
TURSIMO DE LOS ALPES

ISIS-3301 - 2024

INTEGRANTES

Esteban Moreno

Gastón Metzger

Carlos Monroy

PROYECTO ANALÍTICA DE TEXTOS

ÍNDICE

Resumen

Abstract

Introducción

ETAPA 1:

Entendimiento del negocio y enfoque analítico.

Mapa de actores relacionado con el producto de datos creado.

Trabajo en equipo

Sustentación y evaluación del aporte individual.

ETAPA 2:

Desarrollo de la aplicación y justificación

Trabajo en equipo

Sustentación y evaluación del aporte individual.

Bibliografía

Resumen

Este proyecto desarrolla un modelo de analítica de textos para evaluar y predecir las calificaciones de sitios turísticos basadas en reseñas de usuarios. A través de dos etapas fundamentales, primero construimos el modelo analítico y luego lo automatizamos y preparamos para su uso en una aplicación final.

Utilizamos técnicas de aprendizaje automático para procesar y analizar datos de reseñas de turistas en Colombia, México y Cuba, proporcionados por importantes actores del sector turístico como el Ministerio de Comercio, Industria y Turismo de Colombia y cadenas hoteleras como Hilton.

El objetivo principal es mejorar la estrategia turística y la popularidad de los destinos evaluados, identificando factores clave que influyen en las recomendaciones turísticas.

Abstract

This project focuses on developing and deploying a text analytics model to assess and predict tourist ratings for sites based on user reviews. The project is executed in two main phases: model construction and model automation for application use.

Machine learning techniques are applied to analyze review data from tourist sites in Colombia, Mexico, and Cuba, provided by key industry stakeholders such as Colombia's Ministry of Commerce, Industry, and Tourism and hotel chains including Hilton. The primary aim is to enhance tourism strategies and the popularity of assessed destinations by identifying crucial factors influencing tourist recommendations.

Introducción

El turismo es un sector vital para la economía global y especialmente significativo para países con ricos destinos culturales y naturales como Colombia, México y Cuba. Este proyecto nace de la necesidad de entender y mejorar la percepción de los turistas sobre estos destinos a través de sus reseñas. En colaboración con el Ministerio de Comercio, Industria y Turismo de Colombia y diversas cadenas hoteleras, se propone desarrollar un modelo de analítica de textos que pueda predecir la calificación de un sitio turístico basada en las reseñas de los usuarios.

La primera etapa del proyecto se centra en la construcción de un modelo analítico, utilizando reseñas de sitios turísticos y calificaciones proporcionadas. Se explorará la relación entre el contenido de las reseñas y las calificaciones recibidas, buscando patrones y términos recurrentes que influyen en las percepciones positivas o negativas. La segunda etapa del proyecto aborda la automatización y la implementación del modelo en una aplicación funcional que será utilizada por los stakeholders para tomar decisiones informadas sobre cómo mejorar la oferta turística y aumentar la satisfacción del cliente.

Este enfoque no solo busca mejorar la calidad de los servicios turísticos, sino también proporcionar una herramienta estratégica para el desarrollo sostenible del turismo en la región.

ETAPA 1:

Entendimiento del negocio y enfoque analítico.

- a. Definición de los objetivos y criterios de éxito desde el punto de vista del negocio. En este punto debe incluir para uno de los actores del sector turismo en el que se enmarca este proyecto, el impacto que puede tener en Colombia este proyecto.

En cuanto a los objetivos, se menciona la búsqueda de características compartidas entre sitios turísticos que presentan calificaciones positivas y negativas con el propósito de analizarlas y compararlas; con esas características se diseñaría implementar un mecanismo para predecir o determinar la calificación que un turista podría asignarle a un local.

Los criterios de éxito definidos son principalmente, lograr un alto nivel de precisión en la calificación predicha, identificar cómo mejorar la popularidad de los locales para mejorar la popularidad entre turistas y aconsejarle que cambiar a los locales que presenten calificaciones negativas.

Esto podría impactar positivamente en todo el sector turístico de Colombia, ya que se podrían fomentar prácticas que generen buenas calificaciones (atrayendo a más turistas) a lo largo del país, mejorando todo el turismo en general, Colombia será un destino mucho más atractivo, y al trabajar orientados a las calificaciones de los turistas y a las experiencias podría volverse un destino importante a nivel mundial, mejorando los ingresos del país.

b. Descripción del enfoque analítico para alcanzar los objetivos del negocio.

Viendo los datos, es claro que para el proyecto se va a necesitar usar técnicas de procesamiento de lenguaje natural. La finalidad es poder extraer información relevante para la predicción de las reseñas: sentimientos, opiniones, temas y/o aspectos destacados de cada sitio. Dentro de estas técnicas estarían la tokenización, lematización, identificación de entidades entre otros, aunque los mencionados anteriormente consideramos son más apropiados para este proyecto.

En cuanto al aprendizaje automático, una primera propuesta es construir modelos analíticos que nos permitan la clasificación automática de reseñas nuevas. Tanto el análisis de sentimientos como de texto es lo que nos permitirá llegar a la calificación. Frecuencia de palabras, longitud de la reseña y/o presencia de términos o expresiones destacadas se considerarán para la ingeniería de las características.

c. Nombres de los estudiantes del curso de estadística con los cuales va a interactuar para validar el enfoque que le está dando el proyecto, a nivel de actor del sector que se verá beneficiado del proyecto, los impactos esperados y la formalidad a utilizar al momento de presentar resultados del modelo analítico construido. Adicionalmente, debe incluir a nivel de planeación, la fecha en que se reunirá con esos estudiantes y el canal que utilizará para socializar el resultado de esta etapa 1 e iniciar el trabajo de la etapa 2.

Ana e Isabella pertenecen al grupo de expertos (Estudiantes del curso de estadística) a los cuales le haremos presentación del modelo que estamos desarrollando con el fin de demostrar que existe una oportunidad de negocio al implementar este modelo y ser capaces de no sólo analizar sino entender y aprender de lo que los usuarios y

clientes del hotel crean y/o afirman que se debe mejorar de ser un error o en todo caso mantener de ser un acierto, esperando impactos positivos como mayor confianza Compañía - Cliente para mejorar la imagen del hotel en cuestión, la planeación se llevará a cabo en grupo, siendo designadas las fechas por el líder de proyecto y/o por el grupo de expertos, en este caso la primera reunión se llevó a cabo el día 25 de marzo del 2024, utilizando la herramienta zoom para posteriormente reunirse de manera presencial el día 2 de abril del 2024.

Mapa de actores relacionado con el producto de datos creado.

Rol dentro de la empresa	Tipo de Actor	Beneficio	Riesgo
Junta directiva	Usuario-Cliente	Apoya en una buena toma de decisiones de negocio	Si el modelo llegara a fallar, esto repercute en la estabilidad del negocio del hotel debido a la toma de decisiones
Dirección de ingeniería de sistemas y computación	Financiador	Modelo de aprendizaje que permita obtener un análisis acertado sobre lo que los clientes del hotel piensan y comentan sobre el mismo	En caso el modelo no funcionase esto repercute directamente en la gestión de la dirección y sería una mala inversión que pudo ser beneficiosa para la empresa en caso se haya dado satisfactoriamente.
Servicios tecnológicos e informáticos	Proveedor	Busca garantizar que se cumplan las necesidades del producto a desarrollar, en ámbitos como la privacidad de los	Manejo incorrecto de los datos que lleven a una violación de la privacidad del usuario.

		usuarios y la seguridad de los datos	
Personal del hotel	Beneficiado	Recibe un feedback constante de cómo se puede mejorar y hacer un trabajo más impecable con menor esfuerzo gracias a los comentarios de los clientes	Afecta psicológicamente de manera negativa debido a la presión de no poder atender correctamente a los clientes.
Cliente del hotel	Beneficiado	Permite dar una opinión con el fin de mejorar el servicio que está consumiendo	Incomodidad y/o molestia de tener que increpar al hotel por algún mal servicio y dar una mala imagen del mismo

Trabajo en equipo

Líder de proyecto: Esteban Rico Se encuentra a cargo de la gestión del proyecto en general, define fechas de reuniones, gestiona avances de los integrantes, supervisa la integridad del proyecto con el fin de hacer entrega de un proyecto claro y fácil de entender para quien está dirigido.

Líder de negocio: Alessandro Metzger Es el encargado de brindar una vista del proyecto alineado a la estrategia de negocio con el fin de poder comunicar los resultados del modelo de aprendizaje de manera eficiente y clara para el grupo de expertos.

Líder de datos: Esteban Rico Se encarga de gestionar los datos recibidos y hacer un entendimiento y una limpieza de estos para un siguiente uso de estos en un modelo de aprendizaje optimo y eficiente el cual requiere de un manejo de datos impecable por parte de su responsable.

Líder de analítica: Carlos Monroy Se encarga de la creación del modelo de analítica con el cual se puede analizar de manera más ordenada y eficiente los datos obtenidos previamente con el fin de la gestión de previsión de reseñas del hotel gracias a un modelo de aprendizaje.

ETAPA 2:

Desarrollo de la aplicación y justificación:

Trabajo en equipo:

- Líder de proyecto:

Gastón lideró la coordinación del proyecto asegurándose de que todos los entregables estuvieran alineados con los objetivos del equipo. Organizó las reuniones semanales, definió las fechas de entrega de pre-entregables y se aseguró de que las tareas estuvieran equitativamente distribuidas entre los miembros del equipo.

Además, fue responsable de subir las entregas finales del grupo y tomó decisiones clave en los momentos en que el equipo no lograba un consenso.

- Ingeniero de datos:

Esteban se enfocó en la calidad y la eficiencia de los procesos de automatización utilizados para construir el modelo analítico del proyecto. Desarrolló y afinó algoritmos para asegurar que los datos fueran procesados de manera óptima, lo que

incluyó la limpieza, integración y análisis de estos para garantizar resultados precisos y confiables.

- Ingeniero de software responsable del diseño de la aplicación y resultados:

Carlos lideró el diseño de la interfaz de usuario de la aplicación, asegurándose de que fuera intuitiva y eficiente para los usuarios finales. También estuvo a cargo de la creación de un video dinámico que presentara los resultados del análisis de datos.

- Ingeniero de software responsable de desarrollar la aplicación final:

Gastón gestionó el desarrollo técnico de la aplicación. Implementó el código necesario para convertir los diseños y prototipos en una aplicación funcional. Trabajó estrechamente con Carlos para asegurarse de que la implementación se alinea con el diseño propuesto y cumplió con todos los requerimientos técnicos y funcionales establecidos.

- Grupo Extendido:

El grupo extendido, compuesto por Ana e Isabella, apoyaron en diversas áreas según fuera necesario, incluyendo pruebas de la aplicación, documentación técnica y ayudando en la preparación de la presentación final del proyecto. Su versatilidad permitió al equipo adaptarse rápidamente a cualquier cambio o desafío que surgiera durante el desarrollo del proyecto.

Sustentación y evaluación del aporte individual:

Gastón Metzger - Líder de proyecto:

Sustentación: Gastón demostró excelentes habilidades de liderazgo y gestión al coordinar todas las fases del proyecto. Su capacidad para mantener al equipo alineado con los objetivos y su eficacia en la resolución de conflictos fueron cruciales para mantener el proyecto en curso.

Evaluación: Muy satisfactoria. Gastón no solo cumplió con sus responsabilidades básicas, sino que también se aseguró de que el equipo trabajara de manera cohesionada y eficiente, lo que resultó en entregas puntuales y de alta calidad.

Esteban Rico - Ingeniero de datos:

Sustentación: Esteban fue clave en la construcción del modelo analítico, aportando su conocimiento técnico para asegurar la integridad y precisión de los datos. Su enfoque en la optimización de procesos resultó en un modelo robusto que apoyó efectivamente las decisiones del proyecto.

Evaluación: Excelente. La calidad del trabajo de Esteban permitió que el equipo confiara plenamente en los datos analizados, lo que es fundamental en la toma de decisiones basadas en datos.

Carlos Monroy - Ingeniero de software responsable del diseño de la aplicación y resultados:

Sustentación: Carlos lideró el diseño y la usabilidad de la aplicación, asegurándose de que fuera accesible para los usuarios finales. Además, la creación de un video para presentar los resultados fue una innovación que agregó valor al proyecto.

Evaluación: Muy satisfactoria. Carlos no solo cumplió con las expectativas de su rol, sino que también añadió un componente creativo que mejoró la presentación del proyecto.

Ana - Miembro del grupo extendido:

Sustentación: Ana participó activamente en las pruebas y la documentación técnica, asegurando que la aplicación funcionara correctamente antes del lanzamiento. Su trabajo fue esencial para mitigar riesgos de errores y mejorar la calidad del producto final.

Evaluación: Satisfactoria. Aunque su rol fue de soporte, Ana fue una pieza fundamental en las etapas finales del proyecto, garantizando que el producto final fuera de alta calidad y estuviera libre de errores importantes.

Isabella - Miembro del grupo extendido:

Sustentación: Isabella fue instrumental en ayudar con la documentación técnica y la organización de la presentación final. Su habilidad para adaptarse y apoyar en diferentes áreas permitió al equipo enfrentar eficientemente los últimos desafíos del proyecto.

Evaluación: Satisfactoria. Isabella desempeñó un papel crucial en la documentación y presentación del proyecto, áreas que a menudo son subestimadas pero vitales para la comunicación efectiva de los resultados.

Bibliografía:

Ortiz, Y. V., Martínez, E. E. V., Rogel, R. M. N., & Nechar, M. C. (s. f.). *Los stakeholders de la industria hotelera: una clasificación a partir de sus intereses ambientales.*

<https://www.redalyc.org/journal/1872/187244133008/html/>

Peralta, N., Castellaro, M., & Santibáñez, C. (2020). *El análisis de datos textuales como metodología para el abordaje de la argumentación: una investigación con estudiantes de pregrado en universidades chilenas**.

<https://www.redalyc.org/journal/2550/255066212012/html/>