

**INTELIGENCIA DE NEGOCIOS**

**Proyecto final Inteligencia de Negocios**

**G30**

Santiago Yesser Najar

Daniel Alfonso Garcia Pilimur

**Universidad de los Andes**

**Departamento de Ingeniería de Sistemas y Computación**

**Bogotá D.C., abril 6 de 2024**

**Informe Proyecto Inteligencia de Negocios**

**Entendimiento del negocio y enfoque analítico:**

|  |  |
| --- | --- |
| **Elemento** | **Descripción** |
| **Oportunidad/problema Negocio** | Mejorar la experiencia del turista en Colombia mediante soluciones analíticas avanzadas |
| **Enfoque analítico (Descripción del requerimiento desde el punto de vista de aprendizaje automático) e incluya las técnicas y algoritmos que propone utilizar** | Análisis predictivo para identificar patrones de comportamiento del turista. Se propone utilizar técnicas de clustering para segmentar a los turistas de acuerdo con sus Reviews, y algoritmos de recomendación basados en filtrado colaborativo para personalizar las sugerencias. En adición, también se va a realizar la utilización de algoritmos de Machine Learning para personalizar recomendaciones y optimizar la oferta de servicios turísticos. Por último, se propone utilizar algoritmos de clasificación para segmentar a los turistas según sus preferencias, y algoritmos de optimización para mejorar la asignación de recursos. |
| **Organización y rol dentro de ella que se beneficia con la oportunidad definida** | Agencia de viajes, Hoteles, Turistas, Atracciones turísticas, Empresas de turismo, Autoridades locales, Organizaciones de eventos. |
| **Contacto con experto externo al proyecto y detalles de la planeación** | La primera reunión fue realizada el 24 de marzo de 2024, se mantiene contacto diario con el experto externo al proyecto, es decir, con el grupo de estadística respectivo. |

**Mapa de actores relacionado con el producto de datos creado:**

|  |  |  |  |
| --- | --- | --- | --- |
| Rol dentro de la empresa | **Tipo de actor** | **Beneficio** | **Riesgo** |
| Científico de datos | Profesional especializado en análisis de datos y modelado estadístico | Desarrolla modelos predictivos para identificar patrones de comportamiento del turista, lo que permite personalizar las recomendaciones y optimizar la oferta de servicios turísticos. | El riesgo radica en la calidad de los datos utilizados para el modelado y la interpretación adecuada de los resultados |
| Ingeniero de datos | Profesional encargado de la gestión y procesamiento de datos | Garantiza la integridad y disponibilidad de los datos necesarios para el análisis y modelado. | Posibles problemas de escalabilidad, calidad de datos y seguridad de la información |
| Ingeniero Software | Profesional encargado del desarrollo de aplicaciones informáticas. | Diseña y desarrolla interfaces de usuario intuitivas para las recomendaciones personalizadas y la interacción con los servicios turísticos | Errores de implementación que podrían afectar la usabilidad y la experiencia del usuario final |
| Analista marketing | Profesional especializado en estrategias de marketing | Utiliza los insights proporcionados por el análisis de datos para diseñar campañas de marketing dirigidas a diferentes segmentos de turistas | Falta de alineación entre las estrategias de marketing y las recomendaciones generadas por el sistema, lo que podría afectar la efectividad de las campañas |
| Gestor Producto Turístico | Profesional encargado de la gestión y promoción de productos turísticos | Utiliza las recomendaciones personalizadas para mejorar la oferta de servicios turísticos y aumentar la satisfacción del cliente | Resistencia al cambio por parte de los proveedores de servicios turísticos o dificultades en la implementación de las recomendaciones |

**Trabajo en equipo:**

* **Líder del proyecto:** Daniel Garcia, el estudiante realiza las reuniones con el grupo de Inteligencia de Negocios y con el grupo respectivo de Estadística. Las reuniones fueron programadas para que fueran realizadas cada lunes de la semana para poder mostrarle al grupo de Estadística los avances que se iban dando con el código del proyecto y con los resultados de este. Además, el estudiante realiza dos algoritmos, Random Forest y Nayive Bayes para poder obtener las posibles conclusiones con los cuales se realizan los resultados del proyecto.
* **Líder del negocio:** Daniel Garcia, el estudiante fue el encargado de velar por la solución del problema y de poder identificar las distintas formas con las cuales abordar la problemática y poder mostrarle al cliente las posibles soluciones que se tienen sobre el proyecto. El estudiante realizó el contacto con los estudiantes de Estadística y establecieron que los resultados serán mostrados de una forma clara y sencilla, en donde los estudiantes de Estadística logren entender los resultados que fueron obtenidos, y además, reconozcan que los algoritmos usados son necesarios seguirlos entrenando para que los índices de precisión suban su valor, dado que todavía en una primera iteración el modelo tiene índices de precisión bajos y es necesario seguir iterando el modelo para que reconozca más valores y aumente la precisión del mismo.
* **Líder de datos:** Santiago Najar, el estudiante fue el encargado de realizar la limpieza de datos y poner a disposición del grupo de Inteligencia de negocios los datos para proceder a realizar los algoritmos y que sean analizados por parte del grupo de Inteligencia de Negocios. En adición, el estudiante realiza la distribución de los cargos sobre el código para poder reconocer los resultados. El estudiante realiza un algoritmo el cual es Arboles de decisión, y asigna al estudiante Daniel Garcia a realizar los otros dos algoritmos, los cuales fueron mencionados en una parte más arriba en este documento, Random Forest y Nayive Bayes, los cuales se considera que son dos algoritmos que sirven para poder solucionar el problema de reconocer el texto en los Reviews de los hoteles.
* **Líder de analítica:** Santiago Najar, el estudiante fue el encargado de analizar los algoritmos realizados por el grupo y reconocer cual es el algoritmo que tiene la mayor precisión para el problema dado. Además, el estudiante realiza el análisis sobre la lectura de los Reviews, en donde se realiza la vectorización del texto y se eliminan las palabras que son consideradas Stopwords. Por último, el estudiante analiza los distintos algoritmos realizados, y reconoce que el algoritmo de Random Forest es el de mayor precisión para poder ser usado en los resultados que serán mostrados al cliente respectivo.

**Conclusiones Halladas:**

Nuestras recomendaciones estratégicas para los hoteles del sector de Comercio, Industria y Turismo de Colombia son las siguientes:

1. Priorizar la excelencia en el servicio: Los análisis demuestran que un servicio de alta calidad es el factor más decisivo en las opiniones de los usuarios. Instamos a los hoteles a concentrarse en la eficiencia y la calidad en todos los servicios ofrecidos, como la limpieza y la gastronomía. Más que agregar lujos superfluos, lo fundamental es minimizar los errores y asegurarse de que los servicios básicos se ejecuten a la perfección.

2. Optimizar la experiencia de la habitación: La habitación es un aspecto crítico en la percepción del huésped. Recomendamos que las habitaciones sean espaciosas, limpias, confortables, prácticas y modernas, ya que esto influye significativamente en la satisfacción del cliente y, por consiguiente, en sus evaluaciones.

3. Adaptar la oferta gastronómica al perfil del hotel: La comida es un elemento esencial para lograr valoraciones positivas. En lugar de optar por menús excesivamente elaborados y costosos que no siempre cumplen con las expectativas de los clientes, aconsejamos adaptar la oferta gastronómica a las capacidades y el nicho del hotel. Para hoteles de alta gama, se deben ofrecer platos de calidad premium que justifiquen su precio. En cambio, los hoteles más económicos podrían beneficiarse de incluir opciones de comida rápida bien ejecutadas y asequibles que satisfagan a sus clientes.

Con estos ajustes, los hoteles podrán mejorar las calificaciones, aumentar la satisfacción de sus huéspedes y potenciar su competitividad en el mercado turístico.

De requerirse más información, anexaremos una gráfica donde se muestra las 20 palabras o temas más influyentes que determinan cómo los turistas califican su experiencia en los hoteles y lugares turísticos, basado en las reseñas. Aquí te explico qué significa cada parte:

Barras más largas: Las palabras en la parte superior con barras más largas, como "excelent" (excelente) y "mas" (más), son los factores que más impactan positivamente las opiniones de los huéspedes. Por ejemplo, si una reseña tiene la palabra "excelente", es muy probable que la calificación sea alta.

Barras intermedias: Las palabras en el medio, como "buen" (bueno) y "habit" (habitación), son importantes, pero tienen un impacto menor en comparación con las primeras. Aun así, son esenciales para una experiencia positiva.

Barras más cortas: Las palabras con barras más cortas, hacia el final, como "pued" (puede) y "ciud" (ciudad), aunque tienen influencia, son menos decisivas en la percepción global del turista.

