



**MINERÍA DE
DATOS, DE
TEXTOS, DE
SENTIMIENTOS Y
SUS
DIFERENTES
APLICACIONES
EN LA
ACTUALIDAD**

ARTÍCULO

NIDIA RODRÍGUEZ MAZAHUA

FIA

El autor de este artículo agradece al Instituto Tecnológico de Orizaba máxima casa de estudios a la que le enorgullece pertenecer, a CONACYT por brindar los recursos para la realización de esta investigación, a la Maestría en Ingeniería Administrativa por el voto de confianza y al profesor Fernando Aguirre Y Hernández por la dedicación y compromiso por forjar talentos.

Índice

Resumen.....	4
.Minería de datos, de textos, de sentimientos	5
Introducción.....	5
Minerías de texto	6
El proceso de minería de datos consiste en las siguientes etapas:.....	7
Aplicación de la minería de textos.....	7
Desafíos de la minería de textos.....	9
Minería de sentimientos	10
Objetivos de la minería de sentimientos.....	11
Algo de historia de la minería de sentimientos / opiniones.....	12
Conclusión.....	13
Referencias bibliográficas	14

Resumen

Antes de que surgiera el internet y de que se creara la web , la mayor parte de la comunicación se llevaba a cabo de forma oral , en la cual no se podía tener prueba ni evidencia de estas conversaciones, sin embargo en la actualidad esto es muy diferente más que nada debido a la revolución del internet .

Se han creado sistemas informáticos que se ven representados por una gran cantidad de herramientas y procesos mediante el cual las personas pueden aprender de forma más rápida y mejor.

Sin duda representa todo un reto el poder descubrir conocimientos en grandes volúmenes de datos que se generan en las conversaciones por tal motivo surgió la minería de textos, y representa el área de investigación más nueva del procesamiento automático de texto

Previamente al surgimiento de la web social las personas se hacían recomendaciones unas a otras fuese el ámbito que fuese, sin embargo en nuestros días internet hace posible que los usuarios encuentren opiniones de todo ámbito y de personas de todo el mundo.

En el aspecto de compañías que se dedican a la fabricación de productos y servicios, analizar las opiniones de esta misma que circulan por internet se hace una estrategia de vital importancia hoy en día.

De esta forma las empresas están consientes de que existe una relación muy estrecha entre la inversión que se hace en la optimización del clima social de la marca y los rendimientos de los productos que elaboran.

.Minería de datos, de textos, de sentimientos

Introducción

Antes de que surgiera el internet y de que se creara la web , la mayor parte de la comunicación se llevaba a cabo de forma oral , en la cual no se podía tener prueba ni evidencia de estas conversaciones, sin embargo en la actualidad esto es muy diferente más que nada debido a la revolución del internet .

En el presente los datos de las conversaciones se almacenan a un ritmo acelerado, las personas y las organizaciones participan en intercambios de información y esto se realiza por medio de correos electrónicos, chats virtuales, mensajes de texto, blogs y foros.

Los avances que se han hecho en el procesamiento del lenguaje natural representan oportunidades para que los documentos puedan ser analizados, además la web permite que áreas como las escuelas presenten avances significativos en poco tiempo.

Por esta razón se han creado sistemas informáticos que se ven representados por una gran cantidad de herramientas y procesos mediante el cual las personas pueden aprender de forma más rápida y mejor.

La minería de datos se refiere al proceso de descubrimiento de conocimientos a través del análisis de grandes volúmenes de datos.

Las principales aplicaciones de las técnicas de minería de datos relacionadas con la educación son:

- **Sistemas de personalización:** a través de estos se puede obtener la mayor cantidad de datos posible de los usuarios con el objetivo de mejorar el servicio que se le está brindando.
- **Sistemas recomendadores:** la utilidad de estos se manifiesta en ofrecer al usuario diferentes temas que pueden ser de su interés, esto se hace comparando perfiles de estos con características de los temas.
- **Sistemas de modificación:** se refiere a la detección de necesidades realizando modificaciones necesarias para mejorar la experiencia del usuario.
- **Sistemas de detección de irregularidades:** estos se basan en la detección de errores.

Minerías de texto

La minería de texto se trata de una disciplina especializada en la obtención de información que no está definida de manera explícita a través de identificar patrones en un texto

Sin duda representa todo un reto el poder descubrir conocimientos en grandes volúmenes de datos que se generan en las conversaciones por tal motivo surgió la minería de textos, y representa el área de investigación más nueva del procesamiento automático de textos.

La minería de texto, “se define como aquel proceso automático de descubrimiento de patrones interesantes en una colección de textos, estos patrones no deben estar explícitamente en ningún texto de la

colección, y deben surgir de relacionar el contenido de varios de ellos
“ (Barrio, 2014)

Por ejemplo en Twitter existe el hasta que aparentemente pudiera ser un indicador de temas de interés, sin embargo gracias a la minería de datos se puede clasificar a los usuarios con base en sus preferencias, ideas y opiniones. (semanticwebbuilder, 2016).

El proceso de minería de datos consiste en las siguientes etapas:

1. Pre procesamiento: en esta etapa los textos se transforman en alguna forma de representación estructurada o semiestructurada que de alguna manera vuelva más fácil el análisis posterior.
2. Descubrimiento: en esta etapa las representaciones intermedias se analizan con el objetivo principal de establecer patrones de interés o nuevos conocimientos.

Aplicación de la minería de textos

Entre las aplicaciones de la minería de datos se pueden encontrar las siguientes:

- Extracción de información: existe un gran número de documentos que se encuentran en la red , la minería de textos permite la extracción de información relevante de cantidades extensas de textos y logra definir entidad y relacionarlas y de esta forma revelar información relevante

de los textos logrando que se definan entidades y sus relaciones , revelando información semántica importante la cual se usa como metadatos que a la vez se agregan a contenidos web y de esta manera se facilita la comprensión de estos documentos.

- Análisis de sentimientos o minería de opiniones: es el análisis de las opiniones que generan los usuarios de redes sociales y ayuda a revelar información de importancia para temas específicos y de esta manera puede apoyar a diferentes campos como la inteligencia en los negocios y la toma de decisiones de consumidores.
- Clasificación de documentos: este es un aspecto clave para facilitar la recuperación y navegación de documentos ya que si se desea obtener información de diferentes áreas de desarrollo de los proyectos resultaría muy difícil, sin embargo gracias al uso de algoritmos de minería de datos se puede agrupar los documentos, para así obtener información que describa a cada grupo para hacer más fácil la comprensión de cada uno de estos.
- Elaboración de resúmenes: tiene como objetivo obtener descripciones generales de un conjunto de documentos que pertenecen a un tema específico los métodos pueden ser clasificados en dos categorías:
 1. Sumarización extractiva : resúmenes que se conforman de unidades de información extraídas de textos

2. Sumarización abstracta, en esta la información sintetizada no está necesariamente formada por unidades de información contenida en los textos.
- Extracción de conocimientos: por medio de minería de datos, es posible que se realicen representaciones de información extraída a través de modelos de conocimiento.

Desafíos de la minería de textos

La minería de textos debe enfrentar los siguientes desafíos:

1. Debe conocer el contexto en el que se generaron los contenidos de las diferentes tareas de la minería de textos, ya que debe saber quién es el autor, la región en la que se encuentra y el momento histórico para así poder comprender la información.
2. Debe ir más allá de obtener información y buscar la obtención de conocimientos esto implica la transformación de la información extraída en diferentes tareas del área a un lenguaje que pueda ser entendido por las máquinas.
3. Es necesario que se analicen cantidades muy grandes de información, esto puede ser muy difícil a la hora de hacerlos procesos y se deben definir algoritmos paralelos que aprovechen mejor la infraestructura de cómputo actual. (Lozada, 2016).

Minería de sentimientos

Previamente al surgimiento de la web social las personas se hacían recomendaciones unas a otras fuese el ámbito que fuese, sin embargo en nuestros días internet hace posible que los usuarios encuentren opiniones de todo ámbito y de personas de todo el mundo.

En el aspecto de compañías que se dedican a la fabricación de productos y servicios, analizar las opiniones de esta misma que circulan por internet se hace una estrategia de vital importancia hoy en día.

De esta forma las empresas están conscientes de que existe una relación muy estrecha entre la inversión que se hace en la optimización del clima social de la marca y los rendimientos de los productos que elaboran.

El analizar los sentimientos en opiniones es un procedimiento que se lleva a cabo aplicando técnicas de minería de datos sobre los textos que las expresan.

La minería de datos consiste en la extracción con un objetivo específico de la información que se encuentra implícita en los datos, esta información es previamente desconocida y puede ser útil para ciertos procesos.

De esta manera se puede comprender que la minería de datos prepara, sondea y explora los datos para extraer información que se encuentra oculta en estos.

La minería de datos engloba el conjunto de las técnicas encaminadas a la extracción de conocimientos procesables que está implícito en las bases de datos.

Las bases de la minería de datos pertenecen a la inteligencia artificial y mediante los modelos extraídos utilizando procesos de minería de datos se aborda la solución a problemas de predicción, clasificación y segmentación.

Objetivos de la minería de sentimientos

La minería de sentimientos tiene como principal objetivo el de crear herramientas web que haga más fácil la extracción de las opiniones de distintas redes sociales para posteriormente poder clasificarlas.

Sin embargo esta herramienta debe ser capaz de otorgar una interfaz que sea entretenida y que resulte amigable para el entrenamiento de sistema que se especializa en la clasificación de opiniones, y permita mostrar resultados de forma más cómoda de interpretar que si fueran simples datos.

De esta manera se podrán demostrar la importancia de las opiniones de los usuarios y como estas influyen en la popularidad

e integración en el mercado o también en el fracaso de ciertos productos. (Pelado, 2011)

Algo de historia de la minería de sentimientos / opiniones

El termino minería de opinión apareció por primera vez en un artículo que se publico en la conferencia de la www¹ en el 2003 , debido a esta publicación el termino es muy popular en comunidades fuertemente asociadas a la búsqueda en web o de recuperación de información.

En el artículo se manifiesta que la herramienta ideal de opinión debe procesar resultados de búsquedas que se realizan sobre el tema establecido.

Sin embargo el termino análisis de sentimientos ha sufrido algo parecido apareció en 2001 en artículos que se relacionaban con el marketing y tomo relevancia en el mundo de las computadoras al publicarse en trabajos posteriores en estas áreas. (Lopez, 2011)

¹ World Wide Web

Conclusión

La minería de datos, de textos y de sentimientos tiene una gran importancia en el mundo actual ya que gracias a esta se pueden sintetizar de gran manera los procesos de obtención y decodificación de información que existe en internet y que contienen datos relevantes para la toma de decisiones de muchas empresas, así como también para mejorar los procesos de atención a clientes y desarrollo de productos.

La minería de datos, de textos y de sentimientos facilita las actividades de todas las personas alrededor del mundo y pretende seguir avanzando a pasos agigantados.

El reto más grande de la minería es lograr otorgar datos precisos que sean confiables y que estén clasificados según su procedencia para así lograr los mejores resultados para las mejores tomas de decisiones de la humanidad.

Referencias bibliográficas

Barrio, E. (2014). Minería de texto para el análisis de la colaboración en SIENA. *Universidad de la Laguna*.

Lopez, L. (2011). Clasificador basado en análisis de sentimiento. *IUMA*.

Lozada, A. (2016). *semanticwebbuilder*. Obtenido de semanticwebbuilder.:
http://www.semanticwebbuilder.org.mx/es_mx/swb/Mineria_de_textos_y_sus_aplicaciones

Pelado, J. (2011). Minería de datos aplicada al análisis de sentimiento en opiniones de internet. *Escuela Universitaria de Informática*.

Semanticwebbuilder. (2016). *semanticwebbuilder*. Obtenido de semanticwebbuilder:
[wb/Procesamiento_de_Lenguaje_Natural_y_Mineria_de_textos](http://www.semanticwebbuilder.org.mx/wb/Procesamiento_de_Lenguaje_Natural_y_Mineria_de_textos)

Propuesta de tesis

Minería de opiniones y su impacto actual en el desarrollo de empresas multinacionales.