CS770: Assignment 2
BMI Classification Based on Gender, Height, and Weight Using Machine Learning.

Max points:100

Assignment Tasks

➢ **Data Exploration and Cleanup**

Conduct initial data exploration to understand dataset characteristics.
Clean the data by handling missing values or anomalies.

➢ **Exploratory Data Analysis (EDA)**

Analyze the distribution of BMI categories and explore the relationship between gender, height, weight, and BMI category.
Investigate gender-specific distributions for height, weight, and BMI categories.

➢ **Data Preprocessing**

Implement strategies to manage imbalanced data, such as oversampling, undersampling, or applying SMOTE.
Normalize or standardize height and weight features for better model performance.

➢ **Model Training and Prediction**

Divide the dataset into training and testing sets, ensuring gender representation is balanced in both.
Train classification models (e.g., Logistic Regression, SVM and KNN) and make predictions on BMI categories.
Gender-Specific Modeling: Train separate models for each gender to evaluate if model performance improves when predicting BMI categories within gender groups.

➢ **Model Evaluation and Comparison**

Utilize metrics such as Accuracy, Precision, Recall, and F1 Score for model evaluation.
Compare model performances across the general and gender-specific models to assess differences in predictive accuracies or other metrics
Discuss the impact of hyperparameter tuning on model performances.

➢ **Gender-Based Prediction Analysis**
Analyze and discuss the outcomes of gender-specific models in comparison to the general model.
Explore how gender differences might affect BMI predictions and the implications of these differences.