

Comparative analysis using data mining in the prediction of the academic performance of teenagers based on emotions

Gleny Paola Gamarra Ramos, Estudiante Universitario¹, María Elisabeth Farfán Choquehuanca Estudiante Universitario¹

¹Universidad Nacional de San Agustín de Arequipa, Perú,
ggamarrara@unsa.edu.pe, mfarfanc@unsa.edu.pe

Abstract— All research related to performance academic is of great importance in our society due to the decadence that our educational system suffers, that is why it proposes to carry out a comparative analysis using techniques of data mining to predict the academic performance of adolescents based on emotions, for this we make use of data mining techniques such as decision trees and networks Bayesian, reaching in this way a deep analysis through those two techniques. To make this analysis, use was made of data collected selectively from an educational institution of Arequipa, obtaining as a result that the students who tend to have negative emotions during the first years of secondary school students have low academic performance compared to the rest of the students.

Keywords— Data mining, Weka, Bayesian networks, Decision trees

Digital Object Identifier (DOI):
<http://dx.doi.org/10.18687/LACCEI2019.1.1.77>
ISBN: 978-0-9993443-6-1 ISSN: 2414-6390

Análisis Comparativo usando minería de datos en la predicción del rendimiento académico de adolescentes basado en emociones

Gleny Paola Gamarra Ramos, Estudiante Universitario¹, María Elisabeth Farfán Choquehuanca Estudiante Universitario¹

¹Universidad Nacional de San Agustín de Arequipa, Perú,
ggamarrara@unsa.edu.pe, mfarfanc@unsa.edu.pe

Abstract– *Toda investigación relacionado a la educación y el rendimiento académico es de gran importancia en nuestra sociedad debido a la variabilidad que tiene nuestro sistema educativo, Las metodologías de educación que suelen ser usadas desconocen de las nuevas problemáticas que surgen y afectan a los estudiantes de diferentes niveles educativos, es por ello que se propone realizar un análisis comparativo usando técnicas de minería de datos para predecir el rendimiento académico de adolescentes basado en emociones, de esta manera se buscó identificar las deficiencias y puntos cumbre con el motivo de aportar información a nuevas metodologías, para ello hacemos uso de técnicas de minería de datos como árboles de decisión y redes bayesianas, alcanzando de esta forma un análisis profundo mediante esas dos técnicas. Este análisis requirió del uso de datos recolectados selectivamente de una institución educativa de Arequipa, obteniendo como resultado que los alumnos que tienden a tener emociones negativas durante los tres primeros años de secundaria y mostrando declive en sus notas en comparación a los últimos dos años.*

Palabras clave- Minería de datos, WEKA, redes bayesianas, Árboles de decisión, Rendimiento Académico, Educación.

I. INTRODUCCIÓN

La educación es uno de los temas de mayor importancia a nivel mundial, existen diversos estudios que buscan ayudar en el análisis de aspectos que intervienen en el rendimiento académico, como las relaciones familiares, amorosas, económicas, entre otras. Todos estos aspectos intervienen directa e indirectamente durante el desarrollo de aprendizaje de los adolescentes, siendo esta una etapa vulnerable, llena de emociones.

Se considera que el rendimiento académico viene siendo estudiado y analizado desde tiempos memorables [1], y por diversas áreas. Es así que, aprovechando el avance tecnológico, se presenta la realización de predicciones que permitirán anteponerse a la problemática a la cual se presenta en la actualidad. Como se conoce bien la adolescencia es una etapa de cambios en los jóvenes, las emociones son diversas de acuerdo a las circunstancias en las que viven. ¿Estas emociones influyen positiva o negativamente en el rendimiento académico

de los jóvenes? es una pregunta que se presentan de forma común ahora.

Es por ello que nuestro trabajo busca analizar el rendimiento académico de los adolescentes basado en sus emociones. Nuestro objetivo es realizar un análisis comparativo de minería de datos para la predicción del rendimiento académico en adolescentes basado en emociones, mostrando la importancia de las emociones en el rendimiento académico.

El artículo se encuentra organizado de la siguiente manera: En la sección II se hace referencia a los diferentes trabajos que presentan investigación con técnicas de minería de datos y el rendimiento académico. En la sección III realizamos una definición de los diferentes conceptos de relevancia para el análisis. En la sección IV definimos la propuesta de solución, requerimientos encontrados y el modelado en la herramienta utilizada.

Finalmente presentamos nuestras conclusiones referentes al análisis comparativo con técnicas de minería de datos para la predicción del rendimiento académico de adolescentes basado en emociones, mostrando los datos inferidos de ambas técnicas usadas.

II. TRABAJOS RELACIONADOS

La minería de datos es un campo que viene ganando mayor relevancia con el pasar de los años, a continuación, revisaremos diversas investigaciones referentes al rendimiento académico de estudiantes, con diversas técnicas.

Amaya en [5] realiza un Modelo predictivo de deserción estudiantil usando técnicas de minería de datos, en dicho estudio realiza el análisis comparativo en algoritmos árboles de decisión como: El algoritmo ID3 y el algoritmo C4.5. Llegando a la conclusión de que más de la mitad de alumnos de dicha institución podrían desertar de la carrera.

En [6] desarrollan una metodología de estudio del Rendimiento académico Mediante la Minería de Datos donde

Digital Object Identifier (DOI):
<http://dx.doi.org/10.18687/LACCEI2019.1.1.77>
ISBN: 978-0-9993443-6-1 ISSN: 2414-6390

hacen uso de los siguientes modelos de agrupamiento, clasificación y asociación. En este estudio buscan encontrar comportamientos que lleven a los académicos a obtener éxito o fracaso en el rendimiento académico, se busca obtener información para revertir casos en desventaja que puedan tener los estudiantes.

La Red en [7] analiza los Perfiles de Rendimiento Académico: Un Modelo basado en Minería de dato. Donde analiza las situaciones socio-económicas de los estudiantes del primer semestre de la escuela de ingeniería de sistemas de información. Analizando el desempeño en las materias, así como la procedencia de instituciones educativas anteriores a la universidad tecnológica.

Rico en [8] analiza las calificaciones parciales y finales de estudiantes que pertenecen al área de ingeniería del instituto politécnico nacional, haciendo uso de técnicas de minería de datos. Para este análisis además de las calificaciones, consideraron las actividades académicas.

Al obtener los resultados fueron corroborados con la información que ya se tenía, alcanzando una exactitud de 73% de aprobación. Este trabajo le permitió al instituto predecir el rendimiento académico de los estudiantes y conocer cuáles son los factores que influyen en ello.

En [9] Menacho realiza una predicción del rendimiento académico de alumnos matriculados en el curso de Estadística General de la UNALM en los semestres 2013 II y 2014 I, para lo cual hace uso de las siguientes técnicas de minería de datos:

Regresión logística, árboles de decisión, redes bayesianas y redes neuronales, lo que buscan es predecir la cantidad de alumnos aprobados y desaprobados que se van a matricular en dicho curso. Los resultados que obtuvieron indican que la red Naive de Bayes obtuvo la mayor tasa de buena clasificación alcanzando un 71,0% de una correcta clasificación.

Si bien las diversas investigaciones que se han realizado hasta el momento está basado en calificaciones, procedencia de centros educativos. En [10] realizan una investigación específica de la relación que existe entre las relaciones familiares, el auto concepto y el rendimiento académico que tiene el estudiante, tanto la psicología como la educación se reunieron para investigar, hasta dónde influyen estos aspectos en el rendimiento académico.

Es así como hemos visto diversas investigaciones no solo científicas, sino también psicológicas que se encargan de estudiar este tema tan sensible del rendimiento académico de los adolescentes. Nuestro trabajo es diferente porque busca

analizar el rendimiento académico en función a las emociones que puedan presentar los adolescentes durante esta etapa.

III. METODOLOGÍA

Para el desarrollo del análisis, es necesario considerar ciertos conceptos que nos permitirán tener mayor conocimiento acerca del tema y de la importancia del uso de las herramientas usadas.

A. Emociones

Se define que un individuo tiene reacciones a diferentes estímulos del medio ambiente. Estas reacciones son conocidas como emociones. La calidad emocional de un ser vivo es diferente, los seres Humanos son calificados como individuos que pueden controlar y manejar emociones, dependiendo de su control y la edad en la que se encuentren [11].

No siendo el mismo caso de los animales, individuos que poseen emociones sin control, donde predomina el instinto salvaje. Cada una de estas emociones corresponden a estímulos que pueden dar información sobre la situación del individuo, tales como bienestar emocional, influencia en la educación, desarrollo personal, entre otros [11].

En el presente trabajo buscamos que mediante los datos obtenidos de las emociones de los estudiantes podamos identificar la influencia principal en su rendimiento académico, los datos son referidos de una encuesta previa realizada.

B. Rendimiento Académico

Se puede entender por rendimiento académico como el resultado de la capacidad que un alumno posee para realizar las actividades en una etapa académica. Dicha capacidad es variable debido a las diferentes exigencias a las que sea sometido el alumno en la institución donde realice su desenvolvimiento de estudio [11, 13].

El rendimiento académico viene acompañado de técnicas de estudio que adoptan los estudiantes, donde se considera las influencias primarias, desde las estrategias motivacionales donde las emociones tienen mayor control y no afectan el rendimiento, hasta las tendencias específicas de estudio por compromiso [14].

Mediante los trabajos revisados se puede definir que la forma de evaluar el rendimiento académico de alumno es mediante las notas obtenidas, dado que ahí se refleja el resultado de los esfuerzos aplicados. Un promedio de notas elevado daría a significar que el alumno da un buen aprovechamiento de su educación, por el contrario, un promedio de notas bajo sería un mal aprovechamiento [6].

La influencia del rendimiento académico se considera mediante emociones que pueden ser favorecedoras y no serlo, consideramos una encuesta basada en emociones personales y emociones de entorno, como son la situación económica y emocional que pueda percibir el estudiante en su entorno familiar.

C. Minería de Datos

La minería de datos se puede clasificar en dos categorías: minería de datos descriptiva y minería de datos predictiva. En el siguiente trabajo buscamos trabajar con dos técnicas que realizan predicción, es decir se hace uso de la minería de datos descriptiva [1]. Cada una se desarrolla de forma independiente con uno o más atributos y son usados para dos tipos de análisis en torno a una variable diferente o más variables].

i. Redes Bayesianas

Las redes bayesianas son herramientas que son usadas para realizar modelado estadístico, debido a que representan un conjunto de incertidumbres relacionadas. Se construyen mediante su fundamento probabilístico, esta estructura las hace apropiadas para diseñar sistemas de variación orientados a la clasificación, el diagnóstico y la toma de decisiones [15].

Una red bayesiana está compuesta por una estructura de cuatro niveles. En el nivel superior, se presenta un conjunto de variables representadas por nodos y un conjunto de flechas que relacionan estas variables en términos de influencia. Continuando con un nivel inferior estarían los niveles o estados, también conocidos como espacio de estados [4].

El uso de variables conocidas permite a las redes bayesianas hallar relaciones de dependencia e independencia. En base a ello, se hace uso de métodos de razonamiento probabilístico que permiten realizar predicciones determinando el valor de cualquier variable desconocida basados en los valores de las variables conocidas.

ii. Árbol de Decisión

Es una de las técnicas más conocidas en predicción debido a su facilidad de interpretación dada su estructura y que puede ser convertida directamente en un conjunto de reglas, las cuales te permiten definir conceptos mediante un SI ENTONCES [1].

Esta técnica de aprendizaje inductivo supervisado no paramétrico, te permite trabajar una base de datos mediante diagramas de construcción lógica, los mismos que son basados en reglas que sirven para identificar y clasificar condiciones repetitivas que se den para la solución de un problema [2].

La modalidad de trabajo de los árboles de decisión inicia con un nodo raíz que inicia con una evaluación y en base a ella se avanzan a las ramas u hojas, donde las hojas representan

clases y el nodo raíz representa todos los patrones de entrenamiento los cuales se han de dividir en clases [2].

En este trabajo mediante el uso esta técnica se busca diferenciar las emociones que influyen de forma negativa y positivamente, los alumnos son sometidos a encuestas sobre sus emociones y son comparadas con la información de su promedio académico que se realizan en base a condiciones que son perfectas según el desarrollo de los árboles de decisión, cada condición está dada por la pregunta previa realizada a los alumnos en encuestas.

D. Herramienta WEKA

WEKA, acrónimo de Wakaito Environment for Knowledge Analysis, es un software de código abierto bajo términos de GNU GPL y posee muchas herramientas para minería de datos y aprendizaje automático [12].

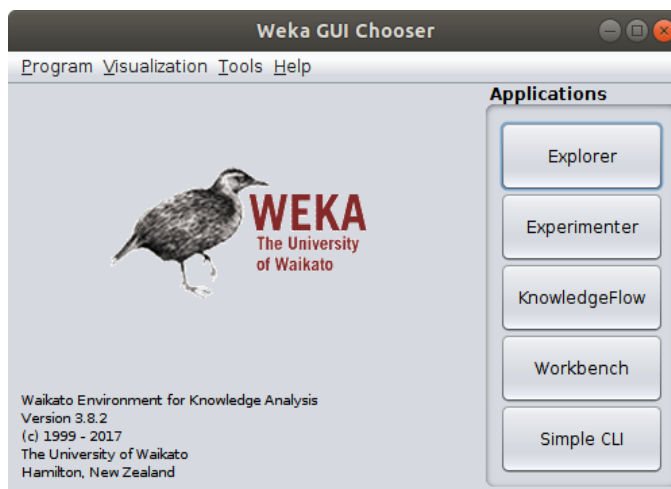


Figura 1. Plataforma virtual WEKA. Fuente Propia

V. CASO DE APLICACIÓN: COLEGIO DE AREQUIPA

A. Descripción del problema

En el departamento de Arequipa se conoce que muchos estudiantes se han visto que el Rendimiento Académico de los mismos son influenciados por diversos problemas emocionales que viven día a día, todo esto se ve reflejado en el promedio de notas de cada uno de ellos.

La educación en la actualidad es considerada de acuerdo a estudios [10] muy influenciada por la estabilidad emocional de la familia, condición económica y muchos otros factores que se ve reflejados en los distintos comportamientos de los estudiantes, consideramos diferentes emociones que pueden expresar y darnos un significado, esta información es procesada mediante encuestas.

B. Propuesta

Nuestra propuesta tiene por objetivo realizar un análisis comparativo de las diversas técnicas de minería de datos para poder determinar el rendimiento académico de los adolescentes basado en emociones, de esta manera poder anteponerse a la problemática y tomar medidas para poder contrarrestarlas aportando información nueva para las metodologías de educación.

De esta manera poder manejar la problemática de alumnos que sufren diversos alumnos, y obtener un mejor rendimiento académico, enfocarnos no solo en cuáles son las emociones más identificadas, también en los años de estudio donde se puede observar con mayor frecuencia este cúmulo de emociones que influyen con mayor intensidad.

C. Requerimientos

Los requerimientos para realizar este trabajo son:

- a. **Datos:** Los datos con los que vamos a trabajar durante esta investigación, fueron recopilados de una institución de la ciudad de Arequipa. Estos datos fueron recopilados brevemente por encuestas que se realizaron, donde se fueron considerados los siguientes aspectos: Edad, sexo, depresión, estado anímico (feliz, triste), grado que cursan, grado académico de los padres.

Se consideró estudiantes de diferentes grados de estudio que cursen el nivel secundario, debido a que los de nivel primario son más volubles y no es fácil controlar o definir las emociones que pueden expresar.

Toda esta información fue cuidadosamente seleccionada y colocada en un archivo .csv con el cual se va a trabajar en la herramienta WEKA.

- b. **Instrumentos:** Es importante identificar de qué técnicas se hace uso de acuerdo a los datos recopilados. Es por ello que se decidió trabajar con WEKA, debido a su contenido que permite trabajar con una gran mayoría de técnicas de minería de datos, las cuales están disponibles y son de fácil uso, de acuerdo a un análisis previo se determinó que los árboles de decisión y las redes bayesianas son las técnicas más óptimas para la predicción.

- c. **Pre procesamiento de los datos:** Weka nos permite realizar filtros para el pre procesamiento de los datos ingresados, de esta forma se encuentran listos para la clasificación. Estos datos son verificados de forma que no se encuentren datos erróneos o datos no encontrados, Weka también nos permite poder completar datos que no se encuentren con valores promedio.

- d. Construcción y prueba del modelo del rendimiento académico de adolescentes basado en emociones.
- e. Validación de las técnicas usadas en el modelo de rendimiento académico y análisis de los resultados.

D. Construcción y Pruebas del Modelo de Rendimiento Académico

Inicialmente cargamos los datos con los que han sido seleccionados de las encuestas realizadas a los alumnos, la siguiente imagen muestra la carga de datos en WEKA.

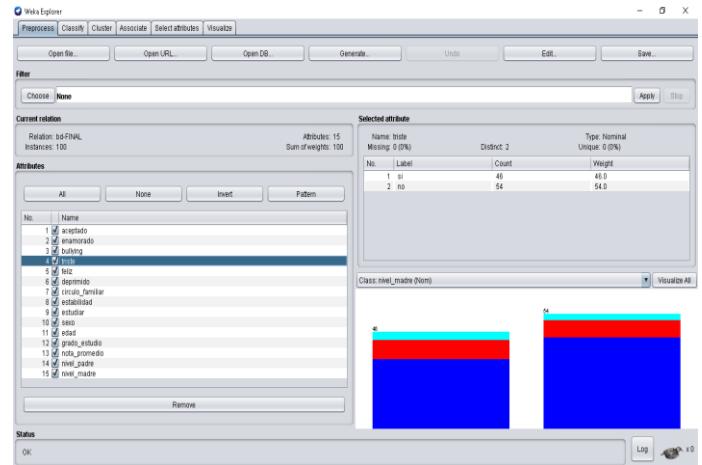


Figura 2. Datos cargados en WEKA

Una vez cargados los datos, realizamos el pre-procesamiento realizando los filtros, posteriormente se trabajó con el Árbol de decisión J4.8 obteniendo el siguiente diagrama.

El modelo se construyó a partir de 100 instancias, las que utilizamos para poder entrenar y probar correctamente el modelo a obtener. Una vez que cargamos los datos y se realizó la clasificación con el algoritmo J4.8 de Weka, obtuvimos el siguiente árbol de decisión.

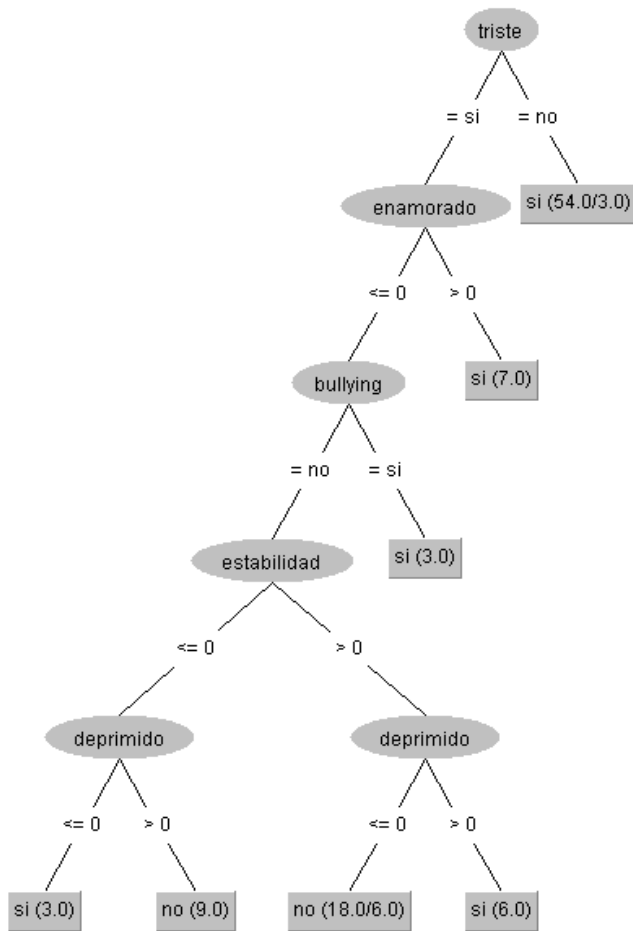


Figura 3. Árbol de decisión del modelo obtenido (Fuente: Propia).

A continuación, observamos los resultados adquiridos del algoritmo J48.

```

==== Run information ====

Scheme:   weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: bd-FINAL
Instances: 100
Attributes: 15
  aceptado, enamorado, bullying, triste, feliz, deprimido,
  circulo_familiar, estabilidad, estudiar, sexo, edad,
  grado_estudio, nota_promedio, nivel_padre, nivel_madre

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      85      85 %
Incorrectly Classified Instances    15      15 %
Kappa statistic                    0.6057
  
```

```

Mean absolute error      0.1911
Root mean squared error  0.3555
Relative absolute error  51.8969 %
Root relative squared error 83.101 %
Total Number of Instances 100

==== Detailed Accuracy By Class ====
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC
Area Class
0,882 0,250 0,918 0,882 0,899 0,608 0,825
0,911 si
0,750 0,118 0,667 0,750 0,706 0,608 0,825
0,640 no
Weighted Avg. 0,850 0,218 0,858 0,850 0,853
0,608 0,825 0,846

==== Confusion Matrix ====

a b <-- classified as
67 9 | a = si
6 18 | b = no
  
```

Figura 4. Datos resultantes de la evaluación con el árbol de decisión J48 en función al estado de ánimo.

Este árbol de decisión nos da probabilidad en exactitud de un 85%. Como se observa en los siguientes resultados, la mayoría de alumnos que sufren de emociones negativas tiene esta relacionado con que: Sufren de bullying, tienen enamorado, la estabilidad económica de su familia existe, generando en ellos una depresión.

Comparando con la técnica de redes bayesianas, como se puede apreciar en la imagen, los resultados son en base a los datos y en función principal a su grado de estudios.

```

Naive Bayes Classifier

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      98      98 %
Incorrectly Classified Instances     2       2 %
Kappa statistic                    0.9738
Mean absolute error                 0.0121
Root mean squared error              0.0696
Relative absolute error              3.9506 %
Root relative squared error          17.7663 %
Total Number of Instances           100

==== Detailed Accuracy By Class ====

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC
  
```

```

Area Class
1,000 0,000 1,000 1,000 1,000 1,000 1,000
1,000 qui
0,909 0,000 1,000 0,909 0,952 0,941 1,000
1,000 seg
1,000 0,000 1,000 1,000 1,000 1,000 1,000
1,000 terce
1,000 0,000 1,000 1,000 1,000 1,000 1,000
1,000 cuart 1,000 0,031 0,946 1,000
0,972 0,958 1,000 1,000 prime
Weighted Avg. 0,980 0,011 0,981 0,980 0,980
0,972 1,000 1,000

==== Confusion Matrix ====

a b c d e <-- classified as
11 0 0 0 0 | a = quinto
0 20 0 0 2 | b = segundo
0 0 17 0 0 | c = tercero
0 0 0 15 0 | d = cuarto
0 0 0 0 35 | e = primero

```

Figura 5. Datos resultantes de la evaluación con el algoritmo Redes Bayesianas en weka en función al grado de estudio

La confiabilidad de las redes bayesianas alcanza un 98%, siendo altamente confiable. Como se puede observar dado los primeros años de estudios los adolescentes presentan más insatisfacción en sus emociones lo que provoca un promedio limitado. Mientras transcurren los años el estudiante disminuye la sensación de presión y el ataque centralizado en para sí mismo, no cuenta tanto su depresión mejorando su calidad y estabilidad emocional por lo tanto también su promedio presenta un incremento.

Así mismo se evaluaron los datos de acuerdo al estado anímico: Feliz o triste como se analizó anteriormente con el árbol de decisión. Se obtuvieron los siguientes resultados, como se muestra en la imagen.

```

==== Stratified cross-validation ====
==== Summary ====

Correctly Classified Instances      78      78  %
Incorrectly Classified Instances    22      22  %
Kappa statistic                    0.3969
Mean absolute error                 0.284
Root mean squared error             0.4249
Relative absolute error             77.1147 %
Root relative squared error         99.3341 %
Total Number of Instances          100

==== Detailed Accuracy By Class ====

TP Rate FP Rate Precision Recall F-Measure MCC   ROC Area PRC Area
Class

```

```

0,855 0,458 0,855 0,855 0,855 0,397 0,743
0,873 si 0,542 0,145 0,542 0,542 0,542
0,397 0,743 0,426 no
Weighted Avg. 0,780 0,383 0,780 0,780 0,780
0,397 0,743 0,766

==== Confusion Matrix ====

a b <-- classified as
65 11 | a = si
11 13 | b = no

```

Figura 6. Datos resultantes de la evaluación con el algoritmo Redes Bayesianas en weka en función al estado anímico.

El análisis de acuerdo al estado emocional en función al sexo de los estudiantes, se observa que su tendencia se inclina a tener más sentimientos tristes que las mujeres, eso se representa durante los tres primeros años de secundaria, los resultados son reflejados en sus notas como se demostró en el análisis.

triste	masculino	femenino
si	0,606	0,394
no	0,427	0,573

Figura 7. Datos resultantes de la evaluación con el algoritmo Redes Bayesianas en weka en función al estado anímico vs sexo.

E. Análisis de Resultados

En el primer análisis por la técnica de Árboles de decisión nos centramos en las emociones principales de cada estudiante evaluado y el porcentaje en el que se incrementan, los estudiantes consideran en estado de ánimo feliz como momentáneo, es decir no es un estado emocional estable.

La influencia de tener un enamorado o no se incluye dado que en la adolescencia se suele promover el compartir experiencias, lo cual también genera distracciones que podrían influir en el desempeño académico. Analizando los resultados obtenidos en el árbol de decisión se observa que los alumnos que sufren emociones negativas, asociadas con tener enamorado y sufrir bullying tienden a tener un rendimiento académico menor al resto de alumnos.

Por la técnica de Redes Bayesianas nos centramos en todas las emociones incluidas, pero en qué años se considera más la influencia total, considerando el año académico en donde se muestra mayor intensidad de influencias. Según el análisis los primeros años de estudios los adolescentes presentan mayor

interacción en sus emociones lo que provoca un promedio condicionado.

Con el pasar de grados de estudio los estudiantes no presentan la misma intensidad de los sentimientos negativos como depresión, alteración, etc. Se observa una mejora de su calidad de estabilidad emocional y por lo tanto también su promedio presenta un incremento y es mejor en comparación al principio, se demostró que los estudiantes varones durante estas tres primeras etapas, tienden a sentirse más tristes que las estudiantes mujeres.

VI. CONCLUSIONES

DESPUÉS de analizar los dos métodos en minería de datos, se llega a las siguientes conclusiones:

- ❑ El árbol de decisión nos ayudó a predecir que los alumnos asociados a emociones negativas tienden a tener un rendimiento académico menor al resto de alumnos.
- ❑ Se demostró que los alumnos de los tres primeros años, tienden a tener un rendimiento académico más bajo, así como una inestabilidad emocional constante.
- ❑ Así mismo se presentó muestras de que los varones durante estos tres primeros años son los que tienen sentimientos más tristes que las mujeres.
- ❑ El tener conocimiento de estos resultados permite a las autoridades de los centros educativos, tomar medidas para poder contrarrestarlas y/o anteponerse a los problemas académicos y sociales.
- ❑ Las metodologías que se realicen en un futuro pueden ser realizadas basadas en esta investigación dando prioridad a los tres primeros años de educación secundaria, considerando una tutoría en los temas donde los alumnos tienden a presentar más inestabilidad emocional.
- ❑ Este análisis sirve como colaboración para futuras investigaciones, relacionadas al tema de rendimientos académico y salud emocional.

REFERENCIAS

[1] Eiriku Yamao, "Predicción Del Rendimiento Académico Mediante Minería De Datos En Estudiantes Del Primer Ciclo De La Escuela Profesional De Ingeniería De Computación Y Sistemas, Universidad De San Martín De Porres, Lima-Perú", Perú, 19-30.

[2] Guillermo Roberto Solarte Martínez, José A. Soto Mejía, "Árboles de decisiones en el diagnóstico de enfermedades cardiovasculares", Universidad Tecnológica de Pereira, Pereira, Colombia

[3] Jorge López Puga, Juan García García, Leticia de la Fuente Sánchez y Emilia Inmaculada de la Fuente Solana. "Las redes bayesianas como herramientas de modelado en psicología", España, 2007.

[4] Pablo Felgaer, "Optimización de Redes Bayesianas basado en Técnicas de Aprendizaje por Inducción", Argentina, 2005.

[5] Y. Amaya, E. Barrientos and D. Heredia, "Student Dropout Predictive Model Using Data Mining Techniques", IEEE Latin America Transactions, Volumen: 13, Número: 9, septiembre de 2015.

[6] David Luis La Red, Corrientes, Carlos Enrique Podestá, Curuzú Cuatía, "Methodology Study of Academic Performance Using Data Mining", Campus Virtuales, n° 01, v. III, 2014, Revista Científica de Tecnología Educativa; ISSN: 2255-1514 www.revistacampusvirtuales.es

[7] La Red Martínez, d. L., Karanik, M., giovannini, M., y Pinto, n. (2015). Perfiles de Rendimiento Académico: Un Modelo basado en Minería de datos. Campus Virtuales, Vol. IV, núm. 1, pp. 12-30. Consultado el [15/01/2019] en www.revistacampusvirtuales.es

[8] Andrés Rico Páez, Daniel Sánchez Guzmán, "Proyecto de modelo para automatizar a previsão do desempenho acadêmico em estudantes do IPN", Revista iberoamericana para la investigación y el desarrollo educativo, Vol. 8, Núm. 16, Enero – junio 2018, DOI: 10.23913/ride.v8i16.340

[9] Cesar Higinio Menacho Chiok, "Prediction of academic performance applying data mining techniques", Anales Científicos, 78 (1): 26-33 (2017), ISSN 2519-7398, DOI: <http://dx.doi.org/10.21704/ac.v78i1.811> © Universidad Nacional Agraria La Molina, Lima - Perú

[10] Ana Álvarez, Natalia Suárez, Elián Tuerro, José C. Núñez, Antonio Valle y Bibiana Regueiro, "Implicación familiar, auto concepto del adolescente y rendimiento académico", European Journal of Investigation in Health, Psychology and Education 2015, Vol. 5, N° 3 (Págs. 293-311) e-ISSN 2254-9625 www.ejihpe.es doi: 10.1989/ejihpe.v5i3.133

[11] Rafael Bisquera Alzina, "Educación emocional y bienestar", Editorial Wolters Kluwer, 2009

[12] Diego García Morate, "MANUAL DE WEKA", diego.garcia.morate@gmail.com

[13] José C. Núñez Pérez, Julio A. González-Piend, Marta García Rodríguez, Soledad González-Pumariega, Cristina Rocas Montero, Luis Álvarez Pérez y M^a del Carmen González Torres, "Estrategias De Aprendizaje, Autoconcepto Y Rendimiento Académico", 1996

[14] Evelyn Ezpinoza, "Impacto del maltrato en el rendimiento académico", 2006

[15] Jorge López Puga, "Cómo Construir y Validar Redes Bayesianas con Netica", 2012