# Project Overview

This is an End-to-End pipeline for the **Telco Customers Churn** data set Covering both the **data analyst** and **machine learning**

## Project Motivation

The **Customers Churn** is one of the most relevant business matters currently with the popularity of **SasS** and **Subscription based platforms** , Which is the reason why **Telco Customers Churn** Data set fits perfectly the relevant modern business problems.

## Dataset

This project will use **cell2cell (duke university)** dataset Link , the motivation to go with this data set is the following :

- **Messy and Unprocessed** : This isn't the **pre-processed** version of the data set Link which can will highlight more the **data engineering** part
- **Perfect for Learning** : Since This is my Portfolio project to show case the full data pipeline in action
- **Challenging** : There is no prior notebooks and work done on this dataset which is a **Realistic** scenario in data jobs

## Overview

This project gonna cover the following :

- **ETL** pipeline
- Exploratory Data Analysis (**EDA**)
- Feature Encoding
- **Logistic Regression** implementation from scratch (Numpy + Pandas only)
- Evaluation and insights
- Flask API **Deployment**

## Objectives

- Perform detailed **EDA**
- Handle data cleaning and preprocessing
- **Feature** engineering
- Implement **Logistic Regression** from scratch
- Preform **Predictions** and Model **Evaluation**
- Construct a clear pipeline of work From **Raw Data** to meaningful **Business Insights**
- Present the whole workflow and the thinking process in a short **YouTube** video
- Deploy the model with a lightweight **Flask API** and construct a minimal UI/UX

## Stack & Tools

| Phase | Tools Used |
|---|---|
| Data Handling | Pandas, Numpy, SQL |
| Visualization | Seaborn, Matplotlib |
| Modeling | Numpy , Pandas only |
| Data Storage | SQLite |
| API Deployment | Flask |
| Documentation | Markdown, LaTex |

## Pipeline Diagram

## Project Structure

# Resources

### Books

- "An introduction to Statistical Learning with Python"
- "Elements of Statistical Learning"
  **Articles**

### Repositories

- Isl-python