

Aplicación del Algoritmo K means para la Tipificación de Trastornos Psicológicos mediante la escala DASS -21 en Alumnos, durante el aislamiento en tiempos de Covid19

Application of the K means Algorithm for the Typing of Psychological Disorders using the DASS -21 scale in Students, during isolation in the time of Covid19

Felipe J Núñez-Cárdenas ^a, Gustavo A Cruz-Rosas ^b, Heidi Tolentino-Bautista ^c, Victor T Tomas-Mariano ^d, Ana M Felipe-Redondo ^e

Abstract:

This project focuses on the description of psychological problems caused by the isolation of the students of the Higher School Huejutla in their homes using the DASS-21 Scale tool, for which data mining techniques and the K-means algorithm are used to the grouping and typification of the levels of depression, anxiety and stress caused by the isolation and confinement in the homes of the students, derived from the COVID19 pandemic

Keywords:

Data Mining, Clustering, K-Means, DASS-21 Scale

Resumen:

El presente proyecto se centra en la descripción de problemas psicológicos causados por el aislamiento en sus hogares de los alumnos de la Escuela Superior Huejutla utilizando la herramienta Escala DASS-21, para ello se usan técnicas de minería de datos y el algoritmo K-means para la agrupación y tipificación de los niveles de depresión, ansiedad y estrés provocados por el aislamiento y confinamiento en los hogares de los alumnos, derivados de la pandemia de COVID 19.

Palabras Clave:

Minería de Datos, Clustering, K means, Escala DASS-21

1. Introducción

Actualmente a causa de la pandemia por COVID-19 que aqueja al mundo entero se han tenido que cerrar las escuelas como medida precautoria para la propagación del virus, por este motivo los alumnos de la Escuela Superior Huejutla han tenido que tomar sus clases de

manera virtual desde sus hogares, el hecho de adaptarse a esta nueva manera de aprendizaje aunado a estar en cuarentena (encerrados) en sus hogares ha creado en la gran mayoría un alto índice de tendencia a padecer algún trastorno psicológico, que van desde el estrés, la ansiedad o la depresión, es por este motivo que es de vital importancia realizar un test el cual nos permita determinar la presencia de algún trastorno, y desde esta

^a Autor de Correspondencia, Profesor de Tiempo Completo, Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Huejutla, ORCID: 0000-0002-2462-3654, email: felipe_nunez@uaeh.edu.mx

^b Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Huejutla, email: cr352305@uaeh.edu.mx

^c Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Huejutla, email: to352874@uaeh.edu.mx

^d Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Huejutla, ORCID: 0000-0001-6623-860X, email: victor_tomas@uaeh.edu.mx

^e Universidad Tecnológica de la Huasteca Hidalguense. Email: ana.felipe@uthh.edu.mx

perspectiva tomar acciones para minimizar o en su efecto erradicar este problema.

El siguiente proyecto busca describir los trastornos psicológicos en los alumnos de la Escuela Superior de Huejutla perteneciente a la Universidad Autónoma del Estado de Hidalgo provocados por la cuarentena a causa del Covid-19, este análisis se desarrolla de manera general (todos los alumnos), nivel de estudio (licenciatura y bachillerato) y sexo (hombres y mujeres). Para este proyecto se usó la Escala DASS-21 (recolección de datos) para determinar los trastornos de estrés, ansiedad y depresión, se complementó con el uso de la herramienta WEKA y específicamente el algoritmo K-means para la determinación de los clusters o agrupamientos.

Mediante la minería de datos se buscó describir patrones a partir de la encuesta de la escala DASS-21. La Escala DASS-21 se encuentra compuesta por 21 ítems estos a su vez divididos en tres paquetes con 7 ítems cada uno de tipo Likert es decir que permite medir actitudes y conocer el grado de conformidad del encuestado, cada grupo con cuatro puntos de severidad que permite medir el grado de los estados, que ha experimentado el alumno [1].

1.1 Estado del Arte

En el año 2012 Zayra Antúnez en Chile en su artículo titulado Escalas de Depresión, Ansiedad y Estrés (DASS - 21): Validación de la Versión abreviada en Estudiantes Universitarios Chilenos en el cual describe la aplicación de una muestra que estuvo constituida por 484 estudiantes de 1ro a 4to año de distintas carreras de la Universidad Austral de Chile, seleccionados mediante un muestreo probabilístico por conglomerados, utilizando un diseño correlacional transversal [2].

Cristina Hernández en abril de 2020, en su trabajo el Impacto Psicológico de la Pandemia COVID 19 en los Profesionales Sanitarios de Coruña, uso el instrumento DASS 21 para diagnosticar los niveles de estrés, ansiedad y depresión en los trabajadores de la salud como consecuencia de la atención a pacientes de la pandemia [3].

Francisca Román en su artículo Uso de las Escalas de Depresión Ansiedad Estrés (DASS-21) como Instrumento de Tamizaje en Jóvenes con Problemas Clínicos describe la aplicación de esta Escala para detectar jóvenes en riesgo de problemas de salud mental en una muestra de 393 jóvenes no consultantes y una muestra clínica de 77 jóvenes consultantes en fase inicial de tratamiento, obteniendo como muestra de población general en Depresión un rango 0 -16 con una media de 2,95 y una desviación típica de 2,513, en ansiedad un rango de 0 -13 una media de 2,52 y una desviación típica de 2,167 y en

estrés un rango de 0 -17 una media de 4,21 y una desviación típica de 2,937 [4].

En el año 2015, Angélica Jaramillo, Henry Patricio y Paz Arias en su artículo titulado Aplicación de Técnicas de Minería de Datos para Determinar las Interacciones de los Estudiantes en un Entorno Virtual de Aprendizaje, en el cual se realizó un análisis a la base de para seleccionar los atributos necesarios que permitieron generar un modelo de minería de datos. Se desarrolló un análisis comparativo tomando en cuenta características de las herramientas de minería de datos donde se seleccionó RapidMiner para realizar los procesos mediante algunos algoritmos conjuntamente con los datos de los estudiantes, los mismos que se dividieron en dos conjuntos, para entrenamiento y validación, obteniendo como resultado que el mejor algoritmo fue el decision tree, ya que clasificó las instancias correctamente con un margen de error mínimo, así mismo presentó un árbol con los diferentes atributos dando las mejores reglas de las interacciones de los estudiantes, de tal forma se pudo generar el modelo mediante el cual se determinó que en gran mayoría los estudiantes tienen un nivel de interacción medio en el curso virtual de inglés, donde los factores que más influyen son las interacciones en las tareas, exámenes, recursos, situación laboral y estado civil del estudiante [5].

En ese mismo año Javier Hernández Cáceres de Colombia en su artículo Clustering basado en el algoritmo K-means para la Identificación de grupos de Pacientes Quirúrgicos en cual busca apoyar la Gestión Estratégica del Proceso de Cirugía en un hospital público de alta complejidad, a partir del agrupamiento de pacientes quirúrgicos cuyos diagnósticos estuvieron asociados a tumores malignos los Resultados obtenidos de la descripción exploratoria de atributos y clústeres que identifican grupos de pacientes que fueron sometidos a intervenciones quirúrgicas durante los períodos 2007 al 2015. Los resultados obtenidos permitieron evidenciar la presencia de diagnósticos asociados a cáncer, agrupando la población por: edad, grupo etario, régimen de salud, género, zona de residencia, estrato, etnia, estado civil, grupo CIE-10, grupo quirúrgico y tipo de atención [6].

1.2 Minería de Datos

Las técnicas de minería de datos han surgido a partir de sistemas de aprendizaje inductivo en computadoras, siendo la principal diferencia entre ellos, los datos sobre los que se realiza la búsqueda de nuevo conocimiento.

La Minería de Datos es un proceso que, a través del descubrimiento y cuantificación de relaciones predictivas en los datos, permite transformar la información disponible en conocimiento útil, surgió como una integración de múltiples tecnologías tales como la

estadística, el soporte a la toma de decisiones, el aprendizaje automático, la gestión y almacenamiento de bases de datos y procesamiento en paralelo. Para la realización de estos procesos se aplican técnicas procedentes de muy diversas áreas, como pueden ser los algoritmos genéticos, las redes neuronales, los árboles de decisión, etc.

La minería de datos descubre relaciones, tendencias, desviaciones, comportamientos atípicos, patrones y trayectorias ocultas, con el propósito de soportar los procesos de toma de decisiones con mayor conocimiento. La minería de datos se puede ubicar en el nivel más alto de la evolución de los procesos tecnológicos de análisis de datos [7].

1.3 Algoritmo K-Means

K-means es un algoritmo de aprendizaje simple no supervisado que utiliza para resolver problemas de agrupamiento. Sigue un procedimiento simple de clasificación de un conjunto de datos dado en una serie de grupos, definidos por la letra "k", que se fija antes de proceder. Los clusters se posicionan como puntos y todas las observaciones o puntos de datos se asocian con el cluster más cercano, se calculan. Se ajustan y luego el proceso comienza de nuevo utilizando los nuevos ajustes hasta que se alcanza un resultado deseado.

El algoritmo K-Means es un método utilizado para el análisis de clusters, especialmente en minería de datos y estadísticas. Su objetivo es dividir un conjunto de observaciones en una serie de grupo (k). Se puede considerar un método para descubrir a que grupo pertenece realmente un determinado objeto [8].

El algoritmo:

- A. K puntos se colocan en el espacio de datos del objeto que representa el grupo inicial de los centroides.
- B. Cada objeto o punto de datos se asigna a la k más cercana.
- C. Después de asignar todos los objetos, se recalculan las posiciones de los k centroides.
- D. Los pasos B y C se repiten hasta que las posiciones de los centroides ya no se muevan.

1.4 Escala DASS-21

Las Escalas de Depresión, Ansiedad y Estrés (DASS) fueron creadas con el objetivo de evaluar la presencia de afectos negativos de depresión y ansiedad, así como lograr una discriminación máxima entre estas condiciones, cuya superposición clínica ha sido reportada por clínicos y por investigadores P. F. Lovibond and S. H. Lovibond. Para esto se incluyeron síntomas esenciales de cada condición y se excluyeron los que pueden estar presentes en ambos trastornos, por ejemplo, cambios en el apetito. Durante los sucesivos análisis factoriales

surgió el tercer factor de estrés, que reúne síntomas referidos a la dificultad para relajarse, tensión nerviosa, irritabilidad y agitación [9].

El DASS fue desarrollado para cumplir las exigencias psicométricas de una escala de autoreporte desde una concepción dimensional más que categórica de los trastornos psicológicos. Esto implica que las diferencias entre sujetos normales y con alteraciones clínicas radican en la severidad en que experimentan los estados afectivos de depresión, ansiedad y estrés.

Su versión abreviada, denominada Escalas de Depresión Ansiedad y Estrés - 21 (DASS-21) tiene como ventajas el ser un instrumento de autor reporte, breve, fácil de responder y que ha mostrado adecuadas propiedades psicométricas en estudios de validación en adultos de población general, en muestras clínicas, en adolescentes y estudiantes universitarios [10].

2. Metodología

La ilustración 1 muestra la metodología CRISP-DM, que son siglas de Cross-Industry Standard Process for Data Mining, es un método probado para orientar sus trabajos de minería de datos [11].

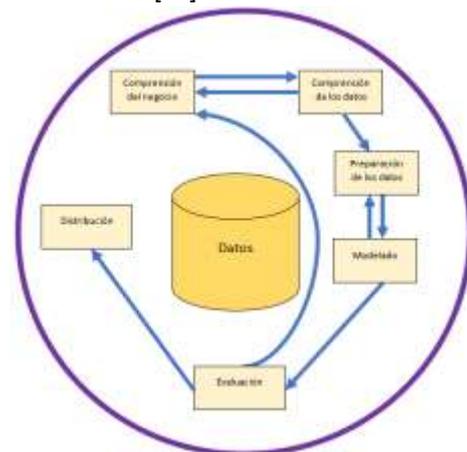


Ilustración 1 Metodología CRISP -DM

- Comprensión del negocio (Objetivos y requerimientos desde una perspectiva no técnica)
- Comprensión de los datos (Familiarizarse con los datos teniendo presente los objetivos del negocio)
- Preparación de los datos (Obtener la vista minable o dataset)
- Modelado (Aplicar las técnicas de minería de datos a los dataset)
- Evaluación (De los modelos de la fase anteriores hará determinar si son útiles a las necesidades del negocio)

- Despliegue (Explotar utilidad de los modelos, integrándolos en las tareas de toma de decisiones de la organización) [12]

3. Desarrollo

A continuación, se realiza la aplicación de la metodología CRIS-DM en el presente proyecto en cada una de sus fases.

3.1 Comprensión del Negocio

El principal objetivo de este proyecto es el de detectar por medio de la Escala DASS-21 la cual se muestra en la ilustración 2, trastornos psicológicos como ansiosos, depresivos y estrés, que podrían presentar los Alumnos de la Escuela Superior de Huejutla debido a la contingencia sanitaria a causa del COVID-19 que aqueja no solo a nuestro país sino al mundo entero en la actualidad y a la necesidad de permanecer en cuarentena en nuestros hogares. Utilizaremos el algoritmo K-means para realizar una agrupación de los K grupos de acuerdo a similitudes encontradas por la escala antes mencionada, llevando de este modo una descripción, cual nos permitirá conocer el o los trastornos que podrían estar presentes en los alumnos y de este modo realizar una intervención de manera oportuna por la institución o por padres de familia.

3.2 Comprensión de Datos

Los Datos utilizados se obtuvieron a partir de una encuesta virtual realizada a 154 alumnos de la Escuela Superior de Huejutla utilizando la escala DASS-21 la cual permite el diagnóstico de problemas psicológicos causados por el confinamiento en sus hogares, dicha encuesta consta de alrededor de 24 ítems los 3 primeros referentes a datos personales del alumno y los 21 ítems restantes al cuestionario se dividen en 3 grupos de 7 ítems referentes a cada padecimiento y cada ítem a su vez dividido en 4 posibles valores.

Para el presente trabajo solo se consideraron las encuestas resultas totalmente, con el objeto de eliminar outlier y datos ausentes que originan diagnósticos erróneos dentro del modelado.

DASS-21

La siguiente encuesta es con la finalidad de recabar datos para llevar a cabo un proyecto para la asignatura de Minería de datos del programa educativo de Ciencias Computacionales el cual consiste en determinar posibles trastornos psicológicos causados por la necesidad de permanecer en cuarentena en nuestros hogares a causa del COVID-19.

Sexo *

Mujer

Hombre

Nivel de estudio *

Bachillerato

Licenciatura

Fecha *

Día mes año

Por favor lee las siguientes afirmaciones y colócale un círculo alrededor de un número (0, 1, 2, 3) que indica en qué grado le ha ocurrido o usted esta afirmación durante la semana pasada. La escala de calificación es la siguiente:

0: No me ha ocurrido; 1: Me ha ocurrido un poco, o durante parte del tiempo; 2: Me ha ocurrido bastante, o durante una buena parte del tiempo; 3: Me ha ocurrido mucho, o la mayor parte del tiempo.

1.	Me ha costado mucho descargar la tensión	0	1	2	3
2.	Me di cuenta que tenía la boca seca	0	1	2	3
3.	No podía sentir ningún sentimiento positivo	0	1	2	3
4.	Se me hizo difícil respirar	0	1	2	3
5.	Se me hizo difícil tomar la iniciativa para hacer cosas	0	1	2	3
6.	Reaccioné exageradamente en ciertas situaciones	0	1	2	3
7.	Sentí que mis manos temblaban	0	1	2	3
8.	He sentido que estaba gastando una gran cantidad de energía	0	1	2	3
9.	Estaba preocupado por situaciones en las cuales podía tener pánico o en las que podría hacer el ridículo	0	1	2	3
10.	He sentido que no había nada que me ilusionara	0	1	2	3
11.	Me he sentido inquieto	0	1	2	3
12.	Se me hizo difícil relajarme	0	1	2	3
13.	Me sentí triste y deprimido	0	1	2	3
14.	No toqué nada que no me permitiera continuar con lo que estaba haciendo	0	1	2	3
15.	Sentí que estaba al punto de pánico	0	1	2	3
16.	No me pude entusiasmar por nada	0	1	2	3
17.	Sentí que valía muy poco como persona	0	1	2	3
18.	He sentido a sentirme confundido con facilidad	0	1	2	3
19.	Sentí los latidos de mi corazón a pesar de no haber hecho ningún esfuerzo físico	0	1	2	3
20.	Tuve estado sin razón	0	1	2	3
21.	Sentí que la vida no tenía ningún sentido	0	1	2	3

Ilustración 1 Encuesta DASS - 21

El DASS-21 posee tres subescalas, Depresión (ítems: 3, 5, 10, 13, 16, 17 y 21), Ansiedad (ítems: 2, 4, 7, 9, 15, 19 y 20) y Estrés (ítems: 1, 6, 8, 11, 12, 14 y 18). Para evaluar cada subescala por separado, se deben sumar las puntuaciones de los ítems correspondientes a cada una. Pueden obtenerse un indicador general de síntomas emocionales sumando las puntuaciones de todos los ítems.

Interpretación:

A mayor puntuación general, mayor grado de sintomatología.

Puntos de corte comúnmente utilizados:

- Depresión:
 - 5-6 depresión leve

- 7-10 depresión moderada
- 11-13 depresión severa
- 14 o más, depresión extremadamente severa.
- Ansiedad:
 - 4 ansiedad leve
 - 5-7 ansiedad moderada
 - 8-9 ansiedad severa
 - 10 o más, ansiedad extremadamente severa.
- Estrés:
 - 8-9 estrés leve
 - 10-12 estrés moderado
 - 13-16 estrés severo
 - 17 o más, estrés extremadamente severo.

3.3 Preparación de Datos

De los datos obtenidos anteriormente por la encuesta se realizaron 2 grupos en donde en el primer grupo encontramos la fecha de elaboración de cuestionario, sexo, nivel de estudios de los alumnos y por el otro los datos que tomamos en cuenta es decir solo los datos relativos a la encuesta es decir los resultados de la misma.

Lo elaborado en esta parte se refirió a la transformación de texto a número. Tomando en cuenta la siguiente escala:

- 0: No me ha ocurrido.
- 1: Me ha ocurrido un poco, o durante parte del tiempo.
- 2: Me ha ocurrido bastante, o durante una buena parte del tiempo.
- 3: Me ha ocurrido mucho, o la mayor parte del tiempo

En la ilustración 3 se muestra como se obtiene el archivo después de aplicar la encuesta a los alumnos, posteriormente con la finalidad de estandarizar los resultados se transforman en números enteros de acuerdo a la descripción anterior, utilizando el 0,1,2 y 3. Esta transformación de los datos podemos observarla en la ilustración 4.

Ilustración 2 Registro de Encuesta DASS -21

Sexo	Nivel de estudio	Fecha	1	Me	2	Me	3	No p.d.	Se m	Se	6	Re	7	Senti	q	0	Me	
Mujer	Licenciatura	23/05/2020	0	0	1	0	2	1	0	1	0	0	0	0	0	1	0	1
Mujer	Licenciatura	23/05/2020	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
Mujer	Licenciatura	23/05/2020	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mujer	Licenciatura	23/05/2020	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Mujer	Licenciatura	23/05/2020	1	2	2	0	1	3	0	0	0	0	0	0	0	0	0	1
Hombre	Licenciatura	23/04/2020	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hombre	Licenciatura	24/05/2020	2	0	2	2	3	2	2	2	2	2	2	2	2	2	2	2
Mujer	Licenciatura	24/05/2020	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mujer	Licenciatura	08/07/1998	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Mujer	Licenciatura	24/05/2020	1	1	1	1	1	2	0	1	2	0	1	0	1	0	1	1
Mujer	Licenciatura	24/05/2020	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1
Mujer	Bachillerato	24/05/2020	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hombre	Licenciatura	24/05/2020	1	1	3	0	3	1	0	3	1	0	3	0	3	0	3	3
Mujer	Licenciatura	24/05/2020	2	1	1	1	2	2	0	0	0	0	0	0	0	0	0	0
Hombre	Licenciatura	24/05/2020	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mujer	Licenciatura	24/05/2020	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mujer	Licenciatura	24/05/2020	1	1	1	1	1	0	2	1	0	2	1	0	2	1	2	1

Ilustración 3 Transformación de Registros de Encuesta DASS -21

Se reordenaron los datos de acuerdo a cada patología, en el cual cada una tuvo 7 ítems, cabe mencionar que cada patología está compuesta por 4 niveles de acuerdo a la gravedad del trastorno. En los registros generados se eliminaron espacios innecesarios y se sustituyeron por comas entre registros. Una vez obtenidos los grupos referentes a cada patología se procedió a dividirlos en sexo y escolaridad, esto con la intención de indagar más en la información recabada y de este modo conocer los niveles de trastorno de cada alumno.

A continuación, en la ilustración 5 se muestra la manera en la cual se formateo los datos en archivos de texto con extensión arff, esto con la intención explotarlo posteriormente en la herramienta WEKA para la aplicación del método correspondiente dentro de minería de datos.

```

DASS-21-Depresion...
Archivo Edición Formato Ver Ayuda
@relation DASS-21-Depresion
@attribute var1 numeric
@attribute var2 numeric
@attribute var3 numeric
@attribute var4 numeric
@attribute var5 numeric
@attribute var6 numeric
@attribute var7 numeric
@data
1,2,1,0,0,0,0
0,1,0,0,0,0,0
0,0,0,1,0,0,0
0,0,0,0,0,0,0
2,1,1,3,1,1,0
0,0,0,1,0,0,0
2,3,2,2,2,2,1
0,0,0,0,0,0,0
1,0,0,1,1,1,1
1,1,1,2,0,1,1
1,1,1,1,1,0,0
0,0,0,0,0,0,0
3,3,3,3,3,0,3
    
```

Ilustración 4 Preparación de Datos para Herramienta WEKA

3.4 Modelado

Como técnica de modelado para el proyecto se utilizará clustering con el algoritmo K-Means para realizar el agrupamiento lo cual conlleva identificar grupos en los datos, para esto se utilizará la herramienta WEKA. Se determino el parámetro k (número de Clúster) a cuatro esto derivado de los niveles de gravedad de la patología. En la ilustración 6 podemos observar la configuración del modelado de acuerdo a los parámetros establecidos.

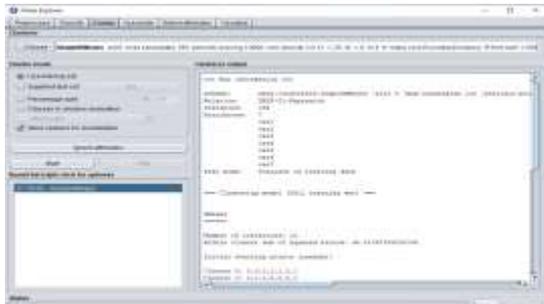


Ilustración 5 Modelado del Algoritmo K means en WEKA

Una vez ejecutado dentro de la herramienta WEKA de acuerdo al modelo establecido podemos visualizar en la Ilustración 7 la conformación de los cuatro cluster resultantes.

```
Initial starting points (random):
Cluster 0: 0,0,0,1,1,1,1
Cluster 1: 0,1,1,0,0,0,1
Cluster 2: 1,1,1,1,2,0,2
Cluster 3: 1,3,2,2,1,3,0

Missing values globally replaced with mean/mode

Final cluster centroids:
Attribute   Full Data   Cluster#
            (154.0)    (40.0)    (55.0)    (34.0)    (25.0)
-----
var1        0.9223      1         0.8545    1.6471    1.72
var2        1.0974      1.05      0.0727    2.8882    2.88
var3        0.9805      0.725    0.0727    1.6471    2.48
var4        1.0113      1.275    0.0909    2         1.28
var5        1.0065      1.25     0.0727    2.8882    1.2
var6        0.9223      0.825    0.0364    1.8824    1.72
var7        0.8806      0.8       0.0727    2.4118    0.76

Time taken to build model (full training data) : 0.01 seconds
=== Model and evaluation on training set ===

Clustered Instances
0: 40 ( 26%)
1: 55 ( 36%)
2: 34 ( 22%)
3: 25 ( 16%)
```

Ilustración 6 Corrimiento del Modelado en WEKA

3.5 Evaluación

Evaluación Depresión

En los resultados finales del algoritmo ejecutado por WEKA para el trastorno depresión se puede visualizar en la tabla 1, la cual muestra que se ingresaron 154 instancias correspondientes al 100% de los alumnos y 7 atributos correspondientes a los 7 ítems, el número de interacciones es de 10.

- En el cluster 0 correspondiente al 26% se puede observar que el nivel de depresión está por debajo de los puntos de corte determinados para esta patología. Por lo tanto, se puede concluir que los alumnos que se encuentran dentro de este cluster no sufren de depresión.
- En el cluster 1 correspondiente al 36% se puede observar que el nivel de depresión está por debajo de los puntos de corte determinados para esta patología. Por lo tanto, se puede concluir que los alumnos que se encuentran dentro de este cluster no sufren de depresión.
- En el cluster 2 correspondiente al 22% de los alumnos se puede observar que el promedio de puntos obtenido es de 8 por tanto su nivel de depresión corresponde a un nivel de depresión moderada.
- En el cluster 3 correspondiente al 16% de los alumnos se puede observar que el promedio de puntos obtenido es de 8 por tanto su nivel de depresión corresponde a un nivel de depresión severa.

Cluster 0	0,0,0,1,1,1,1
Cluster 1	0,1,1,0,0,0,1
Cluster 2	1,1,1,1,2,0,2
Cluster 3	1,3,2,2,1,3,0

Tabla 1 Conformación de Cluster DEPRESION

Evaluación Ansiedad

Para el trastorno ansiedad se puede visualizar en la tabla 2, se ingresaron 154 instancias correspondientes al 100% de los alumnos y 7 atributos correspondientes a los 7 ítems, el número de interacciones es de 6.

- En el cluster 0 correspondiente al 40% se puede observar que el nivel de ansiedad está por debajo de los puntos de corte determinados para esta patología. Por lo tanto, se puede concluir que los alumnos que se encuentran dentro de este cluster no sufren de Ansiedad.
- En el cluster 1 correspondiente al 25% de los alumnos se puede observar que el promedio de puntos obtenido es de 5 por tanto su nivel de ansiedad corresponde a un nivel de ansiedad moderada.
- En el cluster 2 correspondiente al 19% de los alumnos se puede observar que el promedio de puntos obtenido es de 9 por tanto su nivel de ansiedad corresponde a un nivel de Ansiedad severa.

- En el cluster 3 correspondiente al 17% de los alumnos se puede observar que el promedio de puntos obtenido es de 12 por tanto su nivel de ansiedad corresponde a un nivel de ansiedad extremadamente severa

Cluster 0	0,0,0,0,1,1,0
Cluster 1	0,1,0,1,1,1,1
Cluster 2	2,1,2,1,1,1,1
Cluster 3	2,2,0,0,3,2,3

Tabla 2 Conformacion de Cluster ANSIEDAD

Evaluación Estrés

Por último, el trastorno estrés se puede visualizar en la tabla 3, en donde ingresaron 154 instancias correspondientes al 100% de los alumnos y 7 atributos correspondientes a los 7 ítems, el número de interacciones es de 10.

- En el cluster 0 correspondiente al 38% se puede observar que el nivel de estrés está por debajo de los puntos de corte determinados para esta patología. Por lo tanto, se puede concluir que los alumnos que se encuentran dentro de este cluster no sufren de estrés.
- En el cluster 1 correspondiente al 10% se puede observar que el nivel de estrés está por debajo de los puntos de corte determinados para esta patología. Por lo tanto, se puede concluir que los alumnos que se encuentran dentro de este cluster no sufren de estrés.
- En el cluster 2 correspondiente al 30% de los alumnos se puede observar que el promedio de puntos obtenido es de 13 por tanto su nivel de estrés corresponde a un nivel de estrés severo.
- En el cluster 3 correspondiente al 21% de los alumnos se puede observar que el promedio de puntos obtenido es de 8 por tanto su nivel de estrés corresponde a un nivel de estrés leve

Cluster 0	0,0,0,1,1,1,1
Cluster 1	1,0,0,3,0,0,0
Cluster 2	2,2,2,1,1,2,3
Cluster 3	0,2,0,0,1,3,2

Tabla 3 Conformación de Cluster ESTRES

3.6 Despliegue (Resultados)

Depresión

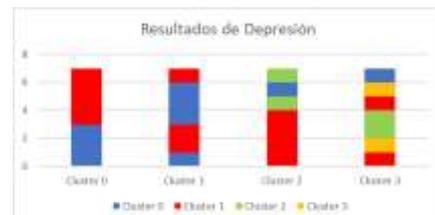
Tras haber ejecutado el algoritmo K-means a los datos obtenidos con la encuesta de la escala DASS-21

obtuvimos gran cantidad de conocimiento sobre los problemas psicológicos por los cuales los alumnos de la ESH tienden a padecer, en la gráfica 1 se puede observar la distribución de los clústers resultantes para el trastorno Depresión. El clúster 0 está conformado por el 26% y tiene un nivel de depresión leve, el clúster 1 está conformado por el 36%, con un nivel de depresión moderada y como se puede ver es el cluster con mayor porcentaje, el clúster 2 está conformado por el 22%, nivel de depresión severa y el clúster 3 está conformado por el 16%, el cual representa un nivel de depresión extremadamente severa.



Gráfica 1 distribución de Clúster DEPRESION

En la siguiente gráfica 2, se presenta la conformación de cada clúster.



Gráfica 2 Conformación de Cluster DEPRESIÓN

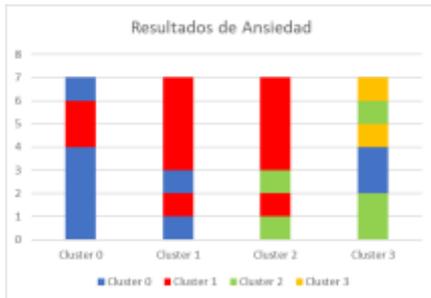
Ansiedad

En cuanto al trastorno Ansiedad se puede observar en la gráfica 3 la distribución de los clústers resultantes. El clúster 0 está conformado por el 39% y tiene un nivel de ansiedad leve y como se puede ver es el cluster con mayor porcentaje, el clúster 1 está conformado por el 25%, con un nivel de ansiedad moderada, el clúster 2 está conformado por el 19%, nivel de ansiedad severa y el clúster 3 está conformado por el 17%, el cual representa un nivel de ansiedad extremadamente severa.



Gráfica 3 Distribución Clúster ANSIEDAD

En la gráfica 4, se presenta conformación de los clúster resultantes.



Gráfica 4 Conformación Clúster ANSIEDAD

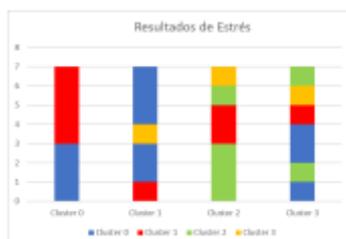
Estrés

Por último, en cuanto al trastorno del estrés, en la gráfica 5 se puede observar la distribución de los clústers resultantes. El clúster 0 está conformado por el 39% y tiene un nivel de estrés leve y como se puede ver es el cluster con mayor porcentaje, el clúster 1 está conformado por el 10%, con un nivel de estrés moderada, el clúster 2 está conformado por el 30%, nivel de estrés severa y el clúster 3 está conformado por el 21%, el cual representa un nivel de estrés extremadamente severa.



Gráfica 5 Distribución Clúster ESTRÉS

La conformación de los clústers de este trastorno lo podemos observar en la gráfica 6.



Gráfica 6 Conformación Clúster ESTRÉS

En resultados más específicos se pudo determinar que las mujeres en Depresión, un 19% tienen un nivel leve, un 38% presentan un nivel moderado, el 28% presentan un nivel severo y un 15% de la población presenta un nivel extremadamente severo.

Continuando con la misma población de mujeres en cuanto a la ansiedad se determinó que el 13% padecen un nivel leve, el 43% presentan un nivel moderado, el 16% padece un nivel severo y el 28% presentan un nivel extremadamente severo.

Por último, el estrés dentro de este mismo género se determinó que el 13% presentan un nivel leve, el 51% un nivel moderado, el 22% presentan un nivel severo y el 14% padecen de un nivel extremadamente severo.

De igual manera se realizó el análisis en los hombres en el cual podemos observar con un porcentaje de 29% muestran depresión severa, el 8% un nivel de depresión extremadamente severa y 39% presentan depresión moderada.

En cuanto la ansiedad se pudo observar que un 19% tiene un nivel leve, mientras que el 42%, presentan un nivel moderado, el 32% presentan un nivel severo y el 7% un nivel extremadamente severo.

En cuanto al estrés, el 19% de los hombres padecen de un nivel leve, el 39% presentan un nivel moderado, el 22% tiene un nivel severo y el 20% de hombres presentan un nivel extremadamente severo.

Otro de los análisis que se realizaron fue con respecto al nivel de escolaridad que cursan los alumnos, específicamente el nivel bachillerato y nivel licenciatura.

Primeramente, en licenciatura se determinó que el 17% padecen de un nivel de depresión leve, el 36% presentan un nivel de depresión moderada, el 26% padecen de un nivel de depresión severa y el 21% un nivel de depresión extremadamente severa.

En cuanto a la ansiedad el 26% padecen un nivel leve, el 37% presentan un nivel moderada, el 21% padecen de un nivel severo y el 16% de presenta un nivel extremadamente severo.

Para el trastorno del estrés, el 20% un nivel de leve, el 39% un nivel moderado, el 27% un nivel severo y el 18% un nivel extremadamente severo.

En cuanto al bachillerato, se presenta una depresión moderada en la mayoría de los alumnos, el 41% representan un nivel de ansiedad extremadamente severa, y el 37% de presentan un nivel de estrés moderada.

4. CONCLUSIONES Y TRABAJOS FUTUROS

Del proyecto realizado sobre la descripción de trastornos psicológicos en alumnos de la Escuela Superior de Huejutla, es importante comentar que las graficas mostradas en este documento obedecen a la incidencia general de los trastornos, sin embargo, se llevo a cabo muestras que representan el 10% del total de alumnos que nos permitieron conocer otros factores de incidencia.

La implementación de la escala DASS-21 para la evaluación de los alumnos se llevó a cabo de manera virtual. Dentro de los resultados podemos concluir que el 60% de los alumnos está padeciendo Ansiedad, el 52% está padeciendo de Estrés y el 38% padece Depresión desde el punto de vista General.

En el contexto de género se pudo observar que las mujeres tienden más a mostrar trastornos de estrés, siguiendo la ansiedad y por ultimo la depresión. Mientras que en los hombres el trastorno más recurrente es la ansiedad, y de igual forma y en segundo sitio la depresión y el estrés.

En cuanto al nivel de escolaridad se puede observar que en los alumnos de bachillerato se muestra en primer lugar el trastorno de ansiedad, siguiendo el estrés y por ultimo la depresión. Para el caso de Licenciatura primeramente se da el estrés, después la ansiedad y por último la depresión.

Es indudable que con esta información se pueden emprender diferentes acciones encaminadas a abatir estos trastornos psicológicos que actualmente ocurren con nuestros alumnos. Se pueden implementar acciones para apoyar desde el Programa Institucional de Tutorías y a través de línea Garza apoyo psicológico para los alumnos.

De igual manera desde el punto de vista de la Plataforma Garza se podrían describir los comportamientos que tienen los alumnos dentro de ella, esto para sugerir o incluir diferentes recursos didácticos que nos permitan como universidad lograr mejores aprendizajes significativos en nuestros educandos.

Referencias

- [1] Portilla, F. E. P., Galvan, D. A. C., Blandón, K. M. H., Silva, M. F. R., Chaustre, J. A. J., & León, M. A. A. (2020). Depresión, ansiedad y estrés: Una situación en estudiantes de Enfermería en una Universidad Pública. 16 de Abril, 59(278), 1002.
- [2] Antúnez, Z., & Vinet, E. V. (2013). Problemas de salud mental en estudiantes de una universidad regional chilena. Revista médica de Chile, 141(2), 209-216
- [3] Hernández Melián, C. (2020). Impacto psicológico de la pandemia del COVID-19 en los profesionales sanitarios del Área Sanitaria de Coruña.
- [4] Román, F., Santibáñez, P., & Vinet, E. V. (2016). Uso de las escalas de depresión ansiedad estrés (DASS-21) como instrumento de tamizaje en jóvenes con problemas clínicos. Acta de investigación psicológica, 6(1), 2325-2336.
- [5] Jaramillo, A., & Arias, H. P. P. (2015). Aplicación de Técnicas de Minería de Datos para Determinar las Interacciones de los Estudiantes en un Entorno Virtual de Aprendizaje. Revista Tecnológica-ESPOL, 28(1).
- [6] Cáceres, J. H. (2015). Clustering basado en el algoritmo K-means para la identificación de grupos de pacientes quirúrgicos. Trabajo de investigación, Universidad Santo Tomás, seccional Bucaramanga. Bucaramanga, Colombia..
- [7] Orallo, J. H., Quintana, M. J. R., & Ramírez, C. F. (2004). Introducción a la Minería de Datos. Pearson Educación.
- [8] tachopedia. (s.f.). Obtenido de <https://www.techopedia.com/definicion/32057/k-means-clustering>
- [9] Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behaviour research and therapy, 33(3), 335-343.
- [10] Tully, P. J., Zajac, I. T., & Venning, A. J. (2009). The structure of anxiety and depression in a normative sample of younger and older Australian adolescents. Journal of abnormal child psychology, 37(5), 717.
- [11] Cárdenas, F. D. J. N., Hernández, E. A., Zenil, A. E. G., & Redondo, A. M. F. (2020). Application Method of Data Mining Using the K means Algorithm for the Determination of Stress Level in High School Students Using the Beck Depression Inventory. Ciencia Huasteca Boletín Científico de la Escuela Superior de Huejutla, 8(15), 1-8.
- [12] Goicochea, A. (2009). CRISP-DM. Una Metodología para Proyectos de Minería de Datos. [En línea]. Disponible en la web:<<http://anibalgoicochea.com/2009/08/11/crisp-dm-una-metodologia-paraproyectos-de-mineria-de-datos/>>..