

Adicción a las Redes Sociales en Tiempos de Covid: Un Enfoque desde la Minería de Datos

Addiction to Social Networks in times of Covid: an approach from data mining

*Felipe de Jesús Núñez Cárdenas^a, Ana María Felipe-Redondo^b, Mayte Guerrero-Ruiz^c,
Abigail Rivera-Hernandez^d, Victor T Tomas-Mariano^e*

Abstract:

In current times of confinement of the population as a result of the COVID pandemic, adolescents have become more vulnerable in the use of different applications on the web, among the most used are social networks. The present project aims to know the degree of addiction to social networks of high school students of the Higher School of Huejutla in times of COVID, understanding this with the levels of high, moderate and without addiction, For this the algorithm was used Kmeans within the Clustering technique. This algorithm was modeled in the WEKA tool, the methodology used is the so-called CRISP-DM. Data collection was carried out with the application of the Addiction Risk Scale Test in Social Networks (ERA-RSI). The results obtained are shown in three different clusters, one oriented to discharge, the second to moderate and a last to the population without addiction.

Keywords:

ERA RSI test, Kmeans algorithm, Datamining

Resumen:

En tiempos actuales del confinamiento de la población como consecuencia de la pandemia de COVID, ha vuelto a los adolescentes más vulnerables en el uso de diferentes aplicaciones en la web, dentro de los más utilizados se encuentran las redes sociales. El presente proyecto tiene como objetivo conocer el grado de adicción a las redes sociales de los alumnos de bachillerato de la Escuela Superior de Huejutla en tiempos de COVID, entendiéndose ésta con los niveles de alta, moderada y sin adicción, Para ello se utilizó el algoritmo Kmeans dentro de la técnica de Agrupamiento. Este algoritmo fue modelado en la herramienta WEKA, la metodología empleada es la denominada CRISP-DM. La recogida de datos se llevó a cabo con la aplicación del Test Escala de Riesgos de Adicción en las Redes Sociales (ERA-RSI). Los resultados obtenidos se muestran en tres diferentes cluster, uno orientado a la alta, el segundo a la moderada y un último a la población sin adicción.

Palabras Clave:

Test ERA RSI , Algoritmo kmeans, Minería de Datos

Introducción

En la Actualidad a causa de la Pandemia Mundial ocasionada por el COVID se han modificado muchos procesos con la finalidad de evitar el contagio de este virus. Algunos de ellos se han visto afectados de tal

manera que se ha optado por realizarlos desde la casa. Uno de ellos es el Educativo el cual se lleva a cabo de manera virtual. Dentro de la Escuela Superior Huejutla no ha sido la excepción, sin embargo, esta cuarentena en el hogar ha ocasionado que se disponga de la tecnología para aminorar este aislamiento.

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

^b Universidad Tecnológica de la Huasteca Hidalguense. ORCID: <https://orcid.org/0000-0002-8579-6532> Email: ana.felipe@uthh.edu.mx

^c Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Huejutla, email: gu359600@uaeh.edu.mx

^d Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Huejutla, email: ri362160@uaeh.edu.mx

^e Universidad Autónoma del Estado de Hidalgo, Escuela Superior de Huejutla, ORCID: <https://orcid.org/0000-0001-6623-860X>, email: victor_tomas@uaeh.edu.mx

Este trabajo se enfoca sobre la utilización de herramientas de minería de datos aplicando una técnica de agrupamiento y el algoritmo k-means, para determinar el nivel de dependencia a las redes sociales en los alumnos de bachillerato de la Escuela Superior Huejutla, usando como herramienta de evaluación el test ERA RSI [1], el cual nos puede diagnosticar el nivel de dependencia o adicción a este recurso.

Por otro lado, la minería de datos es el conjunto de técnicas de análisis que nos ayuda a predecir y describir el comportamiento de los datos. También extraer patrones, tendencia y regularidades [2].

Bajo este contexto se pretende diagnosticar bajo estas técnicas, el grado de adicción de los alumnos a dichas plataformas.

Adicción a las Redes Sociales

Las redes sociales como Instagram, Facebook, Twitter y los sistemas de mensajería como WhatsApp son atractivos para los jóvenes porque su funcionamiento implica la interactividad inmediata. Su uso es de carácter social para comunicarse en un ámbito popular, pero esto sin un control educativo puede llegar a trastornar la realidad de algunos, sobre todo cuando hablamos de jóvenes adolescentes. La adicción a las redes sociales se vincula a la mayoría de jóvenes que tienen carencias emocionales y acaban sufriendolas mediante las redes sociales o perfiles irreales.

Su uso nació bajo principios positivos, pero el problema ocurre cuando se detecta que se dejan de lado el resto de las obligaciones propias de una vida social normal. (estudiar o trabajar, hacer deporte u otras aficiones, salir con amigos o relacionarse con la familia). Además, los últimos estudios han detectado que el abuso de las redes sociales provoca alejamiento del entorno de la vida real, provoca niveles altos de ansiedad, afecta a la autoestima y puede acabar causando la pérdida de la capacidad de autocontrol [3].

Hay personas que son más vulnerables que otras, por lo que los perfiles de los adictos a las redes sociales son bastante determinados. Si bien la mayoría de redes sociales están disponibles para todo el mundo, solo una pequeña parte pueden presentar problemas de abuso y adicción. Generalmente, los adolescentes constituyen un potencial colectivo de riesgo porque son los que más se conectan a Internet, más horas pasan en línea y además, también están más familiarizados con las nuevas tecnologías actuales.

Existen ciertas características de personalidad o estados emocionales que tienden a aumentar la vulnerabilidad psicológica del adicto. El Dr. Fábregas, psiquiatra especialista en adicciones del centro CITA, explica que cuando hablamos de adicciones a redes sociales suele haber un problema subyacente que debe

tratarse psicológicamente también. La timidez excesiva, baja autoestima, rechazo de la imagen corporal, depresión o hiperactividad son solo algunos ejemplos que pueden manifestarse tras la adicción.

En estos casos el abuso de Internet es como una cortina de humo que evidencia la existencia de un fuego para controlar, un problema de personalidad o un trastorno mental, que es lo que habría que abordar en realidad. Hay otros casos en los que se trata a personas que muestran una insatisfacción personal con su vida y que intentan ocultar esas carencias familiares o sociales. En estos casos Internet o las redes sociales actúan como una prótesis tecnológica que facilita el escape de la vida real, aunque a largo plazo tenga consecuencias negativas [4].

Agrupamiento: Algoritmo Kmeans

Los algoritmos de clustering son considerados de aprendizaje no supervisado. Este tipo de algoritmos de aprendizaje no supervisado busca patrones en los datos sin tener una predicción específica como objetivo (no hay variable dependiente). En lugar de tener una salida, los datos solo tienen una entrada que serían las múltiples variables que describen los datos. K-means necesita como dato de entrada el número de grupos en los que vamos a segmentar la población. A partir de este número k de clusters, el algoritmo coloca primero k puntos aleatorios (centroides). Luego asigna a cualquiera de esos puntos todas las muestras con las distancias más pequeñas.

A continuación, el punto se desplaza a la media de las muestras más cercanas. Esto generará una nueva asignación de muestras, ya que algunas muestras están ahora más cerca de otro centroide. Este proceso se repite de forma iterativa y los grupos se van ajustando hasta que la asignación no cambia más moviendo los puntos. Este resultado final representa el ajuste que maximiza la distancia entre los distintos grupos y minimiza la distancia intragrupo.

El algoritmo k-Means sigue los siguientes pasos:

- Inicialización: se elige la localización de los centroides de los K grupos aleatoriamente
- Asignación: se asigna cada dato al centroide más cercano
- Actualización: se actualiza la posición del centroide a la media aritmética de las posiciones de los datos asignados al grupo.

Los pasos de asignación y actualización se siguen interactivamente hasta que no haya más cambios [5].

Adicción a las Redes Sociales Aplicando Minería de Datos

En el año 2015 Roció Leticia Cortes Campos, Alfredo Zapata Gómez, Víctor Hugo Menéndez y Pedro José Canto Herrera en su trabajo de investigación titulado el estudio de los hábitos de conexión en redes sociales virtuales, para ello utilizaron la minería de datos, su propósito fue presentar una visión general de los hábitos de conexión a las redes sociales virtuales entre los estudiantes de pregrado. La información se obtuvo por medio de un cuestionario, aplicando a 458 alumnos de nivel licenciatura de la universidad Autónoma de Yucatán, México. El análisis y la interpretación se realizaron utilizando técnicas de minería de datos, mediante las cuales se obtuvieron grupos de estudiantes definidos por el género, la frecuencia de conexión y las redes sociales que consultan. También se extrajeron reglas de asociación que sugieren importantes condiciones socioeconómicas y laborales, sobre todo con respecto a los lugares y a los dispositivos utilizados por los estudiantes el momento de conectarse. [6]

Por otro lado, en el año 2015 Elena Asensio Blasco, en su trabajo de investigación titulado Aplicación de minería de datos en redes sociales/web, en el cual se pudo obtener información de las redes sociales, como Twitter. Su análisis correcto proporciona un valor adicional a la recuperación de información. Se basa en la realización de una aplicación que gestione y obtenga datos sobre el uso de los lenguajes oficiales de las comunidades autonómicas en Twitter. [7]

Metodología

El proceso del presente trabajo se llevó a cabo bajo la metodología CRISP-DM ya que contempla el ciclo de vida de un Proyecto de minería de datos [8].

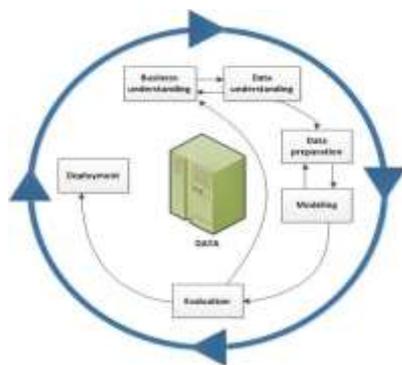


figura 1 Ciclo de Vida de un Proyecto de Minería de Datos, metodología CRISP-DM

Desarrollo

Compresión del Caso de Estudio

La Escuela Superior de Huejutla de la Universidad Autónoma del Estado de Hidalgo actualmente cuenta con

los programas educativos de, Ciencias Computacionales, Derecho, Administración, Medicina y Enfermería en el nivel superior, además de ofertar el bachillerato General, es en esta sección donde se aplicó este proyecto. En promedio cuenta con una matrícula de 550 alumnos.

A partir de marzo del año 2020 y como consecuencia de la Pandemia del COVID 19, se tomó la determinación que las clases se llevaran a cabo de manera virtual por medio de plataformas tecnológicas. Por lo que el contexto se sitúa en educación a distancia bajo esta primicia. Este nivel de estudios de la ESH tiene dos horarios uno matutino y uno vespertino en donde los horarios de atención son de 7 am – 14:00 para el matutino y de 12:00 pm a 19:00 pm para el vespertino.

Entendimiento de los Datos

Como instrumento para la recolección de la información fue la encuesta de Escala de riesgo de adicción-adolescente a las redes sociales e internet (ERA-RSI) [9], el cual consta de 29 ítems, dicha encuesta se elaboró en Google Forms y se difundió por las redes sociales.

Encuesta de Riesgo de Adicción-Adolescente a las Redes Sociales e Internet (ERA-RSI)

Algunas personas de tu edad se conectan a internet o través de teléfonos inteligentes smartphones, iPhone, tablets, portátiles o otros sistemas digitales. Pasa prácticamente todo el día conectado por diversión o por los deberes, la diversión, la seguridad, se conectan entre amigos, etc. Sin embargo, en ocasiones puede crear ansiedad por la conexión a internet o de las redes sociales. Sin una parte del cuestionario respondiendo sobre el consumo a internet o redes sociales y cuál es el uso que realizan de ellas. Es importante que contestes a todas las afirmaciones que tienes a continuación y que lo hagas con sinceridad, marcando la respuesta que consideres adecuada de acuerdo a como te sientas.

1 = Nunca o casi nunca; 2 = Algunas vez o pocas; 3 = Bastante veces o bastante; 4 = Siempre o Mucha

Ítem	1	2	3	4
1. ¿Conectas en internet todas las veces que tienes un Wi-Fi o Smartphone? (marcando el tiempo que pasó) (marcando en ellas la)				
1.1. Al menos de 1 hora (2) Una 2 horas (3) Entre 3 y 4 horas (4) Más de 4 horas				
2. ¿Tienes un Wi-Fi en tus horas de estudio y/o trabajo?				
3. ¿Algunas veces cuando estás en internet o cuando estás conectado de las Wi-Fi				
4. ¿Acabas o las Wi-Fi cuando estás leyendo o estudiando?				
5. ¿Cada vez que estás en las Wi-Fi las usas para ver videos o entretenimiento?				
6. ¿Te gustaría tener un móvil que conectarse a las redes sociales y por los deberes?				
7. ¿Cuándo estás en un momento que estás conectado a las Wi-Fi				
8. Si estás conectado a internet en tus horas de estudio				
9. ¿Algunas veces cuando estás conectado a las Wi-Fi				
Total puntaje de adicción				
10. ¿Qué es más habitual que realizar de las redes sociales en tu tiempo libre?				
11. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
12. ¿Qué redes sociales usas más?				
13. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
14. ¿Has usado alguna vez las redes sociales en tu tiempo libre?				
15. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
16. ¿El tiempo de las redes sociales que utilizas en las Wi-Fi o internet es:				
17. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
18. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
Total puntaje de las Wi-Fi (Riesgo Alto)				
19. ¿Qué es más habitual que realizar de las redes sociales en tu tiempo libre?				
20. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
21. ¿Qué redes sociales usas más?				
22. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
23. ¿Has usado alguna vez las redes sociales en tu tiempo libre?				
24. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
25. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
26. ¿El tiempo de las redes sociales que utilizas en las Wi-Fi o internet es:				
27. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
28. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
29. ¿Cuánto tiempo usas las redes sociales en tu tiempo libre?				
Total puntaje de adicción				

figura 2 Encuesta ERA RSI

Esta encuesta muestra como resultado tres niveles de adicción: Alta, Moderada y Sin Adicción. Y se trabaja bajo cuatro esquemas de determinación sobre estos niveles, es importante comentar que en el uso del instrumento de la encuesta ERA-RSI Podemos encontrar o se agrupan los ítem en cuatro factores que originan el riesgo, estos

son síntomas-adicción, uso-social, rasgos-frikis y nomofobia.

El factor síntomas-adicción checa conductas de adicción a sustancias no tóxicas, el factor uso-social evalúa la socialización virtual de un adolescente, el factor rasgos-frikis recoge aspectos propios como reunirse con personas afines a algo y el factor nomofobia están relacionados a la ansiedad y control del uso del móvil [2].

Preparación de los Datos

Una vez obtenido los resultados de la encuesta, se preparan los datos iniciales para poder usarlos, para lo cual se puede utilizar la aplicación de Bloc de notas para su edición y se guarda el archivo con la extensión .arff, para así utilizarlos en la herramienta de explotación Weka.

```

Proyecto.arff: Bloc de notas
Archivo Edición Formato Ver Ayuda
relation Proyecto
@attribute P1{1hrs, 2hrs, 3hrs, 4hrs}
@attribute P2{Nada, Poco, Bastantes, Mucho}
@attribute P3{No, Poco, Bastante, Mucho}
@attribute P4{Nada, Poco, Bastantes, Mucho}
@attribute P5{Nada, Poco, Bastantes, Mucho}
@attribute P6{Nada, Poco, Bastantes, Mucho}
@attribute P7{Nada, Poco, Bastantes, Mucho}
@attribute P8{Nada, Poco, Bastantes, Mucho}
@attribute P9{Nada, Poco, Bastantes, Mucho}
@attribute P10{Nada, Poco, Bastantes, Mucho}
@attribute P11{Nada, Poco, Bastantes, Mucho}
@attribute P12{Nada, Poco, Bastantes, Mucho}
@attribute P13{Nada, Poco, Bastantes, Mucho}
@attribute P14{Nada, Poco, Bastantes, Mucho}
@attribute P15{Nada, Poco, Bastantes, Mucho}
@attribute P16{Men100, 101a1000, 1001a3000, Mas3000}
@attribute P17{Nada, Poco, Bastantes, Mucho}
@attribute P18{Nada, Poco, Bastantes, Mucho}
@attribute P19{Nada, Poco, Bastantes, Mucho}
@attribute P20{Nada, Poco, Bastantes, Mucho}
@attribute P21{Nada, Poco, Bastantes, Mucho}
@attribute P22{Nada, Poco, Bastantes, Mucho}
@attribute P23{Nada, Poco, Bastantes, Mucho}
    
```

figura 3 preparación de datos en archivo .arff

Modelado

En esta etapa, se aplicó la técnica descriptiva de minería de datos, aplicando la herramienta Weka. Dentro de esta herramienta se utilizó la técnica de Agrupamiento en donde se optó por la creación de tres cluster, esto debido a la clasificación de la encuesta en sus tres niveles de adicción: alta, moderada y sin adicción. El algoritmo utilizado para tal fin es el Simple Kmeans.

```

Final cluster centroids:
Attribute  Full Data  Cluster#
              (58.0)  (13.0)  (8.0)  (37.0)
-----
P1          2hrs    2hrs    4hrs    2hrs
P2          Bastantes  Poco  Bastantes  Bastantes
P3          Poco    Poco    Poco    No
P4          Poco    Poco    Poco    Poco
P5          Poco    Poco    Mucho    Poco
P6          Poco    Poco    Mucho    Poco
P7          Nada    Poco    Poco    Nada
P8          Nada  Bastantes  Nada    Nada
P9          Poco  Bastantes  Mucho    Poco
P10         Poco    Poco    Poco    Nada
P11         Poco  Bastantes  Poco    Poco
P12         Poco  Bastantes  Poco    Poco
P13         Poco  Bastantes  Mucho    Poco
P14         Poco    Poco  Bastantes  Nada
P15         Poco    Poco    Nada    Poco
P16         Men100  101a1000  Men100  Men100
P17         Poco    Poco  Bastantes  Poco
P18         Poco    Poco  Bastantes  Poco
P19         Poco    Poco    Poco    Nada
P20         Nada  Bastantes  Mucho    Nada
P21         Nada    Poco    Poco    Nada
P22         Poco    Poco  Bastantes  Poco
P23         Nada    Poco    Mucho    Nada

Time taken to build model (full training data): 0.09 seconds
=== Model and evaluation on training set ===
Clustered Instances
0      13 (.22%)
1       8 (.14%)
2      37 (.64%)
    
```

figura 4 Resultados del Modelo Ejecutado

Resultados y Discusion

Para la elaboración del modelo K-Means, con los 58 registros que corresponden al total de alumnos que contestaron la encuesta, los resultados que se obtuvieron fueron a través de tres cluster:

- Cluster 0: Representa una adicción moderada a las redes sociales que obtuvieron 13 de 58 que es equivalente a 22%.
- Cluster 1: Representa una adicción alta a las redes sociales que obtuvieron 8 de 58 que es equivalente a 14%.
- Cluster 2: Representa no tener una adicción a las redes sociales que obtuvieron 37 de 58 que es equivalente a 64%.

las principales variables con que se obtuvieron estos resultados son el tiempo de uso, el acompañamiento virtual, la frecuencia de conexión, consulta de perfiles de amigos o conocidos, el uso más habitual para usar las redes sociales, entre otras, todas ellas referenciadas en los cuatro apartados del instrumento ERA -RSI. Podemos visualizar estas características en la figura 4.

CLUSTER	POBLACION	%
0	13	22%
1	8	14%
2	37	64%
TOTAL	58	100%

Tabla 1 resultados a nivel de cluster del modelado

Es importante comentar que este Proyecto solo describe el indicador total de la adicción a las redes sociales sin adentrarse a los diferentes factores síntomas-adicción, uso-social, rasgos-frikis y nomofobia. Con la intención de poder concluir un grado de adicción dentro de los diferentes grupos.

Conclusiones y Trabajos Futuros

Como conclusión de la aplicación del cuestionario ERA-RSI se obtuvo información importante sobre alumnos de bachillerato de la UAEH ESH, se aplicó el algoritmo K-Means para la agrupación de los datos, se eligieron 3 cluster, debido a que el ERA-RSI consta de 3 posibles diagnósticos que pueden dar de acuerdo con el puntaje obtenido en la evaluación. En base a los mismo Podemos decir que el 64% de la población no representa ningún síntoma de adicción a las redes sociales, mientras que el 22% presenta una adicción moderada a las mismas y el 14% presenta un alto grado de adicción a las redes sociales. Podemos a través de estos resultados el cual arroja la atención al 36% de la población estudiantil de bachillerato implementar a través de técnicas de asociación y algoritmos como A priori encontrar correlaciones entre el desempeño académico y estos grados de adicción encontrados.

Referencias

- [1] Florez, B. R., González, R. M., & López, M. B. Valoración del riesgo de adicción a redes sociales e internet en adolescentes de la zona básica de salud de Mondéjar. CONDUCTAS DE RIESGO, 119.
- [2] Orallo, J. H., Quintana, M. J. R., & Ramírez, C. F. (2004). Introducción a la Minería de Datos. Pearson Educación.
- [3] Cajas Ixco de Puluc, P. J. (2018). Experiencias psicológicas en el uso de la tecnología en adolescentes, estudio realizado en el Instituto Adrián Zapata durante el año 2017 (Doctoral dissertation, Universidad de San Carlos de Guatemala).
- [4] Josep, D. M. (s.f.). ADICCIÓN A LAS REDES SOCIALES. Obtenido de <https://www.clinicascita.com/adiccion-redes-sociales/>
- [5] Duck2. (9 de Enero de 2019). K-Means Clustering: Agrupamiento con Minería de datos. Obtenido de <https://estrategiastrading.com/k-means/>
- [6] Rocío Leticia Cortes Campos, A. Z. (2015). el estudio de los hábitos de conexión en las redes sociales virtuales, por medio de la minería de datos. *ISSN*, 99-114
- [7] Blasco, E. A. (2015). Aplicación de técnicas de minería de datos en redes sociales/web. Valencia.
- [8] Goicochea, A. (11 de Agosto de 2009). CRISP-DM, Una metodología para proyectos de Minería de Datos. Obtenido de Tecnologías de la Información y Estrategia: <https://anibalgoicochea.com>
- [9] Peris, M., Maganto, C., & Garaigordobil, M. (2018). Escala de riesgo de adicción-adolescente a las redes sociales e internet: fiabilidad y validez (ERA-RSI).