**🦷 Forensic Dentistry: Gender Prediction Using Dental Metrics**

**📖 Project Documentation**

**📌 Project Title: Using Dental Metrics to Predict Gender**

**📌 Author: Latika Dattaram Bane**

**📌 Date: 31/03/2025**

**1️⃣ Introduction**

**🔹 Problem Statement**

Forensic dentistry is a branch of forensic medicine that plays a crucial role in identifying individuals, especially in cases where bodies are unrecognizable due to natural disasters or other incidents. Dental structures, particularly **teeth**, are highly durable and can serve as reliable indicators of identity. This project leverages **machine learning** to predict an individual’s **gender** based on dental measurements.

**🔹 Objective**

* To analyze dental metrics and their relationship with gender.
* To implement **machine learning models** for gender classification.
* To evaluate the performance of different classifiers and identify the best-performing model.

**2️⃣ Dataset Description**

**🔹 Dataset Name: Dentistry Dataset.csv**

This dataset contains various dental measurements, including inter-canine distance, canine width, and canine index, used to classify gender.

**🔹 Key Features:**

| **Feature Name** | **Description** |
| --- | --- |
| **Age** | The age of the individual |
| **Gender (Target Variable)** | Male (1) / Female (0) |
| **Sample ID & SL No.** | Unique identifier (not used for prediction) |
| **Inter-canine distance intraoral** | Measurement between upper canine teeth |
| **Right & Left Canine Width Casts** | Width of the right and left canines |
| **Canine Index** | Canine index measurement |

**🔹 Target Variable:**

* **Gender** (Male = 1, Female = 0)
* **Independent Variables:** Various dental measurements

**3️⃣ Methodology**

**🔹 Step 1: Data Preprocessing**

* Handle **missing values** (if any).
* Encode categorical variables (**Label Encoding** for Gender).
* Drop unnecessary columns (**Sample ID, SL No.**).
* Normalize numerical features (**Scaler or Normalizer**).

**🔹 Step 2: Exploratory Data Analysis (EDA)**

* **Heatmap to check feature correlations**
* **Histograms to visualize distributions of features**

**🔹 Step 3: Model Building**

* Split data into **training (80%) and testing (20%)**.
* Train the following classification models:
  1. **Logistic Regression**
  2. **Decision Tree**
  3. **Random Forest**
  4. **XGBoost**

**🔹 Step 4: Model Evaluation**

* **Accuracy Score**
* **Confusion Matrix**
* **ROC-AUC Curve**

**4️⃣ Results & Analysis**

**🔹 Model Performance Comparison**

| **Model** | **Accuracy (%)** |
| --- | --- |
| XGB | 0.8% |
| Decision Tree | 0.82% |
| Random Forest | 0.84% |

📌**Best Performing Model: Random Forest**  
📌**Key Findings:**

* The best model, **Random Forest**, achieved an accuracy of **0.84%**.
* The **heatmap analysis** showed that certain **dental features are strongly correlated with gender**.

**5️⃣ Conclusion & Future Work**

**✅ Conclusion**

* This project successfully built a **gender classification model** using dental metrics.
* The best model achieved an **accuracy of 0.84%**, demonstrating potential for forensic applications.

**🔮 Future Work**

* **Expand dataset**: More diverse samples for better accuracy.
* **Use Deep Learning**: Implement Neural Networks for improved classification.
* **Test on real-world forensic datasets**: Validate the model in forensic investigations.

**6️⃣ References**

* **Scikit-learn Documentation**: [https://scikit-learn.org](https://scikit-learn.org/)
* **Research Papers on Forensic Dentistry**

**7️⃣ Installation & Usage**

**🔹 Clone the Repository**

git clone <https://github.com/banelatika/MLPROJECTS_CAPSTONE.git>

**🔹 Install Dependencies**

pip install -r requirements.txt

**🔹 Run the Jupyter Notebook**

jupyter notebook

**📌 License**

This project is licensed under the MIT License.

**📢 Contribution**

Feel free to contribute to this project by submitting a pull request.

📌**Prepared by:***latika Dattaram Bane*   
📌**Date: 31/03/2025**

📌**GitHub Repository:** https://github.com/banelatika/MLPROJECTS\_CAPSTONE.git  
📢*For contributions, feel free to submit a pull request!*🚀