



## DATATHON 2024

# PROPUESTA FINAL

### **Analytic Avengers:**

Ana Paola Almeida Pérez | A00833937

Viviana Alanis Fraige | A01236316

Cynthia Cristal Quijas Flores | A01655996

David Antonio Matamoros Alvarado | A01622865



# Contenido

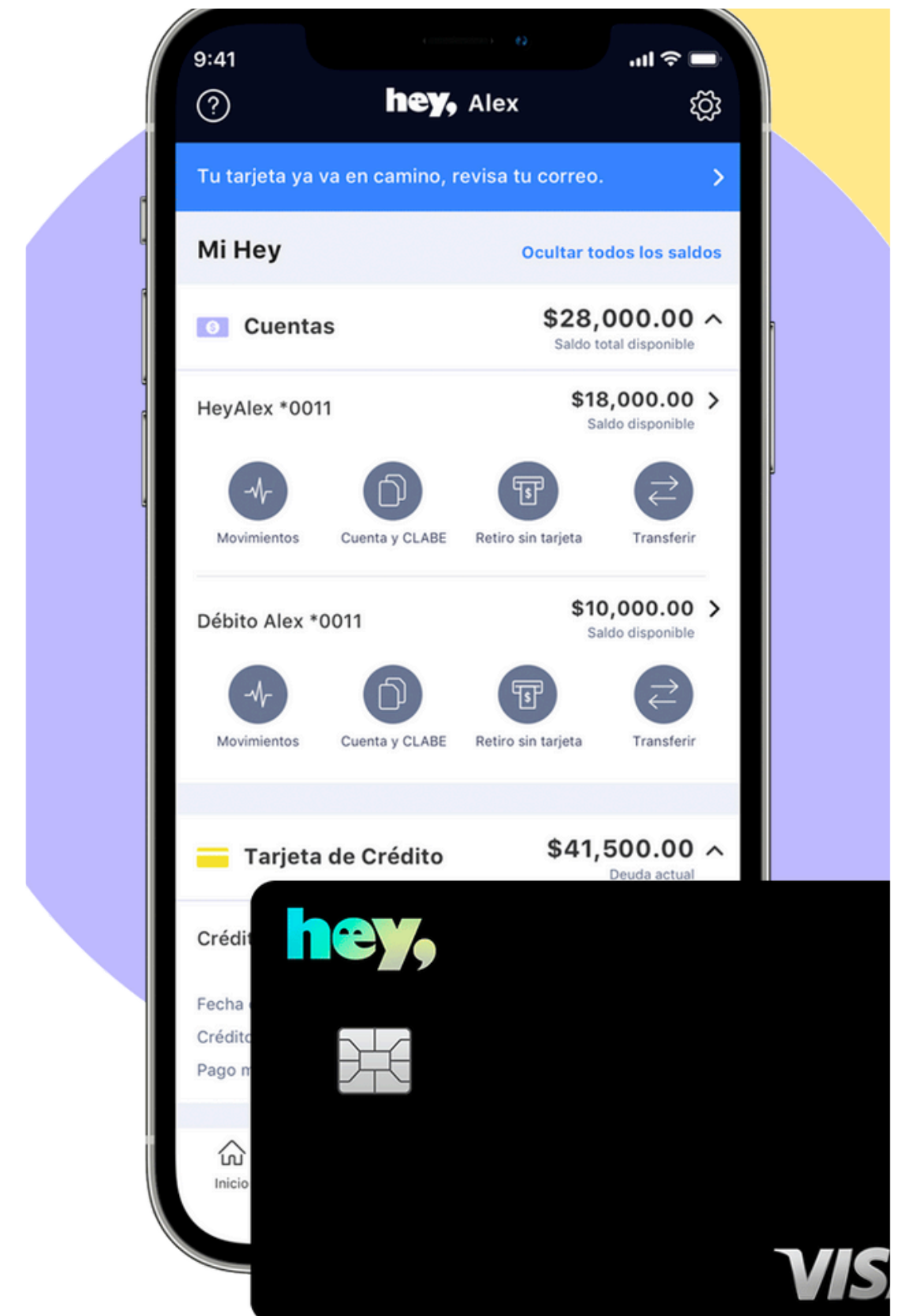


01.	Problemática
02.	Propuestas descartadas
03.	Propuesta final

04.	Implementación
05.	¿Por qué nosotros?

# Problemática

Automatizar el procesamiento de bases NLP para que se obtengan insights de la interacción de los usuarios en "X" con ayuda de modelos de machine learning



# Propuestas descartadas

01.

Uso de método topológico a NLP tokenizado para observar la relación entre un grupo de palabras para ver como se relacionan entre si.

02.

Aplicación de método Random Forest (aprendizaje supervisado).

03.

Implementación de nuevas características a la base de datos para crear un modelo predictivo de necesidades del usuario en base a su perfil.



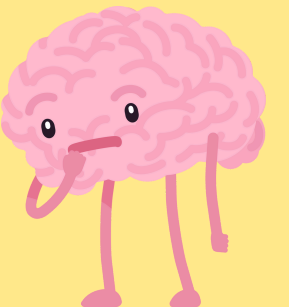
# Propuesta final

Creación de una interfaz de uso interno para analizar el comportamiento del servicio al cliente de los usuarios en Twitter o 'X' utilizando el algoritmo Tfidf Vectorizer.

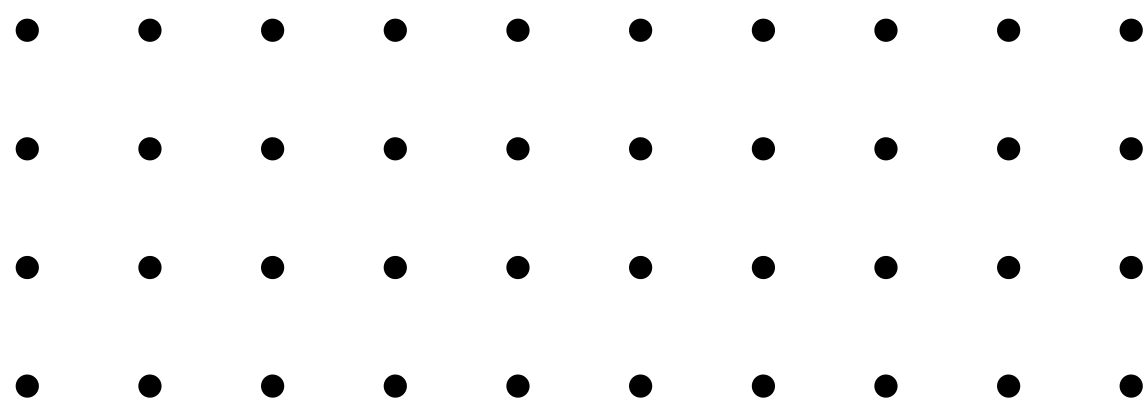
- Limpieza de datos
- Análisis de sentimientos (NLTK)
- Modelo NLP de clasificación
- Interfaz

**¿A quién ayuda esto?**

- Business Analyst
- Data Engineer
- Mercadotecnia

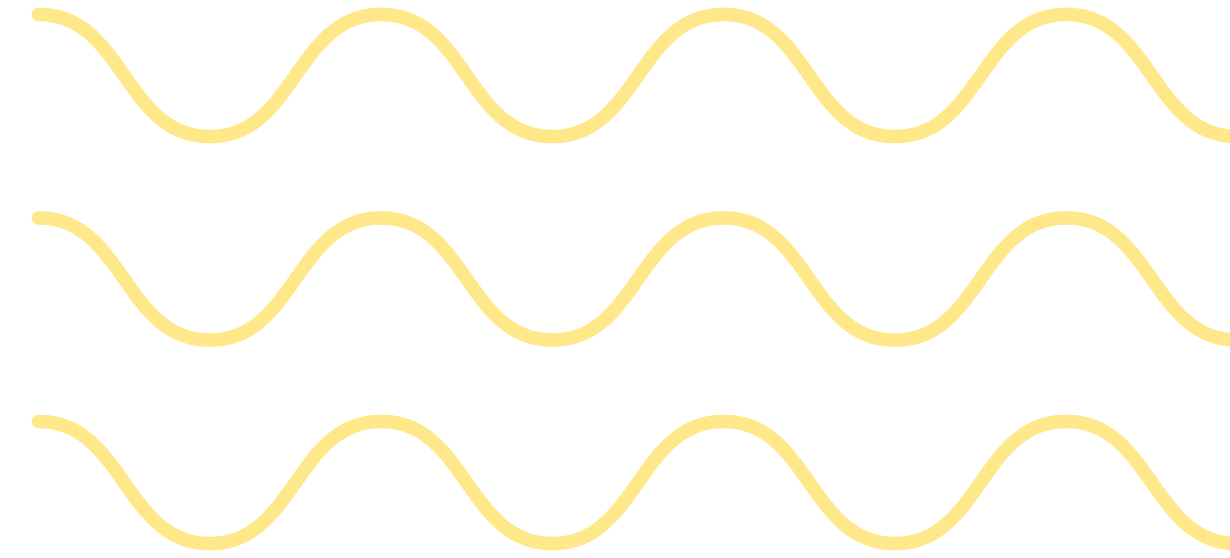
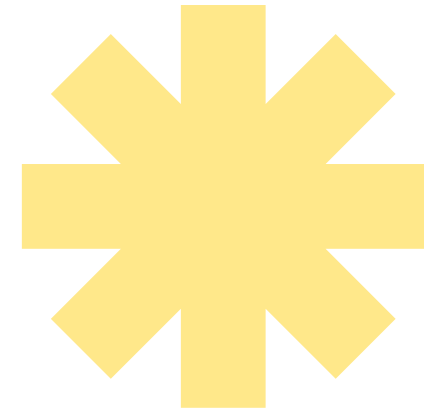


# Implementación (Interfaz)



# Mejoras

- Optimizar proceso de traducción de la base de datos
- Pasar a PRD la interfaz para uso global de usuarios internos
- A partir de la clasificación obtenida, obtener más características para hacer predicciones de las necesidades de los mismos
- Optimizar el proceso de asignar pesos a las palabras de acuerdo con su frecuencia



¡Muchas  
gracias!

