

Proyecto Capstone

Sistema de Análisis de Reseñas y Sentimientos de Usuarios

Desarrollo de una solución integral para el análisis de opiniones y sentimientos de usuarios mediante técnicas de Procesamiento de Lenguaje Natural

Grupo #3

- Julio César Hernández Santiago
- David Eliud Hernández Vidal
- Keneth Uriel Chinchilla
- Fernando Mejía Pinto
- José Inestroza

Bases de Datos Organizacionales

Docente: Servio Palacios, Ph.D.

Fecha de entrega: 15 de noviembre de 2025



Planteamiento del Problema

Las empresas enfrentan la necesidad crítica de:



Monitorear y comprender la percepción

de clientes sobre productos, servicios y marcas



Procesar datos no estructurados

comentarios y reseñas que contienen información valiosa



Transformar en conocimiento accionable

para tomar decisiones estratégicas de manera rápida

Ejemplo: The Walt Disney Company requiere herramientas para:

- ✓ Identificar tendencias de satisfacción
- ✓ Detectar áreas de mejora
- ✓ Optimizar la experiencia del cliente
- ✓ Desarrollar estrategias de marketing efectivas



Solución Propuesta

Sistema diseñado para abordar la problemática mediante la recopilación, procesamiento y visualización de opiniones de usuarios de diversas fuentes digitales.

Flujo de Trabajo



Ingesta de Datos

Recopilación desde APIs y archivos



Procesamiento

Análisis con algoritmos NLP



Almacenamiento

Base de datos NoSQL



Visualización

Dashboards interactivos

Características



Clasificación de Polaridad

Positiva, neutra o negativa



Procesamiento en Tiempo Real

Análisis inmediato de datos



Visualización Interactiva

Dashboards intuitivos

Objetivos del Proyecto



Objetivo General

Desarrollar un sistema de base de datos organizacional para capturar, almacenar y analizar datos no estructurados en formato JSON, utilizando técnicas de NLP para realizar análisis de sentimientos en reseñas de usuarios.

Objetivos Específicos



Diseñar un modelo de base de datos NoSQL

Almacenar comentarios de usuarios y resultados de análisis de sentimientos.



Implementar un sistema de alto rendimiento

Garantizar eficiencia en procesamiento y inserción de resultados analíticos.



Integrar herramientas analíticas

Conectar con Python y Power BI para visualización e interpretación de información.



Establecer un flujo de trabajo ETL

Facilitar extracción, transformación y carga de datos desde fuentes externas.



Evaluar eficiencia y escalabilidad

Procesar datos en tiempo real y adaptarse al crecimiento del volumen de información.

Justificación Tecnológica



MongoDB

Base de datos NoSQL documental

- ✓ Flexibilidad de esquema
- ✓ Escalabilidad horizontal
- ✓ Alto rendimiento
- ✓ Integración sencilla

NoSQL



Python y NLP

Procesamiento analítico

- ✓ Ecosistema de análisis
- ✓ Librerías NLP (spaCy, NLTK)
- ✓ Pipelines ETL
- ✓ Análisis de sentimiento

NLP



Power BI

Visualización de datos

- ✓ Visualización interactiva
- ✓ Métricas clave
- ✓ Gráficos interactivos
- ✓ Consultas directas

Visualización



Arquitectura del Sistema

Sistema estructurado en cuatro capas con flujo de trabajo secuencial que transforma reseñas de usuarios en información analítica valiosa.



Ingesta de Datos

- ✓ APIs
- ✓ Archivos (CSV/JSON)
- ✓ Formularios web



Procesamiento NLP

- ✓ Limpieza y normalización
- ✓ Análisis de sentimiento
- ✓ Extracción de keywords



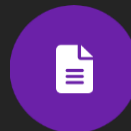
Almacenamiento

- ✓ Base de datos NoSQL
- ✓ Documentos JSON
- ✓ Metadatos enriquecidos



Visualización

- ✓ Power BI
- ✓ Gráficos interactivos
- ✓ Dashboards intuitivos



1. Recopilación



2. Procesamiento



3. Almacenamiento



4. Visualización

Modelo de Datos

MongoDB ofrece una estructura flexible y escalable que relaciona:



Usuarios



Reseñas



Parques

NoSQL

Documental

Flexible

Estructura de Documento JSON

● Usuarios ● Reseñas

```
{
  "users" : {
    "id" : "0504200198071" ,
    "name" : "John Doe" ,
    "email" : "johndoe@gmail.com"
  },
  "reviews" : {
    "id" : "45928" ,
    "rate" : 3,
    "text" : "Fue un excelente paseo por los castillos." ,
    "date" : "2024-06" ,
    "userid" : "0504200198071"
  },
  "parks" : [
    {
      "id" : 6,
      "name" : "Hong Kong Disneyland" ,
      "thematic" : "Magic Kingdom"
    }
  ]
}
```

Implementación y Resultados

MongoDB Atlas

Configuración de cluster para optimizar el rendimiento de las consultas en la colección de reseñas.

 Creación de Cluster



 Configuración de Índices



 Seguridad y Acceso


Integración con Power BI


Pasos para integrar MongoDB Atlas con Power BI mediante ODBC:


- 1 Crear conexión SQL en MongoDB Atlas
- 2 Elegir método de conexión adecuado
- 3 Instalar controlador ODBC de Atlas SQL
- 4 Configurar origen de datos en Windows
- 5 Crear instancia de base de datos federada

Dashboard Interactivo

Características del dashboard final para visualización de resultados:


 **Distribución de Sentimientos**
Gráficos de porcentaje de reseñas positivas, negativas y neutras

 **Nubes de Palabras Clave**
Visualizaciones de las palabras más frecuentes por sentimiento

 **Tendencias Temporales**
Gráficos que ilustran la evolución de los sentimientos a lo largo del tiempo


Dashboard y Visualización

Tipos de Visualizaciones




Distribución de Sentimientos

Gráficos mostrando % de reseñas positivas, negativas y neutras



Nubes de Palabras Clave

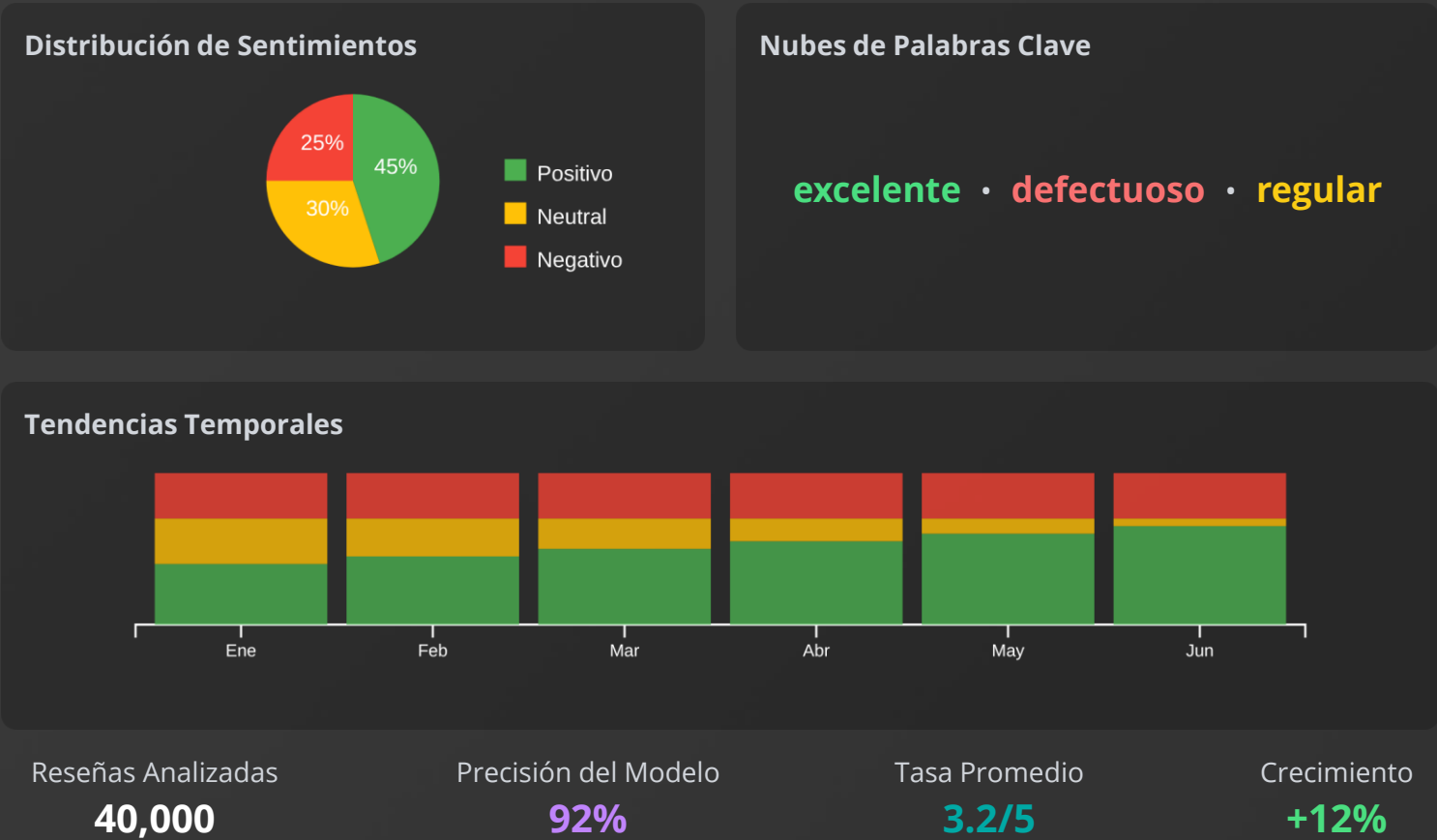
Visualizaciones de palabras frecuentes por tipo de sentimiento



Tendencias Temporales





Gráficos que ilustran evolución de sentimientos con el tiempo

Dashboard Power BI



Conclusiones y Futuro

✓ Conclusiones

-  Aplicación exitosa de los conocimientos adquiridos en Bases de Datos Organizacionales
-  Diseño e implementación de un sistema robusto para el análisis de reseñas y sentimientos de usuarios
-  Elección estratégica de tecnologías: MongoDB, Python y NLP para una solución escalable y flexible
-  Transformación eficaz de opiniones de usuarios en información valiosa para toma de decisiones

🚀 Futuro



Modelos Predictivos

Integración de modelos predictivos avanzados para anticipar tendencias de satisfacción



Análisis Multilingüe

Desarrollo de capacidades para analizar opiniones en múltiples idiomas



Integración en Tiempo Real

Implementación de capacidades para procesar y analizar reseñas de manera inmediata



Expansión de Casos de Uso

Aplicación del sistema a nuevas industrias y escenarios de análisis