

Proyecto Yelp & Google Maps



SMART RESOLVE ANALYTICS

Nuestro Cliente

'Active Life Strategies' es una firma de inversiones en busca de oportunidades estratégicas en el sector de servicios orientados a una vida sana, en el estado de Florida, USA.

Desean abrir una cadena de locales en los que se realizarán actividades de movimiento como danza, yoga y pilates, acompañados de ofertas de alimentos y productos saludables .

Cuentan con nosotros para obtener una perspectiva informada y datos sólidos sobre el mercado actual, que respalden las decisiones de inversión.





Equipo de trabajo



Felipe Rodríguez

Data Engineer



Oscar Francovich

Data Engineer



Alan García

Data Scientist



Gala Taddeo

Data Analyst



Ana Paulina Maya

Data Analyst



Objetivos

Analizar el mercado para determinar la presencia o ausencia de establecimientos que ofrezcan servicios dirigidos a promover la vida saludable, en el estado de Florida, USA.

Evaluar las reseñas de los clientes, registradas en Google Maps y Yelp, sobre este tipo de negocios para entender la percepción de estos sobre los servicios recibidos.

Ofrecer a nuestro cliente recomendaciones sobre ubicaciones estratégicas, en las cuales posicionar sus locales y ofrecer sus servicios, buscando así asegurar el éxito en la inversión realizada.

Comprender la interacción de factores sociales, las tendencias emergentes y las preferencias del consumidor en esta región.



Alcance del proyecto

Se encuentra dentro del alcance de este proyecto:

Negocios objetivo: aquellos enfocados en el bienestar y mejora de la vida saludable de los clientes.

Datos: La fuente que utilizamos como información son las reviews en Google Maps y Yelp, así como la data externa a través de las API's.

Ubicación: El proyecto se centrará en el estado de Florida, USA. Esta elección se fundamenta en la demanda de este tipo de ofertas, gracias a que sus habitantes tienen una creciente interés en la salud y el bienestar. También por la infraestructura ya existente al igual que su potencial de crecimiento



KPI's

- Tasa de retención de clientes (aumento del 5% en 6 meses)
- Promedio de calificación de negocios (aumento del 10% en 6 meses)
- Porcentaje de reseñas positivas (aumento del 10% en 6 meses)
- Tasa de conversión de recomendaciones en visitas (aumento del 10% en 1 año)
- Tasa de crecimiento de la industria (aumento del 10% en 1 año)

Pipeline y Stack Tecnológico



Calidad de los datos

- 2 Grupos de datos provenientes de Google Maps
- Reviews 2'850.000 datos con un 62.1% tienen una review
- Text y Rating datos más importantes
- Escala de rating de [1-5]
- Unión mediante "gmap_id"
- Category como dato para clasificar los negocios
- MISC como los servicios adicionales ofrecidos por los negocios
- Palabras Claves Health, Organic, Gym, Spa, Vegan, etc
- 9.971 nulos en category
- Longitud y Latitud para localizarlos geográficamente
- 91.530 datos finales

Metadata	Reviews
<ul style="list-style-type: none">- name- address- gmap_id- description- latitude- longitude- category- avg_rating, num_of_reviews- price- hours- MISC- state- relative_results- url	<ul style="list-style-type: none">- user_id- name- time- rating- text- pics- resp- gmap_id



Recomendaciones

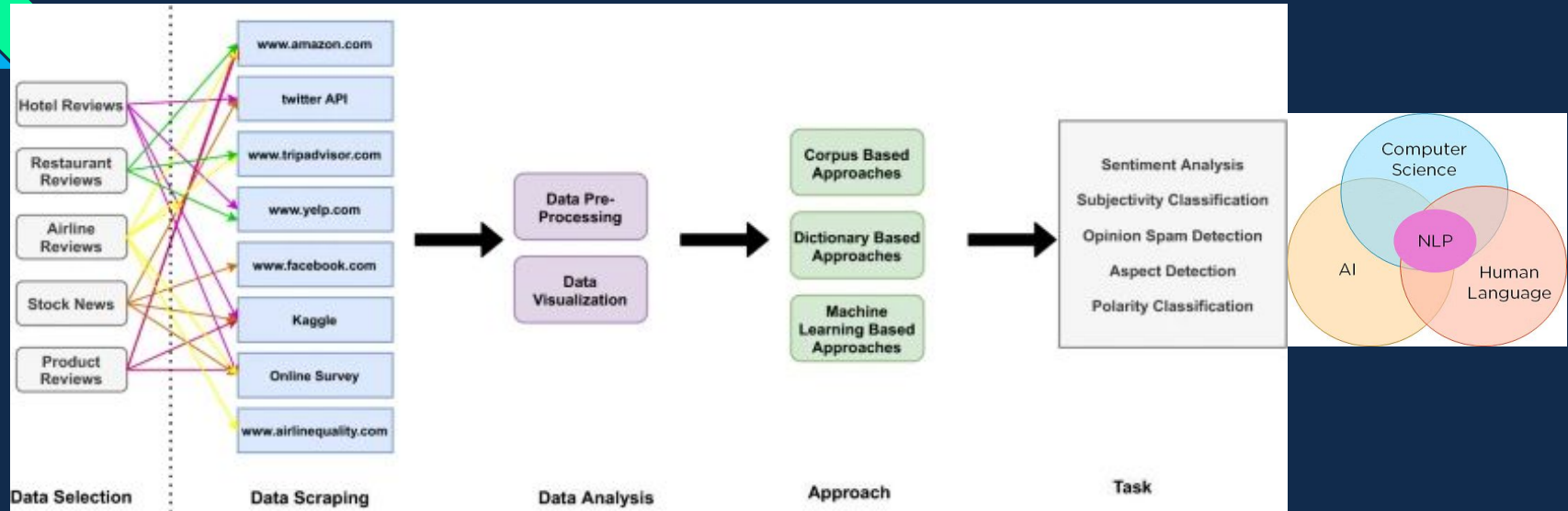
Tratamiento Avanzado de Datos: Realizar un tratamiento exhaustivo a los datos, enfocado en la estandarización del formato de los datos, facilitando así su manipulación y análisis posterior. (pandas, numpy, SQL)

Análisis Estadístico Detallado: Realizar un análisis estadístico detallado con el fin de observar el comportamiento de los dato y su valor. Esto incluye la identificación y tratamiento de valores nulos, duplicados, y outliers.

Enriquecimiento del Conjunto de Datos: Se sugiere integrar información adicional proveniente del dataset de Yelp, así como realizar tareas de web scraping en la plataforma de Google Maps. Mejorando así la robustez y profundidad de nuestro análisis.

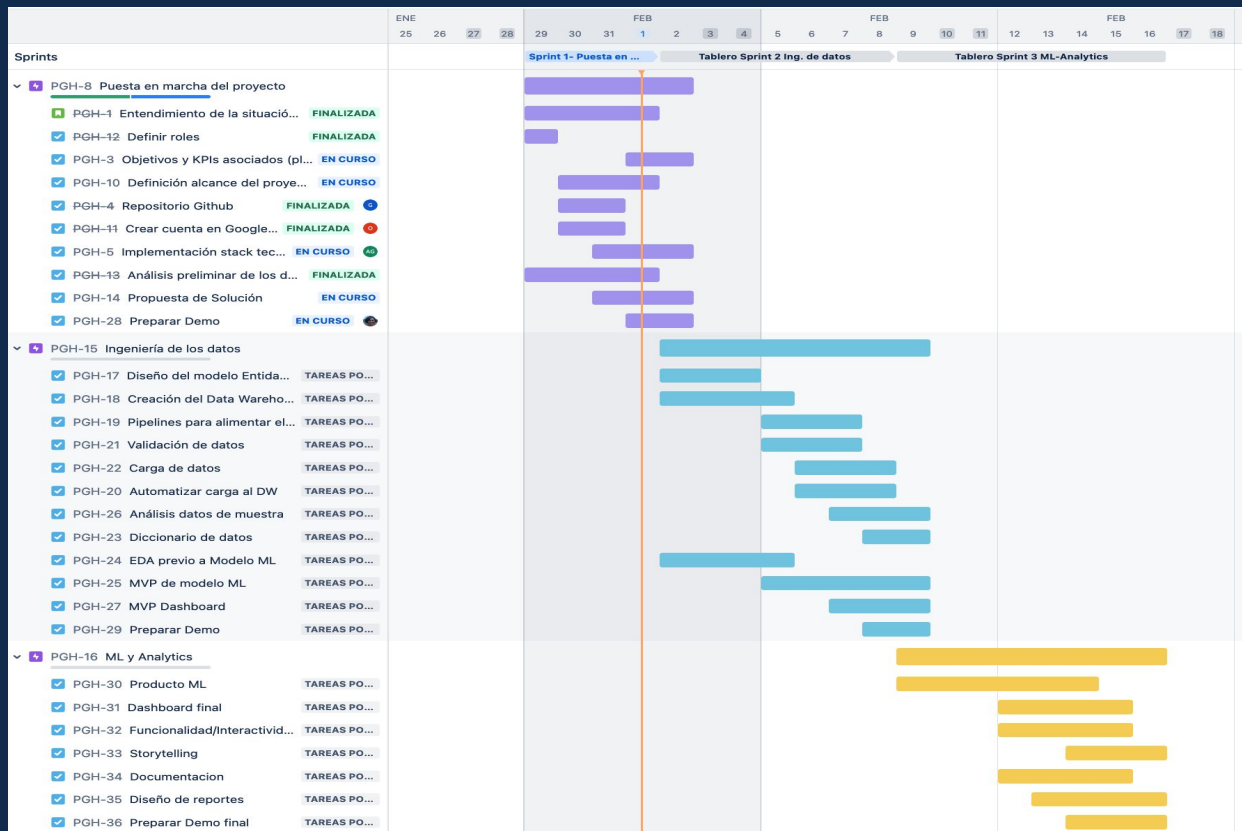
Verificación de Información Geográfica: Se recomienda el uso de librerías de Python especializadas como Folium o GeoPandas. Estas herramientas no solo facilitan la verificación de la veracidad de los datos geográficos, sino que también permiten la visualización avanzada de los mismos, ofreciendo así una perspectiva geoespacial enriquecedora para nuestro análisis.


Modelo ML



El Procesamiento Natural del Lenguaje (NLP) es una rama de la inteligencia artificial que permite a las computadoras comprender, interpretar y generar lenguaje humano de manera natural. En el ámbito de la inteligencia de negocios, el NLP se utiliza para realizar análisis de sentimiento en opiniones de clientes, extraer información relevante de grandes cantidades de texto, automatizar respuestas en servicios de atención al cliente y realizar una segmentación de mercado más precisa. Esta tecnología ha transformado la forma en que las empresas utilizan el lenguaje humano para tomar decisiones basadas en datos, mejorar la satisfacción del cliente y optimizar sus operaciones.

Planificación del proyecto





Gracias por elegirnos para explorar, junto a ustedes, las posibilidades ilimitadas que ofrece el mercado de alimentos saludables y vida sana en Florida. Con nuestro enfoque basado en datos y nuestra pasión por la excelencia, estamos listos para ayudarles a alcanzar sus metas financieras y empresariales.



SMART RESOLVE ANALYTICS

