# Gestión del proyecto

En este documento se describen las fases por las que ha pasado el desarrollo de este proyecto así como una planificación inicial.

## Descripción de las fases del proyecto

El proyecto se ha dividido en varias fases que han sido revisadas con posterioridad y completadas o modificadas hasta llegar a completar el proyecto.

* + - Análisis de la tecnología que se ha utilizado o se ha pensado utilizar en este proyecto. Tecnologías inicialmente planteadas para el desarrollo del proyecto:
      * Inicialmente se planteó usar un lenguaje de programación conocido como es JavaScript para la usabilidad del sistema a nivel de usuario, así como Python para el desarrollo del modelo KNN y un entorno de programación también conocido como es Visual Studio Code y Anaconda respectivamente para los lenguajes antes mencionados. Se realizó un análisis y una primera implementación del sistema pero se llegó a la conclusión de que no era la implementación más apropiada después de recabar más información sobre estos métodos.

Tecnologías utilizadas finalmente en el desarrollo del proyecto:

* + - * Finalmente se optó por usar solamente el lenguaje de programación Python, que es un lenguaje planteado para este tipo de análisis de datos y generación de modelos de clasificación. Se ha utilizado el entorno de programación Anaconda y Jupyter Notebooks para realizar la implementación del script desarrollado.
    - Toma de requisitos y datos necesarios. Se accede a la página Kaggle que contiene toda la información necesaria para realizar el proyecto, tanto los datos utilizados como un detalle del entorno del proyecto y una explicación de los conjuntos de datos.
    - Aprendizaje de la tecnología utilizada. Se hace necesario adquirir los conocimientos necesarios para utilizar el entorno de desarrollo Jupyter Notebooks y para desarrollar en el lenguaje de programación Phyton con los paquetes utilizados.
    - Diseño del análisis de los datos y de los modelos utilizados. Se realiza una estrategia de análisis de los datos del conjunto de datos para sacar conclusiones acerca de las variables más aptas para ser incluidas en el diseño y se evalúan los algoritmos que se pueden utilizar.
    - Implementación de los análisis realizados sobre los datos incluyendo las gráficas que muestran la distribución y el comportamiento de los datos. Se implementan a continuación las funciones que generarán los modelos de clasificación y la predicción sobre los conjuntos de test.
    - Experimentación de los modelos generados. Se realizan pruebas con varios modelos generados con diferentes configuraciones de las funciones de entrenamiento y con diferentes fracciones de los conjuntos de datos iniciales.
    - Documentación del proyecto. Se elabora la memoria que detalla el trabajo realizado y se prepara la presentación del proyecto.

## Planificación

A continuación se va a realizar una breve indicación sobre la duración de las fases y sus tareas:

* + - Análisis de la tecnología (30 días)
      * Análisis de JavaScript -> 6 días
      * Análisis de VSC -> 5 días
      * Análisis de Python -> 10 días
      * Análisis de Jupyter Notebooks -> 9 días
    - Toma de requisitos y datos necesarios (3 días)
      * Alta en cuenta de Kaggle y acceso al reto -> 1 día
      * Análisis del proyecto y de los datos propuestos -> 2 días
    - Aprendizaje de la tecnología utilizada (40 días)
      * Aprendizaje de JavaScript -> 5 días
      * Aprendizaje de VSC -> 6 días
      * Aprendizaje de Python -> 17 días
      * Aprendizaje de Jupyter Notebooks -> 12 días
    - Diseño del sistema (15 días)
      * Diseño del análisis de los datos -> 9 días
      * Diseño de los modelos utilizados -> 6 días
    - Implementación de análisis de datos y modelos (25 días)
      * Implementación de los análisis de datos y los gráficos -> 15 días
      * Implementación de las funciones de generación de los modelos -> 10 días
    - Experimentación (15 días)
      * Realización de experimentos con distintos algoritmos -> 10 días
      * Informe de experimentos -> 5 días
    - Documentación del proyecto (27 días)
      * Elaboración de la memoria -> 22 días
      * Elaboración y preparación de la presentación -> 5 días

## Medios empleados

Para la realización de este proyecto se ha contado con los siguientes medios:

* + - Un ordenador de escritorio, con el sistema operativo Windows 10.12.6, con un procesador AMD Ryzen 5 3600X y 16 GB de memoria RAM. Este equipo se ha utilizado para realizar las tareas de programación del código necesario para realizar el análisis de los datos, generar los gráficos y generar los modelos de clasificación con sus correspondientes predicciones. Es importante que el equipo tenga una capacidad de procesamiento elevada teniendo en cuenta la clase de algoritmos de clasificación que se van a utilizar.
    - Se ha usado como entorno de programación Jupyter Notebooks de Anaconda para facilitar la programación en Python y la descarga de los paquetes de funciones necesarias.
    - La versión 3 de Anaconda para Python que es el entorno de desarrollo utilizado en este tipo de desarrollos.