

Sede San Joaquín

| PTY4614 – Capstone |
| --- |
| **Tótem recomendador de productos con detección por cámara y búsqueda por voz.** |

Integrantes: Julián Martínez, Leonardo Pavez, Sebastián Tapia

Fecha: Agosto – 2025

# Índice

1. Abstract  
2. Desarrollo de Ingeniería  
 2.1 Descripción y relevancia  
 2.2 Relación con competencias del perfil de egreso  
 2.3 Relación con intereses profesionales  
 2.4 Factibilidad del proyecto  
 2.5 Indicadores de calidad (IL 1.5)  
3. Conclusiones individuales  
4. Reflexión  
5. Bibliografía  
6. Anexos

# 1. Abstract

## Español

El proyecto considera implementar un tótem inteligente con cámara y micrófono para tiendas de retail. El sistema detecta prendas, color dominante y rango etario. Permite búsquedas por voz para consultar disponibilidad. Con esta información genera recomendaciones personalizadas y captura métricas de interacción en un Data Warehouse. Se utiliza visión por computador, ASR/NLU, FastAPI y Flutter Web en modo kiosk. El objetivo es mejorar la experiencia del cliente y entregar insights relacionados al negocio.

## English

This project implements an intelligent in-store kiosk with camera and microphone for retail. The system detects garments, dominant color and age bucket, and enables voice search for stock availability. Using this data, it produces personalized recommendations and logs interaction metrics into a Data Warehouse. It leverages computer vision, ASR/NLU, FastAPI, and Flutter Web in kiosk mode. The goal is to enhance customer experience and deliver actionable insights for the business.

# 2. Desarrollo de Ingeniería

## 2.1 Descripción y relevancia

Ubicado en la entrada de la tienda, el tótem detecta atributos del cliente (prenda, color, rango etario) y responde consultas por voz (por ejemplo: “¿hay zapatillas rojas 42?”). Muestra recomendaciones con stock y registra las interacciones para análisis posterior. Su relevancia está en la innovación en Retail, mejorando la conversión en tienda y entregando métricas valiosas al negocio.

## 2.2 Relación con competencias del perfil de egreso

• Construcción de soluciones de software (4.1–4.3): Frontend Flutter + Backend FastAPI; integración de CV/ASR; despliegue en tótem.  
• Modelamiento e implementación de datos (3.1–3.2): Modelo OLTP + DW Kimball; ETL y consultas KPI.  
• Gestión de proyectos informáticos (2.1–2.2): Plan en cascada 14 semanas, hitos por fases, control con Jira.  
• Pruebas y certificación (1.1–1.3): Plan de pruebas API, precisión CV (F1, matriz de confusión), WER/latencia ASR, pruebas de usabilidad.

## 2.3 Relación con intereses profesionales

Los intereses del equipo están en IA aplicada, retail tech e interfaces interactivas. Este proyecto permite una experiencia real en visión por computador, voz, arquitectura de datos y BI.

## 2.4 Factibilidad del proyecto

El cronograma se organiza en 14 semanas con metodología tradicional en cascada. Se disponen recursos accesibles: mini-PC, cámara, micrófono, pantalla táctil, frameworks libres (YOLO, Whisper/Vosk, FastAPI, Flutter, PostgreSQL/Oracle). Se anticipan riesgos como precisión CV/ASR y latencia, que se mitigarán con simulaciones iniciales, optimización de modelos y pruebas de integración tempranas.

## 2.5 Indicadores de calidad (IL 1.5)

ER y DDL OLTP versionado.  
 Modelo DW Kimball con dimensiones y hechos.  
 Procesos ETL definidos.  
 Arquitectura de software documentada.  
 Contratos API con Swagger.  
 Plan de despliegue en tótem.  
 Plan de pruebas API, métricas CV y ASR.  
 Carta Gantt 14 semanas + Jira.  
 Plan de mitigación de riesgos.

# 3. Conclusiones individuales (English)

• Student A: I learned how to design the OLTP and DW models and understood the importance of data for KPIs.  
• Student B: I learned how to integrate computer vision and ASR models and saw their impact on real retail use cases.  
• Student C: I focused on frontend and BI dashboards, realizing how UX and KPIs combine to create value for the business.

# 4. Reflection (English)

This phase allowed us to align our project idea with our career competencies and professional interests. We confirmed the feasibility of building the kiosk within the timeframe using pre-trained models and open-source frameworks. Our main challenge is ensuring accuracy and latency in computer vision and ASR, but by planning simulations and incremental integration we feel confident. The feedback from phase 1 will guide improvements for the next development stages.

# 5. Bibliografía

• MediaPipe Documentation: https://developers.google.com/mediapipe  
• Ultralytics YOLOv8: https://docs.ultralytics.com  
• OpenAI Whisper ASR: https://github.com/openai/whisper  
• FastAPI Documentation: https://fastapi.tiangolo.com  
• Kimball, R. (2013). The Data Warehouse Toolkit.

# 6. Anexos

• Anexo 1: Modelo de datos ER.

<https://drive.google.com/file/d/1BWTAug1Cf_n4DSK5CYssGYe59fPGR86i/view?usp=sharing>  
• Anexo 2: Script DDL Oracle 11g.

<https://drive.google.com/file/d/1fYlMdcGvzbt8uQQ4wO64fhPp4XyyJxb6/view?usp=drive_link>  
• Anexo 3: Carta Gantt 14 semanas.

<https://drive.google.com/file/d/1cDGGZ7gjopem5v5g6Ae6dVBvADVm8c8g/view?usp=sharing>

• Anexo 4: Capturas Jira/Trello.

<https://drive.google.com/file/d/102afCLDBVHtWS7AWKazEyfJE55MSHpv6/view?usp=sharing>  
<https://drive.google.com/file/d/1DrE10gEu46RjfRXqA3zZfq-YT95_nIPr/view?usp=sharing>