

Main Variables in Teaching Quality: A Data Mining Approach

Omar Danilo Castrillón, Ph.D.¹ , odcastrillong@unal.edu.co

¹Universidad Nacional de Colombia, Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Industrial, Campus La Nubia Bloque Q piso 2, Manizales – Colombia.

Abstract – This research is based on a student perception survey that was applied to students about their university professors. This survey consists of 1 dependent variable (teaching quality) and 20 independent variables: 1) time dedicated to the subject, 2) reasons for enrollment, 3) adequate class preparation, 4) adequate mentorship, 5) fostering of relationships with other subjects, 6) pleasant work methods, 7) encouragement of argumentation, 8) promotion of autonomous learning, 9) motivation of the topics covered, 10) respectfulness, 11) attention that everyone learned, 12) adaption of teaching methods, 13) respect of the rules, 14) fairness, 15) students learnt with evaluations, 16) depth of learning, 17) the professor transformed their way of thinking in regards to the course, 18) the professor taught the students another course, 19) learning was facilitated by the professor's strengths, 20) the professor endeavored for students to learn. Through the J48 algorithm of the free machine learning and data mining platform called Weka, the main variables are identified. As a result of this research, three groups of students are found, with three different sets of independent variables capable of predicting the teaching quality. In the first group of students, with an effectiveness greater than 90%, the following variables are found: 3, 12, 2, 14, 19, 16, 10. In the second group of students, with an effectiveness greater than 80%, the following variables are found: 19, 16, 3. Finally, in the third group, with an effectiveness greater than 75%, the following variables are established: 19, 4.

Keywords-- Student perception, teaching quality, student surveys, data mining, Weka.

Principales Variables en la Calidad Docente: Un Enfoque desde la Minería de Datos

Main Variables in Teaching Quality: A Data Mining Approach

Omar Danilo Castrillón, Ph.D.¹,  odcastrillong@unal.edu.co

¹Universidad Nacional de Colombia, Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Industrial, Campus La Nubia Bloque Q piso 2, Manizales – Colombia.

Resumen– Esta investigación, parte de una encuesta de la percepción estudiantil que los estudiantes tienen sobre los docentes universitarios. Esta encuesta consta de una variable dependiente (calidad docente) y 20 variables independientes: 1) Tiempo dedicado a la asignatura, 2) motivo de la inscripción, 3) preparación adecuada de clases, 4) asesoría adecuada, 5) propicio relaciones con otras asignaturas, 6) trabajo de forma agradable, 7) promovió la argumentación, 8) promovió el aprendizaje autónomo, 9) motivo los temas tratados, 10) fue respetuoso, 11) se preocupó por que cada uno aprendiera, 12) adecuo sus métodos de enseñanza, 13) respeto las reglas del juego, 14) fue justo, 15) se aprendió en sus evaluaciones, 16) se aprendió con profundidad, 17) transformó su forma de pensar, 18) vería otra asignatura con este docente, 19) las fortalezas del docente le ayudaron a aprender, 20) hizo todo lo posible por que aprendieran. Por medio del algoritmo J48 de la plataforma libre de aprendizaje automático y minería de datos denominada Weka, se identifican las principales variables más importantes. Como resultado de esta investigación se encuentra tres grupos de estudiantes, con tres conjuntos diferentes de variables independientes capaces de predecir la calidad docente. En el primer grupo de estudiantes, con una efectividad superior al 90% se encuentran las siguientes variables: 3, 12, 2, 14, 19, 16, 10. En el segundo grupo de estudiantes, con una efectividad superior al 80%, se encuentran las siguientes variables: 19, 16, 3. Finalmente, en el tercer grupo, con una efectividad superior al 75%, se establecen las siguientes variables: 19, 4.

Palabras claves– Percepción estudiantil, calidad docente, encuestas estudiantiles, minería de datos, Weka.

Abstract – This research is based on a student perception survey that was applied to students about their university professors. This survey consists of 1 dependent variable (teaching quality) and 20 independent variables: 1) time dedicated to the subject, 2) reasons for enrollment, 3) adequate class preparation, 4) adequate mentorship, 5) fostering of relationships with other subjects, 6) pleasant work methods, 7) encouragement of argumentation, 8) promotion of autonomous learning, 9) motivation of the topics covered, 10) respectfulness, 11) attention that everyone learned, 12) adaption of teaching methods, 13) respect of the rules, 14) fairness, 15) students learnt with evaluations, 16) depth of learning, 17) the professor transformed their way of thinking in regards to the course, 18) the professor taught the students another course, 19) learning was facilitated by the professor's strengths, 20) the professor endeavored for students to learn. Through the J48 algorithm of the free machine learning and data mining platform called Weka, the main variables

are identified. As a result of this research, three groups of students are found, with three different sets of independent variables capable of predicting the teaching quality. In the first group of students, with an effectiveness greater than 90%, the following variables are found: 3, 12, 2, 14, 19, 16, 10. In the second group of students, with an effectiveness greater than 80%, the following variables are found: 19, 16, 3. Finally, in the third group, with an effectiveness greater than 75%, the following variables are established: 19, 4.

Keywords– Student perception, teaching quality, student surveys, data mining, Weka.

I. INTRODUCCIÓN

Los resultados de evaluaciones estudiantiles son empleados en muchos contextos para evaluar la calidad de la enseñanza [1], variable que en este contexto será denotada como desempeño docente, entendido este como la percepción que un estudiante tienen sobre un docente. En diversos estudios, la anterior variable (desempeño docente) está regulada por el comportamiento de una serie de variables independientes (<https://edificando.unal.edu.co/portal/docentes>) las cuales presentan una incidencia directa sobre la variable dependiente. Este conjunto de variables independientes es: Tiempo dedicado a la asignatura, motivo de la inscripción, preparación adecuada de clases, asesoría adecuada, propicio relaciones con otras asignaturas, trabajo de forma agradable, promovió la argumentación, promovió el aprendizaje autónomo, motivo los temas tratados, fue respetuoso, se preocupó por que cada uno aprendiera, adecuo sus métodos de enseñanza, respeto las reglas, fue justo, se aprendió en sus evaluaciones, se aprendió con profundidad, transformó su forma de pensar, vería otra asignatura con él, las fortalezas del docente le ayudaron a aprender, hizo todo lo posible por que aprendieran

No obstante, no todos los estudiantes les dan la misma importancia a las variables independientes. Existe diversos grupos de estudiantes los cuales le dan más importancia a unas determinadas variables independientes que a otras. Siendo de gran trascendencia identificar las variables más relevantes por cada grupo de estudiantes, con el fin de que el desempeño docente puede ser medido de la forma más objetiva posible. Este trabajo investigativo es desarrollado por medio del algoritmo J48 de la plataforma libre de aprendizaje automático

y minería de datos denominada Weka (algoritmo de comportamiento bayesiano) [2]. En esta plataforma se consideran 20 variables independientes (ilustradas en el párrafo anterior) y una variable dependiente denominada desempeño docente, evaluada en una escala cualitativa de cuatro valores: pésimo, malo, bueno o excelente. Siendo el objetivo fundamental de esta artículo identificar los diferentes grupos de estudiantes e identificar las diferentes variables independientes (por cada grupo de estudiantes) conducentes a una valoración objetiva de la variable dependiente.

Se resalta que, en este campo de estudio, la inteligencia artificial y en especial los árboles de decisión tienen una amplia aplicabilidad en la calidad docente [3], e incluso el desempeño docente puede ser comparado aun entre diferentes países, sin embargo, es necesario tener cuidado sobre la forma como este proceso debe ser realizado [4], siendo importante considerar otros aspectos diferentes a la percepción estudiantil [5]. No obstante, este aspecto será analizado en futuras líneas de investigación. Diferentes aspectos relacionados con la calidad docente pueden ser encontrados en las diversas revisiones literarias: [6,7,8]. Finalmente, para una comprensión este documento ha sido organizado de la siguiente forma: Metodología, Resultados, Discusiones y Referencias.

II. METODOLOGIA

Con el fin de definir la metodología [9], se define una base de datos tomando como referencia la estructura de la encuesta empleada en <https://edificando.unal.edu.co/portal/docentes>. Con esta base de datos se hace una medición (número de encuestas) estadísticamente válida. Lo anterior permite estructurar los siguientes pasos de la metodología: 1. Estructuración y aplicación de la encuesta. 2. Estructuración de la base de datos. 3. Construcción de los respectivos archivos Weka. 4. Selección de los grupos de estudiantes. 5. Árboles de decisión y variables estadísticas.

Paso 1. Estructuración y aplicación de la encuesta.

Partiendo de la encuesta referenciada en el párrafo anterior, se estructura la encuesta que será aplicada en una población estudiantil, estadísticamente válida. La tabla 1, ilustra la encuesta definida.

Con base en la siguiente fórmula se determina el número de encuestas que debe ser aplicada [10]:

$$n = \frac{I^2 pqM}{ME^2 + I^2 pq} \quad \text{Ec (1)}$$

Dónde: n: Tamaño de la muestra; I: Intervalo de confianza, para una distribución normal estándar invertida; p y q: representan las variaciones éxito y fracaso, respectivamente; M: Población de estudiantes; E: Error muestral.

TABLA I
ENCUESTA – SATISFACCIÓN DOCENTE

#	PREGUNTA	1	2	3	4	5
P1	Cuál fué el tiempo que le dedicó a la asignatura.? 1. Muy bajo 2. Bajo. 3. El Justo 4. Alto 5. Muy Alto					
P2	Cual fué el motivo de inscripción de la asignatura. ? 1. Obligatoria. 2. No había otra disponible. 3. Por que era optativa. 4. Porque era electiva. 5. Por el docente					
P3	El/La docente preparó adecuadamente las clases. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P4	La asesoría fué adecuada. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P5	El/La docente propició relaciones con otras asignaturas y/o contextos. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P6	El/La docente trabajó de forma agradable. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P7	El/La docente promovió la argumentación y reflexión crítica. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P8	El/La docente promovió el aprendizaje autónomo. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P9	El/La docente motivó los temas tratados. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P10	El/La docente fué respetuoso. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P11	El/La docente se preocupó por que cada uno aprenda. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P12	El/La docente adecuó sus métodos de enseñanza. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P13	El/La docente respetó las reglas del juego. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P14	El/La docente fué justo. ? 1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.					
P15	Se aprendió en sus evaluaciones. ? 1. No 5. Si					
P16	Se aprendió con profundidad. ? 1. No 5. Si					
P17	El/La docente Transformó su forma de pensar con respecto a la asignatura. ? 1. No 5. Si					
P18	Vería otra asignatura con este docente. ? 1. No. 5. Si					
P19	Las fortalezas del docente le ayudaron a aprender. ? 1. No 5. Si					
P20	El/La docente hizo todo lo posible por que aprendiera. ? 1. No 5. Si					
P21	El desempeño del docente fué. ? 1. Pésimo 2. Malo 4. Bueno 5. Excelente					

Paso 2. Estructuración de la base de datos.

Con las encuestas realizadas en la fase anterior, se define una base de datos conforme la estructura ilustrada en la Tabla 1. En la Tabla 2, las variables P1 – P20 son las variables independientes, mientras la variable P21 es la variable dependiente.

TABLA II
ESTRUCTURACIÓN BASE DE DATOS

#	Nombre	Escala
P1	Cuál fué el tiempo que le dedicó a la asignatura.?	1. Muy bajo 2. Bajo. 3. El Justo 4. Alto 5. Muy Alto
P2	Cual fué el motivo de inscripción de la asignatura. ?	1. Obligatoria. 2. No había otra disponible. 3. Por que era optativa. 4. Porque era electiva. 5. Por el docente
P3	El/La docente preparó adecuadamente las clases. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P4	La asesoría fué adecuada. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P5	El/La docente propició relaciones con otras asignaturas y/o contextos. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P6	El/La docente trabajó de forma agradable. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P7	El/La docente promovió la argumentación y reflexión crítica. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P8	El/La docente promovió el aprendizaje autónomo. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P9	El/La docente motivó los temas tratados. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P10	El/La docente fué respetuoso. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P11	El/La docente se preocupó por que cada uno aprenda. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P12	El/La docente adecuó sus métodos de enseñanza. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P13	El/La docente respetó las reglas del juego. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P14	El/La docente fué justo. ?	1. Nunca 2. Rara vez. 3. Algunas veces. 4. Casi siempre. 5. Siempre.
P15	Se aprendió en sus evaluaciones. ?	1. No 5. Si
P16	Se aprendió con profundidad. ?	1. No 5. Si
P17	El/La docente Transformó su forma de pensar con respecto a la asignatura. ?	1. No 5. Si
P18	Vería otra asignatura con este docente. ?	1. No. 5. Si
P19	Las fortalezas del docente le ayudaron a aprender. ?	1. No 5. Si
P20	El/La docente hizo todo lo posible por que aprendiera. ?	1. No 5. Si
P21	El desempeño del docente fué. ?	1. Péximo 2. Malo 4. Bueno 5. Excelente

Paso 3. Construcción de los respectivos archivos Weka.

Tomando como referencia la tabla 2 definida en la sección anterior, se elaboran los respectivos archivos ARFF, cabeza y detalle, con el fin de que los mismos sean interpretados desde la plataforma de aprendizaje automático y minería de datos denominada Weka [11].

Paso 4. Selección de los grupos de estudiantes.

Por medio del algoritmo J48 se interpretan los archivos anteriores con el fin de lograr la clasificación de la variable dependiente J21. Con base en esta clasificación se identifican

los registros que fueron clasificados erróneamente por el algoritmo J48. Con los registros erróneos se construirán dos nuevos archivos ARFF (cabeza y detalle) con el fin de realizar una nueva clasificación. El proceso descrito en este paso, se repete hasta que no sea posible seleccionar nuevos grupos.

Paso 5. Arboles de decisión y variables estadísticas.

Para cada uno de los archivos ARFF definidos en la sección anterior se define el respectivo árbol de decisiones (un árbol por cada grupo), el cual muestra (para cada grupo) las diferentes interrelaciones de las variables independientes conducentes a cada uno de los estados específicos de la variable dependiente.

III. RESULTADOS

Como consecuencia de aplicar cada una de los diferentes pasos de la metodología, se obtienen los siguientes resultados:

Paso 1. Estructuración y aplicación de la encuesta.

La ejecución de la Ecuación 1, establece un tamaño de muestra de $n = 121$. Este resultado, es obtenido bajo los siguientes parámetros: $M = 460$; $I = 1.2815$, intervalo de confianza, obtenido con la función de Excel para distribución normal estándar invertida (0.9); $p = 0.5$; $q = 0.5$; $E = 0,05$ (error muestral del 5%, con un intervalo del 90%). No obstante, en la realización de esta investigación se aplicaron 253 encuestas. Un tamaño muestral estadísticamente valido.

Paso 2. Estructuración de la base de datos.

Con los resultados del paso 1 (encuestas elaboradas), se define la base de datos, ilustrada en la siguiente Tabla 3 y 4: (Por razones de espacio solo se ilustra un fragmento)

TABLA III.
BASE DE DATOS. PARTE I

P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
2	1	4	3	1	4	2	5	2	5	2
4	5	4	4	3	5	4	4	5	5	5
.

TABLA IV. BASE DE DATOS.
PARTE II

P12	P13	P14	P15	P16	P17	P18	P19	P20	P21
2	4	5	1	1	1	1	1	1	4
4	5	5	5	5	1	5	5	5	5
.
4	2	4	1	1	1	1	1	1	2
4	3	3	5	1	5	1	5	5	4

Paso 3. Construcción de los respectivos archivos Weka.

Como resultado de esta fase se obtienen los respectivos archivos de clasificación para la plataforma Weka (cabeza y detalle). Las tablas 5, 6 y 7 ilustran estos archivos:

TABLA V.
CABEZA ARCHIVO ARFF

Variables	Rango
@attribute P1	{1,2,3,4,5}
@attribute P2	{1,2,3,4,5}
@attribute P3	{1,2,3,4,5}
@attribute P4	{1,2,3,4,5}
@attribute P5	{1,2,3,4,5}
@attribute P6	{1,2,3,4,5}
@attribute P7	{1,2,3,4,5}
@attribute P8	{1,2,3,4,5}
@attribute P9	{1,2,3,4,5}
@attribute P10	{1,2,3,4,5}
@attribute P11	{1,2,3,4,5}
@attribute P12	{1,2,3,4,5}
@attribute P13	{1,2,3,4,5}
@attribute P14	{1,2,3,4,5}
@attribute P15	{1,5}
@attribute P16	{1,5}
@attribute P17	{1,5}
@attribute P18	{1,5}
@attribute P19	{1,5}
@attribute P20	{1,5}
@attribute P21	{1,2,4,5}

TABLA VI.
ARCHIVO DETALLE PARTE I

P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
2	1	4	3	1	4	2	5	2	5	2
.
5	1	4	3	2	3	3	3	4	5	2
3	3	2	2	4	3	4	4	4	5	3

TABLA VII.
ARCHIVO DETALLE PARTE II

P12	P13	P14	P15	P16	P17	P18	P19	P20	P21
2	4	5	1	1	1	1	1	1	4
.
3	5	4	5	5	1	5	5	1	4
4	2	4	1	1	1	1	1	1	2

Paso 4. Selección de los grupos de estudiantes.

Bajo una interpretación de los archivos anteriores por medio de la plataforma Weka bajo el algoritmo J48, con una validación cruzada 90%-10%, se encuentran 185 registros clasificados correctamente y 68 registros clasificados incorrectamente. Los registros clasificados correctamente formaran el primer grupo de estudiantes. Al realizar una nueva clasificación con solo este grupo, es posible obtener un porcentaje de predicción de la variable dependiente superior al 96%. Posteriormente los 68 registros restantes que fueron clasificados incorrectamente vuelven a ser clasificados encontrándose un nuevo grupo de 52 registros clasificados correctamente mediante un proceso de validación cruzada 90% - 10%. De forma similar al realizar una nueva clasificación con solo este grupo de 52 registros, es posible lograr un porcentaje de acierto superior al 84%. Finalmente, cuando los 16 registros restantes son vueltos clasificar bajo el mismo procedimiento se obtiene un porcentaje de éxito superior al 78%. Las tablas 8 - 16 y las figuras 1, 2, 3 ilustran las respectivas matrices de

clasificación, precisión y confusión para cada uno de los grupos encontrados en esta fase.

Paso 5. Arboles de decisión y variables estadísticas.

Como resultado del paso anterior, se obtiene las siguientes tablas de clasificación y arboles de dedición para cada uno de los grupos obtenidos:

Grupo 1:

TABLA VIII.
CLASIFICACIÓN GRUPO 1

Variables	#	%
Correctly Classified Instances	177	95.6757 %
Incorrectly Classified Instances	8	4.3243 %
Kappa statistic	0.9201	
Mean absolute error	0.0181	
Root mean squared error	0.122	
Relative absolute error	6.5831 %	
Root relative squared error	32.981 %	
Total Number of Instances	185	

TABLA IX.
MATRIZ DE PRECISIÓN. GRUPO I

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.5	0.011	0.5	0.5	0.5	0.869	1
0.75	0	1	0.75	0.857	0.936	2
0.98	0.048	0.962	0.98	0.971	0.987	4
0.972	0.018	0.972	0.972	0.972	0.993	5
0.957	0.034	0.957	0.957	0.956	0.985	← Pro

TABLA X.
MATRIZ DE CONFUSIÓN. GRUPO I

a	b	c	d	classified as
2	0	1	1	a = 1
1	6	1	0	b = 2
1	0	100	1	c = 4
0	0	2	69	d = 5

Grupo 2:

TABLA XI.
CLASIFICACIÓN GRUPO 2.

Variables	#	%
Correctly Classified	44	84.6154 %
Incorrectly Classified Instances	8	15.3846 %
Kappa statistic	0.736	
Mean absolute error	0.0821	
Root mean squared error	0.2306	
Relative absolute error	26.9351 %	
Root relative Instances squared error	59.4723 %	
Total Number of Instances	52	

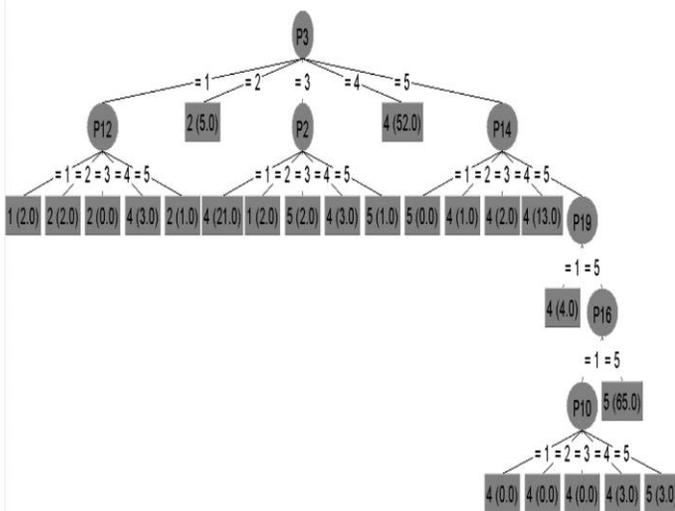


Figura 1. Árbol de decisión Grupo 1

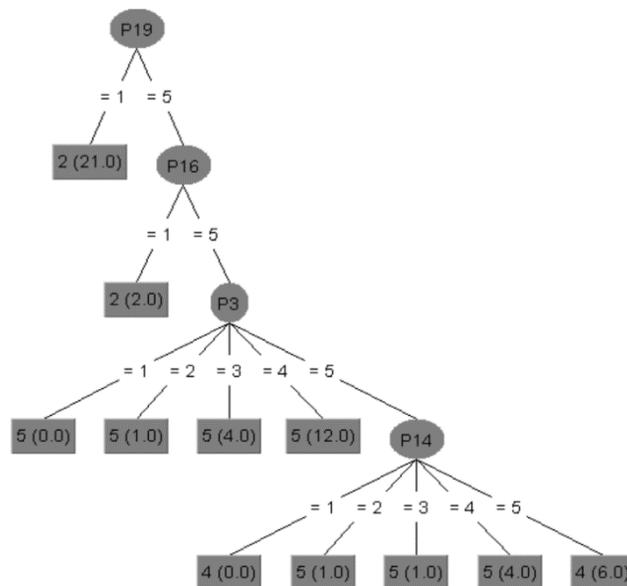


Figura 2. Árbol de decisión Grupo 2

Grupo 2:

TABLA XII.
MATRIZ DE PRECISIÓN. GRUPO 2

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0	0	0	0	0		1
0.913	0	1	0.913	0.955	0.994	2
0.333	0.043	0.5	0.333	0.4	0.813	4
0.913	0.207	0.778	0.913	0.84	0.927	5
0.846	0.097	0.844	0.846	0.84	0.943	← P

TABLA XIII.
MATRIZ DE CONFUSIÓN. GRUPO 2

a	b	c	d	classified as
0	0	0	0	a = 1
0	21	0	2	b = 2
0	0	2	4	c = 4
0	0	2	21	d = 5

Grupo 3:

TABLA XIV.
CLASIFICACIÓN GRUPO 3

Correctly Classified Instances	11	78.5714 %
Incorrectly Classified Instances	3	21.4286 %
Kappa statistic	0.6613	
Mean absolute error	0.1389	
Root mean squared error	0.3219	
Relative absolute error	40.0067 %	
Root relative squared error	76.9645 %	
Total Number of Instances	14	

Grupo 3:

TABLA XV.
MATRIZ DE PRECISIÓN. GRUPO 3

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.833	0.125	0.833	0.833	0.833	0.875	1
0	0	0	0	0	0	2
0.667	0	1	0.667	0.8	0.833	4
0.8	0.222	0.667	0.8	0.727	0.833	5
0.786	0.133	0.81	0.786	0.788	0.851	← P

TABLA XVI.
MATRIZ DE CONFUSIÓN. GRUPO 3

a	b	c	d	Clas as
5	0	0	1	a = 1
0	0	0	0	b = 2
0	0	2	1	c = 4
1	0	0	4	d = 5

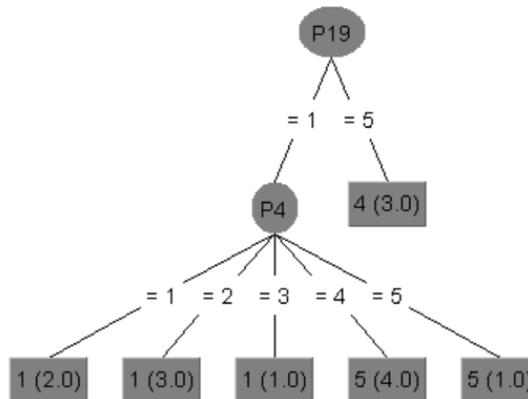


Figura 3. Árbol de decisión Grupo 3

IV. DISCUSIONES

El análisis de los resultados anteriores muestra que este trabajo investigativo no solo identifica si un docente es bueno o malo, sino también las variables independientes que producen una adecuada valoración del docente. Según la figura 1, se encuentra que el primer grupo de estudiantes califican adecuadamente un docente con base en las siguientes variables: P3, P12, P2, P14, P19, P16, P10. Con una efectividad superior al 96%, en un proceso de validación cruzada 90%-10%. De forma similar en el grupo 2, predominan las siguientes variables en el momento de valorar un docente: P19, P16, P3, P14, con una efectividad superior al 84%. Finalmente, para el tercer grupo las variables predominantes son: P19, P4, lográndose una efectividad superior al 78%.

En las diferentes revisiones literarias [1,3,9], las valoraciones docentes solo permiten predecir si un docente es excelente, bueno, regular o malo. Sin embargo, en este trabajo investigativo vas más allá y establece las diferentes variables por grupos de estudiantes, las cuales permiten determinar una valoración objetiva de un docente. Igualmente, se expresa que el hecho de saber con más precisión las variables más importantes por cada grupo de estudiantes que puedan surgir, permite definir políticas de valoración docente más centradas en las prioridades de los diferentes grupos de estudiantes.

En este sentido las políticas estudiantiles y profesoriales, partirán de las necesidades reales de los estudiantes y docentes. Sobre todo tratando de identificar que esperan los estudiantes de los docentes y viceversa. Algo que puede variar considerablemente según el grupo de estudiantes que se esté analizando. Por ende, como se expresó el diseño de estas políticas debe estar focalizado en las características que identifiquen cada grupo, reconocer estas diferencias permite que los docentes puedan obtener un mejor desempeño.

Como futuras líneas de investigación se deben tomar un mayor número de registros, con el fin de evitar sesgos en las encuestas, e incluso identificar y eliminar datos anómalos. Siendo importante incluir algunas otras variables como: socioeconómicas, ambientales del aula o condiciones psicológicas de los estudiantes. Finalmente, se resalta que la metodología propuesta produce un modelo el cual no está sobre ajustado, algo que se identifica en la simplicidad de los árboles de decisión generados, figuras 1,2 y 3.

V. CONCLUSIONES

Con una efectividad promedio superior al 96%, 84% y 78% se identifican tres grupos de estudiantes, los cuales valoran los docentes bajo tres conjuntos de variables diferentes. En el grupo uno, predominan las variables: Preparación adecuada de clases, adecuación de los métodos de enseñanza, motivo inscripción asignatura, el docente fue justo, las fortalezas del docente, se aprende con profundidad, el docente es respetuoso. Para el grupo dos predominan las siguientes variables: las fortalezas del docente, se aprende con profundidad, preparación adecuada de clases, el docente es justo. En el grupo tres

predominan las siguientes variables: fortalezas del docente y asesoría adecuada.

Finalmente, se expresa que para repetir este trabajo solo es necesario seguir cada uno de los pasos descritos en la metodología y aplicarlos sobre la base de dato deseada. Algo factible dado la plataforma libre sobre la cual se desarrolla la investigación. Si bien esto último requiere algunos conceptos técnicos, es de resaltar que también le da una gran versatilidad a la investigación y facilita su aplicabilidad en áreas como la docencia. Área en la cual la aplicación de estas técnicas de minería de datos ha sido muy escasa.

AGRADECIMIENTOS

Este trabajo se realiza como parte del año sabático autorizado mediante resolución Nro. 0032 DE 2024 del 16 de enero de la Facultad de Ingeniería y Arquitectura de la Universidad Nacional de Colombia – Sede Manizales

REFERENCIAS

- [1] B. Senden, T. Nilsen, N. Teig, "The validity of student ratings of teaching quality: Factorial structure, comparability, and the relation to achievement", *Studies in Educational Evaluation*, 78, 101274, 2023. <https://doi.org/10.1016/j.stueduc.2023.101274>
- [2] M. Valencia, J. Correa, F. Díaz, "Métodos Estadísticos Clásicos y Bayesianos para el Pronóstico de Demanda. Un Análisis Comparativo", *Revista Facultad de Ciencias Universidad Nacional de Colombia*, 4(1), 52-67 (2015). <https://doi.org/10.15446/rev.fac.cienc.v4n1.49775>
- [3] Y. Cui, "Optimizing decision trees for English Teaching Quality Evaluation (ETQE) using Artificial Bee Colony (ABC) optimization", *Heliyon* 9, e19274, 2023, <https://doi.org/10.1016/j.heliyon.2023.e19274>
- [4] B. Senden, N. Teig, T. Nilsen, "Studying the comparability of student perceptions of teaching quality across 38 countries", *International Journal of Educational Research Open* 5, 100309, 2023, <https://doi.org/10.1016/j.ijedro.2023.100309>
- [5] F. Boyle, E.J. Cook, "Developmental evaluation of teaching quality: Evidence practice", *Journal of University Teaching & Learning Practice*, Vol 20, 1, 2023, <https://ro.uow.edu.au/jutlp>
- [6] Jaekel, T. Fütterer, R. Gollner, "Teaching characteristics in distance education associations with teaching quality and students' learning experiences", *Teaching and Teacher Education*, 132, 104174, 2023, <https://doi.org/10.1016/j.tate.2023.104174>
- [7] T. Daltoè, E. Ruth-Herbein, B. Brucker, A. Jaekel, U. Trautwein, B. Fauth, P. Gerjets, R. Göllner, "Immersive insights: Unveiling the impact of 360-degree videos on preservice teachers' classroom observation experiences and teaching-quality ratings", *Computers & Education*, article in press. <https://doi.org/10.1016/j.compedu.2023.104976>
- [8] E. BJORHEIM A. PATTAMAWAN J. JAN T. KVALØY, H. BJORHEIM ABRAHAMSEN, "Utilizing Students' Feedback to Ensure Quality in Teaching and Learning: A Lesson From COVID-19", *Reflections*, 30, 1, 2023
- [9] O. Castrillón, W. Sarache, S. Ruiz-Herrera, "Predicción del Rendimiento Académico por Medio de Técnicas de Inteligencia Artificial", *Formación Universitaria*, 13(1), 93-102 (2020) <http://dx.doi.org/10.4067/S0718-50062020000100093>
- [10] V. Osorio, Impacto de las Tecnologías de la Información y las Comunicaciones (TIC) en la Competitividad de las Mi-Pymes, Universidad Nacional de Colombia, Tesis Maestría en Administración, Repositorio Universidad Nacional, (2016)
- [11] Witten, E. Frank., y otros dos autores, "Data Mining Practical Machine Learning Tools and Techniques", *Morgan and Kaufman publication* (Elsevier), ISBN-13: 978-0128042915, Cambridge, USA (2017)