



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL
Facultad de Ingeniería en Electricidad y Computación

“Uso de información académica histórica para predicción de
rendimiento estudiantil”

TRABAJO DE TITULACIÓN

Previo a la obtención del Título de:
MAGISTER EN CIENCIAS DE LA COMPUTACIÓN

Presentado por:
Gladys Carrillo

GUAYAQUIL - ECUADOR

AÑO: 2018

AGRADECIMIENTOS

A mi familia por su apoyo en la etapa de estudios y en todos los momentos difíciles.

A todos quienes hicieron posible que este trabajo llegue a su fin.

DEDICATORIA

Dedico este trabajo a mi hijo
Andrés, mi motivación por siempre.

TRIBUNAL DE SUSTENTACIÓN

PhD. Xavier Ochoa
Tutor

PhD. Katherine Chiliza
Co-Tutora

PhD. César Martín
Subdecano de la Facultad

PhD. Enrique Peláez
Revisor 1

PhD. Otilia Alejandro
Revisor 2

DECLARACIÓN EXPRESA

“La responsabilidad por los hechos, ideas y doctrinas expuestas en este informe me corresponde exclusivamente; y, el patrimonio intelectual de la misma, a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”

(Reglamento de Graduación de la ESPOL).

Gladys Carrillo Bastidas

RESUMEN

Por la inversión estatal en la educación superior y para mejorar la calidad de la educación universitaria, las instituciones de educación superior han puesto gran interés en reducir los fracasos académicos de sus estudiantes para lograr que avancen en su formación y completen su carrera. Varias técnicas de minería de datos se han utilizado para predecir el rendimiento estudiantil con el objetivo de que puedan ser utilizadas por las instituciones de educación superior para asistir a los estudiantes con problemas en alguna etapa de su formación universitaria. En el presente trabajo de tesis se evalúa el rendimiento de las técnicas de minería de datos: Regresión Logística, Máquina de soporte vectorial, Red Neuronal Profunda y Regresión Lineal para la predicción de rendimiento estudiantil utilizando solo la información académica de los estudiantes. Los modelos de predicción propuestos buscan predecir la aprobación de la materia y su nota promedio final. Dos conjuntos de variables se utilizan en los modelos, el primero consiste en las notas promedio finales de las materias de semestres anteriores y el otro conjunto de variables combina aspectos relacionados al semestre, a la materia y al rendimiento académico del estudiante obtenido a través de sus calificaciones. Los resultados de los experimentos de predicción de aprobación de una materia obtienen buenos valores de accuracy, siendo la red neuronal profunda la que obtiene el valor mayor con un 82%. Sin embargo, los experimentos para la predicción de calificaciones no lograron buenos resultados en la estimación del promedio final de una materia.

ÍNDICE GENERAL

RESUMEN	VI
INDICE GENERAL	VII
INDICE DE TABLAS	X
INDICE DE FIGURAS	XI
1. INTRODUCCIÓN	1
1.1. Antecedentes y Justificación	1
1.2. Objetivos	4
1.2.1. Objetivo General	4
1.2.2. Objetivos Específicos	4
1.3. Alcance	5
1.4. Organización del documento	5
2. MARCO TEÓRICO	7
2.1. Algoritmos de predicción y clasificación	7
2.2. Predicción de rendimiento estudiantil	8
2.3. Diferencias de los trabajos existentes con el trabajo propuesto	13
2.4. Métricas de evaluación	14
2.5. Medida de dificultad de la materia	15
2.6. Factores del rendimiento del estudiante	16

2.7. Herramientas de implementación (TensorFlow)	17
3. METODOLOGÍA Y DATOS	18
3.1. Recolección de datos	18
3.2. Análisis descriptivo de los datos	19
3.3. Descripción de predictores	21
3.4. Preprocesamiento de los datos	22
3.4.1. Convalidación de materias	22
3.4.2. Sustitución de valores ausentes	23
3.4.3. Preprocesamiento de predictores	23
3.5. Experimentos	24
Experimento 1: Predicción de aprobación de una mate- ria usando predictores básicos	24
Experimento 2: Predicción de aprobación de una mate- ria usando predictores extendidos	25
Experimento 3: Predicción de promedio final de una materia usando predictores básicos	25
Experimento 4: Predicción de promedio final de una materia usando predictores extendidos	25
4. RESULTADOS	26
4.1. Evaluación de los resultados	26
4.1.1. Experimento 1: Predicción de aprobación de una mate- ria usando predictores básicos	26
4.1.2. Experimento 2: Predicción de aprobación de una mate- ria usando predictores extendidos	28
4.1.3. Experimento 3: Predicción de nota promedio final de una materia usando predictores básicos	30

4.1.4. Experimento 4: Predicción de nota promedio final de una materia usando predictores extendidos	32
4.1.5. Comparativa de resultados entre experimentos	33
CONCLUSIONES	36
BIBLIOGRAFÍA	38

ÍNDICE DE TABLAS

Tabla 3.1. Datos utilizados del historial académico del estudiante	19
Tabla 3.2. Información sobre los datos utilizados en los modelos de predicción	19
Tabla 3.3. Predictores extendidos empleados en las técnicas de predicción	23
Tabla 4.1. Valores de media, mediana, desviación estándar para resultados de accuracy de los modelos de predicción utilizados en el Experimento 1	27
Tabla 4.2. Resultados de accuracy en predicción de aprobación para cinco materias con las tasas mayores de reprobación usando predictores básicos	28
Tabla 4.3. Valores de media, mediana, desviación estándar para resultados de accuracy de los modelos de predicción utilizados en el Experimento 2	29
Tabla 4.4. Resultados de accuracy en predicción de reprobación para cinco materias con las tasas mayores de fracaso usando predictores extendidos	30

ÍNDICE DE FIGURAS

3.1.	Promedio general de la carrera	20
3.2.	Promedio de número de materias tomadas por semestre .	20
3.3.	Promedio de número de materias perdidas por semestre .	21
4.1.	Resultados de accuracy para los modelos de predicción utilizados en el Experimento 1	27
4.2.	Resultados de accuracy para los modelos de predicción utilizados en el Experimento 2	29
4.3.	Resultados de error cuadrático medio (MSE) para Experi- mento 3	31
4.4.	Comparación de promedios reales y estimados para la materia Fundamentos de Programación con la Red Neu- ronal Profunda y Regresión Lineal usando predictores básicos	31
4.5.	Comparación de promedios reales y estimados para la materia Estructuras de Datos con la Red Neuronal Pro- funda y Regresión Lineal usando predictores básicos . . .	32
4.6.	Resultados de error cuadrático medio (MSE) para Experi- mento 4	33

4.7.	Comparación de promedios reales y estimados para la materia Fundamentos de Programación con la Red Neuronal Profunda y Regresión Lineal usando predictores extendidos	33
4.8.	Comparación de promedios reales y estimados para la materia Estructuras de Datos con la Red Neuronal Profunda y Regresión Lineal usando predictores extendidos .	34
4.9.	Comparación de los valores de accuracy en la predicción de aprobación de materias para experimentos 1 y 2	34
4.10.	Comparación de los valores de error cuadrático medio para predecir nota promedio final	35

CAPÍTULO 1

1. Introducción

1.1. Antecedentes y Justificación

En los últimos años en el Ecuador se ha puesto gran interés en mejorar la calidad de la educación universitaria ecuatoriana para satisfacer la demanda de profesionales requeridos en las áreas de interés de desarrollo productivo del país[1].

El rendimiento académico de los estudiantes es un factor fundamental en la medición de la calidad de la educación universitaria[2]. Este indicador, medido a través de las calificaciones obtenidas y reflejando las materias aprobadas y perdidas[3], permite estimar la situación actual de los estudiantes frente al currículum académico de la carrera que esta cursando.

En la educación superior existen diversas variables que influyen

en el rendimiento académico tales como relacionadas al entorno personal, factores socio-familiar y entorno académico[2][3]. Entre las variables relacionadas al entorno personal se encuentran la edad, el género, el estado civil, los problemas de salud o discapacidad y la situación laboral del estudiante. Las variables relacionadas con los factores socio-familiares son el el estrato social, el apoyo afectivo que recibe de su familia, el ingreso familiar, las actividades recreativas y la cantidad de amigos. Las variables relacionadas al entorno académico y universitario incluyen el promedio de calificaciones del bachillerato, el tipo de universidad, los resultados de pruebas vocacionales, el nivel de dedicación del estudiante, el número de asignaturas cursadas, la relación con el profesor, la carga curricular y el promedio general. Cualquiera de las variables mencionadas puede tener un impacto negativo sobre el rendimiento académico si no son manejados adecuadamente, por lo tanto, es importante que los directivos de las universidades estén conscientes de los efectos de los factores y busquen estrategias para detectar a tiempo cualquier situación que ponga en riesgo los logros académicos de sus estudiantes.

Además, por la inversión estatal en la educación superior, las universidades están llamadas a brindar excelente formación académica y reducir los fracasos académicos de su estudiantado, lograr que avancen en su formación y completen su carrera. Las universidades se verían beneficiadas con mecanismos de predicción del rendimiento académico de sus estudiantes para poder identificar a aquellos que tienen problemas de rendimiento académico en alguna etapa de su formación universitaria para poder asistirlos.

El uso de técnicas de minería de datos ha sido explorado por instituciones de educación superior en la predicción del rendimiento académico [4]. La minería de datos es el proceso de descubrir patrones en grandes conjuntos de datos [5] utilizando diversos métodos como estadística, inteligencia artificial y aprendizaje automático. La minería de datos aplicada a la educación o EDM ¹ es un área activa de investigación [6] y está relacionada a la aplicación de métodos de minería en conjuntos de datos originados en entornos educativos, para descubrir nuevos conocimientos sobre cómo los estudiantes aprenden [7].

Una de las aplicaciones más antiguas y populares de la minería de datos en la educación ha sido la predicción del rendimiento estudiantil, empleando para ello diversas técnicas entre las que se encuentran las redes bayesianas [8], los árboles de decisión [9] y máquinas de soporte vectorial (SVM)[10].

Las aplicaciones de las técnicas ya sea para predecir la nota promedio final o predecir aprobación de una materia, pueden ser parte de herramientas de software que utilizan las universidades para dar apoyo a sus estudiantes. Por ejemplo, el sistema de consejerías académicas que está orientado a proveer a los alumnos el apoyo de un tutor –quien realiza recomendaciones en base a su experiencia y al historial académico del estudiante–, podría incluir información adicional generada a través de técnicas de predicción que permitiría al tutor contar con datos útiles para adecuar la estrategia de recomendación para minimizar el riesgo de reprobación de una materia. También la información de predicción podría ser parte de nuevos sistemas

¹Educational Data Mining

o sistemas existentes que le permitan a un profesor consultar en cualquier momento el riesgo que tiene un estudiante de reprobado una materia.

1.2. Objetivos

1.2.1. Objetivo General

Proponer un modelo para predecir el rendimiento académico de un estudiante en una materia mediante el uso de su historial académico y el de sus pares.

1.2.2. Objetivos Específicos

Los objetivos específicos del presente trabajo son los siguientes:

1. Diseñar una metodología para predecir la aprobación de una materia usando solo el historial académico del estudiantes.
2. Diseñar una metodología para predecir la nota promedio final de una materia usando solo el historial académico del estudiante.
3. Implementar las metodologías usando las técnicas de regresión logística, máquina de soporte vectorial, redes neuronales profundas y regresión lineal.
4. Evaluar el rendimiento de las técnicas al ser aplicadas al mismo conjunto de datos.

5. Comparar los resultados de rendimiento entre las diferentes técnicas.

1.3. Alcance

El alcance del presente trabajo es el uso de técnicas de minería de datos en la educación para proponer un modelo para predecir el rendimiento académico usando sólo el historial académico del estudiante y sus pares. No se consideran las variables personales y socio-familiares de los estudiantes como parte de las variables empleadas en las técnicas de predicción.

1.4. Organización del documento

- En el Capítulo 1 se exponen los antecedentes y la justificación del presente trabajo, además se plantean los objetivos que se pretenden alcanzar.
- En el Capítulo 2 se hace una revisión de los algoritmos de predicción y de los trabajos existentes relacionados al uso de minería de datos en la educación para la predicción del rendimiento académico a nivel universitario.
- En el Capítulo 3 se explica el proceso de recolección de datos, el diseño de los experimentos y se presentan las variables de predicción a utilizar.
- En el Capítulo 4 se exponen los resultados de los experimentos y se comparan los valores de rendimiento obtenidos por cada una de las técnicas empleadas en cada experimento.

- Finalmente se presentan las conclusiones del trabajo realizado.

CAPÍTULO 2

2. Marco Teórico

2.1. Algoritmos de predicción y clasificación

Una máquina de soporte vectorial(SVM)[11] es un algoritmo de aprendizaje de máquina supervisado usado principalmente para problemas de clasificación. Es una representación de los ejemplos como puntos en el espacio de n dimensiones (donde n es el número de variables), mapeado de manera que los ejemplos de cada clase se dividan por hiperplanos que definan una brecha clara lo más amplia posible. SVM tiene ventajas en optimización global, capacidad de generalización y aprendizaje con muestras de tamaño pequeño.

La regresión logística es un método estadístico para analizar un conjunto de datos y predecir una salida (variable dependiente con solo dos valores posibles) en función de las variables independientes. El

objetivo de la regresión logística es encontrar el mejor modelo de ajuste para describir la relación entre la variable de interés(dependiente) y el conjunto de variables independientes, utilizando una función logística. Una ventaja de la regresión lineal es que no requiere la normalidad estricta de los datos.

Una red neuronal se refiere a un conjunto de neuronas o unidades conectadas entre sí que funcionan en paralelo para completar una tarea. Cuando la red neuronal tiene varias capas se denomina red neuronal profunda. La idea principal del uso de una red neuronal profunda es que con más capas intermedias –con más neuronas cada una– se pueden mejorar las predicciones en conjuntos de datos más complejos[12].

La regresión lineal es un método estadístico para aproximar la relación entre una variable dependiente y una o más variables independientes. La variable dependiente debe tener un valor continuo. Cuando el modelo lineal involucra varias variables dependientes –con el fin de mejorar la precisión en la estimación– se denomina modelo de regresión lineal múltiple.

2.2. Predicción de rendimiento estudiantil

La predicción del rendimiento estudiantil es un tema que ha sido objeto de estudio durante varios años. Varias técnicas como modelos estadísticos o matemáticos, minería de datos y aprendizaje automático se utilizan para analizar el comportamiento de los estudiantes y predecir el éxito o fracaso durante sus estudios. Las técnicas pueden aplicarse para predecir notas en una materia, promedio general, o

para predecir la aprobación de materias, semestres o carrera. A continuación se presenta un revisión de trabajos de predicción de rendimiento estudiantil a nivel universitario.

Un modelo con regresión logística es utilizado por Sule y Saporu [13] para la predicción del riesgo de reprobación del curso Cálculo Elemental de la Facultad de Ciencias de la universidad de Maiduguri (Nigeria). Las variables utilizadas para la predicción incluyen el “GPA” del estudiante y variables relacionadas a la motivación del estudiante y su actitud hacia el curso, las cuales fueron recolectadas mediante una encuesta realizada a los estudiantes. El modelo de regresión logística pudo realizar la predicción con un 70.8% de precisión.

Otro modelo con regresión logística es utilizado por Urrutia et al. [14] para la predicción del rendimiento académico durante el primer año de la carrera de medicina de la Universidad Nacional Autónoma de México. El modelo utiliza 72 variables generadas a partir de los resultados de una encuesta realizada a los estudiantes para conocer sobre su historial académico, factores sociales, su interés en la carrera y sus estrategias de estudio. El modelo puede predecir con un 71% si los estudiantes aprueban todas las materias o falla alguna de ellas al final del primer año de estudios.

Strecht et al.[15] evaluaron los algoritmos de clasificación k-Nearest Neighbors (kNN), Random Forest (RF), AdaBoost(AB), Clasificación and Regression Trees (CART) y Máquina de Soporte Vectorial(SVM) para predecir la aprobación de los estudiantes en una materia. Los experimentos se realizaron utilizando la información de alrededor de 700 materias de diversos programas de la Universidad de Porto (Portugal). Las variables utilizadas en el proceso de predic-

ción están relacionadas al estudiante como “*edad*”, “*género*”, “*estado civil*”, “*nacionalidad*”, “*beca*”, “*necesidades especiales*”, “*tipo de admisión*”, entre otras; la variable a predecir era la aprobación de la materia. Los resultados muestran que SVM y los árboles de decisión tienen un mejor rendimiento con más del 60 % de precisión, aunque no existe una diferencia estadística significativa entre los demás algoritmos.

Ramanathan [16] propone un algoritmo basado en árboles de decisión mediante una variación del algoritmo ID3 al asignar pesos a cada atributo al momento de cada toma de decisión. Para la evaluación de su propuesta utilizaron datos académicos de 304 estudiantes de la Escuela de Ciencias Computacionales de la Universidad VIT (India), para predecir la aprobación de las materias de un semestre en el año 2011. Las variables utilizadas incluían la “*edad*”, “*género*”, “*las notas de la educación secundaria*” y “*CGPA*”¹. Para medir el rendimiento de su propuesta compararon los resultados de ID3 con los algoritmos J48 y Naïve Bayes, logrando el algoritmo modificado de ID3 un rendimiento mayor con un valor de 93 % de precisión sobre el 78.6 % para J48 y 75 % para Naïve Bayes.

Un modelo usando prototipos difusos y clasificación no supervisada es propuesto por Vásquez et al. [17] para predecir el riesgo que tiene un estudiante de reprobar un semestre. Las variables utilizadas describen la carga académica del semestre mediante medidas referentes a cada materia como la rigurosidad de calificación y la dificultad de aprobación; además incluyen cinco factores que describen las habilidades del estudiante cuando aprueban una materia. Con este conjunto de características se realiza el proceso de clustering para

¹Cumulative Grade Point Average

clasificar de manera no supervisada y predecir el riesgo de reprobación del semestre en base a las características de estudiantes con semestres similares. En la evaluación el modelo propuesto tuvo una tasa de acierto superior al 75 % y al ser comparado con un modelo de regresión logística obtuvo un valor de 0.01 de confiabilidad.

Sharabiani et al. [18] proponen un modelo de predicción basado en redes bayesianas (BN) con el objetivo de obtener un pronóstico de las notas de los estudiantes en tres cursos principales de la carrera de ingeniería de la Universidad de Illinois en Chicago. Las notas de los cursos Física General II, Cálculo II y Programación C/C++ que se toman en el segundo semestre, tienen gran impacto en la retención de los estudiantes ya que son considerados cursos difíciles y la mayoría obtiene notas bajas. El modelo para la predicción consiste en incluir como variables principales la información demográfica (edad, género, raza, estado civil), el GPA del estudiante en el primer semestre y las notas de las materias requisitos de las asignaturas que se van a predecir. En el proceso de configuración de la red bayesiana también definen nuevas variables intermedias para indicar la capacidad del estudiante en cada semestre y el nivel de dificultad de cada materia. El rendimiento del modelo con la red bayesiana fue comparado con otros modelos: DT (árboles de decisión), NB (clasificador bayesiano) y KNN (k vecinos más próximos). El modelo propuesto obtiene un mejor rendimiento en la predicción de las notas para las materias de Física II y Cálculo II; sin embargo, no pudo superar a los modelos de árboles de decisión en la predicción de la nota para la materia de Programación C/C++.

Ahmad et al. [19] proponen un framework para predecir el GPA del

primer año de los estudiantes de la Carrera de Ciencias Computacionales de UniSZA (Malasia), evaluando las técnicas de clasificación Naive Bayes, Árboles de decisión y Basado en reglas. Las variables utilizadas para la predicción incluyen el género, ingreso familiar, modo de ingreso a la universidad y las notas de tres materias: Lengua, Inglés y Matemáticas. La evaluación se realizó utilizando la herramienta WEKA con los datos de 497 estudiantes. Los resultados evidencian que la clasificación basada en Reglas tiene mayor rendimiento que los otros dos algoritmos, logrando un valor de precisión de 71.3%.

Arsad et al. [20] describen un modelo basado en una red neuronal para predecir el CGPA de los estudiantes al final de la carrera de Ingeniería Eléctrica de UiTM (Malasia). En el trabajo también se realiza una comparación de la red neuronal con un modelo de Regresión Lineal utilizando para la evaluación el MSE (Error cuadrático medio) y el coeficiente de Correlación R. Las variables utilizadas en el modelo incluyen el CGPA en cada semestre, el promedio de las notas de las materias de cada semestre y el promedio de los cursos de inglés. El análisis incluye también una comparación entre los efectos de las materias principales y los cursos de inglés en la nota final de la carrera. Los resultados muestran que los modelos de predicción obtuvieron resultados similares en el error cuadrático medio y que existe una fuerte correlación entre las notas de las materias principales y el CGPA final, mientras que los cursos de inglés tienen poco efecto.

Tekin [21] implementó modelos con Redes Neuronales, Máquinas de soporte vectorial y Máquinas de aprendizaje extremo para la predicción del GPA de graduación. Los datos utilizados en el análisis

corresponden a las notas de 127 estudiantes en 47 materias de la carrera de Educación informática de la Universidad Firat (Turquía). Para la evaluación de los modelos se utilizaron los métodos: media cuadrática (RMSE), el coeficiente de correlación (R) y el coeficiente de variación (COV). Los tres modelos de predicción tuvieron un buen desempeño en la predicción del GPA final con valores de precisión superior al 80 %. SVM fue el método que obtuvo la mejor precisión llegando a obtener un coeficiente de correlación de 93.06 %[21].

2.3. Diferencias de los trabajos existentes con el trabajo propuesto

Como se puede ver existe variedad de modelos que combinan diferentes variables personales, sociales y académicos. Como en algunas universidades no se puede tener toda la información de los estudiantes actualizada y disponible, algunos autores han empleado encuestas o cuestionarios para recolectar la información que les permite construir las variables utilizadas en la predicción. Ya que la elaboración y ejecución de estos instrumentos agrega preprocesamiento extra al proceso de predicción, los modelos propuestos en el presente trabajo solo hacen uso de la información académica disponible de los estudiantes, la misma que corresponde a las notas promedio de las materias cursadas durante la carrera universitaria.

Los trabajos revisados construyen modelos utilizando técnicas populares como árboles de decisión y SVM que generalmente están disponibles en herramientas como WEKA y SPSS. Hasta donde conocemos, no se ha encontrado algún trabajo que haga uso de redes

neuronales profundas para la predicción del rendimiento estudiantil, quizás porque se sabe que las redes neuronales profundas son utilizadas generalmente en áreas donde se maneja grandes cantidades de datos como por ejemplo el reconocimiento de imágenes y reconocimiento de voz. En el presente trabajo se evalúa un modelo con una red neuronal profunda utilizando el framework TensorFlow, debido a las ventajas que ofrece esta herramienta para la construcción de modelos para aprendizaje automático.

Los modelos de predicción de notas en los trabajos examinados se utilizan principalmente para la predicción de calificaciones finales de materias específicas o para la predicción de GPA en un semestre, en el primero año o al finalizar la carrera. No se observa que se definan modelos que puedan ser aplicables en cualquier etapa de la carrera o para cualquier materia. El modelo de predicción de notas propuesto en el presente trabajo busca realizar la predicción de la nota promedio final de cualquier materia y explorar como las notas de materias pasadas pueden ayudar a predecir las notas de cursos futuros.

2.4. Métricas de evaluación

Accuracy es una métrica utilizada para evaluar un clasificador y proporciona el porcentaje de ejemplos clasificados correctamente. Su fórmula está dada por[22]:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FP} \quad (2.1)$$

donde:

TP = verdaderos positivos.

TN = verdaderos negativos.

FP = falsos positivos.

FN = falsos negativos.

El error cuadrático medio (MSE, por sus siglas en inglés) es una métrica utilizada en un modelo de predicción para medir la dispersión del error de pronóstico, proporciona la diferencia entre el estimador y el valor estimado. Su fórmula está dada por[23]:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.2)$$

donde:

Y es el vector de valores reales.

\hat{Y} es el vector de n predicciones.

2.5. Medida de dificultad de la materia

La medida de dificultad de la materia (α) introducida por Caulkins et al. [24] está definida por:

$$\alpha_j = \frac{\sum_i GPA_i^2}{\sum_i (r_{ij} \cdot GPA_i)} \quad (2.3)$$

donde:

α_j es el nivel de dificultad de la materia j

GPA_i es el rendimiento académico general del estudiante i

r_{ij} es la nota obtenida por el estudiante i en la materia j

2.6. Factores del rendimiento del estudiante

El rendimiento del estudiante se representa mediante el análisis de factores propuesto por Mendez et. al [25], estos cinco factores describen las habilidades que adquieren los estudiantes al aprobar cada materia. Los factores son:

1. Factor de preparación básica. Comprende dos categorías. La primera categoría está relacionada con los cursos de ciencias básicas de diferentes áreas como Matemáticas, Física, Química y Estadística. La segunda categoría agrupa cursos fundamentales de Ciencias Computacionales como Fundamentos de Programación, Estructura de Datos, Análisis de Algoritmos, Computación y Sociedad
2. Factor de temas avanzados de Ciencias Computacionales. Aquí se agrupan los cursos relacionados a la identificación de los componentes de un sistema y sus relaciones. Entre los cursos se encuentran Programación Orientada a Objetos, Arquitectura y Organización de Computadoras, Sistemas de Bases de datos, Ingeniería de Software, entre otras.
3. Factor de interacción con el cliente . Está relacionado a los cursos que cubren aspectos referente a las habilidades de comunicación necesarias para el trato con los clientes en el proceso de desarrollo de software. Los cursos agrupados en este factor son Ingeniería de Software I y II, Interacción Humano Computadora y Técnicas de Comunicación Oral y Escrita.
4. Factor de programación. Comprende los cursos dedicados ex-

clusivamente a desarrollar habilidades de programación como Estructuras de datos, Programación Orientada a Objetos y Lenguajes de Programación.

5. Factor de cursos no estrechamente relacionados con las Ciencias Computacionales. Se incluyen los cursos que están más relacionados a materias de Ingeniería Eléctrica y que son vistos por los estudiantes como cursos que no son aplicados directamente para resolver problemas de ciencias computacionales.

2.7. Herramientas de implementación (TensorFlow)

TensorFlow es un framework de Google liberado como software de código abierto en 2015 que provee un conjunto de herramientas para el manejo de algoritmos de inteligencia artificial y aprendizaje profundo. Tensorflow puede correr en CPU, GPUs, en Linux de 64 bits, macOS y plataformas móviles que incluyen Android e iOS. Está escrito en C++ y ofrece una API de Python, C++, Java y Go. El modelo de programación de Tensorflow simplifica la complejidad del manejo de algoritmos de aprendizaje profundo, es altamente escalable y con un rendimiento mucho mayor [26].

CAPÍTULO 3

3. Metodología y Datos

3.1. Recolección de datos

La información fue recolectada de los datos académicos históricos de 2214 estudiantes de la carrera de Ingeniería en Ciencias Computacionales de la ESPOL registrados entre 1998 y 2016. Los datos académicos utilizados se detallan en la tabla 3.1. Esta información académica corresponde a las 26 materias que conforman la malla curricular vigente desde el año 2009, como lo presenta Méndez et al. [25].

	Campo	Descripción
1	Año	Año en que la materia fue cursada Término en que la materia fue cursada.
2	Término	Existen tres términos: el primero de mayo a septiembre, el segundo de octubre a febrero y el tercero de marzo a abril
3	Código del estudiante	Número de identificación del estudiante
4	Código de materia	Código de identificación de la materia
5	Promedio	Nota final promedio de la materia

Tabla 3.1: Datos utilizados del historial académico del estudiante

3.2. Análisis descriptivo de los datos

En esta sección se describen los datos académicos históricos con la finalidad de conocer el tamaño de los datos recolectados y resumir la actividad académica del estudiante.

La tabla 3.2 muestra información sobre el conjunto de datos utilizados para las técnicas de predicción.

	Cantidad
Total de estudiantes	2214
Hombres	1663
Mujeres	551
Número de semestres	54
Número de materias	26
Notas promedio finales	93552

Tabla 3.2: Información sobre los datos utilizados en los modelos de predicción

A continuación se muestra la distribución de los valores para el promedio general de la carrera 3.1, el promedio de número materias tomadas por semestre 3.2 y el promedio de materias perdidas por semestre 3.3.

Como se puede ver en la figura 3.1 el promedio general de la

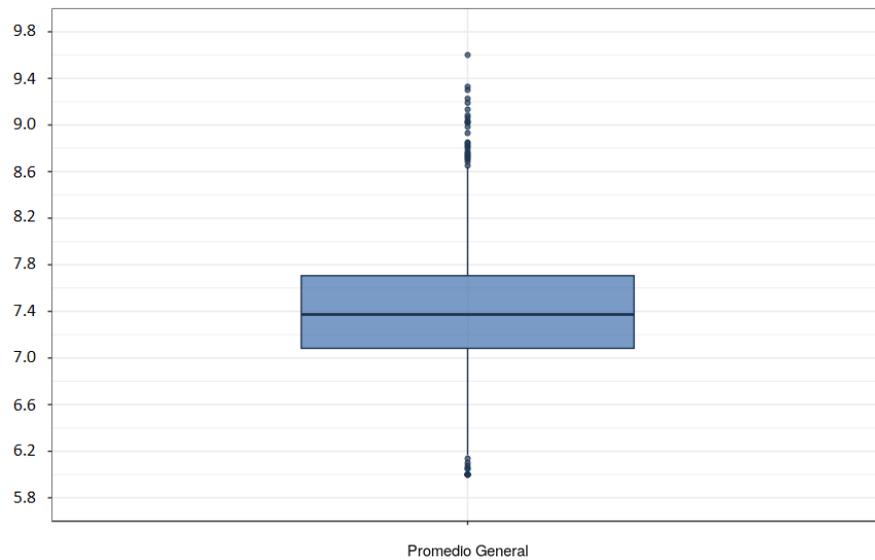


Figura 3.1: Promedio general de la carrera

carrera es de 7.4, la mitad de la muestra tiene un promedio entre 7.1 y 7.7. Existen estudiantes que destacan con promedios superiores a 8.8 y muy pocos con promedio bajo 6.2.

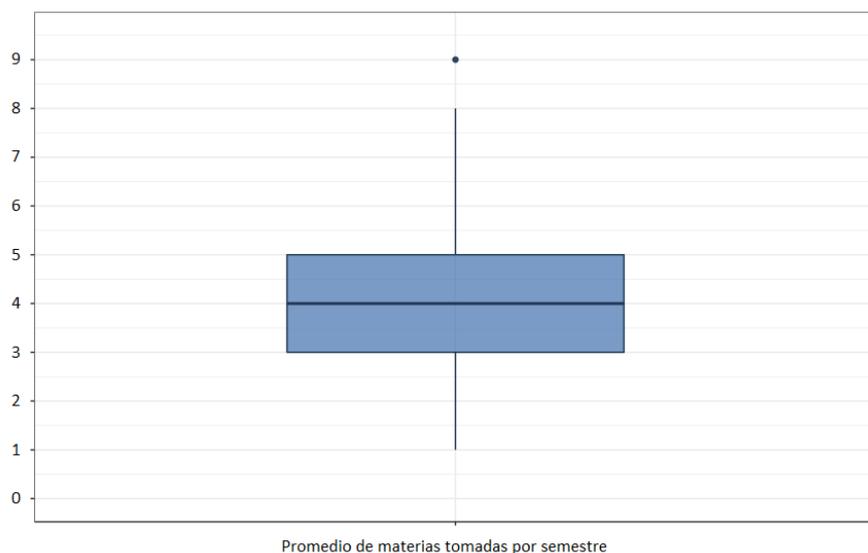


Figura 3.2: Promedio de número de materias tomadas por semestre

La figura 3.2 muestra que el 50% de los estudiantes toman de 3

a 5 materias por semestre, el 25 % toma más de 5 materias y el otro 25 % menos de 3 materias.

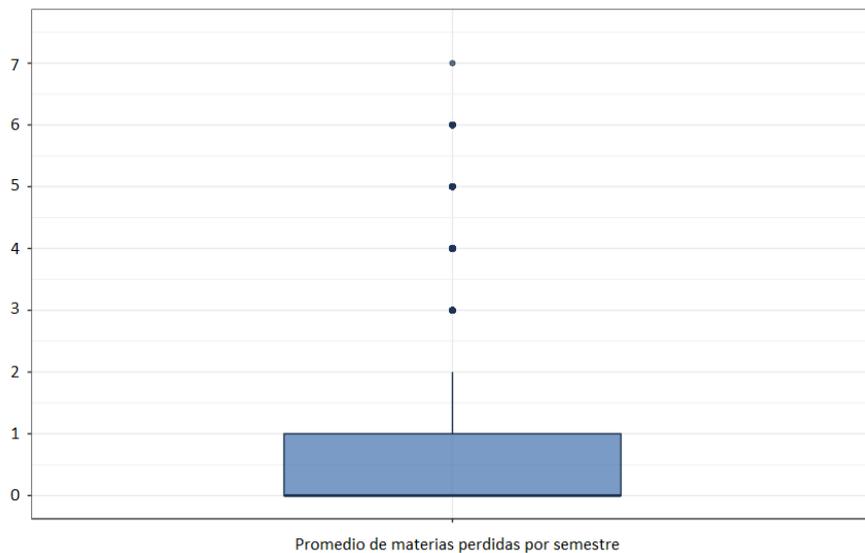


Figura 3.3: Promedio de número de materias perdidas por semestre

En la figura 3.3 se muestra que aproximadamente la mitad de la muestra tiene cero materias perdidas por semestre, el 25 % pierde una materia y el otro 25 % dos materias, existiendo casos de estudiantes que llegan a perder todas las 6 o 7 materias tomadas por semestre.

El porcentaje considerable de estudiantes que pierden al menos una materia cada semestre, afirma la importancia de buscar mecanismos de apoyo como la predicción de rendimiento estudiantil para minimizar el riesgo de reprobación.

3.3. Descripción de predictores

Para los análisis de predicción se utilizan dos conjuntos de predictores: básicos y extendidos.

Los predictores básicos corresponden a la nota promedio final de cada uno de cursos tomados por el estudiante antes de la materia que se analiza.

Los predictores extendidos incluyen:

1. El número del semestre en que se toma la materia, un valor entre 1 y 10 para una carrera con finalización de 5 años.
2. El número de materias tomadas en el periodo académico que se analiza.
3. Información de la materia. La información de la materia está representada por el promedio de notas de todos los estudiantes que han cursado la materia y la medida de dificultad de la materia (α).
4. El rendimiento del estudiante durante sus estudios medido a través de los 5 factores presentados en la sección 2.6.

La tabla 3.3 resume los predictores extendidos empleados en las técnicas de predicción.

3.4. Preprocesamiento de los datos

3.4.1. Convalidación de materias

Un proceso de convalidación de materias se realizó con el fin de evitar la pérdida de información debido a los cambios efectuados en la malla de la carrera de Ingeniería en Ciencias Computacionales durante el periodo al que pertenecen los

	Variable	Descripción
1	promedio de materia	Promedio de notas de todos los estudiantes en esa materia
2	α de materia	medida de dificultad de la materia
3	factor 1	Factor de preparación en ingeniería básica
4	factor 2	Factor de temas avanzados de Ciencias Computacionales
5	factor 3	Factor de interacción con el cliente
6	factor 4	Factor de programación
7	factor 5	Factor de cursos no estrechamente relacionados con las Ciencias Computacionales.
8	carga semestral	Número de materias registradas en el semestre
9	número de semestre	Número del semestre en que se toma la materia

Tabla 3.3: Predictores extendidos empleados en las técnicas de predicción

datos académicos utilizados. El proceso consistió en identificar materias que hayan cambiado de nombre o que se hayan dividido en dos o más cursos, para encontrar la materia correspondiente en la malla académica vigente.

3.4.2. Sustitución de valores ausentes

El GPA del estudiante es utilizado para cubrir valores ausentes de los factores en los predictores extendidos, principalmente para el análisis de las materias en los primeros niveles de la malla.

3.4.3. Preprocesamiento de predictores

Para los predictores extendidos, se realiza el proceso de estandarización de rango para que cada dato tenga un valor entre 0 y 1. Este proceso se realizó debido a que las

variables “*carga semestral*”, “*número de semestre*” y “*alpha de la materia*” tienen un rango diferente a las variables “*factores*” y “*promedio*” cuyos valores van de 0 a 10. La fórmula utilizada para reescalar es:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (3.1)$$

3.5. Experimentos

Los experimentos que se detallan a continuación buscan analizar el comportamiento de cuatro técnicas de minería de datos para la educación con los dos conjuntos de predictores explicados en la sección 3.3. Cada conjunto de predictores se utiliza para predecir la aprobación de una materia y su nota promedio final. Para evaluar el rendimiento y la aplicabilidad de las técnicas para predecir el rendimiento estudiantil se utilizan las métricas de “*accuracy*” y “*error cuadrático medio*” de acuerdo al tipo de predicción que se realiza.

Experimento 1: Predicción de aprobación de una materia usando predictores básicos

El primer experimento consiste en predecir la aprobación de una materia utilizando predictores básicos. Se entrenaron tres modelos de clasificación: Máquinas de soporte vectorial (SVM), Regresión Logística y Red Neuronal Profunda. Para la evaluación del rendimiento de los clasificadores se utiliza la métrica de *accuracy*.

Experimento 2: Predicción de aprobación de una materia usando predictores extendidos

Este experimento consiste en predecir la aprobación de una materia utilizando predictores extendidos. Se entrenaron también tres modelos de clasificación: Máquinas de soporte vectorial (SVM), Regresión Logística y Red Neuronal Profunda. Para la evaluación del rendimiento de los clasificadores se utiliza la métrica de accuracy.

Experimento 3: Predicción de promedio final de una materia usando predictores básicos

El tercer experimento consiste en predecir la nota promedio final que un estudiante obtiene en una materia usando los predictores básicos. Para este experimento se entrenó una Red Neuronal Profunda y un modelo de Regresión Lineal. Los resultados se compararon mediante el error medio cuadrático (MSE).

Experimento 4: Predicción de promedio final de una materia usando predictores extendidos

Este experimento consiste en predecir la nota promedio final que un estudiante obtiene en una materia utilizando predictores extendidos. Para este experimento se entrenó una Red Neuronal Profunda y un modelo de Regresión Lineal y los resultados se compararon mediante el error medio cuadrático (MSE).

CAPÍTULO 4

4. Resultados.

4.1. Evaluación de los resultados

A continuación se detallan los resultados de la evaluación de los experimentos implementados en TensorFlow utilizando el API de Python.

4.1.1. Experimento 1: Predicción de aprobación de una materia usando predictores básicos

La figura 4.1 muestra la distribución de los resultados de accuracy de las tres técnicas de predicción aplicadas al conjunto de materias de la malla curricular presentes en el primer término del año 2016. Como se puede observar en la tabla 4.1, todos los modelos de predicción obtienen un valor

medio de accuracy superior a 0.75, aunque el modelo con la Red Neuronal Profunda obtiene mejores valores de accuracy reflejados en los valores de media, mediana y desviación estándar.

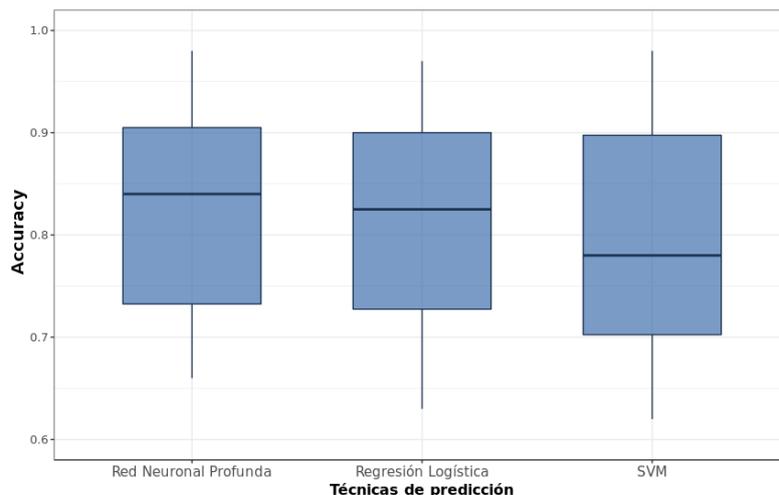


Figura 4.1: Resultados de accuracy para los modelos de predicción utilizados en el Experimento 1

	Media	Mediana	Desviación Estándar
SVM	0.79	0.78	0.12
Regresión Logística	0.81	0.83	0.10
Red Neuronal Profunda	0.82	0.84	0.10

Tabla 4.1: Valores de media, mediana, desviación estándar para resultados de accuracy de los modelos de predicción utilizados en el Experimento 1

La tabla 4.2 muestra los resultados de accuracy en la predicción de reprobación para cinco materias de la categoría capacitación profesional de la malla curricular con los porcentajes más altos de reprobación. A pesar de que el valor de accuracy de las tres técnicas en la predicción general (aprobación/reprobación) se encuentra alrededor de 0.75, los valores de accuracy para reprobación son bajos, alrededor

de 0.40. Estos valores de accuracy en la clasificación de reprobación, afectan el valor de accuracy general que para todas las técnicas está por debajo de la media como se muestra en la tabla 4.1.

Nombre de Materia	SVM	Regresión Logística	Red Neuronal Profunda
Fundamentos de Programación	AP/RP: 0.68 RP: 0.37	AP/RP: 0.69 RP: 0.31	AP/RP: 0.69 RP: 0.39
Estructuras de datos	AP/RP: 0.75 RP: 0.43	AP/RP: 0.75 RP: 0.42	AP/RP: 0.75 RP: 0.54
Lenguajes de Programación	AP/RP: 0.75 RP: 0.44	AP/RP: 0.74 RP: 0.41	AP/RP: 0.75 RP: 0.48
Programación Orientada a Objetos	AP/RP: 0.74 RP: 0.43	AP/RP: 0.75 RP: 0.44	AP/RP: 0.76 RP: 0.51
Métodos de Investigación aplicados a la Computación	AP/RP: 0.78 RP: 0.71	AP/RP: 0.79 RP: 0.44	AP/RP: 0.81 RP: 0.51

Tabla 4.2: Resultados de accuracy en predicción de aprobación para cinco materias con las tasas mayores de reprobación usando predictores básicos

4.1.2. Experimento 2: Predicción de aprobación de una materia usando predictores extendidos

La figura 4.2 muestra la distribución de los resultados de accuracy de las tres técnicas de predicción utilizando los nueve predictores extendidos. Como se puede observar en la tabla 4.3, todos los modelos de predicción obtienen un valor medio de accuracy superior al 0.80, siendo el modelo con la Red Neuronal Profunda el que obtiene mejores valores para esta métrica.

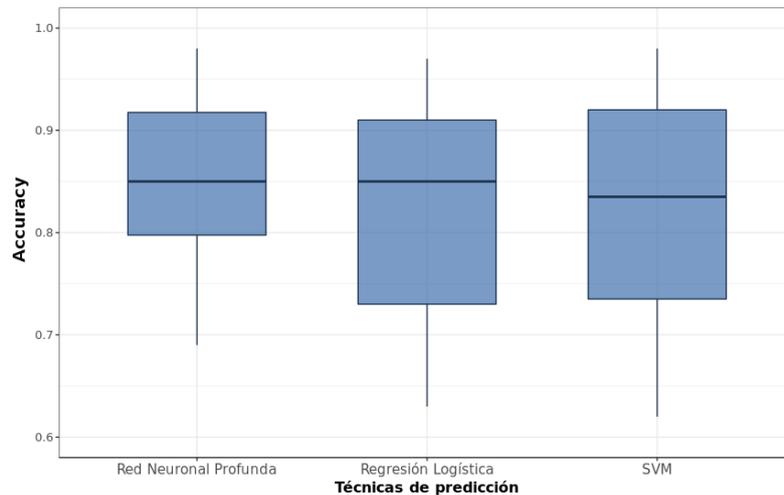


Figura 4.2: Resultados de accuracy para los modelos de predicción utilizados en el Experimento 2

	Media	Mediana	Desviación Estándar
SVM	0.82	0.84	0.11
Regresión Logística	0.82	0.85	0.11
Red Neuronal Profunda	0.85	0.85	0.08

Tabla 4.3: Valores de media, mediana, desviación estándar para resultados de accuracy de los modelos de predicción utilizados en el Experimento 2

La tabla 4.4 muestra los resultados de accuracy en la predicción de aprobación para las cinco materias con los porcentajes más altos de reprobación. Como se puede observar, se evidencia la diferencia de la técnica de Red Neuronal Profunda sobre las técnicas de SVM y Regresión Logística, tanto en el valor de accuracy general de aprobación y reprobación como en el valor de accuracy en la predicción de reprobación.

Nombre de Materia	SVM	Regresión Logística	Red Neuronal Profunda
Fundamentos de Programación	AP/RP: 0.68 RP: 0.39	AP/RP: 0.69 RP: 0.46	AP/RP: 0.77 RP: 0.77
Estructuras de datos	AP/RP: 0.83 RP: 0.67	AP/RP: 0.76 RP: 0.42	AP/RP: 0.85 RP: 0.76
Lenguajes de Programación	AP/RP: 0.73 RP: 0.42	AP/RP: 0.76 RP: 0.43	AP/RP: 0.84 RP: 0.62
Programación Orientada a Objetos	AP/RP: 0.77 RP: 0.50	AP/RP: 0.75 RP: 0.49	AP/RP: 0.80 RP: 0.58
Métodos de Investigación aplicados a la Computación	AP/RP: 0.77 RP: 0.49	AP/RP: 0.85 RP: 0.49	AP/RP: 0.92 RP: 0.75

Tabla 4.4: Resultados de accuracy en predicción de reprobación para cinco materias con las tasas mayores de fracaso usando predictores extendidos

4.1.3. Experimento 3: Predicción de nota promedio final de una materia usando predictores básicos

La figura 4.3 muestra la distribución del error cuadrático medio (MSE) de la predicción realizada con la Red Neuronal Profunda y con Regresión Lineal. Como se puede ver en la figura 4.3 la Red Neuronal Profunda tiene valores mayores de error comparado con la técnica de Regresión Lineal, pero ninguno de los dos métodos obtiene valores menores a 1.

Para visualizar mejor el comportamiento de las técnicas de predicción se muestran las figuras 4.4 y 4.5 que corresponden a la predicción de notas promedio para dos de las materias con mayor porcentaje de reprobación: Fundamentos de Programación y Estructuras de datos. Como se puede

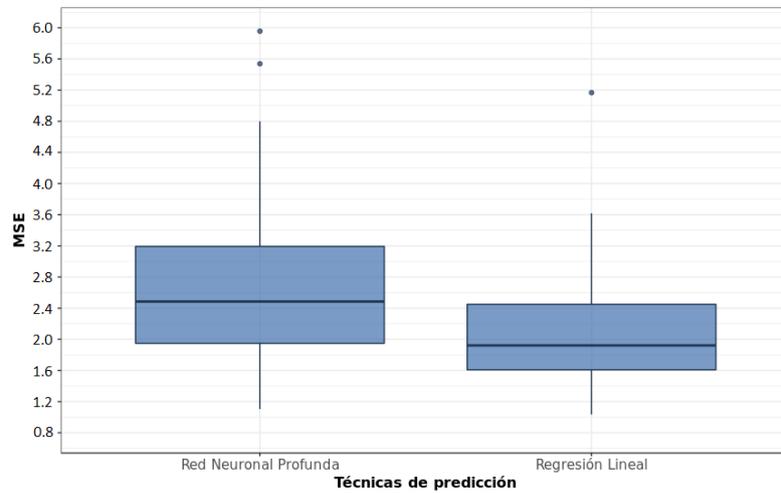


Figura 4.3: Resultados de error cuadrático medio (MSE) para Experimento 3

observar en las figuras muchas de las notas promedio estimadas se alejan de la nota promedio real, especialmente en la predicción de notas promedio bajas.

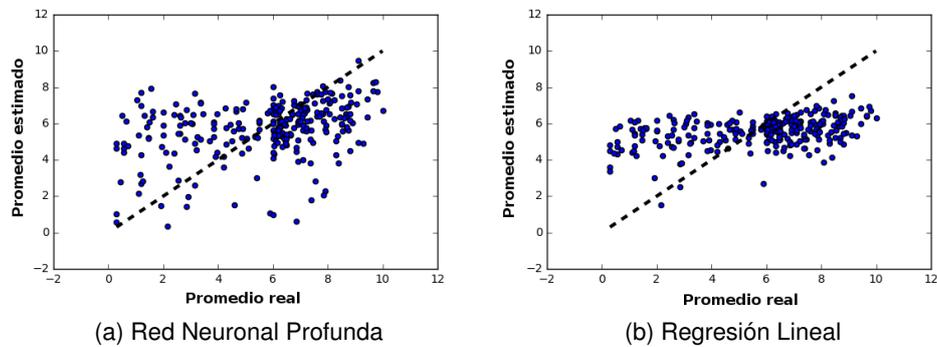


Figura 4.4: Comparación de promedios reales y estimados para la materia Fundamentos de Programación con la Red Neuronal Profunda y Regresión Lineal usando predictores básicos

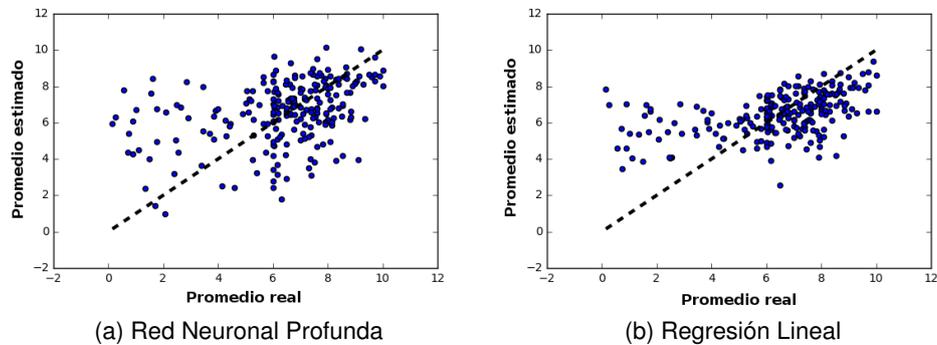


Figura 4.5: Comparación de promedios reales y estimados para la materia Estructuras de Datos con la Red Neuronal Profunda y Regresión Lineal usando predictores básicos

4.1.4. Experimento 4: Predicción de nota promedio final de una materia usando predictores extendidos

La figura 4.6 muestra la distribución del error cuadrático medio (MSE) de la predicción realizada con la Red Neuronal Profunda y con Regresión Lineal. Como se puede ver en la figura la técnica de Regresión Lineal tiene valores menores de MSE en relación a la Red Neuronal Profunda, con el 25 % de los resultados con valores por debajo de 1.2.

Las figuras 4.7 y 4.8 muestran una comparación de los promedios reales y los promedios estimados para las materias Fundamentos de Programación y Estructuras de datos. Como se puede ver en las figuras los valores estimados no guardan relación con los valores reales, incluso los resultados de Regresión Lineal parecen concentrarse horizontalmente entre los valores 4 y 8.

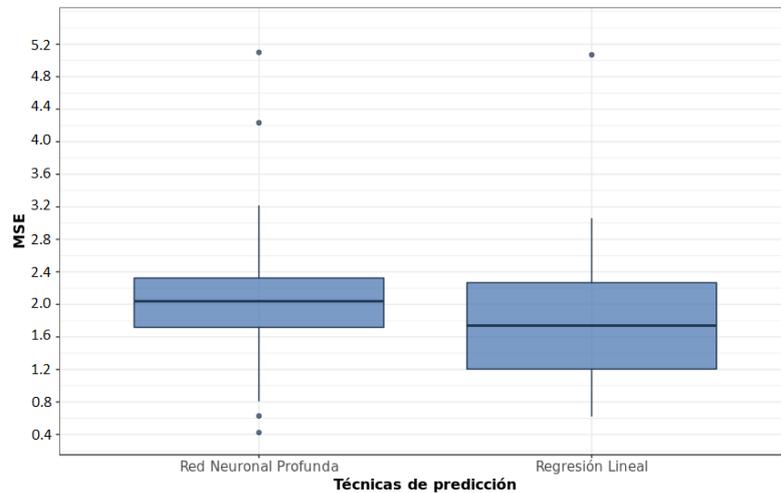


Figura 4.6: Resultados de error cuadrático medio (MSE) para Experimento 4

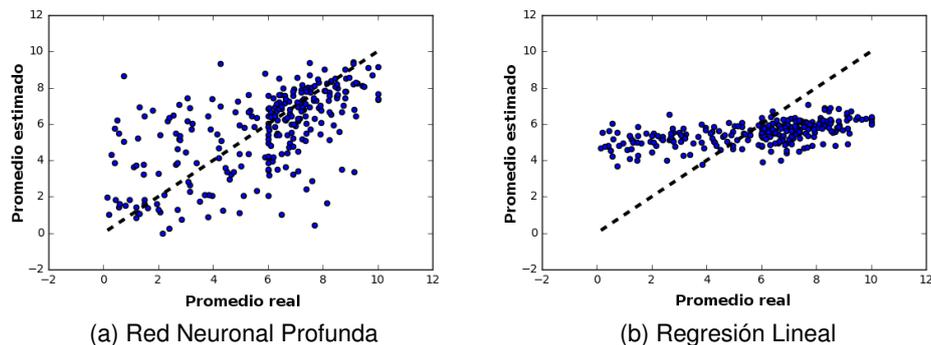


Figura 4.7: Comparación de promedios reales y estimados para la materia Fundamentos de Programación con la Red Neuronal Profunda y Regresión Lineal usando predictores extendidos

4.1.5. Comparativa de resultados entre experimentos

La figura 4.9 muestra la distribución de los valores de accuracy para los modelos de clasificación utilizados en los experimentos 1 y 2.

Como se puede observar los valores de accuracy para el experimento 2 donde se utilizan los predictores extendidos son mayores a los valores del experimento 1. La mayor

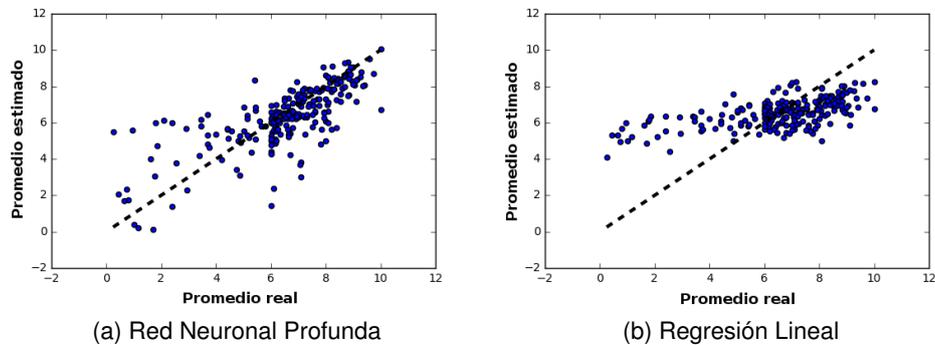


Figura 4.8: Comparación de promedios reales y estimados para la materia Estructuras de Datos con la Red Neuronal Profunda y Regresión Lineal usando predictores extendidos

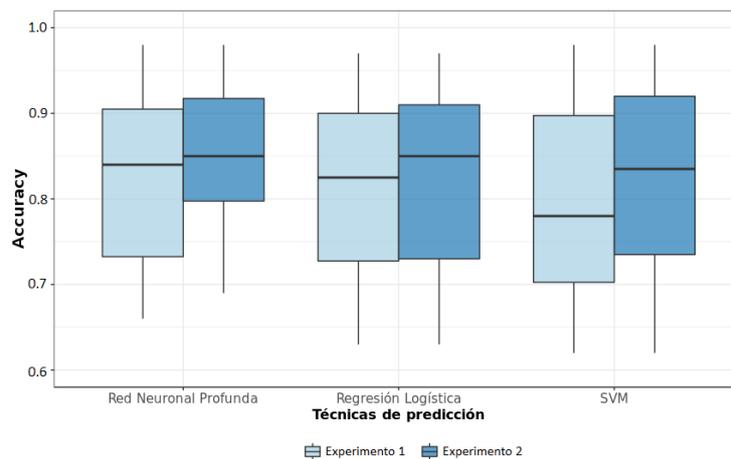


Figura 4.9: Comparación de los valores de accuracy en la predicción de aprobación de materias para experimentos 1 y 2

diferencia se evidencia en la Red Neuronal Profunda donde el 75 % resultados son mayores que 0.8.

La figura 4.10 muestra la distribución de los valores de error cuadrático para los experimentos 3 y 4. Como se puede observar los valores de error para el experimento 4 que utiliza los predictores extendidos son menores a los resultados del experimento 3 que usa los predictores básicos.

La principal reducción del error se evidencia en los resultados del experimento 4 con la Red Neuronal Profunda donde el 75% de los valores son menores que 2.4. De las figuras 4.4, 4.5, 4.7 y 4.8 se puede observar que en los experimentos 3 y 4 los promedios estimados para las materias de Fundamentos de Programación y Estructuras de datos no guardan relación con los promedios reales. A pesar que los valores de MSE mejoran en el experimento 4, la predicción de los promedios finales conservan el mismo comportamiento que los resultados del experimento 3.

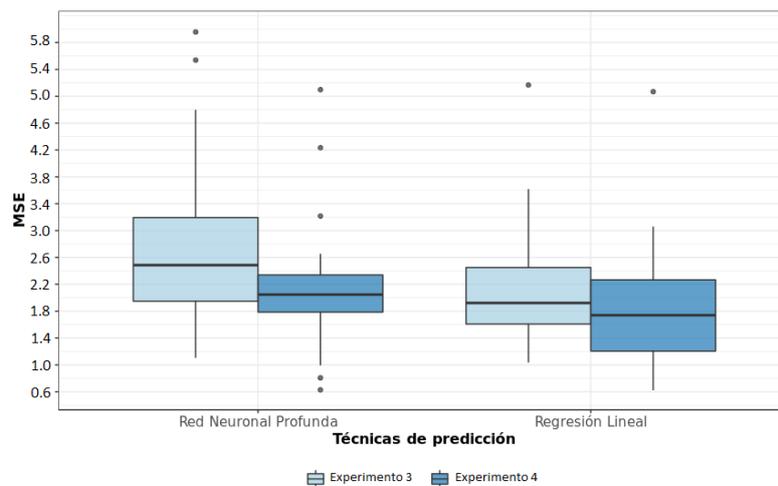


Figura 4.10: Comparación de los valores de error cuadrático medio para predecir nota promedio final

CONCLUSIONES

En el presente trabajo se ha examinado el uso de técnicas de minería de datos para la predicción del rendimiento estudiantil a nivel universitario. Principalmente se explora el uso de una máquina de soporte vectorial, regresión logística, una red neuronal profunda y regresión lineal, sobre un conjunto de datos académicos de los estudiantes para la predicción de aprobación de materias y su nota promedio final. Luego del desarrollo de la tesis las conclusiones son las siguientes:

1. Es posible obtener buenos resultados de predicción de aprobación de materias utilizando calificaciones de materias cursadas en semestres anteriores, por medio de la implementación de las técnicas SVM, red neuronal profunda y regresión logística siendo el modelo implementado con la red neuronal profunda el que obtiene mejor resultado de accuracy con el 82 %.
2. Mejores resultados de precisión para predecir la aprobación de materias se pueden obtener mediante el uso de predictores extendidos que corresponden a variables que combinan información de la materia y

rendimiento del estudiante. Con este modelo, la red neuronal profunda también presenta un mejor rendimiento con el 85 % de accuracy.

3. Para predecir la reprobación de las materias, los resultados muestran que el modelo con predictores extendidos es mejor al modelo con predictores básicos y que la red neuronal profunda es la única técnica que puede dar mejores resultados de accuracy.
4. Aunque en los dos modelos para predecir la aprobación de materias los resultados son prometedores, no ocurre lo mismo para la predicción de la nota promedio final, ya que los resultados muestran que a pesar de que el error de predicción es menor con los predictores extendidos, las notas que se predicen no guardan relación con los promedios reales, por lo que se podría explorar con otros predictores para tratar de reducir la variabilidad de los resultados.

BIBLIOGRAFÍA

- [1] Jaime E Rojas. Reforma universitaria en el Ecuador. etapa de transición. *Innovación Educativa*, 11(57), 2011.
- [2] Guiselle María Garbanzo Vargas. Factores asociados al rendimiento académico en estudiantes universitarios, una reflexión desde la calidad de la educación superior pública. *Educación*, 31(1), 2007.
- [3] A Pérez, J Ramón, and J Sánchez. Análisis exploratorio de las variables que condicionan el rendimiento académico. *Sevilla: Universidad Pablo de Olavide*, 2000.
- [4] Alejandro Peña-Ayala. Educational data mining: A survey and a data mining-based analysis of recent works. *Expert systems with applications*, 41(4):1432–1462, 2014.
- [5] Oded Maimon and Lior Rokach. *Data mining and knowledge discovery handbook*, volume 2. Springer, 2005.
- [6] Amirah Mohamed Shahiri, Wahidah Husain, et al. A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, 72:414–422, 2015.

- [7] Cristobal Romero and Sebastian Ventura. Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1):12–27, 2013.
- [8] Nir Friedman, Dan Geiger, and Moises Goldszmidt. Bayesian network classifiers. *Machine learning*, 29(2-3):131–163, 1997.
- [9] Sreerama K Murthy. Automatic construction of decision trees from data: A multi-disciplinary survey. *Data mining and knowledge discovery*, 2(4):345–389, 1998.
- [10] Christopher JC Burges. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167, 1998.
- [11] Corinna Cortes and Vladimir Vapnik. Support vector machine. *Machine learning*, 20(3):273–297, 1995.
- [12] Simon Haykin. *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1994.
- [13] BO Sule and FWO Saporu. A logistic regression model of students' academic performance in university of maiduguri, maiduguri, nigeria.
- [14] María Esther Urrutia-Aguilar, Ruth Fuentes-García, Vargas Danae Mirel Martínez, Edward Beck, Silvia Ortiz León, and Rosalinda Guevara-Guzmán. Logistic regression model for the academic performance of first-year medical students in the biomedical area. *Creative Education*, 7(15):2202, 2016.
- [15] Pedro Strecht, Luís Cruz, Carlos Soares, João Mendes-Moreira, and Rui Abreu. A comparative study of classification and regression algorithms for modelling students' academic performance. *International Educational Data Mining Society*, 2015.

- [16] L Ramanathan, Saksham Dh, and Suresh Kumar. Predicting students' performance using modified id3 algorithm. 2013.
- [17] Aníbal Vásquez, Enrique Peláez, and Xavier Ochoa. Predictor basado en prototipos difusos y clasificación no-supervisada. *Revista Latinoamericana de Ingeniería de Software*, 3(3):135–140, 2015.
- [18] Ashkan Sharabiani, Fazle Karim, Anooshiravan Sharabiani, Mariya Atanasov, and Houshang Darabi. An enhanced bayesian network model for prediction of students' academic performance in engineering programs. In *Global Engineering Education Conference (EDUCON), 2014 IEEE*, pages 832–837. IEEE, 2014.
- [19] Fadhilah Ahmad, Nur Hafieza Ismail, and Azwa Abdul Aziz. The prediction of students' academic performance using classification data mining techniques. *Applied Mathematical Sciences*, 9(129):6415–6426, 2015.
- [20] Pauziah Mohd Arsad, Norlida Buniyamin, et al. Neural network and linear regression methods for prediction of students' academic achievement. In *Global Engineering Education Conference (EDUCON), 2014 IEEE*, pages 916–921. IEEE, 2014.
- [21] Ahmet Tekin. Early prediction of students' grade point averages at graduation: A data mining approach. *Eurasian Journal of Educational Research*, 54:207–226, 2014.
- [22] David L Olson and Dursun Delen. *Advanced data mining techniques*. Springer Science & Business Media, 2008.
- [23] Dennis Wackerly, William Mendenhall, and Richard Scheaffer. *Mathematical statistics with applications*. Nelson Education, 2007.

- [24] Jonathan P Caulkins, Patrick D Larkey, and Jifa Wei. Adjusting gpa to reflect course difficulty. 1996.
- [25] Gonzalo Méndez, Xavier Ochoa, and Katherine Chiluiza. Techniques for data-driven curriculum analysis. In *Proceedings of the fourth international conference on learning analytics and knowledge*, pages 148–157. ACM, 2014.
- [26] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *OSDI*, volume 16, pages 265–283, 2016.