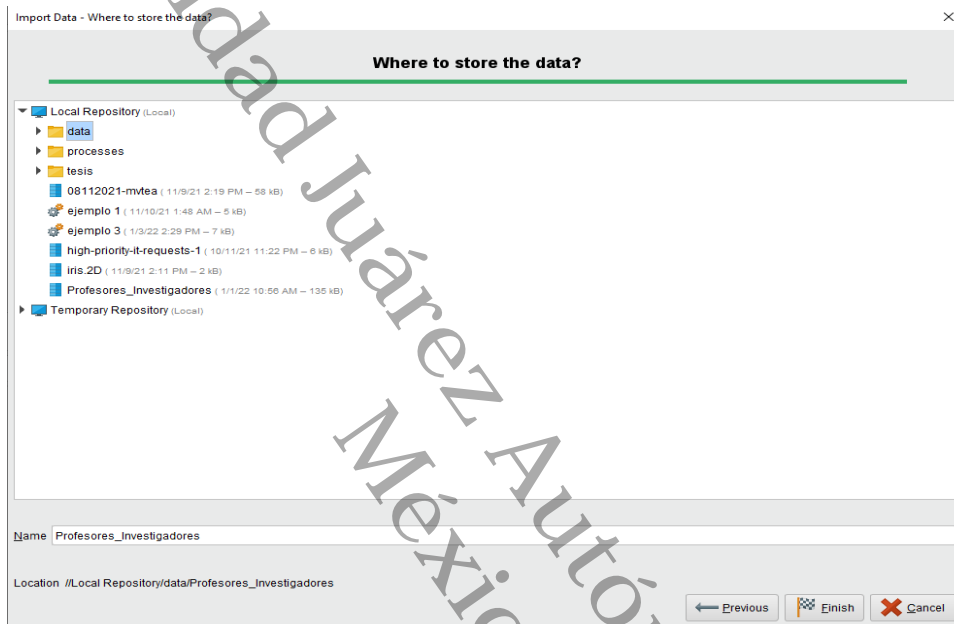
Importación del dataset



Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Nota: Elaboración propia

Figura 14
Almacenamiento del dataset.

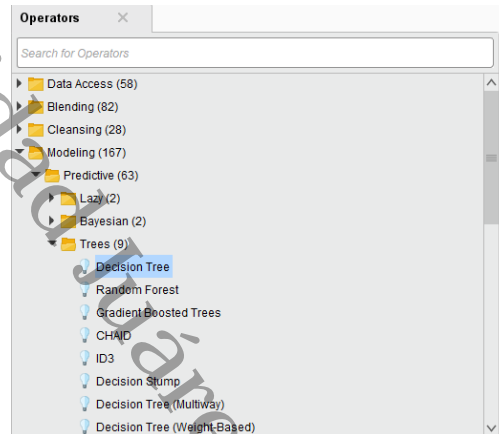


Nota: Elaboración propia

Ya que el dataset ha sido almacenado dentro de *RapidMiner*, es posible empezar a trabajar con la minería de datos, el primer algoritmo que se implementó fue el de árbol de decisión, para ello se utilizan los operadores con los que trabaja *RapidMiner*, los operadores utilizados para el desarrollo de este algoritmo fueron *Replace Missing Values*, el cual permite reemplazar valores nulos o vacíos por valores promedio, *Set Role*, este operador se utiliza para cambiar el rol de uno o más atributos y por último el operador de *Decisión Tree*, este último es operador que construye el árbol de decisión, estos operadores se encuentran en la pestaña de *Operators* de *RapidMiner* (ver Figura 15).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

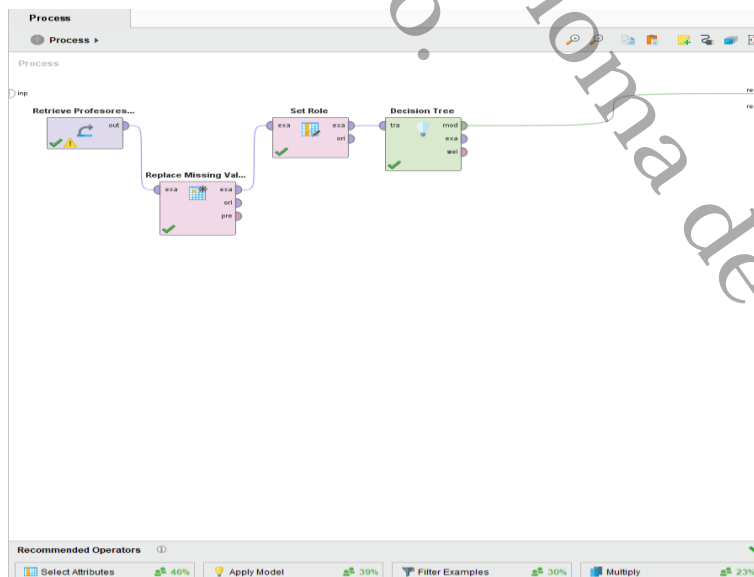
Figura 15
Lista de operadores.



Nota: elaboración propia

Una vez que se eligieron los operadores en *RapidMiner* para crear el árbol de decisión hay que arrastrar al área de procesos cada uno de ellos y unirlos (ver Figura 16).

Figura 16
Proceso de creación de árbol de decisión.



Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Para crear el árbol de decisión se tomaron aspectos como si el profesor puede manejar el estrés en su ambiente laboral como el *target* principal ya que esta pregunta ayudó a identificar si el profesor investigador puede manejar el estrés durante sus labores como investigador y sus actividades académicas.

El árbol de decisión nos da una visión gráfica de los diferentes problemas que enfrenta el profesor, ya que permite decidir, si un profesor enfrenta algún tipo de problema al realizar investigaciones (ver Figura 17).

Universidad Juárez Autónoma de Tabasco.
68

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos.

Figura 17
Árbol de decisión.



Nota: Elaboración propia

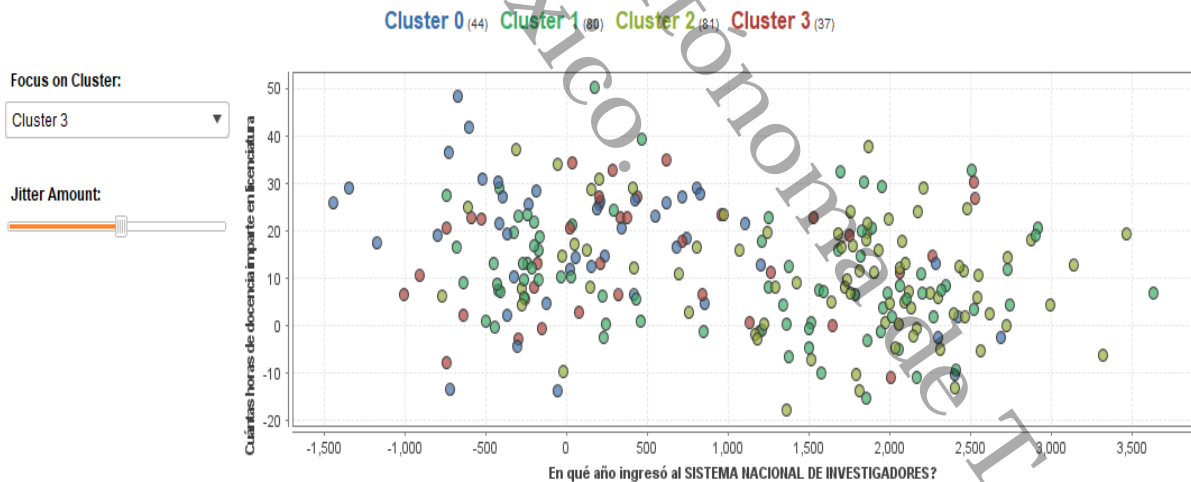
Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Al concluir la minería de datos con el algoritmo de árbol de decisión se realizó el método de agrupación con el algoritmo *K-means*, el cual es un algoritmo que se basa en distancias, para este algoritmo se emplearon cuatro clústeres, la herramienta de *RapidMiner* genera automáticamente el agrupamiento y permite la creación de centroides, en la Figura 18 se observa el comportamiento de los centroides con base en los clústeres creados, cada clúster se identifica con un color al igual que cada centroide, para identificar a que clúster pertenece.

Figura 18

Algoritmo de agrupación k-means.

k-Means - Scatter Plot



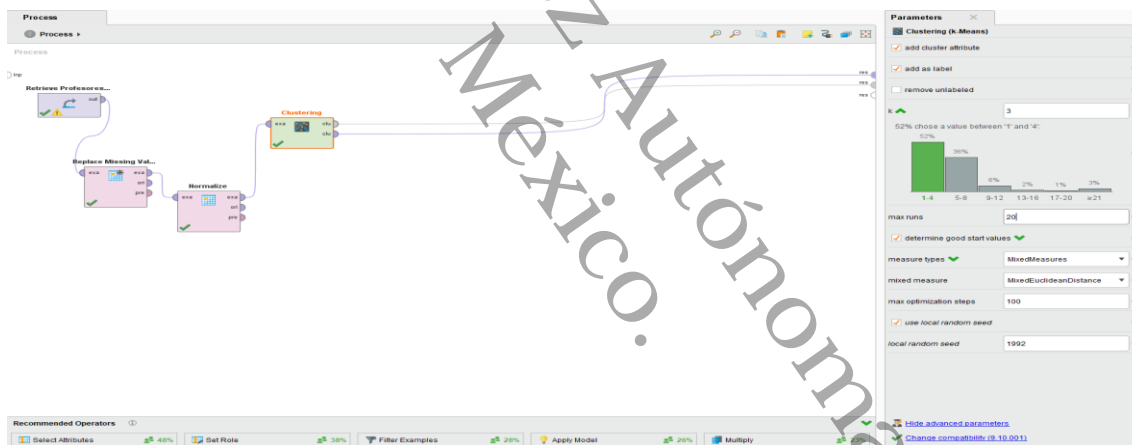
Nota: Elaboración propia a través de *RapidMiner*

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Al igual que con el algoritmo anterior, primero se seleccionó el *dataset*, seguido del operador *Replace Missing Value*, y se agregaron los operadores de *Normalize*, los cuales permiten calcular la media y la desviación estándar, por último, el operador de *Clustering k-means*, el cual se encarga del agrupamiento de los datos. *RapidMiner* tiene la función de elegir automáticamente cuál es la mejor opción en cada uno de sus algoritmos, en la Figura 19 se observa el proceso para la implementación de *K-means*.

Figura 19

Proceso para algoritmo K-means.



Nota: Elaboración propia

También se observa que del lado izquierdo se encuentran los valores que se utilizaron en este algoritmo, con ayuda de *RapidMiner* se optó por elegir cuatro clústeres, esto debido a que entre menor número de clúster se crean más grupos de datos similares y al realizar un análisis a los datos se obtuvo que el 52 % de factibilidad se encontraba dentro de estos primeros cuatro.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

K-means es un algoritmo de clasificación no supervisada, este asocia en K grupos objetos, dependiendo de sus características, esto se realiza gracias a la distancia que existe entre cada objeto y el centroide de cada clúster (González, 2020), en la Figura 20 se observan los cuatro clústeres creados y sus puntos más cercanos, en el clúster cero se encuentra con 44 centroides, en el clúster uno contiene 80, el clúster dos con 81 y el clúster tres con 37.

Figura 20

Clústeres creados con RapidMiner.

k-Means - Summary

Number of Clusters: 4

Cluster 0 44

En qué año ingresó al SISTEMA NACIONAL DE INVESTIGADORES? is on average 72.95% smaller, Edad en que obtuvo el grado de Doctor is on average 71.33% larger, Años laborados en la UJAT is on average 63.12% larger

Cluster 1 80

Años laborados en la UJAT is on average 13.61% larger, Edad is on average 10.70% larger, En qué año ingresó al SISTEMA NACIONAL DE INVESTIGADORES? is on average 9.05% larger

Cluster 2 81

En qué año ingresó al SISTEMA NACIONAL DE INVESTIGADORES? is on average 49.40% larger, Edad en que obtuvo el grado de Doctor is on average 46.35% smaller, Años laborados en la UJAT is on average 43.55% smaller

Cluster 3 37

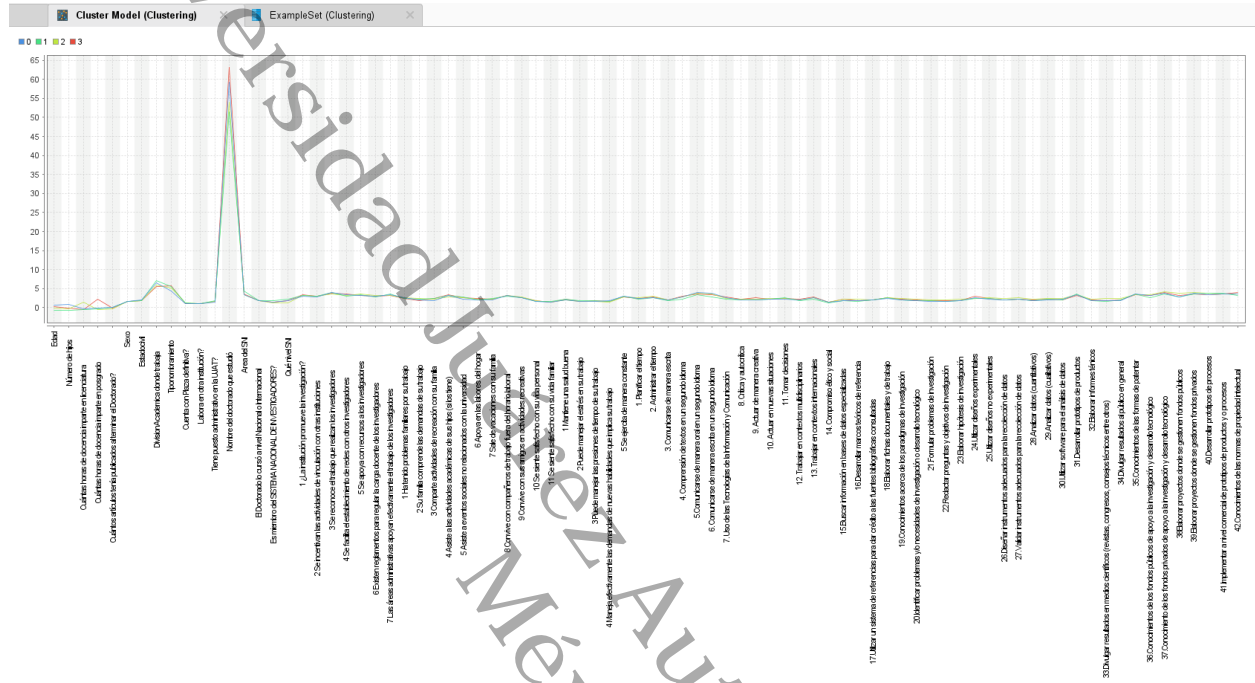
En qué año ingresó al SISTEMA NACIONAL DE INVESTIGADORES? is on average 40.98% smaller, Cuántas horas de docencia imparte en licenciatura is on average 28.73% larger, Cuántas horas de docencia imparte en posgrado is on average 17.72% smaller

Nota: Elaboración propia a través de la herramienta *RapidMiner*

En este mismo proceso, la herramienta permite crear un gráfico *Plot* donde se visualiza como los centroides se agrupan como punto más alto, en el nombre del doctorado que estudiaron los profesores investigadores (ver Figura 21).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Figura 21 Visualización de pico centroides agrupados en un Gráfico Plot.



Nota: Elaboración propia

Para este último algoritmo se tomaron en cuenta, las variables de afrontamiento, condiciones institucionales y habilidades para la investigación y se crearon asociaciones entre dichas variables, las cuales se presentan en la siguiente tabla se presentan cinco grupos de hasta cinco asociaciones.

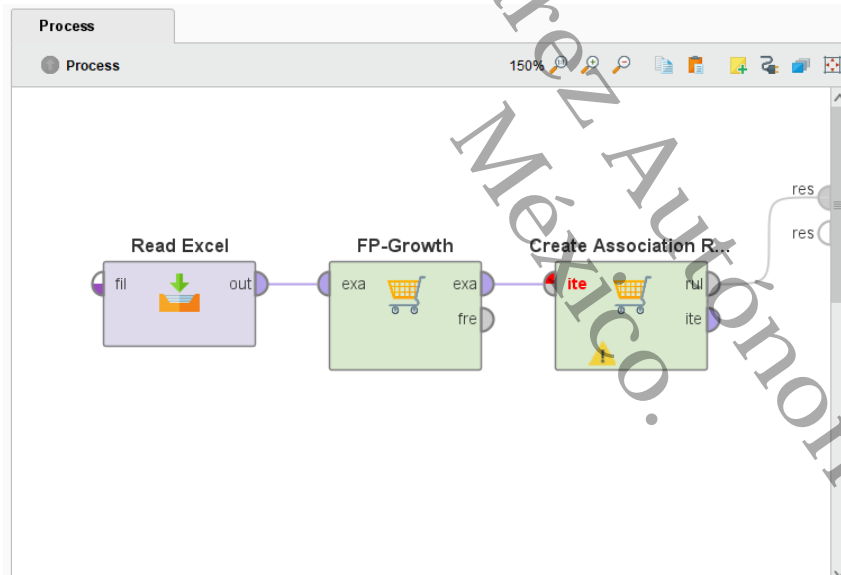
Para crear las reglas de asociación se utilizó el algoritmo Apriori, el proceso de este algoritmo comienza con adquirir ítems frecuentes, este algoritmo crea conjuntos de datos frecuentes estos están conformados por cuyos ítems tienen soportes superiores al soporte mínimo que solicite el usuario, las reglas que crea Apriori tiene en primer lugar los conjuntos de tamaño 1, luego de tamaño 2 y va creciendo hasta que el algoritmo ya no encuentre más conjuntos frecuentes (Sáenz et al., 2017).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Los operadores que se utilizaron en este algoritmo fueron *FP-Growth*, el cual se encarga de encontrar patrones frecuentes dentro del conjunto de datos y *Create Association Rules*, que como su nombre lo dice es el que permite la creación de las reglas de asociación a partir de un conjunto de patrones frecuentes, estos dos operadores ayudaron a descubrir relaciones entre los datos que aparentemente no tenían relación alguna, en la Figura 22 se observa el proceso con dichos operadores.

Figura 22

Proceso de creación de reglas de asociación.



Nota: Elaboración propia

Dentro de las reglas de asociación existen parámetros, estos parámetros se utilizan para medir la validez de dichas reglas con respecto al conjunto de datos, los principales parámetros con mayor calidad en las reglas de asociación son el soporte y la confianza,

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Para Sáenz (2017) el soporte de un *ítem* se calcula entre el número de transacciones que contiene un *ítem* y el número de transacciones en la base de datos, lo mismo para obtener una regla de decisión se divide un conjunto de ítems entre el número total de ítems. Para calcular la confianza el algoritmo divide el soporte obtenido dentro del conjunto de datos. En la Figura 23 se observan las reglas de asociación creadas.

Figura 23

Reglas de asociación.

Asociación Rules

[0] --> [No, Siempre] (confidence: 0.90)
 [Si] --> [Algo desarrollada, Casi siempre] (confidence: 0.900)
 [Desarrollada, Casi siempre] --> [A veces] (confidence: 0.900)
 [Algo en desacuerdo] --> [Desarrollada, Casi siempre] (confidence: 0.900)
 [0] --> [Desarrollada, Casi siempre] (confidence: 0.900)
 [Casi nunca] --> [Desarrollada, A veces] (confidence: 0.900)
 [Poco desarrollada] --> [Si, Casi siempre] (confidence: 0.900)
 [Casi siempre, Algo en desacuerdo] --> [A veces] (confidence: 0.900)
 [Nacional, 0] --> [Siempre] (confidence: 0.900)
 [Algo en desacuerdo] --> [No, Algo desarrollada, Casi siempre] (confidence: 0.900)
 [0] --> [No, Algo desarrollada, Casi siempre] (confidence: 0.900)
 [Desarrollada, Casi siempre] --> [No, Siempre] (confidence: 0.900)
 [No, Nada desarrollada] --> [Desarrollada, Poco desarrollada] (confidence: 0.900)
 [No, Si, 0] --> [A veces] (confidence: 0.900)
 [No, Bastante desarrollada, Masculino] --> [Siempre] (confidence: 0.900)
 [No, Bastante desarrollada, Bastante en desacuerdo] --> [A veces] (confidence: 0.900)
 [Nacional, 0] --> [No, A veces] (confidence: 0.900)
 [Poco desarrollada, Bastante en desacuerdo] --> [No, A veces] (confidence: 0.900)
 [Casi nunca] --> [Algo desarrollada, Desarrollada, Si] (confidence: 0.900)
 [Algo desarrollada, Desarrollada, Bastante en desacuerdo] --> [A veces] (confidence: 0.900)
 [Algo desarrollada, Nada desarrollada] --> [Desarrollada, Poco desarrollada] (confidence: 0.900)
 [Algo desarrollada, Nada desarrollada] --> [Casi siempre, A veces] (confidence: 0.900)
 [Nacional, 0] --> [Algo desarrollada, A veces] (confidence: 0.900)
 [Poco desarrollada, Bastante en desacuerdo] --> [Algo desarrollada, A veces] (confidence: 0.900)
 [Desarrollada, Algo en desacuerdo] --> [Si, Casi siempre] (confidence: 0.900)
 [Casi siempre, Algo en desacuerdo] --> [Desarrollada, Si] (confidence: 0.900)
 [Desarrollada, Nacional, Bastante en desacuerdo] --> [A veces] (confidence: 0.900)
 [Desarrollada, Poco desarrollada, Casado] --> [Nacional] (confidence: 0.900)

Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

posteriores iteraciones y según Timarán *et al.*, (2016) puede incluir algunos de estos aspectos: “visualización de los patrones extraídos, remoción de los patrones redundantes o irrelevantes y la traducción de los patrones útiles en términos que sean entendibles para el usuario” (p.67).

3 Por otra parte, se consolida el conocimiento descubierto para incorporarlo en otro sistema para posteriores acciones o, simplemente, para documentarlo y reportarlo a las partes interesadas; también para verificar y resolver conflictos potenciales con el conocimiento previamente descubierto.

Con la herramienta de minería de datos se analizaron e interpretaron, el resultado de cada uno de los algoritmos utilizados, los cuales brindaron la información necesaria para responder la pregunta que originó esta investigación.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Capítulo 4. Resultados

En este apartado se muestran los resultados obtenidos asociadas a las variables de investigación, trayectoria académica, condiciones institucionales, vida personal y familiar, afrontamiento y habilidades para la investigación de los profesores investigadores, a través de la herramienta de minería de datos *RapidMiner*, aplicando los algoritmos de árbol de decisión, agrupación y reglas de asociación.

4.1 Árbol de decisión

El primer resultado que se presenta es el de árbol de decisión ya que fue el primer algoritmo que se utilizó, se hicieron varias pruebas ya que con este algoritmo se inicia con un atributo en particular, de donde se van desprendiendo las ramas del árbol y se van observando diferentes posibilidades y resultados.

Trayectoria académica

El primer árbol que se construyó fue el de trayectoria académica, y como atributo particular se tomó si cuenta con plaza definitiva, para poder identificar si existe algún patrón que indique si la trayectoria del profesor afecta en su desarrollo como investigador.

El resultado que se obtuvo con este algoritmo ejecutado en la herramienta de *RapidMiner*, el árbol se construyó con una profundidad máxima de 15 para restringir el tamaño del árbol, este tamaño automáticamente la herramienta lo define como la mejor opción por eso se optó por dicha profundidad, el nivel de confianza que se utilizó fue de 0.1 para minimizar el margen de error.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Con la aplicación de dicho algoritmo también se obtuvo una matriz de confusión con una precisión del 87.60% de aciertos y un 3.90% de margen de error, donde 199 instancias indicaron correctamente que sí cuentan con plaza definitiva y 21 instancias fueron clasificadas en el margen de error, por otra parte 9 instancias indicaron que el caso de que no cuentan con plaza definitiva se analizó correctamente, y 13 instancias se encuentran en el margen de error como se observa en la Figura 24.

Figura 24

Matriz de confusión si el profesor cuenta con plaza definitiva

Accuracy:87.60% +/- 3.90% (micro average: 87.60%)

	True Si Plaza	True no plaza
Pred. Si plaza	199	21
Pred. No plaza	9	13

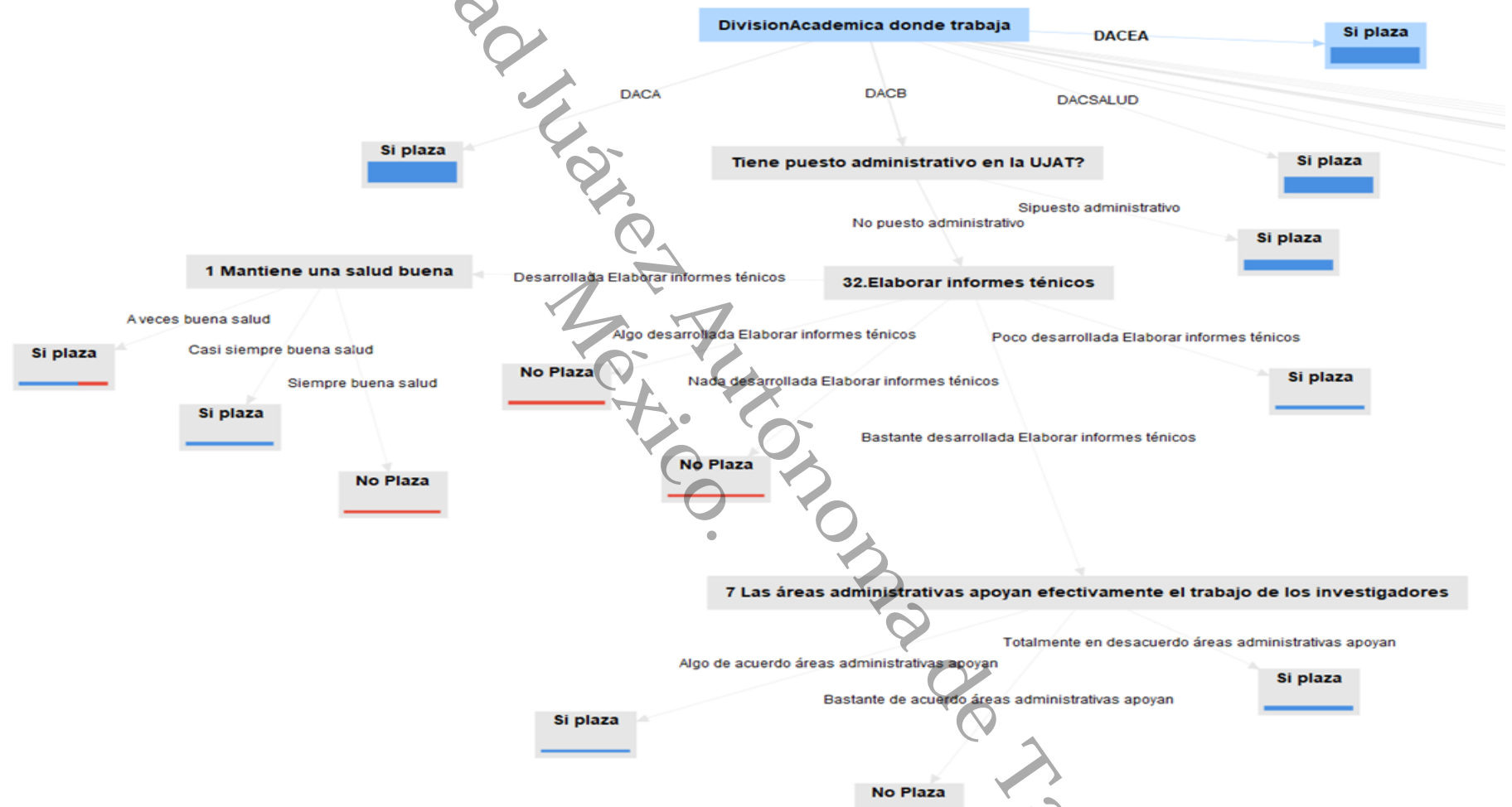
Nota: Elaboración propia

El patrón de conducta que se encontró en este árbol con base a la pregunta de atributo principal de que si el profesor cuenta con plaza definitiva fue que el profesor para que cuente con plaza definitiva necesita tener desarrollada las habilidades para la investigación de algo desarrollada a bastante desarrollada. En las Figuras 25 y 26 se observa el árbol de decisión obtenido, donde se visualizan de manera gráfica todas las ramas, conexiones y predicciones de los patrones encontrados.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Figura 25

Árbol de decisión trayectoria académica primera parte.

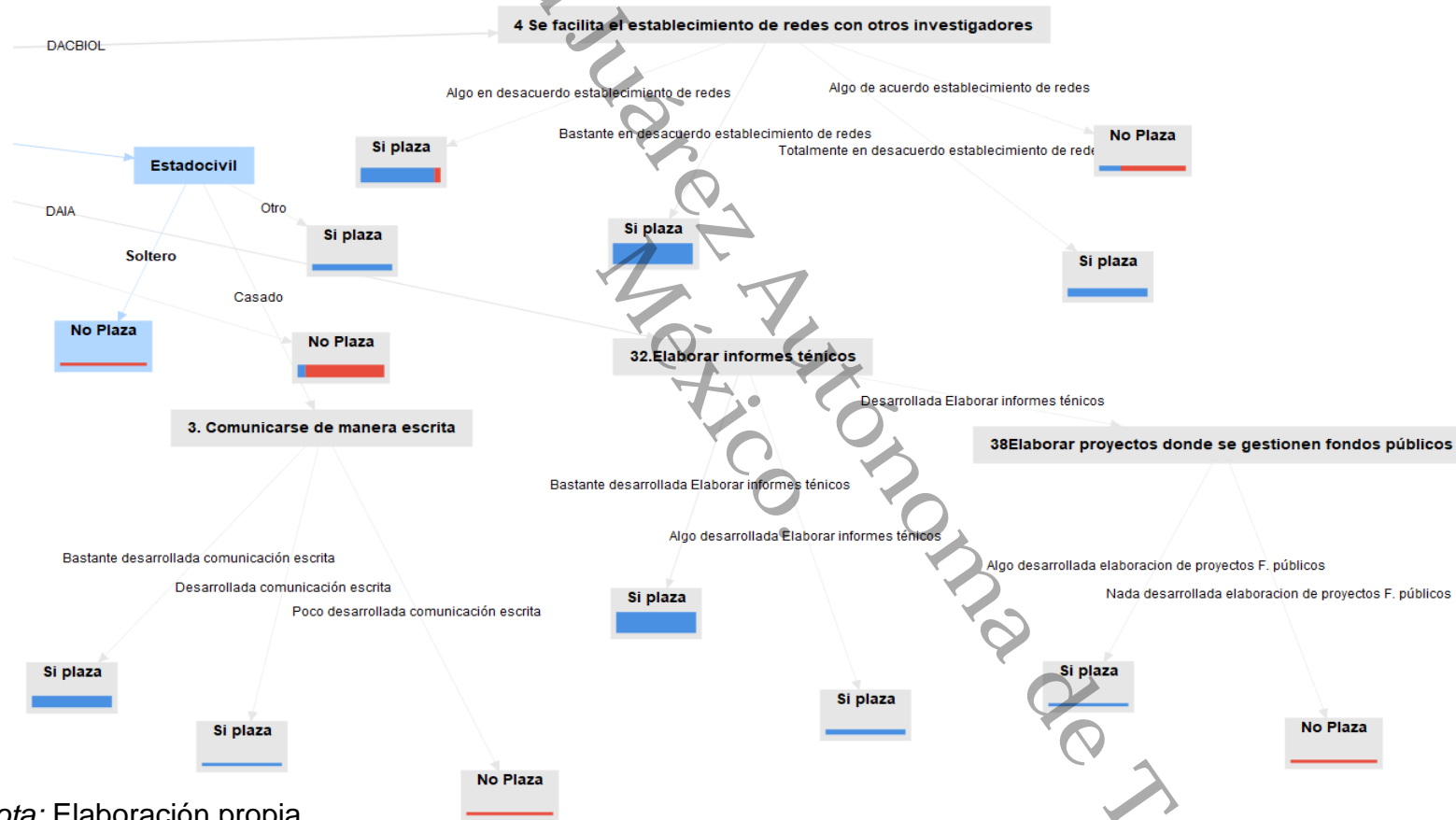


Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Figura 26

Árbol de decisión trayectoria académica segunda parte



Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Condiciones institucionales

En el segundo árbol se encuentran las condiciones institucionales las cuales tienen como atributo principal si la universidad promueve la investigación, para identificar los patrones de conducta asociadas a la variable de condiciones institucionales.

Con los resultados obtenidos respecto a esta variable se pudo observar que la mayoría de los docentes se encuentran bastante de acuerdo con el apoyo que reciben por parte de la institución, esto se observa en la matriz de confusión donde se observa un 39.82% de precisión cuando la respuesta es acertada y 14.94% cuando la clasificación fue incorrecta ver Figura 27.

Figura 27

Matriz de confusión condiciones institucionales

Accuracy:39.82% +/- 14.94% (micro average: 39.67%)

	True Algo en desacuerdo institución	True Bastante de acuerdo institución	True Algo de acuerdo institución	True Totalmente en desacuerdo institución	True Bastante en desacuerdo institución
Pred. Algo en desacuerdo institución	20	6	17	6	20
Pred. Bastante de acuerdo institución	7	33	9	3	1
Pred. Algo de acuerdo institución	9	10	15	3	5
Pred. Totalmente en desacuerdo institución	2	2	1	0	11
Pred. Bastante en desacuerdo institución	15	3	4	12	28

Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

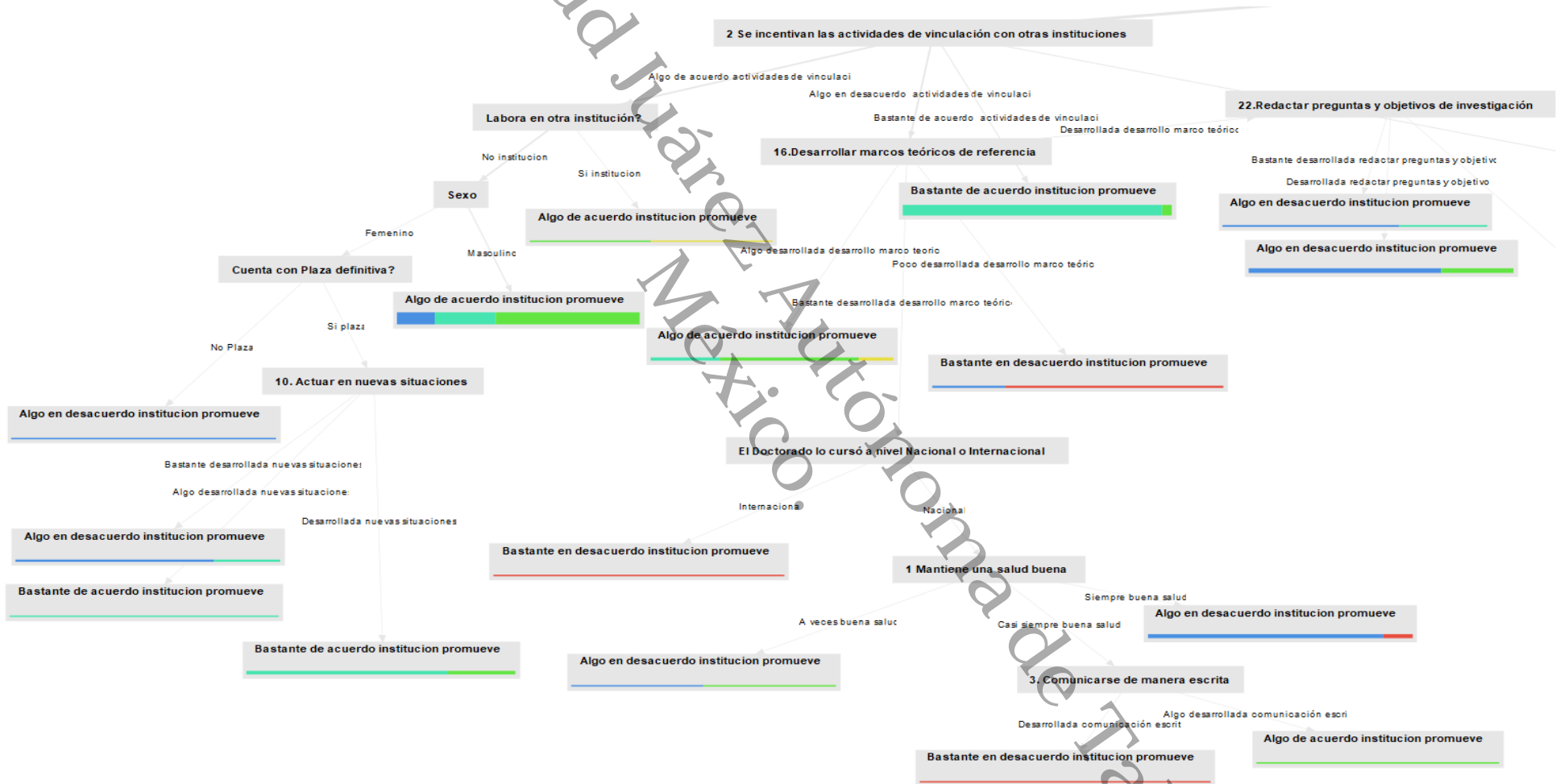
En los datos analizados por este árbol la clase de mayor precisión fue que los profesores se encuentran satisfechos con las condiciones institucionales ya que un 59.65% de los profesores están bastante de acuerdo en que la institución promueve la investigación, con esta variable también se creó un árbol de decisión el cual puede observarse en las Figuras 28 y 29.

Universidad Juárez Autónoma de Tabasco.
México.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Figura 28

Árbol de decisión condiciones institucionales primera parte

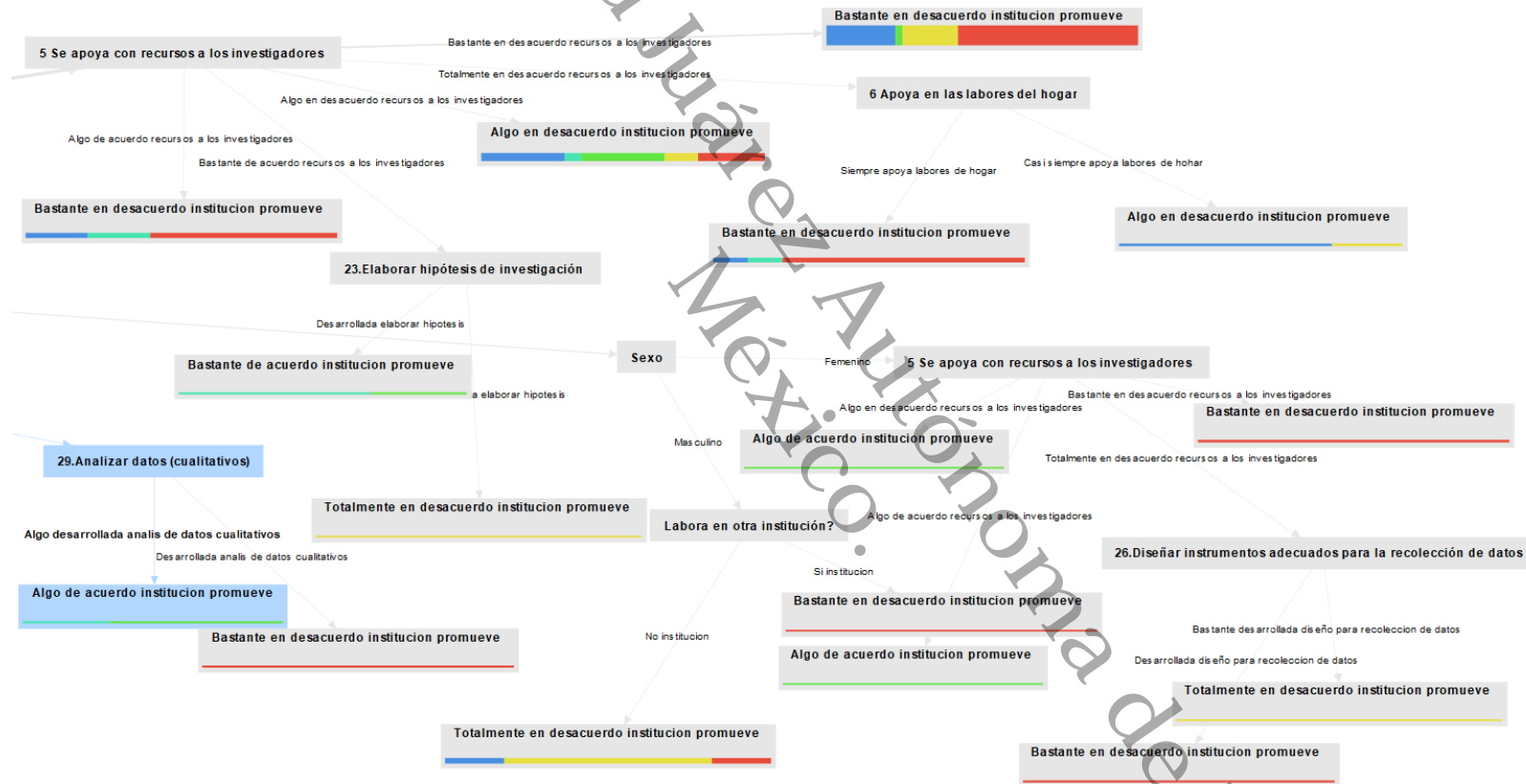


Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

.Figura 29

Árbol de decisión condiciones institucionales segunda parte.



Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Vida personal y familiar

La tercera variable cuestiona si el profesor ha tenido problemas familiares o personales debido a su trabajo, el árbol que se creó cuenta como atributo principal dicha pregunta.

Como resultado se creó tanto una matriz de confusión como un árbol de decisión (ver Figuras 30 y 31) para identificar los patrones que se asocian con esta tercera variable, en la matriz de confusión se observa un 50.07% de precisión en los aciertos y un 8.66% en clasificaciones erróneas, dando como resultado un 66% de predicción en el que el profesor considera que los problemas familiares no representan ningún obstáculo para realizar sus investigaciones.

Figura 30

Matriz de confusión vida personal y familiar.

Accuracy:50.07+/- 8.66% (micro average: 50.00%)

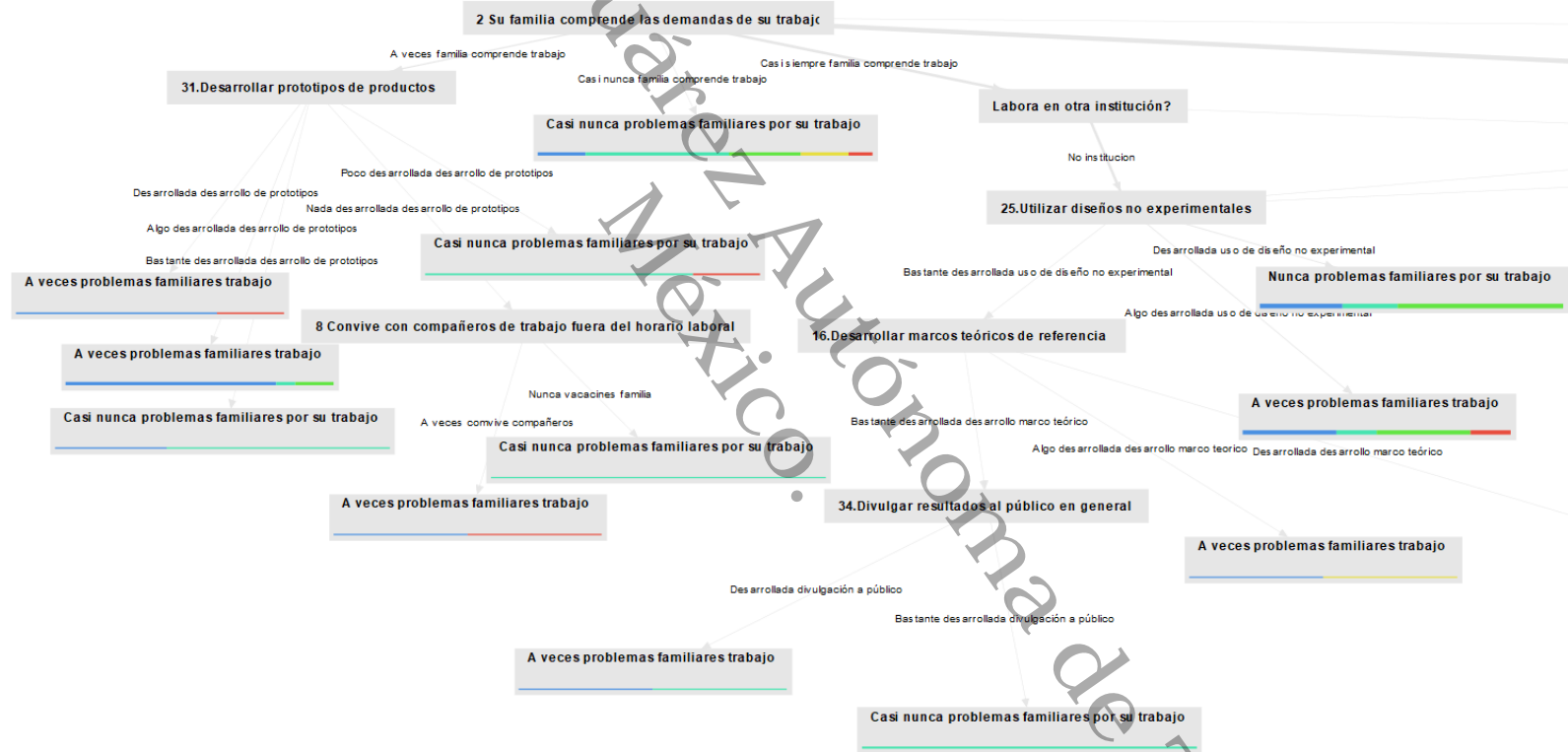
	True A veces problemas familiares	True Casi nunca problemas familiares	True Nunca problemas familiares	True Siempre problemas familiares	True Casi siempre problemas familiares
Pred. A veces problemas familiares	17	16	13	1	4
Pred. Casi nunca problemas familiares	10	10	13	2	3
Pred. Nunca problemas familiares	20	25	94	5	3
Pred. Siempre problemas familiares	0	1	0	0	1
Pred. Casi siempre problemas familiares	1	1	2	0	0

Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Figura 31

Árbol de decisión vida personal y familiar.



Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Afrontamiento

El atributo principal con el que se construyó este árbol de decisión fue si el profesor puede mantener una buena salud a pesar de sus actividades.

Como resultado del algoritmo de árbol de decisión se pudo observar un comportamiento favorable para los profesores ya que la mayoría de ellos pueden afrontar los problemas de estrés mayormente, lo que dio como resultado positivo ya que el 47.7 % de los profesores encuestados dieron como respuesta que la mayoría de las veces pueden manejar situaciones que le generen estrés.

Esto se puede ver en la matriz de confusión donde el 51.30% de instancias resultó ser positiva y un 10.85% erróneas, el algoritmo predice que en un 51.24% de las veces el profesor logra mantener una buena salud, y puede afrontar problemas como el estrés laboral y social, en la Figura 32 se observa la matriz.

Figura 32

Matriz de confusión afrontamiento.

Accuracy:51.30% +/- 10.85% (micro average: 51.24%)

	True A veces buena salud	True Casi siempre buena salud	True Siempre buena salud	True Casi nunca buena salud	True Nunca buena salud
Pred. A veces buena salud	20	26	2	3	1
Pred. Casi siempre buena salud	29	61	15	6	0
Pred. Siempre buena salud	6	26	43	0	1
Pred. Casi nunca buena salud	0	3	0	0	0
Pred. Nunca buena salud	0	0	0	0	0

Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

También se observa el árbol, ramas y predicciones de dicha variable, donde se logra visualizar de una mejor manera, como el algoritmo fue encontrando patrones, ver Figuras 33 y 34.

Universidad Juárez Autónoma de Tabasco.
México.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Figura 33

Árbol de decisión afrontamiento primera parte



Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Habilidades para la investigación

Para la variable de habilidades para la investigación se usó como atributo principal, si el profesor sabe formular problemas de investigación, esto para poder obtener patrones que indiquen si el profesor investigador, tiene dificultades en el conocimiento de la elaboración de investigaciones.

El resultado obtenido de este análisis indica que el profesor cuenta con un buen desarrollo de esta habilidad, en la matriz de confusión indica que hay una instancia de valores de precisión de un 77.32% con 5.95% de valores erróneos ver Figura 35.

Figura 35.

Matriz de confusión habilidades para la investigación.

Accuracy:79.32% +/- 5.95% (micro average: 79.34%)

	True Bastante desarrollada formular	True Desarrollada formular	True Algo desarrollada formular	True Poco desarrollada formula	True Nada desarrollada formular
Pred. Bastante desarrollada formular	94	10	2	0	1
Pred. Desarrollada formular	7	77	15	1	0
Pred. Algo desarrollada formular	1	7	19	4	0
Pred. Poco desarrollada formula	0	0	1	2	1
Pred. Nada desarrollada formular	0	0	0	0	0

Nota: Elaboración propia

En la Figura 36 se observa el árbol de decisión construido con esta variable, dando como resultado un árbol de 15 nodos y ramificaciones.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Figura 36

Árbol de decisión habilidades para la investigación.



Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

4.2 Clustering Algoritmo K-means

Para este algoritmo en el cual también se realizaron diferentes pruebas, cambiando el número de clúster hasta llegar al óptimo, empezando con diez clústeres en la primera prueba, seis clústeres en la segunda y cuatro clústeres en la última prueba este último con la ayuda de la herramienta de *RapidMiner* que como una de sus funciones es mostrar cuál es el mejor algoritmo para el dataset que se está analizando, como resultado se implementaron dos clústeres.

En la Figura 37 se observa el resultado que se obtuvo implementando la técnica de *K-means* como técnica de agrupamiento, en dicha figura se logra apreciar los grupos de instancias con datos aleatorios con los cuales el algoritmo realizó las agrupaciones.

Figura 37

Resultado de agrupamiento de trayectoria académica con *RapidMiner*.

Attribute	Cluster_0	Cluster_1
Edad	15.937	16.798
Sexo	1.658	1.524
Estado civil	1.861	1.821
División académica donde trabaja	6.734	6.048
Tipo nombramiento	5.367	4.369
Cuenta con plaza definitiva	1.158	1.107
Labora en esta institución	1.051	1.119
Tiene puesto administrativo en la UJAT	1.665	1.631
Cuántas horas de docencia imparte en la licenciatura	11.095	11.679
Cuántas horas de docencia imparte en la posgrado	3.247	4.655
Nombre del doctorado que estudió	57.747	52.048
Área del SNI	3.924	3.429
EL doctorado lo curso a nivel nacional o internacional	1.854	1.845
Es miembro del Sistema Nacional de Investigadores	1.456	1.583

Nota: Elaboracion propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Respecto al agrupamiento en la variable de condiciones institucionales se observa que los centroides dentro de ambos cluster tienen tamaños similares siendo más notoria una diferencia en el atributo de si se reconoce el trabajo que realizan los investigadores en el cluster 0 con 2.538 y el clúster 1 con 3.048 ver Figura 38.

Figura 38

Resultados de agrupamiento de Condiciones institucionales con RapidMiner.

Attribute	Cluster_0	Cluster_1
1 ¿La institución promueve la investigación?	2.975	2.976
2 Se incentivan las actividades de vinculación con otras instituciones	2.513	2.631
3 Se reconoce el trabajo que realizan los investigadores	2.538	3.048
4 Se facilita el establecimiento de redes con otros investigadores	2.538	2.5
5 Se apoya con recursos a los investigadores	3.215	3.119
6 Existen reglamentos para regular la carga docente de los investigadores	2.646	2.536
7 Las áreas administrativas apoyan efectivamente el trabajo de los investigadores	3.405	3.405

Nota: Elaboración propia

De igual forma para la variable de vida personal y familiar se observan que el tamaño entre el clúster 1 y 0 no hay mucha diferencia ver Figura 39.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Figura 39

Resultado de agrupamiento de vida personal y familiar con RapidMiner.

Attribute	Cluster 0	Cluster 1
1 Ha tenido problemas familiares por su trabajo	2.424	2.667
2 Su familia comprende las demandas de su trabajo	2.272	2.036
3 Comparte actividades de recreación con su familia	2.329	1.893
4 Asiste a las actividades académicas de sus hijos (si los tiene)	3.222	3.06
5 Asiste a eventos sociales no relacionadas con la universidad	2.500	2.583
6 Apoya en las labores del hogar	2.234	1.833
7 sale de vacaciones con su familia	2.228	2.298
8 Convive con compañeros de trabajo fuera del horario laboral	3.089	3.274
9 Convive con sus amigos en actividades recreativas	2.665	2.785
10 Se siente satisfecho con su vida personal	1.873	1.663
11 Se siente satisfecho con su vida familiar	1.582	1.35

Nota: Elaboracion propia

En la variable de afrontamiento en el atributo de si el profesor se ejercita constantemente se observa la diferencia más notoria entre el tamaño del clúster siendo el clúster 1 más grande que el cluster 0 (ver Figura 40).

Figura 40

Resultado de agrupamiento de afrontamiento con RapidMiner.

Attribute	Cluster 0	Cluster 1
1 Mantiene una salud buena	2.032	2.285
2 Puede manejar el estrés en su trabajo	1.728	1.881
3 Puede manejar las presiones de tiempo en su trabajo	1.709	1.905
4 Maneja efectivamente las demandas de nuevas habilidades que implica su trabajo	1.658	1.833
5 Se ejercita de manera constante	2.747	3.512

Nota: Elaboracion propia

En la variable de habilidades para la investigación existen mas diferencias de tamaños entre los clusters, ya que es la variable con más atributos las diferencias más notorias que se observan son administrar tiempo clúster 0 con 2.949 y el clúster 1 con

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

2.333, comprensión de textos en un segundo idioma cluster 0 con 2.804 y clúster con 2.167, toma de decisiones clúster 0 con 2.722 clúster 1 con 1.750. Se observa que en la mayoría de los casos el clúster 0 es mas grande que el clúster 1 ver Figura 41.

Figura 41

Resultado de agrupamiento de habilidades para la investigación en RapidMiner.

Attribute	Cluster_0	Cluster_1
1 Planificar tiempo	2.525	2.095
2 Administrar tiempo	2.949	2.333
3 Comunicarse de manera escrita	1.994	1.819
4 Comprensión de textos en un segundo idioma	2.804	2.157
5 Comunicarse de manera oral en un segundo idioma	3.082	2.631
7 Uso de las tecnologías de la información y comunicación	2.867	2.255
8 Crítica y autocrítica	2.677	1.952
9 Actuar de manera creativa	2.304	1.702
10 Actuar en nuevas situaciones	2.430	1.69
11 toma de decisiones	2.525	1.754
12 trabajar en contextos multidisciplinares	2.722	1.75
13 Trabajar en contextos internacionales	2.133	1.405
14 Compromiso ético y social	1.563	1.94
15 buscar información en base de datos especializadas	2.177	1.143
16 Desarrollar marcos teóricos de referencia	2.139	1.429
17 Utilizar un sistema de referencias para dar crédito a las fuentes bibliográficas	2.019	1.205
18 Elaborar fichas documentales y de trabajo	2.481	1.94
19 Conocimientos acerca de los paradigmas de investigación	2.443	2
20 Identificar problemas y/o necesidades de investigación o desarrollo tecnológico	2.203	1.417
21 Formular problemas de investigación	2.063	1.457
22 Redactar preguntas y objetivos de investigación	1.968	1.345
23 Elaborar hipótesis de investigación	2.209	1.298
24 Utilizar diseños experimentales	2.285	1.457

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

25 Utilizar diseños no experimentales	2.354	2.488
26 Diseñar instrumentos adecuados para la recolección de datos	2.285	1.859
27 Validar instrumentos adecuados para la recolección de datos	2.399	1.679
28 Analizar datos (cuantitativos)	1.987	1833
29 Analizar datos (cualitativos)	2.342	1.69
30 Utilizar software para el análisis de datos	2.373	1.725
31 Desarrollar prototipos de productos	2.671	2.952
32 Elaborar informes técnicos	2.152	1.345
33 divulgar resultados en medios científicos (revistas, congresos, consejos técnicos)	2.177	1.457
34 Divulgar resultados al público en general	2.361	1.512
35 Conocimientos de las formas de patentar	2.639	2.957
36 Conocimientos de los fondos públicos de apoyo a la investigación y desarrollo	3.095	2.69
37 Conocimientos de los fondos privados de apoyo a la investigación y desarrollo	3.063	2.655
38 Elaborar proyectos donde se gestionen fondos públicos	3.089	2.679
39 Elaborar proyectos donde se gestionen fondos privados	3.057	2.59
40 Desarrollar prototipos de procesos	3.133	3.262
41 Implementar a nivel comercial de prototipos de productos y/o procesos	3.418	3.405
42 Conocimientos de las normas de propiedad intelectual	3.272	3.31

Nota: Elaboracion propia

Como resultado de la implementación se obtuvo que en el clúster 0 se encuentran 158 ítems y el clúster 1 solamente 84 ítems agrupados, siendo el clúster 0 el que contiene la mayoría de los datos agrupados ver Figura 42.

Figura 42

Asignación de ítems en los clústeres.

Cluster Model

```
Cluster 0: 158 items
Cluster 1: 84 items
Total number of items: 242
```

Nota: Elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

De igual forma en la Figura 43 se muestra cómo están distribuidos los ítems que conforman los clústeres, de acuerdo con la variable de la división donde labora el profesor y qué tanto manejan el estrés dentro de su trabajo, lo que da como resultado que la mayoría de los profesores sí pueden manejar el estrés laboral.

Figura 43

Gráfica de agrupamiento en RapidMiner



Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Nota: elaboración propia

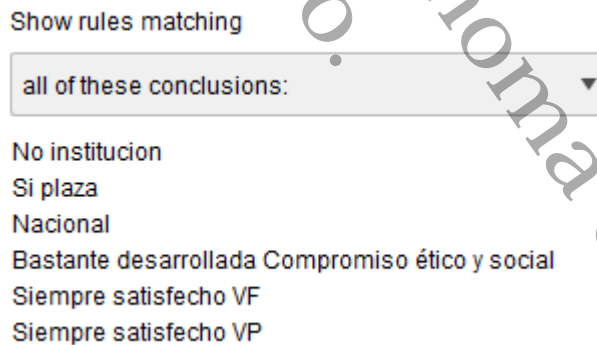
4.3 Reglas de asociación

Con el algoritmo de reglas de asociación se utilizó un soporte mínimo del 95% y un 80% de confianza mínima para el conjunto de datos que se analizó dando como resultado que el número de conjunto de frecuencias fueron 164 instancias, con un máximo de 4 ítems, las cuales están divididas en 38 con un solo ítems, 75 con dos ítems, 44 con tres ítems, y 7 con cuatro ítems, creando así 100 reglas de asociación.

Con el análisis de este algoritmo se encontraron 6 coincidencias de reglas estas coincidencias son las conclusiones que encontró la herramienta, con las que se construyeron cada una de las reglas de asociación ver Figura 44.

Figura 44

Resultados *coincidencia de reglas de asociación.*



Nota: elaboración propia

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

En la Figura 45 se muestran las reglas de asociación creadas con *RapidMiner* donde se observan los diferentes parámetros de confianza y soporte entre otros parámetros para cada una de las reglas.

Figura 45

Reglas de asociación con RapidMiner.

Premises	Conclusión	Support	Confidence	LaPlace	Gain	p-8	Lift
Si plaza, No puesto administrativo	No institución, Nacional	0.430	0.819	0.930	-0.62	0.018	1.043
Siempre satisfecho VP	No institución, Siempre satisfecho VF	0.508	0.828	0.934	0.723	0.152	1.427
Nacional, Siempre satisfecho VP	No institución, Siempre satisfecho VF	0.434	0.833	0.943	0.607	0.133	1.440
No puesto administrativo	No institución, Nacional	0.554	0.848	0.940	0.752	0.133	1.080
Si plaza, Nacional, Siempre satisfecho VF	No institución	0.409	0.884	0.963	0.517	0.041	0.955
No SNI	No institución	0.442	0.884	0.961	0.558	0.019	0.955
Si plaza, Siempre satisfecho VF	No institución	0.483	0.888	0.960	0.607	0.021	0.958
Si plaza, Siempre satisfecho VF, Siempre satisfecho VP	No institución	0.430	0.889	0.964	0.537	0.018	0.960
Siempre familia comprende trabajo	No institución	0.434	0.890	0.964	0.541	0.018	0.961
Nacional, Siempre satisfecho VF	No institución	0.500	0.896	0.963	0.916	0.016	0.968
Bastante desarrollada Compromiso ético y social	No institución	0.430	0.897	0.966	0.529	0.014	0.969
Siempre satisfecho VF	No institución	0.579	0.897	0.960	0.711	0.018	0.970
Nacional, Siempre satisfecho VF, Siempre satisfecho VP	No institución	0.434	0.897	0.967	0.533	0.014	0.970
Siempre satisfecho VF, Siempre satisfecho VP	No institución	0.508	0.898	0.963	0.624	0.016	0.970
Casado, Siempre satisfecho VF	No institución	0.438	0.898	0.967	0.537	0.013	0.970

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Si plaza, Siempre satisfecho VP	No institución	0.479	0.899	0.965	0.587	0.014	0.971
Si plaza, Masculino	No institución	0.463	0.903	0.967	0.562	0.011	0.976
Casado, Masculino	No institución	0.430	0.904	0.969	0.521	0.010	0.970
Bastante desarrollada Compromiso ético y social	No institución	0.430	0.904	0.969	0.521	0.010	0.970
Nacional, Siempre satisfecho VP	No institución	0.471	0.905	0.967	0.570	0.011	0.970
Siempre satisfecho VP	No institución	0.558	0.908	0.964	0.674	0.012	0.979
Bastante desarrollada redactar preguntas de investigación	No institución	0.409	0.908	0.972	0.492	0.008	0.811
Nacional, Bastante desarrollada buscar en BD	No institución	0.409	0.908	0.972	0.492	0.008	0.811
Bastante desarrollada divulgación a público	No institución	0.413	0.909	0.972	0.496	0.008	0.818
Si plaza, Nacional, Bastante desarrollada compromiso ético y social	No institución	0.465	0.909	0.970	0.545	0.008	0.818
Nunca problemas familiares por su trabajo	No institución	0.459	0.910	0.970	0.550	0.008	0.825

Nota: Elaboración propia

Estos parámetros sirven para limitar el número de reglas que se encontraron dentro del *dataset*, es porque para comprender mejor estas reglas de asociación con ayuda de la herramienta de minería de datos *RapidMiner* se crearon las asociaciones entre los datos analizados y cada una de las variables utilizadas en esta investigación en la Figura 46 se observa la descripción de cada una de las reglas.

Figura 46

Descripción de las reglas de asociación con *RapidMiner*.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Association Rule:

1. [Si plaza, PIC] --> [No institución, Nacional] (confidence: 0.800)
2. [PTC] --> [No institución, Nacional] (confidence: 0.803)
3. [No puesto administrativo] --> [Si plaza] (confidence: 0.804)
4. [Bastante desarrollada buscar en BD] --> [No institución, Nacional] (confidence: 0.805)
5. [Casado] --> [No institución, Nacional] (confidence: 0.811)
6. [Si plaza, Siempre satisfecho VP] --> [No institución, Siempre satisfecho VF] (confidence: 0.806)
7. 0.806)
8. [No institución, Nacional, Siempre satisfecho VF] --> [Si plaza] (confidence: 0.818)
9. [Nacional, Masculino] --> [Si plaza] (confidence: 0.815)
10. [Si plaza, No puesto administrativo] --> [No institución, Nacional] (confidence: 0.819)
11. [Nunca problemas familiares por su trabajo] --> [Siempre satisfecho VF] (confidence: 0.820)
12. [No institución, Si plaza, Bastante desarrollada Compromiso ético y social] [Nacional] (confidence: 0.821)
13. [No institución, Masculino] --> [Si plaza] (confidence: 0.824)
14. [Siempre satisfecho VP] --> [No institución, Siempre satisfecho VF] (confidence: 0.826)
15. [No institución, Si plaza] --> [Nacional] (confidence: 0.827)
16. [Si plaza, Bastante desarrollada Compromiso ético y social] --> [Nacional] (confidence: 0.829)
17. [Si plaza, Siempre [Nacional, Siempre (confidence: 0.829)
18. [satisfecho VF] --> [Si plaza] (confidence: 0.830)
19. [No institución, Nacional] --> [Si plaza] (confidence: 0.832)
20. [Si plaza] --> [Nacional] (confidence: 0.832)
21. [No institución, Nacional, Casado] --> [Si plaza] (confidence: 0.832)
22. [Nacional, Siempre satisfecho VD] --> [No institución, Siempre satisfecho VF] (confidence: 0.833)
23. [No institución, Siempre satisfecho VF] --> [Si plaza] (confidence: 0.836)
24. [Bastante desarrollada buscar en BD] --> [Bastante desarrollada Compromiso ético y social] (confidence: 0.837)
25. [Masculino] --> [Si plaza] (confidence: 0.838)
26. [No institución, Bastante desarrollada Compromiso ético y social] --> [Nacional] (confidence: 0.839)
27. [Si plaza, Masculino] --> [Nacional] (confidence: 0.839)
28. [Nacional, Casado] --> [Si plaza] (confidence: 0.839)
29. [Siempre familia comprende trabajo] --> [Nacional] (confidence: 0.838)
30. [Nacional] --> [Si plaza] (confidence: 0.840)
31. [Bastante desarrollado uso de sistema de referencias] --> [Bastante desarrollada Compromiso ético y social] (confidence: 0.840)
32. [No institución, PIC] --> [Si plaza, Nacional] (confidence: 0.840)
33. [Si plaza, No puesto administrativo] --> [Nacional] (confidence: 0.843)
34. [PTC] --> [Si plaza, Nacional] (confidence: 0.843)
35. [Si SNI] --> [Si plaza] (confidence: 0.843)
36. [Si SNI] --> [Nacional] (confidence: 0.843)
37. [Nunca problemas familiares por su trabajo] --> [Si plaza] (confidence: 0.844)
38. [No institución, Siempre satisfecho VP] --> [Nacional] (confidence: 0.844)
39. [Bastante desarrollada Compromiso ético y social] --> [Nacional] (confidence: 0.846)
40. [Bastante desarrollada buscar en BD] --> [Si plaza] (confidence: 0.846)
41. [No institución, Siempre satisfecho VF, Siempre satisfecho VP] --> [Si plaza] (confidence: 0.846)
42. 0.846)
43. [Siempre satisfecho VP] --> [Nacional] (confidence: 0.846)
44. [Siempre satisfecho VF] --> [Si plaza] (confidence: 0.846)
45. [No institución, Nacional, Bastante desarrollada Compromiso ético y social] --> [Si plaza] (confidence: 0.846)
46. [No institución, Si plaza, Siempre satisfecho VF] --> [Nacional] (confidence: 0.846)
47. [Casado, Siempre satisfecho VF] --> [Si plaza] (confidence: 0.847)
48. [No puesto administrativo] --> [No institución, Nacional] (confidence: 0.848)
49. [No institución] --> [Nacional] (confidence: 0.848)
50. [Si plaza, Siempre satisfecho VE] --> [Nacional] (confidence: 0.848)
51. [Nacional, Siempre satisfecho VP] --> [Si plaza] (confidence: 0.849)

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

52. [Nacional, Bastante desarrollada Compromiso ético y social] --> (Si plaza) (confidence: 0.852)

53. [Nunca problemas familiares por su trabajo] --> [Nacional] (confidence: 0.52)

54. [No institución, Si plaza, No puesto administrative) --> [Nacional] (confidence: 0.852)

56. [No institución, Casado) --> [Si plaza) (confidence: 0.853)

57. [No institución) --> [Si plaza) (confidence: 0.853)

58. [No institución, Masculino) --> [Nacional] (confidence: 0.853)

59. [Bastante desarrollada Compromiso ético y social, Siempre satisfecho VF] --> [Si plaza) (confidence: 0.853)

60. [No institución, Siempre satisfecho VF, Siempre satisfecho VP] --> [Nacional] (confidence: 0.856)

61. [Siempre satisfecho UF, Siempre satisfecho VF) --> [Si plaza] (confidence: 0.854)

62. [Siempre satisfecho VE, Siempre satisfecho VP) --> [Nacional] (confidence: 0.854)

64. [No institución, Si plaza, PIC) --> [Nacional] (confidence: 0.855)

65. [Si plaza, PTC] => [Nacional] (confidence: 0.356)

66. [Bastante desarrollado uso de sistema de referencias] --> [Nacional] (confidence: 0.857)

67. [No institución, PIC) --> [Nacional] (confidence: 0.857)

68. [No institución, Si plaza, Casado] --> [Nacional] (confidence: 0.857)

69. [Casado] --> (Si plaza) (confidence: 0.858)

70. [Masculino) ==> (Nacional) (confidence: 0.858)

71. [PTC] --> [Nacional] (confidence: 0.858)

72. [No institución, Siempre satisfecho VP) --> [Si plaza) (confidence: 0.853)

73. [No NT] --> [Nacional] (confidence: 0.860)

74. [Bastante desarrollada Compromiso ético y social, Siempre satisfecho VD) --> [Nacional] (confidence: 0.661)

75. [Si plaza, Casado] --> [Nacional] (confidence: 0.862)

76. [No institución, Siempre satisfecho VF] --> [Nacional] (confidence: 0.864)

77. [No institución, Bastante desarrollada Compromiso ético y social] --> [Si plaza) (confidence: 0.865)

79. [Siempre satisfecho VE) --> [Nacional] (confidence: 0.865)

80. [Bastante desarrollado uso de sistema de referencias] --> [Si plaza) (confidence: 0.866)

81. [Siempre satisfecho VD) => [Si plaza) (confidence: 0.866)

82. [Nacional, Siempre satisfecho VI) --> [Siempre satisfecho VD] (confidence: 0.867)

84. [No puesto administrativo) --> [Nacional] (confidence: 0.867)

85. [No institución, Nacional, Siempre satisfecho VE] --> [Siempre satisfecho VP) (confidence: 0.563)

86. [Bastante desarrollada Compromiso ético y social] ==> (Si plaza) (confidence: 0.263)

87. [Bastante desarrollada Compromiso ético y social, Siempre satisfecho VD) --> [Si plaza) (confidence: 0.870)

88. [Bastante desarrollada Compromiso ético y social, Siempre satisfecho VE) --> [Nacional] (confidence: 0.871)

90. [Siempre familia comprende trabajo) --> [Si plaza) (confidence: 0.873)

91. [Casado, Bastante desarrollada Compromiso ético y social] --> [Si plaza) (confidence: 0.376)

92. [No institución, No puesto administrativo

93. [No SNI) --> (Si plaza) (confidence: 0.876)

94. [Siempre satisfecho VF) --> [Siempre satisfecho UP) (confidence: 0.878)

95. [No institución, Casado) --> [Nacional] (confidence: 0.878)

96. [Casado, Masculino) --> (Nacional) (confidence: 0.878)

97. [No institución, Siempre satisfecho VE] --> [Siempre satisfecho VE] (confidence: 0.878)

98. [Casado, Siempre satisfecho VF) --> [Siempre satisfecho VP) (confidence: 0.881)

99. [Casado] --> (Nacional) (confidence: 0.882)

100.

Nota: Elaboración propia

103

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Al ser una cantidad de reglas grande, se optó por realizar un filtrado de las mejores tomando en cuenta las que tuvieran mayor confianza y mayor soporte, también se utilizó como referencia las 6 coincidencias entre las conclusiones encontradas por *RapidMiner*

La primera conclusión que se encontró fue si el profesor labora en otra institución dando como resultado las siguientes reglas.

1. Cuando el profesor cuenta con plaza definitiva en el 78% de los casos, el profesor no labora en otra institución, esto se evalúa con un soporte 0.789 y una confianza de 0.918.
2. Cuando el profesor cuenta con un doctorado nacional, en el 78% de los casos el profesor no labora en otra institución, con un soporte de 0.785 y una confianza de 0.922.

La segunda conclusión que arrojó la herramienta fue si el profesor cuenta con plaza definitiva, a lo cual se construyeron las siguientes reglas.

1. Cuando el profesor cuenta con doctorado nacional en el 71% de los casos el profesor si cuenta con plaza definitiva, con un soporte de 0.715 y una confianza de 0.840.
2. Cuando el profesor no labora en otra institución y cuenta con doctorado nacional, en el 65% de los casos si cuenta con plaza definitiva con un soporte de 0.653 y una confianza de 0.832.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

3. Cuando el profesor tiene bastante desarrollada el compromiso ético y social, en el 60% de los casos el profesor si cuenta con plaza definitiva, con un soporte de 0.603 y una confianza de 0.869.

La tercera conclusión que se obtuvo con las reglas de asociación fue si el profesor cuenta con doctorado nacional o internacional. Creando las siguientes reglas.

1. Cuando el profesor no labora en otra institución el 78% de los casos cuenta con un doctorado nacional. Con un soporte de 0.785 y una confianza de 0.848.
2. Cuando el profesor cuenta con plaza definitiva el 71% de los casos cuenta con un doctorado nacional, con un soporte de 0.715 y una confianza de 0.832.
3. Cuando el profesor no labora en otra institución y cuenta con plaza definitiva, el 65% de los casos el profesor tiene plaza nacional. Con un soporte de 0.653 y una confianza de 0.827.
4. Cuando el profesor es casado el 61% de los casos cuenta con doctorado nacional. Con un soporte de 0.616 y una confianza de 0.882.

Como cuarta conclusión que la herramienta arrojó se encuentra si el profesor tiene desarrollado el compromiso ético y social, en este resultado solamente se crearon dos reglas, las cuales son.

1. Cuando el profesor tiene bastante desarrollada la búsqueda de información en bases de datos especializadas, en el 42% de los casos el profesor tiene bastante desarrollado el compromiso ético y social. Con un soporte de 0.426 y una confianza de 0.837.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

2. Cuando el profesor tiene bastante desarrollada la habilidad en el uso del sistema de referencias para dar crédito a las fuentes bibliográficas consultadas, en el 41% de los casos el profesor tiene bastante desarrollado el compromiso ético y social. Con un soporte de 0.413 y una confianza de 0.840.

Como quinta conclusión que la herramienta arrojó se encuentra si el profesor se siente satisfecho con su vida familiar, en este resultado se crearon 11 reglas, de las cuales las de mejor soporte están entre el 56 y 48% de los casos.

1. Cuando el profesor siempre se siente satisfecho con su vida personal en el 56% de los casos siempre se encuentra satisfecho con su vida familiar. Con un soporte de 0.566 y una confianza de 0.919.
2. Cuando el profesor siempre se siente satisfecho con su vida personal en el 50% de los casos no labora en otra institución y siempre se encuentra satisfecho con su vida familiar. Con un soporte de 0.508 y una confianza de 0.826.
3. Cuando el profesor no labora en otra institución siempre se siente satisfecho con su vida personal en el 50% de los casos siempre se encuentra satisfecho con su vida familiar. Con un soporte de 0.508 y una confianza de 0.911.
4. Cuando el profesor cuenta con plaza definitiva y siempre se siente satisfecho con su vida personal en el 48% de los casos el profesor se

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

encuentra siempre satisfecho con su vida familiar, con un soporte de 0.483 y una confianza de 0.911.

Como última conclusión que se obtuvo fue de si el profesor se encuentra satisfecho con su vida personal, de la cual se crearon las siguientes reglas.

1. Cuando el profesor no labora en otra institución y siempre se encuentra satisfecho con su vida familiar, en el 50% de los casos el profesor se encuentra siempre satisfecho con su vida personal, con un soporte de 0.508 y una confianza de 0.879.
2. Cuando el profesor cuenta con doctorado nacional y siempre se encuentra satisfecho con su vida familiar, en el 48% de los casos el profesor se encuentra satisfecho con su vida personal, con un soporte de 0.483 una confianza de 0.867.
3. Si el profesor cuenta con plaza definitiva y siempre se encuentra satisfecho con su vida familiar, en el 48% de los casos el profesor siempre estará satisfecho con su vida personal, con un soporte de 0.483 y una confianza de 0.866.

El soporte representa el porcentaje de los casos en el que se aplica cada una de las reglas de asociación y la confianza representa el porcentaje de las veces en el que se aplica cada uno de los casos es cierto.

Como resultado de la creación de reglas de asociación se obtuvieron muchas reglas similares, se observa que un comportamiento en específico es que si el profesor se encuentra satisfecho con su vida personal entonces igual se encuentra satisfecho con su vida familiar, que la mayoría de los profesores obtuvieron un doctorado nacional el cual está muy relacionado a que cuenten con plaza definitiva.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

De las variables utilizadas en esta investigación la que se encuentra más frecuente dentro de las reglas creadas es la vida personal y familiar, habilidades para la investigación y condiciones institucionales.

1. Si el profesor tiene plaza definitiva y no tiene puesto administrativo, entonces no labora en otra institución y su doctorado es nacional.
2. Si el profesor no tiene problemas familiares por su trabajo, entonces siempre está satisfecho con su vida familiar.
3. Si el profesor no labora en otra institución, cuenta con plaza definitiva, y compromiso ético y social, el resultado será que cuenta con un doctorado nacional
4. Si el profesor no labora en otra institución, y es hombre, entonces si cuenta con plaza definitiva.
5. Si el profesor siempre se encuentra satisfecho con su vida personal, entonces el profesor no labora en otra institución y se encuentra satisfecho con su vida familiar.
6. Si el profesor no labora en otra institución y cuenta con plaza definitiva, entonces el doctorado es nacional.
7. Si el profesor cuenta con plaza definitiva, y el compromiso ético y social lo tiene bastante desarrollado, entonces cuenta con doctorado nacional.
8. Si el profesor cuenta con plaza definitiva, y se siente satisfecho con su vida personal, entonces tiene un doctorado nacional.
9. Si el profesor cuenta con doctorado nacional, y siempre se encuentra satisfecho con su vida familiar, entonces cuenta con plaza definitiva.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

10. Si el profesor no labora en otra institución y tiene doctorado nacional entonces, si cuenta con plaza definitiva.
11. Si el profesor cuenta con plaza definitiva, entonces tiene doctorado nacional.
12. Si el profesor no labora en otra institución, cuenta con doctorado nacional y es casado, entonces si cuenta con plaza definitiva.

Como resultado de la creación de reglas de asociación se obtuvieron muchas reglas similares, se observa que un comportamiento en específico es que si el profesor se encuentra satisfecho con su vida personal entonces igual se encuentra satisfecho con su vida familiar, que la mayoría de los profesores obtuvieron un doctorado nacional el cual está muy relacionado a que cuenten con plaza definitiva.

De las variables utilizadas en esta investigación la que se encuentra más frecuente dentro de las reglas creadas es la vida personal y familiar, habilidades para la investigación y condiciones institucionales.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Capítulo 5. Conclusiones, recomendaciones y trabajos futuros

En esta investigación se identificaron patrones de conducta de profesores investigadores, relacionados con las variables de trayectoria académica, afrontamiento, vida personal y familiar y habilidades para la investigación, empleando diferentes técnicas de minería de datos, dichas técnicas fueron aplicadas con el proceso KDD y la herramienta de *RapidMiner*.

Se utilizó este proceso ya que es una metodología que contiene una serie de pasos, que permite aplicar la minería de datos y utilizar los algoritmos de árboles de decisión, agrupamiento y reglas de asociación, los cuales se utilizan para el descubrimiento de información y patrones.

La evaluación de los métodos utilizados se realizó por medio de *RapidMiner*, obteniendo un 0.95 de precisión en el nivel de confianza respectivamente, para esto se descartaron patrones irrelevantes, en la representación gráfica en forma de árbol, donde a partir de un nodo raíz como lo fue, si es el profesor puede manejar el estrés en su trabajo, se empezaron a crear ramas cada una con una cantidad de nodos diferentes, que permiten comprender los problemas que enfrenta el profesor.

Se realizó un árbol de decisión para cada variable para identificar patrones de conducta dentro de cada una de las variables utilizadas en esta investigación, por consecuencia se identificaron patrones de conducta similares dentro de cada uno de los árboles creados, permitiendo identificar correlaciones.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

De acuerdo con los resultados de clasificación se presentan las estancias efectuadas correctamente, para la variable de trayectoria académica es de un 87.60%, condiciones institucionales 39.82%, vida personal y familiar 50.77%, afrontamiento 51.30% y por último la variable de habilidades para la investigación con un porcentaje del 79.32 %, siendo la trayectoria académica y las habilidades para la investigación las mejor valoradas por *RapidMiner*.

Se implementó el algoritmo de K-means como método de agrupamiento, en el cual se obtuvieron dos grupos que el algoritmo agrupó según sus similitudes y la distancia entre los centroides, por lo que las instancias clasificadas de las variables utilizadas en esta investigación comparten similitudes y relaciones esto, aunque en el primer clúster se encuentran el 65% de los datos agrupados y en el segundo clúster el 35%.

Por último, con la técnica de reglas de asociación se encontraron que dentro del *dataset* existen 100 reglas asociadas con las cinco variables de investigación, se identificó que entre los atributos que contiene el *dataset*, existen hasta cuatro números de conjuntos, esto representa el tamaño de la regla de asociación, para determinar las mejores reglas se optó por utilizar una confianza del 85% y un soporte del 95%.

Por consecuente con el análisis de las respuestas obtenidas por parte de los profesores investigadores, y la aplicación de minería de datos, y para dar respuesta a la pregunta de investigación ¿Cómo identificar cuáles son los patrones de conducta que afectan a los profesores investigadores, al realizar sus investigaciones?

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Y de acuerdo a los resultados obtenidos, con la aplicación de los algoritmos de minería de datos, se pudieron identificar patrones de conducta, reglas de asociación y agrupaciones, asociadas a las variables de trayectoria académica, condiciones institucionales, afrontamiento, vida personal y familiar y habilidades para la investigación.

Como parte de la variable trayectoria académica la mayoría de los profesores tiene un doctorado nacional lo que les permite contar con una plaza definitiva dentro de las diferentes divisiones con las que cuenta la institución, esto no quiere decir que solo por el hecho de que su doctorado sea nacional el profesor tendrá asegurada una plaza definitiva, pero si permite conocer que el profesor cuenta con experiencia laboral la cual indica que puede desarrollarse como profesor investigador.

En la variable de vida personal y familiar, los patrones encontrados en los profesores indica que la mayoría de los profesores se encuentran satisfechos con su vida personal y familiar, en el ámbito familiar el profesor no tiene problemas familiares al momento que realizan su trabajo tanto como para realizar investigaciones, así como para sus actividades laborales cotidianas, de igual forma el profesor se encuentra muy satisfecho con su vida personal aunque la mayoría de los encuentros fuera del ámbito profesional son ocasionales.

En las condiciones institucionales se demuestra que existen algunas inconformidades de los profesores en cuando el apoyo por parte la institución, pero no representan un impedimento para realizar sus investigaciones.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Con respecto a la variable de afrontamiento y de cómo se encuentra el profesorado en salud, el análisis demuestra que el profesor suele ejercitarse en ocasiones y cuenta con buena salud en la mayoría de los casos, lo que le permite tener un manejo de presión y estrés laboral, con lo que podría realizar sus investigaciones sin problemas.

Por último, la variable de habilidades para la investigación permitió identificar como el profesor investigador se encuentra preparado para realizar investigaciones, en las diferentes preguntas que se realizaron de acuerdo a dicha variable, se pudo observar que la mayoría de los profesores cuentan con diferentes habilidades muy desarrolladas así como la experiencia necesaria para realizar sus investigaciones, con lo que se puede concluir que con respecto a esta variable no representan impedimento para realizar investigaciones.

5.1 Recomendaciones

Como parte del desarrollo de esta investigación y en el proceso de aprendizaje del uso de la herramienta y la aplicación de minería de datos, se fueron implementando métodos que ayudaron a realizar esta investigación. En este apartado se mencionan un conjunto de buenas prácticas y recomendaciones.

Recomendaciones referentes a la metodología

Existen diferentes metodologías para aplicar la minería de datos tales como CRISP-DM o la metodología SEMMA que están más enfocadas al análisis en negocios, en esta investigación se utilizó la metodología KDD porque era la que más se adecuaba a la investigación, esta metodología permite desde la preparación de los datos, hasta la

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

aplicación de la minería de datos. Algunas de las recomendaciones para este proceso son:

- Es indispensable cumplir con cada una de las etapas del proceso KDD, ya que permite el debido análisis de los datos.
- Para aplicar el proceso KDD la información ya tiene que estar recolectada dentro de un dataset o base de datos.

Recomendaciones referentes a *RapidMiner*

La principal recomendación referente a la herramienta de minería de datos RapidMiner es en referencia a la información ya que el uso de minería de datos puede facilitar el trabajo de analizar las conductas, comportamientos y patrones, pero para el ser humano es muy tedioso realizar esta tarea con grandes volúmenes de datos, también hay que tomar en cuenta, que, si no se tiene definido que es lo que se quiere encontrar, esto arrojará datos sin relevancia, lo cual sería un desperdicio del uso de la información.

Es por eso que, con el desarrollo de esta investigación, se implementaron diferentes prácticas que facilitaron a llegar a los resultados obtenidos, algunas de esas recomendaciones son:

- La herramienta de *RapidMiner* tiene una versión para estudiantes que permite usar la herramienta gratuitamente.
- RapidMiner tiene una función de auto modelo la cual facilita la construcción de los modelos, aunque solo implementa predicción, *clústeres* y *outliers*.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

- Utilizar los operadores que ofrece *RapidMiner*, arrojan resultados más específicos.

Recomendaciones referentes a los resultados

Los resultados que se obtuvieron pertenecen únicamente a profesores investigadores de la UJAT, la población de estudio con la cual se realizó esta investigación, permitió conocer la perspectiva de los profesores, y los patrones de conducta que tienen al desarrollar investigaciones, aunque solo se usaron cinco variables, esto no impide que puedan existir estudios que permitan extender o realizar otra investigación utilizando otro tipo de variable. Algunas recomendaciones son:

- Utilizar algún identificador en los datos, ya que al ser una encuesta en escalas de Likert las respuestas son similares y al ser analizado el *dataset* no se distingue a que atributo pertenece.
- Analizar la información que se va utilizar en la investigación, algunos datos pueden ser redundantes.
- El uso del filtrado de datos en *Excel* permite identificar datos vacíos o nulos, de forma eficaz.

Con respecto a los hallazgos en esta investigación, se recomienda al profesor que en las habilidades para la investigación que no tengan tan desarrolladas, intente fortalecerlas para que al momento de formarse como investigador tenga más facilidades para realizar diferentes tipos de investigaciones.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

También se recomienda a la UJAT poner más atención a sus profesores ya que están inconformes con el apoyo que la institución les está brindando y esto a largo plazo puede llegar a ser un impedimento para realizar investigaciones otras recomendaciones hacia la institución son:

- Apoyar a los investigadores con los recursos necesarios para realizar sus investigaciones.
- Otórgale el reconocimiento adecuado a cada profesor por sus investigaciones realizadas.
- Incentivar las actividades de vinculación que tienen los profesores con otras instituciones.

5.2 Trabajos futuros

Durante el desarrollo de este trabajo han surgido algunas líneas de investigación futuras las cuales se han dejado abiertas para que en un futuro puedan ser abordadas y puedan extender esta investigación o dar inicio a una nueva investigación.

6 A continuación, se presentan algunos trabajos futuros que pueden desarrollarse como resultado de esta investigación

- Analizar e implementar minería de datos con el *dataset* utilizando diferentes algoritmos.
- Utilizar otras herramientas de minería de datos como lo son WEKA y MOA, por mencionar algunas, para comparar si existe diferencias u otros resultados a los obtenidos en esta investigación.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

- Utilizar esta investigación para replicar el estudio en otra universidad, recopilando sus propios datos con la ayuda del cuestionario que se encuentra dentro de los anexos generando un propio *dataset* y así poder comparando los resultados con esta investigación
- Utilizar más variables que ayuden a identificar barreras que afecten el desarrollo de los investigadores.
- Derivado de esta investigación es posible utilizar algoritmos que permitan conocer la producción científica que ha tenido cada profesor y si esto impacta en su desarrollo como investigador.

Universidad Juárez Autónoma de Tabasco.
México.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Referencias

Acosta, M. (2016). Conceptos de calidad de vida laboral en el ámbito docente universitario. *Revista Salud Bosque*, 5(2), 89-100.

<https://revistasaludbosque.unbosque.edu.co/article/view/1469>

2 Aguilar, N., Magaña, D. E., y Guzmán C. (2015). Satisfacción laboral en profesores investigadores universitarios. *Revista Internacional Administración & Finanzas*, 6(8), 11-28. <http://www.theibfr2.com/RePEc/ibf/riafin/riaf-v8n6-2015/RIAF-V8N6-2015-2.pdf>

Arias, F. (2006). El proyecto de investigación: Introducción a la metodología científica. Quinta Edición Editorial. Episteme.

54 Arriola, O., Tecuatl, G. y González, G. (2011). Software propietario vs software libre: una evaluación de sistemas integrales para la automatización de bibliotecas. *Investigación bibliotecológica*, 25(54), 37-70. <http://rev-ib.unam.mx/ib/index.php/ib/article/view/27480>

49 Bain, M. (2009). Comentarios breves sobre la GNU General Public License v3. *Revista de internet, derecho y política*, 8, 14-24.

<https://www.raco.cat/index.php/IDP/article/view/220771>

Belmonte, O. (2004). *Introducción al lenguaje de programación Java. Una guía básica*. https://www.academia.edu/38874914/Introducci%C3%B3n_al_lenguaje_de_programaci%C3%B3n_Java

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Beltrán, M. (2016). *Prácticas y condiciones institucionales para el desarrollo de la docencia*. (1ra ed.). IISUE, Instituto de Investigaciones sobre la Universidad y la Educación.

37 Bifet, A., Holmes, G., Pfahringer, B., Read, J. Kranen, P., Kremer, H., Jansen, T. y Seid, T. (2010). MOA: A Real-Time Analytics Open Source Framework *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. 617-620. https://link.springer.com/chapter/10.1007/978-3-642-23808-6_41

Brownlee, J. (10 de diciembre de 2020). *How to Run Your First Classifier in Weka*. <https://machinelearningmastery.com/how-to-run-your-first-classifier-in-weka/>

Camacho, P., Zapata, A., Menéndez, V. y Canto, P. (2018). Análisis del desempeño del profesorado universitario en el uso de MOODLE a través de técnicas de minería de datos: propuestas de necesidades formativas. *RED. Revista de Educación a Distancia*, 18(58). 1-41. <https://revistas.um.es/red/article/view/351411>

5 Cámara de Diputados del Honorable Congreso de la Unión. (2011, 5 de julio). Reglamento de la *Ley Federal de Protección de Datos Personales en Posesión de los Particulares*. *Diario Oficial de la Federación*. <http://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf>

2 Cárdenas, M., Méndez, L. M. y González, M. T. (2013). Evaluación del desempeño docente, estrés y burnout en profesores universitarios. *Actualidades Investigativas en educación*, 14(1), 1-22. <https://revistas.ucr.ac.cr/index.php/aie/article/view/13210>

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Cardona, D. y Agudelo, H. (2007). Satisfacción Personal como Componente de la Calidad de Vida de los Adultos de Medellín. *Revista de salud pública*, 9 (4), 541-549.

<https://saludpublica.mx/index.php/spm/issue/archive>

36 Consejo Nacional de Ciencia y Tecnología. (2020). *Reglamento del Sistema Nacional de Investigadores*. [https://conacyt.mx/sistema-nacional-de-investigadores/archivo-](https://conacyt.mx/sistema-nacional-de-investigadores/archivo-historico/reglamentos-antteriores/)

[historico/reglamentos-antteriores/](https://conacyt.mx/sistema-nacional-de-investigadores/archivo-historico/reglamentos-antteriores/)

2 Corredor, M. y Monroy, J. (2009). Descripción y comparación de patrones de conducta, estrés laboral y burnout en personal sanitario. *Revista Hacia la Promoción de la Salud*, 14(1), 109-123.

<https://revistasojs.ucaldas.edu.co/index.php/hacialapromociondelasalud/article/view/2051>

42 Chamba, S. F. (2015). *Minería de Datos para segmentación de clientes en la empresa tecnológica Master PC*. [Trabajo de titulación, Universidad Nacional de Loja].

<https://dspace.unl.edu.ec/jspui/bitstream/123456789/10462/1/Chamba%20Jimenez%2C%20Sairy%20Fernanda.pdf>

22 Choque, V. (2019). *Minería de datos aplicada a la identificación de factores de deserción universitaria en programas de pre grado* [tesis de maestría, Universidad Nacional de San Agustín de Arequipa escuela de posgrado]. Repositorio Institucional.

<http://repositorio.unsa.edu.pe/handle/UNSA/11015>

13 Damodaran, A. (13 de enero 2014). Uno de los mayores errores es asumir que el crecimiento de una compañía es gratis o muy barato / Entrevistado por Javier

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

García, Sintetia. https://www.sintetia.com/aswath-damodaran-stern-finance/?utm_content=buffera2ae1&utm_medium=social&utm_source=twitter

Delgado, L. R. (10 de agosto de 2017). *Patrones de comportamiento*. <https://www.milenio.com/opinion/luis-rey-delgado-garcia/para-reflexionar/patrones-de-comportamiento>

18

ESIC Business & Marketing School. (2018). *Minería de datos: qué es, como es el proceso y a qué áreas se puede aplicar*. <https://www.esic.edu/rethink/tecnologia/mineria-datos-proceso-areas-se-puede-aplica>

[Espigares, M., García, R. y Quiñones, C. \(2009\). Minería de datos educativos en plataformas de teletransformación en Boza, A., Méndez, J.M., Monescillo, M., Toscano, M., Aguaded, M.C., Ávila, J.A., Tello, J. y Salas, M. \(Coord.\), *Educación, investigación y desarrollo social. Acta del XIV Congreso de Modelos de Investigación Educativa*. \(1179-1187\). \[https://www.uv.es/aidipe/congresos/XIV_Congreso.pdf\]\(https://www.uv.es/aidipe/congresos/XIV_Congreso.pdf\)](#)

7

Esteras, J., Chorot, P., y Sandín, B. (2014). Predicción del burnout en los docentes: papel de los factores organizacionales, personales y sociodemográficos. *Revista de Psicopatología y Psicología Clínica*, 19(2), 79-92. <https://doi.org/10.5944/rppc.vol.19.num.2.2014.13059>

2

Fernández, V. (2014). *Evaluación de la salud laboral docente: estudio psicométrico del cuestionario de salud docente*. (Tesis doctoral, Facultad de Psicología Universidad Ramon Llull).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

<https://www.tdx.cat/bitstream/handle/10803/283976/TESIS%20%2014-7-14.pdf?sequence=1&isAllowed=y>

Fernández, A. (2012). *Python 3 al descubierto*. RC Libros.
https://www.academia.edu/34249596/Python_3_al_descubierto?pls=RHCVB DY4D

6 Figuroa, M. (2016). *Análisis e interpretación de los datos*. Saber metodología.
<https://sabermetodologia.wordpress.com/2016/03/06/analisis-interpretacion-datos/>

Free Software Foundation, Inc. (2007, 19 de noviembre). *Licencia Pública General GNU Affero*. <https://www.gnu.org/licenses/agpl-3.0.html>

Fresneda, M., Muñoz, J., Mendoza, E. y Carballo, G. (2012). La práctica basada en la evidencia en la logopedia española: actitudes, usos y barreras. *Revista de Investigación Educativa*, 30(1), 29-52. <https://doi.org/10.6018/rie.30.1.111551>

2 Gómez, M. A. Camargo, A., García, M. F. Estrada, J. L. Calderón, A. y Saucillo, B. (2016, 30 de abril). *Estilo de vida y problemas de salud de los docentes universitarios*. <https://www.revista-portalesmedicos.com/revista-medica/estilo-vida-problemas-salud-docentes-universitarios/>

7 Guerrero, E. (2002). Modos de afrontamiento de estrés laboral en una muestra de docentes universitarios. *Revista Interuniversitaria de Formación del Profesorado*, (43) ,93-112. <https://recyt.fecyt.es/index.php/RIFOP/issue/archive>

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Guevara, H., y Belelli, S. (2013). Las trayectorias académicas dimensiones personales de una trayectoria estudiantil. Testimonio de un actor. *Revista De Ciencias Sociales Y Humanas*, 4(4), 45-56.

<http://www.ojs.unsj.edu.ar/index.php/reviise/article/view/40>

50 [Gutiérrez, R. Díaz, K. y Román, R. \(2015\). El concepto de familia en México: una revisión desde la mirada antropológica y demográfica. *Ciencia ergo-sum*, 23 \(3\), 219-228.](#)

53 <http://ri.uaemex.mx/handle/20.500.11799/76903https://dialnet.unirioja.es/servlet/articulo?codigo=5713921>

Hernández, J. A. (2015). *Modelo de minería de datos para identificación de patrones que influyen en el aprovechamiento académico*. (Tesis maestría, Instituto tecnológico Tecnológico de la Paz).

<http://posgrado.lapaz.tecnm.mx/uploads/archivos/TesisHdzCedano.pdf>

26 [Hinestroza, D. \(2018\). El Machine Learning a través de los tiempos, y los aportes a la humanidad. <https://repository.unilibre.edu.co/handle/10901/17289>](#)

44 [Hodgson, G. \(2011\). ¿Qué son las instituciones? *Revista CS*, \(8\), 17 – 53. <https://doi.org/10.18046/recs.i8.1128>](#)

25 [Ionos. \(2018\). *Software de data mining: realiza análisis de datos más efectivos*](#)

<https://www.ionos.mx/digitalguide/online-marketing/analisis-web/software-de-data-mining-las-mejores-herramientas/>

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Izquierdo, A., Bravo, L., Ceruto, T. y Martin, D. (2015). Nuevos plugins para la herramienta Knime para el uso de sus flujos de trabajo desde otras aplicaciones. *Ciencias de la información*, 46(1), 47-52. <http://cinfo.idict.cu/index.php/cinfo/article/view/673>

30 Jaramillo, S., Augusto, S. y Fernández, A. (2015). Minería de datos sobre streams de redes sociales, una herramienta al servicio de la Bibliotecología. *Información, Cultura y Sociedad*, (33), 63-74. <https://doi.org/10.34096/ics.i33.1182>

Jiawei, H., Kamber, M. y Pei. J. (2012). Data Mining concepts and techniques. El servier. <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>

6 Lemos, M., Calle, G., Roldán, T., Valencia, M., Orejuela, J. J. y Calderón, J. P. (2019). Factores psicosociales asociados al estrés en profesores universitarios colombianos. *Diversitas: Perspectivas en Psicología*, 15(1), 59-70. <https://doi.org/10.15332/s1794-9998.2019.0001.05>

51 Lozano, I. (2016). Las trayectorias formativas de los formadores de docentes. *Actualidades Investigativas en Educación*, 16(1), 1-25. <https://revistas.ucr.ac.cr/index.php/aie/article/view/22671>

Luan, J. (2002). Data mining and its applications in higher Education. *New Directions for Institutional*, (113), 17-36. <https://doi.org/10.1002/ir.35>

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Magaña, D. E y Sánchez, P. A. (2008). Síndrome del Desgaste Emocional en Investigadores Mexicanos. *Revista Interamericana de Psicología/Interamerican Journal of Psychology*, 42(2), 5-62.

<https://journal.sipsych.org/index.php/IJP/issue/archive>

20 Manzano, R. y García, H. A. (2016). Sobre los criterios de inclusión y exclusión. Más allá de la publicación. *Revista Chilena de Pediatría*, 87, 511-512.

<https://www.sciencedirect.com/science/article/pii/S0370410616300511>

Mariño, J. D. (2014). RapidMiner Studio 5.3 aplicado a la econometría y la investigación de mercados. [Trabajo de curso, Universidad Nacional de Colombia].

<http://www.fce.unal.edu.co/unidad-de-informatica/proyectos-de-estudio/economia/1167-rapidminer-studio-5-3-aplicado-a-la-econometria-y-la-investigacion-de-mercados.html>

12 Martínez, D. L., Mirtha, M. K. y Pinto, N. (2015). Perfiles de rendimiento académico: Un modelo basado en minería de datos. *Campus Virtuales*, 4(1), 12–30.

<http://uajournals.com/ojs/index.php/campusvirtuales/article/view/66/65>

19 Martínez, D. y Márquez D. (2014). Las habilidades investigativas como eje transversal de la formación para la investigación. *Tendencias Pedagógicas*, 24, 347-360.

<https://revistas.uam.es/tendenciaspedagogicas/article/view/2110>

43 Martínez, F. y Hernández, J.P. (2017). Técnicas de minería de datos con software libre para la detección de factores asociados al rendimiento. *Revistas de Estudios y*

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Experiencias en Educación, 2(2), 135-145.

<http://www.rexe.cl/ojournal/index.php/rexe/article/view/555>

Moreno, M.G. (2005). Potenciar la educación. Un currículum transversal de formación para la investigación. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 3 (1), 520-540. <https://revistas.uam.es/reice/issue/archive>

40 Navas, F. (2016), *Introducción a la minería de datos con weka: aplicación a un problema económico*. [Trabajo de fin de grado, Universidad de Jaén].

<http://tauja.ujaen.es/bitstream/10953.1/6984/1/TFG%20-%20navas%20moreno%2C%20%20Francisco.pdf>

Núñez, M. (2007), *Las variables: estructura y función en la hipótesis*. *Investigación Educativa*. 11(20), 163-179. <https://revistasinvestigacion.unmsm.edu.pe/index.php/educa/article/view/478>
5

8 Núñez, P. S. (2016). *Utilización de técnicas de minería de datos para la identificación de rasgos de comportamiento en procesos d aprendizaje colaborativo en modelos de E-learning y B-learning*. (Tesis doctoral, universidad nacional Santiago antunezAntúnez de mayolo).

12 Ortega, D. (2016). Barreras para la utilización de la investigación. Estudio descriptivo en profesionales de Enfermería en un hospital del sudoeste de Madrid. *Enfermería Global*, 15(3), 261-288. <https://doi.org/10.6018/eglobal.15.3.215331>

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

[Parra, M. J. \(2019\). *Servicios de minería de datos en Cloud Computing*. \(Tesis doctoral, Universidad de Granada\). <https://digibug.ugr.es/handle/10481/56592>](#)

34 [Peña, A. \(2014\). Review: Educational data mining: A survey and a data mining based analysis of recent works. *Expert Systems with Applications*, 41\(4\) ,1432-1462. <https://doi.org/10.1016/j.eswa.2013.08.042>](#)

[Pérez, J. \(2016\). Trayectoria académica e historia institucional. Análisis de los casos de dos profesores investigadores de la Universidad Juárez Autónoma de Tabasco. *Perspectivas docentes*, 30, 21-30. <https://doi.org/10.19136/pd.a0n30.1291>](#)

[Ramón, P., Silva, M. P., García, V., y Estay, J. \(2020\). Afrontamiento, vida personal y familiar de profesores que pertenecen al Sistema Nacional de Investigadores \(SNI\) de una universidad del sureste de México. *Revista de psicología educativa / Journal of Educational Psychology*, 8, 1-14. <http://revistas.usil.edu.pe/index.php/pyr/article/view/492>](#)

27 [Reyes, M., y Hernández, E. \(2014\). Productividad y condiciones para la investigación: el caso de los profesores de lenguas extranjeras. *Sinética revista electrónica de educación*, 42, 1-17. <https://sinectica.iteso.mx/index.php/SINECTICA/article/view/72>](#)

41 [Riascos, J. y Molina, J. E. \(2016\). Breves consideraciones acerca de la importancia de los árboles de decisión en el análisis de carteras. *Revista de la facultad de ciencias económicas y administrativas*, 16 \(1\), 11-33. <https://revistas.udenar.edu.co/index.php/rtend/article/view/2604>](#)

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Rico, A. y Sánchez, D. (2017). Diseño de un modelo para automatizar la predicción del rendimiento académico en estudiantes de IPN. *Revista iberoamericana para la investigación y el desarrollo educativo*, 8 (16), 246-266.

<https://www.ride.org.mx/index.php/RIDE/issue/view/16>

Rodríguez, G. (2008). El software libre y sus implicaciones jurídicas. *Revista de derecho*, (30), 164-199.

<http://rcientificas.uninorte.edu.co/index.php/derecho/article/view/2825>

Rodríguez, Y. y Díaz, A. (2009). Herramientas de Minería de Datos. *Revista Cubana de Ciencias Informáticas*, 3(3-4), 73-80.

<https://rcci.uci.cu/?journal=rcci&page=issue&op=view&path%5B%5D=9>

Rueda, M. (2016). *Prácticas y condiciones institucionales para el desarrollo de la docencia*. (1th Ed.). IISUE

educación. <http://www.iisue.unam.mx/publicaciones/libros/practicas-y-condiciones-institucionales-para-el-desarrollo-de-la-docencia>

Rueda, M., Canales, A., Leyva, Y. y Luna, E. (2014). Condiciones contextuales para el desarrollo de la práctica docente. *Revista Iberoamericana de Evaluación Educativa*, 7(2), 171-183.

<https://revistas.uam.es/riee/article/view/3123>

Ruiz, E. M. y Romero, C. P. (2018). Resultados obtenidos en un proceso de minería de datos aplicado a una base de datos que contiene información bibliográfica referida a cuatro segmentos de la ciencia. *Journal of Information Systems and Technology*

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Management, 15, 1-11.

<http://www.iistem.tecsi.org/index.php/iistem/article/view/2540/680>

Ruiz, E. M. y Romero, C. P. (2018). Resultados obtenidos en un proceso de minería de datos aplicado a una base de datos que contiene información bibliográfica referida a cuatro segmentos de la ciencia. *Journal of Information Systems and Technology*

Management, 15, 1-11.

<http://www.iistem.tecsi.org/index.php/iistem/article/view/2540/680>

Sáenz, A, Cortés, F. y Betancourt, J.R. (2017). Reglas de asociación en una base de datos del área médica. *Revista de arquitectura e ingeniería*, 11(2), 1-8.

<https://dialnet.unirioja.es/servlet/articulo?codigo=6451125>

Sánchez, M. y Maldonado, L. (2003). Estrés en docentes universitarios. Caso LUZ, URBE Y UNICA. *Revista de ciencias sociales*, 9(2), 323-335.

<https://produccioncientificaluz.org/index.php/rcs/article/view/25218>

Sandoval, M., Magaña, D. y Surdez, E. (2013). Clima organizacional en profesores investigadores de una institución de educación superior. *Actualidades Investigativas en educación*,

Investigativas en educación, 13(3), 1-24.

<https://revistas.ucr.ac.cr/index.php/aie/article/view/12035>

Sanjuán, P. y Ávila, M. (2016). Afrontamiento y motivación como predictores del bienestar subjetivo y psicológico. *Revista de Psicopatología y Psicología Clínica*, 21(1), 1-

10. <https://doi.org/10.5944/rppc.vol.21.num.1.2016.15401>

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

[Santizo, C. \(2016\). Condiciones institucionales del trabajo colaborativo como estrategia de cambio del sistema educativo. *Perfiles Educativos*, 38\(153\), 154-167. \[http://www.perfileseducativos.unam.mx/iisue_pe/index.php/perfiles/article/view/57642\]\(http://www.perfileseducativos.unam.mx/iisue_pe/index.php/perfiles/article/view/57642\)](#)

16 [Silva, M. P., García, V. y Ramón, P. \(2020\). La Teoría del Capital Humano y su incidencia en la Educación. Un análisis desde la perspectiva mexicana. *Revista de Psicología y Ciencias del Comportamiento de la Unidad Académica de Ciencias Jurídicas y Sociales*, 11 \(2\), 214-225. <https://doi.org/10.29059/rpcc.20201215-125>](#)

16 [Silva, M. P., García, V. y Valdés, A. A. \(2020\). Factores que influyen en la productividad científica en la universidad pública. La perspectiva de los académicos. En B. I. Sánchez y R. Hinojosa \(coord.\) *Trazas de la investigación educativa en la experiencia de sus quijotes reflexiones y aportes*. \(pp. 263-274\). Red de Investigadores Educativos Chihuahua A.C.](#)

[Soriano, J. \(2002\). Reflexiones sobre el concepto de afrontamiento en psicología. *Boletín de Psicología*, \(75\), 73-85. <https://www.uv.es/seoane/boletin/previos/N75.htm>](#)

[Tardif, M. \(2014\). Trabajo docente pedagogía y enseñanza en *Los saberes del lado docente y su desarrollo profesional*. Vozes en Petrópolis RJ, Brasil. \(82-110\). <https://flomige.files.wordpress.com/2019/04/los-saberes-del-docente-y-su-desarrollo-profesional.pdf>](#)

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

[Timarán, S.R., Hernández, I. Caceido, S.J., Hidalgo, A. y Alvarado, J.C. \(2016\). El proceso del descubrimiento de conocimiento en base de datos. *En descubrimiento de patrones de desempeño académico* \(pp. 63-86\). Bogotá: Ediciones Universidad Cooperativa de Colombia. <http://dx.doi.org/10.16925/9789587600490>](#)

47 [Vieira, L.; Ortiz, L. y Ramírez, S. \(2009\). Introducción a la minería de datos. Rio de Janeiro: E-Paper, 2009.](#)

[Villamar, D., Juárez, A., González, I. G. y Osnaya, M. \(2019\). Factores psicosociales y síndrome de Burnout en académicos de una universidad pública de México. *Revista de Psicología Educativa / Journal of Educational Psychology*, 7\(3\), 111-140. <http://dx.doi.org/10.20511/pyr2019.v7n3.360>](#)

[Villarruel, M., Chávez, R., Hernández, I., Naranjo, F., Salazar, J., Roque, E. y Tejeda, R. \(2018\). Estrés y desgaste profesional en maestros de educación superior tecnológica en Veracruz, México. *REXE: Revista de estudios y experiencias en educación*, 3\(2\), 1-15. <http://www.rexe.cl/ojournal/index.php/rexe/article/view/568>](#)

55 [Zuniga, S. y Pizarro, V. \(2018\). Mediciones de estrés laboral en docentes de un colegio público regional chileno. *Información tecnológica*, 29 \(1\), 171-180. <http://dx.doi.org/10.4067/S0718-07642018000100171>](#)

[Obando, J. A. \(2015\). *Las escalas Likert y el análisis de correspondencias*. <https://www.ucc.edu.co/prensa/2015/Paginas/escalas-Likert-y-el-analisis-de-correspondencias.aspx>](#)

46

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Glosario

A

AGPL: Affero General Public License

C

CONACYT: Consejo Nacional de Ciencia y Tecnología

G

GPL: General Public License

GNU: Licencia Pública General

I

IES: Instituciones de Educación Superior

J

JAVA: Just Another Vague Acronym

K

KNIME: Konstanz Information Miner

KDD: Knowledge Discovery in Database

M

MECC: Modelo de Evaluación de la Condición Contextual

MOODLE: Modular Object Oriented Dynamic Learning Environment

MOA: Massive Online Analysis

S

SNI: Sistema Nacional de Investigadores

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

P

PROMEP: Programa de Mejoramiento al Profesorado

U

UNESCO: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura

UADY: Universidad Autónoma de Yucatán

W

WEKA: Waikato Environment for Knowledge Analysis

Y

YALE: Yet Another Learning Environment

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Apéndice A. Cuestionario aplicado a profesores investigadores

1. Número _____

Cuestionario de Formación de Investigadores

Instrucciones: El propósito del siguiente cuestionario es conocer las características de su perfil académico relacionadas con su desempeño como investigador. Se preguntarán datos tanto personales como institucionales, su participación es voluntaria y anónima. Usted decide si desea contestar, pero no le tomará más de 20 minutos.

I. Datos Personales

2. Edad: _____

3. Sexo: Masculino() Femenino ()

4. Estado civil: Soltero () Casado () Otro (Indique): _____

5. Número de hijos: _____

II. Datos Laborales

6. División Académica: _____

7. Tipo de nombramiento: _____

8. Tiene plaza definitiva de tiempo completo: Si () No ()

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

9. Tiempo en años con la plaza definitiva: _____

10. ¿Cuántos años ha laborado en la institución?: _____

11. ¿Trabaja en otra institución? Si () No ()

12. En caso afirmativo, cuántas horas a la semana: _____

13. ¿Ha ocupado puestos administrativos? Si () No ()

14. ¿Por cuánto tiempo en total (en años)? _____

15. Año inicio puesto administrativo: _____

16. Año fin puesto administrativo: _____

17. Cantidad de horas semanales dedicadas a la docencia en el último semestre de licenciatura

18. Cantidad de horas semanales dedicadas la docencia en el último semestre en posgrado: _____

FACTOR 1 Trayectoria académica

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

19. Nombre del Doctorado: _____

20. Año en que obtuvo el grado: _____

21. Institución en que obtuvo el grado: _____

22. País: _____

23. Número de años en que hizo el Doctorado _____

24. ¿Obtuvo el grado en un programa adscrito al PNPC?: Sí () No ()

25. ¿Contó con alguna beca para sus estudios de doctorado? Si ____ No ____

¿Qué tipo de beca contó para sus estudios de doctorado?:

26. Conacyt: Si () No ()

27. Prodep: Si () No ()

28. Otra: Si () No ()

29. ¿Cuál (Nombre)? _____

30. Número de artículos publicados al concluir al Doctorado: _____

31. Su Director de Tesis contaba con el Reconocimiento del Sistema Nacional de Investigadores: Si () No () No lo sé ()

32. Si su respuesta fue afirmativa, indique el nivel del SNI de su Director de Tesis:

Candidato () I () II () III () Emérito ()

33. Usted pertenece al SNI: Sí () No ()

34. ¿En qué año ingresó? _____

35. Nivel SNI: Candidato () I () II () III () Emérito ()

36. ¿A qué área del SNI pertenece (I-VII)? _____

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

37. ¿Cuántos intentos realizó para de obtener su reconocimiento del SNI?

() No lo he intentado 1 () 2 () Más de 2 ()

38. ¿En caso negativo, por qué no ha intentado ingresar al SNI?

39. Pertenece a un Cuerpo Académico: Sí () No ()

40. En caso afirmativo en qué nivel se encuentra:

En Formación _____ En Consolidación _____ Consolidado _____

41. Pertenece al Núcleo Académico de un posgrado en el PNPC?: Sí () No ()

Productividad: Últimos tres años.

42. Número de artículos publicados en revistas indizadas: _____

43. Número de artículos publicados en revistas no indizadas: _____

44. Capítulos de libro (editorial externa a su institución): _____

45. Número de libros publicados (editorial externa a su institución): _____

46. Número de patentes: _____

47. Número de proyectos con financiamiento interno (UJAT):

Responsable _____ Colaborador _____

48. Número de proyectos con financiamiento externo:

Responsable _____ Colaborador _____

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

FACTOR 2 Condiciones institucionales: Indique en qué medida está de acuerdo o en desacuerdo con las afirmaciones siguientes relativas a su universidad (marque con una X la respuesta que refleje su opinión).

Totalmente en desacuerdo (1), Bastante en desacuerdo (2), Algo en desacuerdo (3), Algo de acuerdo (4), Bastante de acuerdo (5), Totalmente de acuerdo (6)	1	2	3	4	5	6
1. Se reconoce el trabajo que realizan los investigadores						
2. Se facilita el establecimiento de redes con otros investigadores						
3. Se apoya con recursos a los investigadores						
4. Existen reglamentos para regular la carga docente de los investigadores						
5. Las áreas administrativas apoyan efectivamente el trabajo de los investigadores						

FACTOR 3 Vida personal y familiar: Indique en qué medida está de acuerdo o en desacuerdo con las afirmaciones siguientes relativas a su universidad (marque con una X la respuesta que refleje su opinión).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Nunca (1), Casi Nunca (2), A veces (3), Casi siempre (4), Siempre (5)	1	2	3	4	5
1. Ha tenido problemas familiares por su trabajo					
2. Su familia comprende las demandas de su trabajo					
3. Comparte actividades de recreación con su familia					
4. Asiste a las actividades académicas de sus hijos (si los tiene)					
5. Asiste a eventos sociales no relacionados con la universidad					
6. Apoya en las labores del hogar					
7. Sale de vacaciones con su familia					
8. Convive con compañeros de trabajos fuera del horario laboral					
9. Convive con sus amigos en actividades recreativas					
10. Se siente satisfecho con su vida personal					
11. Se siente satisfecho con su vida familiar					

FACTOR 4 Afrontamiento: Indique en qué medida está de acuerdo o en desacuerdo con las afirmaciones siguientes relativas a su manejo del estrés (marque con una X la respuesta que refleje su opinión).

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Nunca (1), Casi Nunca (2), A veces (3), Casi siempre (4), Siempre (5)	1	2	3	4	5
1. Mantiene una salud buena					
2. Puede manejar el estrés de su trabajo					
3. Puede manejar las presiones de tiempo de su trabajo					
4. Maneja efectivamente las demandas de nuevas habilidades que implica su trabajo (Uso de las nuevas tecnologías)					
5. Se ejercita de manera constante					

III. Autovaloración de sus condiciones como investigador: Por favor conteste las siguientes preguntas de manera honesta.

¿Qué actividades de su trabajo le parecen más valiosas? (Asigne un valor en orden de importancia, siendo el **1 la más importante y el 5 la menos importante**).

- 6. Docencia ()
- 7. Investigación ()
- 8. Gestión ()
- 9. Formación ()
- 10. Divulgación ()

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

11. ¿Cuáles son las metas permanentes más importantes que desea alcanzar en su vida profesional?

12. ¿Cuáles cree usted que son los aspectos que le facilitan desempeñarse como investigador?

13. ¿Qué dificultades encuentra usted para desempeñarse como investigador?

FACTOR 5 Habilidades para la investigación

1 A continuación, se le ofrecen una serie de habilidades y se le pide que valore el desarrollo que posee de las mismas. Debe marcar con una X la opción que mejor refleje su opinión al respecto.

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Habilidad	1	2	3	4	5	6
1. Planificar el tiempo						
2. Administrar el tiempo						
3. Comunicarse de manera escrita						
4. Comprensión de textos en un segundo idioma						
5. Comunicarse de manera oral en un segundo idioma						
6. Comunicarse de manera escrita en un segundo idioma						
7. Uso de las Tecnologías de la Información y Comunicación						
8. Crítica y autocrítica						
9. Actuar de manera creativa						
10. Actuar en nuevas situaciones						
11. Tomar decisiones						
12. Trabajar en contextos multidisciplinares						
13. Trabajar en contextos internacionales						
14. Compromiso ético y social						
15. Buscar información en bases de datos especializadas						
16. Desarrollar marcos teóricos de referencia						
17. Utilizar un sistema de referencias para dar crédito a las fuentes bibliográficas consultadas						
18. Elaborar fichas documentales y de trabajo						

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Habilidad	1	2	3	4	5	6
19. Conocimientos acerca de los paradigmas de investigación						
20. Identificación de problemas y/o necesidades de investigación o desarrollo tecnológico						
21. Formular problemas de investigación						
22. Redactar preguntas y objetivos de investigación						
23. Elaborar hipótesis de investigación						
24. Utilizar diseños experimentales						
25. Utilizar diseños no experimentales						
26. Diseñar instrumentos adecuados para la recolección de datos.						
27. Validar instrumentos adecuados para la recolección de datos.						
28. Analizar datos cuantitativos						
29. Analizar datos cualitativos						
30. Utilizar software para el análisis de datos.						
31. Desarrollar prototipos de productos						
32. Elaborar informes técnicos						
33. Divulgar resultados en medios científicos (revistas, congresos, consejos técnicos entre otros)						

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Habilidad	1	2	3	4	5	6
34. Divulgar resultados al público en general (foros empresariales, instituciones y usuarios y revistas de divulgación científica)						
35. Conocimientos de las formas de patentar						
36. Conocimientos de los fondos públicos de apoyo a la investigación y desarrollo tecnológico						
37. Conocimientos de los fondos privados de apoyo a la investigación y desarrollo tecnológico						
38. Elaborar proyectos donde se gestionen fondos públicos						
39. Elaborar proyectos donde se gestionen fondos privados						
40. Desarrollar prototipos de procesos						
41. Implementar a nivel comercial de prototipos de productos y/o procesos						
42. Conocimientos de las normas de propiedad intelectual						

Desea agregar algo más:

¡Muchas gracias por su participación!

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

Apéndice B. Relación de resultados de tamaños de clúster

Atribute	Cluster_0	Cluster_1
Edad	15.937	16.798
Sexo	1.658	1.524
Estado civil	1.861	1.821
División académica donde trabaja	6.734	6.048
Tipo nombramiento	5.367	4.369
Cuenta con plaza definitiva	1.158	1.107
Labora en esta institución	1.051	1.119
Tiene puesto administrativo en la UJAT	1.665	1.631
Cuántas horas de docencia imparte en la licenciatura	11.095	11.679
Cuántas horas de docencia imparte en la posgrado	3.247	4.655
Nombre del doctorado que estudio	57.747	52.048
Área del SNI	3.924	3.429
EL doctorado lo curso a nivel nacional o internacional	1.854	1.845
Es miembro del Sistema Nacional de Investigadores	1.456	1.583
1 ¿La institución promueve la investigación?	2.975	2.976
2 Se incentivan las actividades de vinculación con otras instituciones	2.513	2.631
3 Se reconoce el trabajo que realizan los investigadores	2.538	3.048
4 Se facilita el establecimiento de redes con otros investigadores	2.538	2.5
5 Se apoya con recursos a los investigadores	3.215	3.119
6 Existen reglamentos para regular la carga docente de los investigadores	2.646	2.536
7 Las áreas administrativas apoyan efectivamente el trabajo de los investigadores	3.405	3.405
1 Ha tenido problemas familiares por su trabajo	2.424	2.667
2 Su familia comprende las demandas de su trabajo	2.272	2.036
3 Comparte actividades de recreación con su familia	2.329	1.893
4 Asiste a las actividades académicas de sus hijos (si los tiene)	3.222	3.06
5 Asiste a eventos sociales no relacionadas con la universidad	2.500	2.583
6 Apoya en las labores del hogar	2.234	1.833
7 sale de vacaciones con su familia	2.228	2.298
8 Convive con compañeros de trabajo fuera del horario laboral	3.089	3.274
9 Convive con sus amigos en actividades recreativas	2.665	2.785
10 Se siente satisfecho con su vida personal	1.873	1.663
11 Se siente satisfecho con su vida familiar	1.582	1.35
1 Mantiene una salud buena	2.032	2.285
2 Puede manejar el estrés en su trabajo	1.728	1.881

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

3 Puede manejar las presiones de tiempo en su trabajo	1.709	1.905
4 Maneja efectivamente las demandas de nuevas habilidades que implica su trabajo	1.658	1.833
5 Se ejercita de manera constante	2.747	3.552
1 Planificar tiempo	2.525	2.095
2 Administrar tiempo	2.949	2.333
3 Comunicarse de manera escrita	1.994	1.819
4 Comprensión de textos en un segundo idioma	2.804	2.157
5 Comunicarse de manera oral en un segundo idioma	3.082	2.631
7 Uso de las tecnologías de la información y comunicación	2.867	2.255
8 Crítica y autocrítica	2.677	1.952
9 Actuar de manera creativa	2.304	1.702
10 Actuar en nuevas situaciones	2.430	1.69
11 Toma de decisiones	2.525	1.754
12 Trabajar en contextos multidisciplinares	2.722	1.75
13 Trabajar en contextos internacionales	2.133	1.405
14 Compromiso ético y social	1.563	1.94
15 buscar información en base de datos especializadas	2.177	1.143
16 Desarrollar marcos teóricos de referencia	2.139	1.429
17 Utilizar un sistema de referencias para dar crédito a las fuentes bibliográficas	2.019	1.205
18 Elaborar fichas documentales y de trabajo	2.481	1.94
19 Conocimientos acerca de los paradigmas de investigación	2.443	2
20 Identificar problemas y/o necesidades de investigación o desarrollo tecnológico	2.203	1.417
21 Formular problemas de investigación	2.063	1.457
22 Redactar preguntas y objetivos de investigación	1.968	1.345
23 Elaborar hipótesis de investigación	2.209	1.298
24 Utilizar diseños experimentales	2.285	1.457
25 Utilizar diseños no experimentales	2.354	2.488
26 Diseñar instrumentos adecuados para la recolección de datos	2.285	1.859
27 Validar instrumentos adecuados para la recolección de datos	2.399	1.679
28 Analizar datos (cuantitativos)	1.987	1833
29 Analizar datos (cualitativos)	2.342	1.69
30 Utilizar software para el análisis de datos	2.373	1.725
31 Desarrollar prototipos de productos	2.671	2.952
32 Elaborar informes técnicos	2.152	1.345
33 divulgar resultados en medios científicos (revistas, congresos, consejos técnicos)	2.177	1.457

Identificación de patrones de conducta, asociadas con variables de investigadores de educación superior, aplicando técnicas de minería de datos

34	Divulgar resultados al público en general	2.361	1.512
35	Conocimientos de las formas de patentar	2.639	2.957
36	Conocimientos de los fondos públicos de apoyo a la investigación y desarrollo	3.095	2.69
37	Conocimientos de los fondos privados de apoyo a la investigación y desarrollo	3.063	2.655
38	Elaborar proyectos donde se gestionen fondos públicos	3.089	2.679
39	Elaborar proyectos donde se gestionen fondos privados	3.057	2.59
40	Desarrollar prototipos de procesos	3.133	3.262
41	Implementar a nivel comercial de prototipos de productos y/o procesos	3.418	3.405
42	Conocimientos de las normas de propiedad intelectual	3.272	3.31

Universidad Juárez Autónoma de Tabasco.
México.

Angel Gustavo González Sánchez.pdf

Universidad Juárez Autónoma de Tabasco

Detalles del documento

Identificador de la entrega

trn:oid:::3117:582711755

Fecha de entrega

24 abr 2026, 1:59 p.m. GMT-6

Fecha de descarga

24 abr 2026, 4:22 p.m. GMT-6

Nombre del archivo

Angel Gustavo González Sánchez.pdf

Tamaño del archivo

3.9 MB

170 páginas

25.157 palabras

177.593 caracteres




10% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe



- ▶ Bibliografía
- ▶ Texto citado
- ▶ Coincidencias menores (menos de 15 palabras)
- ▶ Abstract

Fuentes principales

- 9%  Fuentes de Internet
- 4%  Publicaciones
- 0%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad




N.º de alertas de integridad para revisión

-  **Caracteres reemplazados**
90 caracteres sospechosos en N.º de páginas
Las letras son intercambiadas por caracteres similares de otro alfabeto.
-  **Texto oculto**
6 caracteres sospechosos en N.º de página
El texto es alterado para mezclarse con el fondo blanco del documento.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

Fuentes principales

- 9%  Fuentes de Internet
- 4%  Publicaciones
- 0%  Trabajos entregados (trabajos del estudiante)

Fuentes principales

Las fuentes con el mayor número de coincidencias dentro de la entrega. Las fuentes superpuestas no se mostrarán.

1	Internet	ciad.repositorioinstitucional.mx	1%
2	Internet	repositorio.ucv.edu.pe	<1%
3	Internet	renati.sunedu.gob.pe	<1%
4	Internet	repositorio.unsa.edu.pe	<1%
5	Internet	ri.ujat.mx	<1%
6	Internet	hdl.handle.net	<1%
7	Internet	www.coursehero.com	<1%
8	Internet	alicia.concytec.gob.pe	<1%
9	Internet	oldri.ues.edu.sv	<1%
10	Internet	revistas.usil.edu.pe	<1%
11	Internet	www.ionos.mx	<1%

12	Internet	revistas.um.es	<1%
13	Internet	docplayer.es	<1%
14	Internet	core.ac.uk	<1%
15	Internet	www.ceupe.com	<1%
16	Internet	repositorio.uarm.edu.pe	<1%
17	Publicación	Alvarez Rozas, Teresa Paola. "Patrones de comportamiento en el uso de las aulas ..."	<1%
18	Internet	unade.edu.mx	<1%
19	Internet	ess.iesalc.unesco.org	<1%
20	Internet	tesis.pucp.edu.pe	<1%
21	Internet	www.ujat.mx	<1%
22	Internet	repositorio.unjfsc.edu.pe	<1%
23	Internet	www.juntadegobierno.unam.mx	<1%
24	Internet	repositorio.ulead.edu.ec	<1%
25	Internet	repositorio.utn.edu.ec	<1%

26	Internet	hemeroteca.unad.edu.co	<1%
27	Internet	repositorioinstitucional.uson.mx	<1%
28	Internet	dgutyp.sep.gob.mx	<1%
29	Internet	dl.dropboxusercontent.com	<1%
30	Internet	files.eric.ed.gov	<1%
31	Internet	polodelconocimiento.com	<1%
32	Internet	www.archivos.ujat.mx	<1%
33	Internet	repositorio.utc.edu.ec	<1%
34	Internet	www.springerprofessional.de	<1%
35	Internet	e-archivo.uc3m.es	<1%
36	Internet	riaa.uaem.mx	<1%
37	Internet	acikbilim.yok.gov.tr	<1%
38	Internet	doczz.es	<1%
39	Internet	fdocuments.es	<1%

40	Internet	www.mdpi.com	<1%
41	Internet	ciencialatina.org	<1%
42	Internet	cybertesis.unmsm.edu.pe	<1%
43	Internet	iopscience.iop.org	<1%
44	Internet	www.scielo.org.co	<1%
45	Internet	producciocientifica.uv.es	<1%
46	Internet	repositoriodspace.unipamplona.edu.co	<1%
47	Internet	scielo.isciii.es	<1%
48	Internet	1library.co	<1%
49	Internet	accedacris.ulpgc.es	<1%
50	Internet	biblioteca.uny.edu.ve:9443	<1%
51	Internet	repositorio.so.ucr.ac.cr	<1%
52	Publicación	Quispe Mamani, Edgar. "Investigación formativa en el aprendizaje-enseñanza virt..."	<1%
53	Internet	alcance.unesum.edu.ec	<1%

54 Internet

conrado.ucf.edu.cu

<1%

55 Internet

educacionyeducadores.unisabana.edu.co

<1%

56 Internet

produccioncientifica.ucm.es

<1%