

Colorectal Cancer Risk & Survival Prediction

Objective

The goal of this project is to predict the **survival status** of colorectal cancer patients based on their demographics, medical history, and lifestyle factors.

Dataset Overview

The dataset contains records of **colorectal cancer patients**, with features related to risk factors, treatment, and outcomes.

Instructions

1. **Data Exploration:** Load and examine the dataset.
2. **Data Preprocessing:** Handle missing values, encode categorical variables, and normalize numerical features.
3. **Feature Selection:** Identify important predictors for survival.
4. **Model Training:** Train a classification model to predict survival status.
5. **Evaluation:** Assess model performance using accuracy and classification reports.

Exploratory Data Analysis (EDA)¹

- Check for missing values
- Visualize key distributions
- Identify correlations between features

Data Preprocessing

- Handle missing values
- Encode categorical variables
- Normalize numerical features

Model Training

Train a **Logistic Regression** model to predict patient survival.

Conclusion