

TIPIFICACIÓN DE PERSONALIDAD EN LOS ALUMNOS QUE INGRESAN A LICENCIATURA EN PSICOLOGÍA MEDIANTE TÉCNICAS DE MINERÍA DE DATOS.

Felipe de Jesús Núñez Cárdenas¹, Alain Ezau Guerrero Zenil², Ana María Felipe Redondo³,
Eduardo Aguirre Hernández⁴

¹*Maestro de tiempo completo de la Universidad Autónoma del Estado de Hidalgo, México, *
correspondencia: felipe_nunez@uaeh.edu.mx*

²*Alumno de la Licenciatura en Ciencias Computacionales de la Universidad Autónoma del Estado de Hidalgo, México, * correspondencia:
alaguerrero96@gmail.com*

³*Maestro de tiempo completo de la Universidad Tecnológica de la Huasteca Hidalguense, México, *
correspondencia: ana.felipe@uthh.edu.mx*

⁴*Alumno de la Licenciatura en Ciencias Computacionales de la Universidad Autónoma del Estado de Hidalgo, México, * correspondencia: eduar030696@gmail.com*

Resumen— En este artículo se describe el tipo de personalidad de los alumnos (hombres y mujeres) de nuevo ingreso a la licenciatura en psicología, basándose en el cuestionario factorial de personalidad de Raymond B. Cattell 16PF, con la técnica de clustering de minería de datos. Esto con el objetivo de conocer más a fondo la diversidad de personalidades que hay dentro de los alumnos de dicha carrera, y cuáles son los estándares de cada uno de las personalidades, tomando en cuenta el género (hombre, mujer). con el algoritmo K means se analizaron los resultados del test para poder descubrir patrones de un conjunto de datos, para aportar información valiosa en la toma de futuras decisiones sobre dichos alumnos..

también conocido como 16PF, creado originalmente por el psicólogo Raymond Cattell.

El presente artículo busca describir la personalidad de los alumnos (hombres y mujeres) de nuevo ingreso a la licenciatura en psicología, basándonos en el cuestionario factorial de personalidad de Raymond B. Cattell 16PF, con la técnica de clustering de minería de datos. Con el cual con esta técnica se busca descubrir patrones de un conjunto de datos, para aportar información en la toma de futuras decisiones encaminadas a atender a estos alumnos.

Palabras clave — Cuestionario 16PF, CRIPS – DM, Minería de datos, Personalidad.

Con el objetivo de conocer la diversidad de personalidades dentro de los alumnos de la carrera de psicología y cuáles son los estándares en cada una de las personalidades, tomando en cuenta el género (hombre, mujer), sin importar la edad, con el único requisito de ser nuevo ingreso en dicha carrera.

I. INTRODUCCIÓN

Cada uno de nosotros tiene su propia forma de ser. Observamos el mundo de determinada manera, nos relacionamos con los demás de formas concretas cada uno de distinta manera y, en general, expresamos una tendencia a hacer determinadas cosas y a responder de maneras más o menos estables.

Dicho de otra manera, y aunque suene redundante, cada persona tiene su propia personalidad. Este concepto, que define quién y cómo somos, a lo largo de los años ha sido un objeto de estudio clásico de la psicología, habiéndose creado numerosos instrumentos de medida para evaluarlo conocidos como tests de personalidad.

De entre todos ellos en este artículo se usará el “Cuestionario Factorial de la Personalidad o Test de los 16 Factores de la Personalidad”,

Por otro lado, la minería de datos es una herramienta que ayudara a comprender el contenido de los datos, con el fin de sacar información valiosa que pueda contribuir a la psicología de tal manera que aporte la información que a simple vista no se pueda observar, puesto que se puede contener grandes cantidades de información y no saber qué hacer con ello, es por eso que con la minería de datos se podrá descubrir cuáles son las personalidades que más radican en la licenciatura de psicología.

Como ya sabemos la personalidad puede variar en algún aspecto en respuesta a situaciones y evoluciones vitales concretas, pero por lo general se mantiene a lo largo del ciclo vital, observándose constante en la mayor parte de los ámbitos y a través de las diferentes situaciones que vivimos. Esto no quiere decir que aspectos

concretos sean inmodificables, pero requiere de un elevado nivel de esfuerzo y trabajo, manteniéndose en general el conjunto de características que configuran la personalidad.

Por lo tanto, se cree importante el tomar atención a este test para poder observar y trabajar en conjunto con la minería de datos, mediante sus indicaciones y poder darnos cuenta de la magnitud de las personalidades y cuales predominan más en cada uno de los géneros.

II. ESTADO DEL ARTE

En el año 2006, José C. Riquelme de España en su artículo Minería de datos: conceptos y tendencias, menciona que la minería de datos ha sido la revolución para descubrir la información o conocimiento de enormes volúmenes de datos, buscando la explosión de información que actualmente pudiese ser almacenada en una base de datos, tales como: imagen, video, texto y otros datos numéricos, facilitando el procesamiento multimedia como resultado, obteniendo información y el conocimiento de los datos distribuidos por todo el mundo, puesto que la demanda continuara creciendo y el acceso a los grandes volúmenes de datos multimedia traerá la mayor transformación para la sociedad. Así mismo utilizando las herramientas de minería de datos KDD se puede obtener conocimiento de cualquier información. [1]

Posteriormente en el 2007, Elena Durán Rosanna Costagua de Argentina en su proyecto minería de datos para descubrir estilos de aprendizaje, utiliza la minería de datos para la determinación de las características del perfil de aprendizaje de los estudiantes con la finalidad de alinear las estrategias de enseñanza, se puede llegar a descubrir el conocimiento que a simple vista no se puede percibir, para poder detectar el estilo de aprendizaje dominante en la carrera de la licenciatura en sistemas de información. Mediante técnicas de minería de datos, en especial uso clustering. [2]

Por otro lado, en el 2010, Álvaro Jiménez Gilindo de Madrid en su trabajo minería de datos en la educación, menciona que la minería de datos, también conocida como descubrimiento de conocimiento de bases de datos, es el campo que nos permite descubrir información nueva y potencialmente útil de grandes cantidades de datos, el cual se han empleado en numerosos campos incluyendo desde los ya conocidos casos de cesta de la compra hasta la bioinformática, educación, psicología o investigaciones contra el terrorismo, etc. puesto que se ha incrementado el interés en utilizar la minería de datos en el estudio educacional, centrándose en el desarrollo

de métodos de descubrimiento que utilicen los datos de plataformas educacionales y en el uso de esos métodos para comprender mejor a los estudiantes y el entorno en el que aprenden. Utilizando predicción, agrupamiento, minería de relaciones, descubrimiento mediante modelos, entre otros. [3]

Henry Fernando Vallejo Ballesteros en el 2017, en su artículo minería de datos, hace mención en su idea principal que la minería de datos tiene como objetivo analizar los datos desde todas las perspectivas estratégicas para la organización, con el fin de transformarla en información útil y conocimiento, siendo de utilidad general para aumentar la facturación, ampliar el margen operativo, etc. usando técnicas de minería de datos, se pueden potencializar las empresas. [4]

En el proyecto identificación de estilos de aprendizaje en los alumnos universitarios de computación de la huasteca hidalguense mediante técnicas de minería de datos, habla sobre la aplicación de técnicas de minería de datos para descubrir las combinaciones de estilos de aprendizaje mostradas por estudiantes de educación superior de computación o áreas a fines en la región de la huasteca hidalguense. Utilizando el modelo VARK y la herramienta de minería de datos Agrupamiento, clustering. [5]

2.2. PERSONALIDAD

A principios de 1999, Esperanza Herrera García en su investigación titulada factores implicados en el rendimiento académico de los alumnos universidad de salamanca, en vista que su idea principal es describir, conocer y analizar al alumnado de la universidad de salamanca, para considerar el rendimiento académico universitario, así mismo conocer distintas circunstancias el cual provoquen el rendimiento académico de los alumnos de dicha universidad, para ello utilizo dos tipos de encuestas para diagnosticar la personalidad, las cuales son 16PF y TIG. [6]

Posteriormente en el 2001 J. Cebria de Barcelona en su investigación titulada rasgos de personalidad y burnout en médicos de familia, da a conocer una investigación sobre los rasgos de personalidad que podrían asociarse con puntuaciones altas en las escalas de burnout (Maslach Burnout Inventory) en médicos de atención primaria, utilizando el método de encuesta directa por entrevista personal o grupal mediante las herramientas: Maslach Burnout Inventory (MBI) y el test de personalidad de Cattell 16 PF. [7]

En cuanto al proyecto predictores de rendimiento académico en la escuela media:

habilidades, autoeficacia y rasgos de personalidad, que se realizó en el 2005 por Edgardo Pérez de Argentina, da a conocer la contribución de estos constructos psicológicos a la explicación de la variable del rendimiento académico en lengua y matemática, en una muestra de estudiantes secundarios de la ciudad de Córdoba, Argentina. Utilizando una encuesta mediante el cuestionario 16PF. [8]

En el 2011, la psicóloga Nora Elena Gil de Colombia, en su investigación revisión sistemática del cuestionario factorial de personalidad (16PF), menciona que se debe establecer el uso e impacto de la prueba de personalidad 16PF, para determinar la vigencia de dicho test, puesto que la utilidad de las pruebas psicológicas han sido ampliamente demostradas en la historia de la psicología especialmente la de las pruebas de personalidad las cuales han tenido un puesto privilegiado en la configuración actual de la evaluación psicológica. [9]

2.3. KMEANS

En el 2015, M. Amador García de México, en su proyecto titulado Determinación de perfiles delictivos en el estado de Jalisco utilizando WEKA a través de minería de datos, el cual su idea principal, nos comenta que en dicho proyecto utiliza la herramienta WEKA con sus algoritmos Random Forest, KMeans y A priori para determinar los perfiles delictivos en los estados de Jalisco. [10]

Por otro lado, en el 2016, Gutiérrez, H. G. de México en su proyecto Análisis de un caso práctico aplicando el algo K-means mediante weka, en donde hace uso de un algoritmo de agrupamiento k-means, utilizado para encontrar los k grupos más representativos de una colección de datos provenientes de una BD, la intención de utilizar este algoritmo radica en el interés por determinar los niveles de reforestación en la República Mexicana durante el transcurso del periodo 1994 – 2011, tomando como fuente de información fidedigna un banco de datos proporcionados por INEGI, para ello utilizo el algoritmo k-means y la herramienta WEKA. [11]

En el mismo año Pineda Torres de Colombia en su proyecto titulado Discriminación de texturas en imágenes inalámbricas utilizando el Algoritmo KMeans, hace mención que el procesamiento de imágenes y la clasificación de texturas es uno de los principales retos, imágenes y videos son enviados de forma inalámbrica, se plantea el uso de los espacios tridimensionales sobre el análisis y adecuación del algoritmo KMeans, cuyos valores de los centroides son

ubicados para la respectiva discriminación de las texturas. [12]

Palomino, R. P. en el 2016, en su proyecto Model clustering neural network monitoring application No K-Means algorithm for the segmentation of the peruvian university students union, habla sobre la identificación del perfil del alumno moroso utilizando el algoritmo de clusters basado en la red neuronal de k-means para la segmentación de los alumnos de la UPeU, de manera que existe mayor homogeneidad posible dentro de los grupos, con respecto a ciertos atributos que pueden ser cuantitativas o cualitativas. [13]

En el 2017, Carmita del Rocío Echeverría Ruiz de Ecuador, en su proyecto Caracterización del desarrollo de los indicadores de patentes en América Latina, el cual trata de identificar el desempeño de los indicadores fundamentales en relación con las patentes y tales como: solicitudes de patentes, patentes otorgadas, tasa de dependencia y tasa de autosuficiencia, para este propósito se desarrollara un análisis estadístico empleando el software SPSS y el algoritmo KMeans, empleando los datos proporcionado por la red de indicadores de ciencia. [14]

2.4. MINERÍA DE DATOS EN LA PERSONALIDAD

En el 2007, Francisco Xavier Campaña Naranjo de Ecuador en su proyecto aplicación de técnicas de data mining a bases de datos de contenido musical para identificar rasgos de personalidad de los usuarios en el distrito metropolitano de Quito, habla sobre el impacto que ha influenciado la música en la civilización humana haciendo énfasis en las notorias secuelas que marcan a las personas durante su vida tanto a nivel personal como social, cultural, etc. para posteriormente identificar un mecanismo para poder evaluar los rasgos de personalidad. Utilizando el modelo de OCEAN aplicando técnicas de minería de datos e inteligencia de negocios. [15]

Por otro lado, en el 2017 Elena Álvarez de España en su proyecto de investigación titulado estilos de aprendizaje y rendimiento académico en asignaturas universitarias orientadas al aprendizaje de herramientas informáticas impartidas en formato virtual. Da a conocer un análisis de la relación entre los estilos de aprendizaje, y el rendimiento académico y la actividad de los estudiantes de la asignatura no presencial Introducción a los métodos multimedia impartida en la universidad de Cantabria. Con la finalidad de determinar si es común en la formación virtual que el rendimiento se relacione con un estilo de aprendizaje común

o si bien el estilo de aquellos alumnos que aprueban aparecer vinculado a la materia que trate el curso. Para ello utiliza el modelo de estilo de Felder y Silverman, en conjunto con la técnica de minería de datos KMeans. [16]

III. METODOLOGÍA

Para el presente proyecto se utilizó la Metodología CRIPS – DM (Cross Industry Standard Process for Data Mining). La cual consiste en seis fases como se muestra en la siguiente ilustración.

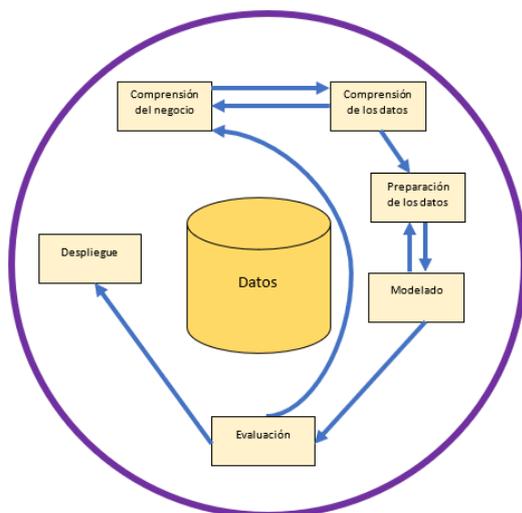


Ilustración 1 Metodología CRIPS – DM.

1. Fase 1 Comprensión del negocio.
2. Fase 2 Entendimiento de datos.
3. Fase 3 Preparación de datos.
4. Fase 4 Construcción del modelo.
5. Fase 5 Evaluación.
6. Fase 6 Despliegue.

IV. DESARROLLO

La metodología que se usó para trabajar la presente investigación fue la CRIPS – DM, la cual se puede trabajar con cualquier herramienta para desarrollar proyecto de minería de datos.

4.1. COMPRESIÓN DEL CASO DE ESTUDIO

El Centro Universitario Vasco de Quiroga se encuentra ubicado en la ciudad de Huejutla de Reyes Hidalgo, en el se imparten diversas licenciaturas, las cuales son: Arquitectura, Ciencias Políticas y administración pública, Educación, Trabajo Social, y Psicología.

Población: para llevar a cabo el presente trabajo se consideraron todos los alumnos de nuevo ingreso que a través de las diferentes

generaciones entran a la carrera de psicología, por lo que se cuenta con una población total de 441 alumnos, hombres y mujeres.

El instrumento para la recolección de la información fue el cuestionario factorial de la personalidad 16PF, el cual consiste en responder 187 preguntas, una vez finalizado el cuestionario se procederá a evaluar su puntuación bruta para posteriormente sacar la puntuación estándar, dependiendo de su edad es con la norma que se obtendrá su puntuación, las normas para mujeres y hombres de 17 años de edad son las normas para estudiantes de educación media y media superior, para hombres y mujeres de 20 años de edad son las normas para estudiantes universitarios, y para adultos con 30 años de edad son las normas para población general. Para posteriormente obtener su puntuación en estenes (Promedio).

4.2. ENTENDIMIENTO DE DATOS

Para esta etapa se determinaron los resultados para la obtención de los datos del test 16PF, para poder extraer los datos del test se creó un vector el cual nos permitirá visualizar el promedio estándar, que está en un rango de 1 al 10, y el factor, dependiendo del promedio es como se clasifico su factor, cabe mencionar que cada factor tiene su propia descripción, también cabe resaltar que si el promedio llegara hacer mayor a 5 tendrá factores descriptivos positivos caso contrario la descripción será negativo. A continuación, se presenta un ejemplo de cómo se asignaron los valores.

A	Afabilidad
B	Razonamiento
C	Estabilidad
E	Dominancia
F	Animación
G	Atención a las normas
H	Atrevimiento
I	Sensibilidad
L	Vigilancia
M	Abstracción
N	Privacidad
O	Aprensión
Q1	Apertura al cambio
Q2	Autosuficiencia
Q3	Perfeccionismo
Q4	Tensión

Ilustración 2 Asignación de los factores a variables

La siguiente ilustración muestra el instrumento de evaluación, el cual es el cuestionario factorial de la personalidad 16 PF, únicamente se toma el resultado del perfil del

cuestionario, los valores obtenidos de este perfil fueron las puntuaciones estándar.

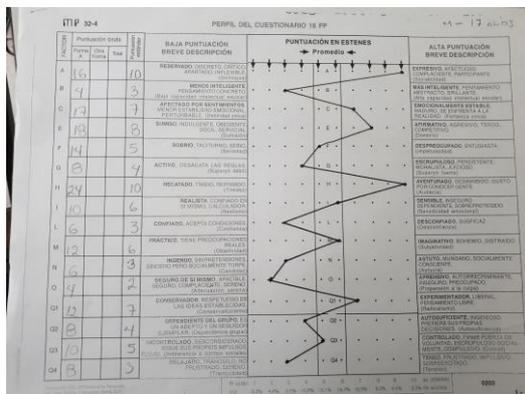


Ilustración 3 Perfil del cuestionario 16 PF

Posteriormente se obtendrá un vector con los resultados del promedio estándar, el cual la primera columna pintada de color verde, serán los alumnos hombres y la columna pintada de color rosa será para las alumnas, la primera fila que está pintada de color marrón claro son las variantes de los factores, por último, los restantes que están de color blanco son los promedios que obtuvieron los alumnos, en sus diferentes factores.

	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4
1	10	3	7	8	5	4	10	6	3	6	3	2	7	4	5	3
2	9	5	8	3	4	8	8	7	4	3	6	5	4	1	6	1
3	5	1	5	1	3	7	6	4	6	3	6	6	6	6	3	3
4	6	6	7	4	5	6	6	5	4	2	9	3	7	7	8	5

Ilustración 4 Vector con los resultados del promedio estándar para los diferentes factores, en los alumnos hombres

	A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4
15	7	2	10	4	4	5	10	7	3	4	9	4	5	5	5	5
16	7	4	10	3	2	9	8	4	2	5	10	5	4	9	9	4
17	3	5	4	3	4	4	4	4	6	3	2	3	9	8	8	8
18	8	5	6	9	3	5	6	2	6	7	5	5	8	5	6	6
19	7	5	8	2	4	8	9	2	5	3	2	3	4	6	6	5
20	8	3	6	3	4	7	7	4	7	8	10	6	5	9	9	4

Ilustración 5 Vector con los resultados del promedio estándar para los diferentes factores, en las alumnas.

Con estos resultados podremos describir los diferentes factores de personalidad que posee cada alumno.

4.3. PREPARACIÓN DE LOS DATOS

Una vez que ya se hayan obtenido los resultados del cuestionario 16PF, estos resultados se acumularan en un archivo de texto, para el análisis grupal, posteriormente se utilizó un editor para crear un formato de tipo (. arff) el cual es necesario para manejar la herramienta WEKA.

Para crear el archivo (.arff), primeramente se colocó en la primera línea la relación seguido del nombre del archivo que en este caso se llama Personalidad (@relation Personalidad), como segundo paso se crean los atributos seguido del nombre (que en este caso son las variables de los factores) y el tipo de dato que es (@attribute A

integer), si en dado caso llegase ser un atributo que solo tenga 2 atributos de tipo String se hará de la siguiente manera puesto que en este caso se utilizó para el sexo (@attribute sexo {H,M}), por último, los datos el cual se generan en orden de acuerdo a los atributos pero de manera horizontal. Podemos tomar de ejemplo la siguiente ilustración.

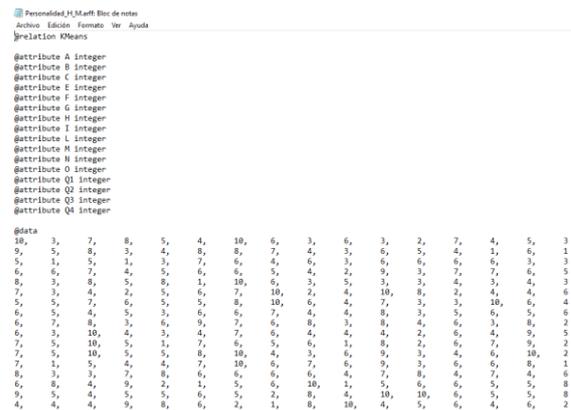


Ilustración 6 Integración de los datos, a un archivo (. arff)

4.4. MODELADO

Para este proyecto se decidió utilizar técnicas de minería de datos del tipo descriptivo, en lo particular se utilizó clustering para identificar grupos homogéneos dentro de la muestra de los alumnos encuestados. Para ello se utilizó el software WEKA.

Para la ejecución se seleccionó la opción Clúster, por lo que se seleccionó el algoritmo kmeans, por lo que solo se pretende detectar las personalidades que más destacan en la carrera de psicología y cuál es la que menos carece.

A continuación, en la siguiente ilustración se presentan el agrupamiento de los alumnos (hombres y mujeres) que ingresaron a dicha carrera, en el cual se observa que tipo de personalidad hay entre el total de esos alumnos.

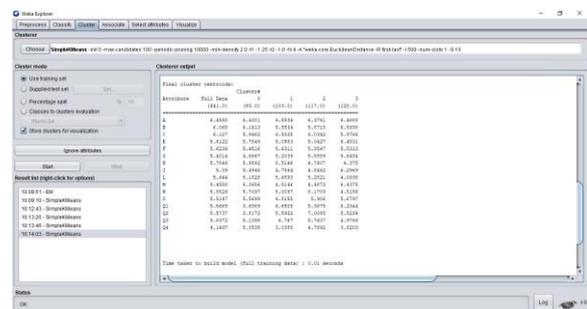


Ilustración 7 Corrida del Modelo KMeans, Agrupamiento entre los alumnos (Hombres y Mujeres), destacando la personalidad que más abunda.

4.5. EVALUACIÓN

Dentro de la construcción de los clústers, WEKA creo 22 iteraciones con una semilla de 10, el clúster 0 está conformado por el 21% de los encuestados, el clúster 1 está conformado por el 23%, el clúster 2 está conformado por el 27% y el clúster 3 está conformado por el 29%, lo cual en dicho resultado podremos observar que la mayor parte de los alumnos está en el clúster 3.

En la construcción del modelo KMeans, se registraron 441 datos de los cuales conformaba la vista minable, con la herramienta WEKA, en donde el software creo 22 iteraciones con puntos de partida inicial aleatorio, una semilla de 10 y utilizando una distancia euclidiana con 500 iteraciones, la segunda pasada con una semilla de 20 el software creo 18 iteraciones en la cuales, se crearon 15 agrupamientos de las cuales corresponden al tipo de personalidad. La evaluación se desarrolló sobre datos de entrenamiento representada en la siguiente tabla del primer recorrido:

Atributo / Clúster	Full Data	0	1	2	3
A	6.4558	6.4301	6.5534	6.3761	6.4688
B	6.068	6.1613	5.5534	5.8718	6.5938
C	6.127	5.9462	6.5825	6.0342	5.9766
E	5.6122	5.7849	5.0583	5.0427	6.4531
F	5.6236	5.4516	5.6311	5.8547	5.5313
G	5.4014	4.6667	5.2039	5.8889	5.6484
H	5.7846	5.9892	8.5146	4.7607	4.375
I	5.39	5.4946	4.7864	4.8462	6.2969
L	5.644	8.1828	5.6893	5.2821	4.0938
M	5.4558	6.3656	4.5146	4.4872	6.4375
N	5.8526	5.7097	5.0097	8.1709	4.5156
O	5.5147	5.5699	4.8155	5.906	5.6797
Q1	5.5669	3.6989	6.6505	5.3675	6.2344
Q2	5.5737	3.8172	5.5922	7.0085	5.5234
Q3	5.6372	5.1398	6.767	5.7607	4.9766
Q4	4.1497	5.0538	3.0388	4.7692	3.8203

Ilustración 8 Evaluación del desarrollo sobre datos de entrenamiento.

4.6. DESPLIEGUE

A continuación, se presenta la integración de los diferentes agrupamientos, así como su composición de las diferentes personalidades que más destacaron durante la explotación de los resultados de los cuestionarios aplicados a los alumnos (Hombres) de nuevo ingreso a la carrera de psicología.

Antes se presenta una gráfica de pastel en donde se muestran los porcentajes de cada clúster, primeramente, se muestra el de los hombres.

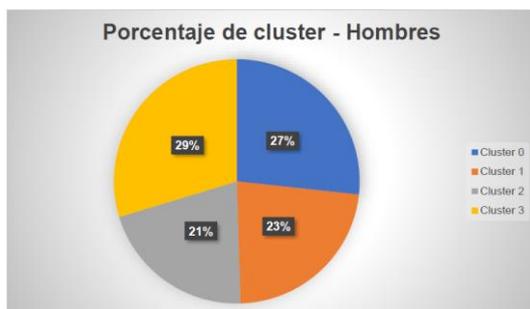


Ilustración 9 Porcentaje de clúster – Hombres.

Dentro de la construcción de los clústers, WEKA creo 12 iteraciones con una semilla de 10, el clúster 0 está conformado por el 27% de los alumnos (hombres) encuestados, el clúster 1 está conformado por el 23%, el clúster 2 está conformado por el 21% y el clúster 3 está conformado por el 30%, lo cual en dicho resultado podremos observar que la mayor parte de los alumnos está en el clúster 3.

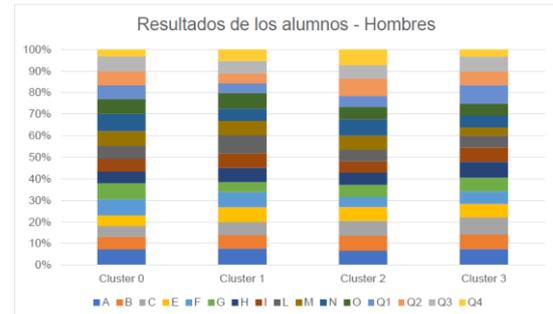


Ilustración 10 Resultado de los alumnos – Hombres.

A continuación, se mostrará la personalidad que más destaca de los alumnos (hombres) hasta la personalidad que más carece: en el puesto número uno se encuentra el Atravimiento, posteriormente le sigue; la Abstracción, Sensibilidad, Vigilancia, Privacidad, Animación, Razonamiento, Apertura al cambio, Estabilidad, Aprensión, Afabilidad, Perfeccionismo, Dominancia, Autosuficiencia, Tensión, Atención a las normas.

Dentro de la construcción de los clústers, WEKA creo 25 iteraciones con una semilla de 10, el clúster 0 está conformado por el 28% de los alumnos (mujeres) encuestados, el clúster 1 está conformado por el 30%, el clúster 2 está conformado por el 19% y el clúster 3 está conformado por el 23%, lo cual en dicho resultado podremos observar que la mayor parte de los alumnos está en el clúster 1.



Ilustración 11 Porcentaje de clúster - Mujeres.

Ahora se presenta la integración de los diferentes agrupamientos, así como su composición de las diferentes personalidades que más destacaron durante la explotación de los resultados de los cuestionarios aplicados a los

alumnos (Mujeres) de nuevo ingreso a la carrera de psicología.

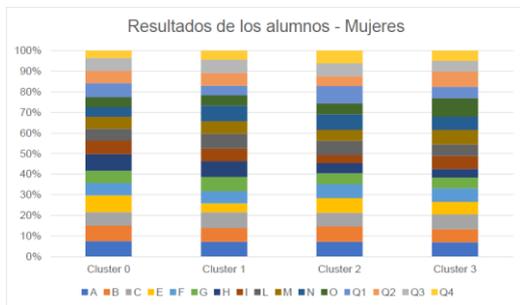


Ilustración 12 Resultados de los alumnos - Mujeres.

A continuación, se mostrará la personalidad que más destaca de los alumnos (Mujeres) hasta la personalidad que más carece: en el puesto número uno se encuentra la privacidad, posteriormente se encuentra el Razonamiento, Afabilidad, Apresión, Tensión, Dominancia, Atención a las normas, Abstracción, Perfeccionismo, Vigilancia, Apertura al cambio, Sensibilidad, Estabilidad, Animación, Autosuficiencia, Atrevimiento. Por último, se creó otra grafica más para que se pudiera describir y saber que personalidad es la que más destaca entre los alumnos (hombres y mujeres) de nuevo ingreso a la carrera de psicología.

Dentro de la construcción de los clústers, WEKA creo 25 iteraciones con una semilla de 10, el clúster 0 está conformado por el 21% de los alumnos (hombres y mujeres) encuestados, el clúster 1 está conformado por el 23%, el clúster 2 está conformado por el 27% y el clúster 3 está conformado por el 29%, lo cual en dicho resultado podremos observar que la mayor parte de los alumnos está en el clúster 3, dicho de otra manera podremos observarlo en la siguiente gráfica.

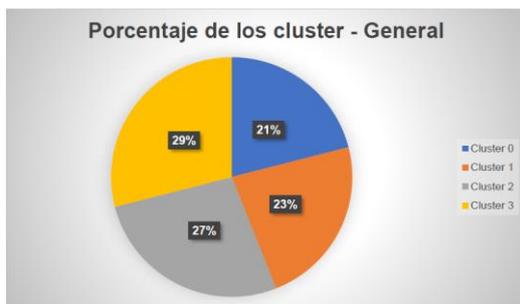


Ilustración 13 Porcentaje de clústeres - General.

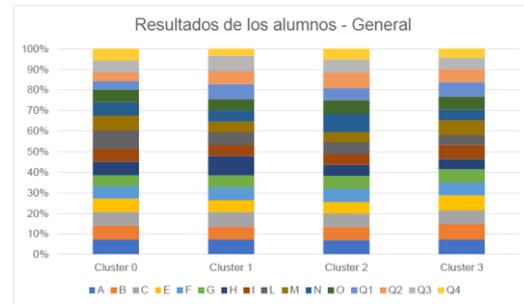


Ilustración 14 Resultados de los alumnos - General.

A continuación, se mostrará la personalidad que más destaca hasta la personalidad que más carece: en el puesto número uno está la Estabilidad, posteriormente le sigue la dominancia, atención a las normas, Abstracción, Razonamiento, Animación, Sensibilidad, Apertura al cambio, Autosuficiencia, Tensión, Afabilidad, Atrevimiento, Vigilancia, Privacidad, Apresión, Perfeccionismo.

V. ANALISIS DE RESULTADOS

Es importante comentar que la aplicación del Test de Personalidad 16PF, se hizo en condiciones estables y de igualdad entre todos los participantes dentro de la muestra tomada para tal fin, por lo que podemos apreciar que, desde el punto de vista de género, podemos afirmar que los hombres que entran a la licenciatura en Psicología tienden a tener una personalidad de Atrevimiento, mientras que las Mujeres es la Privacidad.

Sin embargo, cuando sacamos el conocimiento de los datos históricos de nuevo ingreso a esta licenciatura podemos observar de manera genérica que la personalidad que mas predomina es la Estabilidad, mientras que la de menor predominancia es la Perfeccionista.

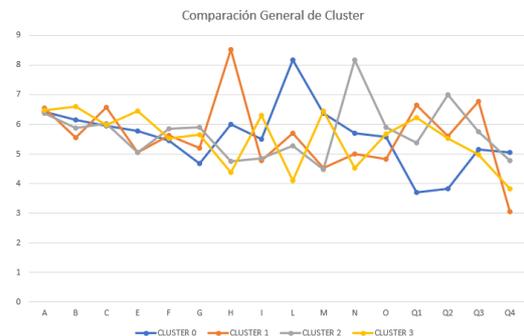


Ilustración 15 Comparación General de cluster.

VI. CONCLUSIONES

Mediante los resultados obtenidos en la investigación realizada y la ejecución de dicho proyecto, se puede observar que en los aspirantes a la carrera de psicología del Centro Universitario Vasco de Quiroga existe un alto

índice en la personalidad de estabilidad teniendo como porcentaje del 10% siendo este el porcentaje más alto entre todos, de la misma forma se puede observar que las personalidades neutras son: Afabilidad, Atrevimiento, Vigilancia, Privacidad y Aprensión, y por último la personalidad Perfeccionismo que cuenta con un porcentaje del 4% ocupando la última posición el cual es la que menos predomina dentro del alumnado de psicología.

Por otro lado, el entender la personalidad se puede llegar a realizar diversas decisiones en cuestiones académicas o sociales, inclusive poder trabajar mejor las habilidades de cada uno, puesto que, con la personalidad podemos saber las habilidades que posee dicha persona, en conjunto con la minería de datos podremos obtener la información que a simple vista no podemos detectar, así mismo poder trabajar a las personas con diferentes personalidades a futuro en las cuestiones y a antes mencionadas.

En lo que concierne al uso de la Minería de Datos en este artículo se puede llegar a la aplicación de técnicas de predicción con algoritmos de árboles de decisión y ID3, con la finalidad de predecir no solamente el tipo de personalidad, sino inclusive poder detectar las habilidades que poseen los alumnos y su permanencia dentro de la carrera de psicología.

REFERENCIAS

- [1] Riquelme Santos, J. C., Ruiz, R., & Gilbert, K. (2006). Minería de datos: Conceptos y tendencias. *Inteligencia artificial: Revista Iberoamericana de Inteligencia Artificial*, 10(29), 11-18.
- [2] Dur, E., & Costaguta, R. (2007). Minería de datos para descubrir estilos de aprendizaje. *Revista Iberoamericana de Educación*, 42(2), 1-10.
- [3] Galindo, A., & García, H. (2010). Minería de Datos en la Educación. *Universidad Carlos III*, 1-8.
- [4] Ballesteros, H. F. V., Iñiguez, E. G., & Velasco, S. R. M. (2018). Minería de Datos. *RECIMUNDO*, 2(1 (Esp)), 339-349.
- [5] Núñez Cárdenas, F. D. J. (2013). Identificación de estilos de aprendizaje en alumnos universitarios de computación de la huasteca hidalguense mediante técnicas de minería de datos.
- [6] Herrera, M., Nieto, S., Rodríguez, M. J., & Sánchez, M. C. (1999). Factores implicados en el rendimiento académico de los alumnos, *Universidad de Salamanca. Revista de investigación educativa*, 17(2), 413-421.
- [7] Cebrià, J., Segura, J., Corbella, S., Sos, P., Comas, O., García, M., ... & Pérez, J. (2001). Rasgos de personalidad y burnout en médicos de familia. *Atención primaria*, 27(7), 459-468.
- [8] Pérez, E., Cupani, M., & Ayllón, S. (2005). Predictores de rendimiento académico en la escuela media: habilidades, autoeficacia y rasgos de personalidad. *Avaliação Psicológica*, 4(1).
- [9] Sánchez-Gallego, N. J., Gómez-Macías, C., & Zambrano-Cruz, R. (2011). Revisión sistemática del Cuestionario Factorial de Personalidad (16PF). *Pensando Psicología*, 7(12), 1-23.
- [10] García, M. A., Martínez, M. C., Camacho, M. C., & Perales, C. R. (2015). Determinación de perfiles delictivos en el estado de Jalisco utilizando WEKA a través de minería de datos. *TECTZAPIC*, (2).
- [11] Gutiérrez, H. G. L., Ruiz, M. G. L., Hernández, V. H., Hernández, B. H., & Hernández, G. H. (2016). Análisis de un caso práctico aplicando el algoritmo K means mediante weka (Waikato environment for knowledge analysis). *Ciencia Huasteca Boletín Científico de la Escuela Superior de Huejutla*, 4(7).
- [12] Pineda-Torres, F., Molina-Sanches, S. J., & Afanador-Montañez, L. E. Discriminación de texturas en imágenes inalámbricas, utilizando el algoritmo k-means.
- [13] Palomino, R. P., Montero, W. D. R., Ibañez, T. A., Vargas, E. P., & Apaza, G. M. (2011). Model Clustering Neural Network Monitoring Applications No K-Means Algorithm for the Segmentation of the Peruvian University Students Union. *Revista de Investigación Business Intelligence*, 1(1).
- [14] Ruíz, C. D. R. E. (2017). Caracterización del desarrollo de los indicadores de patentes en América Latina. *Revista Publicando*, 4(13 (2)), 1-13.
- [15] Campaña Naranjo, F. X. (2017). Aplicación de técnicas de Data Mining a bases de datos de contenido musical para identificar rasgos de personalidad de los usuarios en el Distrito Metropolitano de Quito (Master's thesis, Universidad de las Fuerzas Armadas ESPE. Maestría en Gestión de Sistemas de Información e Inteligencia de Negocios).
- [16] Sáiz, E. Á., & Pantaleón, M. E. Z. (2012). Estilos de aprendizaje y rendimiento académico en asignaturas universitarias orientadas al aprendizaje de herramientas informáticas impartidas en formato virtual. In *Estilos de aprendizaje: investigaciones y experiencias*: [V Congreso Mundial de Estilos de Aprendizaje], Santander, 27, 28 y 29 de junio de 2012.