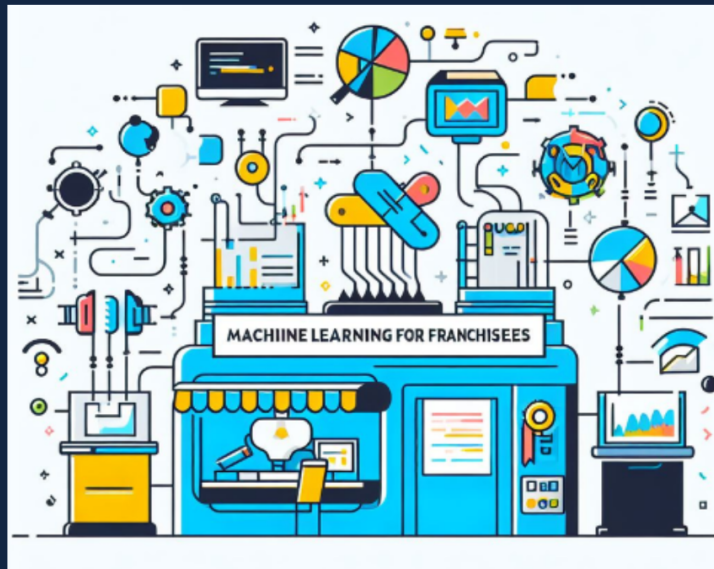


RECOMMENDATION SYSTEM: FAST FOOD FRANCHISES



MINDFUL DATA

INFORME 3

Indice

Indice.....	2
Introducción.....	3
Etapas de trabajo.....	3
Diagrama de Gantt - Semana 3.....	4
Implementación en la nube: Carga inicial.....	4
Implementación en la nube: Carga incremental.....	4
Implementación en la nube: Automatización.....	5
KPIs.....	5
Panel de control.....	7
Conclusiones - Dashboard.....	8
Modelos de ML.....	9
Modelo de ML en producción.....	9
Selección del modelo, feature engineering.....	9
Recomendaciones.....	9

Introducción

El objetivo de este informe es presentar los resultados del proyecto de desarrollo e implementación de un sistema de recomendación y un dashboard personalizado para el cliente (inversor). Este proyecto se realizó con el fin de ofrecer soluciones innovadoras que faciliten la toma de decisiones estratégicas basadas en datos y tecnología.

En este informe se describen las principales etapas del proyecto, desde la planificación hasta el despliegue. Se muestra el Diagrama de Gantt que ilustra los hitos alcanzados durante la Semana 3 del proyecto, así como las implementaciones en la nube que permitieron la carga y automatización de los datos. Se explica el diseño del Panel de Control, que integra los Indicadores Clave de Rendimiento (KPIs) y proporciona una visión detallada de las inversiones. Se analizan los modelos de Machine Learning, que se utilizan para generar recomendaciones personalizadas y optimizar las estrategias de inversión futuras.

Etapas de trabajo

En línea con el informe anterior, nos encontramos en la última semana de trabajo correspondiente a la etapa de Data Analytics y Machine Learning.

Durante esta etapa se presenta el dashboard y el storytelling correspondiente a los KPIs planteados en la puesta en marcha del proyecto. Además, se presenta el modelo de Machine Learning y los entregables finales.

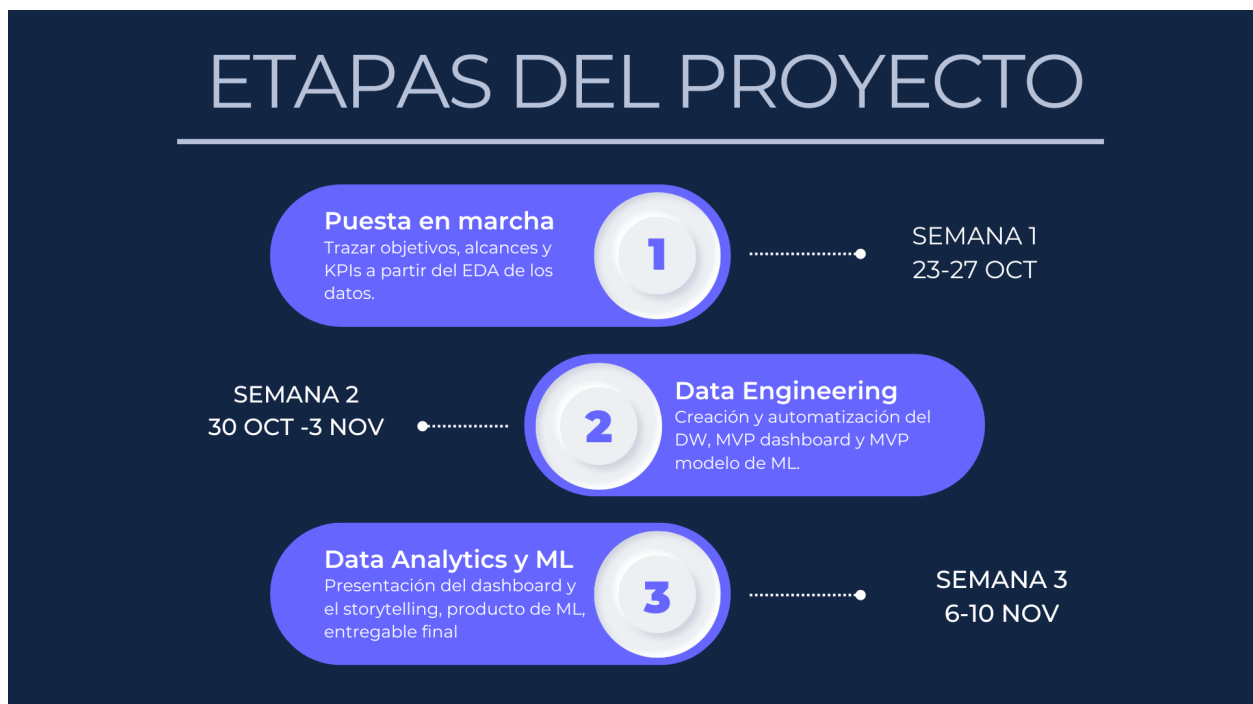
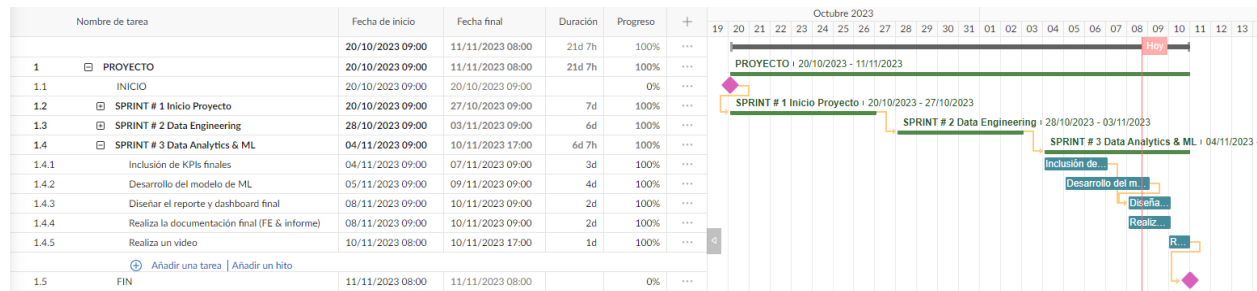


Diagrama de Gantt - Semana 3

A continuación se presenta el diagrama de Gantt con las actividades planificadas para la tercera semana de trabajo.



Además compartimos el enlace para acceder al diagrama de Gantt en la aplicación GanttPro:

[Diagrama de Gantt](#)

Implementación en la nube: Carga inicial

Después de crear el Data Warehouse, el siguiente paso crucial es realizar la carga inicial de datos. Esto implica mover grandes volúmenes de información desde diferentes fuentes hacia el Data Warehouse. Durante esta etapa, es esencial diseñar cuidadosamente el proceso de extracción, transformación y carga (ETL) para garantizar la calidad y la integridad de los datos.

La carga inicial puede ser un proceso intensivo, ya que estamos transfiriendo una gran cantidad de datos en una sola vez. Sin embargo, establecer una base sólida es fundamental para asegurar que los análisis posteriores sean precisos y significativos.

Implementación en la nube: Carga incremental

Los datos cambian constantemente y es crucial mantener nuestro Data Warehouse actualizado para obtener insights en tiempo real. Aquí es donde entra en juego la carga incremental.

La carga incremental implica identificar y transferir solo los datos nuevos o modificados desde las fuentes de origen al Data Warehouse. Esto reduce el tiempo y los recursos necesarios para mantener los datos actualizados.

Implementación en la nube: Automatización

Para la automatización del proceso de ETL, implementamos Google Cloud Functions, al ingresar un archivo al Bucket indicado se inicia el proceso de ETL, que incluye la validación, transformación, limpieza y carga de datos al Data Warehouse en Big Query.

KPIs

Para la creación del dashboard se retoman los Indicadores clave de rendimiento (KPIs) presentados la semana anterior. Cada KPI presenta su métrica, su objetivo y su fórmula para lograr calcularlo.

- ❖ Crecimiento de Reseñas Positivas: Este KPI se centra en el aumento porcentual en el número de reseñas positivas en comparación con el año anterior
Objetivo: Aumento del 5% en el número de reseñas positivas de Yelp y Google para los negocios de Fast food en comparación al año anterior.

Fórmula: $(\text{Cantidad de reseñas positivas año actual} - \text{Cantidad de reseñas positivas año anterior}) / (\text{reseñas anterior})$

- ❖ Disminución de Reseñas Negativas: Este KPI se centra en la disminución porcentual en el número de reseñas positivas en comparación con el año anterior.
Objetivo: Disminución del número de reseñas negativas del 5% de Yelp y Google para los negocios de Fast food es menor en comparación al año anterior .

Fórmula: $(\text{Cantidad de reseñas negativas año anterior} - \text{Cantidad de reseñas negativas año actual}) / (\text{reseñas anterior})$

- ❖ Aumento de la tasa anual de retención de usuarios: Mide la tasa de usuarios que escriben reseñas año a año.
Objetivo: Aumentar la tasa anual de usuarios que dejan reseñas en un 5% en comparación al año anterior.

Fórmula: $(\text{Usuarios que dejan reseñas en el año actual} - \text{usuarios que dejan reseñas en el año anterior}) / \text{usuarios que dejan reseñas en el año anterior}$

- ❖ Índice de satisfacción de usuarios de los principales negocios de fast food (según Yelp y Google) para recomendar una cadena de fast food: Mide la satisfacción global de los usuarios con un negocio, combinando las calificaciones y ponderaciones de las reseñas de los usuarios en las plataformas Yelp y Google.

Objetivo: Aumentar el Índice de Satisfacción del Usuario en un 3% durante el próximo año.

Fórmula:

INFORME DE AVANCE: PROYECTO FINAL - SEMANA 2

Índice de Satisfacción del Usuario = $(\text{Ponderación_Yelp} * \text{Rating_Yelp} + \text{Ponderación_Google} * \text{Rating_Google}) / (\text{Ponderación_Yelp} + \text{Ponderación_Google})$

- Ponderación_Yelp: Peso en proporción al total (1- total google)
- Rating_Yelp: Puntuación promedio de Yelp para un restaurante o categoría de restaurantes.
- Ponderación_Google: Peso en proporción del total (1- total yelp)
- Rating_Google: Puntuación promedio de Google para un restaurante o categoría de restaurantes.

❖ **Aumento de la cantidad de reseñas por restaurante fast-food:** Mide la interacción en el restaurante (o restaurantes, si tiene varias sucursales). Suma el total de reseñas que se obtuvieron.

Objetivo: Elevar la cantidad de reseñas por restaurante por lo menos un 15% en comparación con el año anterior.

Fórmula: $(\text{Cantidad de reseñas año anterior} - \text{Cantidad de reseñas año actual}) / (\text{reseñas anterior})$

KPI	Descripción	Fórmula	Objetivo
Crecimiento de reseñas positivas	Este KPI se centra en el aumento porcentual en el número de reseñas positivas en comparación con el año anterior	$[(\text{Cantidad de Reseñas Positivas en el Año Actual} - \text{Cantidad de Reseñas Positivas en el Año Anterior}) / \text{Cantidad de Reseñas Positivas en el Año Anterior}] \times 100$	Aumento del 5% en el número de reseñas positivas para los negocios de Fast food en comparación al año anterior.
Disminución de reseñas negativas	Este KPI se centra en la disminución porcentual en el número de reseñas positivas en comparación con el año anterior	$[(\text{Cantidad de Reseñas negativas en el Año Anterior} - \text{Cantidad de Reseñas negativas en el Año Actual}) / \text{Cantidad de Reseñas Negativas en el Año Anterior}] \times 100$	Disminución del 5% en el número de reseñas negativas de los negocios de Fast food es menor en comparación al año anterior .

INFORME DE AVANCE: PROYECTO FINAL - SEMANA 2

Aumento de la tasa anual de retención de usuarios	Mide la tasa de usuarios que escriben reseñas año a año.	(Usuarios que dejan reseñas en el año actual - usuarios que dejan reseñas en el año anterior) / usuarios que dejan reseñas en el año anterior.	Aumentar la tasa anual de usuarios que dejan reseñas en un 5% en comparación al año anterior.
Satisfacción de Clientes en Fast Food	Mide la satisfacción global de los usuarios con un negocio, combinando las calificaciones y ponderaciones de las reseñas de los usuarios en las plataformas Yelp y Google.	Índice de Satisfacción del Usuario = $\frac{\text{Ponderación_Yelp} * \text{Rating_Yelp} + \text{Ponderación_Google} * \text{Rating_Google}}{\text{Ponderación_Yelp} + \text{Ponderación_Google}}$	Aumentar el Índice de Satisfacción del Usuario en un 3% durante el próximo semestre
Aumentar la cantidad de reseñas por restaurante	Mide la interacción en el restaurante (o restaurantes, si tiene varias sucursales). Suma el total de reseñas que se obtuvieron.	(Cantidad de reseñas año anterior - Cantidad de reseñas año actual) / (reseñas anterior).	Eleva la cantidad de reseñas por restaurante por lo menos un 15% en comparación con el año anterior.

Panel de control

A continuación presentamos el dashboard de visualizaciones para un inversor de comida rápida en el estado de Florida. Este tablero está diseñado para proporcionar a los inversores una visión clara y concisa del mercado de comida rápida en Florida, lo que les permitirá tomar decisiones informadas y estratégicas para su inversión en este próspero sector.

El sector de comida rápida en Florida es dinámico y de rápido crecimiento, con una población diversa y en constante movimiento. Nuestro dashboard ofrece un resumen completo de datos clave, que incluye:

- ❖ **Ubicaciones Potenciales:** Localiza las áreas con mayor demanda y competencia en la industria de comida rápida, lo que te ayudará a identificar las ubicaciones óptimas para tus negocios.
- ❖ **Analiza las preferencias de restaurantes, tipos de comida rápida, y tendencias para adaptar tu oferta a las necesidades cambiantes de los consumidores.**
- ❖ **Indicadores Clave de Desempeño:** este dashboard presenta indicadores clave de desempeño que te ayudarán a evaluar el éxito de tus inversiones:
- ❖ **Historial de Datos y Previsión:** accede a datos históricos y previsiones para tomar decisiones basadas en tendencias pasadas y futuras.

A continuación se presenta el dashboard:

Se puede además acceder al dashboard interactivo, a través del siguiente enlace:

[Panel de Control - Tableau Cloud.](#)

Conclusiones - Dashboard

Ocurrieron en total 725334 reseñas

32,5 % fueron reseñas de Yelp y 67.5% de Google Maps

El 75.46% de las reseñas resultan en reseñas positivas, el 14,22% fueron reseñas calificadas como neutras y el 10,32% fueron negativas.

- ❖ Aumento del 5% en el número de reseñas positivas de Yelp y Google para los negocios de Fast food en comparación al año anterior.
Podemos evidenciar como en los últimos 5 años se evidencia un aumento de las reseñas positivas realizadas por usuarios del estado de Florida. Durante los años se cumple con el aumento del 5%, salvo en el año 2020 donde ocurre un fuerte descenso posiblemente asociado al año donde ocurre el Covid-19 y es posible que el confinamiento sea la causa de esta tendencia.
- ❖ Disminución del número de reseñas negativas del 5% de Yelp y Google para los negocios de Fast food es menor en comparación al año anterior .
Podemos evidenciar como en los últimos 5 años se evidencia un aumento en vez de una disminución de de las reseñas negativas realizadas por usuarios del estado de Florida, posiblemente asociada a un aumento del número de reseñas totales. Al igual que para las reseñas positivas, se evidencia un descenso en el año 2020, posiblemente asociado al año donde ocurre el Covid-19 y es posible que el confinamiento sea la causa de esta tendencia.
- ❖ Aumentar el Índice de Satisfacción del Usuario en un 3% durante el próximo semestre.
Fórmula:
$$\text{Índice de Satisfacción del Usuario} = (\text{Ponderación_Yelp} * \text{Rating_Yelp} + \text{Ponderación_Google} * \text{Rating_Google}) / (\text{Ponderación_Yelp} + \text{Ponderación_Google})$$
 - Ponderación_Yelp: Peso en proporción al total (1- total google)
 - Rating_Yelp: Puntuación promedio de Yelp para un restaurante o categoría de restaurantes.
 - Ponderación_Google: Peso en proporción del total (1- total yelp)
 - Rating_Google: Puntuación promedio de Google para un restaurante o categoría de restaurantes.

Podemos evidenciar como hay una tendencia al aumento del índice de satisfacción del usuario. Como las tendencias de satisfacción se calculan en función del rating del

restaurante pero para cada reseña, podemos evidenciar un leve aumento año a año, pero que no alcanza el 3% (posiblemente por el número de reseñas tan alto).

- ❖ Aumentar la tasa anual de usuarios que dejan reseñas en un 5% en comparación al año anterior.

Fórmula: $(\text{Usuarios que dejan reseñas en el año actual} - \text{usuarios que dejan reseñas en el año anterior}) / \text{usuarios que dejan reseñas en el año anterior}$.

Podemos ver como el aumento del 5% de la tasa anual de usuarios que dejan reseñas tiene aumentos que superan el 5% para los períodos: 2017-2018, 2018-2019. Para el período 2019-2020 hay un claro descenso posiblemente provocado por el confinamiento a causa de la pandemia del Covid-19. Además, para el período 2020-2021 se comienza a observar una recuperación de las tendencias aún en ascenso aunque no alcanza el 5%.

- ❖ Elevar la cantidad de reseñas por restaurante por lo menos un 15% en comparación con el año anterior.

Fórmula: $(\text{Cantidad de reseñas año anterior} - \text{Cantidad de reseñas año actual}) / (\text{reseñas anterior})$

Podemos evidenciar que el aumento del 15% sólo se cumple para el período 2017-2018, luego el aumento ha quedado por debajo del 15% . Nuevamente se presenta una disminución marcada en el año 2020 posiblemente relacionada con el confinamiento provocado por la pandemia Covid-19.

Modelo de ML

Desarrollo del Sistema de Recomendación de Franquicias de Comida Rápida

1. Introducción:

El presente informe detalla el desarrollo, proceso, objetivos, infraestructura e implementación del Sistema de Recomendación de Franquicias de Comida Rápida en el Estado de Florida. Este sistema tiene como objetivo proporcionar recomendaciones personalizadas a inversores interesados en adquirir franquicias, teniendo en cuenta las preferencias de los usuarios y las tendencias en las reseñas de locales de comida rápida.

2. Objetivos del Sistema:

El Sistema de Recomendación se diseñó con los siguientes objetivos clave:

- Personalización: Proporcionar recomendaciones basadas en las preferencias de los usuarios y las tendencias de reseñas a través de los comentarios en la plataforma Yelp y Gmaps.

INFORME DE AVANCE: PROYECTO FINAL - SEMANA 2

- Eficiencia: Optimizar el tiempo y los recursos de inversión al ofrecer recomendaciones precisas y relevantes.
- Diversificación de Franquicias: Fomentar la diversificación de franquicias en el mercado, proporcionando información sobre diferentes opciones.
- Análisis de Sentimientos: Incorporar análisis de sentimientos en las reseñas para comprender las experiencias de los usuarios.

3. Desarrollo del Proceso:

El desarrollo del sistema involucró varias etapas clave:

- Adquisición de Datos: Recopilación de datos a través del servicio de Cloud Storage de Google.
- Procesamiento de Texto: Limpieza y procesamiento de los textos de reseñas, incluida la eliminación de caracteres especiales y emojis.
- Análisis de Sentimientos: Implementación de un análisis de sentimientos para evaluar la positividad/negatividad de las reseñas.
- Sistema de Recomendación: Desarrollo de algoritmos de recomendación basados en preferencias de usuario y análisis de sentimientos.
- Integración con Información de Franquicias: Integración de datos sobre franquicias, incluyendo nombres, ubicaciones y calificaciones promedio.

4. Infraestructura:

La infraestructura del sistema se basa en tecnologías avanzadas para garantizar la eficiencia y la escalabilidad:

- Python: Utilizado para el desarrollo de algoritmos, análisis de sentimientos y procesamiento de datos a través de Kaggle.
- Pandas y NumPy: Librerías de Python utilizadas para el manejo eficiente de datos y cálculos numéricos.
- Scikit-learn: Empleado para implementar algoritmos de aprendizaje automático para el sistema de recomendación.
- Google Cloud Platform (GCP): Utilizado para el almacenamiento de datos y la implementación en Cloud Storage.

- BigQuery: Plataforma de análisis de datos en la nube de Google utilizada para el almacenamiento y procesamiento eficiente de grandes conjuntos de datos.
- Streamlit: Se usó para la creación de la interfaz de usuario interactiva. Esta herramienta facilitó la presentación y visualización de los resultados de manera amigable para el usuario.

5. Implementación:

Se llevó a cabo en las siguientes fases:

- Desarrollo del Modelo de Recomendación: Creación y validación de modelos de recomendación utilizando datos históricos.
- Implementación del Análisis de Sentimientos: Desarrollo e integración de un módulo de análisis de sentimientos.
- Interfaz de Usuario: Creación de una interfaz de usuario intuitiva para que los inversores ingresen sus preferencias y reciban recomendaciones.
- Pruebas y Ajustes: Realización de pruebas exhaustivas y ajustes basados en la retroalimentación.

6. Resultados y Beneficios:

El Sistema de Recomendación ha demostrado ser eficaz al proporcionar recomendaciones precisas y personalizadas. Algunos de los beneficios claves incluyen:

- Optimización de Inversiones: Los inversores pueden tomar decisiones informadas sobre la adquisición de franquicias, maximizando el rendimiento de sus las mismas.
- Diversificación del Mercado: Fomenta la diversificación del mercado de franquicias, ofreciendo opciones más allá de las marcas convencionales.
- Mejora Continua: La retroalimentación de usuarios y datos en tiempo real permiten mejoras continuas en el sistema.

Recomendaciones

- *De acuerdo a los resultados obtenidos en el panel de control, el cliente debe tomar la siguiente recomendación:*
Para anticiparse a los cambios del mercado y las situaciones imprevistas, se sugiere al inversor, adoptar una actitud dinámica y proactiva. Asimismo, debe enfocarse en estrategias específicas que mejoren la experiencia del cliente, para aquellas franquicias con tendencias positivas, las cuales pretende adquirir. Es esencial participar activamente en programas de retención y evaluar constantemente las estrategias para optimizar el rendimiento a largo plazo de la cartera, asegurando así una adaptabilidad continua y la maximización de oportunidades. En efecto, el dashboard presentado es una herramienta útil para el seguimiento del comportamiento de la satisfacción del usuario.
- *De acuerdo a los resultados obtenidos del sistema de recomendación:* El Sistema de Recomendación de Franquicias de Comida Rápida ha sido un proyecto exitoso que ha cumplido con los objetivos establecidos. Las futuras mejoras pueden incluir la expansión geográfica, la incorporación de más fuentes de datos y la implementación de técnicas avanzadas de aprendizaje automático.