

PARRAS - Prototipo de apoyo para la recomendación de restaurantes basado en análisis de sentimientos de reseñas en línea

Trabajo Terminal No.

*Alumnos: *Chacón Inostrosa Jaime Enrique, Medina Villalpando Josué de Jesús, Sánchez
García Francisco Nicolas.*

*Directores: Moreno Cervantes Axel Ernesto,
Hernández Cruz Macario.*

e-mail:

jchaconi1400@alumno.ipn.mx

jmedinav1500@alumno.ipn.mx

fsanchezg1403@alumno.ipn.mx

Resumen – Se realizará un sistema para el análisis de sentimientos en reseñas encontradas de los restaurantes registrados en línea dentro de la Ciudad de México mediante una aplicación móvil, con el fin de llegar a tener una recomendación basada en diferentes rúbricas, que ayuden al cliente a poder elegir un restaurante mediante retroalimentaciones, sin recurrir a una gran cantidad de las mismas en otras páginas. Usando *Machine Learning*, específicamente Procesamiento del Lenguaje Natural, se analizará los datos obtenidos utilizando *Web Scraping* para hacer un proceso de clasificación, otorgando resultados que pueden apoyar a los clientes en la elección de sus restaurantes.

Palabras clave – Procesamiento del Lenguaje Natural, *Machine Learning*, *Web Scraping*, Aplicación móvil

1. Introducción

Vivimos en la era digital donde las reseñas positivas en línea son un activo valioso para los restaurantes. Las calificaciones de cuatro y cinco estrellas solidifican las decisiones de los clientes potenciales que están considerando sus opciones gastronómicas. Estas calificaciones ofrecen credibilidad a los restaurantes porque cualquiera puede conectarse a través de las redes sociales y revisar sitios para compartir una reseña positiva o negativa.

Hoy en día, los clientes de los restaurantes se basan sobre todo en las opiniones de los clientes existentes, a su vez, los productores y proveedores de servicios analizan las opiniones de los clientes para mejorar la calidad además de los estándares de sus productos y servicios. [1]

Identificar las emociones de los usuarios en las redes sociales es una de las mayores atracciones para el universo de *Social Media Marketing*. Actualmente no existe alguna herramienta que logre vincular a los usuarios con sus respectivas emociones, basándose en la extracción de datos no triviales (tendiendo una reseña positiva o negativa) publicados en las redes sociales. [2]

Se define sentimiento como una sensación positiva o negativa subyacente a la opinión. A veces, el análisis de sentimientos se interpreta como sinónimo de minería de opiniones, sin embargo, estrictamente hablando, la minería de opiniones es una aplicación que hace uso del análisis de sentimientos y contextualiza las clasificaciones de polaridad en temas, aspectos y objetivos. Aunque el análisis de sentimientos está principalmente orientado al texto, también existen enfoques multimodales. [3]

Los motores de búsqueda y las aplicaciones móviles son una excelente manera de alentar a las personas a visitar los diferentes restaurantes que tal vez no conozcan o que estén visitando su área. No solo ayudan a impulsar el negocio en cuestión de los dueños de restaurantes, sino que también le brindan la oportunidad de causar una impresión duradera. Aplicaciones como *Yelp*, *Foursquare*, *Urbanspoon* y *Zomato* brindan formas de mejorar su reputación en línea y le permiten interactuar con sus clientes. [4]

Existen diferentes tipos de análisis de sentimientos, el se va a utilizar es el análisis de sentimientos polarizado, en el cual se busca polarizar una opinión hacia principalmente tres opciones: positiva, neutra o negativa. Aunque este tipo de análisis no permite descubrir los gustos de los clientes con exactitud, si se acerca al sentimiento general sobre un producto o servicio, además de obtener información de características específicas del mismo. [5]

Planteamiento del problema

El cliente toma muchas características al elegir el restaurante tales como la degustación, limpieza, tipo de comida, precio, y qué tal el servicio para cada uno de los clientes. Además, tienen una gran cantidad de opiniones en diversas plataformas que dificultan discernir entre restaurantes. Para resolver este problema, se busca facilitar la toma de decisiones de los clientes para elegir en qué lugar comer. El análisis de sentimiento ayuda a definir la opinión mediante la detección, extracción y clasificación de ésta. [1]

El análisis de sentimiento realiza un procesamiento del lenguaje natural para poder extraer información sesgada de un gran conjunto de datos. Estos enfoques hacen la extracción de atributos y expresiones como polaridad que define la opinión positiva o negativa. A diferencia de un hecho, la opinión se refiere al sentimiento de las personas y al sentimiento hacia el tema.

Las reseñas son la presentación que tienen los restaurantes ante nuevos clientes potenciales. Tenemos diversas fuentes que nos pueden brindar estas opiniones, tales como Google que sabemos que hoy en día se ha convertido en una plataforma top en cuestión de información. Su motor de búsqueda es uno de los más utilizados en la actualidad ofreciendo diversos datos que recopila diferentes sitios web que están indexados. Así, se convierte en un candidato fuerte para poder brindar las reseñas para su análisis y dar una retroalimentación general del restaurante.

Algunos de los TT presentados con anterioridad son “Prototipo para la detección de emociones humanas en redes sociales” que describe el análisis y diseño de la herramienta para la detección de emociones humanas en redes sociales, si bien, este trabajo se enfoca en las emociones, se muestran antecedentes de obtención de datos a partir de interacciones en Facebook. Dicha herramienta se basa en una aplicación web, la cual permite la extracción y clasificación de información de usuarios de la red social Facebook, con el fin de determinar el grado de pertenencia de cada usuario a tres diferentes emociones, Alegría, Enojo y Tristeza. Finalmente la herramienta también permite visualizar los resultados de manera gráfica. [2]

Por otra parte, tenemos “Sistema de análisis de texto mediante reconocimiento de patrones” en el cual, el objetivo es la implementación de un sistema que analice mensajes mediante procesamiento de lenguaje natural y teoría de reconocimiento de patrones. Este sistema pretende servir como herramienta capaz de clasificar la intención de dichas conversaciones como peligrosas o no peligrosas. [4]

SOFTWARE	CARACTERÍSTICAS	PRECIO EN EL MERCADO
<i>Sentiment Analysis - Know The Opinion</i>	Es una aplicación móvil hace un estudio de las opiniones, sentimientos, actitudes y emociones de las personas expresadas en lenguaje escrito Características claves: <ul style="list-style-type: none"> ● Puede analizar máximo 100 tweets en los últimos 7 días. ● Muestra diferentes gráficas que muestra si es una opinión positiva o negativa ● Disponible solo para Android[8] 	Gratuita. [10]
<i>Sentiment Analyzer - (Twitter / Text)</i>	Es una aplicación móvil que hace un análisis de texto o tweets de las opiniones expresadas. Determina la actitud del escritor hacia un tema, producto, etc. en particular sea positiva, negativa o neutral. <ul style="list-style-type: none"> ● Muestra un indicador de que palabras o líneas de texto fueron tomadas como positiva, negativa o neutral ● Da un reporte detallado acerca de la opinión o tweet que se analizó ● Permite hacer un análisis compuesto de entre 1 a 60 tweets como máximo [9] 	Gratuita [11]
<i>A cloud-based tool for sentiment analysis in reviews about restaurants on TripAdvisor</i>	Herramienta de software basada en la nube para el análisis masivo de datos de <i>TripAdvisor</i> . <ul style="list-style-type: none"> ● Habilidad de agregar datos de la plataforma de TripAdvisor ● La posibilidad de realizar análisis combinados de personas y comentarios ● Posibilidad de detectar el sentido (positivo, negativo o neutral) [10] 	Desconocido [12]

Tabla 1. Resumen de productos similares.

2. Objetivo

Implementar una aplicación móvil que permita brindar a los usuarios una retroalimentación basada en las reseñas de los restaurantes que tengan presencia en la red mediante el análisis de procesamiento de lenguaje natural para encontrar información y relaciones en textos a partir del estudio de sentimientos, con el fin de apoyar a la elección de un restaurante mediante diferentes rúbricas enfocadas en escalas de satisfacción del cliente creadas por nosotros.

3. Justificación

¿Por qué los clientes leen reseñas en línea? En los últimos años, según *Hospitality Insights* [7], los clientes de restaurantes tienen cuatro razones principales para usarlas como punto de referencia para tomar decisiones y como un medio alternativo a las recomendaciones de boca en boca [2].

- La primera razón es la reducción de riesgo, donde los consumidores buscan reducir la probabilidad de que una compra los lleve a la decepción.
- La segunda razón es la reducción del tiempo de búsqueda: A los consumidores les cuesta decidir entre varios restaurantes donde comer, debido a la gran cantidad de opciones.
- La tercera razón es la disminución del remordimiento de los clientes tras gastar una considerable suma de dinero en experiencias gastronómicas.
- Como última razón, la influencia de los grupos: La opinión expresada por compañeros pueden afectar la percepción y el comportamiento del cliente, aunque no se haya visitado un restaurante en concreto.

Dentro de los puntos destacables para el desarrollo del mismo debemos de considerar que efectivamente existen ya sistemas que realizan cierta parte del desarrollo dentro del procesamiento de textos con ayuda del lenguaje natural de manera que ahora vamos a lograr brindar un beneficio distinto pero encaminado a que las personas tengan que hacer menos cosas para obtener una función más completa en conjunto lo cual a comparación de algunas aplicaciones ya existentes podemos otorgar una gran diferencia con otras considerando la complejidad que lleva consigo todo el análisis de los datos y crear un camino totalmente distinto al momento de dejar atrás los momentos en los que era necesario tener que estar inclusive dentro del establecimiento para saber las reseñas de aquellas personas que han asistido a ciertos lugares.[4]

Dentro de la actualidad debemos destacar que ahora una aplicación que haga el vínculo entre los clientes y los restaurantes por medio del análisis de reseñas suele ser un proceso beneficioso para las dos partes de forma en que ahora los clientes pueden averiguar adónde ingresar dentro de su hora de comida de una forma sencilla y rápida de manera que puedan posteriormente regresar a ciertos lugares ya analizados con anterioridad.[9]

Considerando que es un proyecto en el cual se aplican distintos conocimientos de las áreas cursadas dentro de la carrera, en el que el principal factor es “*Machine Learning*”, pero más allá de eso podremos lograr una colección de datos a partir de reseñas que van enfocadas a los clientes y a restauranteros, en donde pueden tener acceso a gráficas, datos e inclusive algunos seguimientos personalizados. Así, se requieren muchos aspectos para su desarrollo, además que será una gran alternativa hoy en día que puede mantenerse en el mercado durante mucho tiempo, por lo que él mismo tiene la capacidad de llegar a otorgarnos un título de ingeniero en sistemas computacionales, donde se verá empleado un sin fin de habilidades obtenidas durante nuestra estancia dentro de la Escuela Superior de Cómputo como:

- Creación y consumo de API.
- Diseño de aplicaciones móviles.
- Diseño, creación e implementación de bases de datos.
- Testing de módulos del sistema.

Las mismas que nos servirán para saber de igual forma que tenemos todos los conocimientos, así como capacidades impartidas en la institución que nos ayudarán a dar lo necesario y más logrando este gran proyecto con miradas a futuro de mediano así como largo plazo, dejando un gran impacto en la forma de elegir restaurantes dentro de la Ciudad de México.

Así, se propone una aplicación móvil que brinde a los clientes actuales y clientes potenciales una retroalimentación detallada en diferentes categorías para que tenga una mejor panorama acerca del restaurante con base a las opiniones y reseñas que han dejado clientes durante su estancia en dichos lugares.

Se propone hacer entidades para tener un rubro en:

- Sabor
- Atención
- Ambiente
- Precio
- Variedad
- General (positivo, neutro y negativo)

para que así, dependiendo a las necesidades y gustos de los clientes, ellos pueden ir seleccionando o filtrados restaurantes con base en una o más clasificaciones realizadas a las opiniones. Así, al ofrecer una aplicación móvil, el usuario podrá portarlo en donde sea desde su dispositivo móvil para tener siempre a su alcance un análisis para tomar en cuenta restaurantes que estén cerca de su ubicación o alguno que se tenga de preferencia o recomendados.

Backlog

- Investigación del estado del arte
- Documento de la delimitación del proyecto
- Definición del alcance del proyecto
- Configuración e integración del repositorio Github en el proyecto
- Creación de historias de usuario
- Diseño de *mockups* de la aplicación móvil
- Diseño de base de datos.
- Investigación de herramienta para crear modelos
- Creación del modelo para el análisis de sentimientos
- Entrenamiento de modelo para análisis de sentimientos en reseñas
- Implementación de base de datos.
- Codificación de *mockups*
- Autenticación de usuarios.
- Inicio de sesión de usuarios.
- Registro de usuarios.
- Recuperación de contraseñas.
- Gestión de perfil de usuario.
- Obtención y visualización de restaurantes usando la API de *Google Maps*.
- *Web Scraping* de reseñas.
- Obtención de restaurante.
- Construcción de *feedback* general.
- Gráficas de datos obtenidos del análisis de sentimientos
- Recepción de *feedback* del restaurante.
- Integración de *Google Natural Language API* alimentado por reseñas
- Creación de manual de usuario de la aplicación
- Creación de manual técnico de la aplicación
- Diseño de plan de mantenimiento
- Pruebas de integración del sistema y modelo

5. Metodología

La metodología a emplear es SCRUM. SCRUM es un *framework* que implementa una metodología ágil basada en sprints.

Al ser un equipo conformado por tres personas, al estar enfocado en la parte técnica, de desarrollo y menos en la parte de investigación de campo, nos va a permitir producir productos de valor en pequeños periodos de tiempo.

Vamos a contar con los siguientes roles:

Scrum Master: Director del trabajo terminal.

Product Owner: Sánchez García Francisco Nicolás.

Developers: Chacón Inostrosa Jaime Enrique, Medina Villalpando Josué de Jesús, Sánchez García Francisco Nicolás.

Stakeholders: Restauranteros y clientes actuales y potenciales

Además, siguiendo la metodología, las reuniones diarias se modificarán para ser cada tres días hábiles, en la cual se realiza una revisión del progreso para cumplir con las metas del sprint.

Se utilizará la herramienta Jira para administración de tareas del proyecto que tiene un enfoque ágil. Se escogió Jira porque es una herramienta con muy buena documentación y está respaldada por cientos de grandes empresas, sin mencionar que tiene una interfaz intuitiva y fácil de manejar.

Para el manejo de control de versiones del proyecto se va a usar Github, ya que cuenta con integración con Jira y nos permite trabajar de manera simultánea.

Para el desarrollo de código del proyecto, se utilizará la IDE Visual Studio Code ya que ofrece cientos de extensiones que permiten tener una codificación más veloz, además de que tiene una gran comunidad que la respalda. También se puede conectar con Github para el manejo de los cambios y versiones directamente desde la IDE.

Sprints:

- TT I:
 - Sprint 1 - 9 de agosto al 29 de agosto:
 - Sprint 2 - 1 de septiembre al 26 de septiembre
 - Sprint 3 - 29 de septiembre al 24 de octubre
 - Sprint 4 - 27 de octubre al 21 de noviembre
 - Sprint 5 - 24 de noviembre al 19 de diciembre
- TT II:
 - Sprint 6 - 7 de enero al 30 de enero
 - Sprint 7 - 3 de febrero al 27 de febrero
 - Sprint 8 - 2 de marzo al 27 de marzo
 - Sprint 9 - 30 de marzo al 24 de abril
 - Sprint 10 - 27 de abril al 22 de mayo
 - Sprint 11 - 25 de mayo al 15 de junio

6. Cronograma

Ver anexo 1, 2 y 3.

7. Referencias

- [1] K. Ravi, V. Ravi. “*A survey on opinion mining and sentiment analysis: Tasks, approaches and applications*”. [Online]. Available at: <<https://www.sciencedirect.com/science/article/abs/pii/S0950705115002336?via%3Dihub>> [Accessed 21 March 2021]
- [2] J.C. Acevedo, R.D. Clorio, J Alameda Morales, M.G. Pardo Sixtos. “*Prototipo para la detección de emociones humanas en redes sociales*”. [En línea]. Disponible en: <<https://tesis.ipn.mx/bitstream/handle/123456789/22688/Reporte%20T%c3%a9cnico%20FINAL%20TT2012-B016%20%20emociones%20humanas.pdf?sequence=3&isAllowed=y>> [Accedido 21 Marzo 2021]
- [3] Kim, E., & Klinger, R. (n.d.). “*A Survey on Sentiment and Emotion Analysis for Computational Literary Studies*”. [Online]. Available at: <<https://arxiv.org/ftp/arxiv/papers/1808/1808.03137.pdf>> [Accessed 10 August 2021]
- [4] Reviewmaxer.com. 2019. “*5 reasons why positive online reviews are essential for restaurants*”. ReviewMaxer. [Online] Available at: <<https://www.reviewmaxer.com/blogs/news/5-reasons-why-positive-online-reviews-are-essential-for-restaurants>> [Accessed 19 March 2021].
- [5] Orphée De Clercq. (2016).” *Fine-grained Sentiment Analysis An Overview of the Task and its Main Challenges*”. [Online] Available at: <https://www.iaria.org/conferences2016/filesHUSO16/OrpheeDeClercq_Keynote_ABSA.pdf> [Accessed 10 August 2021]
- [6] A.E. Guerrero Madrigal, S. Hernández Hernández, A. Ruíz Díaz. “Sistema de Análisis de Texto Mediante Reconocimiento de Patrones”. [En línea]. Disponible en: <<https://tesis.ipn.mx/bitstream/handle/123456789/23047/reporteTecnico%20%20TT2013-B017%20Online%20grooming.pdf?sequence=3&isAllowed=y>> [Accedido 21 Marzo 2021]
- [7] P. Masset, P. Weisskopf. “*Online Customer Reviews: Their Impact on Restaurants*”. [Online]. Available at: <<https://hospitalityinsights.ehl.edu/online-customer-reviews-restaurants>> [Accessed 19 March 2021]
- [8] A. Parikh, C. Behnke, M. Vorvoreanu, B. Almanza, and D. Nelson, “Motives for reading and articulating user-generated restaurant reviews on Yelp.com,” *Journal of Hospitality and Tourism Technology*, vol. 5, no. 2, pp. 160–176, Aug. 2014, doi: 10.1108/jhtt-04-2013-0011. [Online]. Available at: <https://www.researchgate.net/publication/265857674_Motives_for_reading_and_articulating_user-generated_restaurant_reviews_on_Yelpcom> [Accessed 19 March 2021]
- [9] Pymnts.com, 2021.“Consumers Add Masks, Tests To Social Gatherings”, Pymnts.com.[Online]. Available at: <https://www.pymnts.com/coronavirus/2020/consumers-add-masks-virus-tests-to-social-gatherings/>. [Accessed: 20-Mar- 2021]
- [10] https://play.google.com/store/apps/details?id=com.aayat.sentiment_analysis&hl=en_US&gl=US
- [11] https://play.google.com/store/apps/details?id=com.kamran.simplesentimentanalyzer&hl=en_US&gl=US
- [12] M.M. Agüero-Torales, M.J. Cobo, E. Herrera-Viedma, A.G. López-Herrera. “A cloud-based tool for sentiment analysis in reviews about restaurants on TripAdvisor”. (2019). . [Online]. Available: <<https://www.sciencedirect.com/science/article/pii/S1877050919320125>>

8. Alumnos y Directores

CARÁCTER:
Confidencial FUNDAMENTO LEGAL: Art. 3, fracc. II,
Art. 18, fracc. II y Art. 21, lineamiento 32, fracc. XVII
de la L.F.T.A.I.P.G. PARTES CONFIDENCIALES: No.
de boleta y Teléfono

Chacón Inostrosa Jaime Enrique. - Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta: 2015010181, Tel. 55 7483 5123, mailto:jchaconi1400@alumno.ipn.mx

Firma: _____

Medina Villalpando Josué de Jesús. - Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta: 2019630064, Tel. 55 3724 1461, email:jmedinav1500@alumno.ipn.mx

Firma: _____

Sánchez García Francisco Nicolas- Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta: 2015090704, Tel. 5545952862, email: mailto:fsanchezgl403@alumno.ipn.mx

Firma: _____

Moreno Cervantes Axel Ernesto. – Dr. en educación del CUGS en 2020, M. en C. del CINVESTAV en 2004, ISC de la ESCOM en 2000. Areas de interés: sistemas distribuidos, redes de computadoras. Contacto: axelernesto@gmail.com. ext. 52032

Firma: _____

Hernández Cruz Macario.- Maestro en Ciencias de la Computación egresado del Centro de Investigación en Computación del IPN, Licenciado en Sistemas Computacionales egresado de la Universidad del Valle de México, Profesor de la ESCOM desde 2005, ha impartido cátedra en la Universidad Autónoma del Estado de México (UAEM) y en la Universidad de las Américas, A. C (UDLA). Areas de interés: Inteligencia Artificial, Tecnologías Web, Tecnologías del Aprendizaje y el Conocimiento. Tel. 57296000 Ext. 57422, e-mail: mahernandezc@ipn.mx.

Firma: _____

Proyecto No.:

Título del Proyecto: PARRAS - Prototipo de apoyo para la recomendación de restaurantes basado en análisis de sentimientos de reseñas en línea

[illegible]

Título del Proyecto: PARRAS - Prototipo de apoyo para la recomendación de restaurantes basado en análisis de sentimientos de reseñas en línea

[illegible]

Proyecto No.:

Título del Proyecto: PARRAS - Prototipo de apoyo para la recomendación de restaurantes basado en análisis de sentimientos de reseñas en línea

[illegible]

Evidencia de confirmación

Director 1: Moreno Cervantes Axel Ernesto

Contacto: axelernesto@gmail.com



Axel Ernesto <axelernesto@gmail.com>

Jue 12/08/2021 10:12 AM

Para: Josue De Jesus Medina Villalpando

CC: Jaime Enrique Chacon Inostrosa; Francisco Nicolas Sanchez Garcia; Macario Hernandez Cruz

Acuso de recibido y doy visto bueno.



Director 2: Hernández Cruz Macario

Contacto: mahernandezc@ipn.mx



Macario Hernandez Cruz

Jue 12/08/2021 10:55 AM

Para: Josue De Jesus Medina Villalpando

Acuso de recibo.

Enterado y aceptado.

Saludos!

Macario Hernández Cruz



Integrante 1: Sánchez García Francisco Nicolas

Contacto: fsanchezg1403@alumno.ipn.mx



Francisco Nicolas Sanchez Garcia

Mié 11/08/2021 12:42 PM

Para: Josue De Jesus Medina Villalpando

Recibido, enterado y aceptado.

...



Integrante 2: Chacón Inostrosa Jaime Enrique

Contacto: jchaconi1400@alumno.ipn.mx



Jaime Enrique Chacon Inostrosa

Jue 12/08/2021 11:25 AM

Para: Josue De Jesus Medina Villalpando

Acuso de recibido, enterado y aceptado.



Integrante 3: Medina Villalpando Josué de Jesús

Contacto: jmedinav1500@alumno.ipn.mx



Josue De Jesus Medina Villalpando

Jue 12/08/2021 12:42 PM

Para: Josue De Jesus Medina Villalpando

Acuso de recibo.

Enterado y aceptado.

