Leela Josna Kona

PROPOSING PREDICTION OF GENDER USING DENTAL MEASUREMENTS

Objective

The goal of this project is to use dental measurements to predict if a person is male or female. By applying **predictive modeling** techniques likes Logistic, Random Forest, we will try to find patterns in tooth data that can help in gender identification. These prediction methods can support in areas like forensic science and anthropology.

Datasets

We will integrate data from two main sources:

1. Forensic Anthropology Data Bank (FDB) / FORDISC
   Source: https://fac.utk.edu/background/
   This dataset contains detailed information such as cranial and postcranial measurements, dental traits, trauma indicators, and demographic data (e.g., height, weight, medical history). It provides rich dental observations useful for gender analysis.
2. Dentistry Dataset.csv
   A structured dataset with 1,100 records including:
   - Intercanine distance
   - Canine width
   - Canine index

The target variable is Gender (Male/Female), which will be encoded numerically for modeling.

Proposed Methods

1. Data Preparation – Clean the data, manage missing values, encode gender, and normalize measurements.
2. Exploratory Analysis – Use charts to understand the data and identify the right features.
3. Model Training – Try different models like:
   - Logistic Regression
   - Decision Tree Classifier
   - Random Forest
   - KNN
4. Model Testing – Check accuracy and performance using test data.
5. Final Selection – Pick the best model and identify key dental features used for prediction.

Through this project, we hope to show that using simple biological data like teeth measurements can help in identifying gender using modern technology.