Osteoporosis Knee Classification

**Project Overview**

Osteoporosis Knee Classification is a machine learning project that automates the diagnosis of osteoporosis using two approaches: **logistic regression for tabular data classification** and **CNN for image-based classification**. The goal is to assist in early detection of osteoporosis using patient data and X-ray images.

**Dataset**

- Source: Kaggle (Preprocessed Osteoporosis Knee Dataset)

- Types:

- Tabular Data: Age, BMI, bone density levels

- Image Data: Knee X-ray images (Resized to 128x256)

- Preprocessing:

- Images: Resizing, Normalization, Data Augmentation

- Tabular Data: Scaling using `MinMaxScaler`

**Models Used**

1. Logistic Regression (Tabular Data Classification)

- Features extracted from patient data

- Training: Used `LogisticRegression` from `sklearn`

- Evaluation Metrics: Accuracy, Precision, Recall, F1-score

2. CNN (Image-Based Classification)

- Architecture:

- Conv2D layers for feature extraction

- MaxPooling for dimensionality reduction

- Dense Layers for classification

3. Training Details:

- Loss Function: Categorical Crossentropy

- Optimizer: Adam

- Epochs: 20-50

- Model Saved As: `cnn\_model.h5`

**Deployment (Streamlit Web App)**

- User chooses between:

- Image Classification: Upload an X-ray for prediction

- Patient Data Classification: Enter medical details

- Pre-trained models loaded using `joblib` and `tensorflow.keras.models.load\_model`

**Performance Metrics**

- Logistic Regression Accuracy: ~85-90%

- CNN Model Accuracy: ~85-95% (varies based on data augmentation)

- Evaluated using Confusion Matrix, Precision, Recall, and F1-score

**Challenges & Solutions**

- Imbalanced Dataset: Used data augmentation & class weighting

- Overfitting in CNN: Applied dropout layers & L2 regularization

- Deployment Issues: Ensured model compatibility with Streamlit

**S - Situation:**

Osteoporosis often goes undiagnosed due to limited access to radiologists and diagnostic delays. Early and accurate detection through AI can significantly aid treatment.

**T - Task:**

Build a machine learning system that can classify osteoporosis using two data types — tabular patient data and knee X-ray images — and deploy it as an easy-to-use web app for real-time prediction.

**A - Action:**

* Used **logistic regression** for classifying tabular features like age, BMI, and bone density.
* Built a **Convolutional Neural Network (CNN)** to classify preprocessed X-ray images resized to 128x256 pixels.
* Applied **data preprocessing**: MinMax scaling for tabular data; augmentation and normalization for image data.
* Tackled data imbalance using **augmentation techniques and class weighting**.
* Deployed the entire system using **Streamlit**, allowing users to upload images or input patient data for prediction.
* Integrated both models in the frontend using **TensorFlow (for CNN)** and **Joblib (for tabular model)**.

**R - Result:**

* Achieved **85–90% accuracy** with logistic regression and up to **95% accuracy** with CNN.
* Delivered an interactive and functional **Streamlit web application**.
* Handled overfitting with **Dropout and L2 regularization**, and improved usability with a clean interface.
* Opened scope for clinical adaptation and cloud-based deployment.