

**UNIVERSIDAD NACIONAL DEL ALTIPLANO**  
**ESCUELA DE POSGRADO**  
**DOCTORADO EN ESTADÍSTICA E INFORMÁTICA**



**TESIS**

**PATRONES DE COMPORTAMIENTO EN EL USO DE LAS AULAS  
VIRTUALES DE LA UNIVERSIDAD NACIONAL DEL ALTIPLANO ÁREA DE  
INGENIERÍAS UTILIZANDO TÉCNICAS DE MINERÍA DE DATOS**

**PRESENTADA POR:**

**TERESA PAOLA ALVAREZ ROZAS**

**PARA OPTAR EL GRADO ACADÉMICO DE:**

**DOCTOR EN ESTADÍSTICA E INFORMÁTICA**

**PUNO, PERÚ**

**2019**

UNIVERSIDAD NACIONAL DEL ALTIPLANO

ESCUELA DE POSGRADO



DOCTORADO EN ESTADÍSTICA E INFORMÁTICA

TESIS

PATRONES DE COMPORTAMIENTO EN EL USO DE LAS AULAS  
VIRTUALES DE LA UNIVERSIDAD NACIONAL DEL ALTIPLANO ÁREA  
DE INGENIERÍAS UTILIZANDO TÉCNICAS DE MINERÍA DE DATOS

PRESENTADA POR:

TERESA PAOLA ALVAREZ ROZAS

PARA OPTAR EL GRADO ACADÉMICO DE:

DOCTOR EN ESTADÍSTICA E INFORMÁTICA

APROBADA POR EL JURADO SIGUIENTE:

PRESIDENTE

  
.....  
Dr. JOSÉ EMMANUEL CRUZ DE LA CRUZ

PRIMER MIEMBRO

  
.....  
Dr. MARCO ANTONIO QUISPE BARRA

SEGUNDO MIEMBRO

  
.....  
Dr. VÍCTOR ALFREDO PANIAGUA GALLEGOS

ASESOR DE TESIS

  
.....  
Dr. EDGAR ELOY CARPIO VARGAS

ÁREA: Estadística e Informática.

TEMA: Enseñanza virtual de aprendizaje con Minería de datos.

LÍNEA: Nuevas tecnologías informáticas, educación y sociedad.

Puno, 02 de octubre de 2019

## DEDICATORIA

**Esta tesis está dedicada a Dios por guiar  
mis pasos, y a los seres que amo.**

A mis padres Juan Alvarez y Olimpia Rozas con inmenso cariño y gratitud porque depositaron su entera confianza en cada reto que se me presentaba.

A mis hermanas Janeth, Patricia y Karin que son un ejemplo de superación, tenacidad y coraje de quienes siempre he recibido cariño y apoyo incondicional.

A mis sobrinos: Sivonne, Milena, Jossue, Valentina y Clarents que con su alegría me motivaron a seguir adelante.

A mi amado hijo Joaquin por estar en mi vida, y ser fuente de inspiración e impulso.

A Fredy, mi compañero, amigo y cómplice de toda mi vida, con quien compartimos metas y me da fuerzas día a día.

## AGRADECIMIENTOS

- A la Universidad Nacional del Altiplano por permitir actualizarme de acuerdo a la era del conocimiento y la globalización, basado en un enfoque científico.
- A los Catedráticos del Doctorado en Estadística e Informática por sus enseñanzas impartidas durante mi formación de posgrado
- Mi reconocimiento al jurado calificador de la presente investigación, en las personas del Dr. José Emmanuel Cruz De La Cruz por sus acertadas exigencias en la culminación del trabajo de investigación, al Dr. Marco Quispe Barra por su constante aliento y al Dr. Víctor Paniagua por la confianza y la predisposición para apoyarme.
- Al Doctor y amigo, Edgar Eloy Carpio Vargas, por apoyarme en la realización del presente trabajo de investigación.

**ÍNDICE GENERAL**

	<b>Pág.</b>
DEDICATORIA	i
AGRADECIMIENTOS	ii
ÍNDICE GENERAL	iii
ÍNDICE DE TABLAS	v
ÍNDICE DE FIGURAS	vi
RESUMEN	vii
ABSTRACT	viii
INTRODUCCIÓN	1

**CAPÍTULO I****REVISIÓN DE LITERATURA**

1.1 Marco teórico	2
1.1.1 ¿Qué es Minería de Datos?	2
1.1.2 Minería de Datos vs KDD	3
1.1.3 Minería de datos y otras disciplinas.	5
1.1.4 Análisis Clustering	7
1.1.5 Cluster Jerárquicos: estrategia aglomerativa vs divisiva	10
1.1.6 Análisis Conglomerados.	11
1.1.7 Elección de la Técnica Clúster:	13
1.1.8 Plataforma Virtual	14
1.1.9 WEKA	16
1.1.10 Aulas Virtuales	17
1.1.11 Moodle	17
1.1.12 Matriz de Confusión	19
1.2 Antecedentes	20

**CAPÍTULO II****PLANTEAMIENTO DEL PROBLEMA**

2.1 Identificación del Problema	25
2.1.1 Planteamiento del Problema	25
2.2 Enunciado del problema	26
2.3 Justificación	26

2.4	Objetivos	27
2.4.1	Objetivo General	27
2.4.2	Objetivos Específicos	27
2.5	Hipótesis	27
2.5.1	Hipótesis General	27
<b>CAPÍTULO III</b>		
<b>MATERIALES Y MÉTODOS</b>		
3.1	Lugar de Estudio	28
3.2	Población y Muestra	28
3.2.1	Población	28
3.2.2	Muestra	28
3.3	Método de Investigación	29
3.3.1	Metodología	29
3.3.2	Descripción detallada de métodos por objetivos específicos	30
3.3.2.1	Análisis del uso de la plataforma virtual	30
3.3.2.2	Asociación de las variables de la plataforma virtual	30
3.3.2.3	Identificación de patrones desconocidos	30
<b>CAPÍTULO IV</b>		
<b>RESULTADOS Y DISCUSIÓN</b>		
4.1	Resultados	31
4.1.1	Fase 1: Selección.	31
4.1.2	Fase 2: Pre proceso.	31
4.1.2.1	Población	34
4.1.2.1.1	Muestra	35
4.1.2.2	Herramientas Virtuales de la Plataforma	38
4.1.2.3	Tiempo de uso de la plataforma.	44
4.1.3	Fase 3: Transformación y Reducción de los datos.	47
4.1.4	Fase 4: Minería de Datos (Data Mining)	63
4.1.5	Fase 5: Interpretación y evaluación del conocimiento obtenido.	70
4.2	Discusión	78
CONCLUSIONES		81
RECOMENDACIONES		83
BIBLIOGRAFÍA		84

## ÍNDICE DE TABLAS

	<b>Pág.</b>
1. Actividades en la Plataforma Virtual	38
2. Precisión detallada por clases	61
3. Matriz de Confusion	62
4. Centroides del grupo final	70

## ÍNDICE DE FIGURAS

	Pág.
1. Minería de Datos	4
2. Clases de Minería de Datos	5
3. Estructura básica del Dendograma	9
4. Ejemplo de distribución de elementos	10
5. Cluster aglomerativo	10
6. Cluster divisiva	11
7. Métodos de Análisis Clúster	14
8. Discretizacion de variables	33
9. Actividad por Área Académica	34
10. Actividad por Área Académica	34
11. Muestra	35
12. Actividad por Facultad	35
13. Actividad por Facultad	36
14. Actividad por Escuela Profesional	36
15. Actividad por Escuela Profesional	37
16. Herramientas Virtuales de la Plataforma.	38
17. Actividad por Uso de Módulos	39
18. Actividad por Uso de Módulos	40
19. Actividad por Uso de Módulo Específico	42
20. Actividad por Uso de Módulos Específicos	43
21. Actividad de Tiempo de Uso de la Plataforma	45
22. Actividad Tiempo de Uso de la Plataforma	46
23. Clases para generar un podado o sin podar	47
24. Horas de acceso al Sistema por Escuela	63
25. Horas de acceso al Sistema por Módulo	64
26. Procedimiento para la asociación A PRIORI	64
27. Proceso para el Análisis K-means	67
28. Procesamiento de los Clústeres	68
29. Selección de Atributos	71
30. Relación entre Escuelas Profesionales y los Módulos	73
31. Relación entre las Escuelas Profesionales y la Actividad.	75
32. Relación entre la Escuela Profesional y la Hora	77
	vi

## RESUMEN

La incorporación de las tecnologías de la información y comunicación, ha permitido diseñar los escenarios donde se producen los procesos de enseñanza-aprendizaje, provocando que sea cada vez más corta la barrera entre educación presencial, semipresencial y virtual; bajo esta introducción el objetivo general fue determinar los patrones de comportamiento en el uso de las aulas virtuales de la Universidad Nacional del Altiplano, considerando para ello como objetivos específicos analizar el uso de la plataforma virtual de los estudiantes de ingenierías, establecer asociaciones de las variables e identificar patrones desconocidos en la plataforma virtual. Los datos fueron obtenidos de la base de datos de la Oficina de Tecnología e Información de la Universidad Nacional del Altiplano, que consta en el uso de la plataforma Moodle utilizada por los estudiantes del área de Ingenierías del año 2017, II semestre. La metodología usada fue KDD, los datos fueron procesados con el software WEKA, concluyendo que de las 16 escuelas de ingenierías presentan mayor actividad: Ingeniería de Minas, Ingeniería Mecánica Eléctrica, Ingeniería de Sistemas e Ingeniería Económica, también sobresalen el uso de los siguientes módulos: course (curso), assign (asignar), resource (recurso), quiz (examen) y fórum (foros), y las horas de mayor uso son: 18, 19 y 20 horas. Con el proceso de clasificación J48 asociamos a las escuelas más relevantes con las variables módulos y horas. Para encontrar patrones desconocidos que nos oriente sobre el comportamiento de datos empezamos a trabajar con Minería de Datos y aplicando la técnica de asociación A PRIORI, encontramos que en el módulo VISTA en las escuelas de Ing. de Minas e Ing. Topográfica y Agrimensura, presentan un comportamiento, con una confiabilidad del 97%. Otro patrón desconocido se presenta en las escuelas de Ingeniería de Minas, Agroindustrial, Económica, Electrónica y Arquitectura y Urbanismo, quienes presentan comportamientos similares en términos de horas de uso, tiempo, uso de recursos y actividades con un nivel de confiabilidad del 98%. Para tener una mayor precisión con la asociación de los datos se trabajó con el algoritmo K-means donde afirmamos en el cluster 0 que la escuela de Ingeniería de Minas tiene más actividad con el módulo Course y View a las 18 horas, y en el cluster 1 con un 23% de la data, se determinó que la escuela de Ing. Económica registra mayor actividad en asignación de tareas visitadas a las 22 horas.

**Palabras clave:** Aulas virtuales, big data, minería de datos, patrones y plataforma virtual.

## ABSTRACT

The incorporation of information and communication technologies has allowed us to design the scenarios where teaching-learning processes occur, causing the barrier between face-to-face, blended and virtual education to be increasingly shortened; Under this introduction the general objective was to determine the behavior patterns in the use of virtual classrooms of the National University of the Altiplano, considering for this purpose as specific objectives to analyze the use of the virtual platform of engineering students, establish associations of the variables and identify unknown patterns on the virtual platform. The data were obtained from the database of the Office of Technology and Information of the National University of the Altiplano, which consists in the use of the Moodle platform used by the students of the Engineering area of 2017, II semester. The methodology used was KDD, the data was processed with the WEKA software, concluding that of the 16 engineering schools have more activity: Mining Engineer, Electrical Mechanical Engineer, Systems Engineer and Economic Engineer, also highlight the use of the following modules: course (course), assign (assign), resource, quiz and forum, and the hours of greatest use are: 18, 19 and 20 hours. With the J48 classification process, we associate the most relevant schools with the variables modules and hours. To find unknown patterns that guide us on data behavior, we started working with Data Mining and applying the A PRIORI association technique, we found that in the VISTA module in the schools of Mining Engineering and Topographic Engineering and Surveying, they present a behavior, with a reliability of 97%. Another unknown pattern is presented in the Mining, Agroindustrial, Economic, Electronic, Architecture, and Urbanism Engineering schools, who exhibit similar behaviors in terms of hours of use, time, use of resources and activities with a 98% reliability level. To have greater precision with the association of the data, we worked with the K-means algorithm where we affirm in cluster 0 that the Mining Engineer school has more activity with the Course and View module at 18 hours, and in the cluster 1 with 23% of the data, it was determined that the School of Economic Engineering registers greater activity in assigning tasks visited at 22 hours.

**Keywords:** Big data, data mining, patterns, virtual classrooms and virtual platform.

## INTRODUCCIÓN

Las universidades almacenan gran volumen de datos e información en diferentes medios o bases de datos, los cuales pueden ser de gran importancia, principalmente en la toma de decisiones. (Camaro, et al, 2014)

Es por ello que la presente investigación se orientó en la búsqueda del comportamiento, e identificación de patrones desconocidos, en una plataforma virtual (MOODLE) con la que se trabaja en la UNA Puno, haciendo uso de big data se puede contemplar la forma como hoy crecen los datos en volumen, velocidad y variedad, esto debido al gran avance y uso de las tecnologías de información, y al uso diario que las personas hacen de ellas. Mediante minería de datos se realizará el tratamiento de datos masivos para poder extraer conclusiones relevantes, que ayuden al docente a monitorear al estudiante por medio del análisis de los registros de navegación.

Para una mejor comprensión del presente aporte, se ha dividido en cuatro capítulos:

El Capítulo I, se ha revisado diferentes conceptos y teorías sobre Minería de Datos, Big data, metodologías, plataformas virtuales y software que ayuden a analizar sobre el comportamiento de los datos. También se ha revisado trabajos de investigación, concernientes al estudio de gran cantidad de datos revisando el comportamiento para encontrar patrones que oriente a un mejor desempeño académico.

En el Capítulo II, se trata el planteamiento de problema, la pregunta que guía la presente investigación, justificación, objetivos y la respectiva hipótesis planteada.

En el Capítulo III explicamos el lugar de estudio, así como se describe los materiales y métodos empleados para realizar la investigación, también se define la población y la muestra, y se describe de manera clara el método de investigación, de tal manera que pueda ser replicada por otros investigadores.

En el Capítulo IV, describe ampliamente los resultados obtenidos, sobre todos los procedimientos empleados, además se realiza su respectiva interpretación y también la discusión respecto a otros autores que realizaron investigaciones similares.

Finalmente se muestra las conclusiones en base a los objetivos definidos previamente y las recomendaciones de la presente investigación.

## CAPÍTULO I

### REVISIÓN DE LITERATURA

#### 1.1 Marco teórico

##### 1.1.1 ¿Qué es Minería de Datos?

La realización de base de datos se ha vuelto una acción fundamental para las empresas, ya que les permiten crear estrategias para conseguir nuevos clientes o fidelizar a los habituales. Pero a consecuencia de la generación masiva de datos, nos encontramos frente a un problema, la infoxicación, disponemos de tanta información, que a veces es imposible organizarla con efectividad. Por ello, la clave está en descubrir patrones o algoritmos para sacarle el máximo partido, en el Data Mining o minería de datos.

El Data Mining es un conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos que expliquen el comportamiento de datos (Ribas, 2018).

Se puede definir a la Minería de Datos como el proceso de búsqueda en grandes bases de datos para encontrar información útil que sirva para la toma de decisiones. También se utiliza el término en inglés “data mining”.

Entender la tecnología y el software utilizado para encontrar patrones de comportamiento dentro de la base de datos. La base fundamental de esto es que esos patrones ayuden a la toma de decisiones.

La ventaja fundamental del análisis de datos es la gran cantidad de escenarios a los que se puede aplicar, como es la Predicción, Probabilidad y Análisis de Secuencias.

Así mismo dentro de un proceso de minería de datos podemos encontrar cinco fases:

- **Objetivo y Recolección de datos:** Se debe centrarse en qué tipo de información se quiere obtener. Imaginemos el ejemplo que un supermercado quiere conocer a qué hora del día es donde más asistencia de clientes hay. Este sería el objetivo y la información que quiere obtener el comercio en este caso.
- **Procesamiento y gestión de los datos:** Al conocer los datos que queremos recopilar nos ponemos a trabajar. Esta quizás sea la fase más complicada del proceso. Pues requiere seleccionar la muestra representativa sobre la que se va a realizar el análisis. Una vez escogida la muestra se debe analizar qué tipo de variables o modelo de regresión se va a realizar sobre la muestra.
- **Selección del modelo:** Crear un modelo o Algoritmo que nos arroje el mejor resultado posible, para hacer un análisis exhaustivo de las variables a incluir en el modelo. Esto se convierte en una tarea complicada ya que dependerá del tipo de información a analizar. Los mineros de datos llevan a cabo distintos exámenes del algoritmo como: regresión lineal, árbol de decisión, series temporales, etc.
- **Análisis y revisión de resultados:** Analiza los resultados y comprueba si se arroja una explicación lógica. La cual debe facilitar la toma de decisiones en base a la información suministrada por los resultados.
- **Actualización del modelo:** Es muy importante ya que se va haciendo con el paso del tiempo para que no quede obsoleto. Las variables del modelo podrían pasar a ser no significativas y requiere un control periódico del mismo (Lopez, 2018).

### 1.1.2 Minería de Datos vs KDD

La Minería de Datos es considerada como el proceso, lo más automatizado posible, que va de los datos elementales disponibles a la toma de decisiones. El objetivo principal es crear un proceso automatizado que toma como punto de partida los datos y cuya meta es la ayuda a la toma de decisiones.

El Descubrimiento de Conocimiento en Bases de Datos (KDD) apunta a procesar automáticamente grandes cantidades de datos para encontrar conocimiento útil en ellos, así permitirá al usuario el uso de esta información valiosa para su conveniencia.

El KDD es el Proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles, comprensibles a partir de los datos (Barbosa, 2018).

Minería de Datos versus KDD (Knowledge Discovery in Databases): Usualmente ambos términos son intercambiables. KDD (Knowledge Discovery in Databases): Es el proceso de encontrar información y/o patrones útiles en los datos.

Minería de Datos: es el uso de algoritmos para extraer información y/o patrones como parte del proceso KDD.

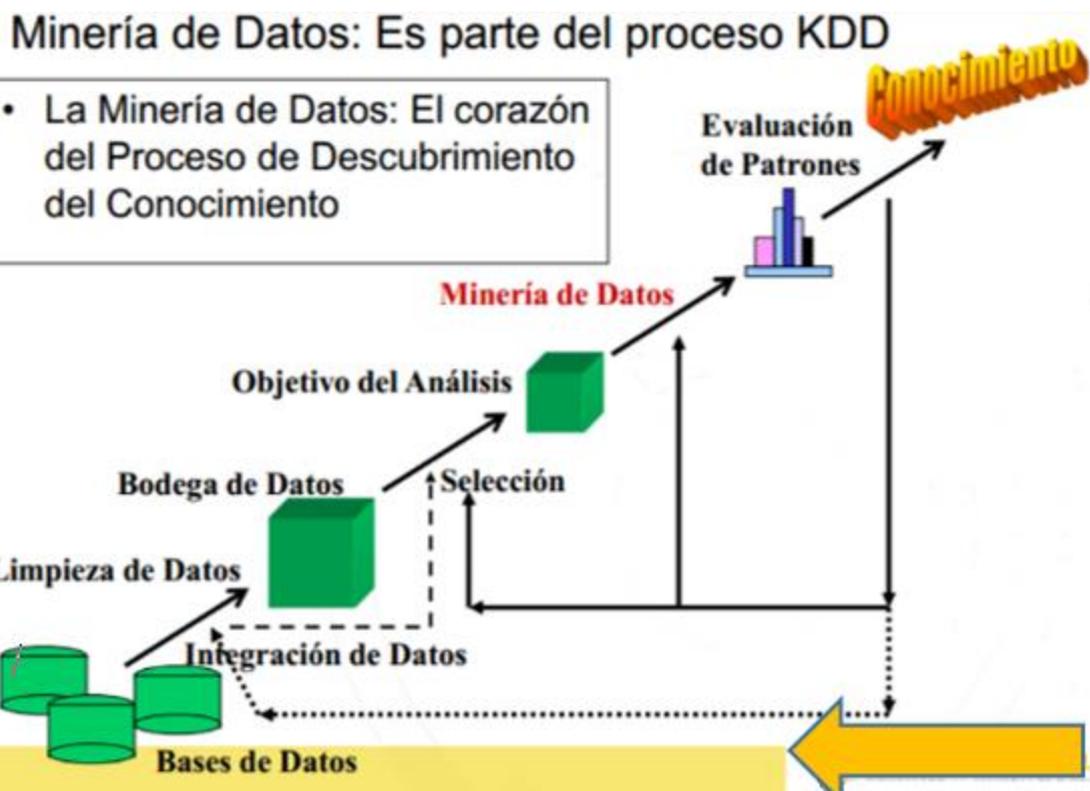


Figura 1. Minería de Datos

Fuente. (González, 2018)

¿Por qué usar Minería de Datos? Muchos datos están siendo generados y almacenados, datos de la Web, comercio electrónico, realizan millones de transacciones, necesitando computadoras más potentes que proporcionen mejores y más servicios personalizados. La tarea de la minería de datos es Descriptivas y Predictivas (Rodríguez, 2016).

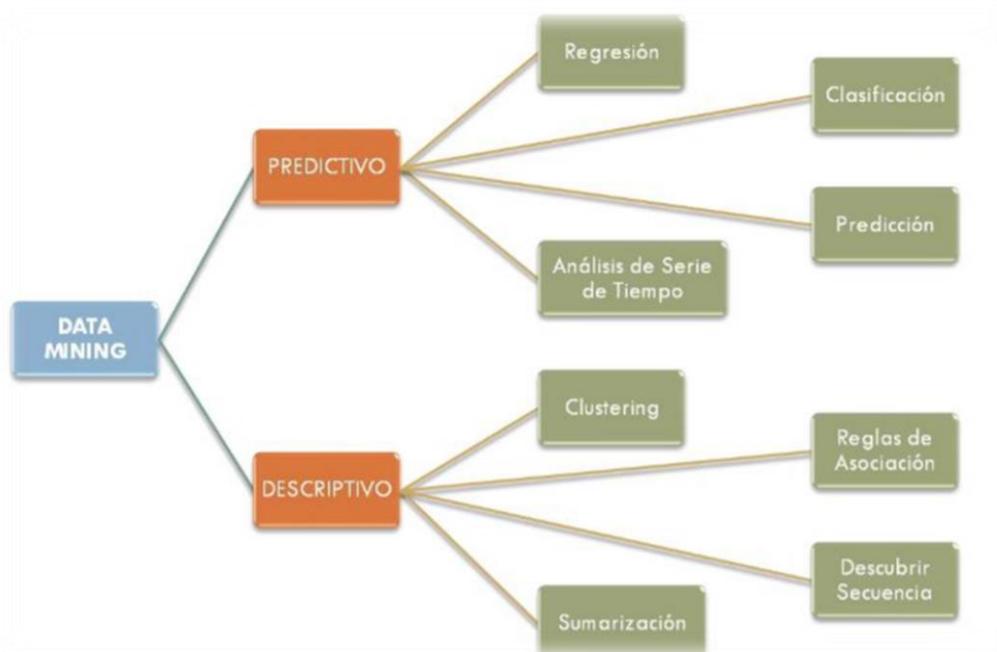


Figura 2. Clases de Minería de Datos  
Fuente: (González 2018)

### 1.1.3 Minería de datos y otras disciplinas.

Existen ciertas fronteras entre la minería de datos y las disciplinas análogas, como pueden serlo la estadística, la inteligencia artificial, etc. Hay quienes sostienen que la minería de datos es una estadística envuelta en una jerga de negocios, convirtiéndola en un producto vendible. Otros, encuentran en ella una serie de problemas y métodos específicos que la hacen distinta de otras disciplinas. Los modelos y algoritmos de uso general en minería de datos, redes neuronales, árboles de regresión y clasificación, modelos logísticos, análisis de componentes principales, etc. gozan de una tradición relativamente larga en otros campos.

#### a. De la estadística.

La minería de datos bebe de la estadística, de la que toma las siguientes técnicas:

- **Análisis de varianza:** donde se evalúa la existencia de diferencias significativas entre las medias de una o más variables continuas en poblaciones distintas.
- **Regresión:** define la relación entre una o más variables y un conjunto de variables predictoras de las primeras.

- **Prueba chi-cuadrado:** por medio de la cual se realiza el contraste de la hipótesis de dependencia entre variables.
- **Análisis de agrupamiento o clustering:** permite la clasificación de una población de individuos caracterizados por múltiples atributos (binarios, cualitativos o cuantitativos) en un número determinado de grupos, con base en las semejanzas o diferencias de los individuos.
- **Análisis discriminante:** clasifica a los individuos en grupos que previamente se han establecido, para encontrar la regla de clasificación de los elementos de estos grupos, e identificar cuáles son las variables que definan la pertenencia al grupo.
- **Series de tiempo:** permite el estudio de la evolución de una variable a través del tiempo para poder realizar predicciones, a partir de ese conocimiento y bajo el supuesto de que no van a producirse cambios estructurales.

**b. De la informática.**

- **Algoritmos genéticos:** Son métodos numéricos de optimización, en los que aquella variable o variables que se pretenden optimizar junto con las variables de estudio constituyen un segmento de información. Aquellas configuraciones de las variables de análisis que obtengan mejores valores para la variable de respuesta, corresponderán a segmentos con mayor capacidad reproductiva. Luego de cierto número de iteraciones, la población estará constituida por buenas soluciones al problema de optimización, pues las malas soluciones han ido descartándose, iteración tras iteración.
- **Inteligencia Artificial:** Mediante un sistema informático que simula un sistema inteligente, se procede al análisis de los datos disponibles. Entre los sistemas de Inteligencia Artificial se encuadrarían los Sistemas Expertos y las Redes Neuronales.
- **Sistemas Expertos:** Son sistemas que han sido creados a partir de reglas prácticas extraídas del conocimiento de expertos. Principalmente a base de inferencias o de causa-efecto.

- **Sistemas Inteligentes:** Son similares a los sistemas expertos, pero con mayor ventaja ante nuevas situaciones desconocidas para el experto.
- **Redes neuronales:** Son métodos de proceso numérico en paralelo, en el que las variables interactúan mediante transformaciones lineales o no lineales, hasta obtener salidas que contrastan con los que tenían que haber salido, dando lugar a un proceso de retroalimentación mediante el cual la red se reconfigura, hasta obtener un modelo adecuado.

Mediante minería de datos, se puede hacer consultas mucho más complejas que utilizando métodos de consulta convencionales. La información que proporciona puede mejorar notablemente la calidad y fiabilidad (Pedraza, 2018).

#### 1.1.4 Análisis Clustering

El Análisis Cluster, también conocido como Análisis de Conglomerados, Taxonomía Numérica o Reconocimiento de Patrones, es una técnica estadística multivariante cuya finalidad es dividir un conjunto de objetos en grupos (cluster en inglés) de forma que los perfiles de los objetos en un mismo grupo sean muy similares entre sí (cohesión interna del grupo) y los de los objetos de clusters diferentes sean distintos (aislamiento externo del grupo).

- Plantear el problema a resolver por un Análisis Cluster
- Establecer medidas de semejanza y de distancia entre los objetos a clasificar en función del tipo de datos analizado
- Analizar algunos de los métodos de clasificación propuestos en la literatura haciendo especial énfasis en los métodos jerárquicos aglomerativos y en el algoritmo de las k-medias, y determinar el número de grupos.
- Interpretar los resultados obtenidos
- Analizar la validez de la clasificación obtenida (Figueras, 2001).

**Clúster:** un número de cosas o personas similares o cercanas, agrupadas.

**Clustering:** es el proceso de particionar un conjunto de objetos (datos) en un conjunto de sub-clases con cierto significado, se conoce también como el proceso

de aprendizaje no supervisado, es un método descriptivo para interpretar un conjunto de datos.

Particionar ejemplos de clases desconocidas en subconjuntos disjuntos de clústers:

- En un mismo clúster sean altamente similares entre si
- En diferentes clústers sean altamente disimiles entre sí.

#### **Aplicaciones:**

- Re conocimiento de patrones
- Procesamiento de imágenes
- Investigación de mercado
- Web mining
- Es también un procesamiento para otras técnicas de data mining.

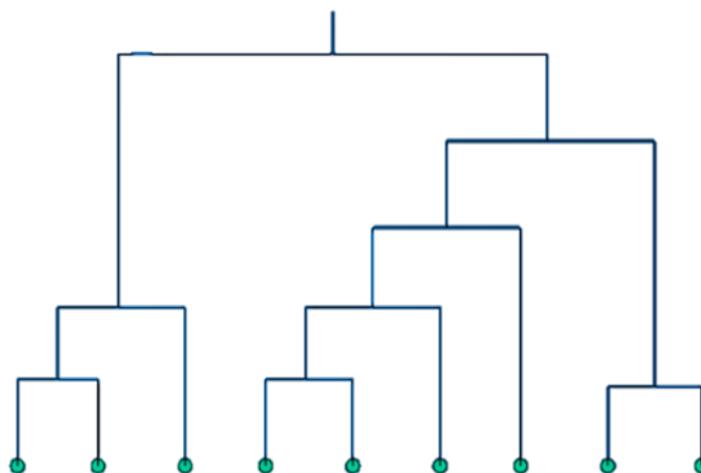
#### **Características:**

- Alta similaridad intra-clúster
- Baja similaridad inter-clúster
- Un buen método descubre algunos o todos los patrones ocultos.
- La calidad clustering depende de la medida de similitud.
- La representación del clúster permite evaluar el método que afecta al resultado.

#### **Principales técnicas de Clustering:**

- Algoritmos de particionamiento
- Algoritmos jerárquicos
- Método basado en densidad (BD espaciales)
- Método basado en grilla
- Método basado en modelo

**Clustering Jerárquico:** Un dendograma muestra cómo se mezclan los clusters de manera que cortando el dendograma en diferentes niveles se consiguen diferentes clusters.



*Figura 3.* Estructura básica del Dendograma  
Fuente: (Isistan 2018)

Construye un árbol binario o dendograma a partir de un conjunto de ejemplos:

- Aglomerativo (bottom-up) métodos que comienzan con cada ejemplo en un cluster diferente y combinan iterativamente para formar clústeres mayores.
- Divisivo (top-down) métodos que comienzan con todos los ejemplos en un mismo clúster y los separan en clusters más chicos.

Determina la similitud de dos clústeres:  $sim(x,y)$ , conteniendo múltiples instancias:

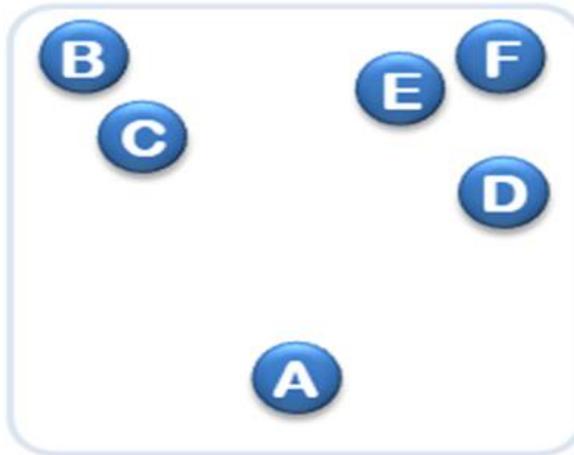
- Single Link: similitud de los dos elementos del clúster más similares
- Complete Link: similitud de los dos elementos del clúster menos similares
- Group Average: promedio de similitudes entre los elementos del clúster

**Ventajas:** No se necesita conocer el número de clústeres  $k$

**Desventaja:** No puede recuperarse de decisiones incorrectas y es computacionalmente costoso (Schiaffino, 2018).

### 1.1.5 Cluster Jerárquicos: estrategia aglomerativa vs divisiva

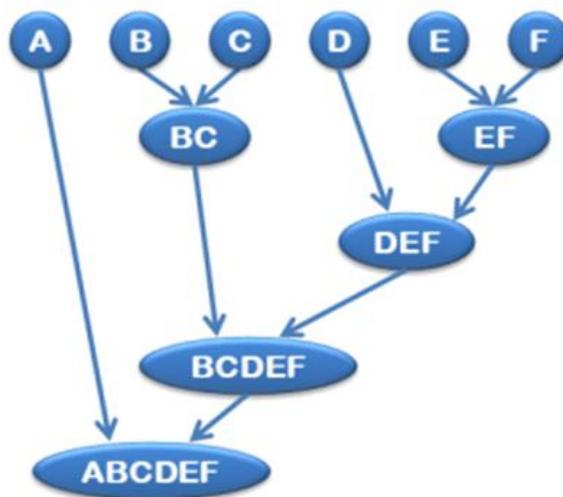
Los métodos jerárquicos tratan de crear grupos de elementos homogéneos entre sí y heterogéneos entre grupos, para conseguirlo principalmente se puede hacer mediante estrategia aglomerativa o divisiva.



*Figura 4.* Ejemplo de distribución de elementos  
Fuente: (Isistan 2018)

#### a. Estrategia Aglomerativa

Las estrategias aglomerativas parten de un conjunto de elementos individuales y van juntando los elementos que más se parezcan hasta quedarse con un número de clusters que se considere óptimo.



*Figura 5.* Cluster aglomerativo  
Fuente: (Isistan 2018)

## b. Estrategia Divisiva

Las estrategias divisivas parten del conjunto de elementos completos y se van separando los grupos que más diferentes sean entre ellos hasta quedarse con un número de clusters que se considere óptimo (Calvo, 2018).

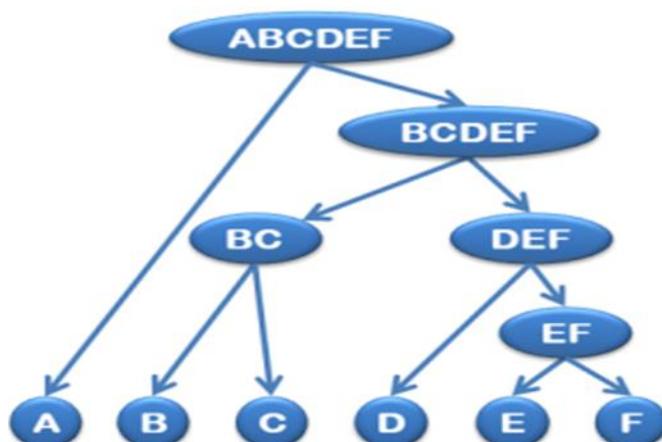


Figura 6. Cluster divisiva  
Fuente: (Isistan 2018)

### 1.1.6 Análisis Conglomerados.

El Análisis Clúster, conocido como Análisis de Conglomerados, es una técnica estadística multivariante que busca agrupar elementos (o variables) tratando de lograr la máxima homogeneidad en cada grupo y la mayor diferencia entre los grupos.

El método estadístico multivariante de clasificación automática de datos, trata de situar los casos (individuos) en grupos homogéneos o clústeres, no conocidos de antemano sean asignados a un mismo clúster, mientras que individuos diferentes (disimilares) se localicen en clústeres distintos.

Sin embargo, en el análisis discriminante es necesario especificar previamente los grupos por un camino objetivo, ajeno a la medida de las variables en los casos de la muestra, mientras que en el análisis clúster define grupos tan distintos como sea posible en función de los propios datos.

La variedad de medir diferentes formas multivariadas proporciona diversas posibilidades de análisis. El empleo de ellas, y el de las que continuamente siguen

apareciendo, así como de los algoritmos de clasificación depende del fenómeno estudiado y del conocimiento previo de posible agrupamiento que de él se tenga.

Puesto que la utilización del análisis clúster ya implica un desconocimiento o conocimiento incompleto de la clasificación de los datos, el investigador ha de ser consciente de la necesidad de emplear varios métodos, ninguno de ellos incuestionable, con el fin de contrastar los resultados.

Existen dos tipos de análisis clústeres: no jerárquicos y jerárquicos.

Los no jerárquicos son aquellos que asignan los casos o grupo diferenciados en los que el propio analista configura, sin que uno dependa del otro, este método puede a su vez producir clústeres disjuntos, o bien clústeres solapados (un caso puede pertenecer a más de un grupo), son de difícil interpretación y poco utilizados.

Los clústeres jerárquicos configuran grupos con estructura arborescente, de forma que clústeres de niveles bajos van siendo englobados en otros clústeres de niveles superiores.

Se dispondrá de una colección de casos agrupada en subconjuntos jerárquicos o no jerárquicos, aplicando técnicas estadísticas comparativas convencionales, así como otras pruebas multivariantes, para las que ya contará con una variable dependiente.

Se puede ampliar el horizonte con la aplicación logística y análisis discriminante con posibles nuevas variables independientes (no sería correcto utilizar las mismas que han servido para la confección de los grupos). También serían aplicables pruebas de asociación y análisis de correspondencias.

El análisis clúster se puede utilizar para agrupar individuos (casos) y también para agrupar variables, pero se debe tener en cuenta tres decisiones:

- Selección de las variables relevantes para identificar a los grupos
- Elección de la medida de proximidad entre los individuos.
- Seleccionar el criterio para agrupar individuos en conglomerados.

Es decisiva la selección de las variables que realmente sean relevantes para identificar a los grupos, de acuerdo con el objetivo que se pretenda lograr en el estudio. De lo contrario, el análisis carecerá de sentido.

Para seleccionar la medida de proximidad es conveniente estar familiarizado con este tipo de medidas, básicamente similitudes y distancias, ya que los conglomerados que se forman lo hacen en base a las proximidades entre variables o individuos. Los grupos que se forman en cada paso dependen de la proximidad, distintas medidas de proximidad pueden dar resultados distintos para los mismos datos.

Para seleccionar el criterio de agrupación conviene conocer como mínimo los principales métodos de análisis clúster (Terrádez, 2017).

### **1.1.7 Elección de la Técnica Clúster:**

#### **a. Método jerárquico**

Siendo su objetivo agrupar clúster para formar un nuevo o separar alguno ya existente para dar origen a otros dos de forma que se maximice una medida de similaridad o se minimice alguna distancia.

#### **b. Método no jerárquico**

Diseñado para la clasificación de individuos en K grupos. El procedimiento es elegir una partición de los individuos en k grupos e intercambiar los miembros de los clústeres para tener una mejor partición (De la Fuentes, 2014).

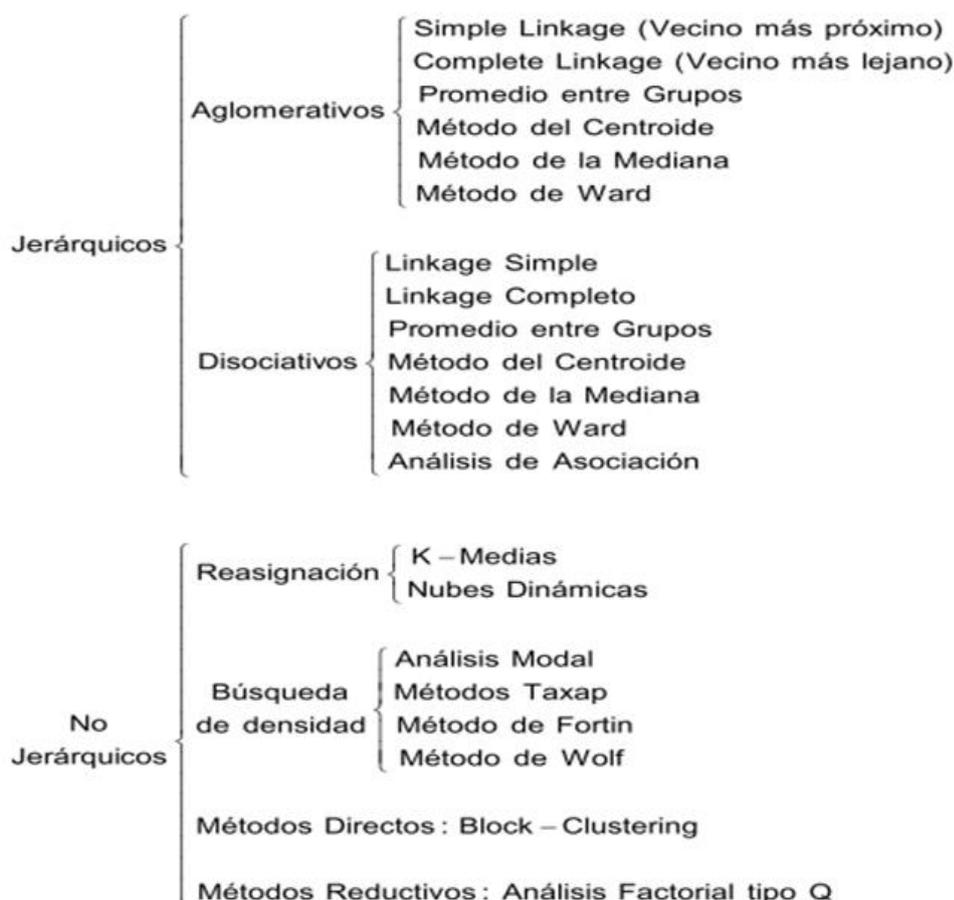


Figura 7. Métodos de Análisis Clúster

Fuente: (De la Fuente 2011)

### 1.1.8 Plataforma Virtual

Plataforma es un concepto con varios usos. Por lo general se trata de una base que se halla a una cierta altura o de aquello que brinda un soporte, ya sea físico o simbólico. El uso más habitual del término virtual, por su parte, está vinculado a lo que existe de manera aparente o simulada, y no físicamente.

Con estas definiciones en claro, podemos introducir el concepto de plataforma virtual, que se emplea en el ámbito de la tecnología. Una plataforma virtual es un sistema que permite la ejecución de diversas aplicaciones bajo un mismo entorno, dando a los usuarios la posibilidad de acceder a ellas a través de Internet.

Esto quiere decir que, al utilizar una plataforma virtual, el usuario no debe estar en un espacio físico determinado, sino que sólo necesita contar con una conexión a la Web que le permita ingresar a la plataforma en cuestión y hacer uso de sus servicios.

Las plataformas virtuales, por lo general, se emplean para la educación a distancia e intentan simular las mismas condiciones de aprendizaje que se registran en un aula. Aunque cada plataforma puede presentar diferentes características, lo habitual es que permitan la interacción de los alumnos entre sí y con los profesores. Para esto, cuentan con diversas vías de comunicación, como chat, foros, etc.

El concepto de plataforma virtual es muy usado en la enseñanza de idiomas a distancia; personas de cualquier parte del mundo pueden estudiar sus lenguas favoritas sin necesidad de realizar grandes inversiones económicas, sin tener que moverse de sus casas y con docentes nativos (una de las mayores ventajas de este sistema educativo).

Los sistemas de permisos sirven para la gestión eficiente de las plataformas virtuales. El administrador, que suele ser el docente o el responsable de la formación, puede acceder a toda la información disponible en la plataforma; los alumnos, en cambio, no.

A grandes rasgos, las plataformas virtuales de enseñanza se organizan de la siguiente manera:

- \* el usuario debe ingresar en el portal de la institución y escoger el plan de estudios que más se ajuste a sus necesidades y posibilidades. Es importante resaltar que existen ciertos requisitos técnicos para que funcione la plataforma; si bien es raro que un alumno no posea la cantidad de memoria y la velocidad del procesador necesarias, sí pueden exigirle que cuente con una cámara o, al menos, con un micrófono;
- \* habiendo escogido el plan deseado, se procede a la reserva de las clases. Esto puede variar en cada caso, pero lo normal es que las lecciones se ofrezcan en paquetes de diferentes cantidades, con tentadores descuentos, y que el pago se deba realizar por adelantado. Algunas instituciones brindan una clase gratis o un paquete inicial a precio reducido para probar sus servicios antes de tomar la decisión final;
- \* la plataforma virtual no sólo ofrece la ventaja de no tener que movilizarse para acceder a ella, sino que también brinda una gran flexibilidad con respecto a los horarios de las lecciones. Dado que los profesores tampoco se ven obligados a

salir de sus casas, es posible concertar turnos fuera del horario comercial, incluso durante la noche, siempre que ambas partes estén de acuerdo;

- \* las clases pueden aprovechar tecnologías tales como las pizarras digitales, para brindar a profesores y alumnos la comodidad de escribir directamente con sus manos, como lo harían en una hoja. Esto es muy útil en el aprendizaje de idiomas con sistemas de escritura diferentes al propio;
- \* a través de la captura de imagen y sonido, se puede crear un ambiente similar al de un aula tradicional, en el cual los alumnos entablen lazos y colaboren para mejorar su experiencia;
- \* luego de cada lección, los profesores suelen enviar a sus alumnos un resumen de los temas vistos, para que los repasen cómodamente (Gardey, 2017).

### 1.1.9 WEKA

WEKA ( Waikato Environment for Knowledge Analysis) es una herramienta de tipo software para el aprendizaje automatizado y minería de datos, diseñado a base de Java y desarrollado en la Universidad de Waikato en Nueva Zelanda en el año 1993, siendo un software libre. Contiene una colección de algoritmos para realizar análisis de datos, modelamiento predictivo y visualización. Posee una interfaz gráfica que unifica las herramientas para que estén a una mejor disposición. En el año 2005 esta herramienta muy flexible y fácil de utilizar, recibe el galardón “*Data Mining and Knowledge Discovery Service*”, por parte de la ACM (Asociación for Computing Machinery) que es la Sociedad Científica Para el Desarrollo de la Computación Educacional. Siendo por esto una herramienta muy versátil que soporta muchas tareas estándares de la minería de datos en especial tareas de procesamiento de datos, regresión, clasificación, clustering entre otras cosas. Las técnicas de WEKA están basadas principalmente en la función de los datos que están disponibles en un fichero plano o una relación, en donde cada registro de datos esta descrito por un número fijo de atributos nominales o numéricos. WEKA permite el acceso a otras instancias de bases de datos por medio de SQL, debido al JDBC, además puede procesar un resultado generado a base de una consulta hecha a una base de datos (Córdoba, 2013).

### 1.1.10 Aulas Virtuales

Es un entorno digital que posibilita el desarrollo de un proceso de aprendizaje. Las tecnologías de la información y la comunicación permiten que el estudiante acceda al material de estudio e interactúe con el profesor y con otros estudiantes.

Un aula no tiene límites físicos: sus limitaciones se vinculan a la disponibilidad de acceso por la vía informática. El alumno, por otra parte, puede ingresar al aula en cualquier momento y desde cualquier lugar para tomar sus clases. Las aulas virtuales pueden presentar: videoconferencias, foros, chats y ejercicio interactivos que son habituales en este tipo de entornos educativos. Si bien hay entidades que ofrecen formación totalmente en línea, existen universidades con aulas virtuales para mejorar la calidad de la propuesta educativa (Perez y Merino, 2017).

### 1.1.11 Moodle

Es un sistema de enseñanza diseñado para crear y gestionar espacios de aprendizaje online para las necesidades de profesores, estudiantes y administradores.

Técnicamente es un sistema web dinámico creado para gestionar entornos de enseñanza virtual, basado en tecnología PHP y bases de datos MySQL.

La primera versión fue creada en el año 2002 por el pedagogo e informático australiano Martin Dougiamas, y su nombre original procede del acrónimo de Module Object-Oriented Dynamic Learning Environment (Entorno Modular de Aprendizaje Dinámico Orientado a Objetos). El carácter gratuito y abierto de Moodle lo convierten en una herramienta muy atractiva, que además cuenta con muchas más ventajas:

#### ✓ Herramienta estable y de confianza

Todo tipo de organizaciones de todos los tamaños confían en ella para desarrollar sus proyectos de formación online.

#### ✓ Intuitiva y fácil de usar

Aprender a gestionarla y utilizarla es muy sencillo. El panel de usuario tiene una interfaz simple, características de arrastrar y soltar, y recursos bien documentados.

### ✓ Siempre actualizada

Moodle es continuamente revisado y mejorado para adaptarse a las necesidades de los usuarios a lo largo del tiempo. En su desarrollo están implicados miles de usuarios de todo el mundo que se organizan en torno a comunidades online.

### ✓ Flexible y personalizable

Al ser un software de código abierto, Moodle puede ser personalizarse y adaptarse a las necesidades individuales gracias a su estructura de funcionamiento modular.

### ✓ Escalable a cualquier tamaño

Es una plataforma que puede dar servicio desde unos cuantos estudiantes a miles de ellos, tanto en organizaciones pequeñas como en grandes.

### ✓ Ubicua y accesible desde cualquier dispositivo

El acceso a Moodle se realiza desde la web, por lo que puede accederse a él desde cualquier lugar del mundo, en cualquier momento y desde cualquier dispositivo. Su interfaz es compatible con móviles y todos los navegadores de internet.

### ✓ Robusta, segura y privada

Los desarrolladores de Moodle están comprometidos con la seguridad de los datos y la privacidad del usuario, por eso los controles de seguridad de la plataforma son actualizados constantemente. Moodle cuenta con sistemas que dan protección frente al acceso no autorizado, la pérdida de datos y el mal uso.

### ✓ Con funcionalidades ampliables

Las posibilidades de Moodle son ilimitadas. Sus funcionalidades pueden extenderse gracias a la instalación de plugins y complementos, fruto de la colaboración de una gran comunidad global.

### ✓ En tu propio idioma

Moodle está traducido a más de 120 idiomas. Su capacidad multilingüe es otra de sus características más apreciadas.

La plataforma Moodle sirve para crear espacios de enseñanza online y administrar, distribuir y controlar todas las actividades de formación no presencial de una entidad educativa u organización.

Por este motivo sus desarrolladores lo consideran un sistema «todo en uno».

Moodle cubre las necesidades de los 3 roles principales implicados en acciones formativas online:

- **Profesores**

Su funcionamiento facilita al máximo las tareas del formador online. Su completo kit de herramientas garantiza el control de todas las actividades del proceso de enseñanza-aprendizaje, desde un único panel de administrador.

- **Alumnos**

Su uso también resulta funcional, simple e intuitivo para los alumnos. Esto les ayuda a centrarse en sus tareas de estudio y no tener que preocuparse por aprender a utilizar una herramienta compleja.

- **Administradores**

Su interfaz gráfica permite crear aulas virtuales y cursos con facilidad, sin aplicar tareas de programación. Es un sistema flexible y totalmente personalizable capaz de adaptarse a los parámetros de cualquier entidad educativa, método de enseñanza, estructura de contenidos, formato de recursos didácticos (texto, imagen, vídeo, presentación, etc.), estética visual, etc. Sus capacidades también pueden ampliarse con la instalación de plugins (Merayo, 2018).

### 1.1.12 Matriz de Confusión

Algoritmo C4.5 (J48) El algoritmo C4.5 construye árboles de decisión de un sistema de datos del entrenamiento de la misma forma que el algoritmo ID3, que usa el concepto de entropía de la información. Los datos del entrenamiento son un sistema  $S = s_1, s_2, \dots$  de muestras ya clasificados. Cada ejemplo  $s_i = \{x_1, x_2, \dots\}$  es un vector donde  $x_1, x_2, \dots$  representan los atributos o características del ejemplo. Los datos de entrenamiento son aumentados con un vector  $C = \{c_1,$

$c_2, \dots$  } donde  $c_1, c_2, \dots$  representan la clase a la que pertenece cada muestra. C4.5 es una extensión del algoritmo ID3 desarrollado anteriormente por Quinlan J. R. en 1993. Los árboles de decisión generados por C4.5 pueden ser usados para clasificación, y por esta razón, C4.5 está casi siempre referido como un clasificador estadístico (Mosquera, Parra, Castrillon, 2016).

## 1.2 Antecedentes

Marulanda, et al. (2017) Los autores resaltan la importancia del uso de la minería de datos para descubrir patrones y tendencias desde la gestión del conocimiento (GC) en las empresas. La investigación presenta la aplicación de técnicas de minería de datos a una base de datos de Gestión de Conocimiento de 322 PYMES de las ciudades de Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales, Pereira y Armenia, para lo cual utilizaron enfoques investigativos como el exploratorio y el descriptivo, conjuntamente con técnicas de minería de datos y el uso de una herramienta de software propietaria. Destacando el avance en competencias, herramientas y TIC para la Gestión de Conocimientos.

Las plataformas e-learning o tutores inteligentes almacenan actividad que los estudiantes realizan, facilitando la aplicación de técnicas de minería de datos para extraer patrones y modelos, permitiendo a los profesores analizar cómo se lleva a cabo el proceso de enseñanza -aprendizaje y proponer acciones dirigidas para su mejora. La contribución de estas tesis, se centra en dos aspectos: por un lado, se ofrecen modelos de predicción del rendimiento de los estudiantes en cursos virtuales atendiendo a distintas métricas y distintos procesos de minería para mejorar su calidad y precisión y, por otra parte, propone un sistema que construye el mejor modelo de clasificación para un conjunto de datos sin interacción de un experto en minería de datos. Para la consecución de sus objetivos, se han seguido líneas de trabajo como la detección y eliminación de anomalías en los conjuntos de entrenamiento para mejorar los modelos de predicción del rendimiento de los estudiantes (García, 2016).

Ruiz y Romero (2018) indican que el trabajo tuvo como objetivo mejorar la calidad de la información que pertenece a la base de datos CubaCiencia, del Instituto de Información Científico y Tecnológico. La base de datos contenía información bibliográfica referida a cuatro segmentos de la ciencia siendo esta la base de datos principal del Sistema de Gestión Bibliotecario. La metodología aplicada estuvo basada en los Árboles de Decisión,

la Matriz de Correlación y el Scatter Plot 3D, que son técnicas utilizadas por la minería de datos, para el estudio de grandes volúmenes de información. Los resultados alcanzados no solo permitieron mejorar la información de la base de datos, sino que también aportaron patrones verdaderamente útiles en la solución de los objetivos propuestos.

Los alumnos ingresantes de las universidades enfrentan problemas de rendimiento, resultando en posible deserción. La minería de datos en educación, aplica técnicas de información generada en el sector educación. El trabajo consistió en realizar la predicción del rendimiento académico de los alumnos que ingresaron a la Escuela Profesional de Ingeniería de Computación y Sistemas de la Universidad de San Martín de Porres en el primer ciclo utilizando minería de datos. Se extrajeron datos de 1304 ingresantes, clasificados en tres factores: sociales, económicos y académicos, realizándose predicciones a través de tres técnicas: regresión lineal, árbol de decisiones y support vector machines, teniendo un mejor resultado con árboles de decisión; se obtuvieron factores que influenciaron en el rendimiento académico y utilizando minería de datos realizaron predicciones del rendimiento académico de los ingresantes, detectando aquellos ingresantes que podrían enfrentar problemas en sus estudios (Yamao, 2018).

Flores (2016) plantea una propuesta para la utilización de aulas virtuales en la plataforma Moodle que permita dinamizar los procesos de enseñanza aprendizaje, de la asignatura de Precálculo I que se imparten a los estudiantes de la Universidad Continental de la ciudad de Huancayo. La metodología se fundamenta en el estudio del proceso enseñanza aprendizaje. Para reforzar el proceso, se ha utilizado la metodología PACIE, y se aprovechó las herramientas que ofrece la plataforma Moodle elaborando un aula virtual, apreciándose una mejora en el rendimiento y propiciando la participación de todos los estudiantes en el proceso de aprendizaje, siendo el 90% de ellos que consideran que el manejo del aula virtual es fácil, permitiendo al docente trabajar sin dificultades. Se recomendó a la institución la ejecución de propuestas planteadas para la utilización de aulas virtuales como apoyo didáctico.

Melo (2018) determinar patrones de predicción mediante el uso de minería de datos del consumo de medicamentos en la Redes Puno, usando arboles de decisión y así obtener la información disponible, aplicando el proceso KDD. Finalmente trabajar con WEKA en algoritmos de minería de datos y obtener arboles de decisión para la extracción de conocimientos. Se hizo un diagnóstico de la situación actual de los Establecimientos de

Salud, seleccionando, procesando y almacenando en un repositorio multidimensional (DataMart), reduciendo el tiempo de cálculo de indicadores y recursos humanos. Con las técnicas de clustering agrupó elementos con datos de clientes con características similares. Finalmente generó un modelo de clasificación asignando elementos no evidenciados, para estimar la necesidad de los medicamentos. Se pudo tomar decisiones informadas y mejorar su capacidad de control de provisión de medicamentos.

Holgado (2018) en su estudio detectó patrones de bajo rendimiento académico de los estudiantes de la Universidad Nacional Amazónica de Madre de Dios (UNAMAD) año 2018, empleándose la metodología CRISP-DM, aplicando el algoritmo Random Forest, e identificando las variables: cantidad de asignaturas cursadas, el servicio de comedor universitario, la carrera profesional, deuda con la universidad, siendo las más influyentes en la predicción del rendimiento académico,. Se emplearon tres algoritmos: Random Forest, C5.0 y CART, obteniendo que el mejor desempeño en el modelo de clasificación de bajo rendimiento académico por los estudiantes de la UNAMAD, fue: “estudiantes que aprobaron más de 6 cursos, pero menos de 62 cursos, que no poseen servicio de comedor universitario y que poseen alguna deuda con la universidad”

Cepeda (2017) indica que la utilización de Aulas Virtuales es un recurso de apoyo para fortalecer los procesos académicos dentro de la enseñanza Universitaria en el Ecuador abordando el caso de la implementación de las estrategias metodológicas. Realizó un estudio cuasi experimental con pos prueba, formando dos grupos, uno de control que trabajo exclusivamente en el salón de clases y otro experimental que trabajo con el apoyo de aulas virtuales en la asignatura de Psicología General.

Trabajó con la prueba T-student para muestras independientes, llegando a la conclusión que el grupo experimental que trabajo con Aulas Virtuales mejoro sus calificaciones y las actividades que fueron asistidas por el docente.

Benalcázar (2017) indica la importancia de la aplicabilidad de la Minería de Datos en todo tipo de empresas ya sea por las exigencias del ritmo de desarrollo o niveles de servicio, brindando a las organizaciones ventajas competitivas como el reconocimiento del comportamiento de patrones y la predicción de tendencias. Concluyó que la metodología que mejor se adapta a la industria de servicios es CRISP-DM (Cross Industry Standard Process for Data Mining) el cual permite llevar de tareas generales a proyectos específicos. Resalta la necesidad de las organizaciones en encontrar información de

acuerdo al comportamiento de sus clientes, acentuándose mas en la Industria de servicios dado por el nivel de demanda del usuario, la metodología de la Minería de Datos plantea actividades en función con los objetivos de la organización y el giro del negocio.

Sulla (2015) manifiesta que la minería de datos se puede aplicar en datos educativos prediciendo la deserción académica, descubre el conocimiento oculto, mediante los métodos de clasificación como los arboles de decisión, las reglas minadas, etc., se puede aplicar para predecir el comportamiento de los estudiantes, el rendimiento en los exámenes, la deserción estudiantil etc. Estas predicciones ayudaron a las autoridades para identificar la deserción estudiantil, con la ayuda del algoritmo del árbol de decisión se pudo conocer el número de estudiantes que son propensos a abandonar la carrera, ayudando a las autoridades a tomar mejores medidas.

Para tener una mejor exactitud se comprobó con el algoritmo Random Tree, encontrando que es eficiente en términos de precisión de los resultados académicos del estudiante.

Laura (2016) Realiza en su trabajo un estudio comparativo de técnicas no supervisadas de minería de datos para la segmentación alumnos en tres agrupaciones (alto, medio, bajo), obteniendo grupos de mejor calidad con menor y mayor distancia (intra e inter cluster), utilizó algoritmos de K-means y PAM dentro del clustering particional y los métodos del clustering jerárquico aglomerativo para realizar la segmentación académica. Una vez elegido el algoritmo de minería de datos donde se obtiene la mejor calidad de agrupamiento se pudo realizar el estudio comparativo para la segmentación académica. Finalmente utilizó el coeficiente de siluetas para medir la calidad de las agrupaciones formadas con las que se obtiene grupos de mejor calidad para la segmentación académica y el reforzamiento de los alumnos.

Hernandez (2016) describe un modelo por capas para la gestión y tratamiento de datos en el dominio ambiental. Aprovecha las ventajas de la tecnología asociadas a Big Data para el tratamiento de estos, identificando estrategias para la captura, almacenamiento y tratamiento de los datos desde un enfoque, que a su vez depende de la naturaleza. El modelo por capas permitió hacer la administración y análisis de datos ambientales.

Los datos suministrados por el Instituto de Estudios Ambientales – IDEA – de la Universidad Nacional de Colombia Sede Manizales, realizó también, análisis relacionado a inteligencia artificial que nos dieron predicciones de comportamientos a futuro de las

variables de precipitación, temperatura y presión barométrica, las cuales quedaron sujetas a validación y agrupamiento de Clustering usando Mahout y Hadoop.

Gutierrez (2017) desarrolló una metodología de extracción de conocimientos, en los datos de prestaciones del Seguro Integral de Salud Región Piura, con cuestionario a los trabajadores sobre el uso de las TIC. Con técnicas de minería de datos, estableció relaciones que ayudó en la toma de decisiones respecto a las prestaciones brindadas en las Instituciones Prestadoras de Servicios de Salud, previa depuración de la información, identificó patrones, obteniendo los siguientes resultados: el 75% trabajadores estuvo de acuerdo con la implementación de la propuesta de mejora. Un 63% opino que la metodología optimizaría el tiempo de trabajo, el 88% de trabajadores les sería útil el uso de la herramienta, un 50% opino que el conocimiento obtenido ayudara a optimizar procesos y el 75% indicó que el conocimiento obtenido ayudará a las autoridades.

Muñoz (2015) indica que el objetivo central en la mejora del RA (rendimiento académico) de los estudiantes de la E.S.O. (Enseñanza Secundaria Obligatoria) utilizando como herramienta el análisis de la Minería de Datos, siguiendo pautas del enfoque Onto-Semiotico, plantea que el rendimiento matemático – lingüístico afecta de forma colateral a su aprendizaje en otras asignaturas de bachillerato o de la universidad.

Mediante técnicas de Minería de Datos estableció modelos e indicadores que mejoraron el RA de los alumnos. El principal aporte es la introducción del aprendizaje computacional para obtener modelos con técnicas de minería de datos, los cuales nos muestran el comportamiento de los estudiantes, permitiendo al docente redirigir las aptitudes de los alumnos tratando de evitar un posible fracaso.

Samaniego et al (2014) describen un procedimiento creado para la identificación de patrones de uso docente en el Sistema de Gestión de Aprendizaje Moodle, identificando patrones tales como: lugar, horario habitual de trabajo, tipología de recursos utilizados, niveles de interacción. Mediante Minería de Datos identifica patrones en función a las siguientes variables: recursos utilizados, acciones realizadas, lugar, horario y mes de trabajo, así como niveles de uso e iteración. Se demostró que el profesor realiza la mayoría de las actividades en el recurso fórum. Así mismo el lugar habitual de trabajo para realizar estas actividades es la ESPOCH generalmente en las mañanas. El estudio ayudo al docente en identificar patrones que generaron conocimiento en torno al uso de la plataforma virtual Moodle.

## CAPÍTULO II

### PLANTEAMIENTO DEL PROBLEMA

#### 2.1 Identificación del Problema

##### 2.1.1 Planteamiento del Problema

La actual sociedad de la información nos plantea un escenario social en el que las Tecnologías de la Información y la Comunicación (TIC) tienen un protagonismo fuertemente marcado en todos los ámbitos incluido el de la educación. El amplio desarrollo del mundo tecnológico y de las comunicaciones hace posible la aparición de nuevas fórmulas educativas entre las que se encuentra el e-learning, b-learning como formación basada en TIC y esto se logra mediante el uso de plataformas virtuales, es el caso de Moodle, como plataforma virtual oficial de nuestra universidad (Sinche,2016).

Hoy en día existen nuevos retos en la educación universitaria, por lo tanto, es muy importante la aplicación de las nuevas tecnologías en el aula. El uso de estas tecnologías en el aula (plataforma virtual), permitiría mayor comunicación y pues una mejor retroalimentación entre el docente y el alumno, problema que se ve en las clases convencionales. Algunos modelos de los sitios webs educativos resulta ser otro gran problema, pues se da que algunos de ellos son poco o nada utilizados en la medida de los usuarios que trabajan en ella. Es así que nuestra investigación desea analizar la tendencia, comportamiento e identificación de patrones desconocidos, los cuales fortalecerán las estrategias de nuestros docentes para un mejor desarrollo y proceso enseñanza-aprendizaje virtual (Gómez, Contreras , y Gutiérrez, 2016).

## 2.2 Enunciado del problema

Entonces, nuestro proyecto genera el siguiente cuestionamiento:

¿Cómo son los patrones de comportamiento en el uso de las aulas virtuales de la Universidad Nacional del Altiplano área de Ingenierías, utilizando técnicas de Minería de Datos?

## 2.3 Justificación

La incorporación de las nuevas tecnologías en todos los aspectos de la vida y la sociedad está demandando nuevos modelos de enseñanza y aprendizaje. El protagonista es el docente profesor, dejando paso a un sistema basado en el aprendizaje, donde el alumno es responsable de su propio aprendizaje y el profesor debe buscar y analizar la metodología y los medios más adecuados que ayuden al alumno en este proceso. En este entorno, “es importante mencionar que al aplicar las plataformas Tics en el desarrollo de las asignaturas, permitirá operar y transformar la información desde fuentes y formatos diversos, para analizar, sintetizar, crear y contrastar hipótesis de investigación” (De la Rosa, 2011).

El estudio tiene relevancia porque analiza la navegación que realizan para mejorar el aprendizaje de los estudiantes universitarios, que puede ser empleado para la codificación de comunicaciones virtuales a través de chats, foros, correos electrónicos etc., con el fin de obtener resultados confiables que faciliten la comprensión de la tendencia, el comportamiento y los patrones desconocidos usados en los procesos de enseñanza–aprendizaje en entornos virtuales.

La plataforma Moodle (Modular Object Oriented Dynamic Learning Environment) es un ambiente educativo virtual, también es un sistema de gestión de cursos (CMS, Course Management Systems) de distribución libre, que ayuda a los educadores a crear comunidades de aprendizaje en línea. Este tipo de plataformas tecnológicas también se conoce como LMS (Learning Management System) (Gonzáles, 2019).

El número de universidades peruanas que utilizan una plataforma como Moodle se ha incrementado. Es por esto que planteamos nuestra investigación para conocer el impacto de uso de esta plataforma en el rendimiento académico, y que sirve como fuente de

conocimiento para que los docentes utilicen los módulos que son relevantes y promuevan en los alumnos el uso de esta tecnología como apoyo de los cursos presenciales.

## 2.4 Objetivos

### 2.4.1 Objetivo General

Determinar los patrones de comportamiento en el uso de las aulas virtuales de la Universidad Nacional del Altiplano área de Ingenierías, utilizando técnicas de Minería de Datos.

### 2.4.2 Objetivos Específicos

- Analizar el uso de la plataforma virtual de los estudiantes de ingenierías de la Universidad Nacional del Altiplano.
- Establecer la asociación de las variables en el uso de la plataforma virtual de los estudiantes de ingenierías de la Universidad Nacional del Altiplano.
- Identificar patrones desconocidos en la plataforma virtual de los estudiantes de ingenierías de la Universidad Nacional del Altiplano.

## 2.5 Hipótesis

### 2.5.1 Hipótesis General

La Minería de Datos permite encontrar patrones desconocidos en el uso de las aulas virtuales de los estudiantes de Ingenierías de la Universidad Nacional del Altiplano.

## CAPÍTULO III

### MATERIALES Y MÉTODOS

#### 3.1 Lugar de Estudio

La investigación se aplica en estudiantes de Ingenierías II semestre año 2017, de la Universidad Nacional del Altiplano ubicado en la ciudad de Puno, perteneciente a la Región Puno - Perú, geográficamente, según las coordenadas, su ubicación es de 15°50'31" de Latitud Sur y 70°01'11" de Longitud Oeste del Meridiano de Greenwich con una altitud sobre el nivel del mar de 3825 msnm.

#### 3.2 Población y Muestra

##### 3.2.1 Población

La población en la investigación son los estudiantes de las escuelas de ingenierías de la Universidad Nacional del Altiplano Puno para lo cual se obtuvieron la base de datos, expresados en 343 217 de registros aproximadamente correspondiente al semestre académico 2017 II.

##### 3.2.2 Muestra

Por las características de la herramienta a usar, se aplicará el muestro no probabilístico, por conveniencia a criterio del investigador, basado en la elección del año y semestre académico, expresado en 191,990 registros, siendo el elegido el segundo semestre 2017.

En este tipo de muestra no existe una oportunidad real de que un elemento en particular de la población sea seleccionado, por lo tanto, no es posible calcular el

error de muestreo, ni la confiabilidad de las inferencias. Se utilizará en su totalidad los registros consignados.

### 3.3 Método de Investigación

La investigación es de carácter cuantitativa, el diseño es no experimental y el tipo de investigación es correlacional debido a que se procesa y evalúa los datos de las Aulas Virtuales de la Universidad Nacional del Altiplano Puno.

#### 3.3.1 Metodología

Para la implementación se trabajó con el proceso KDD, que persigue la preparación o pre procesamiento de los datos para la automatización del conocimiento, debido a que puedan ser inconsistentes o incompletos, o pueda resultar menores que el original y la preparación da lugar a datos de calidad (Herrera y Cano, 2006). La metodología se desarrolló de la siguiente forma:

1. Selección: etapa en la que se elige un conjunto de datos objetivo sobre los que se realizara el análisis.
2. Pre proceso: en este proceso aseguramos la calidad de los datos a analizar ya que de ello dependerá la calidad del conocimiento descubierto, se incluye tareas como el filtrado y la eliminación de ruidos o la normalización de datos.
3. Transformación y Reducción de los datos: Modificamos la estructura de los datos con objetivo de facilitar el análisis de los mismos, esto incluyo la discretizacion de los datos y encontrar espacios de menor dimensión para un mejor conocimiento y por ende una mejor proyección.
4. Minería de datos (Data Mining): Aplicamos las herramientas de Minería de Datos, para lo cual trabajamos con el algoritmo de Clustering y así encontrar la relación entre las variables y extraer información oculta.
5. Interpretación y/o evaluación del conocimiento extraído: Para esta esta evaluamos e interpretamos el conocimiento extraído teniendo como base tres criterios fundamentales: precisión, claridad e interés (Martínez, 2012).

### **3.3.2 Descripción detallada de métodos por objetivos específicos**

#### **3.3.2.1 Análisis del uso de la plataforma virtual**

En este objetivo delimitamos la base de datos con la cual se trabajó, que consistió en la depuración y validación de datos, tratamiento que se realizó con el programa WEKA donde se comparó las escuelas, los módulos de las aulas virtuales y el tiempo en los que fueron ingresados.

#### **3.3.2.2 Asociación de las variables de la plataforma virtual**

En la asociación de las variables de la plataforma virtual se buscó la clasificación y su relación entre estas, aplicándose el método J48, y para la prueba de validación se trabajo con la Matriz de Confusión.

#### **3.3.2.3 Identificación de patrones desconocidos**

Para identificar patrones desconocidos en nuestra investigación aplicamos el análisis K- means del proceso de Minería de Datos, seleccionando las variables, para luego identificar estos patrones en un análisis visual gráfico, El proceso se desarrolló en el programa WEKA.

## CAPÍTULO IV

### RESULTADOS Y DISCUSIÓN

#### 4.1 Resultados

Damos a conocer los resultados obtenidos del estudio de investigación planteado, mediante el proceso KDD.

El proceso KDD persigue la extracción automatizada de conocimientos no triviales, previamente desconocido y potencialmente útil a partir de grandes volúmenes de datos (Timarán et al., 2016).

El software utilizado para el procesamiento de datos es WEKA.

##### 4.1.1 Fase 1: Selección.

Se trabajó con la base de datos de Aulas Virtuales de la Universidad Nacional del Altiplano, identificando patrones que nos hagan comprender el uso de los recursos más comunes y las necesidades de los estudiantes al trabajar en este sistema virtual.

En la selección de datos se optó por tomar solo las escuelas de Ingenierías de la Universidad Nacional del Altiplano debido que presenta la mayor tendencia de alumnos en ingresar y utilizar el Aula Virtual. Así mismo dada la gran cantidad de datos se seleccionó el semestre 2017 II, debiendo decodificarlo para su entendimiento.

##### 4.1.2 Fase 2: Pre proceso.

Se determinó la confiabilidad de la información, es decir, realizar tareas que garanticen la utilidad de los datos. Para esto se realizó el proceso de depuración y

validación de datos; previamente debiendo ser eliminados los campos de área y facultad, para así obtener resultados mediante los procesos de minería de datos, solo a partir del comportamiento de las escuelas profesionales.



Figura 8. Discretizacion de variables

En la Figura 8 se verificó, cada uno de las variables y sus atributos, con los resultados descriptivos iniciales comprobamos que la data está limpia y no tiene error alguno (data outlier).

#### 4.1.2.1 Población

La población con la que trabajamos son las escuelas de ingenierías.

**Current relation**

Relation: toolscourses      Attributes: 8  
 Instances: 343217      Sum of weights: 343217

Name: area      Type: Nominal  
 Missing: 0 (0%)      Distinct: 5      Unique: 0 (0%)

No.	Label	Count	Weight
1	INGENIERIAS	154166	154166.0
2	BIOMEDICAS	16741	16741.0
3	SOCIALES	61549	61549.0
4	EMPRESARIALES	96474	96474.0
5	OTRAS	14287	14287.0

Figura 9. Actividad por Área Académica

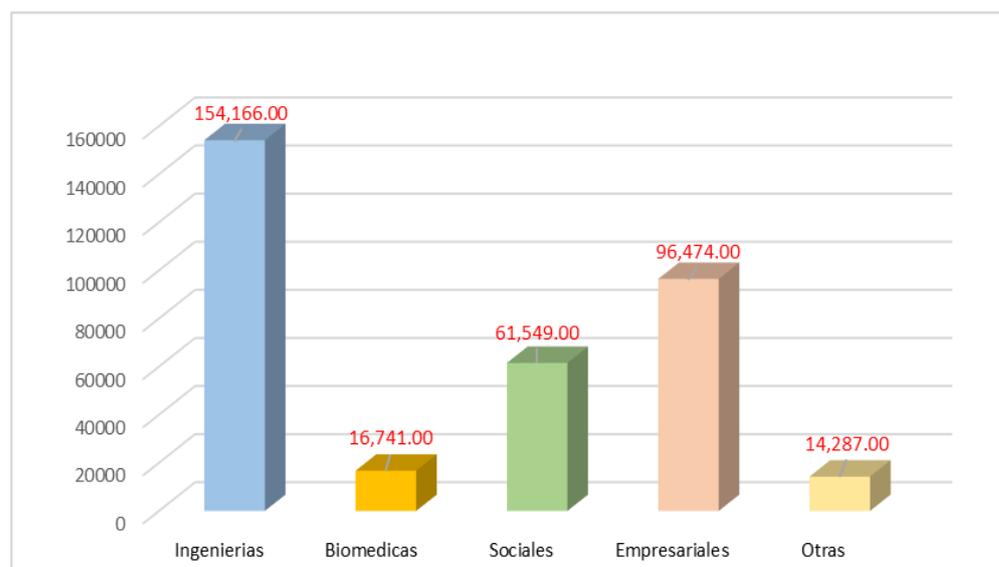


Figura 10. Actividad por Área Académica

Se aprecia que, de un total de 343,217 actividades, el área académica que más resalta es Ingenierías, como se muestra en las figuras 9 y 10, el cual nos interesa para el estudio realizado.

**4.1.2.1.1 Muestra**

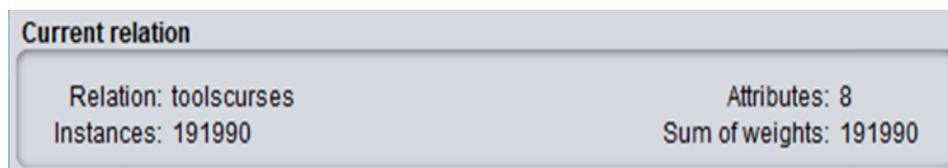


Figura 11. Muestra

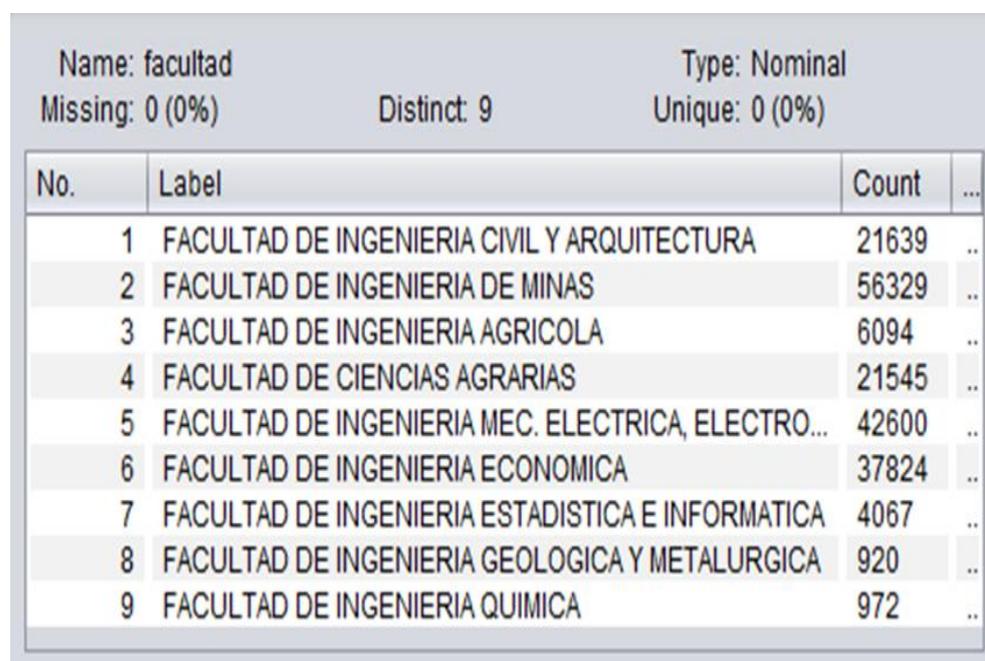


Figura 12. Actividad por Facultad

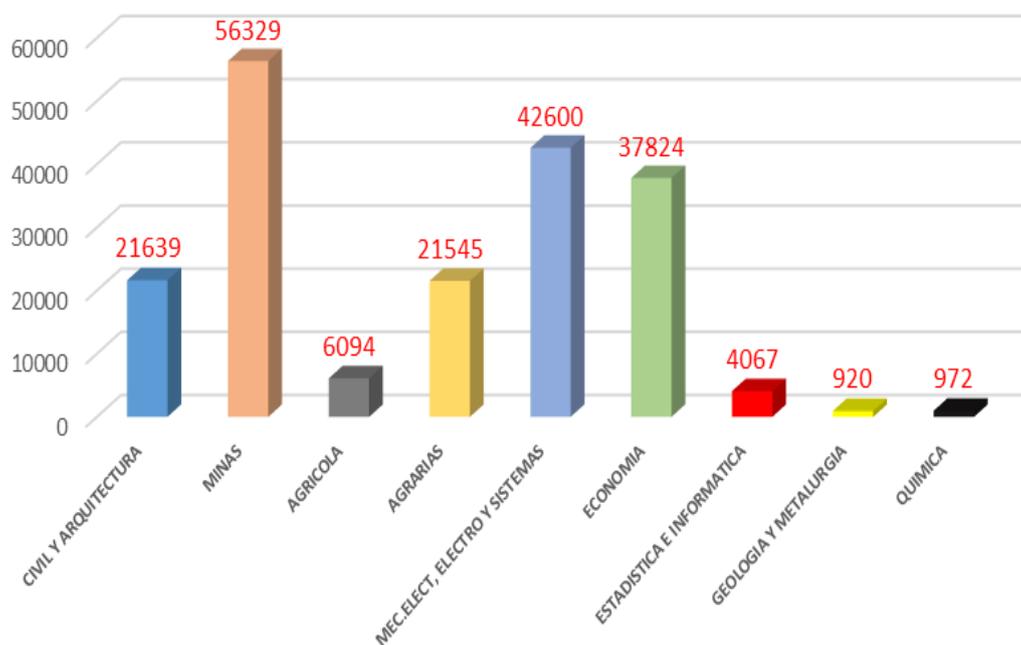


Figura 13. Actividad por Facultad

El tamaño de la muestra es de 191 990 registros, pertenecientes todas al área de ingenierías, resaltando con mayor actividad las facultades de Ing. de Minas, Ing. Mecánica Eléctrica-Electrónica-Sistemas e Ing. Economía, como lo observamos en las figuras 11, 12 y 13

Name: escuela		Type: Nominal	
Missing: 0 (0%)		Distinct: 16	
		Unique: 0 (0%)	
No.	Label	Count	Weight
1	ARQUITECTURA Y URBANISMO	11962	11962.0
2	CIENCIAS FISICO MATEMATICAS	80	80.0
3	INGENIERIA DE MINAS	56329	56329.0
4	INGENIERIA AGRICOLA	6094	6094.0
5	INGENIERIA AGROINDUSTRIAL	10100	10100.0
6	INGENIERIA AGRONOMICA	4127	4127.0
7	INGENIERIA CIVIL	9597	9597.0
8	INGENIERIA DE SISTEMAS	11693	11693.0
9	INGENIERIA ECONOMICA	37824	37824.0
10	INGENIERIA ELECTRONICA	30887	30887.0
11	INGENIERIA ESTADISTICA E INFORM...	4067	4067.0
12	INGENIERIA GEOLOGICA	431	431.0
13	INGENIERIA MECANICA ELECTRICA	20	20.0
14	INGENIERIA METALURGICA	489	489.0
15	INGENIERIA QUIMICA	972	972.0
16	INGENIERIA TOPOGRAFICA Y AGRIM...	7318	7318.0

Figura 14. Actividad por Escuela Profesional

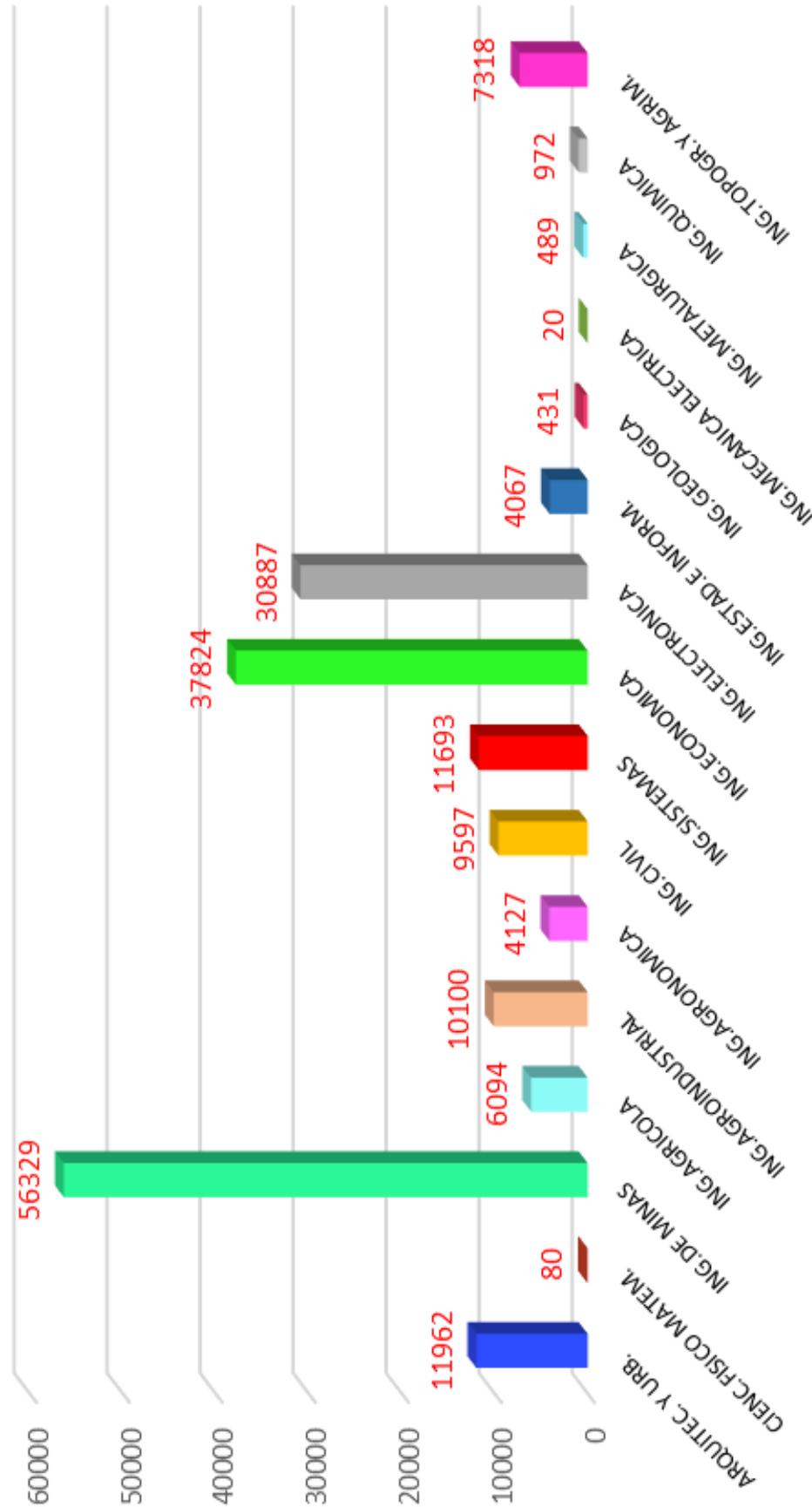


Figura 15. Actividad por Escuela Profesional

Las actividades son disgregadas por Escuelas Profesionales; reafirmamos que la escuela de Ingeniería de Minas sobresale de las demás, seguidamente de Ingeniería Económica y aparece Ingeniería Electrónica, entre las que realiza la mayor cantidad de actividades en las aulas virtuales, como son mostrada en las figuras 14 y 15.

### 4.1.2.2 Herramientas Virtuales de la Plataforma

El siguiente grafico presenta las herramientas virtuales de la plataforma, tanto en inglés y castellano, resaltando las que tienen mayor utilidad.



Figura 16. Herramientas Virtuales de la Plataforma.

Tabla 1  
*Actividades en la Plataforma Virtual*

Ingles	Español			Ingles	Español	
Course	Curso			url	URL	URL
User	Usuario			Calendar	Calendario	
Blog	Blog			Label	Etiqueta	(R) Etiqueta
Fórum	Foro	(A)	Foro	Role	Papel	
Assign	Asignar	(A)	Tarea	Chat	Charla	(A) Chat
Quiz	Examen	(A)	Cuestionario	Discussion	Discusión	
Resource	Recurso	(R)	Archivo	Glosary	Glosario	
Folder	Carpeta	(R)	Carpeta	Notes	Notas	
Page	Pagina			Grade	Grado	

No.	Label	Count	Weight
1	course	81940	81940.0
2	user	7474	7474.0
3	blog	992	992.0
4	forum	14510	14510.0
5	assign	42964	42964.0
6	quiz	19057	19057.0
7	resource	20985	20985.0
8	folder	827	827.0
9	page	324	324.0
10	url	1781	1781.0
11	calendar	638	638.0
12	label	137	137.0
13	role	104	104.0
14	chat	71	71.0
15	discussion	28	28.0
16	glossary	0	0.0
17	notes	30	30.0
18	grade	128	128.0

Figura 17. Actividad por Uso de Módulos

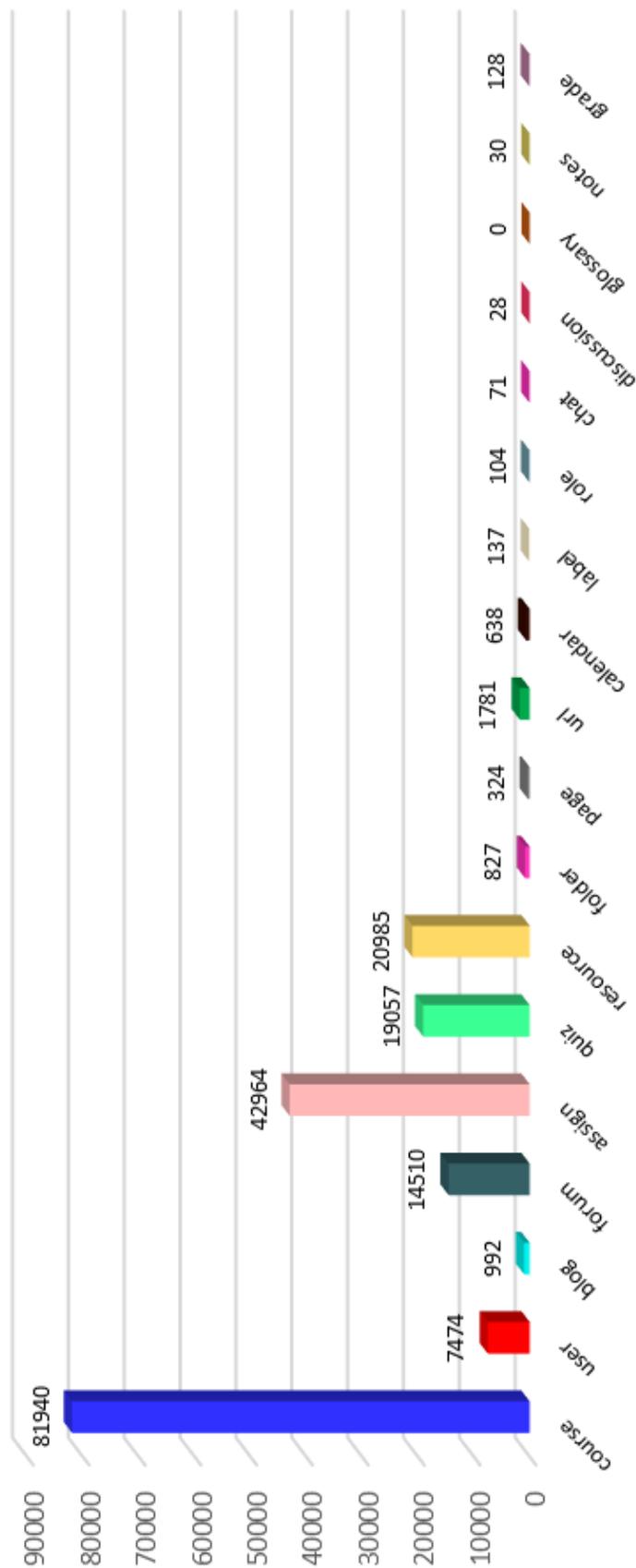


Figura 18. Actividad por Uso de Módulos

Se muestra las actividades que se presentan en la plataforma virtual tanto en el idioma inglés como en el español, tal como se presenta en la tabla 1, las figuras 17 y 18 resaltan aquellas actividades que presentan mayor actividad entre los que se encuentran: el módulo course (curso), seguida de assign (tarea), resource (archivo), quiz (examen) y fórum (foro).

No.	Label	Count	Weight
1	add discussion	1616	1616.0
2	add entry	245	245.0
3	add mod	0	0.0
4	add post	1243	1243.0
5	add category	76	76.0
6	assign	1	1.0
7	attempt	103	103.0
8	close attempt	665	665.0
9	continue attempt	624	624.0
10	delete discussion	9902	9902.0
11	delete mod	1	1.0
12	delete post	189	189.0
13	download all submissions	5	5.0
14	edit	48	48.0
15	editquestions	289	289.0
16	editsection	309	309.0
17	enrol	0	0.0
18	enrol	0	0.0
19	grade submission	1153	1153.0
20	grant extension	0	0.0
21	lock submission	1	1.0
22	mail error	2	2.0
23	manualgrade	117	117.0
24	mark read	41	41.0
25	preview	10	10.0
26	recent	1157	1157.0
27	report	129	129.0
28	report live	0	0.0
29	report log	12	12.0
30	report outline	2	2.0
31	report participation	1	1.0
32	review	2260	2260.0
33	search	112	112.0
34	submission statement accepted	24	24.0
35	submit	3470	3470.0
36	submit for grading	0	0.0
37	subscribe	6	6.0
38	talk	3	3.0
39	unassign	1	1.0
40	unenrol	0	0.0
41	unsubscribe	0	0.0
42	update	737	737.0
43	update mod	1122	1122.0
44	update post	44	44.0
45	user report	57	57.0
46	view	137646	137646.0
47	view all	5023	5023.0
48	view confirm submit assignment form	0	0.0
49	view discussion	5518	5518.0
50	view forum	8387	8387.0
51	view forums	10	10.0
52	view grading form	1238	1238.0
53	view section	0	0.0
54	view submission	3	3.0
55	view submission grading table	1272	1272.0
56	view submit assignment form	6041	6041.0
57	view summary	885	885.0

Figura 19. Actividad por Uso de Módulo Específico



Figura 20. Actividad por Uso de Módulos Específicos

En las figuras 19 y 20 al desgargar los módulos podemos observar que el más frecuente es view (vista), que nos indica que la mayoría de los estudiantes entran solo a visualizar, si los docentes dejan algún contenido, seguido de fóruns y una acción de close attemp (intento).

#### **4.1.2.3 Tiempo de uso de la plataforma.**

En relación al tiempo de uso o permanencia de la plataforma, a pesar que el sistema registra la hora de ingreso al sistema (log in), el problema está en el registro de salida del sistema (log out), puesto que se ha verificado que en la mayoría de los usuarios NO usan el link (salir) recomendado por el sistema, solo cierran la ventana sin salir correctamente, ello perjudica el registro de la hora de salida, de allí que no se puede calcular en forma efectiva el tiempo de uso de la plataforma.

No.	Label	Count	Weight
1	00	2012	2012.0
2	01	911	911.0
3	02	399	399.0
4	03	582	582.0
5	04	1296	1296.0
6	05	1848	1848.0
7	06	3273	3273.0
8	07	6151	6151.0
9	08	9275	9275.0
10	09	11191	11191.0
11	10	9487	9487.0
12	11	10967	10967.0
13	12	9023	9023.0
14	13	11923	11923.0
15	14	12902	12902.0
16	15	12375	12375.0
17	16	10302	10302.0
18	17	13125	13125.0
19	18	16200	16200.0
20	19	14454	14454.0
21	20	11151	11151.0
22	21	10004	10004.0
23	22	7519	7519.0
24	23	5620	5620.0

Figura 21. Actividad de Tiempo de Uso de la Plataforma

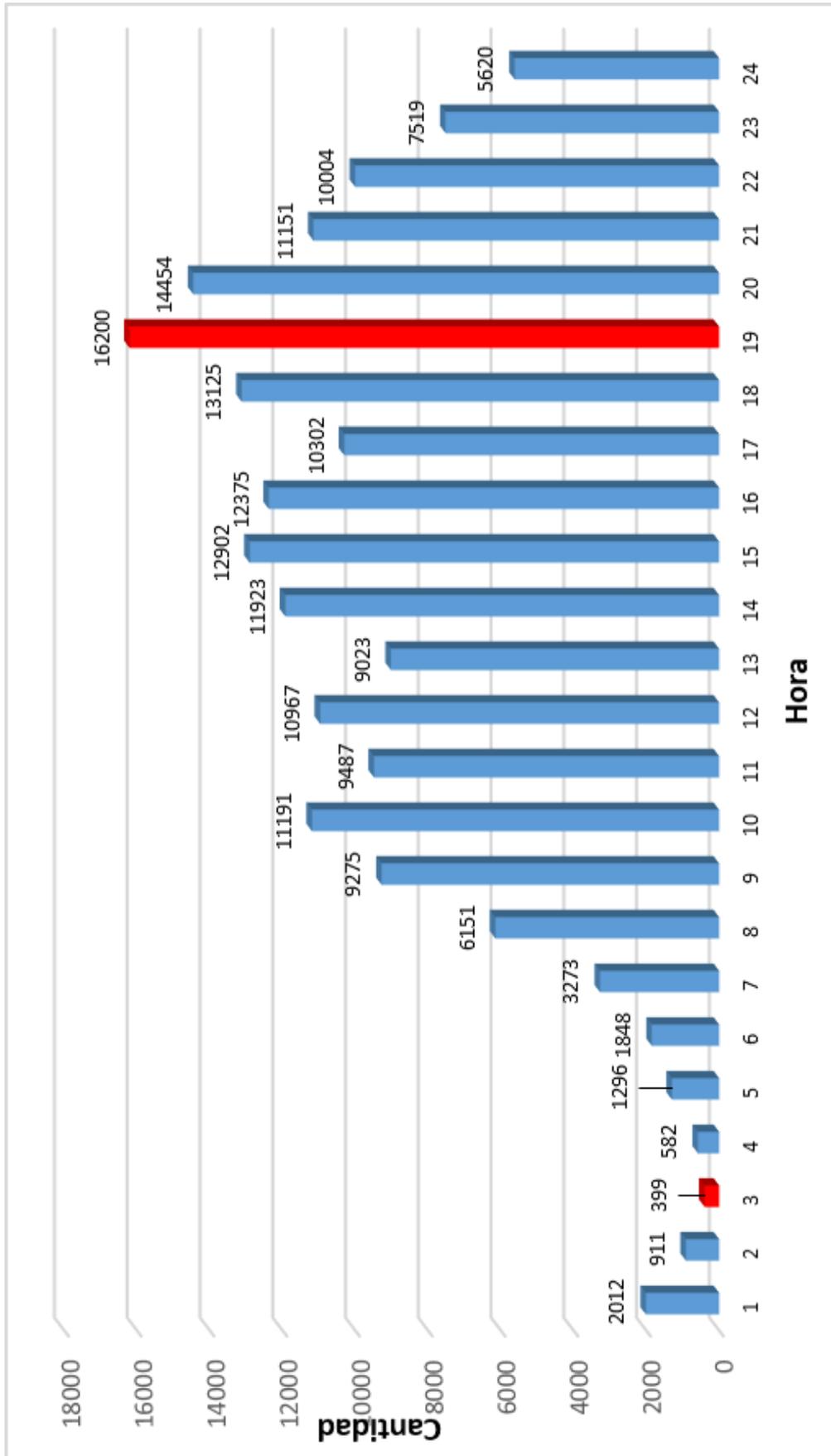


Figura 22. Actividad Tiempo de Uso de la Plataforma

Las figuras 21 y 22 representa las horas en las que con mayor frecuencia ingresan los estudiantes a las aulas virtuales, observamos que con menor frecuencia son 3 de la mañana y con mayor frecuencia son las 19 horas como se observa.

#### 4.1.3 Fase 3: Transformación y Reducción de los datos.

En esta etapa empezaremos con la clasificación y la prueba de validación del proceso, para luego involucrar a las escuelas, los módulos y las horas.

##### a. Proceso de clasificación

Lo que se buscó es la clasificación de las variables y su relación con las otras variables mediante el método de clasificación **J48**, las mismas que mediante un proceso de clasificación discriminó a las escuelas más relevantes en términos de uso y a la vez clasificados por la hora en la que se usó.

```

Class for generating a pruned or unpruned C4
5 decision tree. For more information, see
Ross Quinlan (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA.

CAPABILITIES
Class -- Binary class, Missing class values, Nominal class

Attributes -- Binary attributes, Date attributes, Empty nominal attributes, Missing values, Nominal attributes, Numeric attributes, Unary attributes

Interfaces -- Drawable, PartitionGenerator, Sourcable, WeightedInstancesHandler

Additional
Minimum number of instances: 0

```

Figura 23. Clases para generar un podado o sin podar

```

=== Run information ===
Scheme:   weka.classifiers.trees.J48 -C 0.25 -M 2
Instances: 191990
Attributes: 4
          escuela
          module
          activity
          hour
Test mode: evaluate on training data
=== Classifier model (full training set) ===
J48 pruned tree (arbol podado)
-----

```

**Module = course**

| hour = 00  
 | | activity = add mod: INGENIERIA DE MINAS (4.0/3.0)  
 | | activity = recent: INGENIERIA CIVIL (22.0/12.0)  
 | | activity = view: INGENIERIA DE MINAS (815.0/564.0)  
 | hour = 01  
 | | activity = recent: INGENIERIA CIVIL (15.0/10.0)  
 | | activity = view: INGENIERIA DE MINAS (443.0/321.0)  
 | hour = 02  
 | | activity = view: INGENIERIA DE MINAS (165.0/113.0)  
 | hour = 03  
 | | activity = add mod: ARQUITECTURA Y URBANISMO (2.0/1.0)  
 | | activity = recent: INGENIERIA DE MINAS (5.0/4.0)  
 | | activity = view: INGENIERIA DE MINAS (250.0/175.0)  
 | hour = 04: ARQUITECTURA Y URBANISMO (628.0/305.0)  
 | hour = 05: ARQUITECTURA Y URBANISMO (814.0/535.0)  
 | hour = 06  
 | | activity = add mod: INGENIERIA DE MINAS (37.0/26.0)  
 | | activity = delete mod: INGENIERIA DE MINAS (5.0/2.0)  
 | | activity = editsection: INGENIERIA DE MINAS (6.0/2.0)  
 | | activity = recent: INGENIERIA CIVIL (9.0/6.0)  
 | | activity = update mod: INGENIERIA DE SISTEMAS (42.0/23.0)  
 | | activity = view: INGENIERIA ECONOMICA (1234.0/914.0)  
 | hour = 07  
 | | activity = add mod: ARQUITECTURA Y URBANISMO (85.0/59.0)  
 | | activity = delete mod: INGENIERIA ECONOMICA (69.0/4.0)  
 | | activity = editsection: INGENIERIA DE MINAS (5.0/2.0)  
 | | activity = recent: INGENIERIA AGROINDUSTRIAL (36.0/25.0)  
 | | activity = submission statement accepted: INGENIERIA DE MINAS (0.0)  
 | | activity = update: INGENIERIA ECONOMICA (11.0/7.0)  
 | | activity = update mod: ARQUITECTURA Y URBANISMO (58.0/24.0)  
 | | activity = view: INGENIERIA DE MINAS (2770.0/1848.0)  
 | hour = 08  
 | | activity = add mod: INGENIERIA DE MINAS (72.0/57.0)  
 | | activity = delete mod: ARQUITECTURA Y URBANISMO (2.0/1.0)  
 | | activity = editsection: INGENIERIA CIVIL (2.0)  
 | | activity = recent: INGENIERIA AGROINDUSTRIAL (62.0/52.0)  
 | | activity = report log: INGENIERIA AGROINDUSTRIAL (1.0)  
 | | activity = update: INGENIERIA DE SISTEMAS (4.0/2.0)  
 | | activity = update mod: ARQUITECTURA Y URBANISMO (61.0/46.0)  
 | | activity = view: INGENIERIA ELECTRONICA (3872.0/2750.0)  
 | hour = 09  
 | | activity = add mod: ARQUITECTURA Y URBANISMO (79.0/61.0)  
 | | activity = delete mod: INGENIERIA AGRICOLA (6.0/4.0)  
 | | activity = editsection: ARQUITECTURA Y URBANISMO (11.0/7.0)  
 | | activity = recent: INGENIERIA DE MINAS (50.0/33.0)  
 | | activity = update: INGENIERIA ECONOMICA (3.0)  
 | | activity = update mod: ARQUITECTURA Y URBANISMO (83.0/41.0)  
 | | activity = view: INGENIERIA DE MINAS (4495.0/3360.0)  
 | hour = 10  
 | | activity = add mod: INGENIERIA DE SISTEMAS (97.0/78.0)  
 | | activity = delete mod: INGENIERIA DE SISTEMAS (8.0/3.0)

| | activity = editsection: INGENIERIA ELECTRONICA (10.0/5.0)  
 | | activity = recent: INGENIERIA CIVIL (68.0/51.0)  
 | | activity = report log: INGENIERIA ECONOMICA (9.0/2.0)  
 | | activity = update: INGENIERIA ECONOMICA (4.0/2.0)  
 | | activity = update mod: INGENIERIA AGROINDUSTRIAL (30.0/20.0)  
 | | activity = view: INGENIERIA DE MINAS (4294.0/3240.0)  
 | hour = 11  
 | | activity = add mod: INGENIERIA AGROINDUSTRIAL (65.0/50.0)  
 | | activity = delete mod: INGENIERIA DE MINAS (3.0/2.0)  
 | | activity = editsection: INGENIERIA DE MINAS (5.0/2.0)  
 | | activity = recent: INGENIERIA DE MINAS (87.0/61.0)  
 | | activity = update: INGENIERIA ELECTRONICA (10.0/5.0)  
 | | activity = update mod: INGENIERIA ELECTRONICA (20.0/12.0)  
 | | activity = view: INGENIERIA DE MINAS (5240.0/3733.0)  
 | hour = 12: INGENIERIA ELECTRONICA (4116.0/2910.0)  
 | hour = 13  
 | | activity = add mod: INGENIERIA AGRICOLA (63.0/36.0)  
 | | activity = editsection: INGENIERIA DE MINAS (3.0/1.0)  
 | | activity = recent: INGENIERIA ECONOMICA (69.0/47.0)  
 | | activity = report log: INGENIERIA AGROINDUSTRIAL (1.0)  
 | | activity = update: INGENIERIA AGROINDUSTRIAL (7.0/5.0)  
 | | activity = update mod: INGENIERIA ELECTRONICA (52.0/27.0)  
 | | activity = view: INGENIERIA ECONOMICA (4625.0/3329.0)  
 | hour = 14: INGENIERIA ELECTRONICA (4584.0/3185.0)  
 | hour = 15  
 | | activity = add mod: INGENIERIA DE SISTEMAS (62.0/43.0)  
 | | activity = delete mod: INGENIERIA ECONOMICA (8.0/2.0)  
 | | activity = editsection: ARQUITECTURA Y URBANISMO (22.0/2.0)  
 | | activity = recent: INGENIERIA DE MINAS (71.0/46.0)  
 | | activity = update: INGENIERIA DE MINAS (12.0/5.0)  
 | | activity = update mod: INGENIERIA DE SISTEMAS (64.0/45.0)  
 | | activity = view: INGENIERIA DE MINAS (4929.0/3764.0)  
 | hour = 16: INGENIERIA DE MINAS (4376.0/2943.0)  
 | hour = 17: INGENIERIA DE MINAS (5758.0/3752.0)  
 | hour = 18  
 | | activity = add mod: INGENIERIA ECONOMICA (55.0/38.0)  
 | | activity = delete mod: INGENIERIA AGRICOLA (21.0/5.0)  
 | | activity = editsection: ARQUITECTURA Y URBANISMO (25.0/1.0)  
 | | activity = recent: INGENIERIA ECONOMICA (77.0/60.0)  
 | | activity = update: INGENIERIA ELECTRONICA (1.0)  
 | | activity = update mod: INGENIERIA DE MINAS (44.0/32.0)  
 | | activity = view: INGENIERIA DE MINAS (6096.0/3404.0)  
 | hour = 19  
 | | activity = add mod: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (90.0/62.0)  
 | | activity = delete mod: INGENIERIA DE MINAS (8.0/5.0)  
 | | activity = editsection: ARQUITECTURA Y URBANISMO (18.0/4.0)  
 | | activity = recent: INGENIERIA DE MINAS (73.0/51.0)  
 | | activity = update: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (9.0/5.0)  
 | | activity = update mod: ARQUITECTURA Y URBANISMO (50.0/28.0)  
 | | activity = view: INGENIERIA DE MINAS (5728.0/3161.0)  
 | hour = 20  
 | | activity = add mod: INGENIERIA ECONOMICA (65.0/41.0)  
 | | activity = delete mod: INGENIERIA DE MINAS (3.0)

- | | activity = recent: INGENIERIA ECONOMICA (73.0/51.0)
- | | activity = update: INGENIERIA DE MINAS (2.0)
- | | activity = update mod: INGENIERIA DE SISTEMAS (49.0/34.0)
- | | activity = view: INGENIERIA DE MINAS (4579.0/3029.0)
- | hour = 21
- | | activity = add mod: INGENIERIA DE MINAS (64.0/47.0)
- | | activity = delete mod: INGENIERIA ECONOMICA (8.0/1.0)
- | | activity = editsection: INGENIERIA ECONOMICA (5.0/2.0)
- | | activity = recent: INGENIERIA CIVIL (81.0/59.0)
- | | activity = update: INGENIERIA DE MINAS (2.0/1.0)
- | | activity = update mod: INGENIERIA DE MINAS (49.0/27.0)
- | | activity = view: INGENIERIA DE MINAS (4178.0/2907.0)
- | hour = 22
- | | activity = add mod: INGENIERIA ECONOMICA (43.0/26.0)
- | | activity = delete mod: INGENIERIA DE MINAS (2.0)
- | | activity = editsection: INGENIERIA ECONOMICA (3.0/1.0)
- | | activity = recent: INGENIERIA CIVIL (48.0/31.0)
- | | activity = update mod: INGENIERIA ESTADISTICA E INFORMATICA (62.0/41.0)
- | | activity = view: INGENIERIA DE MINAS (3021.0/2096.0)
  
- | hour = 23: INGENIERIA DE MINAS (2164.0/1486.0)

Durante todo el semestre académico 2017-II a las 00 horas se realizarón 815 actividades de revisión o vista del módulo curso, de las cuales 564 de ellas fueron realizadas por la escuela de Ing. de Minas.

De la misma forma observamos que siendo las 05 horas de las 814 revisiones de los cursos, 535 fueron realizados por la escuela de Arquitectura y Urbanismos.

#### **Modulo Blog**

module = blog: INGENIERÍA ECONÓMICA (992.0/340.0)

#### **Interpretación:**

Ingeniería Económica realiza la mayor cantidad de blog de los 992 publican 340

#### **Modulo Foro**

module = forum

- | hour = 00
- | | activity = add discussion: INGENIERIA AGROINDUSTRIAL (2.0/1.0)
- | | activity = search: INGENIERIA ECONOMICA (2.0)
- | | activity = update post: INGENIERIA ECONOMICA (1.0)
- | | activity = user report: INGENIERIA ECONOMICA (9.0)
- | | activity = view discussion: INGENIERIA ECONOMICA (53.0/22.0)
- | | activity = view forum: INGENIERIA DE MINAS (84.0/49.0)
- | hour = 01: INGENIERIA ECONOMICA (89.0/58.0)
- | hour = 02
- | | activity = add discussion: INGENIERIA AGROINDUSTRIAL (2.0/1.0)
- | | activity = view discussion: INGENIERIA ECONOMICA (16.0/7.0)
- | | activity = view forum: INGENIERIA DE SISTEMAS (23.0/12.0)
- | hour = 03: INGENIERIA DE MINAS (40.0/23.0)
- | hour = 04
- | | activity = view discussion: INGENIERIA ELECTRONICA (28.0/15.0)
- | | activity = view forum: INGENIERIA DE MINAS (29.0/20.0)

| hour = 05  
 | | activity = add discussion: INGENIERIA ECONOMICA (3.0/1.0)  
 | | activity = view discussion: INGENIERIA AGROINDUSTRIAL (45.0/25.0)  
 | | activity = view forum: INGENIERIA ECONOMICA (72.0/51.0)  
 | hour = 06: INGENIERIA ECONOMICA (266.0/170.0)  
 | hour = 07: INGENIERIA ECONOMICA (463.0/299.0)  
 | hour = 08  
 | | activity = add discussion: INGENIERIA ELECTRONICA (30.0/20.0)  
 | | activity = search: INGENIERIA QUIMICA (6.0/4.0)  
 | | activity = update post: INGENIERIA ECONOMICA (3.0/1.0)  
 | | activity = user report: INGENIERIA CIVIL (2.0)  
 | | activity = view discussion: INGENIERIA ECONOMICA (266.0/154.0)  
 | | activity = view forum: INGENIERIA ECONOMICA (379.0/285.0)  
 | | activity = view forums: INGENIERIA AGROINDUSTRIAL (2.0/1.0)  
 | hour = 09  
 | | activity = add discussion: INGENIERIA ELECTRONICA (22.0/5.0)  
 | | activity = add post: INGENIERIA ELECTRONICA (6.0/3.0)  
 | | activity = search: INGENIERIA ELECTRONICA (2.0)  
 | | activity = update post: INGENIERIA ELECTRONICA (2.0/1.0)  
 | | activity = view discussion: INGENIERIA ECONOMICA (371.0/264.0)  
 | | activity = view forum: INGENIERIA ECONOMICA (465.0/347.0)  
 | hour = 10  
 | | activity = add discussion: INGENIERIA DE MINAS (9.0/5.0)  
 | | activity = search: INGENIERIA ECONOMICA (5.0/2.0)  
 | | activity = submission statement accepted: INGENIERIA ECONOMICA (0.0)  
 | | activity = update post: INGENIERIA ECONOMICA (3.0/1.0)  
 | | activity = user report: INGENIERIA DE MINAS (5.0/2.0)  
 | | activity = view discussion: INGENIERIA ECONOMICA (353.0/222.0)  
 | | activity = view forum: INGENIERIA DE MINAS (452.0/325.0)  
 | hour = 11  
 | | activity = add discussion: INGENIERIA ECONOMICA (6.0/3.0)  
 | | activity = add post: INGENIERIA DE MINAS (8.0/2.0)  
 | | activity = search: INGENIERIA DE MINAS (8.0/5.0)  
 | | activity = view discussion: INGENIERIA ECONOMICA (392.0/217.0)  
 | | activity = view forum: INGENIERIA DE MINAS (580.0/379.0)  
 | hour = 12: INGENIERIA ECONOMICA (798.0/485.0)  
 | hour = 13  
 | | activity = add discussion: INGENIERIA ELECTRONICA (21.0/3.0)  
 | | activity = search: INGENIERIA ECONOMICA (4.0/1.0)  
 | | activity = update post: INGENIERIA ELECTRONICA (5.0/1.0)  
 | | activity = user report: INGENIERIA ECONOMICA (3.0)  
 | | activity = view discussion: INGENIERIA ECONOMICA (368.0/124.0)  
 | | activity = view forum: INGENIERIA ECONOMICA (521.0/275.0)  
 | hour = 14: INGENIERIA ECONOMICA (684.0/437.0)  
 | hour = 15: INGENIERIA ECONOMICA (880.0/565.0)  
 | hour = 16  
 | | activity = add: INGENIERIA DE MINAS (4.0/2.0)  
 | | activity = add discussion: INGENIERIA DE MINAS (6.0/4.0)  
 | | activity = add post: INGENIERIA DE SISTEMAS (17.0/1.0)  
 | | activity = search: INGENIERIA ECONOMICA (11.0/5.0)  
 | | activity = update: INGENIERIA DE MINAS (3.0/2.0)  
 | | activity = update post: INGENIERIA DE SISTEMAS (5.0/2.0)  
 | | activity = view discussion: INGENIERIA ECONOMICA (294.0/210.0)

```

| | activity = view forum: INGENIERIA DE MINAS (443.0/305.0)
| hour = 17
| | activity = add discussion: INGENIERIA DE SISTEMAS (10.0/6.0)
| | activity = add post: INGENIERIA DE SISTEMAS (9.0)
| | activity = search: INGENIERIA ECONOMICA (9.0/4.0)
| | activity = view discussion: INGENIERIA ECONOMICA (309.0/201.0)
| | activity = view forum: INGENIERIA ECONOMICA (606.0/436.0)
| hour = 18
| | activity = add discussion: INGENIERIA AGROINDUSTRIAL (9.0/5.0)
| | activity = search: INGENIERIA ECONOMICA (9.0/6.0)
| | activity = update post: INGENIERIA DE MINAS (2.0/1.0)
| | activity = view discussion: INGENIERIA ECONOMICA (407.0/271.0)
| | activity = view forum: INGENIERIA DE MINAS (555.0/369.0)
| hour = 19
| | activity = add: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (14.0/1.0)
| | activity = add discussion: INGENIERIA AGROINDUSTRIAL (13.0/7.0)
| | activity = add post: INGENIERIA DE MINAS (4.0/1.0)
| | activity = search: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (7.0/3.0)
| | activity = user report: INGENIERIA DE MINAS (6.0/2.0)
| | activity = view discussion: INGENIERIA DE MINAS (311.0/224.0)
| | activity = view forum: INGENIERIA DE MINAS (562.0/369.0)
| hour = 20
| | activity = add discussion: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (35.0/20.0)
| | activity = add post: INGENIERIA DE MINAS (5.0/2.0)
| | activity = search: INGENIERIA DE MINAS (5.0/4.0)
| | activity = update post: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (4.0/1.0)
| | activity = view discussion: INGENIERIA ECONOMICA (390.0/227.0)
| | activity = view forum: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (665.0/460.0)
| hour = 21: INGENIERIA ECONOMICA (741.0/476.0)
| hour = 22
| | activity = add discussion: INGENIERIA AGROINDUSTRIAL (15.0/8.0)
| | activity = add post: INGENIERIA DE MINAS (7.0/3.0)
| | activity = search: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (7.0/3.0)
| | activity = view discussion: INGENIERIA ECONOMICA (202.0/142.0)
| | activity = view forum: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (440.0/328.0)
| hour = 23: INGENIERIA ECONOMICA (418.0/301.0)

```

Analizando los resultados en relación al módulo Foro, concluimos que las escuelas profesionales de Ing, de Minas, Ing. Económica, presentan una mayor actividad en relación a otras escuelas.

Lo interesante es una importante interactividad con la plataforma desde las 06 horas, específicamente con este módulo.

#### **Módulo Asignación de Tareas:**

```

module = assign
| hour = 00
| | activity = grade submission: INGENIERIA ESTADISTICA E INFORMATICA (16.0/6.0)
| | activity = submit: INGENIERIA DE MINAS (57.0/39.0)
| | activity = view: INGENIERIA DE MINAS (458.0/333.0)
| | activity = view grading form: INGENIERIA ESTADISTICA E INFORMATICA (17.0/7.0)
| | activity = view submission grading table: INGENIERIA ESTADISTICA E INFORMATICA (5.0/2.0)

```

| | activity = view submit assignment form: INGENIERIA DE MINAS (114.0/78.0)  
 | hour = 01: INGENIERIA DE MINAS (184.0/136.0)  
 | hour = 02: ARQUITECTURA Y URBANISMO (106.0/64.0)  
 | hour = 03  
 | | activity = submit: INGENIERIA DE MINAS (6.0/2.0)  
 | | activity = view: INGENIERIA ECONOMICA (96.0/51.0)  
 | | activity = view submit assignment form: INGENIERIA DE MINAS (16.0/4.0)  
 | hour = 04  
 | | activity = submit: ARQUITECTURA Y URBANISMO (9.0/6.0)  
 | | activity = update: ARQUITECTURA Y URBANISMO (40.0/2.0)  
 | | activity = view: INGENIERIA ECONOMICA (156.0/89.0)  
 | | activity = view submit assignment form: INGENIERIA CIVIL (16.0/10.0)  
 | hour = 05  
 | | activity = submit: ARQUITECTURA Y URBANISMO (6.0/4.0)  
 | | activity = view: INGENIERIA ECONOMICA (186.0/97.0)  
 | | activity = view grading form: INGENIERIA ECONOMICA (123.0/1.0)  
 | | activity = view submission grading table: INGENIERIA ECONOMICA (73.0/16.0)  
 | | activity = view submit assignment form: ARQUITECTURA Y URBANISMO (19.0/11.0)  
 | hour = 06  
 | | activity = add: INGENIERIA DE SISTEMAS (9.0/4.0)  
 | | activity = download all submissions: INGENIERIA DE SISTEMAS (6.0/1.0)  
 | | activity = grade submission: INGENIERIA ECONOMICA (152.0/26.0)  
 | | activity = submit: ARQUITECTURA Y URBANISMO (67.0/21.0)  
 | | activity = update: INGENIERIA ECONOMICA (13.0/6.0)  
 | | activity = view: ARQUITECTURA Y URBANISMO (546.0/338.0)  
 | | activity = view submission grading table: INGENIERIA ECONOMICA (40.0/7.0)  
 | | activity = view submit assignment form: ARQUITECTURA Y URBANISMO (107.0/42.0)  
 | hour = 07  
 | | activity = add: INGENIERIA DE MINAS (5.0/3.0)  
 | | activity = download all submissions: INGENIERIA DE MINAS (2.0/1.0)  
 | | activity = grade submission: INGENIERIA ECONOMICA (139.0/13.0)  
 | | activity = submit: INGENIERIA ECONOMICA (32.0/22.0)  
 | | activity = view: INGENIERIA ECONOMICA (614.0/445.0)  
 | | activity = view all: INGENIERIA DE MINAS (3.0/1.0)  
 | | activity = view grading form: INGENIERIA ECONOMICA (152.0/18.0)  
 | | activity = view submission grading table: INGENIERIA ECONOMICA (70.0/41.0)  
 | | activity = view submit assignment form: INGENIERIA ECONOMICA (82.0/54.0)  
 | hour = 08  
 | | activity = add: INGENIERIA AGROINDUSTRIAL (20.0/15.0)  
 | | activity = grade submission: INGENIERIA ECONOMICA (208.0/24.0)  
 | | activity = submit: INGENIERIA AGROINDUSTRIAL (111.0/70.0)  
 | | activity = update: ARQUITECTURA Y URBANISMO (18.0/11.0)  
 | | activity = view: INGENIERIA AGROINDUSTRIAL (1180.0/873.0)  
 | | activity = view submission grading table: INGENIERIA AGROINDUSTRIAL (57.0/34.0)  
 | | activity = view submit assignment form: INGENIERIA AGROINDUSTRIAL (180.0/129.0)  
 | hour = 09  
 | | activity = add: INGENIERIA AGROINDUSTRIAL (20.0/10.0)  
 | | activity = download all submissions: INGENIERIA DE SISTEMAS (7.0/2.0)  
 | | activity = submit: INGENIERIA AGROINDUSTRIAL (195.0/116.0)  
 | | activity = update: ARQUITECTURA Y URBANISMO (39.0/13.0)  
 | | activity = view: INGENIERIA AGROINDUSTRIAL (1559.0/1184.0)  
 | | activity = view submission grading table: INGENIERIA ECONOMICA (93.0/60.0)  
 | | activity = view submit assignment form: INGENIERIA AGROINDUSTRIAL (303.0/199.0)

| hour = 10

- | | activity = add: INGENIERIA DE SISTEMAS (17.0/12.0)
- | | activity = grade submission: INGENIERIA AGROINDUSTRIAL (61.0/19.0)
- | | activity = submit: INGENIERIA DE MINAS (125.0/99.0)
- | | activity = update: INGENIERIA AGROINDUSTRIAL (8.0/2.0)
- | | activity = view: INGENIERIA ECONOMICA (1267.0/949.0)
- | | activity = view grading form: INGENIERIA AGROINDUSTRIAL (74.0/19.0)
- | | activity = view submission grading table: INGENIERIA DE SISTEMAS (42.0/18.0)
- | | activity = view submit assignment form: INGENIERIA DE MINAS (235.0/186.0)

| hour = 11

- | | activity = add: INGENIERIA AGRICOLA (18.0/12.0)
- | | activity = grade submission: INGENIERIA DE MINAS (4.0/2.0)
- | | activity = submission statement accepted: INGENIERIA ECONOMICA (0.0)
- | | activity = submit: INGENIERIA ELECTRONICA (191.0/101.0)
- | | activity = view: INGENIERIA ECONOMICA (1633.0/1114.0)
- | | activity = view grading form: INGENIERIA ESTADISTICA E INFORMATICA (8.0/3.0)
- | | activity = view submission grading table: INGENIERIA DE MINAS (80.0/52.0)
- | | activity = view submit assignment form: INGENIERIA ELECTRONICA (328.0/212.0)

| hour = 12

- | | activity = add: INGENIERIA AGROINDUSTRIAL (5.0/2.0)
- | | activity = grade submission: INGENIERIA ESTADISTICA E INFORMATICA (14.0/1.0)
- | | activity = submit: INGENIERIA ECONOMICA (87.0/68.0)
- | | activity = update: INGENIERIA AGROINDUSTRIAL (4.0/2.0)
- | | activity = view: INGENIERIA ECONOMICA (1147.0/751.0)
- | | activity = view grading form: INGENIERIA ESTADISTICA E INFORMATICA (14.0/1.0)
- | | activity = view submission grading table: INGENIERIA AGRICOLA (38.0/20.0)
- | | activity = view submit assignment form: INGENIERIA ECONOMICA (179.0/139.0)

| hour = 13

- | | activity = add: INGENIERIA ECONOMICA (10.0/6.0)
- | | activity = grade submission: INGENIERIA AGROINDUSTRIAL (6.0/1.0)
- | | activity = submit: INGENIERIA ELECTRONICA (155.0/99.0)
- | | activity = view: INGENIERIA ECONOMICA (1997.0/1095.0)
- | | activity = view all: INGENIERIA ECONOMICA (4.0/1.0)
- | | activity = view grading form: INGENIERIA AGROINDUSTRIAL (17.0/4.0)
- | | activity = view submission grading table: INGENIERIA AGROINDUSTRIAL (31.0/18.0)
- | | activity = view submit assignment form: INGENIERIA ELECTRONICA (261.0/180.0)

| hour = 14

- | | activity = add: INGENIERIA AGROINDUSTRIAL (11.0/8.0)
- | | activity = submit: INGENIERIA ECONOMICA (240.0/130.0)
- | | activity = update: INGENIERIA AGROINDUSTRIAL (11.0/7.0)
- | | activity = view: INGENIERIA ECONOMICA (1626.0/968.0)
- | | activity = view submission grading table: INGENIERIA DE SISTEMAS (33.0/16.0)
- | | activity = view submit assignment form: INGENIERIA ECONOMICA (351.0/220.0)

| hour = 15

- | | activity = add: INGENIERIA DE SISTEMAS (15.0/9.0)
- | | activity = submit: INGENIERIA ECONOMICA (383.0/276.0)
- | | activity = update: INGENIERIA ECONOMICA (23.0/12.0)
- | | activity = view: INGENIERIA ECONOMICA (2145.0/1327.0)
- | | activity = view all: ARQUITECTURA Y URBANISMO (2.0/1.0)
- | | activity = view submission grading table: INGENIERIA DE MINAS (225.0/100.0)
- | | activity = view submit assignment form: INGENIERIA ECONOMICA (510.0/371.0)

| hour = 16

- | | activity = add: INGENIERIA ELECTRONICA (13.0/6.0)

| | activity = submit: INGENIERIA DE SISTEMAS (187.0/123.0)  
 | | activity = view: INGENIERIA ECONOMICA (1483.0/1058.0)  
 | | activity = view grading form: INGENIERIA DE MINAS (17.0/1.0)  
 | | activity = view submission grading table: INGENIERIA DE MINAS (144.0/45.0)  
 | | activity = view submit assignment form: INGENIERIA AGRICOLA (382.0/295.0)  
 | hour = 17  
 | | activity = add: INGENIERIA AGRICOLA (12.0/8.0)  
 | | activity = grade submission: INGENIERIA AGRONOMICA (34.0/12.0)  
 | | activity = submit: INGENIERIA DE MINAS (213.0/158.0)  
 | | activity = update: ARQUITECTURA Y URBANISMO (31.0/8.0)  
 | | activity = view: INGENIERIA DE MINAS (1934.0/1399.0)  
 | | activity = view grading form: INGENIERIA AGRONOMICA (48.0/20.0)  
 | | activity = view submission grading table: INGENIERIA DE MINAS (81.0/50.0)  
 | | activity = view submit assignment form: INGENIERIA DE MINAS (371.0/254.0)  
 | hour = 18: INGENIERIA DE MINAS (2739.0/1846.0)  
 | hour = 19  
 | | activity = add: INGENIERIA ECONOMICA (12.0/6.0)  
 | | activity = grade submission: INGENIERIA ECONOMICA (35.0/14.0)  
 | | activity = submit: INGENIERIA DE MINAS (234.0/148.0)  
 | | activity = update: ARQUITECTURA Y URBANISMO (11.0/2.0)  
 | | activity = view: INGENIERIA DE MINAS (1837.0/1135.0)  
 | | activity = view all: ARQUITECTURA Y URBANISMO (4.0/2.0)  
 | | activity = view grading form: INGENIERIA ECONOMICA (35.0/15.0)  
 | | activity = view submission grading table: INGENIERIA DE MINAS (32.0/14.0)  
 | | activity = view submit assignment form: INGENIERIA DE MINAS (445.0/251.0)  
 | hour = 20: INGENIERIA DE MINAS (2930.0/1937.0)  
 | hour = 21  
 | | activity = grade submission: INGENIERIA ESTADISTICA E INFORMATICA (68.0/9.0)  
 | | activity = submit: INGENIERIA DE MINAS (253.0/124.0)  
 | | activity = update: ARQUITECTURA Y URBANISMO (11.0/7.0)  
 | | activity = view: INGENIERIA DE MINAS (1993.0/1343.0)  
 | | activity = view all: INGENIERIA DE MINAS (5.0/2.0)  
 | | activity = view grading form: INGENIERIA ESTADISTICA E INFORMATICA (72.0/10.0)  
 | | activity = view submission grading table: INGENIERIA DE MINAS (37.0/11.0)  
 | | activity = view submit assignment form: INGENIERIA DE MINAS (448.0/230.0)  
 | hour = 22  
 | | activity = add: INGENIERIA DE MINAS (8.0/1.0)  
 | | activity = grade submission: INGENIERIA ECONOMICA (11.0/2.0)  
 | | activity = submit: INGENIERIA DE MINAS (228.0/142.0)  
 | | activity = update: ARQUITECTURA Y URBANISMO (16.0/7.0)  
 | | activity = view: INGENIERIA DE MINAS (1515.0/1067.0)  
 | | activity = view all: INGENIERIA DE MINAS (2.0/1.0)  
 | | activity = view grading form: INGENIERIA ECONOMICA (14.0/3.0)  
 | | activity = view submission grading table: INGENIERIA DE MINAS (26.0/13.0)  
 | | activity = view submit assignment form: INGENIERIA DE MINAS (417.0/264.0)  
 | hour = 23  
 | | activity = add: INGENIERIA DE MINAS (9.0/4.0)  
 | | activity = grade submission: INGENIERIA ESTADISTICA E INFORMATICA (16.0/1.0)  
 | | activity = submit: ARQUITECTURA Y URBANISMO (252.0/145.0)  
 | | activity = update: INGENIERIA DE MINAS (6.0/1.0)  
 | | activity = view: ARQUITECTURA Y URBANISMO (1390.0/910.0)  
 | | activity = view grading form: INGENIERIA ESTADISTICA E INFORMATICA (17.0/1.0)  
 | | activity = view submission grading table: INGENIERIA DE MINAS (37.0/3.0)

| | activity = view submit assignment form: ARQUITECTURA Y URBANISMO (446.0/273.0)

En el módulo de Asignación de Tareas otras escuelas ingresan a la actividad virtual como Ing., Estadística e Informática, Arquitectura y Urbanismo, Ing. Agrícola, Ing. de Sistemas e Ing. Electrónica, sin embargo, nuevamente resalta que las escuelas de Ing. de Minas e Ing. Económica son las que más actividad muestran.

### Modulo Examen

module = quiz

| hour = 00: INGENIERÍA DE MINAS (4.0)  
 | hour = 01: INGENIERIA DE MINAS (11.0)  
 | hour = 02: INGENIERIA DE MINAS (10.0)  
 | hour = 03: INGENIERIA DE MINAS (15.0)  
 | hour = 04: INGENIERIA DE MINAS (29.0/1.0)  
 | hour = 05  
 | | activity = add: INGENIERIA DE MINAS (1.0)  
 | | activity = editquestions: INGENIERIA DE MINAS (1.0)  
 | | activity = report: INGENIERIA ELECTRONICA (34.0)  
 | | activity = update: INGENIERIA DE MINAS (1.0)  
 | | activity = view: INGENIERIA DE MINAS (8.0)  
 | hour = 06  
 | | activity = add: INGENIERIA DE MINAS (1.0)  
 | | activity = editquestions: INGENIERIA DE MINAS (39.0)  
 | | activity = report: INGENIERIA ELECTRONICA (9.0)  
 | | activity = review: INGENIERIA DE MINAS (11.0)  
 | | activity = view: INGENIERIA DE MINAS (17.0/5.0)  
 | hour = 07  
 | | activity = add: INGENIERIA DE MINAS (3.0)  
 | | activity = attempt: INGENIERIA DE MINAS (4.0)  
 | | activity = continue attempt: INGENIERIA DE MINAS (31.0)  
 | | activity = editquestions: INGENIERIA DE MINAS (48.0)  
 | | activity = preview: INGENIERIA DE MINAS (2.0)  
 | | activity = report: INGENIERIA DE MINAS (2.0)  
 | | activity = review: INGENIERIA ELECTRONICA (13.0)  
 | | activity = update: INGENIERIA DE MINAS (2.0)  
 | | activity = view: INGENIERIA DE MINAS (146.0/9.0)  
 | | activity = view summary: INGENIERIA DE MINAS (4.0)  
 | hour = 08: INGENIERIA DE MINAS (1050.0/12.0)  
 | hour = 09: INGENIERIA DE MINAS (1216.0/14.0)  
 | hour = 10  
 | | activity = add: INGENIERIA ELECTRONICA (3.0)  
 | | activity = attempt: INGENIERIA DE MINAS (6.0)  
 | | activity = close attempt: INGENIERIA DE MINAS (13.0)  
 | | activity = continue attempt: INGENIERIA DE MINAS (109.0)  
 | | activity = editquestions: INGENIERIA ELECTRONICA (79.0)  
 | | activity = review: INGENIERIA DE MINAS (133.0)  
 | | activity = update: INGENIERIA ELECTRONICA (2.0)  
 | | activity = view: INGENIERIA DE MINAS (55.0/14.0)  
 | | activity = view summary: INGENIERIA DE MINAS (17.0)  
 | hour = 11

```

| | activity = review: INGENIERIA DE MINAS (9.0)
| | activity = view: INGENIERIA ELECTRONICA (123.0/22.0)
| hour = 12: INGENIERIA ELECTRONICA (965.0/47.0)
| hour = 13
| | activity = add: INGENIERIA ESTADISTICA E INFORMATICA (1.0)
| | activity = attempt: INGENIERIA ELECTRONICA (118.0/1.0)
| | activity = close attempt: INGENIERIA ELECTRONICA (112.0/1.0)
| | activity = continue attempt: INGENIERIA ELECTRONICA (1041.0/1.0)
| | activity = editquestions: INGENIERIA ESTADISTICA E INFORMATICA (20.0)
| | activity = manualgrade: INGENIERIA ESTADISTICA E INFORMATICA (2.0)
| | activity = report: INGENIERIA ESTADISTICA E INFORMATICA (3.0)
| | activity = review: INGENIERIA ELECTRONICA (81.0/14.0)
| | activity = update: INGENIERIA ELECTRONICA (5.0/2.0)
| | activity = view: INGENIERIA ELECTRONICA (410.0/52.0)
| | activity = view summary: INGENIERIA ELECTRONICA (111.0/1.0)
| hour = 14: INGENIERIA ELECTRONICA (3453.0/57.0)
| hour = 15: INGENIERIA ELECTRONICA (608.0/65.0)
| hour = 16: INGENIERIA ELECTRONICA (878.0/29.0)
| hour = 17
| | activity = add: INGENIERIA DE MINAS (5.0/2.0)
| | activity = attempt: INGENIERIA DE SISTEMAS (26.0/11.0)
| | activity = close attempt: INGENIERIA ELECTRONICA (11.0)
| | activity = continue attempt: INGENIERIA ELECTRONICA (220.0/18.0)
| | activity = editquestions: INGENIERIA DE SISTEMAS (24.0/9.0)
| | activity = report: INGENIERIA DE MINAS (4.0/1.0)
| | activity = review: INGENIERIA ELECTRONICA (227.0/12.0)
| | activity = update: INGENIERIA ELECTRONICA (3.0/1.0)
| | activity = view: INGENIERIA ELECTRONICA (244.0/144.0)
| | activity = view summary: INGENIERIA ELECTRONICA (13.0/1.0)
| hour = 18: INGENIERIA DE MINAS (3643.0/759.0)
| hour = 19: INGENIERIA DE MINAS (2653.0/65.0)
| hour = 20: INGENIERIA DE MINAS (329.0/11.0)
| hour = 21: INGENIERIA DE MINAS (369.0/13.0)
| hour = 22: INGENIERIA DE MINAS (168.0/34.0)
| hour = 23
| | activity = review: INGENIERIA ELECTRONICA (22.0)
| | activity = view: INGENIERIA DE MINAS (18.0/5.0)

```

Respecto al módulo Examen, la diferencia es abismal, en el uso de esta herramienta por parte de las usuales escuelas profesionales (Ing. Minas, Ing. Economía, Ing. Sistemas e Ing. Electrónica), siendo escasa la presencia en Ing. Estadística e informática, sin embargo, al ser este un módulo muy importante, preocupa la nula actividad de las demás escuelas profesionales.

#### Modulo Archivo

module = resource

```

| hour = 00
| | activity = add: INGENIERIA AGRONOMICA (2.0/1.0)
| | activity = view: INGENIERIA DE MINAS (187.0/121.0)
| hour = 01
| | activity = add: INGENIERIA CIVIL (7.0)
| | activity = view: INGENIERIA ECONOMICA (79.0/59.0)
| hour = 02: INGENIERIA ECONOMICA (44.0/32.0)

```

| hour = 03: INGENIERIA DE MINAS (57.0/23.0)  
 | hour = 04: INGENIERIA ECONOMICA (88.0/57.0)  
 | hour = 05  
 | | activity = add: INGENIERIA ELECTRONICA (12.0/3.0)  
 | | activity = view: INGENIERIA ECONOMICA (158.0/102.0)  
 | hour = 06: INGENIERIA ECONOMICA (284.0/197.0)  
 | hour = 07  
 | | activity = add: INGENIERIA ECONOMICA (44.0/26.0)  
 | | activity = update: INGENIERIA CIVIL (5.0/1.0)  
 | | activity = view: INGENIERIA DE MINAS (884.0/515.0)  
 | hour = 08  
 | | activity = add: INGENIERIA DE MINAS (45.0/36.0)  
 | | activity = update: INGENIERIA CIVIL (4.0)  
 | | activity = view: INGENIERIA DE MINAS (994.0/768.0)  
 | hour = 09  
 | | activity = add: INGENIERIA AGRICOLA (35.0/25.0)  
 | | activity = view: INGENIERIA ECONOMICA (1254.0/989.0)  
 | hour = 10  
 | | activity = add: INGENIERIA DE SISTEMAS (54.0/40.0)  
 | | activity = update: INGENIERIA DE MINAS (3.0/2.0)  
 | | activity = view: INGENIERIA DE MINAS (1165.0/907.0)  
 | hour = 11: INGENIERIA DE MINAS (1336.0/894.0)  
 | hour = 12  
 | | activity = add: INGENIERIA ELECTRONICA (30.0/15.0)  
 | | activity = update: INGENIERIA ELECTRONICA (2.0)  
 | | activity = view: INGENIERIA DE MINAS (988.0/772.0)  
 | hour = 13  
 | | activity = add: INGENIERIA AGRICOLA (38.0/16.0)  
 | | activity = update: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (3.0)  
 | | activity = view: INGENIERIA ELECTRONICA (971.0/781.0)  
 | hour = 14  
 | | activity = add: INGENIERIA QUIMICA (34.0/26.0)  
 | | activity = view: INGENIERIA ECONOMICA (1211.0/935.0)  
 | hour = 15: INGENIERIA DE SISTEMAS (1588.0/1003.0)  
 | hour = 16: INGENIERIA DE MINAS (1369.0/1025.0)  
 | hour = 17  
 | | activity = add: INGENIERIA ECONOMICA (44.0/27.0)  
 | | activity = view: INGENIERIA DE MINAS (1971.0/1027.0)  
 | hour = 18  
 | | activity = add: INGENIERIA ECONOMICA (28.0/17.0)  
 | | activity = view: INGENIERIA DE MINAS (1561.0/847.0)  
 | hour = 19  
 | | activity = add: INGENIERIA DE MINAS (42.0/32.0)  
 | | activity = update: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (5.0/1.0)  
 | | activity = view: INGENIERIA DE MINAS (1291.0/847.0)  
 | hour = 20  
 | | activity = add: INGENIERIA ECONOMICA (47.0/25.0)  
 | | activity = update: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (2.0)  
 | | activity = view: INGENIERIA DE MINAS (1081.0/770.0)  
 | hour = 21  
 | | activity = add: INGENIERIA ELECTRONICA (39.0/30.0)  
 | | activity = update: INGENIERIA DE MINAS (2.0/1.0)  
 | | activity = view: INGENIERIA DE MINAS (814.0/596.0)

```
| hour = 22
| | activity = add: INGENIERIA ECONOMICA (19.0/7.0)
| | activity = view: INGENIERIA DE MINAS (627.0/446.0)
| hour = 23: INGENIERIA DE MINAS (426.0/312.0)
```

Con respecto al modulo Archivo, observamos el escaso uso de esta por las escuelas profesionales, sin embargo, este bajo uso de la plataforma lo contradicen las escuelas habituales (Ing. Minas, Ing. Economía, Ing. Sistemas e Ing. Electrónica), resaltando el “ingreso” de Ing. Agronómica e Ing. Topográfica y Agrimensura.

### Modulo Folder

```
module = folder: INGENIERIA DE MINAS (827.0/137.0)
```

Es interesante este resultado puesto se crea una carpeta con la finalidad de subir o poner a disposición el material académico inmerso en esta. El proceso de clasificación solo muestra a la escuela de Ing. de Minas, esto no significa que las otras escuelas no hayan usado este módulo, sin embargo, el proceso de clasificación resalta no solo la activación de este módulo, sino que de las 827 carpetas (Folder) creados con material académico, solo 137 pertenecientes a Ing. de Minas tuvieron una importante interactividad de parte de sus usuarios, mas no de las otras escuelas.

### Modulo URL (Direccion WEB)

```
module = url
| hour = 00: INGENIERIA DE MINAS (11.0/7.0)
| hour = 01: INGENIERIA DE SISTEMAS (4.0/2.0)
| hour = 02: INGENIERIA CIVIL (1.0)
| hour = 03: INGENIERIA ELECTRONICA (14.0/2.0)
| hour = 04: INGENIERIA ELECTRONICA (6.0/3.0)
| hour = 05: INGENIERIA DE MINAS (23.0/8.0)
| hour = 06: INGENIERIA ELECTRONICA (23.0/13.0)
| hour = 07: INGENIERIA ECONOMICA (41.0/29.0)
| hour = 08: INGENIERIA AGROINDUSTRIAL (68.0/43.0)
| hour = 09
| | activity = add: INGENIERIA CIVIL (6.0/3.0)
| | activity = update: INGENIERIA CIVIL (1.0)
| | activity = view: INGENIERIA DE MINAS (108.0/61.0)
| hour = 10: INGENIERIA DE MINAS (88.0/33.0)
| hour = 11
| | activity = add: INGENIERIA AGROINDUSTRIAL (3.0)
| | activity = view: INGENIERIA DE MINAS (73.0/34.0)
| hour = 12: INGENIERIA DE MINAS (34.0/25.0)
| hour = 13: INGENIERIA DE MINAS (76.0/47.0)
| hour = 14: INGENIERIA DE MINAS (104.0/64.0)
| hour = 15: INGENIERIA DE MINAS (130.0/66.0)
| hour = 16
| | activity = add: INGENIERIA DE SISTEMAS (2.0)
| | activity = view: INGENIERIA DE MINAS (104.0/51.0)
| hour = 17: INGENIERIA DE SISTEMAS (204.0/95.0)
| hour = 18: INGENIERIA DE SISTEMAS (160.0/109.0)
| hour = 19: INGENIERIA DE MINAS (155.0/59.0)
| hour = 20
```

```

| | activity = add: INGENIERIA ELECTRONICA (5.0/1.0)
| | activity = update: INGENIERIA ELECTRONICA (3.0)
| | activity = view: INGENIERIA DE MINAS (126.0/69.0)
| hour = 21
| | activity = add: INGENIERIA TOPOGRAFICA Y AGRIMENSURA (4.0/1.0)
| | activity = update: INGENIERIA ELECTRONICA (6.0/1.0)
| | activity = view: INGENIERIA DE MINAS (94.0/61.0)
| hour = 22: INGENIERIA DE MINAS (61.0/42.0)
| hour = 23: INGENIERIA DE MINAS (43.0/28.0)
    
```

El proponer direcciones web (direcciones URL) a los estudiantes es interesante para ampliar o validar conocimientos, en este módulo en particular la escuela de Ing. de Sistemas, resalta por el uso de éste.

### Modulo Chat

module = chat: INGENIERIA AGROINDUSTRIAL (71.0/24.0)

Interesante resultado, de los 71 chats que se activaron en el semestre 24 fueron implementados por la escuela de Ing. Agroindustrial, mostrando una comunicación fluida con relación a las otras escuelas.

Number of Leaves (número de hojas):	5680
Size of the tree (tamaño del árbol):	5788
=== Summary ===	
Correctly Classified Instances	77399 40.3141 %
Incorrectly Classified Instances	114591 59.6859 %
Kappa statistic	0.2181
Mean absolute error	0.0899
Root mean squared error	0.212
Relative absolute error	86.3521 %
Root relative squared error	92.9263 %
Total Number of Instances	191990

Tabla 2  
*Precisión detallada por clases*  
=== Detailed Accuracy By Class =

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0,193	0,017	0,427	0,193	0,266	0,257	0,809	0,280	ARQUITECTURA Y URBANISMO
0,000	0,000	?	0,000	?	?	0,945	0,009	CIENCIAS FÍSICO MATEMÁTICAS
0,773	0,498	0,392	0,773	0,520	0,254	0,727	0,579	INGENIERÍA DE MINAS
0,033	0,002	0,314	0,033	0,059	0,093	0,755	0,110	INGENIERÍA AGRÍCOLA
0,134	0,017	0,311	0,134	0,188	0,176	0,739	0,178	INGENIERÍA AGROINDUSTRIAL
0,012	0,000	0,607	0,012	0,024	0,084	0,757	0,080	INGENIERÍA AGRONÓMICA
0,021	0,002	0,411	0,021	0,040	0,083	0,720	0,130	INGENIERÍA CIVIL
0,095	0,010	0,388	0,095	0,152	0,168	0,755	0,197	INGENIERÍA DE SISTEMAS
0,396	0,170	0,363	0,396	0,379	0,219	0,702	0,396	INGENIERÍA ECONÓMICA
0,420	0,069	0,537	0,420	0,471	0,388	0,799	0,557	INGENIERÍA ELECTRÓNICA
0,062	0,000	0,755	0,062	0,115	0,213	0,738	0,134	INGENIERÍA ESTADÍSTICA E INFORMÁTICA
0,000	0,000	?	0,000	?	?	0,876	0,019	INGENIERÍA GEOLÓGICA
0,000	0,000	?	0,000	?	?	0,973	0,004	INGENIERÍA MECÁNICA ELÉCTRICA
0,000	0,000	?	0,000	?	?	0,844	0,023	INGENIERÍA METALÚRGICA
0,010	0,000	0,250	0,010	0,020	0,050	0,878	0,037	INGENIERÍA QUÍMICA
0,057	0,005	0,316	0,057	0,096	0,120	0,784	0,139	INGENIERÍA TOPOGRAF. Y AGRIMENSURA
Weighted Avg.								
0,403	0,194	?	0,403	?	?	0,746	0,396	

Tabla 3  
Matriz de Confusion  
==== Confusion Matrix =====

a	b	c	d	e	F	g	h	i	J	k	l	m	n	o	p	Class
2314	0	6203	19	485	0	32	64	2234	577	14	0	0	0	0	20	a = ARQUITECTURA Y URBANISMO
5	0	38	0	1	0	0	0	19	15	0	0	0	0	0	2	b = CIENCIAS FÍSICO MATEMÁTICAS
975	0	43553	89	593	4	77	529	6988	3234	32	0	0	0	2	253	c = INGENIERÍA DE MINAS
116	0	3560	200	196	0	9	122	1271	599	0	0	0	0	5	16	d = INGENIERÍA AGRÍCOLA
243	0	5119	32	1357	6	18	83	2340	789	6	0	0	0	2	105	e = INGENIERÍA AGROINDUSTRIAL
43	0	2633	0	41	51	5	108	769	4590	0	0	0	0	6	12	f = INGENIERÍA AGRONÓMICA
505	0	6216	4	74	0	200	75	1661	797	0	0	0	0	1	64	g = INGENIERÍA CIVIL
132	0	6783	80	119	4	7	1108	2590	811	0	0	0	0	3	56	h = INGENIERÍA DE SISTEMAS
622	0	17875	99	786	18	70	378	14969	2672	29	0	0	0	6	300	i = INGENIERÍA ECONÓMICA
299	0	11737	62	530	1	36	266	4937	12969	1	0	0	0	5	44	j = INGENIERÍA ELECTRÓNICA
69	0	2126	32	160	0	8	53	937	409	253	0	0	0	0	20	k = INGENIERÍA ESTADÍSTICA E INFORMÁTICA
8	0	323	0	0	0	1	0	53	33	0	0	0	0	0	3	l = INGENIERÍA GEOLÓGICA
0	0	16	0	0	0	0	0	4	0	0	0	0	0	0	0	m = INGENIERÍA MECÁNICA ELÉCTRICA
12	0	333	0	3	0	4	6	83	44	0	0	0	0	0	4	n = INGENIERÍA METALÚRGICA
18	0	581	11	3	0	0	9	197	143	0	0	0	0	10	0	o = INGENIERÍA QUÍMICA
64	0	3987	8	19	0	10	56	2149	610	0	0	0	0	0	415	p = INGENIERÍA TOPOGRAF. Y AGRIMENSURA

La MATRIZ DE CONFUSION, es la prueba de validación del proceso de clasificación usando el método J48, el cual nos dice que, si la diagonal principal es mayor que sus extremos, se da por VALIDA el proceso, con la peculiaridad que en aquellos casos de valor CERO de la diagonal principal, implica que su participación en la plataforma es casi nula o estadísticamente NO SIGNIFICATIVA

#### 4.1.4 Fase 4: Minería de Datos (Data Mining)

- Determinamos el problema que necesitamos resolver, refiriéndonos al paradigma apropiado de Minería de Datos, conocer la asociación de las variables y encontrar patrones que nos orienten sobre el comportamiento de los datos.
- Empezamos a analizar la relación de las variables: escuelas profesionales, actividad y horas de trabajo.

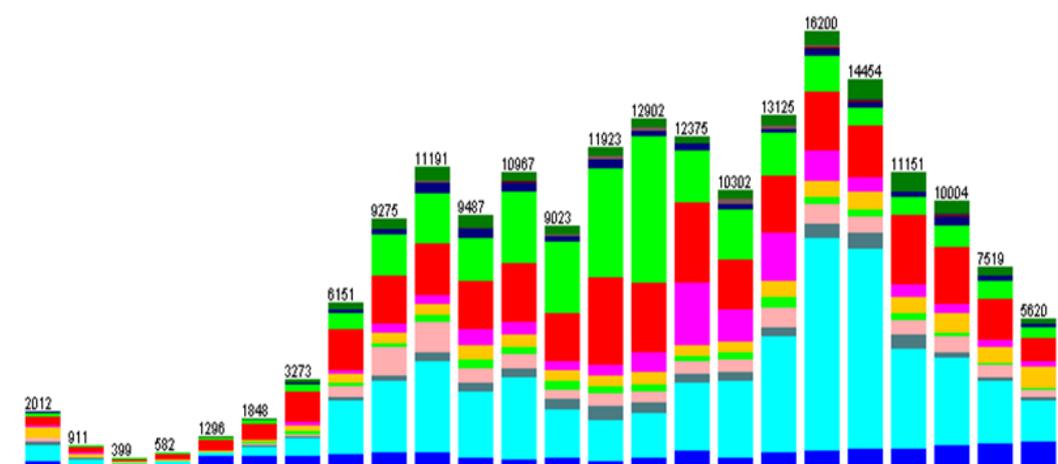


Figura 24. Actividad y horas de acceso al Sistema por Escuela

Resalta en este grafico que las horas que se tiene mayor ingreso a las aulas virtuales son las 19 horas teniendo la mayor actividad las siguientes escuelas: Ing. Agrícola e Ing. Metalúrgica, resaltando nuevamente las escuelas de Ing. de Sistemas, Ing. de Minas, Ing. Económica, Ing. Mecánica Eléctrica e Ing. Electrónica.

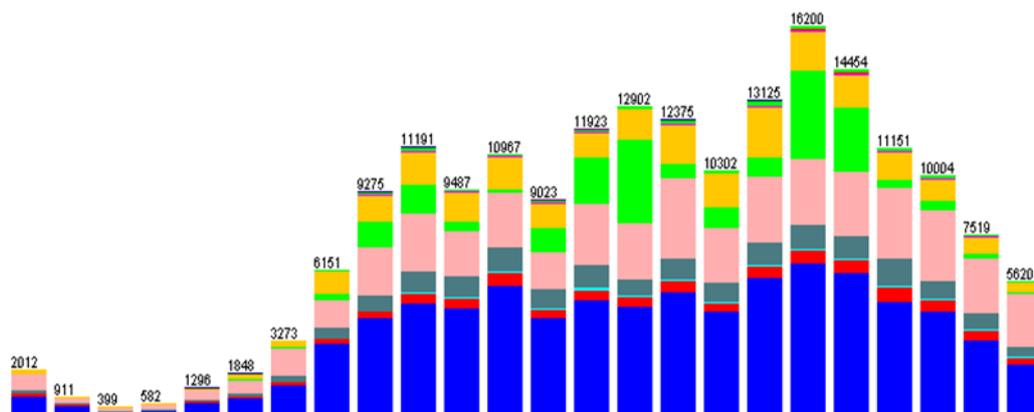


Figura 25. Horas de acceso al Sistema por Módulo

El modulo más visitado es el de curso (course) a las 19 horas, el módulo asignación de tareas (assign) a horas 16 , el módulo exámen (quiz) tiene mayor frecuencia tanto a las 15 y 19 horas, podemos mencionar también al módulo recursos (resource) a las 18 horas.

**a. Análisis de Asociación (Associate)**

Es uno de los procesos que la Minería de Datos nos brinda con la finalidad de conocer los patrones de comportamiento desconocidos.

En un proceso de análisis de datos, con una cantidad apreciable de registros, como es nuestro caso (191 990 registros) pese a los procesos y/o técnicas de análisis es imposible determinar ciertos comportamientos, para ello las técnicas de Minería de Datos nos ofrece algoritmos capaces de lograr esta cualidad.

Aplicamos el procedimiento de asociación (Associate) utilizando específicamente la técnica de asociación **A PRIORI**, mediante:



Figura 26. Procedimiento para la asociación A PRIORI

Obteniendo los siguientes resultados:

=== Run information ===

Scheme: weka.associations.Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1

Relation: toolscourses-weka.filters.unsupervised.attribute.Remove-R1-2,6-7-weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6

Instances: 191990

Attributes: 4

escuela

module

activity

hour

=== Associator model (full training set) ===

Apriori=====

Minimum support: 0.1 (19199 instances) (proporción mínima para inicio del entrenamiento)

Minimum metric <confidence>: 0.95 (Nivel de confianza mínima de los resultados alpha=0.05)

Number of cycles performed: 18 (número de entrenamientos)

Generated sets of large itemsets:

Size of set of large itemsets L(1): 7

Size of set of large itemsets L(2): 7

Size of set of large itemsets L(3): 1

Best rules found:

✓ escuela=INGENIERIA DE MINAS module=view ==>

escuela=INGENIERIA TOPOGRAFICA Y AGRIMENSURA

module=view <conf:(0.97)

El uso del módulo Vista por parte de la escuela profesional del Ingeniería de Minas, y la escuela profesional de Ingeniería Topografía y Agrimensura, presentan el mismo comportamiento en términos de actividad y hora, con un nivel de confianza del 97%.

✓ module=assign ==> activity=view <conf:(0.96)>

Sin considerar alguna escuela profesional, el uso del módulo Asignación de tarea y la actividad Vista, presentan el mismo comportamiento por todas aquellas escuelas que la usan y en cualquier hora, con un nivel de confiabilidad del 96%.

✓ escuela=INGENIERIA DE MINAS hour=ALL ==>

escuela=ARQUITECTURA Y URBANISMO hour=ALL ==>

escuela=INGENIERIA AGROINDUSTRIAL hour=ALL ==>

escuela=INGENIERIA ECONOMICA hour=ALL ==>

escuela=INGENIERIA ELECTRONICA hour=ALL conf:(0.98)

Otro de los patrones de comportamiento desconocidos, se presenta por las escuelas profesionales de Ingeniería de Minas, Agroindustrial, Económica, Electrónica y Arquitectura y Urbanismo, los cuales presentan un patrón de comportamiento similar en el uso de la plataforma, sin considerar una hora en particular. Con un nivel de confiabilidad del 98%.

Elegimos el algoritmo de minería de datos más adecuado.

### **b. Algoritmo K-Means**

Es un algoritmo que se utiliza para realizar agrupamientos, implementando la técnica de minería de datos, su objetivo es colocar todos los atributos en un espacio determinado dependiendo de las características formando grupos con rasgos similares pero diferentes a los demás.

Para nuestra investigación, el propósito es determinar y analizar el comportamiento mediante esta técnica de cada una de las escuelas profesionales en el uso de la plataforma virtual, sin embargo, con los pasos previos a este nivel de análisis, debemos resaltar que las escuelas que prevalecen su actividad en el uso de la plataforma, son Ingeniería de Minas e Ingeniería Económica entre las más importantes. Es por eso que los análisis ya discriminados propios de las técnicas estadísticas aplicadas, nos dan solo en función a esas escuelas.

El proceso de análisis se trabajará mediante el siguiente procedimiento:

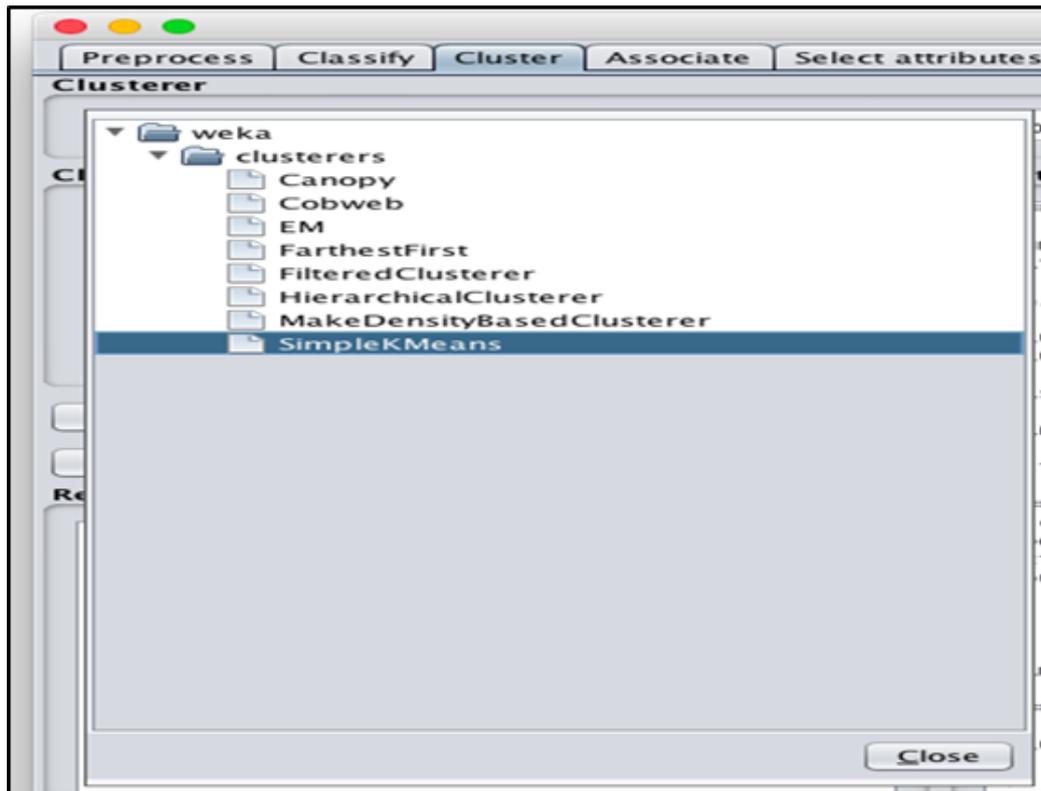


Figura 27. Proceso para el Análisis K-means



Obtenemos los siguientes resultados, que procedemos a analizar en interpretar.

```

=== Run information ===

Scheme:      weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning
10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-
last" -I 500 -num-slots 1 -S 10

Relation:    tools.curses-weka.filters.unsupervised.attribute.Remove-R1-2,6-7-
weka.filters.supervised.attribute.Discretize-Rfirst-last-precision6

Instances:   191990
Attributes:  4
             escuela
             module
             activity
             hour

Test mode:   evaluate on training data

=== Clustering model (full training set) ===
kMeans
=====
Number of iterations: 3
Within cluster sum of squared errors: 418310.0
Initial starting points (random):

```

Las figuras 27 y 28 muestran que de los 191 990 registros y trabajando con los 4 principales módulos y a la vez con las 2 escuelas que más presentan actividad se empieza al Modo de Prueba, procedemos con la evaluación en datos de entrenamiento, mediante el método K-means, obteniendo una suma de cuadrados de errores de 418 310 ya en la 3 iteración.

### c. Listado de Centroides por Escuela Profesional

Este proceso es inicial una vez que se determinó después de 3 iteraciones, el cual consiste en inicializar con los centroides, los cuales recayeron en los siguientes cluster (grupos)

**Cluster 0:** 'INGENIERIA ELECTRONICA', course, view,21

Este algoritmo elige al azar a la escuela profesional de Ing, Electrónica, módulo curso, actividad vista y a las 21 horas.

**Cluster 1:** 'ARQUITECTURA Y URBANISMO', assign, view,22

Del mismo modo para el segundo cluster elige a la escuela profesional de Arquitectura y Urbanismo, módulo asignación de tarea, actividad vista y a las 22 horas.

Missing values globally replaced with mean/mode

Tabla 4  
*Centroides del grupo final*

Attribute	Full Data (191 990.0)	Cluster #	
		0 (147168) 77%	1 (44822.0) 23%
<b>Escuela</b>	INGENIERÍA DE MINAS	INGENIERÍA DE MINAS	INGENIERÍA ECONÓMICA
<b>Module</b>	course	Course	Assign
<b>Activity</b>	View	View	View
<b>hour</b>	18	18	22

Luego de las 3 iteraciones, propias del método cluster k-means, se llegó a los siguientes resultados:

- Utilizando toda la data: (Full data: 191990), determinamos que la escuela profesional de Ingeniería de Minas es la que más actividad tiene en la plataforma virtual, con el uso del Módulo Curso(course), en la Actividad Vista(view) y la hora en que más realizan esta interacción es a las 18 horas.
- Respecto al clúster:
  - Clúster 0: Tomó 147168 registros (77%), el cual es predominado por la Escuela profesional del Ingeniería de Minas con el módulo course, actividad vista, todo esto realizado a las 18 horas.
  - Clúster 1: De los 191990, el algoritmo tomó 44822(23%) de la actividad registrada, determinada por la escuela profesional de Ing. Económica, con el modulo asignación de tarea, en la actividad vista y realizado a las 22 horas.
- Aclaremos también que se realizó pruebas con 3, 4 y 5 clúster, reafirmando los resultados del análisis de 2 clúster, sin embargo, cuanto más clúster se prueban, se observa que la actividad desagregada por módulo predomina en el uso de la plataforma la escuela profesional de Ing. de Minas.

#### **4.1.5 Fase 5: Interpretación y evaluación del conocimiento obtenido.**

En esta última fase se evalúa e interpreta el conocimiento.

##### **a. Selección de Atributos**

Con la finalidad de validar la selección que desde un inicio hizo los demás métodos de Minería de Datos, se procede de la siguiente manera:

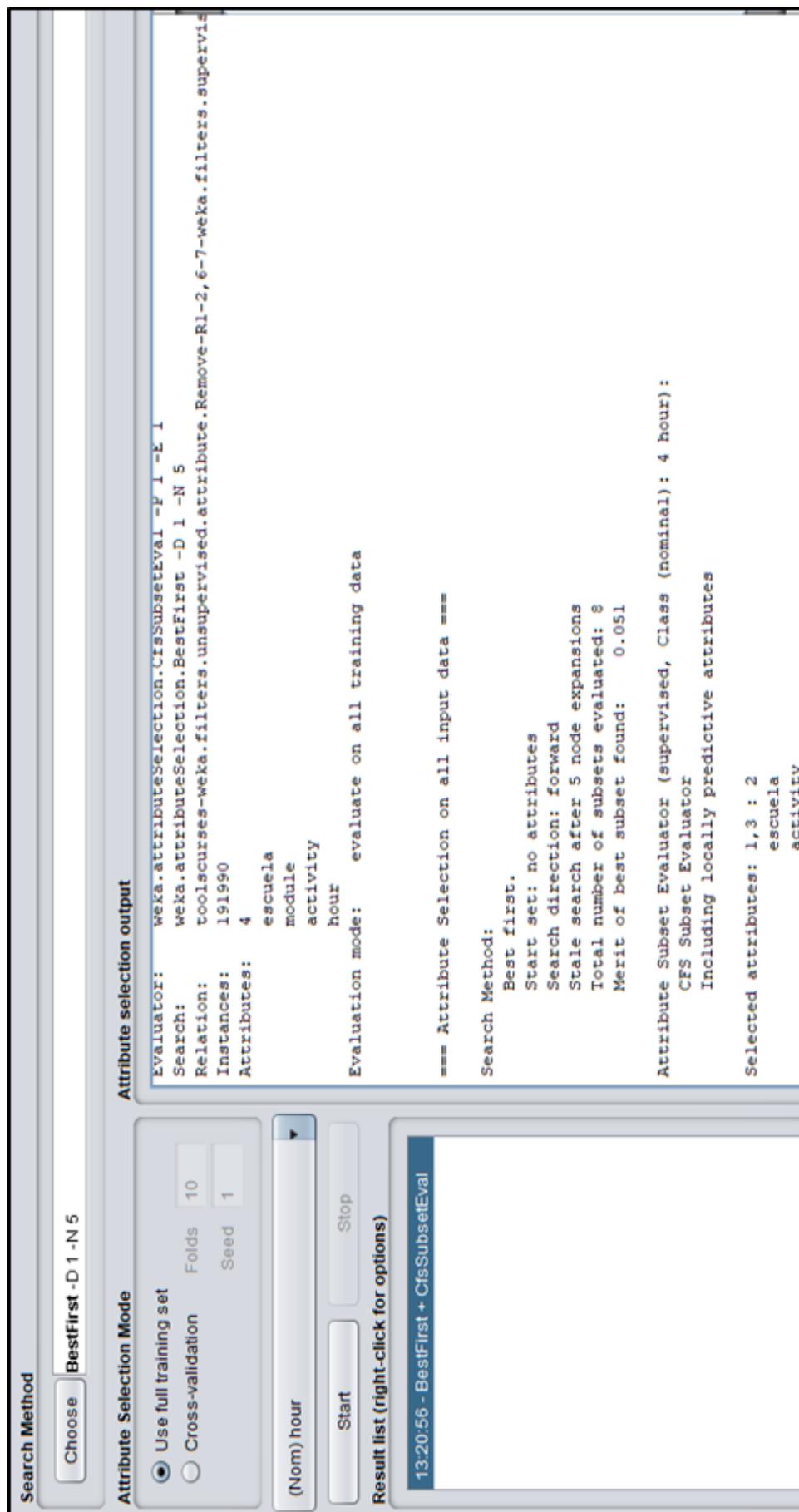


Figura 29. Selección de Atributos

Podemos verificar que, de las variables utilizadas y elegidas por los diferentes métodos de Minería de Datos, tanto por su nivel de influencia y uso frecuente en la plataforma, es la variable Actividad (activity), siendo la más importante e influyente en el uso y análisis de la plataforma virtual.

#### **a. Análisis Gráfico**

Finalmente, es necesario tener el comportamiento de la actividad en la plataforma mediante una representación gráfica, las cuales nos sirvieron para tener una idea real de los resultados propios de las técnicas aplicadas mediante la Minería de Datos.

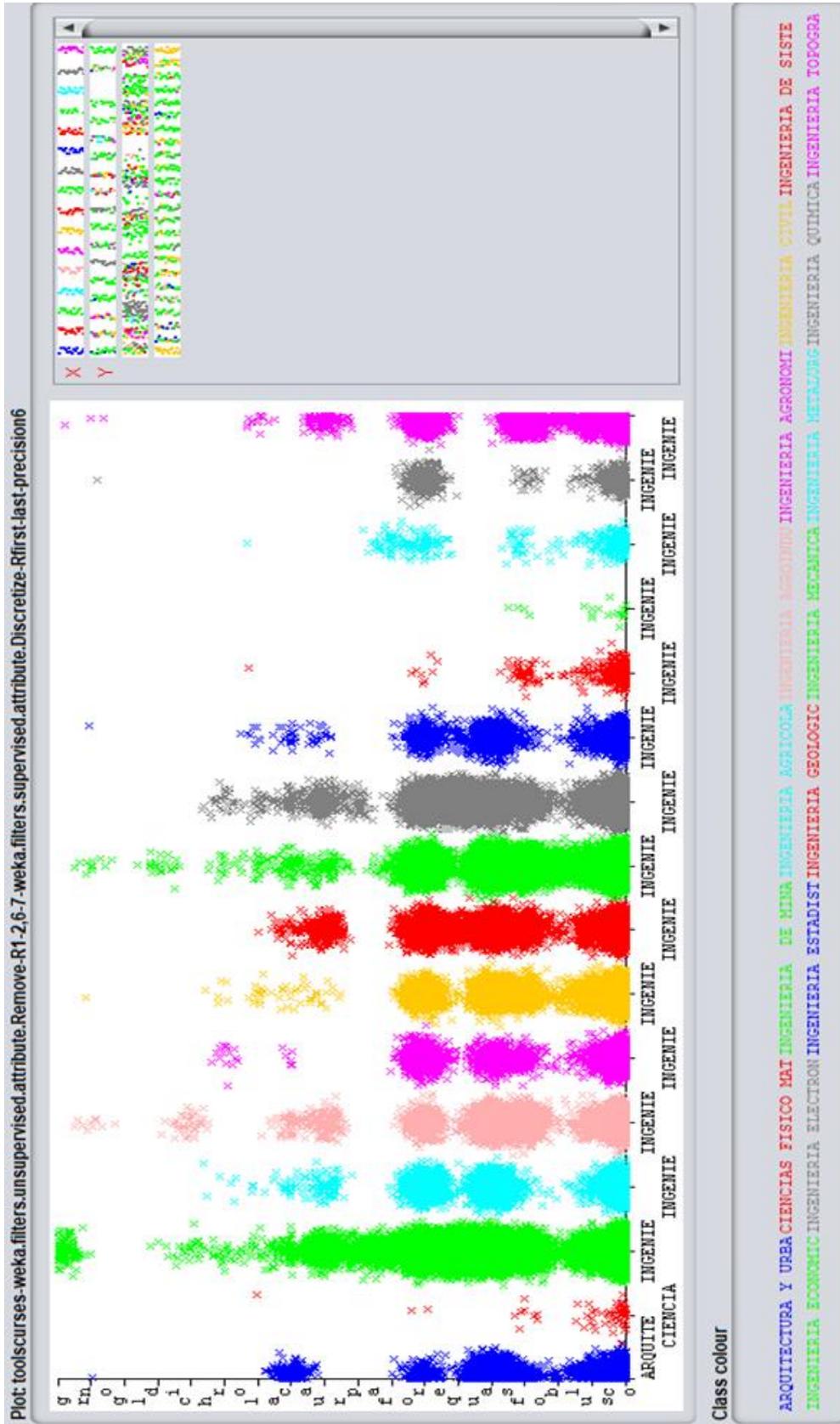


Figura 30. Relación entre Escuelas Profesionales y los Módulos

Al buscar la relación entre escuelas profesionales y la actividad con los módulos de la plataforma virtual encontramos las siguientes asociaciones:

- La escuela de Ingeniería de Minas es la que presenta la mayor actividad resaltando entre las otras escuelas, seguidamente de la escuela de Ingeniería Económica, Ingeniería Electrónica e Ingeniería de Sistemas.
- Mientras que las escuelas de Ciencias Físico Matemático, Ingeniería Geológica e Ingeniería Mecánica son en extremo pobre con su participación y/o uso de la plataforma. Se observa también a las escuelas de Ingeniería Metalúrgica e Ingeniería Química tratando de salir de este grupo.
- El resto de las escuelas tiene un promedio de uso de la plataforma virtual.

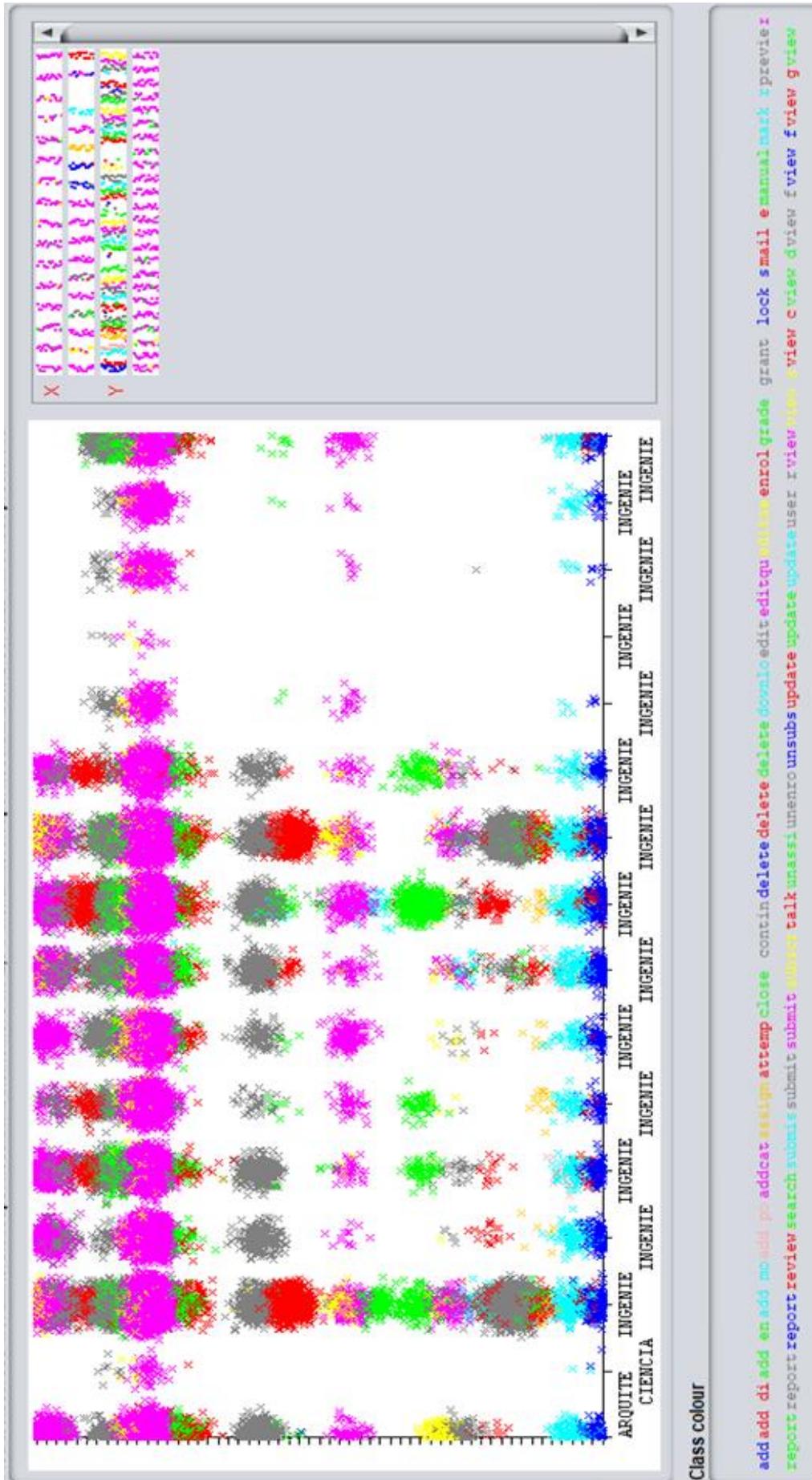


Figura 31. Relación entre las Escuelas Profesionales y la Actividad.

La figura 31 presenta una información muy interesante entre las escuelas profesionales y la actividad que realizan en la plataforma virtual, encontramos la siguiente asociación:

- Más allá de mostrarnos la actividad importante en todos sus módulos por las escuelas profesionales de Ing. de Minas, Ing. Economía, Ing. Electrónica, Ing. de Sistemas que son las que más resaltan entre otras, también consideramos que hay dos actividades que prevalece en casi todas las escuelas profesionales como son Add (adicionar o crear) y View (visualizar contenido), reforzando esto que los puntos de color azul (add) implica que para que se tenga actividad en la plataforma tiene que crear y/o adicionar contenido, pero lo interesante está en los puntos rosados que está presente en casi todas las escuelas y esto es por la actividad View (visualizar contenido), implica que la actividad de hacer una simple vista de la plataforma prevalece en relación a otras actividades que sí podrían mostrar la real interactividad con el contenido por parte de los usuarios-alumnos.



La figura 32 presenta la relación entre la escuela profesional y las horas de uso dentro de la plataforma virtual encontrando las siguientes asociaciones:

- Para validar el uso de la plataforma por parte de los alumnos de Ingenierías, éste gráfico en particular nos muestra la actividad de la plataforma y las horas de uso; resalta notar la poca presencia de uso de la plataforma entre las 00 horas y 05 horas, teniendo mayor acumulación de actividad entre las 09 horas hasta las 23 horas, especialmente en las escuelas que tuvieron presencia con mayor actividad; además la ausencia del uso de la plataforma por las escuelas como Ciencias Físico-Matemáticas, Ingeniería Metalurgia, Ingeniería Geológica e Ingeniería Química continua siendo casi nula su participación a cualquier hora del uso de la plataforma virtual.

#### 4.2 Discusión

En la investigación se coincide con Marulanda et al., (2017), debido a que tanto en nuestra investigación y la de los autores, la técnica de Minería de Datos implica, en sus primeras fases, limpieza, ajuste y transformación, de acuerdo a las necesidades del estudio. Estos procesos son necesarios porque, a pesar de todos los controles para la gestión de la información, los usuarios incurren permanente en errores, para los autores de digitación y para nuestro estudio en intentos fallidos o la mala ejecución del programa especialmente al momento de salir del sistema lo que hacen difícil la coincidencia de los datos.

Así mismo indican que su estudio de las pymes evaluadas se debe avanzar mucho más en el uso de buenas prácticas para la Gestión de Conocimiento, desarrollo de más y mejores competencias personales y organizacionales esto con la ayuda de técnicas para el tratamiento de datos masivos y poder extraer conclusiones relevantes que ayuden a tomar decisiones.

En el trabajo de investigación de García ( 2016), indica que la aplicación de técnicas de minería de datos en el campo educativo es un área de investigación de creciente interés,. El uso de estas técnicas cobra aún mayor relevancia en entornos de aprendizaje virtual como es la plataforma Moodle, donde se almacena información relativa a la actividad del estudiante y mediante técnicas de Minería de Datos se realiza la extracción de conocimientos que permita a los profesores mejorar el proceso de enseñanza –

aprendizaje en sus cursos virtuales, los cuales también fueron sustentados en nuestra investigación.

Se coincide con el autor en indicar que el proceso KDD por ser complejo debe ser implementado para personas que tengan conocimientos en Minería de Datos, por lo que recomienda obtener modelos en un entorno amigable para usuarios no expertos.

También hace referencia que el proceso de clasificación J48 es con el que se obtiene mejores resultados sobre los conjuntos de entrenamiento, el cual fue implementado en nuestra investigación.

Yamao (2018) Concluye que mediante técnicas estadísticas y minería de datos logra predecir el rendimiento académico de los estudiantes ingresantes a la Escuela Profesional de Ingeniería de Computación y Sistemas de la Universidad de San Martín de Porres. Pudo demostrar que, mediante patrones de comportamiento, identificó principales indicadores, extraída de la base de datos, como son: nota de examen de admisión, género, edad, modalidad de ingreso y distancia desde su casa hasta el centro de estudios., y obtuvo perfiles que ayudaron a la identificación de estudiantes que podrían encontrarse con dificultades en sus estudios. Se coincide con el autor, indicando que en las técnicas de Minería de Datos encontramos patrones de comportamiento y a la vez asociamos: escuelas, módulos, actividad y hora, como se trabajó en la investigación.

Melo (2018) Indica que con la aplicación de WEKA y el algoritmo J48 específicamente es ideal para estudios en la obtención del conocimiento, para su investigación, determinó las medicinas más demandadas en relación a las enfermedades más comunes por la que las personas acuden al establecimiento de salud Redes Puno. Se coincide con el autor que, en la obtención de conocimiento, la aplicación y el algoritmo ayudo sobre la similitud de comportamiento entre los módulos, escuelas, actividad y tiempo. Del mismo modo se hace énfasis que el proceso KDD permite la descripción de la información.

En el trabajo de investigación Holgado (2018) indica que con la aplicación de Minería de Datos mediante sus metodologías y algoritmos identificamos las variables más influyentes en un modelo de clasificación, con la cual se está de acuerdo y más aún cuando se indica que el algoritmo C5.0 es el que obtiene mejor desempeño en un modelo predictivo de clasificación, ya que para nuestro estudio se trabajó con el proceso de clasificación J48(C4.5) que es un antecesor de C5.0 con algunas mejoras en la cual

asociamos a las escuelas más relevantes con las variables módulos y horas, para encontrar patrones desconocidos.

Con el estudio de Ruiz y Romero (2018) coincidimos con los autores que al aplicar Minería de Datos mejora la calidad de información de la base de datos, nos ayuda a encontrar patrones ocultos para mejorar la gestión de la información y del conocimiento, aspectos que también fueron validados en la investigación presentada y gracias a las técnicas de Minería de Datos ya sea los arboles de decisión, Matriz de Correlación o la visualización grafica de los datos queda demostrado que en grandes volúmenes de información aplicar técnicas de Minería de Datos es lo más apropiado para las dos investigaciones.

Para el autor Flores (2016) la actualización constante, el contenido dinámico y la administración eficiente de la Plataforma logran mantener el interés de los estudiantes en la Unidad Didáctica, logrando fomentar su autoaprendizaje y aprendizaje colaborativo. Así mismo evidencia que la enseñanza virtual a través de la plataforma Moodle, favorece el desarrollo de las clases presenciales, actualizando los contenidos en la plataforma virtual, mantiene la atención y motivación de los estudiantes. Concluye su investigación en recomendar implementar la enseñanza virtual, con la plataforma Moodle, para medir el grado de incremento del aprendizaje de los estudiantes en comparación con otras escuelas. Se coincide con el autor debido, que nuestra investigación resalta la diferencia entre escuelas sobre la utilización del aula virtual, notando la diferencia entre las que tienen mayor actividad y aquellas en las que hay poca o casi no tienen presencia en el estudio.

## CONCLUSIONES

- ❖ Al identificar el comportamiento de las variables en el uso de aulas virtuales, resaltan los módulos Vista, Foro, Asignación de Tareas y Archivo, siendo la hora más frecuente las 17 horas, que son patrones cuyo comportamiento estadístico es **ALTAMENTE SIGNIFICATIVO**.
- ❖ En la investigación, mediante un análisis descriptivo se concluye que los estudiantes de Ingenierías presentan mayor actividad en las aulas virtuales, resaltando de entre las 16 escuelas Ing. de Minas e Ing. Económica, teniendo la mayor participación de sus estudiantes en la plataforma virtual, también destacan las escuelas de Ingeniería de Sistemas, e Ingeniería Mecánica Eléctrica.
- ❖ Al aplicar el método de clasificación J48 dentro del análisis de minería de datos y mediante la Matriz Confusion, se dio la validación a este proceso, el cual encontró la siguiente asociación de variables:
  - La escuela de Ingeniería de Minas al presentar la mayor actividad en aulas virtuales obtiene una mayor asociación en los siguientes módulos, actividad y horas: curso a las 00 horas, Foro a las 18 horas, Asignación de Tareas a horas 20:00, Examen y Archivo en distintas horas, pero solo el modulo Folder aparece en la escuela de Ing. de Minas.
  - El módulo Blog, solamente está presente con una notable actividad en la escuela de Ingeniería Económica.
  - La escuela de Ing. Económica tiene una participación activa en las aulas virtuales, resaltando en los módulos: Foro, Asignación de Tareas, Exámen, Archivo. Se tiene también la participación de las escuelas: Ing. de Sistemas, Ing.

Electrónica, Ing. Estadística e Informática, Arquitectura y Urbanismo, Ing. Agronómica, Ing. Topográfica y Agrimensura e Ing. Agroindustrial presentando esta vez dicha escuela una actividad fluida en el módulo Chat.

- Ing. de Sistemas, con el modulo direcciones URL, estando solo con esta actividad, ya que se ve muy escasa participación de otras escuelas.
- ❖ Al trabajar con la metodología KDD e implementar Minería de Datos en el programa WEKA y aplicar la técnica de asociación A PRIORI encontramos los siguientes patrones desconocidos:
- En el módulo Vista las escuelas de Ingeniería de Minas e Ingeniería Topográfica y Agrimensura, presentan un mismo comportamiento con una confiabilidad del 97%.
  - Las escuelas de Ingeniería de Minas, Agroindustrial, Económica, Electrónica y Arquitectura y Urbanismo, quienes presentan un comportamiento similar en términos de horas de uso, tiempo, uso de recursos y actividades con un nivel de confiabilidad del 98%.

## RECOMENDACIONES

- ❖ Al ser la Universidad Nacional del Altiplano una Universidad Licencia y con la convicción de presentar una educación de calidad se recomienda hacer cursos de capacitación con la finalidad de que las escuelas utilicen el aula virtual, debido que solo de las 16 escuelas tienen participación activa 10. Debiéndose promover o implementar la utilización de aulas virtuales, considerándose un requisito para sus docentes en su enseñanza – aprendizaje.
- ❖ Se recomienda investigar mediante Minería de Datos, comportamientos similares en las áreas de biomédicas y sociales con el fin de homogeneizar e interrelacionar criterios de capacitación en el uso de plataformas virtuales.
- ❖ Impulsar capacitaciones particularizadas en las diferentes áreas, por la presencia de patrones desconocidos en el comportamiento de uso de la plataforma virtual.

## BIBLIOGRAFÍA

- Barbosa, C. (2018). Concepto de KDD. *Minería de Datos - Información, (Concepto KDD)*. Recuperado de <http://mindatos.blogspot.com/2009/09/12-concepto-de-kdd.html>
- Benalcázar, J. B. (2017). *Análisis comparativo de metodologías de minería de datos y su aplicabilidad a la Industria de Servicios. (Tesis Maestría)*. Universidad de las Americas de Ecuador, Ecuador. Recuperado de <http://dspace.udla.edu.ec/bitstream/33000/7547/6/UDLA-EC-TMGSTI-2017-11.pdf>
- Calvo, D. (2018). *Cluster Aglomerativo divisivo*. Pandas Análisis y Modelamiento. Recuperado de <http://www.diegocalvo.es/cluster-jerarquicos-estrategia-aglomerativa-vs-divisiva/>
- Camargo, J.J., Camargo, J.F., y Joyanes, L. (2014). Conociendo Big Data. *Facultad de Ingeniería, vol.24 no.*(ISSN 0121-1129). Retrieved from [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0121-11292015000100006](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-11292015000100006)
- Cepeda, L. G. (2017). *Estrategia metodológica del uso de aulas virtuales en el proceso de enseñanza aprendizaje universitario*. (Tesis de doctorado), Universidad Nacional Mayor de San Marcos, Lima Perú.
- Córdoba, L. (2013). *WEKA* [mensaje de blog], Minería de Datos. Recuperado de <http://cor-mineriadatos.blogspot.com/2011/06/weka.html>
- De la Fuentes Fernandez, S. (2014). *Análisis de Conglomerados*. Universidad Autónoma de Madrid, Madrid, España. Recuperado de

- <http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/SEGMENTACION/CONGLOMERADOS/conglomerados.pdf>
- De la Rosa Rios, J. C. (2011). *Aplicación de la plataforma moodle para mejorar el rendimiento académico en la enseñanza de la asignatura de cultura de la calidad total en la Facultad de Administración de la Universidad del Callao*. (Tesis de maestría), Facultad de Educación. Universidad Nacional de San Marcos. Lima, Perú. Recuperado de [http://cybertesis.unmsm.edu.pe/bitstream/handle/cybertesis/2542/rosa\\_rj.pdf;jsessionid=8768E74C93A9E66F2366352BE632CB7C?sequence=1](http://cybertesis.unmsm.edu.pe/bitstream/handle/cybertesis/2542/rosa_rj.pdf;jsessionid=8768E74C93A9E66F2366352BE632CB7C?sequence=1)
- Erazo, S., y Molías, L. M. (2014). Procedimiento para la definición de Patrones de Uso docente en un Sistema de gestión de Aprendizaje *eaching Patterns Using in a Learning Management*. *Campus Virtuales*, III(2014), 88–95. Recuperado de <https://www.researchgate.net/publication/273575522> Procedimiento para la definicion de patrones de uso docente en un Sistema de Gestion de Aprendizaje.
- Figueras, M. S. (2001). Análisis de conglomerados o Cluster. España. Recuperado de [ciberconta.unizar.es/LECCION/cluster/inicio.html](http://ciberconta.unizar.es/LECCION/cluster/inicio.html)
- Flores, E. F. (2016). “ *Influencia de la plataforma Moodle en el rendimiento academco de los estudiantes del curso de precalculo I de la Universidad Continenetal*.” (Tesis de maestría). Universidad Nacional del Centro del Perú. Facultad de Ingeniería de Sitemas, Huancayo, Perú.
- Garcia, D. (2016). *Minería de datos aplicada a la enseñanza virtual: nuevas propuestas para la construcción de modelos y su integración en un entorno amigable para el usuario no experto*. (Tesis de maestría). Departamento de Ingeniería Informática y Electrónica. Universidad de Cantabria, Cantabria, España.
- Gardey, A., Perez, J. (2017). Definición de plataforma virtual - Qué es, Significado y Concepto. Recuperado de <https://definicion.de/plataforma-virtual/>
- Gómez, M. E., Contreras, L., y Gutiérrez, D. (2016). El impacto de las tecnologías de la información y la comunicación en estudiantes de ciencias sociales: un estudio comparativo de dos universidades públicas. *Innovación Educativa*, vol.16 no. Recuperado de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1665-26732016000200061](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-26732016000200061)

- González, S. (2019). Revision de plataformas de entorno de aprendizaje. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699. <https://doi.org/10.1017/CBO9781107415324.004>
- Gutiérrez, J. J. (2017). *Propuesta de una metodología de extracción de conocimientos a partir de datos de las prestaciones del seguro integral del salud en la region Piura en el año 2016.* (Tesis de maestría). Facultad de Ingeniería. Escuela Profesional de Ingeniería de Sistemas. Universidad Católica Los Angeles de Chimbote. Piura, Perú.
- Hernandez, E. J. (2016). *Aplicación de técnicas de analisis de datos y administracion de Big Data ambientales*, (Tesis de maestría). Ingeniería Administrativa. Universidad Nacional de Colombia Sede Medellin. Medellin, Colombia.
- Herrera, F., y Cano, J. R. (2006). Técnicas de reducción de datos en KDD. El uso de Algoritmos Evolutivos para la Selección de Instancias. *Actas del I Seminario Sobre Sistemas Inteligentes (SSI'06), Universidad Rey Juan Carlos, Madrid (Spain).*, 165–181.
- Holgado, L. A. (2018). *Detección de patrones de bajo rendimiento académico mediante técnicas de minería de datos de los estudiantes de la Universidad Nacional Amazónica de Madre de Dios 2018.* (Tesis de maestría). Universidad Nacional del Altiplano. Puno, Perú.
- Laura, L., Rosas, K., & Esquicha, J. (2017). Estudio comparativo de Técnicas no Supervisadas de Minería de Datos para Segmentación de Alumno. Universidad Católica Santa María, Arequipa, Perú.
- Lopez, J. (2018). Minería de Datos. Economipedia. Madrid, España. Recuperado de <https://economipedia.com/definiciones/mineria de datos.html>
- Martínez, C. A. (2012). *Aplicacion de Técnicas de Minería de Datos.* ( Tesis de maestría). Facultad de Ciencias Físicas y Matemáticas, Departamento de Ingeniería Industrial. Universidad de Chile. Santiago de Chile.

- Marulanda, E.C., Lopez, M., y Mejia, M. (2017). Minería de datos en gestión del conocimiento de pymes de Colombia. *Universidad Catolica Del Norte*, 224–227. Recuperado de <http://www.redalyc.org/pdf/1942/194250865013.pdf>
- Melo, A. D. (2018). *Patrones para la estimación de consumo de medicamentos con minería de datos Redes Puno*. (Tesis de maestría). Universidad Nacional del Altiplano. Puno, Perú.
- Merayo, P. (2018). ¿Qué es la plataforma Moodle y para qué sirve. [mensaje en un blog].\_ Máxima Formación. Recuperado de <https://www.maximaformacion.es/e-learn/que-es-moodle-y-para-que-sirve/>
- Mosquera, R., Parra, L., y Castrillón, O. D. (2016, ). Metodología para la Predicción del Grado de Riesgo Psicosocial en Docentes de Colegios Colombianos utilizando Técnicas de Minería de Datos. *Información Tecnológica*, [on line] 27(6), 259–272. ISSN 0718-0764. <https://doi.org/10.4067/S0718-07642016000600026>
- Pedraza, A. C. (2018). Ventajas del uso de la Minería de Datos. *Gestiopolis*. Recuperado de <https://www.gestiopolis.com/ventajas-del-uso-de-la-mineria-de-datos/>
- Perez, J. y Merino, María. (2017). Definición de aula virtual - Qué es, Significado y Concepto. *Definicion.de*. Recuperado de <https://definicion.de/aula-virtual/>
- Ribas, E. (2018). *Que es Data Mining*. Barcelona, España. Recuperado de <https://www.iebschool.com/blog/data-mining-mineria-datos-big-data/#>
- Rodriguez O. (2016). Conceptos Básicos. *Consultoria Minería de Datos*. Costa Rica. Recuperado de [http://oldemarrodriguez.com/yahoo\\_site\\_admin/assets/docs/Presentación\\_-\\_Conceptos\\_Básicos.41132532.pdf](http://oldemarrodriguez.com/yahoo_site_admin/assets/docs/Presentación_-_Conceptos_Básicos.41132532.pdf)
- Ruiz, E.M., y Romero, C. P. (2018, ). Resultados obtenidos en un proceso de minería de datos aplicado a una base de datos que contiene información bibliográfica referida a cuatro segmentos de la ciencia. *Journal of Information Systems and Technology Management – Jistem . USP*. Sao Paulo, Brasil. 15, 11. <https://doi.org/10.4301/S1807-1775201815003>

- Samaniego, G., Marqués, L. & Gisbert, M.(2014). *Procedimiento para la definición de Patrones de Uso docente en un Sistema de Gestión de Aprendizaje*. Campus Virtuales, Vol. III Num 2, Revista campus virtuales, pp.88–95.
- Schiaffino, S. (2018). Inteligencia Artificial: Clustering. Argentina. Recuperado de [http://www.exa.unicen.edu.ar/catedras/optia/public\\_html/clustering.pdf](http://www.exa.unicen.edu.ar/catedras/optia/public_html/clustering.pdf)
- Sulla, J. A. (2015). *Aplicación de técnicas supervisadas de minería de datos para determinar la predicción de deserción académica*. (Tesis de segunda especialidad). Facultad de Ciencias e Ingenierías Físicas y Formales.Universidad Catolica de Santa Maria. Arequipa, Perú.
- Terrádez, M. (2017). Análisis de conglomerados. Recuperado de <https://docplayer.es/21470240-Analisis-de-conglomerados.html>
- Timarán, S. R., Hernández, I., Caicedo, S. J., Hidalgo, A., y Alvarado, J. C. (2016). Descubrimiento de patrones de desempeño académico con árboles de decisión en las competencias genéricas de la formación profesional. *Ediciones Universitarias*. Bogotá, Colombia. 63–86. <https://doi.org/10.16925/9789587600490>
- Yamao Eiriku. (2018). *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*. (Tesis de maestría) Lima, Perú.