

**Sistema recomendador para la asignación de materias y carga horaria docente para
una Institución de Educación Superior del Ecuador**

Pérez Villarreal, Jessica Fernanda y Romero Vallejo, Edwin David

Vicerrectorado De Investigación, Innovación Y Transferencia De Tecnología

Centro De Posgrados

Maestría En Gestión De Sistemas de Información e Inteligencia de Negocios

Trabajo de titulación, previo a la obtención del título de Magíster en Gestión De Sistemas
de Información e Inteligencia de Negocios

Ph.D. Cárdenas Delgado, Sonia Elizabeth

11 de agosto de 2020



Document Information

Analyzed document	Rev_SC_Urkund_30jul_Tesis_JessicaPerez-DavidRomero.docx (D77326344)
Submitted	7/30/2020 6:06:00 PM
Submitted by	Cardenas Delgado Sonia Elizabeth
Submitter email	secardenas@espe.edu.ec
Similarity	0%
Analysis address	secardenas.espe@analysis.orkund.com

Sources included in the report

A handwritten signature in black ink, appearing to read "Sonia Elizabeth Cardenas Delgado", written over a horizontal dotted line.

Cárdenas Delgado, Sonia Elizabeth

DIRECTORA



VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA
CENTRO DE POSGRADOS

CERTIFICACIÓN

Certifico que el trabajo de titulación, **“Sistema recomendador para la asignación de materias y carga horaria docente para una institución de educación superior del Ecuador”** fue realizado por los señores **Pérez Villarreal, Jessica Fernanda y Romero Vallejo, Edwin David** el mismo que ha sido revisado y analizado en su totalidad, por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolqui, 11 de agosto de 2020

Cárdenas Delgado, Sonia Elizabeth

Directora

C.C.: 1713261160



**VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA**

CENTRO DE POSGRADOS

RESPONSABILIDAD DE AUTORÍA

Nosotros **Pérez Villarreal, Jessica Fernanda** y **Romero Vallejo, Edwin David**; con cédulas de ciudadanía n° 1716218415 y 1715193940, declaramos que el contenido, ideas y criterios del trabajo de titulación: **Sistema recomendador para la asignación de materias y carga horaria docente para una Institución de Educación Superior del Ecuador** es de mi/nuestra autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Sangolquí, 21 de agosto de 2020

Firma (s)

Pérez Villarreal, Jessica Fernanda

C.C.: 1716218415

Romero Vallejo, Edwin David

C.C.: 1715193940



VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

AUTORIZACIÓN DE PUBLICACIÓN

Nosotros **Pérez Villarreal, Jessica Fernanda** y **Romero Vallejo, Edwin David** autorizamos a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **Sistema recomendador para la asignación de materias y carga horaria docente para una Institución de Educación Superior del Ecuador** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de nuestra responsabilidad.

Sangolquí, 21 de agosto de 2020

Firma

Pérez Villarreal, Jessica Fernanda

C.C.: 1716218415

Romero Vallejo, Edwin David

C.C.: 1715193940

DEDICATORIA

A mis padres,

Por brindarme su confianza, apoyarme con las metas que me planteo y por ser quienes me enseñaron a concluir las sin decaer, superando todas las dificultades que se presenten en el camino.

Jessica

A mi madre,

Con mucho amor y cariño le dedico todo mi esfuerzo puesto para la realización de esta tesis.

David

AGRADECIMIENTO

A mis hermanas,

Por brindarme su apoyo durante el desarrollo de este proyecto.

A mi compañero de tesis,

Por el trabajo en equipo, el tiempo y esfuerzo dedicado para finalizar con éxito este proyecto.

A mis amigos,

Por sus consejos, ayuda y apoyo en los momentos de estrés para no decaer y continuar con este proyecto.

A la Universidad de las Fuerzas Armadas,

Por la importante enseñanza impartida para alcanzar mis metas académicas.

Jessica

A mi querida familia,

Por haber sido mi apoyo a lo largo de toda mi carrera. A ese ser especial por ser fuente de motivación e inspiración para poder superarme cada día más y así lograr mis objetivos propuestos y sobre todo por creer en mi capacidad.

A mi compañera de tesis,

Un agradecimiento por su esfuerzo y dedicación, pero sobre todo por el compromiso que tuvo para concluir con este proyecto.

A mi institución educativa,

Un agradecimiento afectuoso a la institución por permitirme formarme como profesional y así cumplir con un progreso personal.

David

ÍNDICE DE CONTENIDOS

Capítulo I.....	18
Introducción.....	18
Antecedentes	18
Contexto del problema	20
Planteamiento del problema	21
Justificación e importancia.....	22
Alcance	23
Objetivo general	24
Objetivos específicos.....	24
Formulación del problema	25
Capítulo II.....	28
Fundamento teórico.....	28
Técnicas de minería de datos.....	29
Metodología de prototipado rápido	33
Técnicas de búsquedas.....	36
Búsquedas Metaheurísticas	37
Técnicas de recogida de datos.....	39
Encuesta	39
Sistemas basados en conocimiento (SBC).....	44
Sistemas recomendadores (SR).....	45
Herramientas tecnológicas	47
Pentaho	47

RapidMiner	50
Power designer	51
Mysql	52
Apache tomcat	52
Java (JSP)	52
Apache ignite machine learning.....	53
Fundamento legal.....	55
Instituciones de Educación Superior.....	63
Planificación académica	63
Gestión académica.....	64
Asignación de materias y carga horaria docente	65
Capítulo III.....	66
Análisis de la información	66
Pasos aplicados para el proceso de minería de datos.....	66
Recogida de datos	66
Limpieza de datos	72
Integración	73
Selección de los datos.....	78
Transformación de datos.....	82
Análisis de datos	82
Evaluación de patrones	93
Presentación del conocimiento.....	99
Capítulo IV	104

Desarrollo del prototipo	104
Fase 1: Identificación de requisitos.....	104
Fase 2: Diseño del prototipo.....	104
Diagrama de casos de uso	105
Diagrama de secuencia.....	111
Configuración para el prototipo.....	113
Diseño del formulario web	113
Diseño de interfaces.....	114
Fase 3: Construcción del prototipo	119
Diseño de la base de datos	119
Fase 4: Evaluación del prototipo.....	121
Capítulo V	141
Conclusiones y recomendaciones	141
Conclusiones	141
Recomendaciones.....	144
Bibliografía	145

ÍNDICE DE TABLAS

Tabla 1 Preguntas de investigación.....	25
Tabla 2 Niveles de confianza.....	43
Tabla 3 Dimensión de la muestra	43
Tabla 4 Comparación Fabricantes de herramientas de integración de datos.....	49
Tabla 5 Listado de archivos recibidos para el análisis	67
Tabla 6 Exactitud de modelos de clasificación.....	79
Tabla 7 Porcentaje de clasificación correcta.....	94
Tabla 8 Descripción del caso de uso No. 1.....	107
Tabla 9 Descripción del caso de uso No. 2.....	107
Tabla 10 Descripción del caso de uso No. 3.....	108
Tabla 11 Descripción del caso de uso No. 4.....	108
Tabla 12 Descripción del caso de uso No. 5.....	108
Tabla 13 Descripción del caso de uso No. 6.....	109
Tabla 14 Descripción del caso de uso No. 7.....	109
Tabla 15 Descripción del caso de uso No. 8.....	109
Tabla 16 Descripción del caso de uso No. 9.....	109
Tabla 17 Descripción del caso de uso No. 10.....	110
Tabla 18 Descripción del caso de uso No. 11.....	110
Tabla 19 Descripción del caso de uso No. 12.....	110
Tabla 20 Descripción del caso de uso No. 13.....	110
Tabla 21 Descripción del caso de uso No. 14.....	111

Tabla 22 Comparación por materias – Docente Departamento Ciencias de la Computación	128
Tabla 23 Comparación por área de conocimiento – Docente Departamento de Ciencias de la Computación	129
Tabla 24 Comparación por materias – Docente Departamento de Ciencias Exactas	135
Tabla 25 Comparación por área de conocimiento – Docente Departamento de Ciencias Exactas	136
Tabla 26 Criterios evaluados en los resultados	137

ÍNDICE DE FIGURAS

Figura 1 Categorización de Variables	29
Figura 2 Metodología de prototipado rápido	34
Figura 3 Esquema de funcionamiento básico de un SR	46
Figura 4 Cuadrante de Gartner para Data Integration Tools	48
Figura 5 Pasos realizados por el algoritmo genético	53
Figura 6 Limpieza de datos con PDI	73
Figura 7 Lectura de datos con PDI	74
Figura 8 Relaciones entre fuentes de datos	75
Figura 9 Exportación de los datos finales	75
Figura 10 Proceso de preparación de los datos usando PDI	77
Figura 11 Operador Optimize Parameters	78
Figura 12 Operador Cross Validation	79
Figura 13 Operador k-NN	80
Figura 14 Resultados de la asignación de pesos a las variables	81
Figura 15 Formación académica	83
Figura 16 Tiempo de dedicación.....	84
Figura 17 Condición laboral.....	85
Figura 18 Género de docentes	86
Figura 19 Experiencia docente	87
Figura 20 Experiencia profesional	88
Figura 21 Distribución de la carga horaria	89
Figura 22 Carga horaria efectiva	90

Figura 23 Rotación de materias impartidas	91
Figura 24 Sugerencias de mejora para asignación de carga horaria	92
Figura 25 Diagrama de casos de uso	106
Figura 26 Diagrama de secuencia	112
Figura 27 Diseño interfaz de acceso al prototipo	114
Figura 28 Diseño de interfaces para registro de datos.....	115
Figura 29 Diseño de interfaz para generar evaluación de afinidad	116
Figura 30 Reporte de listados de afinidad	117
Figura 31 Reporte de datos del docente.....	118
Figura 32 Diagrama de base de datos.....	120
Figura 33 Creación de usuario en el prototipo	122
Figura 34 Registro de datos del docente del Departamento de Ciencias de la Computación	123
Figura 35 Perfil Docente Ciencias de la Computación	124
Figura 36 Evaluación de afinidad administrador departamento Ciencias de la Computación.....	125
Figura 37 Materias recomendadas docente Departamento de Ciencias de la Computación	126
Figura 38 Proceso de ejecución del algoritmo - Docente Departamento de Ciencias de la Computación	127
Figura 39 Registro de datos del docente Departamento Ciencias Exactas.....	130
Figura 40 Perfil Docente Ciencias Exactas.....	131
Figura 41 Evaluación afinidad administrador departamento Ciencias Exactas	132

Figura 42 Materias recomendadas docente Departamento Ciencias Exactas . 133

Figura 43 Corrida interna del algoritmo Departamento de Ciencias Exactas ... 134

RESUMEN

Uno de los procesos importantes en la planificación de una Institución de Educación Superior (IES), es la asignación de materias. Dicho proceso es complejo pues debe estar alineado al cumplimiento de leyes, reglamentos y políticas internas y nacionales. Además, se deberían tener en cuenta varios indicadores propios del talento humano docente, como su formación académica, experiencia, carga horaria, actividades de docencia e investigación, entre otras. En las IES del Ecuador no se han evidenciado trabajos que permitan gestionar la asignación de materias usando herramientas tecnológicas. Actualmente, el proceso se gestiona de forma manual lo que conlleva tiempo adicional, una posible asignación subjetiva y, por tanto, generación de errores involuntarios.

El presente trabajo pretende contribuir en la mejora del proceso de asignación de materias a docentes en una IES de Ecuador, así como también disminuir el tiempo que este proceso toma actualmente, a través del diseño e implementación de un prototipo de sistema recomendador para obtener sugerencias de materias a docentes basado en reglas que permitan en función del perfil académico y profesional del docente, encontrar las materias que mejor se ajusten a su perfil, y así convertirse en una herramienta de apoyo en la toma de decisiones.

Palabras clave:

- **ASIGNACIÓN DE MATERIAS**
- **INTELIGENCIA ARTIFICIAL**
- **MINERÍA DE DATOS**
- **SISTEMA RECOMENDADOR**

ABSTRACT

One of the important processes in the planning of a Higher Education Institution (IES) is the assignment of subjects. This process is complex because it must be aligned with policies of compliance with laws, regulations and internal and national. In addition, various indicators of teaching human talent must be taken into account, such as their academic training, experience, workload, teaching and research activities, among others. In the IES of Ecuador there have been no evidence of works that allow managing the assignment of technological tools subjects. Currently, the process is managed manually, which entails additional time, a possible subjective allocation and, therefore, generation of involuntary errors. The present work aims to contribute to the improvement of the process of assigning subjects to teachers in an IES in Ecuador, as well as reducing the time that this process currently takes, through the design and implementation of a prototype of recommender system to obtain suggestions of Rules-based subjects to teachers that can, depending on the academic and professional profile of the teacher, find the subjects that best suit their profile, and thus become a support tool in decision-making.

Keywords:

- **ASSIGNMENT OF SUBJECTS**
- **ARTIFICIAL INTELLIGENCE**
- **DATA MINING**
- **SYSTEM RECOMMENDER**

CAPÍTULO I

INTRODUCCIÓN

1.1 Antecedentes

Las organizaciones que son consideradas “más competitivas”, actualmente consideran que el talento humano es el capital intelectual y por ello lo incluyen en sus estados financieros. Estos enfoques modernos han determinado el inicio de una época con tendencia humanista, es decir, realizar un manejo inteligente de los recursos humanos. Dicha consideración surge pues solo el ser humano es capaz de tomar decisiones, multiplicar recursos financieros y optimizar los recursos disponibles. Además, contar con grupos humanos renovados con conocimientos, con formación académica actualizada, con experiencia y con excelentes competencias profesionales permite sostenibilidad, mejora continua y ventaja competitiva en un entorno globalizado (Carpio Solano, 2003).

El papel del capital humano en todo proceso de cambio organizacional, se enmarca en una visión centrada en analizar, prever y potencializar las cualidades y capacidades humanas para mejorar su desempeño. En este contexto, también se enmarca la gestión del talento humano docente en las Instituciones de Educación Superior (IES), ya que permite una visión integral para optimizar recursos, mejorar las dinámicas internas de las relaciones, el conocimiento y el quehacer del personal docente en sus actividades de enseñanza-aprendizaje, investigación y vinculación. Una de las actividades para la gestión del talento humano docente es la asignación de materias y carga horaria. Dicha actividad

actualmente es compleja ya que se lo hace en forma manual, lo cual no permite realizar una correcta asignación que contemple leyes, reglamentos, políticas internas y necesidades institucionales, así como incluyendo indicadores propios del talento humano docente, como formación académica, especialidades, experiencia, preferencias, actividades de docencia e investigación.

En la literatura revisada se ha evidenciado que a nivel mundial y regional las IES realizan el proceso de gestión docente basados en diferentes estrategias y técnicas. Algunos trabajos han aplicado modelos de programación binaria, otros han aplicado modelos de programación lineal de enteros y algoritmos de búsqueda. Además, existen trabajos que han incluido técnicas de inteligencia artificial y han demostrado avances significativos en este ámbito, desarrollando modelos para la asignación de carga académica (Wang, 2002) (López Bruno, 2007)(Abdelhalim & El Khayat, 2016). Adicionalmente, algunos estudios se han enfocado en estrategias para tomar en cuenta las preferencias, afinidades y disponibilidad del docente (Domenech & Lusa, 2016)(Javier Fiallos, Jorge García, 2008).

Hasta donde se ha podido conocer en Ecuador, no se han desarrollado soluciones tecnológicas que permitan gestionar la asignación de materias y carga horaria docente en forma automática (Alvear, 2017). Cada IES realiza diferentes procesos de acuerdo a sus políticas internas y en su mayoría se lo realiza de forma manual, lo cual podría generar subjetividad, malestar en el talento humano docente, errores en dicha asignación y hasta podría afectar en la calidad del

proceso académico. Además, dicha situación no garantiza un proceso riguroso, fidedigno y real, que esté alineado a los principales objetivos de la educación superior.

En este contexto, se propone realizar un estudio exploratorio-descriptivo para conocer a detalle la situación actual de la gestión del talento humano, en lo que respecta al proceso de asignación de materias y carga horaria de los docentes. También se determinarán indicadores relacionados con la afinidad docente, es decir, con la carrera, con las materias, con sus conocimientos de formación académica, especialidad y experiencia. Además, basados en los resultados obtenidos en el estudio, se pretende desarrollar el prototipo de un sistema recomendador que permita realizar la asignación de materias y carga horaria de forma automática, que incluya indicadores relacionados con la afinidad y garantizando un proceso libre de subjetividad. Dicha herramienta tecnológica será desarrollada aplicando técnicas de búsquedas metaheurísticas.

1.2 Contexto del problema

Las IES en general deben realizar la planificación académica por cada semestre o período académico. Dicho proceso es complejo, pues implica optimizar el talento humano docente disponible de acuerdo a las necesidades institucionales, a las leyes y reglamentos existentes y basado en diferentes parámetros propios de un docente como su formación académica, su experiencia, preferencias, afinidades y disponibilidad.

En Ecuador todavía se realiza el proceso de asignación de materias y carga horaria de forma manual. Se utilizan herramientas de ofimática, donde utilizando en una hoja de Excel se completa en función de la experiencia del coordinador académico de cada departamento o unidad responsable de la distribución de la carga horaria docente, de acuerdo con las políticas internas de la IES y/o de la unidad encargada de dicha actividad.

Esta forma de gestionar dicho proceso podría generar errores involuntarios, realizar una asignación subjetiva, llevar tiempo adicional del previsto, inclusive no se consideran algunos parámetros como la formación académica, preferencias, afinidad y experiencia del personal docente. Este proceso es complejo y muy importante en la planificación académica de las IES. Si el proceso se desarrolla de la misma manera, la calidad del proceso enseñanza-aprendizaje también podría tener un impacto negativo.

1.3 Planteamiento del problema

Todo proceso humano es proclive a errores y la asignación de materias y carga horaria docente, no es la excepción. Tal como se realiza actualmente dicho proceso resulta subjetivo y se podrían producir errores involuntarios, lo que afectaría en la calidad del proceso enseñanza-aprendizaje, ya que no existe la participación de los involucrados y no se consideran parámetros que hacen referencia a la afinidad como gustos, preferencias, formación y experiencia.

Otro de los problemas que se puede detectar en el proceso de asignación y que resulta de mucha importancia, es que la materia asignada no sea de su dominio, lo que conlleva a que necesite mayor tiempo de preparación antes de impartir la materia. Dicha asignación se realiza unas pocas semanas antes o inclusive unas horas antes de iniciar las actividades del nuevo período académico, entonces, su rendimiento no será el esperado. Posiblemente los estudiantes también se vean afectados debido a la calidad de conocimientos, experiencias y desafíos que reciben. Tal situación genera criterios errados sobre la calidad de la formación y experiencia de los docentes que forman dicha IES.

1.4 Justificación e importancia

La asignación de materias y carga horaria de los docentes tiene una implicación muy importante en la calidad de la enseñanza que brinda la IES, lo cual además se ve reflejado en las evaluaciones que se realizan a la IES.

En este proceso se debe tener en cuenta la normativa vigente de Ecuador sobre las IES, donde se especifica la distribución del tiempo de dedicación en horas que debe cumplir cada docente de acuerdo a su clasificación (CES, 2012).

La creación del prototipo de un sistema recomendador alineado a los principales objetivos de la educación superior, aplicando técnicas de búsquedas y basado en el análisis previo de la información referente al proceso de asignación de las materias y carga horaria docente, permitirá la asignación de docentes de acuerdo a sus capacidades, conocimientos, preferencias, experiencias y afinidades.

Además, contribuirá para que dicho proceso de asignación garantice rigurosidad, efectividad, confiabilidad, que esté libre de subjetividad y de errores humanos. Dichos resultados permitirán definir mejores métodos de evaluación al docente y al estudiante y hasta participar de mejoras en el escalafón regido por el Consejo de Educación Superior (CES, 2012).

1.5 Alcance

El desarrollo del estudio planteado inició con una revisión literaria sobre temas relacionados con desarrollo de sistemas recomendadores, técnicas de búsquedas metaheurísticas y de esta manera fueron seleccionadas las idóneas.

Se recopiló la información sobre la asignación de materias a docentes, de dos fuentes principalmente, la primera corresponde a la base de datos de la IES y la segunda hace referencia a la aplicación de encuestas a docentes. Posteriormente, se realizó un análisis a los datos aplicando minería de datos. Esto permitió conocer los parámetros de asignación de materias, la carga horaria a los docentes y los resultados obtenidos en el proceso manual que se realizan actualmente.

Se diseñó y desarrolló el prototipo del sistema recomendador siguiendo la metodología de prototipado rápido y aplicando técnicas de búsquedas metaheurísticas, algoritmos genéticos. Se ha realizado el proceso de validación y pruebas. Dichas pruebas fueron aplicadas en un caso de estudio específico, es decir en dos departamentos de la IES. Finalmente, en base a los resultados

obtenidos con la simulación de los algoritmos genéticos se realizó un análisis comparativo para recomendar su uso y aplicación en todas las unidades o departamentos de la IES y así mejorar el proceso de asignación de materias y carga horaria de los docentes.

1.6 Objetivo general

Desarrollar el prototipo de un sistema recomendador para la asignación automática de materias y carga horaria docente en una institución de educación superior, mediante la implementación de una técnica de búsqueda metaheurística.

1.7 Objetivos específicos

OE1: Realizar la revisión de la literatura relacionada con el ámbito de desarrollo de sistemas recomendadores, herramientas y técnicas usadas para este tipo de sistemas.

OE2: Recopilar información referente a la asignación de materias y carga horaria docente. Una parte será la base de datos disponible en la IES y la otra parte se obtendrá mediante la aplicación de encuestas a docentes.

OE3: Analizar la información recopilada aplicando minería de datos, para definir indicadores propios del proceso de asignación de materias y carga horaria docente.

OE4: Diseñar e implementar el prototipo de un sistema recomendador aplicando una metodología de desarrollo y una técnica de búsquedas metaheurísticas.

OE5: Realizar el proceso de pruebas y simulaciones con el prototipo propuesto y validar los resultados con un docente planificador aplicando casos seleccionados.

OE6: Realizar un análisis comparativo de los resultados obtenidos y recomendar los mejores que podrían ser implantados en la IES.

1.8 Formulación del problema

Con el objetivo de delimitar adecuadamente el presente estudio, a continuación, se plantearon preguntas de investigación, las cuales se definieron por cada objetivo específico.

Tabla 1

Preguntas de investigación

OBJETIVO ESPECÍFICO	PREGUNTAS DE INVESTIGACIÓN
OE1: Realizar la revisión de la literatura relacionada con el ámbito de desarrollo de sistemas recomendadores, herramientas y técnicas usadas para este tipo de sistemas.	<p>OE1-RQ1: ¿Qué herramientas son de utilidad para desarrollar un sistema recomendador?</p> <p>OE1-RQ2: ¿Qué técnicas de búsquedas metaheurísticas son apropiadas para desarrollar un sistema recomendador?</p>
OE2: Recopilar información referente a la asignación de materias y carga horaria docente. Una parte será la base de datos disponible en la IES y la	OE2-RQ1: ¿Cómo identificar qué información es relevante para el presente estudio?

OBJETIVO ESPECÍFICO	PREGUNTAS DE INVESTIGACIÓN
<p>otra parte se obtendrá mediante la aplicación de encuestas a docentes.</p>	<p>OE2-RQ2: ¿Qué tipo de encuesta se debe aplicar para recopilar información de los docentes?</p>
<p>OE3: Analizar la información recopilada aplicando herramientas de minería de datos, para definir indicadores propios del proceso de asignación de materias y carga horaria docente.</p>	<p>OE2-RQ3: ¿Cuál es el medio apropiado para aplicar las encuestas?</p> <p>OE3-RQ1: ¿Qué herramienta de minería de datos es la adecuada para el presente estudio?</p> <p>OE3-RQ2: ¿Cuáles técnicas de minería de datos permiten descubrir el comportamiento actual de docentes frente al proceso manual de asignación?</p>
<p>OE4: Diseñar e implementar el prototipo de un sistema recomendador aplicando una metodología de desarrollo y una técnica de búsquedas metaheurísticas.</p>	<p>OE4-RQ1: ¿Qué metodología permitirá guiar el desarrollo del prototipo de sistema recomendador?</p> <p>OE4-RQ2: ¿Qué beneficios aportará a la IES la implementación de un sistema recomendador?</p>
<p>OE5: Realizar el proceso de pruebas y simulaciones con el prototipo propuesto y validar los resultados con un docente planificador aplicando casos seleccionados.</p>	<p>OE5-RQ1: ¿Qué proceso debe seguirse para aplicar protocolo de validación y pruebas?</p> <p>OE5-RQ2: ¿Qué tipo de pruebas se deben utilizar para evaluar el</p>

OBJETIVO ESPECÍFICO	PREGUNTAS DE INVESTIGACIÓN
OE6: Realizar un análisis comparativo de los resultados obtenidos y recomendar los mejores resultados obtenidos que podrían ser implantados en la IES.	<p data-bbox="824 442 1325 527">rendimiento de una técnica de búsqueda?</p> <p data-bbox="824 576 1325 761">OE6-RQ1: ¿Cuáles serán los criterios para el análisis comparativo entre los resultados obtenidos con la técnica de búsqueda aplicada?</p> <p data-bbox="824 810 1325 895">OE6-RQ2: ¿Qué tiempo toma obtener soluciones con la técnica aplicada?</p> <p data-bbox="824 944 1325 1081">OE6-RQ3: ¿Qué técnica se recomienda aplicar para este tipo de sistemas?</p>

CAPÍTULO II

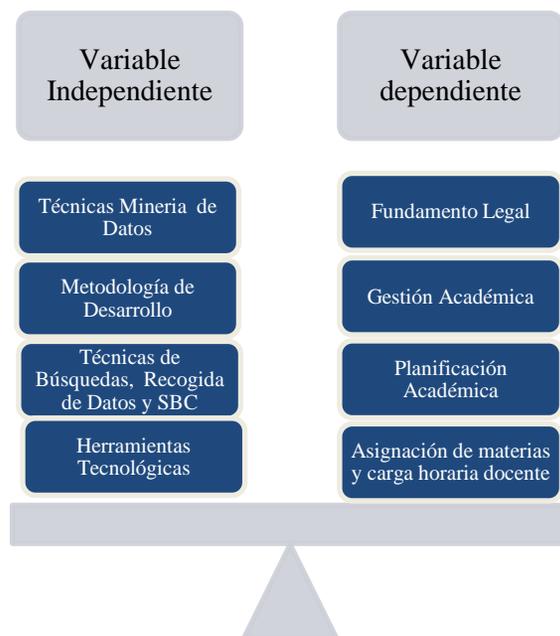
FUNDAMENTO TEÓRICO

En este capítulo se detalla el fundamento teórico que ha sido desarrollado en función a la identificación de variables dependientes e independientes relacionadas con el estudio. Se describe la metodología aplicada, técnicas de minería de datos usada en el estudio, técnica de búsquedas metaheurísticas, herramientas tecnológicas aplicadas y los criterios legales en el país y en la IES sobre asignación de carga horaria docente.

Identificación de variables

Con el propósito de estudiar cada categoría se ha establecido la siguiente categorización de variables dependientes e independientes:

- a. Variable dependiente:** Se describe el fundamento teórico en el que se basa este estudio, técnicas de minería de datos, técnicas de búsquedas y recogida de información, y la metodología para desarrollar un prototipo. Tales variables permiten efectuar un análisis de la información con el objetivo de identificar patrones de asignación de materias más adecuadas respecto al perfil de los señores docentes.
- b. Variable independiente:** Contiene el fundamento legal, la normativa vigente respecto a la asignación de materias y carga horaria docente y los procesos referentes a la gestión académica de la IES en la que se desarrolla el estudio.

Figura 1*Categorización de Variables***a. Descripción de la variable independiente**

A continuación, se describen los fundamentos respecto a la variable independiente:

2.1 Técnicas de minería de datos

Ante el crecimiento acelerado de datos en grandes volúmenes hace que la toma de decisiones sea más compleja, por lo cual el análisis de datos desde una visión predictiva, permite identificar patrones con la finalidad de predecir comportamientos futuros desde diferentes escenarios con el objetivo de transformar datos en conocimientos proactivos.

“La minería de datos comprende una serie de técnicas, algoritmos y métodos cuyo fin es la explotación de grandes volúmenes de datos con vistas al descubrimiento de información previamente desconocida y que pueda servir de ayuda en el proceso de toma de decisiones, formando parte del conjunto de tecnologías de la Inteligencia de Negocios.” (Román, García, Jesús, & Rueda, 2012).

Minería de datos o data mining (DM) es un proceso que hace uso de técnicas, tecnologías y análisis matemáticos sobre grandes conjuntos de datos, con el fin de descubrir información, patrones o reglas que puedan aportar a la toma de decisiones.

Las técnicas de minería de datos son básicamente algoritmos que se aplican sobre conjuntos de datos, las cuales tienen distintas aplicaciones y alcance. Las más representativas son:

- Redes neuronales: imitan el proceso de aprendizaje de las neuronas de organismos vivos, es decir, un conjunto de neuronas que están conectadas entre ellas, las cuales trabajan en conjunto para realizar una tarea. Las conexiones se van reforzando a través del aprendizaje que deja cada tarea realizada.
- Regresión lineal: permite conocer la relación entre variables (X, Y) de una distribución, para predecir un valor Y estimado, para un valor X que no está en la distribución.

- Árboles de decisión: permiten la representación gráfica de varios eventos que pueden suceder partiendo de una decisión.
- Clustering o segmentación: permite simplificar un conjunto de datos, a través de la agrupación de datos que tienen similares características.

Los modelos de minería de datos se pueden aplicar a escenarios específicos, como:

- Previsión: estimación de ventas, predicción de cargas del servidor o tiempo de inactividad del servidor.
- Riesgo y probabilidad: elegir los mejores clientes para envíos dirigidos, determinar el punto de equilibrio probable para escenarios de riesgo, asignar probabilidades a diagnósticos u otros resultados.
- Recomendaciones: determinar qué productos es probable que se vendan juntos, generando recomendaciones.
- Búsqueda de secuencias: análisis de selecciones de clientes en un carrito de compras, prediciendo los próximos eventos probables.
- Agrupación: separación de clientes o eventos en un grupo de elementos relacionados, análisis y predicción de afinidades.

Para lograr los objetivos de la minería de datos se debe obtener conocimiento y descubrir patrones partiendo del análisis de los datos y sus relaciones. Con el fin de llevar a cabo este proceso se necesita seguir en orden algunos pasos. A continuación, se describen dichos pasos:

1. Recogida de datos, consiste en recopilar toda la información necesaria para el estudio.
2. Limpieza de datos, se refiere a quitar la información que no es relevante y las inconsistencias que se detecte en los datos.
3. Integración, en este paso es posible integrar y combinar todos los orígenes de datos. A este paso también se lo conoce como pre-procesamiento.
4. Selección de los datos, en este paso se deben seleccionar los datos más importantes para el estudio.
5. Transformación de datos, se deben unificar los datos de forma adecuada, si es necesario se deben reducir, pero cuidando su integridad, y se deben realizar las operaciones necesarias para lograr la consolidación.
6. Análisis de la información, es un paso muy importante en el proceso, aquí se aplican métodos inteligentes como por ejemplo clasificadores, reglas, aprendizaje automático, estadística, herramientas tecnológicas, para obtener patrones de los datos que fueron procesados.
7. Evaluación de patrones, permite la identificación de los patrones más relevantes para obtener interesantes resultados que permitan representar el conocimiento, esto se realiza mediante medidas y análisis.

8. Presentación del conocimiento, los usuarios pueden comprobar los resultados obtenidos en base a la aplicación de técnicas de representación y visualización del conocimiento.

2.2 Metodología de prototipado rápido

Esta metodología se orienta al desarrollo de diferentes conceptos propuestos mediante prototipos de software o hardware, para su posterior evaluación. El desarrollo de la simulación o prototipado del sistema futuro es de gran ayuda antes de su implementación e inversión. Dicha metodología permite a los usuarios visualizar el sistema (su concepto) e informar sobre el mismo pudiéndose utilizar para aclarar opciones sobre los requerimientos de usuario y para especificar detalles de la interfaz de usuario a incluir en el sistema futuro.

El prototipado rápido se describe como un método que pretende reducir el ciclo iterativo de desarrollo. Los prototipos iterativos desarrollados podrán ser rápidamente reemplazados o modificados según los informes de diversas procedencias, como experiencias previas de usuarios o de diseñadores con experiencia, a medida que se evoluciona en el desarrollo de las tareas a realizar.

Esta metodología se centra en la interpretación, esto permite motivar al experto en el dominio del conocimiento, consiste en adquirir conocimientos y codificar hasta considerar que tenemos un modelo lo suficientemente bueno. Se deben realizar varias sesiones de trabajo con los clientes, usuarios y/o expertos.

Si el proyecto es favorable, se desarrolla un prototipo rápido para mostrar al experto, el mismo que podrá ser refinado y ampliado.

Las ventajas de esta metodología residen en la rapidez en el desarrollo de una primera versión del sistema motiva al experto (pronto se ve algo operativo). Esta metodología permite ver resultados en un corto tiempo y el usuario puede evaluar cada versión del prototipo e indicar las mejoras, las cuales serán aplicadas para las siguientes versiones del prototipo. Las evaluaciones iterativas terminan cuando el prototipo satisface las necesidades del cliente, manteniendo la idea inicial del prototipo que en futuro servirá como una plantilla para definir un sistema definitivo.

La metodología cuenta con cinco fases, las cuales se muestran en la Figura 2.

Figura 2

Metodología de prototipado rápido



Fase 1: Identificación de requisitos

En esta fase se identifican, a través de las diferentes reuniones de trabajo las necesidades de los usuarios.

Fase 2: Diseño del prototipo

Una vez obtenidos los requisitos, se elabora el diseño rápido del prototipo del sistema a través de diagramas UML (Unified Modeling Language) que sean convenientes y aporten para el desarrollo, diagramas de base de datos e interfaces de usuario.

Fase 3: Construcción del prototipo

Consiste en el desarrollo del prototipo en función de los requerimientos de los usuarios que han sido recolectados en la fase 1 y de acuerdo a lo diseñado en la fase 2.

Fase 4: Evaluación del prototipo

En esta fase, los usuarios evalúan que el prototipo desarrollado contemple lo acordado en la fase 1, en caso de identificar diferencias entre lo acordado y lo desarrollado, se regresa a la fase 1 para aclarar los requisitos, a la fase 2 para modificar el diseño y a la fase 3 para modificar el desarrollo del prototipo.

Fase 5: Entrega del prototipo

Una vez que la evaluación del prototipo cumpla con los requisitos del cliente, se realiza la entrega final del prototipo.

2.3 Técnicas de búsquedas

Las técnicas de búsqueda son una serie de esquemas de representación del conocimiento, que mediante diversos algoritmos permiten resolver ciertos problemas.

Los elementos que integran las técnicas de búsqueda son:

- Conjunto de estados: todas las configuraciones posibles en el dominio.
- Estados iniciales: estados desde los que partimos.
- Estados finales: las soluciones del problema.
- Operadores: se aplican para pasar de un estado a otro.

Las técnicas heurísticas permiten resolver un problema de optimización que puede ser más generales o específicas. Los métodos heurísticos específicos deben ser diseñados a propósito para cada problema, utilizando toda la información disponible y el análisis teórico del modelo. Las heurísticas generales presentan ventajas significativas como la sencillez, adaptabilidad y robustez de los procedimientos.

Los conceptos actuales de metaheurísticas son estrategias inteligentes para diseñar o mejorar procedimientos heurísticos muy generales con alto rendimiento. Algunos tipos de metaheurísticas son los de relajación, los procesos constructivos, búsquedas por entornos y para procedimientos evolutivos.

2.3.1 Búsquedas Metaheurísticas

En general, guían los procedimientos que usan transformaciones o movimientos para recorrer el espacio de soluciones alternativas y explorar las estructuras asociadas.

“Son una familia de algoritmos aproximados de propósito general. Suelen ser procedimientos iterativos que guían una heurística subordinada de búsqueda, combinando de forma inteligente distintos conceptos para explorar y explotar adecuadamente el espacio de búsqueda.” (Herrera, 2009).

La utilización de búsquedas metaheurísticas es un procedimiento inteligente que parte del conocimiento relacionado con una tarea, desde una perspectiva de adaptabilidad, flexibilidad sobre la optimización de los recursos utilizados para cada proceso de búsqueda.

Este tipo de metaheurística establece estrategias para recorrer el espacio de soluciones de un problema, obteniendo soluciones iterativas.

Existen distintos métodos metaheurísticos en función de conceptos como son:

- Constructivos: Parten de una solución inicial vacía y van añadiéndole componentes hasta construir una solución dada en: GRASP o Colonias de hormigas.
- Búsquedas o trayectorias: Parten de una solución inicial e iterativamente tratan de reemplazarla por otra solución con mejor calidad dada en: búsqueda local, enfriamiento simulado, búsqueda tabú, BL iterativa.

- Evolutivos: establecen estrategias para conducir la evolución en el espacio de búsqueda de conjuntos de soluciones (poblaciones) con la intención de lograr un acercamiento a la solución óptima. Se distinguen por la forma en que combinan la información proporcionada por los elementos de la población para hacerla evolucionar mediante obtención de nuevas soluciones. Los algoritmos genéticos, meméticos y los de estimación de distribuciones, entre otros.

De los métodos anteriormente mencionados, el que tiene mayor relación con la estructuración del presente estudio son los evolutivos mediante la aplicación de algoritmos genéticos, debido a la capacidad de soluciones que puede presentar para generar recomendaciones.

Algoritmos Genéticos

Los algoritmos genéticos tienen la capacidad de crear soluciones para solventar problemas que ocurren en el mundo real, la evolución tales soluciones hacia valores óptimos del problema depende en buena medida de una adecuada codificación de las mismas. Los principios básicos de los algoritmos genéticos fueron establecidos por Holland 1975 (GIL, 2006).

Este tipo de algoritmo permite obtener soluciones a problemas complejos o abstractos, debido a su complejidad no tiene una resolución estable o definida con anterioridad, al contrario, se basa en hechos empíricos para dar soluciones múltiples en un tiempo aceptable.

Las soluciones se identifican con individuos que pueden formar parte de la población de búsqueda. La codificación de una solución se interpreta como un

cromosoma del individuo compuesto de un cierto número de genes a los que les corresponden ciertos alelos. Se consideran dos operaciones básicas: la mutación y el cruce. La población evoluciona de acuerdo a las estrategias de selección de los individuos, para las operaciones y para la supervivencia.

2.4 Técnicas de recogida de datos

2.4.1 Encuesta

Es una técnica de investigación que a través del instrumento denominado cuestionario, permite la recolección de datos de modo rápido, eficaz y posteriormente a su aplicación se pueden realizar análisis con el fin de identificar características y relaciones de los datos recolectados, frente al problema que se desea resolver.

El objetivo de un cuestionario es formular preguntas concretas, cortas y de fácil interpretación, que permitan obtener respuestas fiables y que puedan ser cuantificadas.

Antes de redactar las preguntas, hay que tener en cuenta:

- Características de la población que será encuestada como por ejemplo nivel cultural, edad, estado de salud.
- Sistema de aplicación a emplear, como por ejemplo vía web, presencial.

Estos aspectos permitirán determinar la cantidad de preguntas, el lenguaje a utilizar, el formato de respuesta y otras características relevantes, que deben componer el cuestionario.

Tipos de preguntas

- Cerradas: son aquellas donde se puede elegir entre dos opciones como, por ejemplo: si/no, verdadero/falso.
- De elección múltiple: este tipo de preguntas se dividen en tres tipos:
 - o Abanico de respuestas: varias opciones de respuestas que son exhaustivas y mutuamente excluyentes.
 - o Abanico de respuestas con un ítem abierto: varias opciones de respuestas que permiten al encuestado incluir opciones distintas a las presentadas.
 - o Preguntas de estimación: varias opciones de respuestas a manera de escala, como por ejemplo las propuestas por Likert.
- Abiertas: son aquellas que permiten al encuestado que se exprese con sus propias palabras.

Es recomendable:

- Incluir instrucciones sencillas que expliquen cómo responder el cuestionario.
- Ubicar al inicio del cuestionario preguntas sencillas y que motiven al encuestado.
- Evitar preguntas ambiguas.
- Si el cuestionario tratará más de un tema, cada tema debe iniciarse con preguntas introductorias.
- Agrupar las preguntas por temas.
- Incluir ejemplos demostrativos para responder una pregunta.

Muestra

La cantidad de personas de las cuales se va obtener información relevante para el desarrollo de la presente investigación, se denomina muestra.

Es importante conocer los siguientes términos antes de calcular el tamaño de una muestra:

- Tamaño de la población: Corresponde a la cantidad total de personas que conforman un grupo de interés.
- Margen de error: Es el porcentaje que representa que tan cercano se encuentra de obtener una respuesta correcta.
- Nivel de confianza: Son intervalos que se usan para delimitar un valor con una determinada probabilidad alta.
- Desviación estándar: Es un índice numérico de la dispersión de un conjunto de datos.

Para calcular el tamaño de la muestra, existen dos fórmulas, las cuales se aplican dependiendo si se conoce o no el tamaño de la población.

Si se desconoce el tamaño de la población, la muestra se calcula de la siguiente manera:

$$n = \frac{z_a^2 * p * q}{d^2}$$

Donde:

z: es el nivel de confianza,

p: es la proporción de personas que poseen la característica de estudio, este valor es generalmente desconocido y se suele suponer que $p=q=0.5$ que es la opción más segura

q: es la proporción de individuos que no poseen la característica de estudio, es decir, es $1-p$

d: es la precisión o error máximo admisible en términos de proporción

Si se conoce el tamaño de la población, la muestra se calcula de la siguiente manera:

$$n = \frac{N * z^2 * p * q}{e^2 * (N - 1) + (z^2 * p * q)}$$

Donde:

N: es el tamaño de la población

z: es el nivel de confianza

p: es la proporción de personas que poseen la característica de estudio, este valor es generalmente desconocido y se suele suponer que $p=q=0.5$ que es la opción más segura

q: es la proporción de individuos que no poseen la característica de estudio, es decir, es $1-p$

e: es la precisión o error máximo admisible en términos de proporción.

Los valores z y sus niveles de confianza son:

Tabla 2*Niveles de confianza*

Z	Nivel de confianza
1,15	75%
1,28	80%
1,44	85%
1,65	90%
1,96	95%
2	95,5%
2,58	99%

Se puede también, considerar como referencia la Tabla 3 a continuación para determinar la dimensión de la muestra.

Tabla 3*Dimensión de la muestra*

Población	Margen de error			Nivel de confianza		
	10 %	5 %	1 %	90 %	95 %	99 %
100	50	80	99	74	80	88

Población	Margen de error			Nivel de confianza		
500	81	218	476	176	218	286
1,000	88	278	906	215	278	400
10,000	96	370	4,900	264	370	623
100,000	96	383	8,763	270	383	660
+ 1,000,000	97	384	9,513	271	384	664

2.5 Sistemas basados en conocimiento (SBC)

Es un sistema que utiliza y representa explícitamente conocimiento sobre un dominio concreto para realizar una tarea que requeriría de un experto, es decir, que es capaz de exportar el conocimiento a través de los mecanismos apropiados de razonamiento para proporcionar un comportamiento de alto nivel en la resolución de problemas en ese dominio (Castro, Souto, Alonso, & Guijarro, 2002).

El desarrollo de un SBC beneficia en varios aspectos, mayor rapidez, mayor productividad y calidad en la toma de decisiones. A partir de la Inteligencia Artificial surgen algunas ramas: los sistemas conexionistas, los sistemas expertos, la robótica, entre otras. Los sistemas expertos es un logro importante ya

que permiten enfrentar problemas reales y complejos utilizando conocimiento específico de varios dominios.

En este contexto, surge la Ingeniería del Conocimiento (IC) que no está incluida en la Ingeniería de Software (IS). En IC se hacen modelos de un ámbito concreto, se hace un análisis exhaustivo de la organización donde se aplicará el modelo. Para desarrollar los SBC es necesario adquirir el conocimiento, y debe ser continuo, no se obtienen especificaciones de requisitos.

El modelado del conocimiento se realiza en base al conocimiento adquirido y razonamiento que debe tener del sistema futuro. El conocimiento forma una base de conocimientos, los cuales pueden estar representados por reglas. Hay conocimiento de tarea (qué y cómo), inferencial el razonamiento que se aplica mediante tareas (contexto del problema) y el dominio del conocimiento y de la información relevante como si fuera un modelo de datos u objetos.

Las reglas son una forma natural y común de representar el conocimiento simbólico. Para modelar las reglas se establece una relación, donde el antecedente y el consecuente no son instancias de conceptos sino expresiones de esas instancias. Se modela una relación entre expresiones acerca de los valores de los atributos. Se modelan un conjunto de reglas de estructura similar. Las relaciones no son estrictamente lógicas.

2.5.1 Sistemas recomendadores (SR)

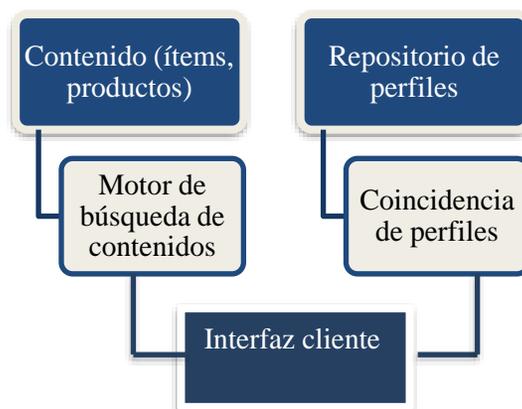
Tratan de ser un paso adelante en el contexto de la recuperación de información tradicional, que se da por palabras claves del tópico que se desea

encontrar a través de los conocidos motores de búsqueda Google, Lycos y Altavista entre otros. Como su nombre lo indica, los SR se encargan de recomendar o sugerir a los usuarios ítems o productos concretos basándose en sus preferencias (Velez-Langs & Santos, 2014).

Los SR son sistemas de filtrado de información que tienen como enfoque generar recomendaciones sobre un determinado objeto de estudio a partir de preferencias, opiniones o afinidades, ingresadas por el usuario o un sistema automático.

Figura 3

Esquema de funcionamiento básico de un SR



Hay diversas maneras de crear una tipología para los SR:

De acuerdo al método usado:

- Basados en modelos: Estos algoritmos primero desarrollan un modelo de las calificaciones de los usuarios utilizando algoritmos de aprendizaje automático como Redes Bayesianas, encasillamiento o enfoques basados en reglas.

De acuerdo a la forma de recopilar el conocimiento del usuario:

- Explícito: El usuario le hace conocer al sistema lo que piensa respecto a la información de los ítems.
- Implícito: Se registran las acciones de los usuarios y se concluyen las calificaciones sobre estas acciones.

De acuerdo a las fuentes de conocimiento del usuario:

- Usuario único
- Muchos usuarios (FC)

De acuerdo a los tipos de acción:

- Búsqueda
- Filtrado de información
- Recomendación de sitios
- Tutoría
- Recomendación de personas y grupos

2.6 Herramientas tecnológicas

Con el fin de alcanzar el objetivo propuesto en el presente estudio, se realizó un estudio sobre las herramientas de software utilizadas para procesamiento de datos o específicamente Minería de Datos.

2.6.1 Pentaho

Es una plataforma que pertenece a la empresa Hitachi Vantara, la cual consta de un conjunto de programas orientados al ámbito de inteligencia de

negocios, la cual incluye herramientas para reportes, análisis, minería, visualización e integración de datos.

Para el presente estudio se hizo uso de Pentaho Data Integration o sus siglas PDI, porque facilita efectuar tareas de limpieza e integración de datos, con el objetivo de mantener un solo repositorio formado a partir de distintas fuentes. PDI tiene una edición comunitaria, así como una edición profesional, la edición comunitaria es apropiada especialmente para soluciones de bajo presupuesto. De acuerdo al cuadrante mágico de Gartner, PDI se encuentra en el cuadrante jugadores de nicho.

Figura 4

Cuadrante de Gartner para Data Integration Tools



Gartner, a través de Peer Insights, proporciona revisiones y calificaciones de software y servicios aportados por profesionales verificados de la comunidad de TI.

Se realizó una comparación de Hitachi Vantara con el mejor fabricante del cuadrante de Líderes y cuadrante Retadores, las puntuaciones van del 1 al 5, donde 1 es la calificación más baja y 5 la más alta y los resultados se encuentran a continuación.

Tabla 4

Comparación Fabricantes de herramientas de integración de datos

Característica	Hitachi Vantara	Microsoft	Informática
Disponibilidad para recomendar	4.1	4.2	4.2
Movimiento de datos por lotes	4.5	4.5	4.6
Virtualización	3.8	3.8	3.9
Movimiento Orientado a Mensajes	3.8	3.8	3.8
Replicación y sincronización de datos	4.4	4.3	4.3
Flexibilidad de precios	4.2	4	3.7
Facilidad de implementación	4.2	4.4	4.1

Característica	Hitachi Vantara	Microsoft	Informática
Oportunidad de respuesta del vendedor	4	4.1	4.1
Calidad de Soporte Técnico	4.2	4.2	4.2
Puntaje Total	37.2	37.3	36.9

De acuerdo a la tabla anterior, se pudo evidenciar que la característica que genera una diferenciación es la flexibilidad de precios.

2.6.2 RapidMiner

Es una herramienta de minería de datos ampliamente usada y probada a nivel internacional en aplicaciones empresariales, de gobierno y academia. Implementa más de 500 técnicas de pre-procesamiento de datos, modelación predictiva y descriptiva, métodos de prueba de modelos, visualización de datos, etc. (Microsystem, 2019).

Es una herramienta que facilita el análisis predictivo de temas complejos partiendo de un análisis sistemático, permitiendo la creación de sistemas predictivos flexibles. Está disponible como una aplicación independiente para el análisis de datos, ya que es una herramienta unificada donde tiene un entorno de programación visual fácil de usar, cuenta con un formato XML basado en interfaz gráfica de usuario, facilita el autoservicio de análisis predictivo permitiendo una avanzada analítica.

Ha evolucionado hasta convertirse en una plataforma cruzada que combina minería de datos, con minería de texto, aprendizaje de máquina, inteligencia de negocios y análisis de negocios. Esta herramienta cuenta con un motor analítico patentado, que potencia los procesos de negocios, ayudando a los usuarios a sacar el mejor partido de sus datos. Con este software los usuarios pueden acelerar la adquisición de datos y el análisis para tomar acciones rápidamente.

2.6.3 Power designer

Power Designer es la herramienta líder de modelización de datos, que permite visualizar, analizar y manipular de manera más fácil los metadatos para tener una arquitectura de información de empresa eficaz. Power Designer mezcla excepcionalmente diferentes técnicas de modelización (modelo conceptual tradicional, físico y lógico con una modelización única de inteligencia de negocios y de traslado de datos) para juntar análisis de negocios con soluciones formales de diseño de base de datos. PowerDesigner funciona con más de 60 sistemas de gestión de base de datos (Novalys, 2019).

Es una herramienta que permite analizar, diseñar y construir una base de datos, además del desarrollo orientado a modelos de datos a nivel conceptual y físico. Orientado a modelos que reduce los esfuerzos de creación, mantenimiento y reingeniería manual de código, mejorando la productividad de los procesos de creación de cualquier base de datos.

2.6.4 Mysql

MySQL es el sistema para gestionar bases de datos de código abierto más conocido en el mercado. Gracias a su rendimiento, fiabilidad y facilidad de uso, se ha convertido en el más popular para las aplicaciones web, es utilizada en aplicaciones, como Facebook, Twitter y YouTube. Además, es una elección muy popular como base de datos integrada (Oracle, 2019b).

2.6.5 Apache tomcat

“Apache Tomcat funciona como un contenedor de servlets desarrollado bajo el proyecto Jakarta en la Apache Software Foundation. Tomcat implementa las especificaciones de los servlets y de JavaServer Pages de Oracle Corporation” (Apache, 2019) .

Apache es muy versátil ya que funciona en cualquier sistema operativo como un servidor web de aplicativos con soporte de servlets y JSPs.

2.6.6 Java (JSP)

Java es un lenguaje de programación y una plataforma informática comercializada por primera vez en 1995 por Sun Microsystems. Actualmente es usada en muchas aplicaciones y sitios web y cada día se crean más. Java es rápido, seguro y fiable (Oracle, 2019a). El lenguaje de programación Java posee una estructura orientada a objetos, la cual permite que sea un sistema flexible, adaptable y fiable.

2.6.7 Apache ignite machine learning

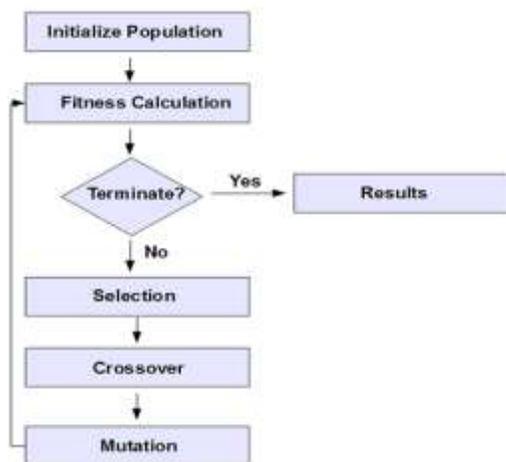
Es una plataforma, la cual es un conjunto de herramientas simples, escalables y eficientes que permiten construir modelos predictivos de aprendizaje automático sin costosas transferencias de datos.

Dentro de Apache Ignite Machine Learning, existe un componente de aprendizaje automático que contiene un conjunto de algoritmos genéticos, los cuales permiten resolver problemas de optimización mediante la simulación del proceso de evolución biológica

En el siguiente diagrama se muestran los pasos realizados por el algoritmo genético.

Figura 5

Pasos realizados por el algoritmo genético (Apache Ignite, 2019)



A continuación, se describen los pasos que se siguieron para implementar el algoritmo genético.

Crear una configuración del algoritmo genético

Es una clase que se utiliza para personalizar el comportamiento de algoritmo genético a través del objeto GAConfiguration.

Definir el gen y el cromosoma

Gen es el bloque de construcción discreto que constituye el cromosoma. En el presente proyecto se definieron como genes a las materias a recomendar, el gen más fuerte será la materia con mejores resultados.

El cromosoma es una secuencia de genes. Un cromosoma representa una solución potencial.

Implementar una función de acondicionamiento físico

El algoritmo genético es lo suficientemente inteligente como para realizar la mayoría del proceso de selección natural, sin embargo, no tiene dominio del problema, por esta razón, se necesita definir una función, para lo cual se debe extender la clase IFitnessFunction del algoritmo para calcular un puntaje de aptitud para un cromosoma potencial. Se utiliza un puntaje de aptitud para determinar qué tan óptima es la solución en relación con otras soluciones potenciales en la población.

Para que el algoritmo saque los mejores individuos se debe dar un ambiente o ecosistema en el cual cada individuo se desarrolle con sus potenciales. Es así que se tiene la función de la aptitud que consiste en generar un bucle que simula las generaciones.

Definir condición de terminación

El siguiente paso es especificar una condición de finalización adecuada para el algoritmo genético. La condición de finalización variará según el dominio del problema.

Evolucionar la población

El paso final es inicializar una instancia de GAGrid usando nuestras instancias de configuración de GAC e Ignite. Luego evolucionamos la población invocando `GAGrid.evolve ()`.

En concordancia con los recursos económicos disponibles para el desarrollo del presente estudio, se eligieron las siguientes versiones de las herramientas descritas anteriormente:

- Edición comunitaria de la herramienta Pentaho Data Integration, versión 8.2.
- Licencia educativa de la herramienta RapidMiner, versión 9.2.
- Licencia gratuita de la herramienta Power Designer versión 16.5.
- Licencia pública de la herramienta Mysql, versión 5.0.

b. Descripción de la variable dependiente

A continuación, se describen los fundamentos respecto a la variable dependiente.

2.7 Fundamento legal

En Ecuador, el sistema de educación superior es regulado por la Ley Orgánica de Educación Superior (LOES).

Uno de los organismos que rige el sistema de educación superior es el Consejo de Educación Superior (CES), el cual tiene como misión la planificación, regulación y coordinación interna del sistema de educación superior de Ecuador.

El CES mediante su “Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior” fijó normas que rigen el ingreso, promoción, estabilidad, evaluación, perfeccionamiento, escalas remunerativas, fortalecimiento institucional, jubilación, y cesación.

En dicho reglamento, entre otras cosas, se especifican las actividades y la cantidad de horas laborales del personal académico. También, se detallan los requisitos mínimos que deben cumplir los distintos tipos de personal académico para laborar en universidades y escuelas politécnicas.

En cuanto a las actividades del personal académico se destacan el siguiente artículo.

“Artículo 6.- Actividades del personal académico. - Los profesores e investigadores de las universidades y escuelas politécnicas públicas y particulares, titulares y no titulares pueden cumplir las siguientes actividades:

a) De docencia.

b) De investigación.

c) De dirección o gestión académica.

Las actividades de vinculación con la colectividad se enmarcan en lo establecido en los artículos 7, 8 Y 9 de este Reglamento.” (CES, 2012).

En el artículo a continuación se detalla la cantidad de horas laborables clasificadas por el tiempo de dedicación.

“Artículo 11.- Del tiempo de dedicación del personal académico. - Los miembros del personal académico de una universidad o escuela politécnica pública o particular, en razón del tiempo semanal de trabajo, tendrán una de las siguientes dedicaciones:

- 1. Exclusiva o tiempo completo, con cuarenta horas semanales;*
- 2. Semi exclusiva o medio tiempo, con veinte horas semanales; y*
- 3. Tiempo parcial, con menos de veinte horas semanales” (CES, 2012).*

Los requisitos que el personal académico titular auxiliar debe cumplir se especifican en el artículo que sigue.

“Artículo 30.- Requisitos del personal académico titular auxiliar de las universidades y escuelas politécnicas. - Para el ingreso como miembro del personal académico titular auxiliar de las universidades y escuelas politécnicas públicas y particulares, además de cumplir los requisitos generales establecidos en este Reglamento, se deberá acreditar:

- 1. Tener al menos grado académico de maestría o su equivalente, reconocido e inscrito por la SENESCYT, en el campo de conocimiento vinculado a sus actividades de docencia o investigación;*
- 2. Ganar el correspondiente concurso público de merecimientos y oposición; y*
- 3. Los demás que determine la institución de educación superior, entre los cuales podrá incluir requisitos de experiencia previa siempre y cuando no sobrepase el 50% del tiempo requerido para el personal académico titular principal 1 a tiempo completo en docencia y/o investigación adquirida en la misma u otra IES, conforme lo determinado en el artículo 32 de este Reglamento y lo establecido en el artículo*

150, literal d) de la Ley Orgánica de Educación Superior (LOES). Adicionalmente se deberá observar las normas constitucionales y legales, así como garantizar los derechos establecidos en el artículo 6 de la LOES” (CES, 2012).

Los requisitos del personal académico titular agregado a cumplir se detallan en el artículo a continuación.

“Artículo 31.- Requisitos del personal académico titular agregado de las universidades y escuelas politécnicas. - Para el ingreso como miembro del personal académico titular agregado de las universidades y escuelas politécnicas públicas y particulares, además de los requisitos generales establecidos en este Reglamento, se deberá acreditar:

- 1. Tener al menos grado académico de maestría o su equivalente, reconocido e inscrito por la SENESCYT, en el campo de conocimiento vinculado a sus actividades de docencia o investigación;*
- 2. Tener al menos tres años de experiencia como personal académico en instituciones de educación superior o en instituciones de investigación de prestigio;*
- 3. Haber creado o publicado al menos tres obras de relevancia o artículos indexados en el campo de conocimiento vinculado a sus actividades de docencia o investigación;*
- 4. Haber obtenido como mínimo el setenta y cinco por ciento del puntaje de la evaluación de desempeño en sus últimos dos periodos académicos;*
- 5. Haber realizado ciento ochenta horas de capacitación y actualización profesional, de las cuales noventa habrán sido en metodologías de aprendizaje e*

investigación, y el resto en el campo de conocimiento vinculado a sus actividades de docencia o investigación;

6. Haber participado al menos doce meses en uno o más proyectos de investigación;

7. Suficiencia en un idioma diferente a su lengua materna;

8. Ganar el correspondiente concurso público de merecimientos y oposición, o ser promovido a esta categoría de conformidad con las normas de este Reglamento;

y

9. Los demás que determine la institución de educación superior, que deberá observar las normas constitucionales y legales, así como garantizar los derechos establecidos en el artículo 6 de la Ley Orgánica de Educación Superior” (CES, 2012).

Los requisitos que el personal académico titular principal debe cumplir se especifican en el artículo siguiente.

“Artículo 32.- Requisitos del personal académico titular principal de las universidades y escuelas politécnicas. - Para el ingreso por concurso como personal académico titular principal de las universidades y escuelas politécnicas públicas y particulares, además de los requisitos generales establecidos en este Reglamento, se acreditará:

1. Tener grado académico de Doctor (PhD o su equivalente), en el campo de conocimiento vinculado a sus actividades de docencia e investigación, reconocido e inscrito por la SENESCYT con la leyenda de "Título de Doctor o PhD válido para el ejercicio de la docencia, investigación y gestión en educación superior". El

incumplimiento de este requisito invalidará el nombramiento otorgado como resultado del respectivo concurso;

2. Tener al menos cuatro años de experiencia en actividades de docencia y/o investigación en instituciones de educación superior o en instituciones de investigación de prestigio;

3. Haber creado o publicado doce obras de relevancia o artículos indexados en el campo de conocimiento vinculado a sus actividades de docencia o investigación, de los cuales al menos tres deberán haber sido creados o publicados durante los últimos cinco años;

4. Haber obtenido como mínimo el setenta y cinco por ciento del puntaje de la evaluación de desempeño en sus últimos dos periodos académicos;

S. Haber realizado ciento noventa y dos horas de capacitación y actualización profesional, de las cuales noventa habrán sido en metodologías de aprendizaje e investigación, y el resto en el campo de conocimiento vinculado a sus actividades de docencia o investigación;

6. Haber participado en uno o más proyectos de investigación con una duración de al menos 12 meses cada uno, por un total mínimo de seis años;

7. Haber dirigido o codirigido al menos una tesis de doctorado o tres tesis de maestría de investigación;

8. Suficiencia en un idioma diferente a su lengua materna;

9. Ganar el correspondiente concurso público de merecimientos y oposición; y

10. Los demás que determine la institución de educación superior, que deberá observar las normas constitucionales y legales, así como garantizar los derechos

establecidos en el artículo 6 de la Ley Orgánica de Educación Superior” (CES, 2012).

Los requisitos del personal académico titular principal investigador son los indicados en el siguiente artículo.

“Artículo 33. · Requisitos del personal académico titular principal investigador de las universidades y escuelas politécnicas. - Para el ingreso como miembro del personal académico titular principal investigador de las universidades y escuelas politécnicas públicas y particulares, además de los requisitos generales establecidos en este Reglamento, se acreditará:

- 1. Tener grado académico de doctorado (PhD o su equivalente), en el campo de conocimiento vinculado a sus actividades de docencia e investigación, reconocido e inscrito por la SENESCYT con la leyenda de "Título de Doctor o PhD válido para el ejercicio de la docencia, investigación y gestión en educación superior". El incumplimiento de este requisito invalidará el nombramiento otorgado como resultado del respectivo concurso;*
- 2. Tener al menos cuatro años de experiencia como personal académico en instituciones de educación superior o en instituciones de investigación de prestigio;*
- 3. Haber creado, publicado o patentado doce obras de relevancia, artículos indexados o resultados de investigación en el campo de conocimiento vinculado a sus actividades de docencia o investigación;*
- 4. Haber obtenido como mínimo el setenta y cinco por ciento del puntaje de la evaluación de desempeño en sus últimos dos periodos académicos;*

5. *Haber realizado ciento noventa y dos horas de capacitación y actualización profesional, de las cuales noventa habrán sido en metodologías de aprendizaje e investigación, y el resto en el campo de conocimiento vinculado a sus actividades de docencia o investigación;*
6. *Haber participado en uno o más proyectos de investigación con una duración de al menos 12 meses cada uno, por un total mínimo de seis años, de los cuales deberá haber dirigido o codirigido al menos dos proyectos de investigación;*
7. *Haber dirigido o codirigido al menos dos tesis de doctorado o cinco tesis de maestría de investigación;*
8. *Suficiencia en un idioma diferente a su lengua materna;*
9. *Ganar el correspondiente concurso público de merecimientos y oposición, o ser incorporado con dedicación exclusiva a las actividades de investigación de conformidad con las normas de este Reglamento; y*
10. *Los demás que determine la institución de educación superior, que deberá observar las normas constitucionales y legales, así como garantizar los derechos establecidos en el artículo 6 de la Ley Orgánica de Educación Superior” (CES, 2012).*

Dentro de las disposiciones generales de la LOES se destaca la primera, la cual manifiesta lo descrito a continuación.

“PRIMERA. - En las instituciones de educación superior públicas y particulares, los profesores e investigadores titulares deberán estar a cargo de al menos el 60% de horas de las actividades de docencia e investigación programadas en cada periodo académico” (CES, 2012).

2.8 Instituciones de Educación Superior

Las instituciones del Sistema Nacional de Educación Superior Ecuatoriano son entidades pluralistas, abiertas a todas las corrientes y formas del pensamiento universal expuestas de manera científica. Su actividad es la formación integral del ser humano para contribuir al desarrollo del país y al logro de la justicia social, en el contexto pluricultural del país, a la afirmación de la democracia, la paz, los derechos humanos, la defensa y protección del medio ambiente (LOES, 2010).

Conforme al macro sistema, la estructura de la IES en la que se realizó el estudio, es de carácter matricial, sustentada en unidades estratégicas académicas denominadas departamentos, las que cumplen con la misión de la IES en un campo específico del conocimiento, para lo cual gestionan el talento humano docente, los materiales, equipos, laboratorios y demás bienes necesarios para la docencia, investigación, innovación y transferencia de tecnología. Con base a la estructura departamental, se desarrollan carreras de tercer nivel técnico-tecnológico y de grado, con la participación del personal académico de los departamentos y los centros, conforme a la planificación curricular correspondiente. Además, los departamentos en sus áreas de conocimiento, planifican y ejecutan programas de posgrado, en los cuales la investigación es el componente fundamental (ESPE, 2019).

2.8.1 Planificación académica

La planificación académica de la institución está alineada a las normas y reglamentos del CEAACES. La institución debe mantener políticas, normativas y

procedimientos que se aplican en la aprobación, seguimiento periódico y evaluación de todas sus carreras y programas (CEAACES, 2018).

La IES en la que se realizó el estudio indica en su estatuto que, para la autoevaluación, acreditación y aseguramiento de la calidad, se cuenta con una Unidad de Autoevaluación y Aseguramiento de la Calidad Académica, que tiene la responsabilidad de planificar, organizar y ejecutar estos procesos; su organización, integración, atribuciones y deberes, constan en el Reglamento Orgánico de Gestión Organizacional por Procesos de la IES.

2.8.2 Gestión académica

Los instrumentos utilizados para la evaluación integral del profesor son transparentes y universales, propician la honestidad, objetividad y seguridad de los evaluadores y permiten identificar el cumplimiento de las obligaciones, responsabilidades y actividades de docencia, investigación y dirección o gestión académica (CEAACES, 2018).

En el caso de la IES en la que se realizó el estudio, la gestión académica comprende todas las actividades relacionadas con los procesos ordinarios de la institución como son: el gobierno y dirección de procesos de docencia e investigación, eventos académicos nacionales e internacionales, desempeño de cargos de director o coordinador, diseño de proyectos de carreras y programas de estudio de grado y posgrado.

2.8.3 Asignación de materias y carga horaria docente

La asignación de carga académica del docente es un componente clave dentro de la planificación estratégica de una IES. Esta programación debe seguir un proceso riguroso, alineado con el principal objetivo institucional, que es proveer una enseñanza de calidad, además se debe considerar la cantidad de horas de clase que dispone cada docente.

En la IES donde se realizó el estudio los docentes con dedicación a tiempo parcial deben impartir de 2 a 9 horas semanales de clases, para los docentes con dedicación a medio tiempo deben impartir 10 horas semanales de clases y los docentes con dedicación a tiempo completo deben impartir de 3 a 16 horas semanales de clases.

Actualmente, esta actividad en la IES donde se realizó el estudio se la hace en forma manual, mediante la utilización de una hoja de Excel que contiene una plantilla la cual se completa en función de la experiencia del coordinador académico de cada departamento, de acuerdo con las políticas internas de la IES.

CAPÍTULO III

ANÁLISIS DE LA INFORMACIÓN

En el presente capítulo se expone el análisis realizado a la información disponible, la cual ha sido recopilada aplicando técnicas de recogida de datos. El análisis se ha realizado mediante técnicas y herramientas de minería de datos.

3.1 Pasos aplicados para el proceso de minería de datos

3.1.1 Recogida de datos

Se realizó el proceso de recogida de datos necesarios para realizar el análisis. Se han aplicado dos técnicas, semiautomática y manual. La técnica semiautomática fue gestionada en la unidad de tecnología de la IES. La técnica manual fue ejecutada mediante la aplicación de encuestas.

a) Recogida semiautomática de datos

Se realizaron las coordinaciones pertinentes en la IES. Se recibieron siete archivos en formato Excel, correspondientes a cuatro periodos académicos detallados a continuación:

- Abril a agosto 2017 (201710)
- Octubre 2017 a febrero 2018 (201720)
- Abril a agosto 2018 (201810)
- Octubre 2018 a febrero 2019 (201811)

Cada archivo tuvo una estructura y campos distintos, dichos archivos contenían información de docentes, materias y horarios. Además, se encontró información útil para el estudio acerca de los señores docentes de la IES, tales

como formación académica, situación laboral, materias que imparten, carga horaria, departamento y área de conocimiento al que pertenecen.

En la Tabla 5 a continuación se muestra un listado que contiene los nombres de los archivos que se recibieron, el número de columnas y el número de registros de cada uno.

Tabla 5

Listado de archivos recibidos para el análisis

Nombre del archivo	Número de columnas	Número de registros
CapacitacionDocentes	12	11546
CargaHoraria201710	11	1163
CargaHoraria201720	11	1264
CargaHoraria201810	11	1266
CargaHoraria201811	11	1159
DATOSDOCENTES_15032018	57	1241
TitulosAcadDocentes	16	3598

b) Recogida manual de datos

Luego de revisar la información entregada por la IES, se evidenció que era necesario complementar la información, pues de acuerdo a los objetivos planteados, eran importantes algunas variables tales como experiencia

profesional, experiencia docente, nivel de satisfacción en cuanto a las actividades que desarrolla y la relación que existe entre las materias asignadas respecto a la formación académica y experiencia profesional. Esta técnica fue aplicada para la recogida de datos mediante encuestas, con el fin de complementar la información recibida.

En base a la información, se planificó un proceso de recopilación de datos mediante la aplicación de una encuesta para los señores docentes de la IES. El diseño de la encuesta fue realizado en conformidad con la Ley Orgánica de Protección de Datos, Art.1, el cual tiene como objetivo garantizar y proteger el tratamiento de los datos personales, las libertades públicas y los derechos fundamentales de las personas físicas y especialmente de su honor e intimidad personal y familiar (Corresponsables, 2019).

Para el presente trabajo, la encuesta se aplicó sobre una muestra que fue calculada en función de los siguientes parámetros:

- Campus: Matriz, el cual fue seleccionado por las facilidades de acceso para los tesisistas.
- Departamentos: Ciencias Exactas y Ciencias de la Computación, los cuales fueron seleccionados debido a la apertura y facilidades que brindaron los diferentes directores de departamento a los tesisistas luego de realizado un acercamiento.
- Cantidad de profesores: 139, de los cuales 64 pertenecen al departamento de Ciencias de la Computación y 75 a Ciencias Exactas, esta cantidad de profesores, fue identificada en los archivos entregados por la IES.

- Nivel de confianza: 90%, el cual calculado de la siguiente manera:

Porcentaje por departamento:

$$\text{Profesores de Ciencias de la Computación} = 64 \Rightarrow 46.04\%$$

$$\text{Profesores de Ciencias Exactas} = 75 \Rightarrow 53.95\%$$

$$n=139$$

$$IC \text{ para la proporción} = \left(\frac{46.04}{100}; \frac{53.95}{100} \right)$$

$$IC \text{ para la proporción} = (0.46; 0.54)$$

$$\text{Amplitud} = 0.54 - 0.46 = 0.08$$

$$E = \text{Amplitud} / 2 \Rightarrow E = 0.08/2 \Rightarrow E = 0.04$$

$$P = (0.46 + E) \vee P = (0.54 - E)$$

$$P = (0.46 + 0.04) \vee P = (0.54 - 0.04)$$

$$P = (0.5) \vee P = (0.5) \Rightarrow P = 0.5$$

Z_c =nivel de confianza

$$Z_c = E / \sqrt{P(1-P)/n}$$

$$Z_c = \frac{0.04}{\sqrt{\frac{0.5(1-0.5)}{139}}}$$

$$Z_c = 0.04/0.04455260166$$

$$Z_c = 0.8978 \cong 0.90$$

$$Z_c\% = 0.90 * 100\% \Rightarrow$$

$$Z_c\% = 90\%$$

- Margen de error: 10%, el cual se obtuvo como se muestra a continuación:

El margen máximo de error es una simple re-expresión del tamaño de la muestra n . Los numeradores de estas ecuaciones se han redondeado a dos decimales.

$$\text{Margen de error en confianza del 90\%} \cong 0.82/\sqrt{n}$$

$$0.82/\sqrt{(139)} = 0.82/11.7898261226 = 0.07 \cong 0.1$$

$$\text{Margen de error \%} = 0.1 * 100\%$$

$$\text{Margen de error \%} = 10\%$$

Con los parámetros detallados anteriormente se realizó el cálculo de la muestra por departamento, mediante la aplicación de la fórmula:

$$n = \frac{N * z^2 * p * q}{e^2 * (N - 1) + (z^2 * p * q)}$$

Departamento de Ciencias de la Computación:

$$n = \frac{64 * 1.65^2 * 0.5 * 0.5}{0.1^2 * (64 - 1) + (1.65^2 * 0.5 * 0.5)}$$

$$n = 33$$

Departamento de Ciencias Exactas:

$$n = \frac{75 * 1.65^2 * 0.5 * 0.5}{0.1^2 * (75 - 1) + (1.65^2 * 0.5 * 0.5)}$$

$$n = 35$$

El resultado obtenido fue una muestra de 68 docentes, pertenecientes a dos departamentos de la IES.

La encuesta diseñada constó de 16 preguntas, de las cuales cuatro preguntas fueron abiertas, siete preguntas cerradas, una pregunta de opinión.

Las preguntas abiertas estuvieron compuestas por variables cuantitativas tales como:

- Años de experiencia docente.
- Años de experiencia profesional.
- Años trabajados en ésta IES.
- Horas semanales por actividad realizada: investigación, vinculación y docencia.

Las preguntas cerradas, fueron planteadas para recoger información académica sobre los docentes, y estas fueron:

- Condición laboral
- Tipo de personal académico: titular o no titular.
- Categoría a la que pertenece
- Tiempo de dedicación
- Formación académica
- Nivel de formación académica

Cuatro preguntas cerradas de selección, aplicando la escala Likert. Dichas preguntas se formularon con el fin de conocer el grado de conformidad de los docentes con el proceso actual de asignación de materias.

- Materias asignadas
- Horario asignado
- Cantidad de horas para docencia asignadas

- Cantidad de horas para vinculación asignadas
- Cantidad de horas para investigación asignadas

Una pregunta de opinión, la cual tuvo el objetivo de recolectar sugerencias de mejora del proceso actual de asignación de materias y carga horaria que los docentes encuestados quisieran que sean consideradas.

3.1.2 Limpieza de datos

Una vez recopilada la información se procedió a preparar los datos antes de iniciar el proceso de análisis.

Este proceso se apoyó en la herramienta Pentaho Data Integration (PDI), la cual permitió la creación de un proceso automatizado, que facilitó la extracción, limpieza y unificación de los datos relevantes, además permitió realizar este proceso con corto tiempo de procesamiento.

Se encontró que en varios de los archivos existían caracteres especiales, datos nulos, tipos de datos erróneos por lo que se realizó una limpieza de datos para controlar dichos caracteres especiales (UTF8), como tildes, uso de la letra ñ, controlar nulos, filtrado, traspuesta y ordenamiento datos, como se muestra en la Figura 6.

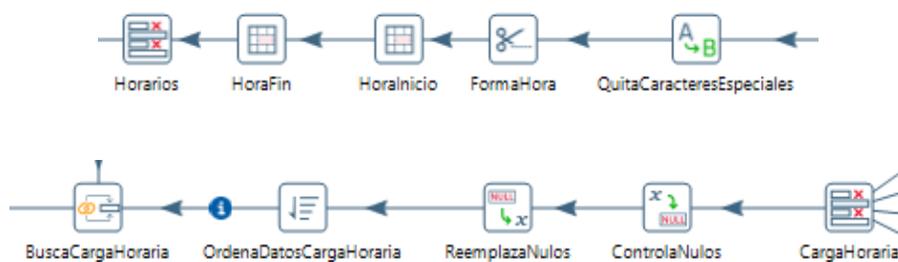
La técnica utilizada fue Data cleaning para eliminar el ruido y resolver las inconsistencias en los datos, a través de sus fases:

- Perfilamiento: para identificar dependencias, registros duplicados, valores perdidos.
- Normalización: para unificar los tipos de datos de las distintas fuentes.

- Valores perdidos: para detectar y tratar valores faltantes.
- Valores atípicos: para identificar los datos que no son consistentes con el resto.

Figura 6

Limpieza de datos con PDI



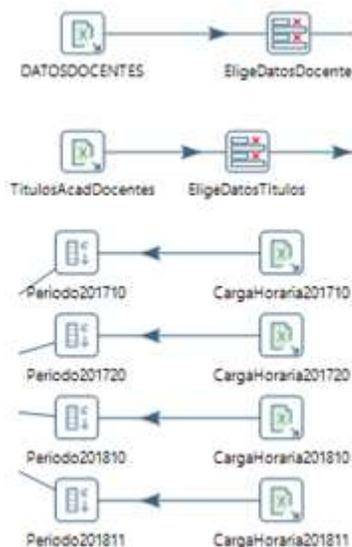
3.1.3 Integración

Se examinaron todos los registros en cada archivo recibido, la información fue subida en memoria para su integración y respectivo procesamiento.

Para combinar todas las fuentes de datos, se inició con la lectura de datos como se evidencia en la Figura 7.

Figura 7

Lectura de datos con PDI



Posteriormente se identificaron las relaciones existentes entre las distintas fuentes, con el fin de conformar una única fuente de datos, esto se puede apreciar en la Figura 8.

La técnica utilizada es la integración de registros, en la cual los registros provenientes de varias fuentes tienen su propia información. Se utilizaron operadores “join” y “union” y técnicas relacionadas con estos operadores. El enfoque “Join” completa todos los atributos faltantes usando las otras fuentes de información.

Figura 8

Relaciones entre fuentes de datos



Una vez que se disponía de datos limpios y unificados, fueron exportados a un archivo en formato Excel, el cual se ha convertido en la fuente completa como insumo para el análisis como se muestra en la Figura 9.

Figura 9

Exportación de los datos finales



El archivo generado contiene 4852 registros, los campos que se eligieron fueron los que hacían referencia a datos de docentes y todos aquellos que estaban acorde a la carga horaria docente, los cuales se detallan a continuación:

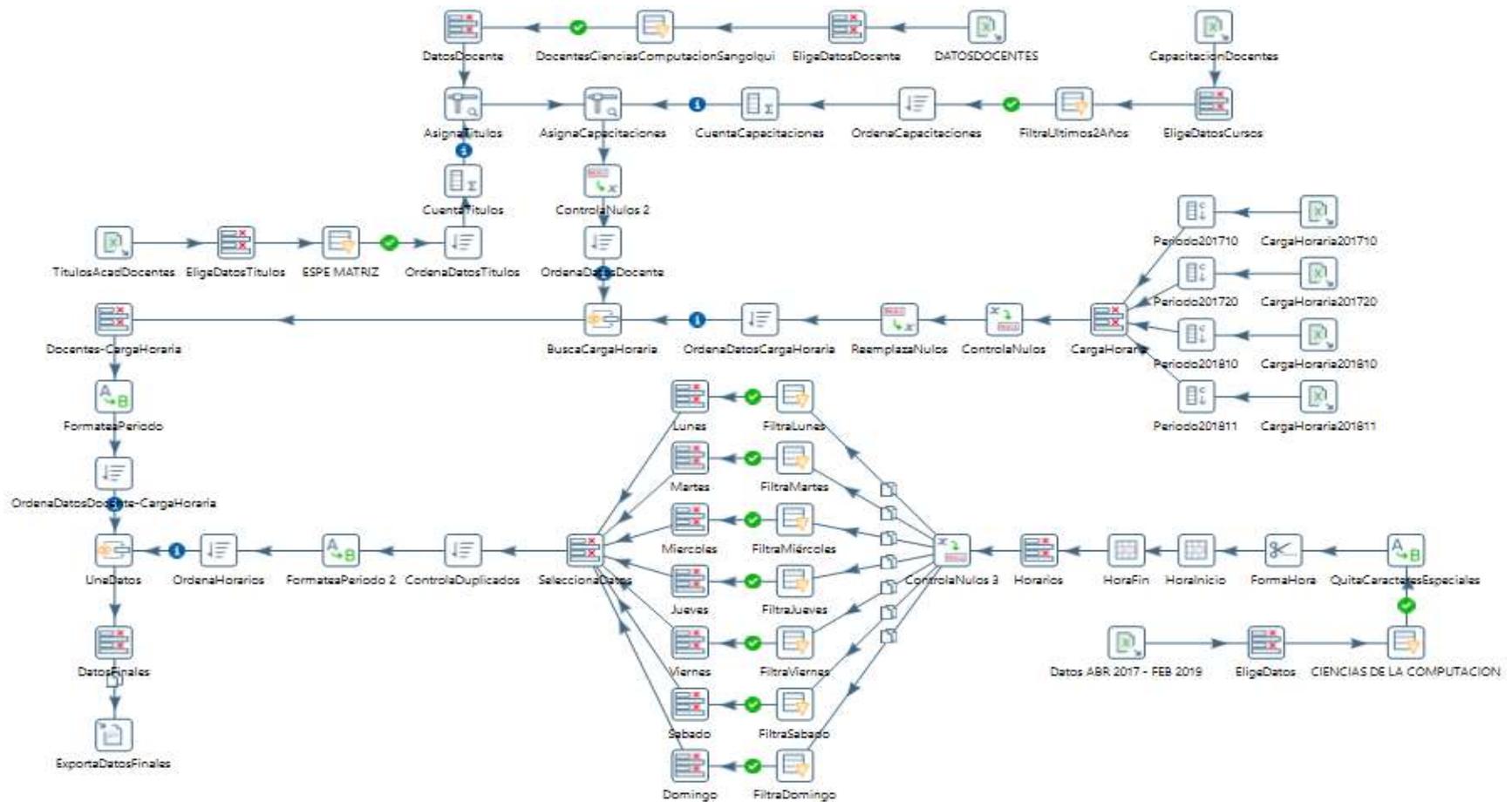
- Departamento al que pertenece el docente
- Código del docente
- Género del docente

- Tiempo de dedicación del docente
- Condición laboral
- Nivel máximo de formación académica del docente
- Período académico
- Cantidad de horas asignadas de forma semanal para gestión académica
- Cantidad de horas asignadas de forma semanal para investigación
- Cantidad de horas asignadas de forma semanal para vinculación
- Cantidad de horas asignadas de forma semanal para docencia
- Área de conocimiento de la materia que imparten
- Materia que imparte el docente
- Créditos de la materia
- Día de la semana en la que imparte la materia
- Hora de inicio y fin de la clase
- Cantidad de alumnos matriculados
- Cantidad de alumnos aprobados
- Cantidad de alumnos reprobados
- Cantidad de alumnos retirados

El proceso completo que se construyó, se visualiza en la Figura 10 a continuación.

Figura 10

Proceso de preparación de los datos usando PDI



3.1.4 Selección de los datos

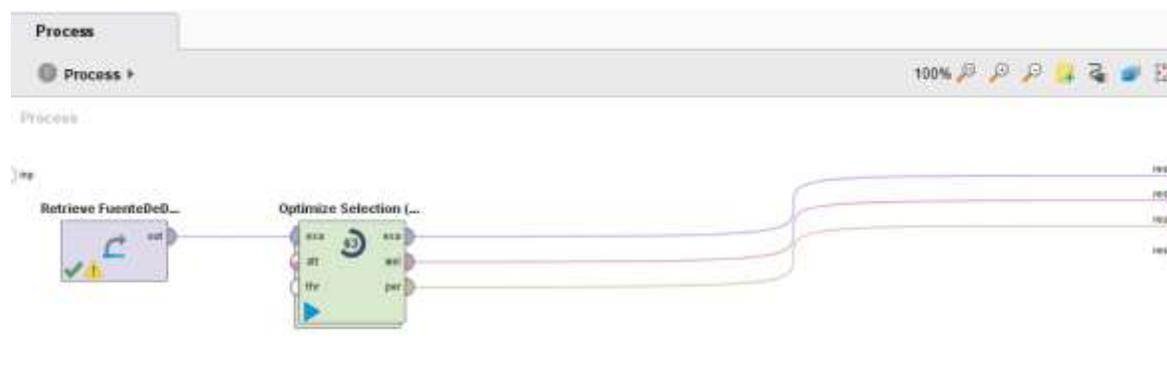
Posterior a la limpieza e integración de datos, se realizó un análisis utilizando la herramienta de minería de datos RapidMiner para conocer los datos o variables que podrían incidir en el desempeño de los docentes en el proceso actual de asignación de carga horaria, en los departamentos de la IES seleccionados para el presente trabajo.

Con el fin de identificar las variables que aporten de mejor manera a la solución del problema de este estudio, se creó un proceso automatizado para la selección mediante la aplicación de un modelo de clasificación, se eligió este tipo de modelo ya que es usado cuando el resultado que se desea es una etiqueta discreta, es decir, es útil cuando la respuesta al problema cae dentro de un conjunto finito de resultados posibles.

Para construir el proceso se inició importando la fuente de datos única, la cual contiene las variables detalladas anteriormente en los numerales 3.2.1 y 3.2.2, se tomó como variable dependiente a la carga horaria, y como variables independientes el resto de variables, luego se hizo uso del operador Optimize Parameters (Evolutionary) el cual permitió seleccionar los atributos más relevantes del conjunto de datos usando un modelo de clasificación.

Figura 11

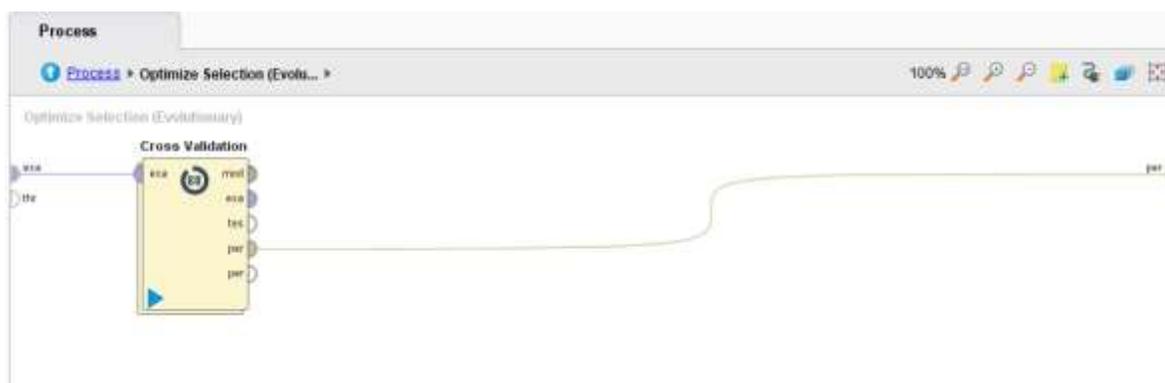
Operador Optimize Parameters



Como subproceso del operador Optimize Parameters se usó una validación cruzada para estimar el rendimiento estadístico del modelo.

Figura 12

Operador Cross Validation



Con el fin de identificar el modelo de clasificación a usar en el presente estudio, se realizaron varias corridas del proceso creado para la selección de variables, usando algunos modelos de clasificación para comparar la exactitud de sus resultados.

Tabla 6

Exactitud de modelos de clasificación

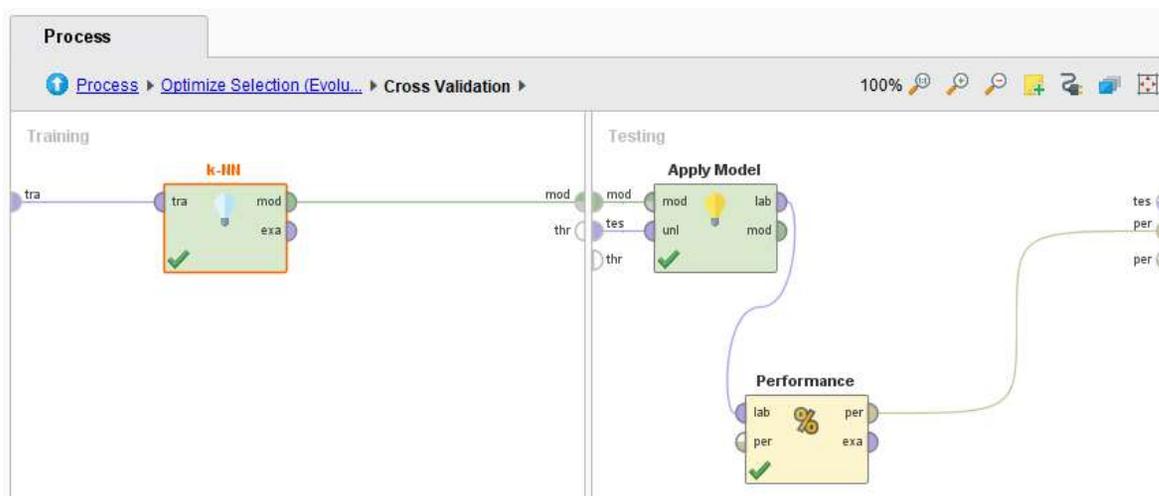
Modelo	Resultado
Random forest	<div style="display: flex; align-items: flex-start;"> <div style="border: 1px solid #ccc; padding: 5px; margin-right: 10px;"> <p>Criterion</p> <p>accuracy</p> </div> <div> <p><input checked="" type="radio"/> Table View <input type="radio"/> Plot View</p> <p>accuracy: 80.22% +/- 2.67% (micro average: 80.22%)</p> </div> </div>
k-NN	<div style="display: flex; align-items: flex-start;"> <div style="border: 1px solid #ccc; padding: 5px; margin-right: 10px;"> <p>Criterion</p> <p>accuracy</p> </div> <div> <p><input checked="" type="radio"/> Table View <input type="radio"/> Plot View</p> <p>accuracy: 97.43% +/- 0.82% (micro average: 97.43%)</p> </div> </div>

Modelo	Resultado
Árbol de decisión	<div style="display: flex; align-items: center;"> <div style="border: 1px solid gray; padding: 5px; margin-right: 10px;"> Criterion accuracy </div> <div> <input checked="" type="radio"/> Table View <input type="radio"/> Plot View </div> </div> <p style="text-align: right;">accuracy: 86.88% +/- 1.95% (micro average: 86.87%)</p>
Naive Bayes	<div style="display: flex; align-items: center;"> <div style="border: 1px solid gray; padding: 5px; margin-right: 10px;"> Criterion accuracy </div> <div> <input checked="" type="radio"/> Table View <input type="radio"/> Plot View </div> </div> <p style="text-align: right;">accuracy: 90.74% +/- 3.28% (micro average: 90.74%)</p>

Como se puede observar en la Tabla 6, el modelo de clasificación con mayor exactitud para este estudio fue k-NN, por esta razón es el modelo seleccionado, el cual basa su funcionamiento en el cálculo de las distancias entre el dato que se desea asignar a una clase, y las clases a las que pertenecen sus k vecinos más cercanos.

Figura 13

Operador k-NN



En la Figura 14, se pueden visualizar las variables y los pesos asignados en la ejecución del proceso construido. Los pesos de uno (1) han sido asignados a las variables que podrían ser parte de la solución y cero (0) cuando posiblemente no son variables consideradas en la solución.

Figura 14

Resultados de la asignación de pesos a las variables

attribute	weight
Cantidad de alumnos reprobados	0
Cantidad de alumnos aprobados	0
Cantidad de alumnos matriculados	0
Cantidad de alumnos retirados	0
Cantidad de años de experiencia docente	1
Cantidad de años de experiencia profesional	1
Cantidad de horas para docencia	1
Cantidad de horas para gestión académica	1
Cantidad de horas para investigación	1
Cantidad de horas para vinculación	1
Código	0
Condición laboral	1
Créditos	0
Departamento	0
Día de la semana	0
Género	0
Hora fin	0
Hora inicio	0
Materia	0
Nivel máximo de formación académica	1
Periodo académico	0
Tiempo de dedicación	1

3.1.5 Transformación de datos

En función del resultado obtenido con la herramienta RapidMiner, se identificaron las variables objeto para el presente estudio, las cuales corresponden a las marcadas con 1 como se mostró en la Figura 14. Esto permitió mejorar el análisis ya que se consideraron únicamente las variables representativas.

Las variables objeto del estudio, se detallan a continuación:

- Nivel de formación académica
- Tiempo de dedicación
- Condición laboral
- Área de conocimiento de la materia
- Años de experiencia docente
- Años de experiencia profesional
- Cantidad de horas de docencia
- Cantidad de horas de investigación
- Cantidad de horas de gestión académica
- Cantidad de horas de vinculación

3.1.6 Análisis de datos

El método que se eligió para realizar el análisis de datos fue la estadística descriptiva. El análisis realizado se enfocó en los siguientes aspectos:

- Datos académicos
- Experiencia docente y profesional
- Carga horaria
- Asignación de materias

- Sugerencia de docentes encuestados sobre la mejora al actual proceso de carga horaria

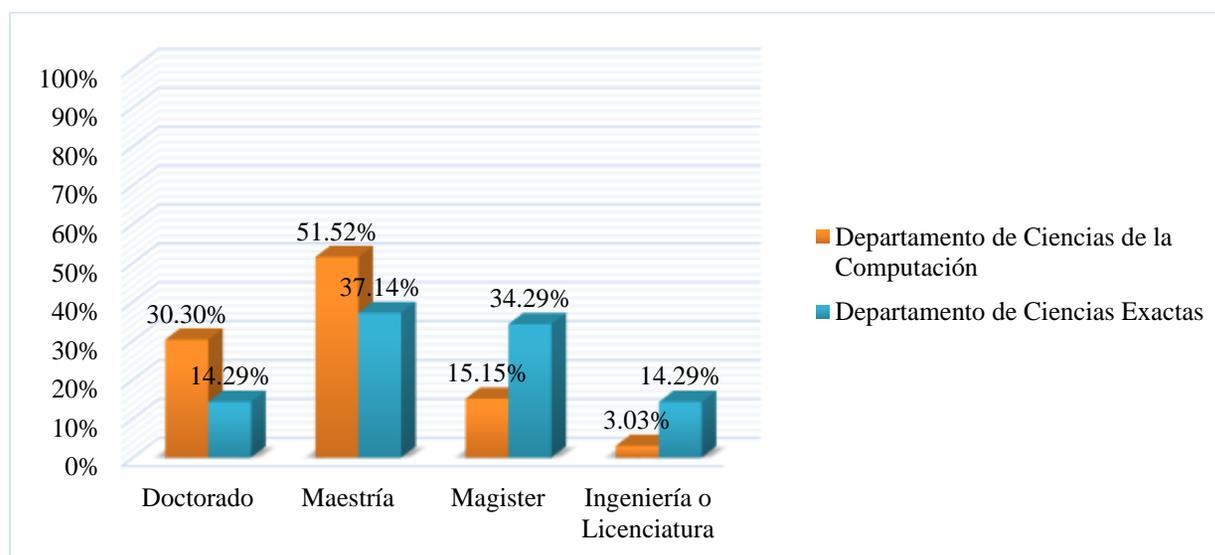
a. Datos académicos

Se inició con un análisis sobre la formación académica, tiempo de dedicación, condición laboral y género.

En cuanto a la formación académica, y de acuerdo a lo dispuesto por el CES, detallado en el capítulo II numeral 2.7, se indica que la formación mínima de un docente debe ser maestría. En la Figura 15 se puede apreciar que en la muestra estudiada solo el 30.30% de docentes del Departamento de Ciencias de la Computación y el 14.29% de docentes del Departamento de Ciencias Exactas han superado la formación mínima, alcanzando la formación académica de Doctorado, lo que quiere decir que es una variable importante a considerar al momento de asignar las materias de acuerdo a la formación académica y especialidad.

Figura 15

Formación académica



De acuerdo a lo dispuesto en el artículo 11 que consta en el documento “Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior” los docentes a tiempo completo cumplen con 40 horas laborables a la semana y los docentes a tiempo parcial cumplen con menos de 20 horas laborables a la semana.

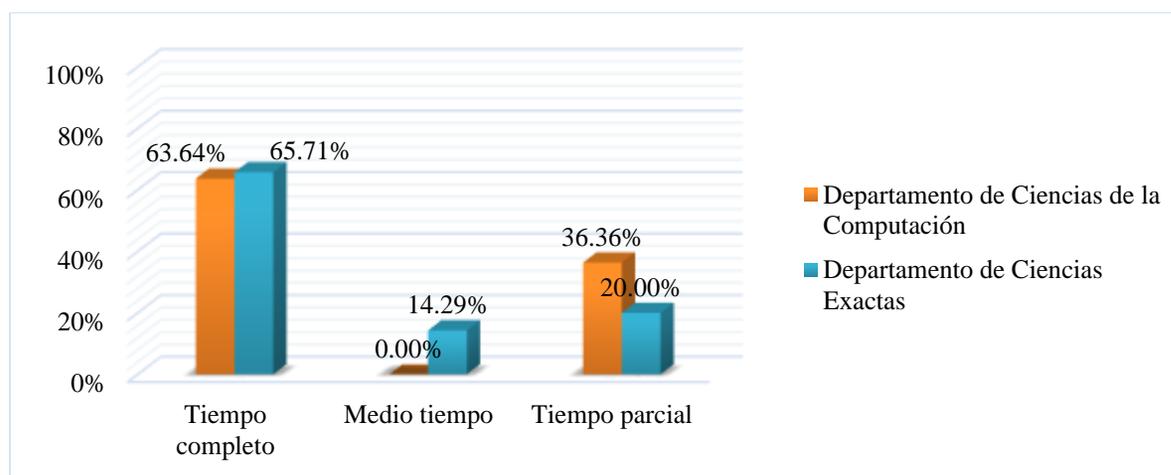
Al contar con porcentajes altos en cantidad de docentes a tiempo completo, como se evidencia en la Figura 16 en el Departamento de Ciencias de la Computación conforman el 63.64%, y en el departamento de Ciencias Exactas el 65.71%, la IES se puede beneficiar ya que dichos docentes pueden desarrollar actividades de gestión académica, vinculación e investigación, además pueden facilitar la participación de estudiantes en proyectos, lo cual aporta experiencia tanto a docentes como alumnos.

Por otro lado, al contar con docentes a tiempo parcial la IES puede ofertar a sus estudiantes materias dictadas por docentes con conocimientos prácticos en el ambiente laboral, lo cual aporta de manera significativa en las clases impartidas.

Por estos motivos la variable tiempo de dedicación debería ser considerada para la asignación de materias y su ponderación podría ser alta.

Figura 16

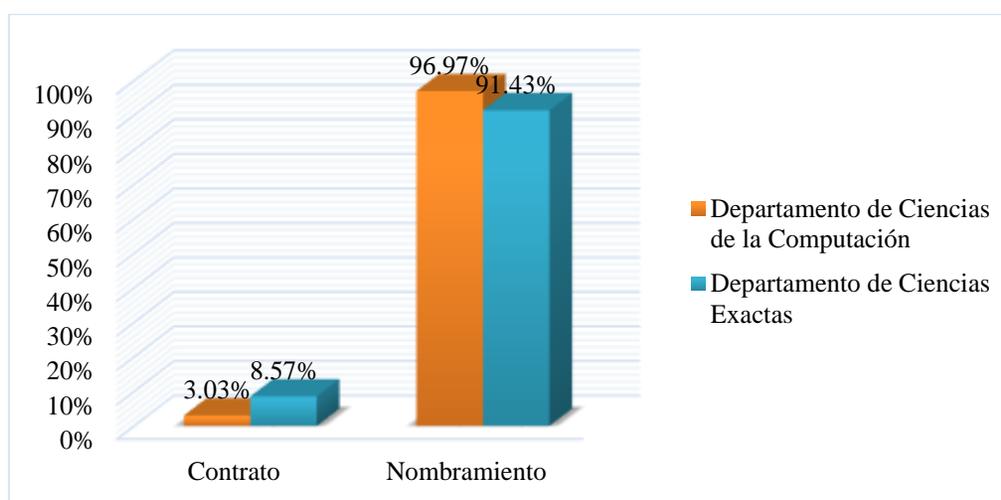
Tiempo de dedicación



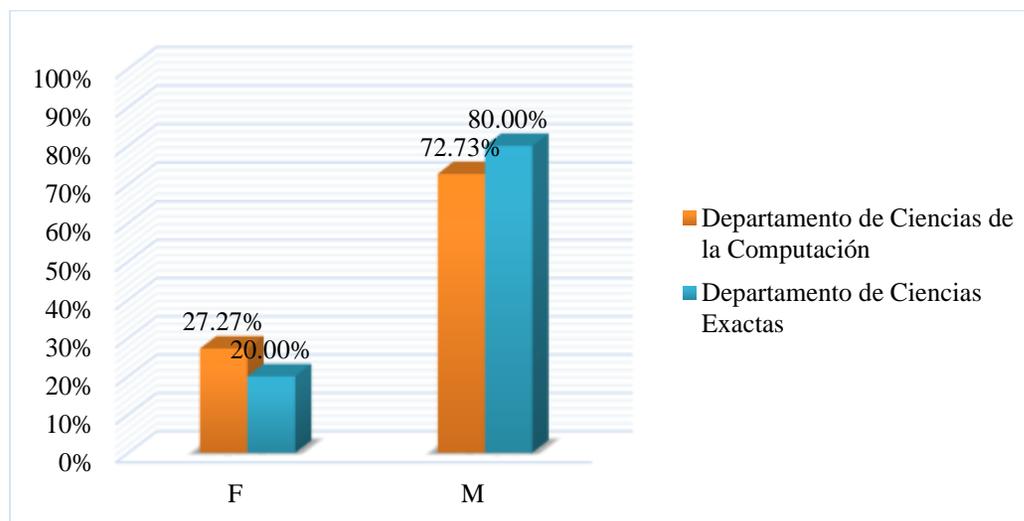
En la Figura 17 se puede identificar que, el 96.97% y el 91.43% de los docentes encuestados, de los departamentos Ciencias de la Computación y Ciencias Exactas respectivamente, tienen condición laboral nombramiento, lo cual garantiza la estabilidad laboral de los docentes y de los procesos de enseñanza.

Figura 17

Condición laboral

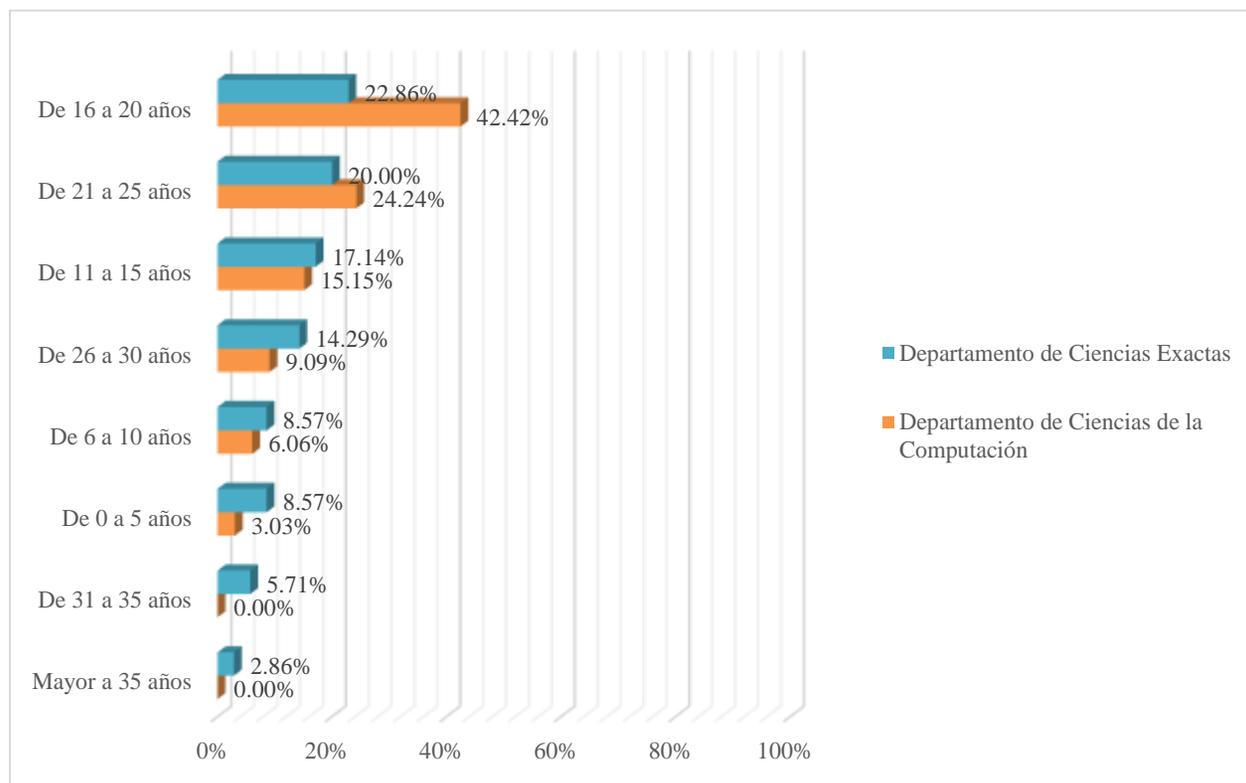


En la Figura 18 se puede notar que, en los dos departamentos existe mayor cantidad de docentes con género masculino, para el departamento de Ciencias de la Computación existe un 72.73% y para el departamento de Ciencias Exactas un 80%, además se denota que la inclusión de docentes de género femenino todavía mantiene porcentajes bajos. Sin embargo, esta variable no se considera relevante debido que el género no condiciona la asignación de materias y carga horaria docente.

Figura 18*Género de docentes***b. Experiencia docente y profesional**

Cuando analizamos el tiempo de experiencia docente entre los dos departamentos escogidos como caso de estudio, podemos ver en la Figura 19, que para el departamento de Ciencias Exactas el 22.86% y para el departamento de Ciencias de la Computación el 42.42% de docentes, cuentan con experiencia en el rango de 16 a 20 años.

De acuerdo a lo analizado, la experiencia docente en la IES es significativa, lo cual contribuye a la calidad de los profesionales que se forman en la IES y, por tanto, esta variable puede ser tomada en cuenta como un indicador para la asignación de la carga horaria docente y su ponderación podría ser alta.

Figura 19*Experiencia docente*

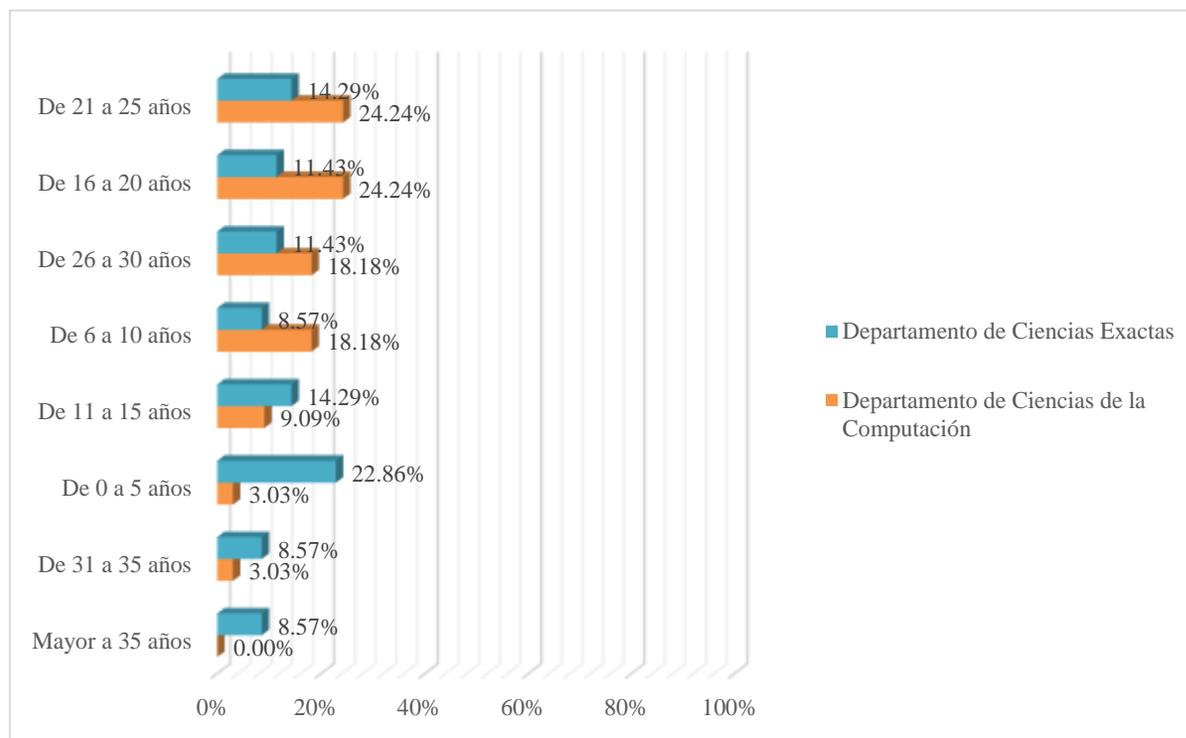
En cuanto al tiempo de experiencia profesional de los docentes pertenecientes a los departamentos del caso de estudio, podemos evidenciar que para el departamento de Ciencias de la Computación el rango de años de experiencia se encuentra entre 16 a 20 representando el 24.24%, en cambio para el departamento de Ciencias Exactas el 22.86% de docentes tiene experiencia profesional de 0 a 5 años, esto se puede también visualizar en la Figura 20.

Se puede observar una diferencia importante entre los departamentos, donde se destaca que el departamento de Ciencias de la Computación posee docentes que cuentan con mayor conocimiento práctico que los docentes del departamento de Ciencias Exactas. El conocimiento práctico es una ventaja en la enseñanza ya que los alumnos se benefician de aprender casos reales y aplicables en el ambiente laboral.

La experiencia profesional es una variable a considerar en el proceso de asignación de materias y su ponderación debería ser alta.

Figura 20

Experiencia profesional



c. Carga Horaria

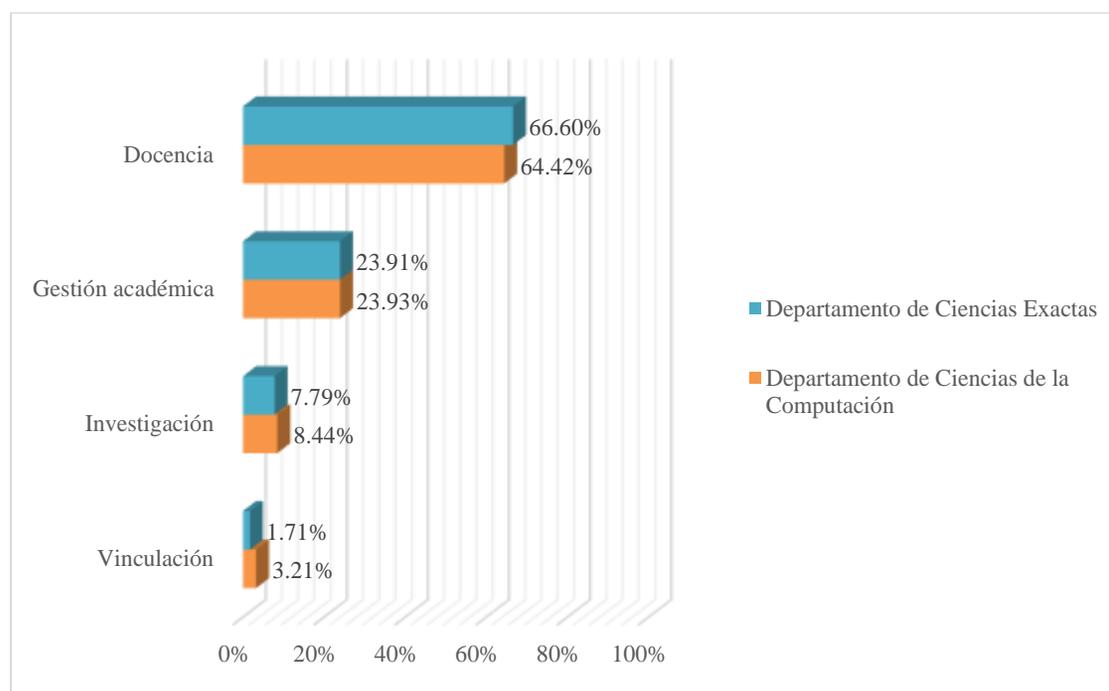
Para este aspecto se analiza la cantidad de horas destinadas para las actividades: vinculación, investigación, gestión académica y docencia a nivel de departamento.

En la Figura 21 se muestra en porcentaje la distribución de la carga horaria por tipo de actividad realizada por los docentes en los departamentos, donde se puede evidenciar que la mayor concentración de la carga horaria se encuentra en la actividad de docencia, para Ciencias Exactas corresponde al 66,60% y para Ciencias de la Computación el 64,42%, por lo tanto, se debería contar con una adecuada gestión para la asignación de materias y carga horaria docente,

la cual permita principalmente a los docentes a tiempo completo cumplir con todas las actividades establecidas por el CES que consta descrito en el artículo 6, del documento “Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior”.

Figura 21

Distribución de la carga horaria



En la Figura 22 se muestra el análisis realizado entre las horas de docencia que debe cumplir cada docente y las horas efectivas, que son aquellas que el docente dedicó a dictar clases.

Se realizó una clasificación de acuerdo al cumplimiento de las horas, donde:

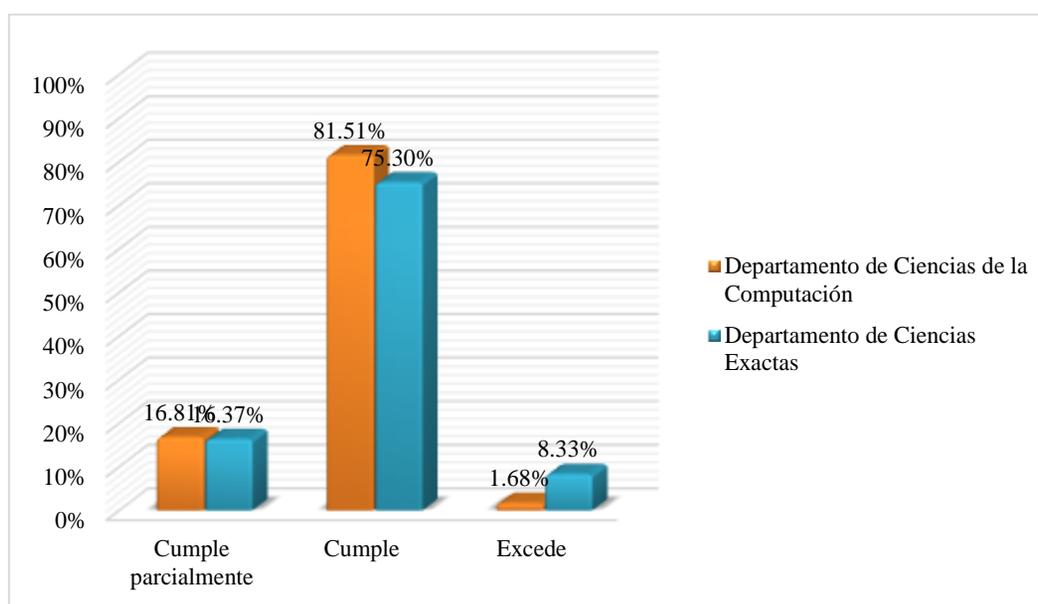
- Si el docente cubrió menos del 100% de las horas de docencia se le etiquetó como Cumple parcialmente.
- Si el docente cubrió el 100% de las horas de docencia se le etiquetó como Cumple.

- Si el docente cubrió más del 100% de las horas de docencia se le etiquetó como Excede.

Este análisis que se realizó para los departamentos del estudio dio como resultado que para el departamento de Ciencias de la Computación el 1.68% y para el departamento de Ciencias Exactas el 8.33% de los docentes tienen sobre carga respecto a las horas de docencia asignadas, por lo que esta variable se podría considerar para la asignación de carga horaria docente y su ponderación sería alta.

Figura 22

Carga horaria efectiva



d. Asignación de materias

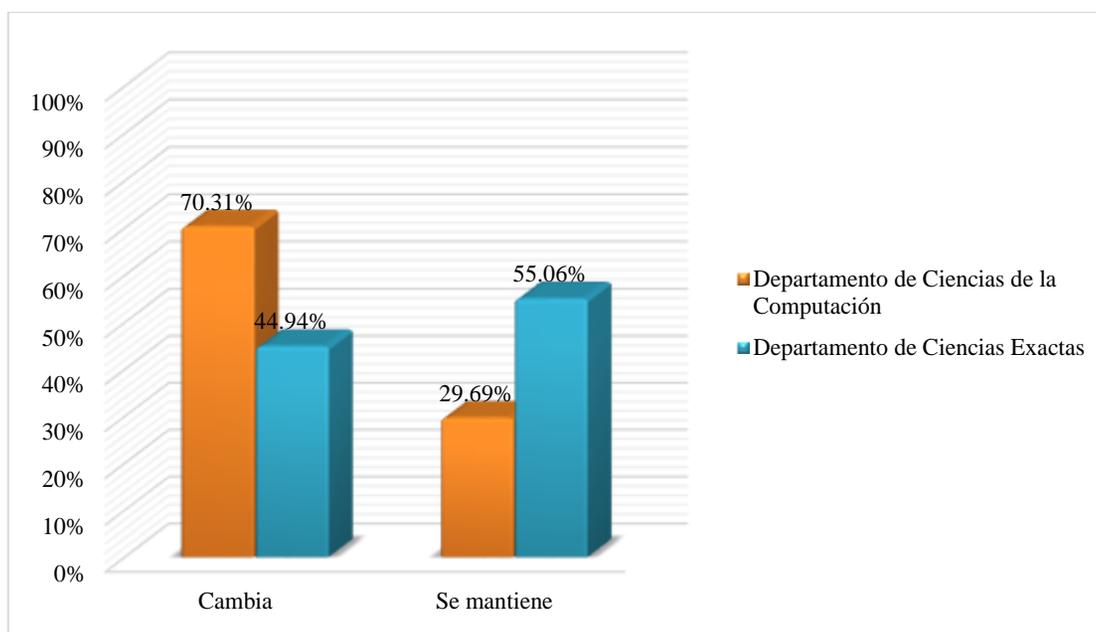
En este análisis se consideró la variación en cuanto al área de conocimiento de cada materia dictada por los docentes en los cuatro períodos académicos disponibles para este estudio, donde se ha encontrado que en ambos departamentos existe una alta rotación de materias impartidas (Figura 23). La alta rotación de materias podría ocasionar varios

inconvenientes, por ejemplo, la nueva materia asignada puede que no tenga relación con la formación académica o con la experiencia del docente, lo cual podría llevar mayor tiempo la preparación de los temas que constan en el sílabo de la materia, por tanto, no se podría garantizar calidad en el proceso de enseñanza – aprendizaje.

En consecuencia, esta variable ha sido considerada con una ponderación alta para la asignación de materias y carga horaria docente.

Figura 23

Rotación de materias impartidas



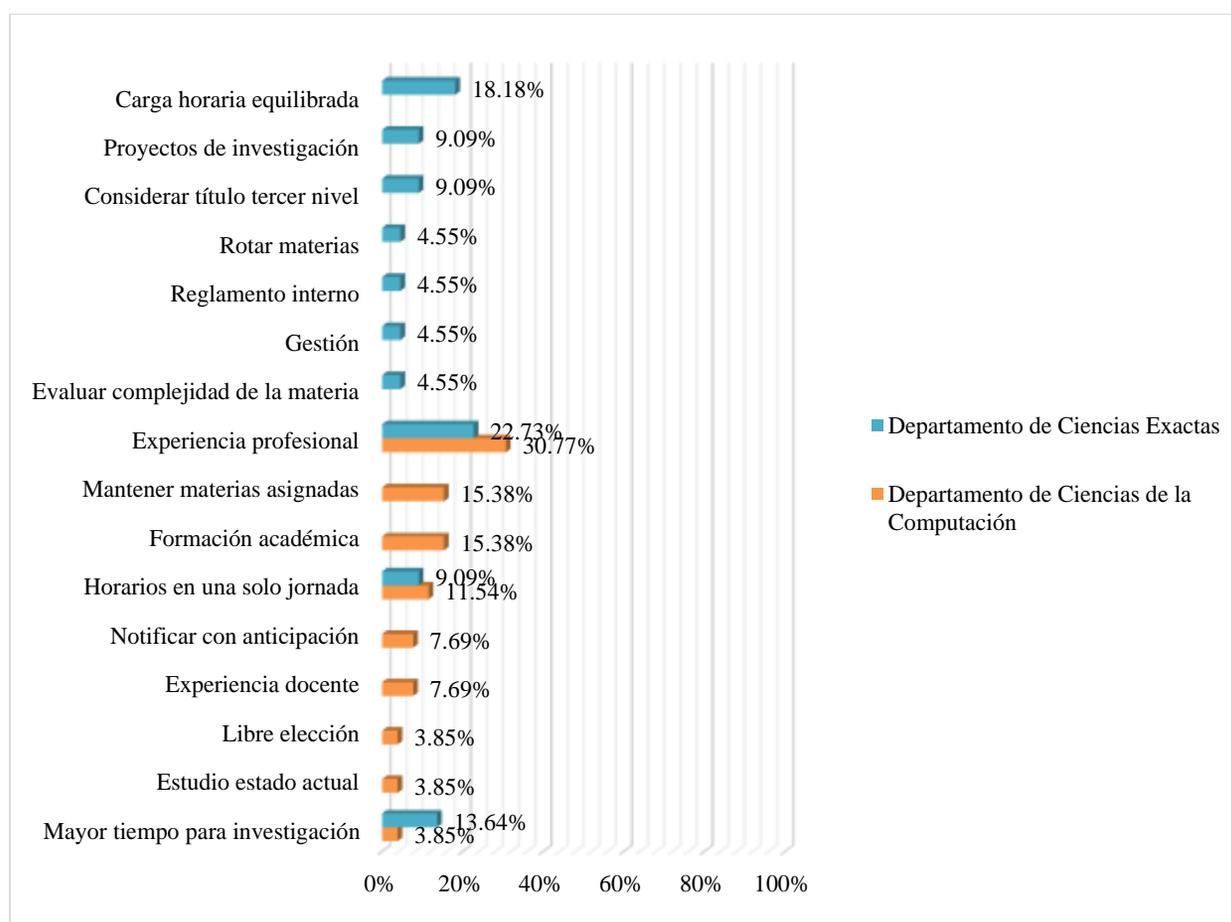
e. Sugerencias de docentes encuestados sobre la mejora al actual proceso de carga horaria

Con el objetivo de obtener ideas de mejora al proceso actual de asignación de carga horaria, se solicitó a los señores docentes escribir sugerencias acerca de dicho proceso en la encuesta realizada, la cual sirvió como complemento al análisis anteriormente detallado.

Para el análisis e interpretación de los datos cualitativos se realizó una tabulación para ambos departamentos, colocando en el eje de las “y” las sugerencias, y en el eje de las “x” la frecuencia con la que los docentes la redactaron, esto se puede observar en la Figura 24.

Figura 24

Sugerencias de mejora para asignación de carga horaria



Se puede notar que la sugerencia que tiene una tendencia muy marcada en los dos departamentos está relacionada con la experiencia profesional, donde el 22.73% en el departamento de Ciencias Exactas y el 30.77% en el departamento de Ciencias de la Computación de los docentes encuestados indica que es importante la inclusión de esta variable en el proceso de asignación de materias y carga horaria docente.

Entre las recomendaciones de soluciones destacadas para mejorar el proceso actual de asignación, el 18,18% de los 68 docentes encuestados sugieren que se les asigne una carga horaria equilibrada y el 13.64% que se les conceda mayor tiempo para desarrollar investigación. Además, el 15,38% propone que se considere la formación académica y que se mantengan las materias que se les asigna cada semestre. Adicionalmente, el 11.54% recomienda que se les asignen horarios en una sola jornada.

3.1.7 Evaluación de patrones

Se eligió un modelo de clasificación, ya que estábamos interesados en encontrar patrones de comportamiento, y éstos suelen ser los usados con mayor frecuencia en minería de datos.

Para identificar el modelo de clasificación adecuado, se realizaron ejecuciones con varios modelos y se evaluó el porcentaje de clasificación correcta. Este proceso se realizó con apoyo de la herramienta de minería de datos Weka.

Tabla 7

Porcentaje de clasificación correcta

Modelo	Resultado
NaiveBayes	<pre> Classifier output === Summary === Correctly Classified Instances 34 50 % Incorrectly Classified Instances 34 50 % Kappa statistic 0.3364 Mean absolute error 0.2784 Root mean squared error 0.4129 Relative absolute error 80.6956 % Root relative squared error 99.6508 % Total Number of Instances 68 </pre>
RandomTree	<pre> Classifier output --- Evaluation on training set --- Time taken to test model on training data: 0 seconds === Summary === Correctly Classified Instances 64 94.1176 % Incorrectly Classified Instances 4 5.8824 % Kappa statistic 0 Mean absolute error 0.0994 Root mean squared error 0.223 Relative absolute error 81.6135 % Root relative squared error 94.6331 % Total Number of Instances 68 </pre>
J48	<pre> Classifier output === Summary === Correctly Classified Instances 43 63.2353 % Incorrectly Classified Instances 25 36.7647 % Kappa statistic 0.4328 Mean absolute error 0.2581 Root mean squared error 0.3593 Relative absolute error 74.8273 % Root relative squared error 86.7127 % Total Number of Instances 68 </pre>

Modelo	Resultado
RandomForest	<pre> Classifier output === Summary === Correctly Classified Instances 35 51.4706 % Incorrectly Classified Instances 33 48.5294 % Kappa statistic 0.2137 Mean absolute error 0.3055 Root mean squared error 0.3908 Relative absolute error 88.5447 % Root relative squared error 94.3265 % Total Number of Instances 68 </pre>
DecisionStump	<pre> Classifier output === Summary === Correctly Classified Instances 39 57.3529 % Incorrectly Classified Instances 29 42.6471 % Kappa statistic 0.3562 Mean absolute error 0.2827 Root mean squared error 0.3719 Relative absolute error 81.956 % Root relative squared error 89.7546 % Total Number of Instances 68 </pre>

De acuerdo a los resultados obtenidos en la Tabla 7, el modelo de clasificación elegido fue RandomTree, que corresponde a un árbol de decisión, el cual permitirá mediante una estructura jerárquica conocer las condiciones que se cumplen para obtener una decisión.

De acuerdo con el análisis realizado con la herramienta Weka a la información recopilada, se han obtenido dos árboles de decisión uno por cada departamento.

Árbol de decisión para el departamento Ciencias de la Computación

RandomTree

=====

Nivel maximo de formacion academica = Doctorado

| Condicion laboral = Contrato : COMP-B GESTION DE INFORMACION Y HCI (2/1)

| Condicion laboral = Nombramiento

| | Cantidad de horas para gestion academica < 6

| | | Cantidad de horas para investigacion < 14

| | | Cantidad de horas para docencia < 31 : COMP-B GESTION DE INFORMACION Y HCI (1/0)
 | | | Cantidad de horas para docencia >= 31 : COMP-5 PROGRAMACION (1/0)
 | | Cantidad de horas para investigacion >= 14 : COMP-B GESTION DE INFORMACION Y HCI (3/2)
 | | Cantidad de horas para gestion academica >= 6
 | | | Cantidad de horas para investigacion < 1 : COMP-D SEGURIDAD Y GESTION DE TIC (5/3)
 | | | Cantidad de horas para investigacion >= 1
 | | | Cantidad de horas para investigacion < 4.5 : COMP-E SEGURIDAD WEB Y TECNOLOGIAS (1/0)
 | | | Cantidad de horas para investigacion >= 4.5
 | | | Cantidad de horas para docencia < 18.5 : COMP-E SEGURIDAD WEB Y TECNOLOGIAS (1/0)
 | | | Cantidad de horas para docencia >= 18.5 : COMP-D SEGURIDAD Y GESTION DE TIC (2/1)
 Nivel maximo de formacion academica = Maestria
 | Tiempo de dedicacion = Tiempo completo
 | | Cantidad de horas para gestion academica < 18
 | | | Cantidad de horas para vinculacion < 5
 | | | Cantidad de horas para docencia < 23
 | | | Cantidad de horas para vinculacion < 1
 | | | Cantidad de horas para gestion academica < 2 : COMP-B GESTION DE INFORMACION Y HCI (2/1)
 | | | Cantidad de horas para gestion academica >= 2 : COMP-5 PROGRAMACION (1/0)
 | | | Cantidad de horas para vinculacion >= 1 : COMP-F CIENCIAS COMPUTACIONALES (1/0)
 | | | Cantidad de horas para docencia >= 23
 | | | Cantidad de horas para gestion academica < 6 : COMP-D SEGURIDAD Y GESTION DE TIC (2/1)
 | | | Cantidad de horas para gestion academica >= 6 : COMP-D SEGURIDAD Y GESTION DE TIC (4/3)
 | | | Cantidad de horas para vinculacion >= 5 : COMP-5 PROGRAMACION (1/0)
 | | Cantidad de horas para gestion academica >= 18 : COMP-F CIENCIAS COMPUTACIONALES (3/2)
 | Tiempo de dedicacion = Tiempo parcial
 | | Cantidad de horas para docencia < 14 : COMP-D SEGURIDAD Y GESTION DE TIC (3/0)
 | | Cantidad de horas para docencia >= 14
 | | | Cantidad de horas para docencia < 17 : COMP-3 ING DE SOFTWARE (3/2)
 | | | Cantidad de horas para docencia >= 17 : COMP-E SEGURIDAD WEB Y TECNOLOGIAS (2/1)
 Nivel maximo de formacion academica = Magister
 | Tiempo de dedicacion = Tiempo completo : COMP-5 PROGRAMACION (1/0)
 | Tiempo de dedicacion = Tiempo parcial
 | | Cantidad de horas para gestion academica < 6
 | | | Cantidad de horas para docencia < 14 : COMP-D SEGURIDAD Y GESTION DE TIC (1/0)
 | | | Cantidad de horas para docencia >= 14 : COMP-5 PROGRAMACION (2/1)
 | | Cantidad de horas para gestion academica >= 6 : COMP-B GESTION DE INFORMACION Y HCI (3/2)
 Nivel maximo de formacion academica = Ingenieria o Licenciatura : COMP-F CIENCIAS COMPUTACIONALES (2/1)
 Size of the tree : 43

El árbol de decisión resultado para el departamento de Ciencias de la Computación realizó una clasificación para identificar el comportamiento de asignación de materias, el cual inició por la formación académica, en el caso de que la formación académica sea:

- Doctorado, realizó evaluaciones posteriores a las variables condición laboral, horas para gestión académica, horas para investigación y horas para docencia.
- Maestría, realizó evaluaciones posteriores a la variable tiempo de dedicación, cantidad de horas para gestión académica, cantidad de horas para vinculación y cantidad de horas para docencia.
- Magister, realizó evaluaciones posteriores a la variable tiempo de dedicación, cantidad de horas para gestión académica y cantidad de horas para docencia.
- Ingeniería o licenciatura, no realizó más evaluaciones posteriores.

Árbol de decisión para el departamento Ciencias Exactas

RandomTree

=====

Nivel maximo de formacion academica = Ingenieria o Licenciatura

| Tiempo de dedicacion = Medio tiempo : EXCT-3 ANALISIS (2/0)

| Tiempo de dedicacion = Tiempo completo

| | Cantidad de horas para docencia < 22 : EXCT-0 FISICA (1/0)

| | Cantidad de horas para docencia >= 22 : EXCT-3 ANALISIS (2/1)

| Tiempo de dedicacion = Tiempo parcial : EXCT-2 QUIMICA (1/0)

Nivel maximo de formacion academica = Maestria

| Cantidad de horas para gestion academica < 7.5

| | Cantidad de horas para docencia < 32

| | | Cantidad de horas para docencia < 14

| | | | Cantidad de horas para gestion academica < 1.5 : EXCT-3 ANALISIS (2/0)

| | | | Cantidad de horas para gestion academica >= 1.5 : EXCT-3 ANALISIS (2/1)

| | | Cantidad de horas para docencia >= 14

| | | | Tiempo de dedicacion = Medio tiempo : EXCT-3 ANALISIS (2/1)

| | | | Tiempo de dedicacion = Tiempo completo

| | | | | Cantidad de horas para docencia < 22 : EXCT-4 ANALISIS FUNCIONAL (2/0)

| | | | | Cantidad de horas para docencia >= 22 : EXCT-0 FISICA (1/0)

| | | | | Tiempo de dedicacion = Tiempo parcial : EXCT-0 FISICA (1/0)

| | | Cantidad de horas para docencia >= 32 : EXCT-3 ANALISIS (4/3)

| | Cantidad de horas para gestion academica >= 7.5

| | Cantidad de horas para vinculacion < 5

```

| | | Condicion laboral = Nombramiento : EXCT-1 ALGEBRA (1/0)
| | | Condicion laboral = Contrato : EXCT-1 ALGEBRA (2/1)
| | Cantidad de horas para vinculacion >= 5 : EXCT-2 QUIMICA (1/0)
Nivel maximo de formacion academica = Magister
| Cantidad de horas para investigacion < 2
| | Tiempo de dedicacion = Medio tiempo
| | | Cantidad de horas para gestion academica < 4 : EXCT-2 QUIMICA (1/0)
| | | Cantidad de horas para gestion academica >= 4 : EXCT-5 ESTADISTICA (1/0)
| | Tiempo de dedicacion = Tiempo completo
| | | Cantidad de horas para gestion academica < 8.5 : EXCT-0 FISICA (1/0)
| | | Cantidad de horas para gestion academica >= 8.5
| | | | Cantidad de horas para gestion academica < 14 : EXCT-3 ANALISIS (4/2)
| | | | Cantidad de horas para gestion academica >= 14 : EXCT-3 ANALISIS (2/1)
| | Tiempo de dedicacion = Tiempo parcial
| | | Cantidad de horas para gestion academica < 7
| | | | Cantidad de horas para gestion academica < 3
| | | | | Cantidad de horas para docencia < 13 : EXCT-3 ANALISIS (2/1)
| | | | | Cantidad de horas para docencia >= 13 : EXCT-3 ANALISIS (1/0)
| | | | Cantidad de horas para gestion academica >= 3 : EXCT-3 ANALISIS (1/0)
| | | Cantidad de horas para gestion academica >= 7 : EXCT-5 ESTADISTICA (1/0)
| Cantidad de horas para investigacion >= 2 : EXCT-0 FISICA (1/0)
Nivel maximo de formacion academica = Doctorado
| Cantidad de horas para gestion academica < 10
| | Condicion laboral = Nombramiento
| | | Cantidad de horas para gestion academica < 4
| | | | Cantidad de horas para investigacion < 2 : EXCT-2 QUIMICA (1/0)
| | | | Cantidad de horas para investigacion >= 2 : EXCT-3 ANALISIS (2/1)
| | | Cantidad de horas para gestion academica >= 4 : EXCT-5 ESTADISTICA (1/0)
| | Condicion laboral = Contrato : EXCT-2 QUIMICA (1/0)
| Cantidad de horas para gestion academica >= 10 : EXCT-3 ANALISIS (1/0)
Size of the tree : 52

```

El árbol de decisión resultado para el departamento de Ciencias Exactas realizó una clasificación para identificar el comportamiento de asignación de materias, el cual inició por la formación académica, en el caso de que la formación académica sea:

- Doctorado, realizó evaluaciones posteriores a las variables horas para gestión académica, condición laboral, horas para investigación.
- Maestría, realizó evaluaciones posteriores a las variables cantidad de horas para gestión académica, cantidad de horas para docencia y tiempo de dedicación.

- Magister, realizó evaluaciones posteriores a las variables cantidad de horas para investigación, tiempo de dedicación, cantidad de horas para gestión académica y cantidad de horas para docencia.
- Ingeniería o licenciatura, realizó evaluaciones posteriores a la variable tiempo de dedicación y cantidad de horas para docencia.

Estas clasificaciones dan a entender que:

- La mayoría de docentes cuentan con formación académica de Maestría ya que la clasificación para ambos departamentos fue más detallada, esta formación académica corresponde a la formación mínima que debe tener un docente de acuerdo a lo expuesto en los artículos del 31 al 37 del reglamento del CES antes mencionado, este comportamiento refleja que pocos docentes han superado la formación mínima requerida y conseguido el nivel académico de doctorado.
- La carga horaria en cuanto a horas de docencia es bastante alta ya que interviene en la mayoría de los niveles de los árboles de decisión para los dos departamentos, lo cual no permite a los señores docentes involucrarse en proyectos de investigación o vinculación, ni en actividades de gestión académica.

3.1.8 Presentación del conocimiento

Luego de haber realizado el análisis de los datos obtenidos en los departamentos seleccionados de la IES del presente trabajo y en función de los patrones identificados, se propone mejorar los porcentajes en la distribución de la carga horaria para que los docentes puedan dedicar el tiempo necesario a las actividades de docencia, vinculación, investigación y gestión académica de forma equilibrada y puedan cumplir así con la reglamentación vigente.

Además, se propone considerar la afinidad que los señores docentes tienen en cuanto al área de conocimiento con el fin de disminuir la rotación en la asignación de materias.

Esta propuesta pretende fomentar la actualización de conocimientos y experiencia profesional de los docentes, ya que se recomendarán las materias que se ajusten a su perfil y permitan una distribución más equitativa en cuanto a las actividades que desarrollan y tomando en cuenta su afinidad, es decir, la formación académica, experiencia docente, experiencia profesional, entre otros que han sido definidos en este estudio. Si en la asignación docente se toman en cuenta los factores considerados de afinidad, se conseguirá que el docente se motive para realizar otras actividades como dirigir o participar en proyectos de investigación y vinculación, generando oportunidades a los estudiantes para realizar prácticas pre profesionales, desarrollo de tesis, publicaciones científicas y de esta manera contribuir en el desarrollo del país.

Para conseguir lo propuesto se crearon reglas basadas en las variables analizadas a las cuales se les asignaron sus respectivas ponderaciones, con el fin de evaluar la información del docente y recomendar las asignaturas de acuerdo a los factores considerados de afinidad y experiencia.

Las variables que se han considerado para conformar las reglas son:

- Años de experiencia docente
- Años de experiencia profesional
- Nivel de formación académica
- Tiempo de dedicación

Se ha propuesto un conjunto de reglas para cada variable listada en el párrafo anterior, cada regla tiene asignado un valor de ponderación en un rango de 0 a 10, donde 0 es el valor más bajo y 10 el más alto, y la suma de las ponderaciones de este conjunto de reglas deberá ser igual a 10 para cada variable.

Dichos valores están asignados de acuerdo al análisis de información realizado en el numeral 3.1.6 del presente documento, y corresponden a los valores óptimos en cuanto al rendimiento y resultados del algoritmo implementado.

Para encontrar la ponderación de cada variable, se han evaluado los datos ingresados por los docentes considerando las reglas definidas a continuación:

Reglas para la variable años de experiencia docente

Los años de experiencia docente indican calidad en el proceso de enseñanza, si un docente cuenta con varios años de experiencia docente se le asignará mayor puntaje.

Esta ponderación se realizó en función a la cantidad de años que el docente ha ejercido la docencia, donde si el docente cuenta:

- De 1 a 5 años de experiencia docente se le asignará el valor de 0.5 puntos
- De 6 a 10 años de experiencia docente se le asignará el valor de 1 punto
- De 11 a 15 años de experiencia docente se le asignará el valor de 1.5 puntos
- De 16 a 20 años de experiencia docente se le asignará el valor de 2 puntos
- Más de 21 años de experiencia docente se le asignará el valor de 5 puntos

Reglas para la variable años de experiencia profesional

La experiencia profesional permite al docente aportar a sus alumnos casos aplicables en el campo laboral, si un docente posee varios años de experiencia profesional se le asignará un puntaje alto.

La ponderación asignada para la experiencia profesional de los docentes se realizó de acuerdo a la cantidad de años con los que el docente cuenta en el ámbito laboral profesional, y se asigna de acuerdo a los siguientes rangos:

- De 1 a 5 años de experiencia profesional se asignará el valor de 0.5 puntos
- De 6 a 10 años de experiencia profesional se asignará el valor de 1 punto
- De 11 a 15 años de experiencia profesional se asignará el valor de 1.5 puntos
- De 16 a 20 años de experiencia profesional se asignará el valor de 2 puntos
- Más de 21 años de experiencia profesional se asignará el valor de 5 puntos

Reglas para la variable formación académica

La formación académica refleja el compromiso que el docente tiene respecto a su capacitación formal, por lo que si el docente cuenta con una formación académica de alto nivel académico se le asignará mayor puntaje.

De acuerdo al análisis realizado a datos obtenidos para este estudio, los docentes de la IES tienen formación académica de tercer o cuarto nivel (maestría o doctorado), por lo que la ponderación se realizó como se indica a continuación:

- Docente con formación académica de tercer nivel se asignará el valor de 4 puntos
- Docente con formación académica de cuarto nivel se asignará el valor de 6 puntos

Reglas para la variable tiempo de dedicación

El tiempo de dedicación refleja de acuerdo a la normativa la cantidad de horas de docencia que puede tener asignada un docente. La ponderación fue asignada a las tres opciones definidas por el CES, dándole mayor puntaje a tiempo completo.

- Tiempo parcial se le asignará el valor de 2 puntos
- Medio tiempo se le asignará el valor de 3 puntos
- Tiempo completo se le asignará el valor de 5 puntos

Evaluación total de la Afinidad

Además de la creación de reglas, se definieron niveles de afinidad, que se obtienen al evaluar cada variable en función de las reglas propuestas, permitiendo así identificar las materias que son las más adecuadas al perfil del docente.

Si un docente obtiene la mayor puntuación por cada regla definida, su puntuación será de 29 puntos, los cuales representan el 100%.

Para identificar los niveles de afinidad se establecieron rangos en función a la puntuación obtenida en la evaluación de las variables, por lo que si el docente obtiene:

- Entre cero y treinta % de la puntuación, se considera una afinidad baja.
- Entre treinta y uno y sesenta y cinco %, se considera una afinidad media.
- Entre sesenta y seis y cien %, se considera una afinidad alta.

CAPÍTULO IV

DESARROLLO DEL PROTOTIPO

En este capítulo se describe por cada fase de la metodología prototipado rápido, el desarrollo realizado alineado de acuerdo a la propuesta realizada en el Capítulo III.

4.1 Fase 1: Identificación de requisitos

Para mejorar el proceso de asignación de materias a los señores docentes de la IES, se ha realizado un estudio analítico aplicando pasos, herramientas y técnicas de Minería de Datos. Los resultados obtenidos han permitido elaborar una propuesta de solución. Dicha propuesta se encuentra descrita en la sección 3.2.7 del capítulo anterior.

Para comprobar la validez de la propuesta, se ha desarrollado un prototipo base el cual evalúa las reglas y variables propuestas relacionadas con la afinidad del docente como la formación académica y formación profesional.

Con el fin de realizar las pruebas del prototipo planteado con datos reales, se ha propuesto el desarrollo de un formulario web que permita recolectar los datos actualizados de los docentes y obtener recomendaciones de asignación de materias a los docentes de los dos departamentos analizados en el presente trabajo.

Los resultados de la recomendación obtenida se mostrarán al usuario con perfil administrador (planificador o coordinador académico) en orden descendente de acuerdo a la afinidad obtenida por cada materia, con el fin de facilitar la toma de decisiones en el proceso de asignación de materias y carga horaria de los docentes de la IES.

4.2 Fase 2: Diseño del prototipo

En esta fase se muestra el diseño del prototipo para la asignación de materias a un docente de acuerdo a los parámetros establecidos de afinidad.

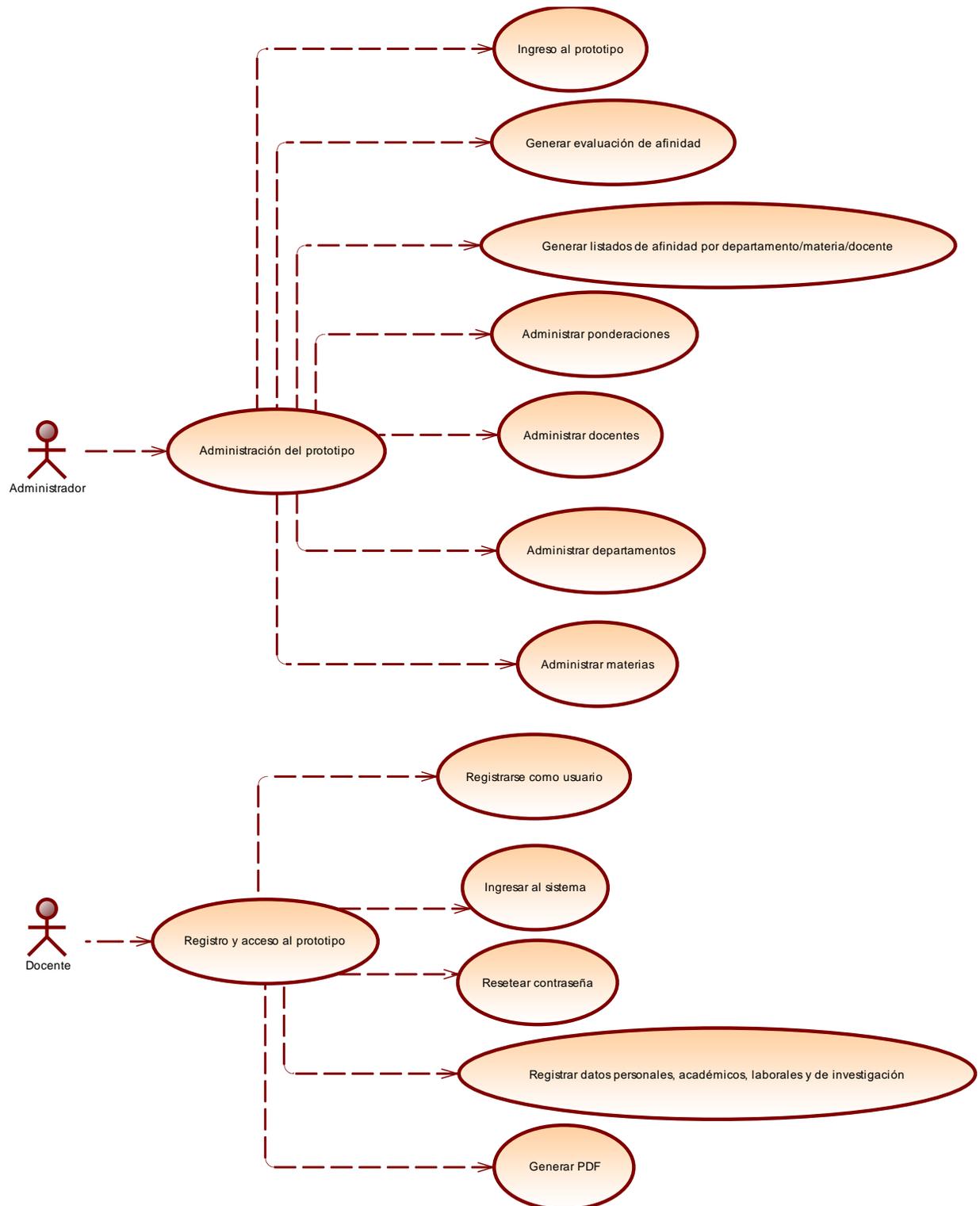
A continuación, se muestran los diagramas elaborados para el diseño del prototipo.

4.2.1 Diagrama de casos de uso

En este diagrama se muestran los actores que intervienen en el manejo del prototipo. Por una parte, el administrador que registra la información y parametriza el prototipo y por otra parte el docente quien registra sus datos.

Figura 25

Diagrama de casos de uso



Se describen los casos de usos a continuación:

Tabla 8

Descripción del caso de uso No. 1

Descripción del caso de uso	
Nombre	Administración del prototipo
Descripción	Permite al actor con perfil administrador, administrar el prototipo en función de las siguientes acciones: <ul style="list-style-type: none"> - Ingreso al prototipo - Generar evaluación de afinidad - Generar listados de afinidad por departamento/materia/docente - Administrar ponderaciones - Administrar docentes - Administrar departamentos - Administrar carreras - Administrar materias

Tabla 9

Descripción del caso de uso No. 2

Descripción del caso de uso	
Nombre	Ingreso al prototipo
Descripción	Permite al actor que tiene perfil administrador ingresar al prototipo con sus respectivas credenciales.

Tabla 10*Descripción del caso de uso No. 3*

Descripción del caso de uso	
Nombre	Generar evaluación de afinidad
Descripción	Permite al actor que tiene perfil administrador, generar la evaluación de afinidad para obtener las materias que se recomiendan con su respectiva afinidad.

Tabla 11*Descripción del caso de uso No. 4*

Descripción del caso de uso	
Nombre	Generar listas por departamento, materias o docente de afinidad
Descripción	Permite al actor que tiene perfil administrador, generar un listado de los resultados de las recomendaciones con su respectiva afinidad donde se muestra el departamento y las materias recomendadas por docente.

Tabla 12*Descripción del caso de uso No. 5*

Descripción del caso de uso	
Nombre	Administrar ponderaciones
Descripción	Permite al actor que tiene perfil administrador crear, modificar y eliminar las ponderaciones asignadas a las reglas.

Tabla 13*Descripción del caso de uso No. 6*

Descripción del caso de uso	
Nombre	Administrar docentes
Descripción	Permite al actor que tiene perfil administrador la creación, modificación y eliminación de datos de docentes.

Tabla 14*Descripción del caso de uso No. 7*

Descripción del caso de uso	
Nombre	Administrar departamentos
Descripción	Permite al actor que tiene perfil administrador crear, modificar y eliminar datos de departamentos.

Tabla 15*Descripción del caso de uso No. 8*

Descripción del caso de uso	
Nombre	Administrar materias
Descripción	Permite al actor que tiene perfil administrador la creación, modificación y eliminación de datos de materias.

Tabla 16*Descripción del caso de uso No. 9*

Descripción del caso de uso	
Nombre	Registro y acceso al prototipo
Descripción	Permite al actor que tiene perfil docente crear un usuario del prototipo, acceder a dicho prototipo e ingresar datos.

Tabla 17*Descripción del caso de uso No. 10*

Descripción del caso de uso	
Nombre	Registrarse como usuario
Descripción	Permite al actor que tiene perfil docente crear su usuario para ingreso al prototipo.

Tabla 18*Descripción del caso de uso No. 11*

Descripción del caso de uso	
Nombre	Ingresar al prototipo
Descripción	Permite al actor que tiene perfil docente ingresar al prototipo usando su usuario.

Tabla 19*Descripción del caso de uso No. 12*

Descripción del caso de uso	
Nombre	Resetear contraseña
Descripción	Permite al actor que tiene perfil docente modificar su contraseña de acceso al prototipo en caso de olvido.

Tabla 20*Descripción del caso de uso No. 13*

Descripción del caso de uso	
Nombre	Registrar datos personales, académicos, laborales y de investigación
Descripción	Permite al actor que tiene perfil docente, luego de haber iniciado sesión en el prototipo, registrar sus datos: personales, académicos, laborales y de investigación.

Tabla 21

Descripción del caso de uso No. 14

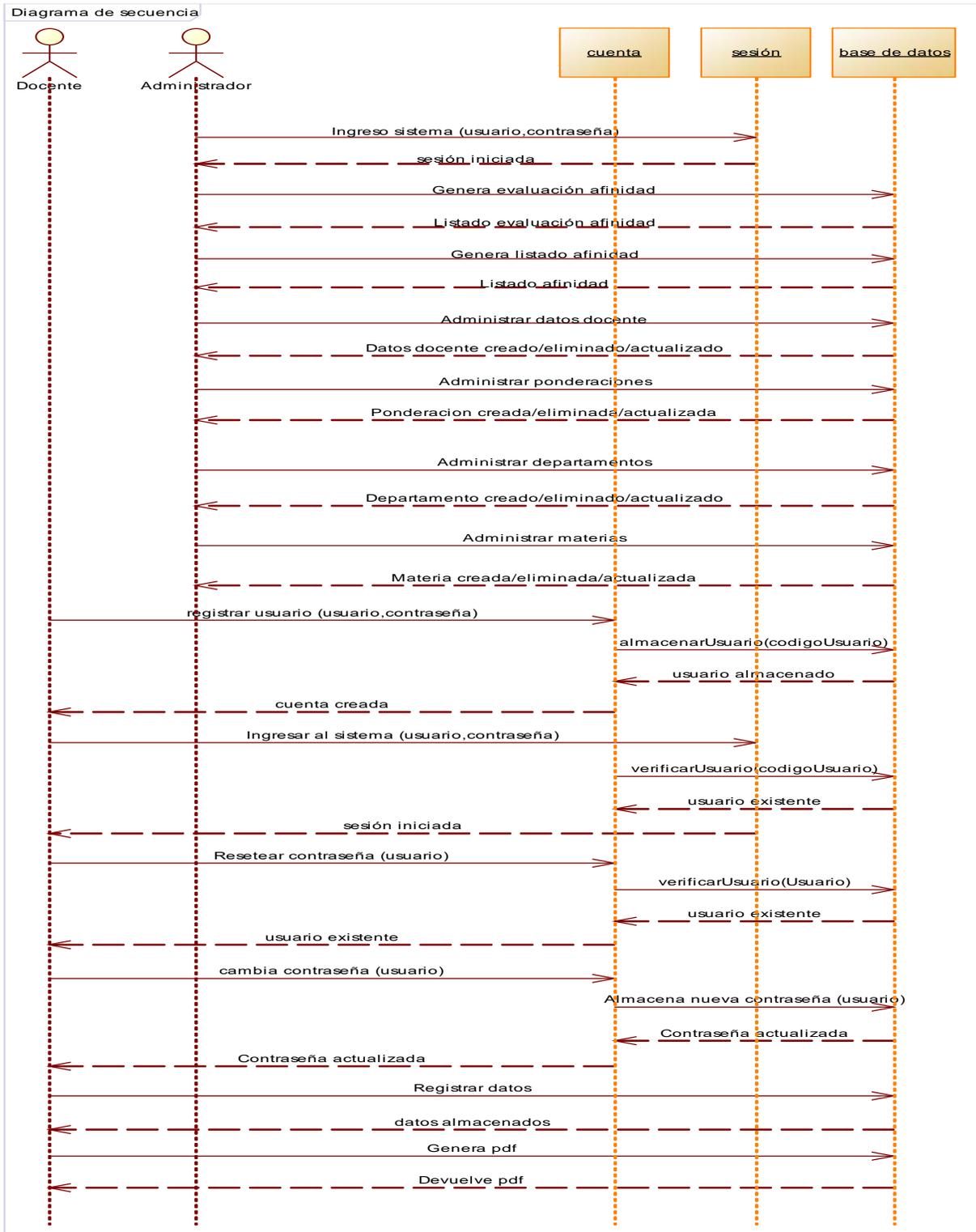
Descripción del caso de uso	
Nombre	Generar PDF
Descripción	Permite al actor que tiene perfil docente generar un archivo PDF, que contiene sus datos ingresados de forma resumida.

4.2.2 Diagrama de secuencia

Se elaboró un diagrama de secuencia con el objetivo de mostrar en orden cronológico las acciones del intercambio de mensajes entre los procesos que intervienen en el prototipo desarrollado.

Figura 26

Diagrama de secuencia



4.2.3 Configuración para el prototipo

Se ha configurado un servidor para alojar el prototipo. Dicho servidor tiene las siguientes características:

- Arquitectura de 64 bits
- Sistema operativo Windows Server 2019 Standard
- Versión 10.0.17763

4.2.4 Diseño del formulario web

El formulario web fue desarrollado en JSP (JavaServer Pages), en tres paquetes de programación de la siguiente manera:

- Web Pages: Este paquete se utilizó para la distribución de las páginas web, la cuales son la interface gráfica hacia el usuario.
- Source Packages: Este paquete se utilizó para distribuir la programación en el diseño lógico de la programación orientada a objetos de java.
- Libraries: Este paquete se utilizó para reunir las librerías que pertenecen a programaciones libres, como en nuestro caso el algoritmo y de esta manera llamamos cada funcionalidad del mismo.

Se crearon dos tipos de perfiles para que los usuarios puedan interactuar con el formulario web:

- Usuario con perfil docente: puede realizar el ingreso y modificación de sus datos académicos y profesionales.
- Usuario con perfil administrador: puede modificar los datos que fueron ingresados por los usuarios con perfil docente y puede además generar la evaluación de afinidad.

4.2.5 Diseño de interfaces

El prototipo tiene varias interfaces para la interacción con el usuario. A continuación, se muestran las principales.

a. Interfaz de acceso al prototipo

En la Figura 27 se muestra la interfaz de acceso al prototipo, que permite crear usuarios, iniciar sesión a usuarios (Docente y/o Administrador) y resetear contraseñas.

Figura 27

Diseño interfaz de acceso al prototipo

El diagrama muestra una interfaz de usuario con un fondo gris claro. En la parte superior izquierda, hay un campo de texto etiquetado 'Usuario:' con un recuadro blanco que contiene el texto 'Usuario/Cedula'. Debajo de esto, hay otro campo de texto etiquetado 'Contraseña:' con un recuadro blanco que contiene el texto 'Contraseña usuario'. En la parte inferior, hay tres botones ovalados con un fondo azul claro y un borde gris: 'Ingresar' a la izquierda, 'Registrar' en el centro y 'Recuperar contraseña' a la derecha.

b. Registro de datos personales, académicos, laborales y de investigación

Para los usuarios que tienen perfil docente se habilitarán 9 interfaces las que permitirán ingresar datos clasificados por tipo de información como son: datos personales, académicos, laborales y de investigación.

En estas interfaces se puede apreciar un cuadro ubicado en la parte superior izquierda, el cual mostrará los datos ingresados previamente por el docente, para cuando sea necesario actualizarlos, y además mostrará un cuadro ubicado en la parte derecha el cual va a contener los campos a ingresar por el docente, como se muestra en la Figura 28.

Figura 28

Diseño de interfaces para registro de datos

El diagrama muestra una interfaz de usuario con los siguientes elementos:

- Encabezado: Título del formulario
- Panel lateral izquierdo: Listado de datos ingresados
- Formulario principal: Título Formulario, con cuatro campos de entrada etiquetados como Dato 1, Dato 2, Dato 3 y Dato 4, y un botón Guardar.
- Pie de página: Título del formulario

c. Interfaz para generar evaluación y asignación de materias

El usuario con perfil administrador podrá generar la evaluación de afinidad mediante la interfaz que se muestra en la Figura 29, de los docentes que han ingresado sus datos. En esta interfaz además se presentará los resultados de la evaluación de afinidad.

La interfaz cuenta en la parte izquierda de una lista donde constan los nombres de los docentes que han ingresado sus datos, debajo de esta lista se dispone de un botón que permite generar la evaluación de afinidad del docente elegido, el resultado se mostrará en la parte derecha de la interfaz.

Figura 29

Diseño de interfaz para generar evaluación de afinidad

Área de conocimiento	Materia recomendada	Afinidad (%)	Descripción afinidad

d. Reportes

Como complemento del formulario web desarrollado, se incluyeron dos reportes, el primero permite al usuario con perfil administrador visualizar de forma resumida los datos de la evaluación de afinidad de todos los docentes, como se aprecia en la Figura 30.

Figura 30*Reporte de listados de afinidad*

Resultados evaluación de afinidad de todos los docentes						
<input type="checkbox"/>	<input type="checkbox"/> Cedula	<input type="checkbox"/> Nombre	<input type="checkbox"/> Apellido	<input type="checkbox"/> Area de conocimiento	<input type="checkbox"/> Departamento	<input type="checkbox"/> Afinidad

El segundo reporte permitirá al usuario con perfil docente visualizar un resumen de sus datos ingresados en formato PDF (Figura 31), este archivo podrá ser exportado.

Figura 31*Reporte de datos del docente*

Logotipo Universidad	Datos personales del docente
Datos personales del docente	
Información laboral del docente	
Títulos académicos del docente	
Títulos académicos del docente	
Capacitaciones realizadas por el docente	
Publicaciones realizadas por el docente	
Participación en investigaciones del docente	
Experiencia docente del docente	
Experiencia profesional del docente	

4.3 Fase 3: Construcción del prototipo

En esta fase se describe el desarrollo del prototipo con el algoritmo implementado.

4.3.1. Diseño de la base de datos

Se ha diseñado una base de datos en Power Designer y se ha implementado en MySQL v5.0 con distribución desarrollador. En dicha base de datos se almacena la información recopilada mediante el formulario web y los resultados recomendados por el algoritmo genético. El diseño realizado se muestra en la Figura 32.

4.4 Fase 4: Evaluación del prototipo

Las pruebas se realizaron con datos reales y se aplicaron para varios docentes, pero se mostrarán a manera de ejemplo los resultados de dos docentes, uno perteneciente al departamento de Ciencias de la Computación y uno del departamento de Ciencias Exactas.

Los docentes con perfil administrador deben ingresar el catálogo de materias del departamento al que pertenecen, para facilitar esta tarea los tesisas realizamos dicha carga para ambos departamentos del caso de estudio.

Se les pidió a los señores docentes que registren sus datos mediante las interfaces desarrolladas en el prototipo propuesto.

El proceso de evaluación inició cuando los docentes crearon su usuario como se muestra en la Figura 33.

Figura 33

Creación de usuario en el prototipo

The image shows a web interface for 'GesDoc'. At the top left is the logo, which consists of a stylized blue and yellow figure resembling a person or a network node, followed by the text 'GesDoc'. To the right of the logo are three navigation links: 'Inicio', 'Contáctenos', and 'Ayuda'. Below the navigation is a horizontal line, and underneath it, the text 'Sistema para el registro y gestión del personal Docente.' is centered. The main content area is a light gray box containing the text 'Debe estar registrado para iniciar sesión' at the top. Below this text are two input fields: the first is labeled 'Usuario/Cedula o Pasaporte' and the second is labeled 'Contraseña'. Below the input fields is a dark gray button with the text 'Ingresar'. At the bottom of the main content area, there are two dark gray buttons: 'Regístrate' on the left and '¿Olvidaste tu contraseña?' on the right.

Proceso de evaluación para el docente del departamento de Ciencias de la Computación

El docente del departamento de Ciencias de la Computación registró sus datos personales, académicos y de investigación (Figura 34).

Figura 34

Registro de datos del docente del Departamento de Ciencias de la Computación

Información Laboral Docente

Información Laboral Registradas

Para editar los datos registrados , haga clic sobre la opción deseada del listado a continuación.

Ciencias de la Computación - Nombramiento **BORRAR**

campos requeridos (*)

Categoría (*): Titular Principal (1, 2 o 3) ▼

Dedicación (*): Tiempo Completo ▼

Relación Laboral (*): Nombramiento ▼

Departamento (*): Ciencias de la Computación ▼

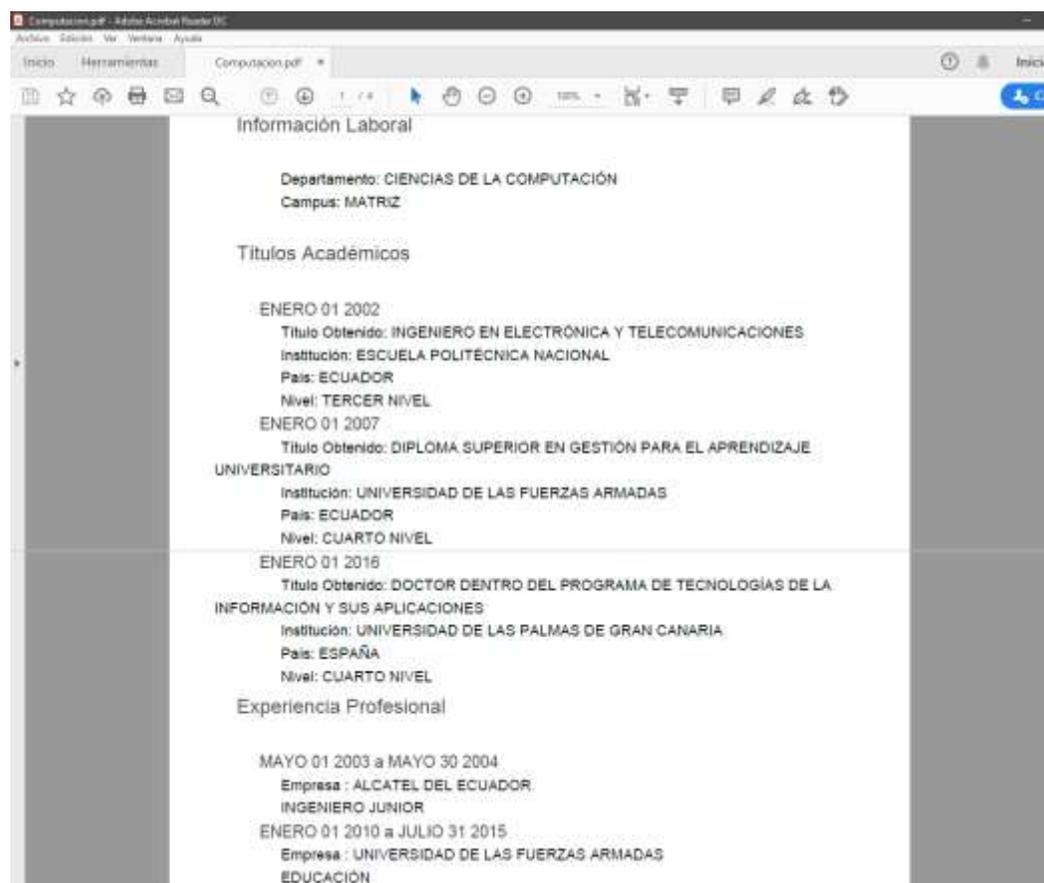
Campus (*): Matriz ▼

Horas Docencia Semanal (*): 16

GUARDAR

<< Atrás 2/9 Siguiente >>

De acuerdo a los datos ingresados por el docente, se cuenta con la siguiente información para realizar la evaluación y propuesta de materias que se ajustan a su perfil.

Figura 35*Perfil Docente Ciencias de la Computación*

En resumen, tenemos lo siguiente:

- Máxima formación académica: Cuarto nivel
- Máximo título obtenido: Doctorado
- Área de conocimiento: Tecnologías de la información
- Años de experiencia profesional: 6
- Ámbito experiencia profesional: Tecnologías de la información
- Años de experiencia docente: 15

Una vez que se cuenta con los datos del docente y las reglas, el administrador del departamento de Ciencias de la Computación accedió al prototipo para ejecutar la evaluación de afinidad mediante la interface “Evaluación de afinidad”, misma que se muestra en la Figura 36, en esta interface el administrador eligió al docente de una lista desplegable y presionó el botón Evaluar, con esta acción se ejecutó el análisis a través del algoritmo implementado.

Figura 36

Evaluación de afinidad administrador departamento Ciencias de la Computación

Área de conocimiento	Materia recomendada	Afinidad (%)	Descripción afinidad
----------------------	---------------------	--------------	----------------------

Se realizaron varias ejecuciones de la evaluación de afinidad para el docente, de forma que se pudiera validar que las recomendaciones realizadas por el algoritmo en cada corrida estaban dentro de su área de conocimiento.

En la Figura 37 se pueden apreciar los resultados de una de las varias ejecuciones realizadas para el docente del departamento Ciencias de la

Computación. Los datos personales fueron ocultados debido a que corresponden a datos reales.

Figura 37

Materias recomendadas docente Departamento de Ciencias de la Computación

Docentes Registrados

EVALUAR

Datos generales

Nombre:

Apellido:

Cédula:

Área de conocimiento	Materia recomendada	Afinidad (%)	Descripción habilidad
COMP-S PROGRAMACION	FUNDAMENTOS DE PROGRAMACIÓN	59	Afinidad media
COMP-C INFRAESTRUCTURA DE TICS	REDES DE DATOS II	69	Afinidad alta
COMP-D SEGURIDAD Y GESTION DE TIC	PLANIFICACION DE TICS	81	Afinidad alta
COMP-D SEGURIDAD Y GESTION DE TIC	TEC INF PARA NEGOCIOS	42	Afinidad media
COMP-G DESA ANALI SOFTWARE Y APLICADI	DESARROLLO WEB PARA LA INTEGRA	13	Afinidad baja
COMP-I DISEÑO Y ADM DE REDES	SISTEMAS OPERATIVOS	92	Afinidad alta

Figura 38

Proceso de ejecución del algoritmo - Docente Departamento de Ciencias de la Computación

```
#####
Generación: 1
Cromosoma más fuerte : Chromosome [fitnessScore=467.4, id=1, genes=[95, 48, 146, 25, 102, 191]]

Codigo Materia:7
Titulo Materia:TEC INF PARA NEGOCIOS
Creditos: 10.0
Genres: [COMP-D SEGURIDAD Y GESTION DE TIC]
Codigo Materia:34
Titulo Materia:SISTEMAS OPERATIVOS
Creditos: 7.5
Genres: [COMP-I DISEÑO Y ADM DE REDES]
Codigo Materia:30
Titulo Materia:FUNDAMENTOS DE PROGRAMACIÓN
Creditos: 10.0
Genres: [COMP-5 PROGRAMACION]
Codigo Materia:22
Titulo Materia:REDES DE DATOS III
Creditos: 4.0
Genres: [COMP-C INFRAESTRUCTURA DE TICS]
Codigo Materia:61
Titulo Materia:PLANIFICACION DE TICS
Creditos: 4.0
Genres: [COMP-D SEGURIDAD Y GESTION DE TIC]
Codigo Materia:37
Titulo Materia:DESARROLLO WEB PARA LA INTEGRA
Creditos: 3.0
Genres: [COMP-G DESA ANALI SOFTWARE Y APLICACI]
#####
>>> Resultado de la evolucion: Chromosome [fitnessScore=467.4, id=1, genes=[95, 48, 146, 25, 102, 191]]
15 1 1 0000 01 00 00 000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
```

Como se puede ver en la figura 38, el algoritmo devuelve algunos resultados:

- Generación: indica el valor de la generación más fuerte.
- Cruce de mutación (fitness): muestra la cantidad de interacciones realizadas en el proceso de selección para encontrar la población más fuerte.
- Población (genes): Corresponde a las materias que resultaron ser las que representaron los genes más fuertes.

Con el objetivo de evaluar los resultados obtenidos por el prototipo se realizó una comparación de las materias asignadas de forma manual en los cuatro períodos académicos que se disponen para este estudio y las materias asignadas mediante el prototipo que mostraron afinidad alta.

Tabla 22

Comparación por materias – Docente Departamento de Ciencias de la Computación

PREGRADO S-I ABR17-AGO17	PREGRADO S-I ABR18-AGO18	PREGRADO S-II OCT17-FEB18	PREGRADO S-II OCT18-FEB19	RECOMENDADOR	AFINIDAD (%)
CLOUD COMPUTING	CLOUD COMPUTING	CLOUD COMPUTING	REDES DE DATOS II	SISTEMAS OPERATIVOS	92
FUNDAMENTOS FISICOS DE TI	REDES DE DATOS II	PROYECTO INTEGRADO R III (SIS)	REDES DE DATOS III	PLANIFICACIÓN DE TICS	85
	SISTEMAS OPERATIVOS I	REDES DE DATOS III	CLOUD COMPUTING	REDES DE DATOS III	69

Tabla 23

Comparación por área de conocimiento – Docente Departamento de Ciencias de la Computación

PREGRADO S-I ABR17- AGO17		PREGRADO S-I ABR18- AGO18		PREGRADO S-II OCT17- FEB18		PREGRADO S-II OCT18- FEB19		RECOMENDADO	AFINIDAD (%)
INFRAESTRU	INFRAESTRU	INFRAESTRU	INFRAESTRU	INFRAESTRU	INFRAESTRU	DISEÑO Y ADM DE	92		
C-TURA DE	C-TURA DE	C-TURA DE	C-TURA DE	C-TURA DE	C-TURA DE	REDES			
TICS	TICS	TICS	TICS	TICS	TICS				
				ING DE		SEGURIDAD Y	85		
				SOFTWARE		GESTIÓN DE TIC			
						INFRAESTRUC-	69		
						TURA DE TICS			

En las tablas 22 y 23 se puede apreciar que tanto en la asignación manual como en la asignación mediante el prototipo existen materias asignadas pertenecientes al área de conocimiento Infraestructura de TICs, lo cual coincide con la formación académica del docente que es doctor dentro del programa de tecnologías de la Información y sus aplicaciones, así como también coincide con su experiencia laboral desarrollada en el mismo ámbito.

Proceso de evaluación para el docente del departamento de Ciencias Exactas

En la Figura 39 se indica el registro de datos personales, académicos y de investigación del docente del departamento de Ciencias Exactas.

Figura 39

Registro de datos del docente del Departamento de Ciencias Exactas

Información Laboral Docente

Información Laboral Registradas
Para editar los datos registrados, haga clic sobre la opción deseada del listado a continuación.

[Ciencias Exactas - Nombramiento](#) **BORRAR**

campos requeridos (*)

Categoría (*): Titular Agregado 1 (1, 2 o 3) ▼

Dedicación (*): Tiempo Completo ▼

Relación Laboral (*): Nombramiento ▼

Departamento (*): Ciencias Exactas ▼

Campus (*): Matriz ▼

Horas Docencia Semanal (*):

GUARDAR

<< Atrás 2/9 Siguiente >>

El perfil que fue ingresado por el docente del caso de prueba se detalla en la Figura 40.

Figura 40*Perfil Docente Ciencias Exactas*

En resumen, tenemos lo siguiente:

- Máxima formación académica: Cuarto nivel
- Máximo título obtenido: Maestría
- Área de conocimiento: Matemática
- Años de experiencia profesional: 0
- Años de experiencia docente: 22

Luego de que el docente registró sus datos, el administrador del departamento de Ciencias Exactas accedió al prototipo para ejecutar la evaluación de afinidad mediante la interface “Evaluación de afinidad” (Figura 41), en esta interface el administrador eligió al docente de una lista desplegable y presionó el botón Evaluar, con esta acción se ejecutó el análisis a través del algoritmo implementado.

Figura 41

Evaluación de afinidad administrador departamento Ciencias Exactas

Área de conocimiento	Materia recomendada	Afinidad (%)	Descripción: afinidad
----------------------	---------------------	--------------	-----------------------

Con el objetivo de verificar que los resultados del prototipo sean adecuados, se ejecutó varias veces la evaluación de afinidad para el docente del departamento de Ciencias Exactas.

En la Figura 42 se puede apreciar los resultados de una de las varias ejecuciones realizadas para el docente del departamento Ciencias Exactas. Los datos personales fueron ocultados debido a que corresponden a datos reales.

Figura 42

Materias recomendadas docente Departamento de Ciencias Exactas

Docentes Registrados

EVALUAR

Datos generales

Nombre:

Apellido:

Cedula:

Área de conocimiento	Materia recomendada	Afinidad (%)	Descripción afinidad
EXCT-0 FISICA	FISICA II (6.0)	35	Afinidad media
EXCT-3 ANALISIS	ALGEBRA LINEAL	53	Afinidad media
EXCT-3 ANALISIS	CALCULO DIFERENCIAL E INTEGRAL	98	Afinidad alta
EXCT-5 ESTADISTICA	ESTADIST DESCRIPT ADMINIST	97	Afinidad alta
EXCT-5 ESTADISTICA	ESTADISTICA DESCRIPTIVA CIJIM	40	Afinidad media
EXCT-5 ESTADISTICA	PROBABILIDAD Y ESTADISTICA TEC	96	Afinidad alta

Figura 43

Corrida interna del algoritmo Departamento de Ciencias Exactas

```
#####
Generación: 1
Cromosoma más fuerte : Chromosome [fitnessScore=323.4, id=1104, genes=[2253, 2245, 2280, 2288, 2285, 2260]]

Codigo Materia:79
Titulo Materia:CALCULO DIFERENCIAL E INTEGRAL
Creditos: 6.0
Genres: [EXCT-3 ANALISIS]
Codigo Materia:84
Titulo Materia:FISICA II (6.0)
Creditos: 6.0
Genres: [EXCT-0 FISICA]
Codigo Materia:103
Titulo Materia:ESTADISTICA DESCRIPTIVA CHUM
Creditos: 4.0
Genres: [EXCT-5 ESTADISTICA]
Codigo Materia:127
Titulo Materia:PROBABILIDAD Y ESTADISTICA TEC
Creditos: 10.0
Genres: [EXCT-5 ESTADISTICA]
Codigo Materia:123
Titulo Materia:ESTADIST DESCRIPT ADMINIST
Creditos: 5.0
Genres: [EXCT-5 ESTADISTICA]
Codigo Materia:117
Titulo Materia:ALGEBRA LINEAL
Creditos: 5.0
Genres: [EXCT-3 ANALISIS]
#####
>>> Resultado de la evolucion: Chromosome [fitnessScore=323.4, id=1104, genes=[2253, 2245, 2280, 2288, 2285, 2260]]
```

Como se puede ver en la figura 43, el algoritmo devuelve algunos resultados:

- Generación: indica el valor de la generación más fuerte.
- Cruce de mutación (fitness): muestra la cantidad de interacciones realizadas en el proceso de selección para encontrar la población más fuerte.
- Población (genes): Corresponde a las materias que resultaron ser las que representaron los genes más fuertes.

Para evaluar los resultados obtenidos por el prototipo se realizó una comparación de las materias asignadas de forma manual en los cuatro períodos

académicos que se disponen para este estudio y las materias asignadas mediante el prototipo que mostraron afinidad alta.

Tabla 24

Comparación por materias – Docente Departamento de Ciencias Exactas

PREGRADO S-I ABR17-AGO17	PREGRADO S-I ABR18-AGO18	PREGRADO S-II OCT17-FEB18	PREGRADO S-II OCT18-FEB19	RECOMENDADO	AFINIDAD (%)
METODOS NUMERICOS	METODOS NUMERICOS	METODOS NUMERICOS	METODOS NUMERICOS	CALCULO DIFERENCIAL E INTEGRAL	98
ESTADISTICA INFERENCIAL	ESTADISTICA INFERENCIAL	LABORATORIO DE MATEMATICA	LABORATORIO DE MATEMATICA	ESTADÍSTICA DESCRIPTIVA ADMINIST	97
		ESTADISTICA INFERENCIAL	ESTADISTICA INFERENCIAL	PROBABILIDAD Y ESTADÍSTICA TEC	96
			CALC DIF E INT ADMINISTRACION		

Tabla 25

Comparación por área de conocimiento – Docente Departamento de Ciencias

Exactas

PREGRADO S-I ABR17- AGO17	PREGRADO S-I ABR18- AGO18	PREGRADO S-II OCT17- FEB18	PREGRADO S-II OCT18-FEB19	RECOMENDADO	AFINIDAD (%)
ANALISIS FUNCIONAL	ANALISIS FUNCIONAL	ANALISIS FUNCIONAL	ANALISIS FUNCIONAL	ANALISIS	98
ESTADISTICA	ESTADISTICA	ESTADISTICA	ESTADISTICA	ESTADISTICA	96

En las tablas 24 y 25 se puede apreciar que tanto en la asignación manual como en la asignación mediante el prototipo existen materias asignadas pertenecientes al área de conocimiento Estadística, lo cual coincide con la formación académica del docente que es Magister en la enseñanza de la matemática.

Adicionalmente se realizó una comparación con base a varios criterios que permitan comparar los resultados del proceso de asignación manual y del prototipo.

Tabla 26*Criterios evaluados en los resultados*

Criterio	Proceso manual	Proceso prototipo	Detalle
Tiempo de análisis, validación y asignación	Indeterminado	1 minuto	En el proceso manual el tiempo depende de la experiencia del coordinador académico para realizar la consolidación de documentos, análisis y toma de decisión, y en el prototipo se refiere al tiempo de ejecución del algoritmo para obtener las materias recomendadas.
Tiempo en detección y corrección de errores en el registro de datos	Indeterminada	Indeterminada	El tiempo del proceso manual va a depender de la agilidad en los procesos académicos para solicitar corrección a las inconsistencias encontradas en los datos de los docentes. En el prototipo corresponde al tiempo que

Criterio	Proceso manual	Proceso prototipo	Detalle
			toma al docente corregir los datos ingresados en el formulario web y volver a ejecutar el algoritmo para obtener una nueva recomendación de materias.
Seguimiento y control de resultados en tiempo real	No	Sí	<p>Con el prototipo existe la posibilidad de controlar en tiempo real los datos y obtener varias recomendaciones de materias antes de tomar una decisión.</p> <p>Mientras que de forma manual el seguimiento y control para la toma de decisión no es en tiempo real.</p>
Afinidad	No	Sí	En el proceso manual no se tienen en cuenta los factores de afinidad de los docentes, por el contrario, en el prototipo tiene una

Criterio	Proceso manual	Proceso prototipo	Detalle
			alta importancia para la asignación.
Facilidad de adaptabilidad a las necesidades cambiantes del negocio	Baja	Alta	<p>Ante cualquier cambio en normativas, leyes, reglamentos que influyan en las variables consideradas para la asignación de materias es sencilla la modificación en el prototipo y volver a realizar una nueva recomendación.</p> <p>Mientras que en el proceso manual depende que cada docente tenga en cuenta los cambios en normativas, leyes, reglamentos.</p>

Como se puede evidenciar en la comparación realizada para el proceso de asignación de carga horaria docente, con el prototipo desarrollado existe una reducción significativa de tiempo, lo cual conlleva a disminuir la carga de trabajo que originan estas actividades manuales a los señores docentes.

Además, este prototipo aporta características como:

- Seguimiento y control de resultados en tiempo real
- Considerar afinidad en el proceso de asignación de materias
- Aporta mayores facilidades de adaptabilidad a las necesidades cambiantes del negocio
- Mayor coherencia en la aplicación de normas

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

En el presente capítulo se presentan las conclusiones sobre el estudio realizado en este trabajo de graduación, así como las recomendaciones para la continuidad del estudio.

De acuerdo al análisis realizado en el presente trabajo, de la información, datos y resultados obtenidos, se llega a las siguientes conclusiones y recomendaciones.

5.1 Conclusiones

Se utilizaron herramientas de software libre y licencias educativas. Para automatizar la selección de variables se usó RapidMiner, para la limpieza e integración de datos se usó PDI, y para identificar patrones en la asignación manual de materias se usó Weka.

Se han aplicado técnicas de recogida de datos para realizar el análisis de la información. Las técnicas aplicadas fueron semiautomática y manual. La técnica semiautomática fue gestionada en la unidad de tecnología de la IES, donde se facilitaron los archivos extraídos de la base de datos institucional. La técnica manual fue ejecutada mediante la aplicación de encuestas a docentes con el fin de completar información que facilite el estudio.

Se realizó una limpieza e integración de datos mediante la aplicación de la técnica data cleaning, con la que se logró realizar el perfilamiento, normalización e identificación de valores perdidos y valores atípicos en los datos disponibles para este estudio.

Se realizó una selección de variables para identificar aquellas que son relevantes para el proceso de asignación de materias en el presente estudio, aplicando el método de clasificación K-NN.

Se identificó que el método de árboles de decisión RandomTree, fue el adecuado en el presente estudio para encontrar el comportamiento del proceso manual de asignación de materias.

Se ha propuesto una variable denominada “afinidad”, la cual se obtuvo al evaluar las reglas en el proceso de pruebas con el algoritmo implementado en el prototipo.

Se realizó el diseño e implementación de un prototipo de sistema recomendador mediante la utilización de la metodología de prototipado rápido debido a que permite ver resultados en corto tiempo, con pocos recursos y con herramientas adecuadas.

Se implementó un algoritmo genético el cual permite obtener sugerencias de materias que podrían dictar los señores docentes de acuerdo a su perfil, basado en reglas que permiten evaluar parámetros de su “afinidad” (formación académica, experiencias docente y profesional, entre otros).

Se realizaron validaciones con el prototipo desarrollado mediante casos de prueba, ejecutando el proceso desde el registro de la información del docente hasta obtener el reporte con las diferentes recomendaciones que incluyen las asignaturas y su respectivo porcentaje de afinidad. Los resultados obtenidos contribuyen a la toma de decisiones en el momento de asignar materias a cada docente de acuerdo al área de conocimiento y departamento.

Los criterios para comparar la asignación de carga horaria manual con el prototipo implementado estuvieron relacionados al tiempo, recursos y cumplimiento de normas. Respecto al tiempo, la recomendación se obtiene en minutos dependiendo de los datos y el procesamiento. Cuando la información se encuentra registrada por cada Docente, se optimizan recursos en el proceso. Además, las recomendaciones que se obtienen están basadas en la normativa vigente y es fácil de actualizar.

5.2 Recomendaciones

Sobre el estudio realizado:

Dar continuidad al desarrollo de este trabajo hasta lograr un prototipo completo, como propuesta de apoyo a la toma de decisiones respecto a la asignación de materias.

Investigar y aplicar otras técnicas usadas para obtener recomendaciones, de forma que se pueda hacer una comparación con lo propuesto en este estudio.

Gestionar el registro de datos en el formulario web con el objetivo de realizar más pruebas y comprobar su validez.

Sobre el programa de maestría:

Profundizar en el módulo que hace referencia al conocimiento estadístico de tal manera que se incluya la utilización de herramientas que faciliten el análisis e interpretación de los datos.

Poner en práctica la aplicación de técnicas de minería de datos con casos aplicables al entorno laboral.

Reforzar la redacción académica – científica con el fin de facilitar el desarrollo de los proyectos de tesis.

BIBLIOGRAFÍA

- Abdelhalim, E. A., & El Khayat, G. A. (2016). A Utilization-based Genetic Algorithm for Solving the University Timetabling Problem (UGA). *Alexandria Engineering Journal*, 55(2), 1395–1409. <https://doi.org/10.1016/j.aej.2016.02.017>
- Alvear, L. (2017). Modelo de programación entera para la asignación de materias a las aulas de la USFQ Jose Luis Alvear Calero Jose Luis Alvear Calero Carlos Iván Sandoval Vargas. *Repositorio.Usfq.Edu.Ec*, 7–34. Retrieved from <http://repositorio.usfq.edu.ec/handle/23000/6524>
- Apache. (2019). Apache Tomcat® - Welcome! Retrieved September 5, 2019, from <http://tomcat.apache.org/>
- Apache Ignite. (2019). Genetic Algorithms. Retrieved September 5, 2019, from <https://apacheignite.readme.io/docs/genetic-algorithms>
- Carpio Solano, J. A. (2003). El talento humano en las organizaciones. Retrieved November 7, 2018, from <http://www.ilo.org/public//spanish/region/ampro/cinterfor/ifp/ina/ina5.htm>
- Castro, L., Souto, Alonso, A., & Guijarro, B. (2002). Ingeniería del Conocimiento. *La Onomástica Asturiana Bajomedieval*, 99. <https://doi.org/10.1515/9783110909456-toc>

- CEAACES. (2018). Modelo de Evaluación Institucional de Universidades y Escuelas Politécnicas. ceaces. <https://doi.org/10.15517/rbt.v63i3.19892>
- CES. Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior, Pub. L. No. RPC-SO-037-No.Z65-2012, 65 (2012). Ecuador. Retrieved from http://www.ces.gob.ec/doc/gaceta_ces/reglamento/rpc-so-037-no-265-2012-b.pdf
- Corresponsables. (2019). Ley Orgánica de Protección de Datos (LOPD) | Corresponsables.com Ecuador. Retrieved September 5, 2019, from <https://ecuador.corresponsables.com/content/ley-orgánica-de-protección-de-datos-lopd>
- Domenech, B., & Lusa, A. (2016). A MILP model for the teacher assignment problem considering teachers' preferences. *European Journal of Operational Research*, 249(3), 1153–1160. <https://doi.org/10.1016/j.ejor.2015.08.057>
- ESPE. (2019). Estatuto y su codificación. *ESPE*.
- GIL, N. (2006). Algoritmos Genéticos_Kuri.pdf.
- Herrera, F. (2009). Introducción a los Algoritmos Metaheurísticos. *Soft Computing and Intelligent Information Systems*, 9–130. Retrieved from <https://sci2s.ugr.es/sites/default/files/files/Teaching/OtherPostGraduateCourses/Metaheuristicas/Int-Metaheuristicas-CAEPIA-2009.pdf>

Javier Fiallos, Jorge García, J. P. (2008). Diseño de un sistema de asignacion automatica de Horario de clases: Caso UNITEC. *LACCEI*, 2–24. Retrieved from <http://www.laccei.org/LACCEI2008-Honduras/Papers/>

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*. <https://doi.org/10.9734/bjast/2015/14975>

LOES. (2010). Ley Orgánica de Educación Superior. *Quito, Ecuador*, 1–39.

López Bruno, J. B. (2007). Modelo de asignación de carga académica usando algoritmos genéticos. *Tecnologico Nacional de Mexico*, 1–7. Retrieved from http://www.itnuevolaredo.edu.mx/maestros/sis_com/takeyas/Tesis_Maestria/articulo.PDF

Microsystem. (2019). RapidMiner | Sistemas de Minería de Datos | Software de Minería de Datos. Retrieved September 5, 2019, from <https://www.microsystem.cl/plataforma/rapidminer/>

Novalys. (2019). Funcionalidades Principales de PowerDesigner. Retrieved September 5, 2019, from <https://www.powerdesigner.biz/ES/powerdesigner/powerdesigner-features.html>

Oracle. (2019a). ¿Qué es Java y para qué es necesario? Retrieved September 5, 2019, from https://www.java.com/es/download/faq/whatis_java.xml

Oracle. (2019b). MySQL | La base de datos de código abierto más popular del mercado | Oracle España. Retrieved September 5, 2018, from <https://www.oracle.com/es/mysql/>

Román, J. V., García, R. M. C., Jesús, J., & Rueda, G. (2012). Definición y conceptos Técnicas y modelos, 39.

Velez-Langs, O., & Santos, C. (2014). Sistemas Recomendadores: Un enfoque desde los algoritmos genéticos. *Industrial Data*, 9(1), 023. <https://doi.org/10.15381/idata.v9i1.5743>

Wang, Y.-Z. Z. (2002). An application of genetic algorithm methods for teacher assignment problems. *Expert Systems with Applications*, 22(4), 302. [https://doi.org/10.1016/S0957-4174\(02\)00017-9](https://doi.org/10.1016/S0957-4174(02)00017-9)