

Manuel → Github Manager

Rafael → Desarrollador DevOps, arquitecto de software

Andrés → Apoyo y soporte

1. **Entendimiento del Negocio. “Revisad el texto” Rafa**

El proyecto para Distromerca S.A. tiene como objetivo desarrollar y optimizar las rutas de entrega de productos perecederos dentro de la Comunidad de Madrid. Esta iniciativa busca no solo reducir costos y mejorar los tiempos de entrega, sino también adaptarse a los desafíos logísticos específicos asociados con la distribución en áreas urbanas y suburbanas densas. Enfocándonos en las necesidades específicas de Distromerca S.A., trabajamos para implementar una herramienta de optimización de rutas en esta región, que es el corazón económico y administrativo de España y presenta desafíos únicos debido a su densidad poblacional y compleja distribución geográfica.

**1. Determinar los Objetivos Empresariales**

* **Antecedentes**: Confrontados con la variabilidad en la demanda y las estrictas expectativas de entrega, es crucial para Distromerca S.A. optimizar sus operaciones logísticas.
* **Objetivos Empresariales**: Minimizar los costes operativos y maximizar la eficiencia y puntualidad de las entregas.
* **Criterios de Éxito Empresarial**: El proyecto será considerado exitoso si se logra una mejora notable en la eficiencia de las rutas y una reducción en los costos operativos.

### 

| **Objetivo** | **Descripción** | **KPI Relacionado** |
| --- | --- | --- |
| Reducción de Costos | Minimizar los gastos operativos asociados con las rutas de entrega de productos perecederos. | Reducción en los costos operativos totales. |
| Maximización de Entregas | Aumentar el número de entregas realizadas dentro de los plazos estipulados. | Aumento en el porcentaje de entregas a tiempo. |
| Optimización de Rutas | Desarrollar rutas más eficientes que disminuyan el tiempo de tránsito y los kilómetros recorridos. | Disminución en el promedio de kilómetros por entrega. |

**2. Evaluar la Situación**

* **Inventario de Recursos**: Datos de vehículos (df\_vehicle.xlsx), ubicaciones de clientes (df\_location.xlsx), pedidos de los clientes(df\_order.xlsx), matrices de distancias y tiempos (df\_distance\_km.xlsx y df\_distance\_min.xlsx), y el histórico de pedidos (df\_historic\_order\_demand.xlsx).
* **Requisitos, Supuestos y Restricciones**: Los vehículos deben operar dentro de las capacidades de carga y autonomía especificadas. Todas las rutas deben comenzar y terminar en el almacén, siguiendo la normativa local de tráfico.
* **Riesgos y Contingencias**: Riesgos tecnológicos, posibles retrasos en la implementación, con planes de mitigación establecidos.
* **Terminología y Costos y Beneficios**: Definiciones claras y análisis de retorno de inversión detallado.

**3. Determinar Objetivos de Minería de Datos**

* **Objetivos de Minería de Datos**: Optimizar rutas basándose en análisis predictivos y en tiempo real de las condiciones de tráfico y patrones de demanda.
* **Criterios de Éxito de la Minería de Datos**: Precisión del modelo de al menos 85% en la simulación y optimización de rutas.

**4. Producir Plan del Proyecto**

* **Plan del Proyecto**: El proyecto seguirá un cronograma detallado que comienza con la recopilación de datos y culmina con la implementación de una solución optimizada para la gestión de rutas. Este plan incluirá fases de evaluación de tecnologías, desarrollo de modelos, pruebas piloto y despliegue final.
* **Supuestos del Proyecto**:
  + Los datos históricos son indicativos de tendencias futuras.
  + La disponibilidad de recursos tecnológicos y humanos conforme a lo planificado sin retrasos significativos.
  + Las condiciones de tráfico y patrones de demanda permanecen consistentes con variaciones predecibles.
  + Las herramientas y técnicas seleccionadas serán eficaces para abordar las necesidades de optimización identificadas.
* **Evaluación Inicial de Herramientas y Técnicas**: Exploración y selección de algoritmos como Algoritmos Genéticos, Tabu Search, Hill Climbing y PuLP para determinar la mejor combinación para optimizar las rutas en función de eficiencia y efectividad.

### **II. Comprensión de Datos**

Esta fase es fundamental para asegurar que los datos que se utilizarán en el proyecto son comprendidos completamente, adecuados y preparados para las etapas subsiguientes de análisis y modelización.

**1. Recolección de Datos Iniciales**

* **Adquisición de Datos**: Asegurarse de que todos los datos necesarios han sido recogidos. Para este proyecto, hemos recopilado varios conjuntos de datos:
  + **Datos de Vehículos** (df\_vehicle.xlsx): Incluye detalles como la capacidad de carga, costo por kilómetro y autonomía de cada vehículo.
  + **Ubicaciones** (df\_location.xlsx): Coordenadas del almacén y los puntos de entrega.
  + **Matrices de Distancia** (df\_distance\_km.xlsx y df\_distance\_min.xlsx): Distancias y tiempos entre las ubicaciones.
  + **Pedidos** (df\_orders.xlsx): Información sobre los pedidos actuales.
  + **Histórico de Pedidos** (df\_historic\_order\_demand.xlsx): Datos para analizar tendencias y patrones pasados.

#### **2. Descripción de los Datos**

* Detalle de cada variable en los conjuntos de datos:
  + **Vehículos (df\_vehicle.xlsx)**
    - **vehicle\_id** (Entero): Identificador único para cada vehículo.
    - **capacity\_kg** (Entero): Capacidad de carga del vehículo en kilogramos.
    - **cost\_per\_km** (Decimal): Costo de operación del vehículo por kilómetro.
    - **range\_km** (Entero): Máxima distancia que el vehículo puede recorrer con una carga completa, en kilómetros.
  + **Ubicaciones (df\_location.xlsx)**
    - **location\_id** (Entero): Identificador único para cada ubicación.
    - **latitude** (Decimal): Coordenada de latitud.
    - **longitude** (Decimal): Coordenada de longitud.
  + **Distancias y Tiempos (df\_distance\_km.xlsx y df\_distance\_min.xlsx)**
    - **from\_location\_id** (Entero): Identificador de la ubicación de origen.
    - **to\_location\_id** (Entero): Identificador de la ubicación de destino.
    - **distance\_km** (Decimal): Distancia entre las ubicaciones en kilómetros.
    - **time\_min** (Decimal): Tiempo estimado de viaje entre las ubicaciones en minutos.
  + **Pedidos (df\_orders.xlsx)**
    - **order\_id** (Entero): Identificador único para cada pedido.
    - **customer\_id** (Entero): Identificador del cliente que realiza el pedido.
    - **product\_type** (Texto): Tipo de producto pedido.
    - **quantity** (Entero): Cantidad de producto pedido.
    - **order\_date** (Fecha): Fecha en que se realizó el pedido.
  + **Histórico de Pedidos (df\_historic\_order\_demand.xlsx)**
    - **order\_id** (Entero): Identificador único para cada pedido histórico.
    - **customer\_id** (Entero): Identificador del cliente que realizó el pedido.
    - **product\_type** (Texto): Tipo de producto pedido en el historial.
    - **quantity** (Entero): Cantidad de producto pedido en el historial.
    - **order\_date** (Fecha): Fecha en que se realizó el pedido histórico.

**3. Exploración de Datos**

* **Análisis Profundo**: Utilizar técnicas estadísticas y de visualización para profundizar en la comprensión de los datos. Esto puede incluir la creación de visualizaciones de distribución de la demanda, análisis de correlaciones entre distancia y tiempo de entrega, y evaluación de la capacidad de carga frente a la demanda de pedidos.

**4. Verificación de la Calidad de los Datos**

* **Limpieza y Preparación de Datos**: Identificar y corregir problemas de calidad en los datos, como valores faltantes, errores de entrada, y discrepancias. Asegurarse de que los datos son consistentes y fiables para su uso en la modelización y toma de decisiones.

i. Descripción detallada de los requisitos funcionales y no funcionales del prototipo.

**Requisitos funcionales**

* **Gestión de Flotas y Vehículos**
  + El sistema deberá permitir el registro y manejo de la **capacidad máxima de carga** de cada vehículo en kilogramos.
  + El sistema deberá gestionar la **autonomía máxima en kilómetros** de los vehículos y garantizar que las rutas generadas no excedan este límite.
  + Deberá ofrecer la posibilidad de simular escenarios con diferentes tamaños de flota, incluyendo la **reducción del 50% de los vehículos disponibles**.
* **Optimización de Rutas**
  + El prototipo deberá implementar algoritmos para generar las **rutas óptimas**, minimizando el coste total y maximizando el número de pedidos entregados.
  + Todas las rutas deberán iniciar y finalizar en el almacén principal.
  + Las rutas generadas deberán cumplir con las **restricciones de capacidad de carga, autonomía de los vehículos y coste por kilómetro**.
* **Manejo de Datos**
  + El sistema deberá permitir la **importación de datos** desde las tablas proporcionadas: vehículos, localizaciones, pedidos, histórico de pedidos, entre otros.
  + Se deberán generar **reportes detallados** que incluyan las rutas optimizadas, estimaciones de coste y número de pedidos entregados.
  + Para la predicción de demanda, el prototipo deberá utilizar el **histórico de pedidos** mediante técnicas de machine learning o modelos estadísticos.
* **Simulación de Escenarios**
  + El sistema deberá soportar la ejecución de los casos de uso predefinidos:
    - Caso 1: Optimización de rutas con datos actuales.
    - Caso 2: Optimización de rutas con una flota reducida al 50%.
    - Caso 3: Simulación de rutas basadas en la predicción de la demanda para el próximo mes.
  + Deberá incluir la funcionalidad para que los usuarios creen y simulen **escenarios personalizados**, ajustando configuraciones como tamaño de la flota o demanda esperada.
* **Interfaz de Usuario**
  + El sistema deberá permitir la **visualización de rutas** generadas en un mapa interactivo.
  + La interfaz deberá proporcionar herramientas para configurar restricciones y parámetros de optimización (por ejemplo, costes por kilómetro, capacidad de los vehículos, autonomía).

**Requisitos no funcionales**

* **Rendimiento**
  + El sistema deberá ser **escalable**, capaz de manejar grandes volúmenes de datos, como cientos de pedidos y decenas de vehículos.
  + El tiempo máximo de respuesta para generar rutas optimizadas no deberá exceder los **5 minutos** con datos a escala real.
* **Usabilidad**
  + La interfaz deberá ser **intuitiva**, adecuada para usuarios sin experiencia técnica.
  + El prototipo deberá incluir documentación completa, como manuales o tutoriales, para guiar a los usuarios en la carga de datos y ejecución de simulaciones.
* **Mantenibilidad**
  + El sistema deberá estar diseñado de manera modular, separando componentes como la optimización, predicción y visualización.
  + El código deberá estar **documentado y estructurado** para facilitar su mantenimiento y actualización.
* **Portabilidad**
  + El prototipo deberá ser compatible con **múltiples plataformas** operativas (Windows, macOS, Linux).
  + El sistema deberá minimizar el uso de dependencias externas para evitar problemas de compatibilidad.
* **Seguridad**
  + Deberán implementarse medidas para **proteger los datos cargados y procesados** por el sistema.
  + El prototipo deberá cumplir con las normativas de **privacidad y protección de datos**, garantizando la confidencialidad de la información de clientes y pedidos.
* **Escalabilidad**
  + El sistema deberá ser **extensible**, permitiendo la incorporación de nuevos casos de uso, restricciones o funcionalidades sin afectar su estructura base.
  + Deberá estar preparado para manejar **incrementos en el volumen de datos**, como un aumento en el número de pedidos, vehículos o ubicaciones.

**ii. Análisis de tecnologías IA implicadas en el proyecto**

Para el diseño de un sistema de optimización de rutas eficiente se van a probar diversas tecnologías IA como son:

Algoritmos de optimización:

-Optimización combinatoria.

-Algoritmo Dijkstra

-Algoritmos geneticos.

**iii. Esquema de los casos de uso y escenarios de usuario.**