

# INFORME DE PROOF OF CONCEPT (PoC) - DATAGUSTO

## Resumen Ejecutivo:

DATAGUSTO ha desarrollado un Proof of Concept (PoC) para dos sistemas de Machine Learning destinados a mejorar la experiencia del usuario y proporcionar insights valiosos para las empresas del sector gastronómico en Florida. Este informe detalla el proceso de desarrollo, desde la definición del problema hasta la implementación y evaluación de los modelos.

## 1. Introducción:

DATAGUSTO se enfrentaba al desafío de optimizar las recomendaciones de restaurantes para los usuarios y proporcionar herramientas predictivas para las empresas del sector gastronómico. Para abordar este desafío, se propuso un PoC que consta de dos partes: un sistema de recomendaciones para usuarios y un sistema de predicción de competencia para el conglomerado gastronómico.

## 2. Definición del Problema y Objetivos:

- **Sistema de Recomendaciones para Usuarios:**
  - **Problema:** Mejorar la experiencia del usuario al proporcionar recomendaciones personalizadas de restaurantes.
  - **Objetivo:** Sugerir restaurantes basados en las experiencias previas de los usuarios.
- **Sistema de Predicción de Competencia para el Conglomerado:**
  - **Problema:** Predecir la competencia para las empresas del sector gastronómico.
  - **Objetivo:** Proporcionar insights para entender la posición competitiva de las empresas y planificar estrategias.

## 3. Recolección y Preparación de Datos:

- Se utilizaron datos de las tablas proporcionadas: **review**, **user**, **business**.
- Los datos se limpiaron y unificaron en un DataFrame para facilitar el análisis y modelado.

#### 4. Sistema de Recomendaciones para Usuarios:

- Se utilizó Collaborative Filtering (Filtrado Colaborativo) con el algoritmo SVD (Singular Value Decomposition) para entrenar el modelo.
- Se generaron recomendaciones para los usuarios basadas en las predicciones de calificación de restaurantes no visitados.
- Se implementó una API para que los usuarios reciban recomendaciones personalizadas.

#### 5. Sistema de Predicción de Competencia:

- Se identificaron variables relevantes como **stars**, **review\_count**, **categories**, y geolocalización para predecir la competencia.
- Se entrenó un modelo de clasificación RandomForestClassifier para predecir si un negocio es competidor.
- Se implementó una API para proporcionar predicciones de competencia a las empresas del conglomerado gastronómico.

#### 6. Implementación y Despliegue (Planificación Futura):

En esta etapa, DATAGUSTO planea implementar y desplegar el prototipo funcional desarrollado en Google Cloud Platform, específicamente utilizando Vertex AI para el desarrollo de modelos y Streamlit para la creación de la interfaz de usuario. Se espera que esta implementación incluya los siguientes pasos:

- **Desarrollo del Prototipo Funcional:**
  - Se completará el desarrollo de una interfaz web intuitiva utilizando Streamlit, que permita a los usuarios recibir recomendaciones personalizadas de restaurantes y a las empresas acceder a predicciones de competencia.
  - Se garantizará que la interfaz sea altamente interactiva y visualmente atractiva, utilizando widgets y gráficos para mejorar la experiencia del usuario.
- **Creación de Modelos con Vertex AI:**
  - Se utilizará Vertex AI para desarrollar y entrenar los modelos de recomendaciones para usuarios y predicción de competencia para el conglomerado gastronómico.
  - Se aprovecharán las capacidades de Vertex AI para automatizar el proceso de entrenamiento y optimización de los modelos, así como para gestionar el ciclo de vida de los mismos.
- **Despliegue en Google Cloud Platform:**
  - Se desplegará el prototipo funcional en Google Cloud Platform, utilizando servicios como Compute Engine para alojar la aplicación web y Vertex AI para servir los modelos de Machine Learning.

- Se aprovechará la escalabilidad y la fiabilidad de Google Cloud Platform para garantizar un despliegue robusto y de alto rendimiento del sistema.
- **Integración de Botones en Streamlit:**
  - Se implementarán botones y controles interactivos en la interfaz de usuario desarrollada con Streamlit, permitiendo a los usuarios realizar acciones como solicitar recomendaciones personalizadas o acceder a análisis detallados.
  - Se aprovecharán las capacidades de Streamlit para facilitar la integración de componentes interactivos y mejorar la usabilidad de la aplicación.
- **Pruebas y Validación:**
  - Se llevarán a cabo pruebas exhaustivas para asegurar que el prototipo funcional cumpla con los requisitos de calidad y rendimiento del cliente.
  - Se realizarán pruebas de integración para garantizar la interoperabilidad entre los diferentes componentes del sistema y su correcto funcionamiento en conjunto.

## **7. Documentación y Presentación:**

- Se documentará detalladamente el proceso de desarrollo, incluyendo los pasos, decisiones y resultados obtenidos.
- Se prepara una presentación que muestra gráficos y métricas de rendimiento de los modelos, junto con una demostración funcional del prototipo.

## **8. Conclusiones:**

El PoC desarrollado por DATAGUSTO demuestra la viabilidad técnica de los sistemas de recomendaciones para usuarios y predicción de competencia para el conglomerado gastronómico. Estos sistemas tienen el potencial de mejorar la experiencia del usuario y proporcionar insights valiosos para las empresas del sector.

## **9. Próximos Pasos:**

Los próximos pasos incluyen la iteración y mejora continua de los modelos, la incorporación de feedback de usuarios y stakeholders, y la implementación de funcionalidades adicionales según sea necesario. Así como su 'Implementación y Despliegue (Planificación Futura)'.