



**MAESTRÍA EN AUDITORIA DE TECNOLOGÍA
DE LA INFORMACIÓN**

POLARIDAD DE LAS OPINIONES EN REDES SOCIALES APLICADA A LA CLASIFICACIÓN DE LOS SERVICIOS UNIVERSITARIOS

Propuesta de artículo presentado como requisito para la obtención del título:

Magíster en Auditoría de Tecnologías de la Información

Por la estudiante:

Mentor Javier SANCHEZ GUERRERO

Bajo la dirección de:

Joe ECHEVERRIA REYES

Universidad Espíritu Santo
Maestría en Auditoría de Tecnología de la Información
Samborondón - Ecuador
Junio del 2017

Polaridad de las opiniones en Redes Sociales aplicada a la clasificación de los servicios universitarios

Polarity of opinions in Social Networks applied to the classification of university services

Mentor Javier SANCHEZ GUERRERO¹

Joe ECHEVERRIA REYES²

Resumen

La presente investigación cumple el objetivo de utilizar el Big Data generado a través de las opiniones de las redes sociales de las Instituciones de Educación Superior, las cuales están libres de sesgos al momento de manifestarse sobre los servicios TI que ofrecen dichas instituciones, proporcionando un banco de datos que sirve de materia prima para realizar una minería de datos y de texto que es utilizado para guiar el proceso de ITIL en su fase de Estrategias de Servicio. El proceso de minería utiliza la metodología conocida como CRISP – DM que obtiene valores que determinan la calidad del servicio estudiado, dando una entrada para la toma de decisiones que permitirán mantener, mejorar o eliminar el servicio TI analizado. Finalmente se concluye que utilizar las redes sociales para conocer las opiniones de los servicios TI ofrecidos por las instituciones y a la vez aplicar un proceso de minería para encaminar la adaptación de mejores prácticas de ITIL es un proceso confiable por cuanto los usuarios se sienten en la libertad de opinar acerca del servicio recibido.

Palabras clave:

Big Data, ITIL, Redes Sociales, Servicios TI, CRISP-DM.

Abstract

The present research fulfills the objective of using the Big Data generated through the opinions of the social networks of the Institutions of Higher Education, which are free of biases when manifesting about the IT services offered by these institutions, providing a bank of data that serves as raw material for data mining and text that is used to guide the ITIL process in its Service Strategy phase. The mining process uses the methodology known as CRISP - DM that obtains values that determine the quality of the service studied, giving an input for the decision making that will allow to maintain, improve or eliminate the IT service analyzed. Finally, it is concluded that using social networks to know the opinions of the IT services offered by the institutions and at the same time applying a mining process to guide the adaptation of ITIL best practices is a reliable process because the users feel in the freedom to comment on the service received.

Key words

Big Data, ITIL, Social Networks, IT Services, CRISP-DM.

¹ Estudiante de Maestría en Auditoría de Tecnología de Información, Universidad Espíritu Santo – Ecuador. E-mail mjsanchez@uees.edu.ec.

² Magíster en Seguridad Informática Aplicada. Docente de la Maestría en Auditoría en Tecnologías de la Información Universidad Espíritu Santo-Ecuador.

INTRODUCCIÓN

La creciente influencia en las actividades cotidianas del ser humano, apoyadas en el uso de la informática, redes y tecnologías, promueven a que se genere una masificación de información derivada de éstas. Según Paredes (2015) se requiere de nuevas formas del tratamiento y procesamiento de la información, que pueda manejar gran cantidad de datos muy diversos a grandes velocidades, con el objetivo de obtener conocimiento que permita tomar decisiones adecuadas, todo esto se ha convertido en una oportunidad para escalar e incrementar su modelo de negocio y sacar ventaja en un entorno globalizado y demasiado competitivo.

Una sociedad que se basa en información y conocimiento soportados por la tecnología originando la sociedad de la información y a medida que este fenómeno ha ido creciendo se transforma de una línea tecnológica a una de más relevancia basada en el usuario y en los contenidos que éste maneja en consumo y generación de información, generando profundos cambios en diferentes contextos (Fernández, 2014). Tomando en cuenta la imperiosa necesidad de alinear las reglas del negocio empresarial o institucional al uso primordial de las TI, optimizándola en los niveles de la organización con la finalidad de mantener procesos eficientes y costes de manera razonable y predecible. El gran problema en la actualidad en la prestación de servicios es garantizar que la prestación del mismo, sea de conformidad con las necesidades del cliente sin la necesidad de incurrir en costos o esfuerzos mayores (Medina & Rico, 2011). Lo que significa que el criterio del usuario acerca del servicio recibido, siempre será lo más importante para la empresa que los oferte.

En las universidades, las actividades docentes, administrativas y estudiantiles, están marcadas por la generación de información que deriva en conocimiento y toma de decisiones. Es de suma importancia saber cómo y dónde se organiza la

información, con el objetivo de que sea fácilmente interpretable (Reyes & Nulez, 2015).

En esta línea, las instituciones de educación superior han venido desarrollando servicios académicos mediante Tecnologías de la Información TI que permiten gestionar y administrar de manera adecuada las actividades de los diferentes actores universitarios. Además el uso constante de las redes sociales, dan paso a generar opiniones acerca de diversos temas de interés dentro de la comunidad, que sirven para capturar opiniones de manera franca y sin sesgos de la funcionalidad de los servicios ofertados (Díaz, 2011). Las universidades del país ya sean de orden público o privado, se apoyan mayormente en redes sociales como Facebook, Twiter o Instgram, para el presente estudio se tomó en cuenta Facebook por ser de mayor uso por parte de los involucrados.

El reto, en realidad no está únicamente en las fuertes inversiones tanto de hardware o de software sino en lo que el cliente piense de la calidad de los servicios recibidos, cuando no se administra adecuadamente simplemente no funciona y no brinda los resultados esperados. La meta debe ser lograr la combinación eficiente de la gente, los procesos y la tecnología para brindar una mejor administración de los servicios de TI, optimizando el uso de los recursos y la mejora permanente de los niveles de servicio (Acevedo, 2010).

Lo anterior, es posible con la incorporación de Buenas Practicas de un Marco Metodológico reconocido, de tal forma que ayude a la empresa a concretar los objetivos planteados. Para tener una guía que permita administrar, gestionar, evaluar y sobre todo un proceso de mejora continua, se puede usar ITIL que tiene amplia documentación y vigencia en todo el mundo. Una de las formas de levantar información necesaria para aplicar ITIL, de las necesidades de los involucrados en el uso de los servicios ofertados, es el uso de la red social Facebook, que da la libertad de expresarse y de opinar en torno a un determinado tema, en nuestro caso particular el uso de los servicios tecnológicos que poseen los

centros de educación superior. Estas opiniones o comentarios emitidos por los usuarios de los servicios pueden ser miles, lo que forma el Big Data de opiniones de un tema en particular.

Según Parra, Gordo, & D'Antonio (2014), la red social Facebook se ha convertido en parte de nuestro imaginario social, aumentando su participación y notoriedad en la nueva era del Big Data; tomando como premisa que el término Big Data aún no posee un concepto definido, puesto que el mismo ha sido descrito como un "término relativo" (Minelli, Chambers, & Dhiraj, 2013).

En este trabajo se analiza las opiniones y se obtiene su polaridad (positiva, negativa), este proceso permite que se pueda clasificar los servicios según la polaridad asignada a cada uno de ellos. La propuesta es identificar la polaridad del texto para luego tratar de interpretar esto como parte de una emoción positiva o negativa. No se pretende asegurar la polaridad de un texto, pues inclusive éste depende del estado de ánimo de la persona, de allí que; nuestros resultados son solamente una guía para una toma de decisión de parte de un experto humano.

La técnica que se utiliza para una parte de este trabajo está basada en minería de datos y en Point Mutual Information (PMI) permitiendo obtener la orientación semántica (SO) de las frases que componen una opinión. La selección de frases se realiza cumpliendo los patrones part of speech de clasificación no supervisada propuesto por (Liu, 2012).

El objetivo general del estudio es clasificar los servicios TI de acuerdo a la polaridad de las opiniones obtenidas de las redes sociales, para la efectiva toma de decisiones.

Sin embargo, esto debe estar enmarcado en un proceso metodológico como ITIL, de tal forma que se tendrá objetivos específicos para alcanzar el general:

- Validar el cumplimiento de los objetivos de los procesos que lleva la institución en cada fase del ciclo de vida del servicio basados en ITIL V3.

- Diseñar una estructura Big Data que permita captar las opiniones acerca de los servicios TI y su calidad
- Diseñar una técnica que permita a través de la polaridad de opiniones sobre los servicios de TI, mejorar la toma de decisiones directivas.

MARCO TEÓRICO

TI Tecnologías de la Información

Como lo indica (Huerta & Larraza, 2001) las Tecnologías de la Información o TI se refiere al conjunto de elementos, ordenadores, fibra óptica, cámaras, robots y algunos otros dispositivos que, unidos con las aplicaciones y sistemas asistidos por computador, tienen la tarea fundamental en la empresa de procesar, almacenar, evaluar y comunicar la información. Existen tres áreas que integran lo que podemos señalar como TI: la informática, las telecomunicaciones y la ofimática, generando una infraestructura de máquinas y sistemas que permiten la relación de las diferentes actividades de la empresa o institución; y entre la empresa y sus clientes y proveedores.

Garcez, Kassouf, Kassouf, Canniatti, & Canhadas (2012), mencionan que las TI o las Tecnologías de la Información son las responsables de la generación de ideas, conocimiento e información en el ambiente empresarial y corporativo por lo que su gestión cumple un rol muy importante en las instituciones. El adecuado funcionamiento de las TI reside en el equilibrio de las demandas reales del desarrollo de los proyectos dentro de los recursos y plazos esperados (Roger & Kuldip, 2014).

Las TI aparecieron hace más de cuatro décadas y si bien representan un gran apoyo en las instituciones, en muchas de ellas siguen representando un costoso problema a resolver; esto se da por cuanto la alta gerencia aún no se involucra de manera efectiva en la adaptación de sus servicios al uso de tecnologías, limitando su

capacidad operativa y de calidad (Feld & Stoddard, 2004)

La integración de TI al modelo empresarial debe tener un enfoque sistemático, basado en tres fases:

- Un proceso de renovación de TI a largo plazo y que esté fuertemente ligado a la estrategia de la empresa.
- Poseer una plataforma tecnológica que resulte simple y unificadora, que se enmarque en una arquitectura abierta con diseño horizontal y sirva a la empresa en su conjunto.
- Una administración de TI funcional y enfocada al servicio y alto desempeño, que trabaje en equipo de acuerdo a los estándares corporativos.

SERVICIOS

Los servicios de TI son una manera de suministrar valor a los clientes, entregándoles resultados que ellos quieren obtener sin tener que asumir costos y riesgos específicos derivados de la transacción (Abad, 2013).

PROCESOS

Se los denomina como un conjunto de acciones que están sistematizadas las que combinan e implementan recursos y habilidades con la finalidad de llegar a un resultado y que de manera directa o indirecta aportan valor para el usuario o el cliente (Abad, 2013).

Al valor se lo entiende como la combinación de la utilidad más la garantía, siendo la utilidad lo que desea o requiere el cliente mientras que la garantía es la seguridad de que los servicios o productos suministrados estén disponibles. Cualquier servicio es protegido por uno o varios procesos, resultando fundamental esta noción para conocer lo que es Servicios de TI y su gestión (Khadige, Mohammad, Peyman, & Ali, 2009).

SERVICIOS DE TI

Se refiere a un grupo de acciones que dan respuesta a una o varias necesidades de un

cliente mediante un cambio de condición sobre los bienes informáticos, potenciando el valor de éstos y reduciendo el riesgo inherente del sistema. Un Servicio de TI se basa en el uso de las Tecnologías de la Información y soporta los procesos de negocio del cliente, se compone de una combinación de personas, procesos y tecnologías y debería estar definido en un Acuerdo de Nivel de Servicio (Abad, 2013).

GESTION DE SERVICIOS DE TI

Para Abad (2013) la Gestión de Servicios de TI es un conjunto de habilidades gerenciales especializadas en otorgar valor al cliente en forma de servicios, atendiendo a un régimen basado en procesos, destinada a alinear los servicios de TI; proporcionados con las necesidades de la empresa o institución, orientando un especial énfasis en las bondades que recibirá el cliente final.

SERVICIOS DE TI EN LAS INTITUCIONES DE EDUCACION SUPERIOR

La infraestructura de las denominadas Tecnologías de la Información, es de importancia fundamental para que las organizaciones construyan su sistema de información. (Laudon & Laudon, 2004). Esta infraestructura está conformada por los equipos físicos (hardware) así como los programas que trabajan en ese hardware (software), esto significa que hay una combinación entre medios de almacenamiento, de comunicación y los programas; además de impresoras, cámaras y demás hardware necesario.

Dentro de las organizaciones también están involucradas las Instituciones de Educación Superior, que manejan grandes cantidades de documentación y a la vez servicios de TI, que son a la vez llamados Servicios Académicos de TI como puede ser el Sistema de Matriculas, Bibliotecas agrupados en un sistema de información global que está al servicio de sus clientes que son principalmente los miles de estudiantes (Siniša, Milan, Jelena, Jelena, & Edit, 2017).

ADMINISTRACIÓN DE PROCESOS

Consiste en la coexistencia con la administración funcional, en la que se tiene personas asignadas a los procesos claves y de interés mayor, dando la posibilidad de crear una gestión inter funcional generadora de valor agregado para sus clientes, lo que significa su satisfacción, dando indicadores de los procesos que ser cambiados o mejorados, crear prioridades y mantener planes de mejora que permitan alcanzar objetivos, mostrando las fortalezas y debilidades (Acevedo, 2010).

MEJORES PRÁCTICAS PARA LA GESTIÓN DE SERVICIOS TI

Según Lopez (2013) las mejores prácticas se definen como un coherente conjunto de acciones que han llevado a obtener el éxito en un determinado contexto y que se prevee que en ambientes similares, tengan los mismos resultados.

MARCOS DE REFERENCIA DE LA GESTIÓN DE SERVICIOS IT

Con la finalidad de que las organizaciones no mantengan una visión basada en el uso de tecnología y en sus propias organizaciones, se debe considerar la mejora continua y calidad de los servicios que ofrecen y enfocarse en las relaciones con sus clientes.

Las instituciones de Educación Superior no están fuera de estos objetivos y la Gestión de los Servicios de TI deben basarse en la promoción y soporte de la aplicación de las mejores prácticas, además de marcos de referencia y demás patrones de aceptación internacional, de los cuales se pueden citar: ITIL, COBIT, ISO.

ITIL

Loayza (2016) conforme se indica en Office of Government Commerce (2008), **ITIL** es el marco de buenas prácticas en la administración del servicio, que permite establecer y mejorar sus capacidades, además de que ITIL ofrece un cuerpo de conocimientos útiles para alcanzar dicho estándar.

ITIL proviene del acrónimo IT Infrastructure Library, que traducido al español sería Biblioteca de Infraestructura de TI, ofrece un grupo de conocimientos de utilidad para alcanzar altos estándares. ITIL se compone de: núcleo y publicaciones complementarias; siendo que el núcleo cuenta con el grupo de mejores prácticas que pueden ser aplicados en cualquier tipo de organización que presta servicios de TI. Las publicaciones complementarias de ITIL tienen un conjunto de publicaciones específicas para industria, modelos operativos y demás arquitecturas tecnológicas.

Existen 5 libros que componen el núcleo de ITIL y a la vez forman su ciclo de vida:

- Estrategia del servicio
- Diseño del servicio
- Transición del servicio
- Operación del servicio
- Mejora continua del servicio

AREAS DE CONOCIMIENTO DE ITIL

ITIL en su versión V3 está compuesta de cinco publicaciones o libros que se describen brevemente en la Tabla 1.

Tabla 1: *Gestión de Servicios TI*

Libro	Descripción
Estrategia del servicio	Se busca mejorar el impacto estratégico a través del diseño, desarrollo, implementación y práctica de la Gestión del Servicio <ol style="list-style-type: none">1. Generación de Estrategia2. Administración de la demanda3. Administración del Portafolio de Servicios4. Administración Financiera de TI
Diseño del Servicio	Proporciona guías sobre cómo diseñar los servicios de TI y cómo evaluar las interacciones entre las soluciones y sus entornos de negocio. Incluye los procesos de: <ol style="list-style-type: none">1. Gestión del Catálogo de Servicios.2. Gestión del Nivel de Servicio.

	3. Gestión de la Capacidad. 4. Gestión de la Disponibilidad. 5. Gestión de la Continuidad del Servicio de TI. 6. Gestión de la Seguridad de la
Transición del Servicio	Proporciona guías sobre cómo entregar un servicio para uso operativo. Incluye los procesos de: 1. Planificación y soporte de la Transición. 2. Gestión de Cambios. 3. Gestión de la Configuración y Activos del Servicio. 4. Gestión de Versiones y Despliegues. 5. Validación y Pruebas del Servicio. 6. Evaluación. 7. Gestión del Conocimiento del Servicio
Operación del Servicio	Proporciona guías sobre la consecución de los niveles de calidad acordados para el servicio entregado. Incluye los procesos de: 1. Gestión de Eventos. 2. Gestión de Incidencias. 3. Gestión de Peticiones. 4. Gestión de Problemas. 5. Gestión de Acceso. 6. Operaciones de TI
Mejora Continua del Servicio	Proporciona guías sobre cómo adaptar los servicios actuales a las necesidades nuevas del cliente. Incluye el proceso de: 1. Proceso de mejora continua del servicio.

Fuente: (Guzmán & García, 2012).

Para el caso de la presente investigación La Estrategia del Servicio de ITIL, se fundamentará en Polaridad de las opiniones en Redes Sociales aplicada a la clasificación de los servicios universitarios, por lo que se tomará la Big Data generada en las redes sociales, de las opiniones de los servicios prestados en una Institución de Educación Superior.

BIG DATA

Paredes (2015) describe al Big Data como la tendencia producto del avance tecnológico lo que ha dado paso a un nuevo paradigma de comprensión y oportuna toma de decisiones que se originan en la gran cantidad de datos

obtenidos de forma no estructurada, estructurada y semi estructurada, la misma que sería muy exigente de analizarla mediante una base de datos relacional demandando tiempo y dinero.

Es un modelo computacional destinado a procesar y tratar ingentes cantidades de datos procedentes de diferentes y variadas fuentes lo que permite que siempre esté disponible, trazable, visible y calculable (Rennó, 2014).

Paniagua (2013) indica que para entender al Big Data primero se debe comprender qué son los datos y donde están, de tal forma que menciona que se encuentran en todas partes y que los seres humanos son quienes los generamos; a través de las formas más variadas tomadas por sensores, redes sociales, latidos del corazón, respiración, medidores, que de paso abren camino también al Internet de las Cosas.

MINERIA DE DATOS

Se la conoce también como Descubrimiento de Conocimiento en Bases de Datos derivada de sus siglas en inglés KDD Knowledge Discovery in Databases y representa un área del conocimiento que permite ir descubriendo nueva información potencialmente útil recolectada de grandes fuentes de datos. (Escobar & Alcivar, 2016).

El gran crecimiento de los datos crea la necesidad de tener información significativa que son extraídas de estos grande volúmenes de datos, es así como, este campo de la minería de datos permite extraer información y a la vez mejorar la calidad de los procesos (Raheela, Agathe, Syed, & Najmi, 2017) tomando en consideración que las técnicas de minería de datos proporcionan un conjunto de varias herramientas muy potentes que permiten generar varios sistemas de clasificación que son impulsados por datos previos, es así como estas técnicas pueden desempeñar un papel demasiado importante en el refinamiento de reglas (Peral, Maté, & Marco, 2017).

En minería de datos un sistema de clasificación que se basa en reglas es muy importante puesto

que el mismo permite tomar decisiones que en su mayoría son asertivos, un árbol de decisión puede usarse en la generación de reglas que son impulsadas por los datos, pero para el tamaño de una muestra estas reglas tienden a describir varios valores atípicos que no necesariamente generalizan a un conjunto de datos más grandes (Singaravelana, Aruna, Arunshunmugama, Jerina, & Muruganb, 2017).

MINERIA DE TEXTO

Mientras el volumen de los datos aumenta considerablemente, la prospectiva y sus métodos serán beneficiados de estos mismos datos, de la misma manera sucede en la minería de textos, la cual se basa en la heterogeneidad de los datos textuales los mismos que abordan un nuevo campo de investigación en la integración (Hussein, Alaaeldin, & Hassan, 2015), permitiendo el acceso y examinación de datos que se los agrega de manera ampliamente automatizada.

Es así como al explotar nuevas fuentes de datos como es el caso Twitter, Facebook, Instagram; entre otras (minería web) (Maarten & Patrick Van, 2014), se van integrando más actores y a la vez puntos de vista haciendo más énfasis en el análisis de los cambios sociales; en la actualidad se pretende explotar el potencial que tiene la minería de texto para la prospectiva considerando diferentes fuentes de datos, como son las redes sociales, los métodos de prospectiva y el enfoque de minería de texto (Vijay Kotu, 2015).

La finalidad de la minería de texto es facilitar la detección y el examen de temas y tecnologías emergentes mediante la ampliación de la base de un conocimiento previo, es decir, la minería de texto proporciona una base sólida para reflexionar sobre futuros posibles.

Las técnicas de minería de texto incluyen la categorización de textos, detección de temas, extracción de conceptos, búsqueda, agrupación y recuperación documentos, y a la vez encontrar alguna información no trivial de una colección de documentos con el objetivo de desarrollar

sistemas de recomendación (Mangi, Jaelim, & Kichun, 2017), mediante un conjunto de criterios de evaluación que midan la efectividad de estas técnicas con la intención de facilitar la selección de una técnica apropiada.

METODOLOGÍA

En los diferentes centros de Educación Superior, existen varios servicios para la comunidad universitaria, muchos de los cuales mayormente están dirigidos a estudiantes y a docentes, en menor cantidad a servidores universitarios y trabajadores. Estos servicios deben ser valorados en la satisfacción al usuario para poder tomar decisiones acerca del mismo, siendo uno de las técnicas más utilizadas las encuestas a los involucrados, que muchas de las veces puede suponer un sesgo en la obtención de datos que desvía la real situación de lo que se está investigando.

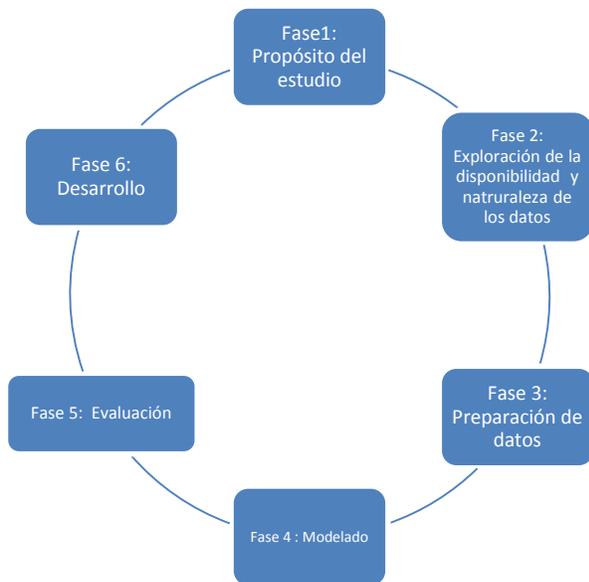
En la presente investigación se plantea usar la Red Social de Facebook para generar opiniones de un determinado servicio, preferentemente de TI, en donde los usuarios comentaran sus opiniones de manera libre y sin sesgos.

Una vez que se ha construido el Big Data de las opiniones descargadas de Facebook acerca del Servicio de Matrículas en el Centro de Idiomas, se procede a establecer la siguiente metodología.

La investigación está enmarcada en el enfoque Cualitativo - Cuantitativo, estos dos enfoques ayudan a la investigación, con lo cualitativo nos enfocamos a los usuarios para realizar consultas de opinión en relación al tema a investigar y cuantitativo nos enfocamos en las dos variables para analizar los resultados obtenidos para determinar la relación y el impacto que existe entre ellas.

Al haber identificado los requerimientos en la construcción del léxico de las opiniones, para la usabilidad se plantean la metodología CRISP-DM.

CRISP-DM es una metodología abierta, gratuita, flexible que se puede personalizar fácilmente para crear modelos que se adapten a las necesidades del proyecto, facilitando la comprensión y manejo de las distintas interacciones entre las etapas involucradas en el proceso, además cabe destacar que una de las principales ventajas de la metodología propuesta es que no requiere necesariamente de personal altamente especializado para su aplicación.



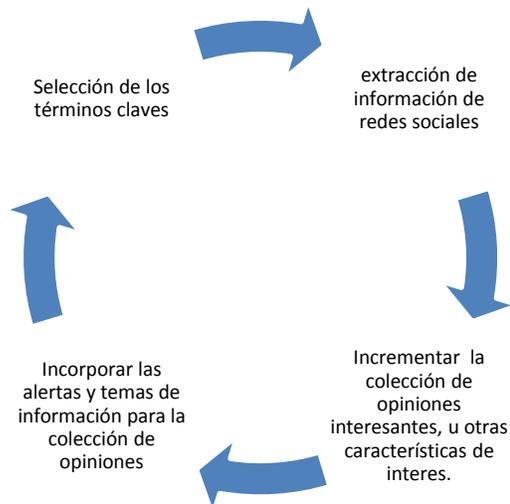
Fase 1: Propósito del estudio

Propone analizar opiniones y obtener su polaridad (positiva, negativa). Este proceso permite que se pueda clasificar los servicios según la polaridad asignada a cada uno de ellos.

Fase 2: Exploración de la disponibilidad y naturaleza de los datos

En esta fase se explica con más detalle el proceso que se ha seguido para recopilar la información. Esencialmente dos fuentes han servido de punto de partida para realizar el análisis de las redes sociales.

El método de extracción de datos se produce mediante la interacción.



Fase 3: Preparación de datos

Se mejora la fiabilidad de los datos seleccionados. Esto incluye clarificar los datos, eliminando valores vacíos y datos no significativos. Dependiendo de la naturaleza de los datos y del estudio, este paso podría involucrar métodos estadísticos complejos e incluso técnicas de predicción, además es necesario realizar una limpieza inicial a los datos para poder trabajar con ellos.

Fase 4: Modelado

El proceso para determinar el tono emocional que hay detrás de una serie de palabras, se utiliza entender las actitudes, opiniones y emociones expresadas en una mención online. El análisis de sentimiento es extremadamente útil en la monitorización de las redes sociales ya que permite hacernos una idea de la opinión pública general sobre ciertos temas.

Para el modelado se aplicará diferentes paquetes para conocer la opinión sobre los servicios universitarios en las redes social Facebook utilizando R. Para ello se mostrará cómo conectarnos desde R a Facebook. Buscaremos las opiniones asociadas de los usuarios asociados a la red social Facebook, donde se clasificarán las emociones y polaridad de las opiniones para luego visualizar la evolución de la opinión y las nubes de palabras, donde se

detectará las opiniones negativas y por qué se deben dichos comentarios del público hacia los servicios universitarios, permitiendo así mejorar los servicios que brindan.

El algoritmo de clusterización se aplicará con el fin de analizar qué tipo de servicios se pueden considerar como positivos y cuales son aquellos que en para los usuarios de esta página de Facebook se consideran como negativos.

Fase 5: Evaluación

Se evalúa e interpreta los patrones descubiertos, en relación a las opiniones que los usuarios tienen hacia los servicios universitarios establecidas al inicio del proceso. Este paso se enfoca en la comprensión y utilidad del modelo inducido. Para la evaluación se selecciona la página y obtención de comentarios, el filtrado de las opiniones, el análisis de sentimientos para luego analizar los resultados de la Clasificación

Fase 6: Desarrollo

Finalmente se incorpora el conocimiento descubierto en el proceso. La idea es incorporarlo al entorno, aplicarlo y medir su impacto. En el caso particular este conocimiento descubierto guía la toma de decisiones en la aplicación de la primera fase (Estrategias del Servicio) de uso de buenas prácticas de ITIL.

ANÁLISIS DE RESULTADOS

Una vez que se ha seleccionado el servicio a analizar, para el presente estudio se utilizó el caso de Matriculación en el Centro de Idiomas de los estudiantes; los usuarios generan opiniones del mismo en Facebook, dando a conocer libremente su opinión acerca de la calidad del servicio presentado, del cual se forma un archivo con el Big Data de las opiniones a través de un archivo con extensión CSV tipo Excel, el cual se conoce como Corpus.

Las opiniones del Corpus CVS, deben ser filtradas antes de empezar a ejecutar la polaridad previa a clasificar por clúster los servicios académicos. El proceso completo se describe:

a. Filtrado de Opiniones

Las opiniones que se difunden en el tema de los servicios universitarios son validadas previamente por un experto en el tema, seleccionando de esta manera las opiniones que van a formar parte del conjunto que serán analizadas mediante Pointwise mutual information (PMI) y el análisis de sentimientos.

1. Selección del tema de los servicios universitarios.

Se ingresó a la página oficial de la UTA y se seleccionó el servicio de Matrícula en Idiomas de los estudiantes universitarios.



Figura 3: Selección caso de estudio: Matriculas en Idiomas.

- b. Obtención de los comentarios.

Se utiliza el Api_Graph de Facebook <https://developers.facebook.com/tools/explorer/>

El número de identificador de la página oficial de la UTA se obtiene ingresando a <http://findmyfbid.com>, en donde se obtiene como resultado 1431460530422719.

Luego se busca la opinión de los servicios de Matriculación en Idiomas, presionamos en la fecha que fue publicado el tema y se copia el número que se genera al final de la url, ejemplo_id: 2012075719027861.

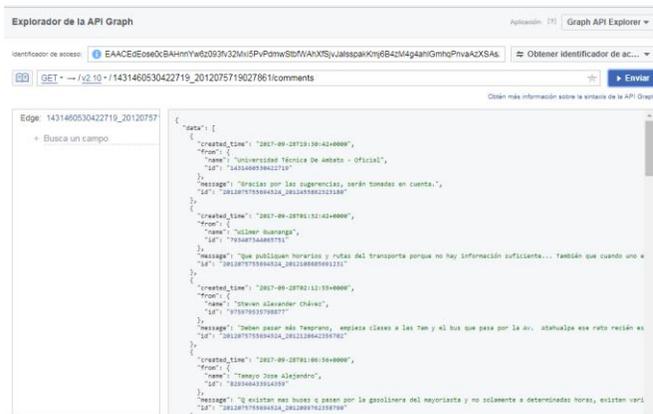


Figura 5: Obtención de comentarios del servicio de matrículas en Idiomas.
Elaborado por: Autor

Al aplicar todo el proceso anteriormente descrito se obtiene los comentarios del servicio de Matriculación en Idiomas en formato JSON. Se pulsa en “Enviar” y se genera el documento con las opiniones.

Para convertir en formato CSV las opiniones obtenidas en formato JSON se utilizó la aplicación Konklone: <https://konklone.io/json/>, se copia los comentarios obtenidos con la API de Facebook.

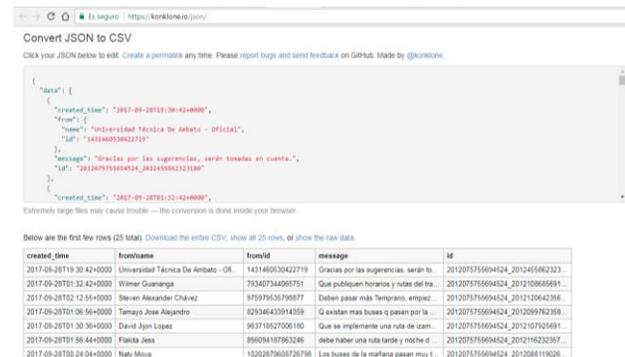


Figura 6: Conversión de JSON a CSV con el tema de estudio
Elaborado por: Autor

Las opiniones del archivo CSV deben ser filtradas por el experto antes de ejecutar la polaridad y analizar el sentimiento previo a la clasificación con el algoritmo de clusterización.

Dicho esto, procedemos a utilizar el software R, elaborado anteriormente y ejecutamos los comentarios en csv.

Figura 7. Resultados del tema de estudio, matrículas para tomar idioma inglés.

Existen mayor cantidad de comentarios positivos que negativos sobre el servicio de Matriculación en Idiomas de la UTA, como se observa en la Fig. 7 existe la probabilidad de que los posibles términos que se incluyen en las opiniones de los servicios de matrícula de la UTA son los términos “inglés”, “proceso” y “web” con una frecuencia alta entre 10.0 y 5.0 en las opiniones de los usuarios del servicio Matrículas de Idiomas.

En la gráfica 8 se muestran los resultados, de las palabras que aparecen al menos 25 veces en las opiniones acerca del servicio de Matriculación de Idiomas, las palabras que resaltan en el grafico denotan mayor presencia en los comentarios de los usuarios y en la gráfica 9 se muestran las 100 palabras con una mayor ocurrencia en los comentarios de los servicios de Matriculación en Idiomas de la UTA que pueden servir a futuro para formar un nuevo archivo o Corpus de palabras en español:

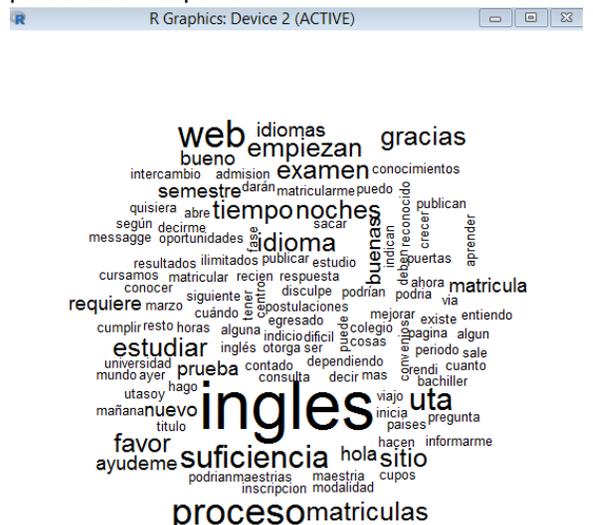


Figura 8: Términos de Frecuencia al menos 25 veces en los servicios de matrículas para tomar idiomas.
Elaborado por: Autor



Figura 9: Términos de Frecuencia de las 100 palabras más usadas en los servicios de matrículas para tomar idiomas.

Elaborado por: Autor

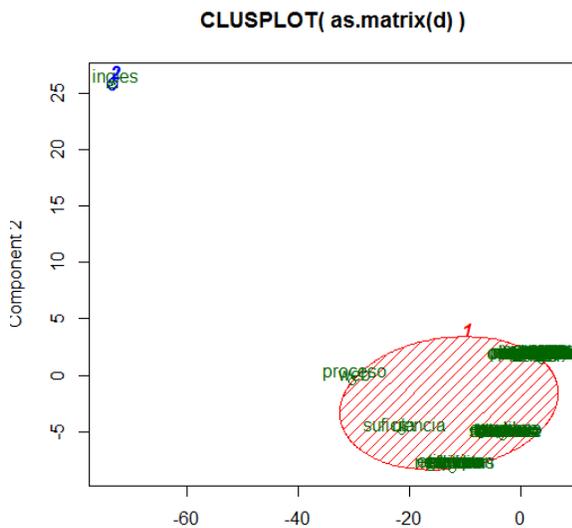


Figura 10: Clusterización del tema de estudio por Kmeans

Elaborado: Autor

En la Figura 10 se obtiene una clusterización con el algoritmo Kmeans en el que nos permitió agrupar un conjunto de observaciones de palabras más cercanas en los comentarios.

CONCLUSIONES

Las decisiones pueden ser tomadas de manera eficiente a través de los resultados obtenidos de la aplicación de la metodología CRISP-DM que genera índices de confianza, porque se basa en el análisis de las opiniones emitidas libremente acerca de un tema en particular por parte de los usuarios, lo que permite tomar decisiones apegadas a la necesidad del cliente adaptadas a las mejores prácticas de ITIL en cada servicio TI que se ofrece en la institución.

Las redes sociales constituyeron una efectiva fuente de levantamiento de datos para armar un Big Data en un archivo Corpus el mismo que sirvió para generar la minería de texto y luego la minería de datos que emitieron resultados confiables para la toma de decisiones.

En las instituciones educativas de nivel superior, la inclusión de estándares y buenas prácticas de ITIL se convierten en una buena opción para gestionar adecuadamente el uso de la tecnología, con la finalidad de mejorar la calidad del servicio y satisfacción al cliente, cumpliendo así con la función de Administración de Niveles de Servicio, dentro del marco ITIL, cuyos beneficios son: concordancia entre los servicios de TI y las necesidades del negocio, y mejorar las relaciones de TI con sus clientes (negocio), en éste caso los estudiantes de la Universidad.

El presente trabajo de investigación estuvo limitado por el requerimiento de una persona experta que analizó de manera inicial las opiniones para clasificar en: positivas o negativas en su polaridad y además que requiere de una limpieza de datos en las opiniones ingresadas en las redes sociales ya que se usan palabras inadecuadas y además el uso de emoticones.

En cuanto al uso del software libre R, es un lenguaje de programación que requiere del manejo a través del uso de su consola en línea de comandos, cuando se muestran mensajes de error resulta complejo la interpretación de los mismos para usuarios novatos.

La Metodología CRISP-DM requiere por parte del operador un amplio conocimiento tanto del

Negocio como de los Datos con los que se trabaja.

Los resultados mostrados en este documento, abren oportunidades a trabajos futuros, para realizar estudios y experimentos diversos temas como educación, infraestructura, economía, sociedad, entre otros. Se fomenta la aplicación de la metodología presentada para aprovechar el uso masivo de las tecnologías y como parte de ésta, el de las redes sociales.

Referencias Bibliográficas

Abad, W. (18 de Febrero de 2013). *Gestión de Servicios de Tecnologías de la Información*. Obtenido de Gestión de Servicios de Tecnologías de la Información:
<http://gsticperu.blogspot.com/2013/02/que-es-servicio-de-ti.html>

Acevedo, H. (16 de junio de 2010). *Magazciturum*. Obtenido de Magazciturum:
http://www.magazciturum.com.mx/?p=50#.Wa39f_jyjIV

Díaz, V. (2011). Mitos y Realidades de las Redes Sociales. *Prisma Social*, 1-26.

Escobar, H., & Alcivar, M. (2016). Aplicaciones de Minería de Datos en Marketing. *Revista Publicando*, 503-512.

Feld, S., & Stoddard, D. (2004). TI a la manera correcta. *Harvard Business Review*, 64-71.

Fernández, Á. (2014). *Sistemas de mobile learning para alumnado con necesidades especiales*. Granada: Universidad de Granada.

Garcez, G., Kassouf, A., Kassouf, N., Canniatti, M., & Canhadas, H. (2012). GESTIÓN DE PROYECTOS INTERNOS DE TI UN

ESTUDIO SOBRE LAS 100 EMPRESAS MÁS IMPORTANTES DE BRASIL. *INVENIO 15*, 81-93.

Guzmán, A., & García, A. (2012). ITIL v3 -Gestión de Servicios de TI. *Revista ECORFAN*, 801-806.

Huerta, E., & Larraza, M. (2001). Tecnologías de la información de la información e innovación organizativa en la empresa española. *Productividad y organización*, 51-60.

Hussein, H., Alaaeldin, H., & Hassan, M. (2015). Selection criteria for text mining approaches. *Computers in Human Behavior*, 729-733.

Khadige, H., Mohammad, F., Peyman, A., & Ali, A. (2009). Information technology policy trends in the world. *Technology in Society*, 125-132.

Laudon, K. C., & Laudon, J. (2004). *Sistemas de informação gerenciais: administrando a empresa digital*. Sao Paulo: Prentice Hall.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.

Loayza, A. (2016). Modelo de gestión de incidentes para una entidad estatal. *INTERFASES*, 221-254.

Lopez, P. (28 de Noviembre de 2013). *Fundamentos de gestión de servicios de tecnologías de información*. Obtenido de
<http://itilunidad5.blogspot.mx/2013/>

Maarten, T., & Patrick Van, E. (2014). Legal aspects of text mining. *Computer Law & Security Review*, 153-170.

- Mangi, K., Jaelim, A., & Kichun, L. (2017). Opinion mining using ensemble text hidden Markov models for text classification. *Expert Systems with Applications*.
- Medina, Y., & Rico, D. (2011). Mejores practicas de gestion para la calidad de los Servicios en Tecnologias de la Informacion. *Gerencia, Tecnologia e Informacion*, 47-57.
- Minelli, M., Chambers, M., & Dhiraj, A. (2013). Big Data, big analytics. *John Wiley & Sons*, 9.
- Paniagua, S. (2013). Un mundo de sensores. De los datos al Big Data. *Revista TELOS (Cuadernos de Comunicación e Innovación)*, 1-3.
- Paredes, A. (2015). Big Data: Estado de la cuestión. En *International Journal of Information Systems and Software Engineering for Big Companies* (págs. 38-59). España.
- Parra, P., Gordo, A., & D'Antonio, S. (2014). La investigación social aplicada en redes sociales. Una innovación metodológica para el análisis de los «Me gusta» en Facebook. *Revista Latina de Comunicación*, 195-212.
- Peral, J., Maté, A., & Marco, M. (2017). Application of Data Mining techniques to identify relevant Key Performance Indicators. *Computer Standards & Interfaces*, 76-85.
- Raheela, A., Agathe, M., Syed, A., & Najmi, G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers & Education*, 177-194.
- Rennó, R. (2014). Smart cities e big data: o cidadão produtor de dados. *URBS. Revista de Estudios Urbanos y Ciencias Sociales*, 13-24.
- Reyes, Y., & Nulez, L. (2015). La inteligencia de negocio como apoyo a la toma de decisiones en el ámbito académico. *GECONTEC: Revista Internacional de Gestión del Conocimiento y la Tecnología*, 63-73.
- Roger, C., & Kuldip, P. (2014). An educational platform to demonstrate speech processing techniques on Android based smart phones and tablets. *Speech Communication*, 13-38.
- Singaravelana, S., Aruna, R., Arunshunmugama, D., Jerina, S., & Muruganb, D. (2017). Inner Interruption Discovery and Defense System by using Data Mining. *Journal of King Saud University - Computer and Information Sciences*.
- Siniša, M., Milan, N., Jelena, J., Jelena, V., & Edit, T. (2017). The impact of information technologies on communication satisfaction and organizational learning in companies in Serbia. *Computers in Human Behavior*, 87-101.
- Vijay Kotu, B. (2015). Chapter 9 – Text Mining. *Predictive Analytics and Data Mining*, 275-303.
-