

PONTIA WORLD

RECONOCIMIENTO DE EMOCIONES

Sistema avanzado de análisis de datos para la detección y clasificación de emociones en parques temáticos.



Nuestro equipo

ANAMARÍA TURDA

Project Manager

Liderazgo del Proyecto, limpieza de datos con python, modelo relacional SQL, desarrollo ML e IA generativa.

INÉS BENITO

Miembro activo

Limpieza de datos con python, modelo relacional SQL, desarrollo ML e IA generativa.

IÑIGO UGIDOS

Miembro activo

Limpieza de datos con python, modelo relacional SQL, desarrollo ML e IA generativa.

WILLIAM GANEM

Miembro activo

Limpieza de datos con python, modelo relacional SQL, desarrollo ML e IA generativa.





Objetivos

OBJETIVO GENERAL

Analizar los datos generales generados en Pontia World para detectar patrones en las emociones y usar esa información para mejorar la experiencia de los clientes y la rentabilidad del negocio.

OBJETIVOS ESPECÍFICOS

OBJETIVO 1

Desarrollar un proceso de limpieza y acondicionamiento de datos completo.

OBJETIVO 2

Diseñar e implementar un modelo de base de datos relacional que organice la información para facilitar consultas.

OBJETIVO 3

Desarrollar, entrenar y comparar diferentes modelos para clasificar emociones y una propuesta de IA Generativa.

OBJETIVO 4

Generar un dashboard de insights para presentar los datos y conclusiones.

Metodología

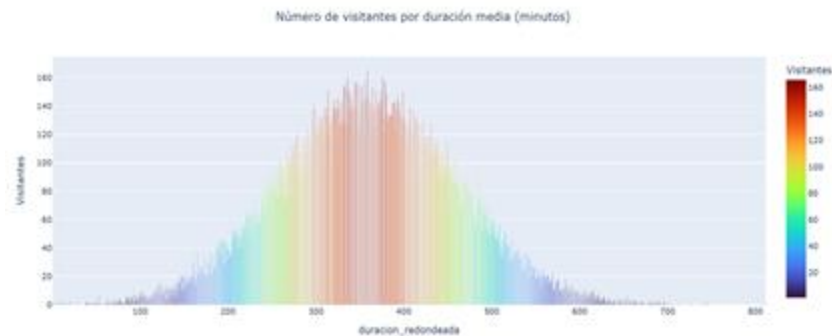
Partiendo de los datos iniciales no estructurados, se identificaron y corrigieron diversos problemas de calidad:

- ★ **FORMATO DE DATOS**
Limpieza y transformación de datos en tablas.
- ★ **VALORES NEGATIVOS**
Tiempo de espera, costes y duraciones negativas transformados a valores positivos.
- ★ **CAMPO ID VISITANTE INCONSISTENTE**
Se creó un nuevo campo id_visitante a partir de los datos de procedencia, duración y día de visita.
- ★ **VALORES EXTREMOS EN DURACIÓN**
Visitas con duración superior a 9 horas imputadas con la duración media de 350 minutos.
- ★ **VALORES ATÍPICOS EN COSTES**
Tickets con valores muy bajos revisados por tipo de entrada.



Análisis exploratorio

Se realizaron visualizaciones como mapamundi de distribución geográfica de visitantes, gráficos de número de visitantes por atracción, distribución de entradas y duraciones medias.



Distribución de Visitantes por Procedencia



Base de datos

PROCESO PREPARACIÓN SQL

★ LIMPIEZA Y TRANSFORMACIÓN

Limpieza y transformación de datos para cargarlos en SQL.

★ SOLUCIÓN DE ERRORES TÉCNICOS

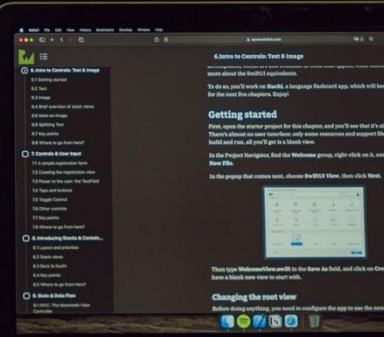
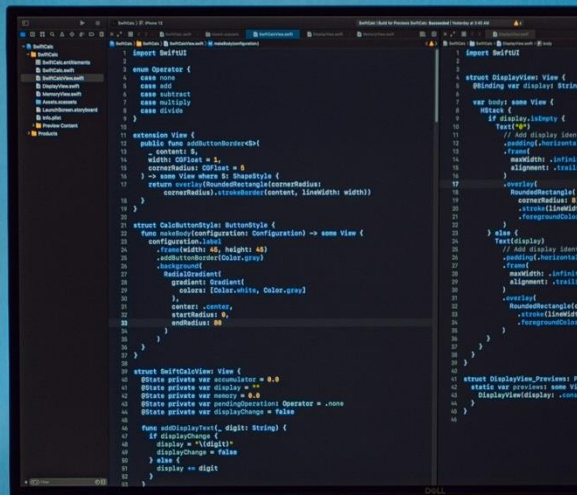
Solución de errores de codificación + validación de tipos + problemas de carga de datos.

★ DEFINICIÓN DE REGLAS DE NEGOCIO

Definición de reglas de negocio para imputar valores anómalos (duración > 540min = 350 min, fast pass comprados con > 3 días de antelación = Pase rápido erróneo).

★ PREPARACIÓN FINAL

Uso de tablas con datos limpios para la realización de consultas de KPIs.



Modelo relacional

Modelo de datos

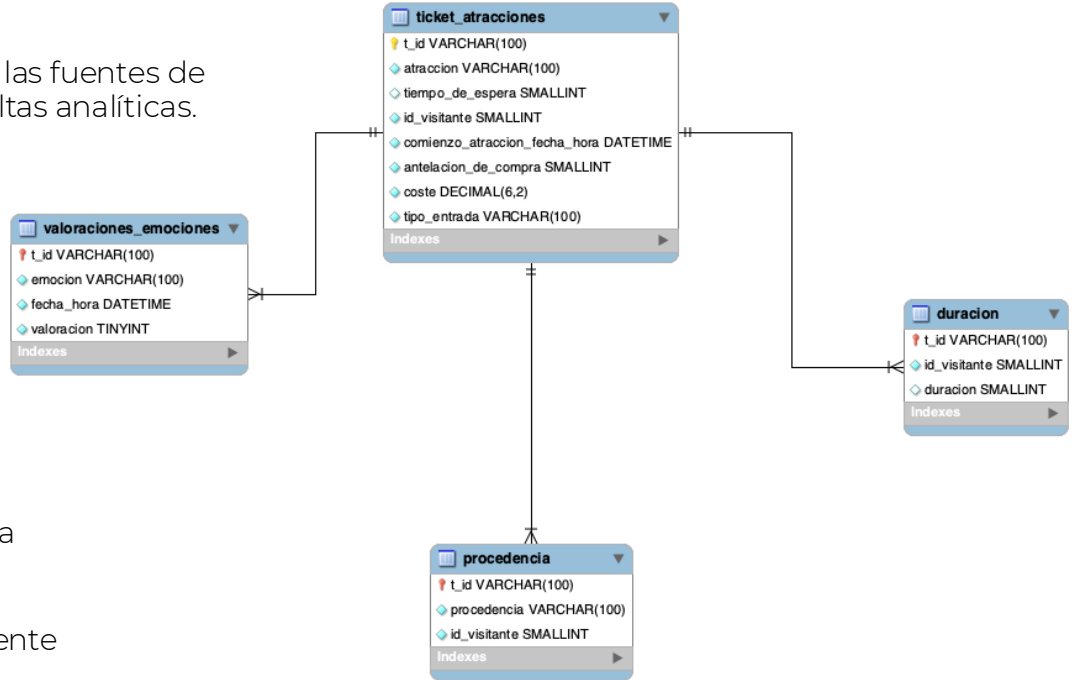
Se diseñó un modelo relacional que conecta las fuentes de datos y tablas principales para facilitar consultas analíticas.

Sistema de Rankings

Se implementó un sistema de ranking para manejar empates en los datos y obtener una mejor visión de conjunto.

Conclusiones

- ★ Los visitantes subían como máximo a 7 atracciones, y como mínimo a 1.
- ★ La emoción más frecuente es "feliz" para todas las atracciones.
- ★ Todas las atracciones tienen prácticamente la misma media de valoración (5/10).



Insights destacados

- ★ El 90,74% de los visitantes se han montado en solo una atracción durante su estancia en el parque.
- ★ La media de valoraciones no varía significativamente por cada país (5/10).
- ★ Las entradas tienen el mismo precio independientemente de la antelación de su compra (15-18€).
- ★ El tiempo de espera no afecta positiva o negativamente a las valoraciones.
- ★ El tiempo de espera es prácticamente igual para todas las atracciones (12-13 min).



Detección de emociones

OBJETIVO

Desarrollar un sistema que clasifique 7 emociones a partir de imágenes faciales pequeñas en blanco y negro.

ANGRY

DISGUST

FEAR

HAPPY

NEUTRAL

SURPRISE

SAD

PREPARACIÓN DE DATOS

- ✓ División de datos para entrenar el modelo y probarlo.
- ✓ Aumentación moderada de imágenes para incrementar el número de fotos por emoción (rotación, traslación, flip, zoom).
- ✓ Normalización de las imágenes para agilizar el procesamiento en el modelo.

MÉTRICAS PRINCIPALES Y CRITERIOS DE EVALUACIÓN

★ KPI principal — Macro-F1

Promedio de la capacidad de acierto y fallos por clase.

★ KPI de referencia — Accuracy

Número de aciertos del modelo general, sin diferenciar por clases.



Evolución de los modelos

DEL PUNTO DE PARTIDA AL MEJOR MODELO

★ FASE 0 – Aprendizaje a partir de un modelo ya entrenado

Empezamos con un modelo ya entrenado (MobileNet), pero no se adaptó bien a nuestras imágenes pequeñas. Resultado: ~48 % de aciertos.

★ FASE 1 – Nuestro primer modelo propio

Creamos desde cero una red sencilla, como una versión básica de la anterior. Esto permitió que el modelo “aprendiera por sí mismo” las emociones. Resultado: mejoró hasta un 54 % de aciertos pero fallaba en sad/angry.

★ FASE 2 – Un modelo más profundo y equilibrado

Añadimos más capas, técnicas para que no se “atasque” y equilibramos las clases para que todas las emociones tuvieran el mismo peso. Resultado: un 63 % de aciertos y empezó a distinguir mejor entre emociones similares.

★ FASE 3 – Combinamos lo mejor de dos mundos

Usamos la red como base y añadimos un clasificador especial para emociones difíciles. Resultado: mejor precisión en angry/fear/sad.

★ FASE 4 – El modelo final: combinación inteligente

Creamos un ensamble, es decir, una fusión suave entre el modelo general y el especializado. Así aprovechamos la fuerza de ambos: uno más equilibrado, el otro más preciso. Resultado final: 65 % de aciertos y 62,4 % de equilibrio entre clases — el mejor rendimiento global del proyecto.

Fase	Modelo	Accuracy	Macro-F1
0	Base	0,48	0,48
1	Simple	0,54	0,50
2	Profunda	0,63	0,61
3	Combinado	0,64	0,61
4	Ensamble	0,65	0,62

Conclusión y modelo final

FUNCIONAMIENTO DEL MODELO ELEGIDO

- 1 Recibe la imagen en tamaño pequeño y en blanco y negro (para quedarse con las formas y gestos importantes).
- 2 Una red neuronal (CNN) actúa como extractor de rasgos: cejas, ojos, comisuras, arrugas. De ahí saca un resumen numérico que luego usamos en nuestro modelo
- 3 Se combinan dos opiniones independientes: generalista (probabilidad por emoción) y especialista (para casos confusos).
- 4 Combinamos dos modelos: uno más general y otro más especializado. El general tiene más peso (75 %) y el especializado aporta un 25 %, de modo que el resultado final aprovecha lo mejor de ambos.
- 5 Se elige la emoción con mayor probabilidad y se muestra también un nivel de confianza.

¿Es bueno el modelo?



¿Es bueno el modelo?



MIEDO



ENFADO



ENFADO

Nuestro modelo

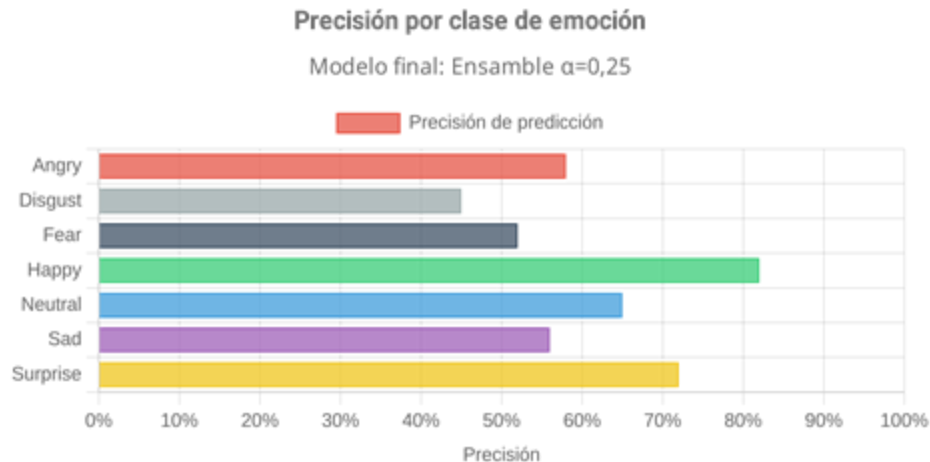
ANGRY (ENFADO)

Aproximadamente 65 % de aciertos.

FEAR (MIEDO)

Aproximadamente 55 % de aciertos.

👉 Es decir, nuestro modelo reconoce bien el enfado en unos 2 de cada 3 casos, y el miedo en algo más de 1 de cada 2



Viabilidad de las métricas

★ SUPERACIÓN DEL NIVEL DE AZAR

El modelo del Proyecto Júpiter (64.54% Accuracy) supera significativamente el nivel de azar para 7 emociones (14.29% Accuracy), siendo aproximadamente 4.5 veces más preciso que una predicción aleatoria.

★ RENDIMIENTO COMPARATIVO CON HUMANOS (FER2013)

En el dataset FER2013, el rendimiento humano estimado es de ~65.5% Accuracy. El modelo del Proyecto Júpiter (64.54% Accuracy) se encuentra muy cerca de este demostrando una capacidad competitiva.

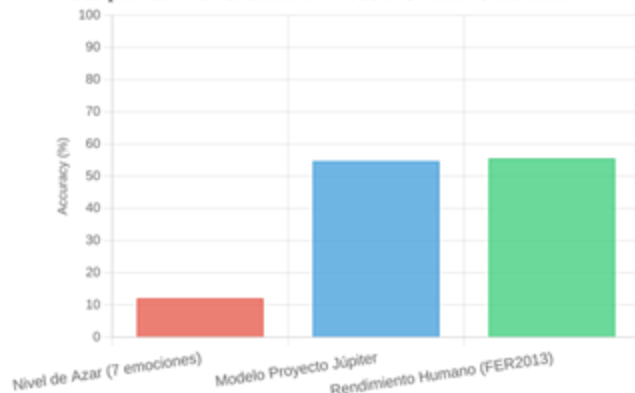
★ CONTEXTO DE COMPLEJIDAD

Estudios muestran que el rendimiento humano en tareas de reconocimiento de emociones puede variar (41%-72% Accuracy) y ser inferior en contextos más complejos o con mayor número de emociones.

★ VIABILIDAD Y POTENCIAL

Las métricas obtenidas demuestran que el modelo es viable y robusto para la clasificación de emociones, ofreciendo una base sólida para la toma de decisiones estratégicas en PontIA World.

Comparación de Rendimiento: Modelo vs. Azar vs. Humano





IA generativa

RESUMEN

★ PROBLEMA

Valoraciones bajas en atracciones (media de 5 sobre 10).

★ OBJETIVO

Mejorar la experiencia y valoraciones de los visitantes.

★ SOLUCIÓN

IA generativa para personalización de la experiencia en tiempo real.

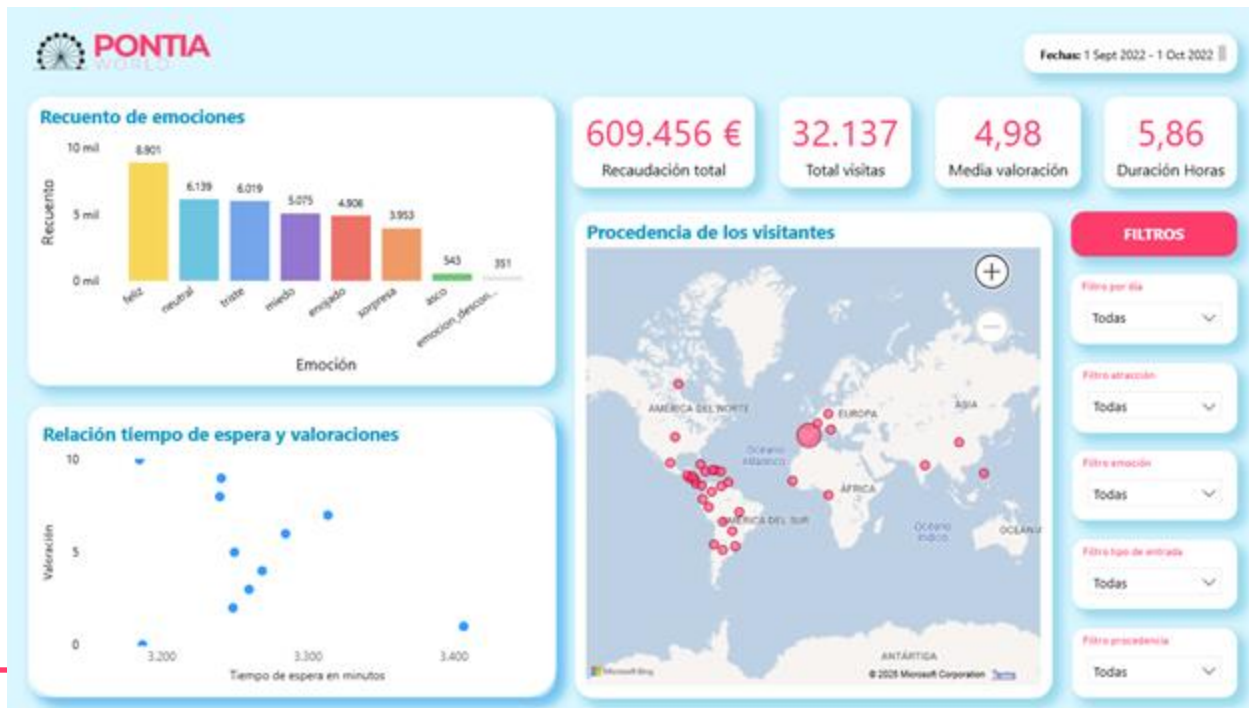
CASOS DE USO

- ✓ Itinerario y recomendaciones personalizadas e info tiempos de espera aproximado para las atracciones en tiempo real.
- ✓ Asistente o chatbot para incidencias y dudas del usuario.
- ✓ Diseño adaptativo y escalable de experiencias basado en emociones y datos recopilados.

BENEFICIOS ESPERADOS

- ✓ Fiabilidad y trazabilidad de datos automatizada.
- ✓ Mejora de valoraciones (objetivo: subir de 5-7 sobre 10 en aprox. 6 meses tras despliegue).
- ✓ Propuesta de valor diferencial frente a competencia.

Dashboard interactivo





Conclusiones

- ✓ Los visitantes no se van del satisfechos con la experiencia: sus valoraciones son bajas y muchos solo se suben a una atracción antes de marcharse.
- ✓ No hay una estrategia clara de precios, lo que hace que algunos pases no resulten atractivos y la recaudación sea baja.
- ✓ El parque recibe pocos visitantes y no logra generar un impacto positivo que los incentive a repetir la experiencia.

Trabajo a futuro

- ★ Reforzar la experiencia del visitante, aprovechando los resultados emocionales para detectar qué zonas o atracciones generan más disfrute o frustración.
- ★ Ajustar la estrategia de precios según el tipo de pase y antelación de compra.
- ★ Diseñar nuevas experiencias basadas en emociones positivas —más momentos “wow” que fomenten la fidelidad.
- ★ Integrar el modelo de emociones en tiempo real.
- ★ Predecir satisfacción y anticipar caídas de afluencia.
- ★ Implementar tablas estáticas en SQL para almacenar métricas clave reduciendo memoria de trabajo.



¡GRACIAS!

