Sri Lanaka Institute of Information Technology

Knee Osteoarthritis Prediction and Progression using Multi-Modal Deep Learning.

Group : 25-26J -112

Data Analysis Report

BSc (Hons) in Information Technology Specializing in Data Science

Department Of Information Technology

Faculty Of Computing

Sri Lanaka Institute of Information Technology

Sri Lanka

January 2026

SLIIT
Discover Your Future

BSc (Hons) in Information Technology Specializing Data Science

Research Project - IT4010 **Data Analysis Report**

Contents

BSc (Hons) in Information Technology Specializing Data Science

Research Project - IT4010 Data Analysis Report

Project ID: 25-26J-112

Project Title: Knee Osteoarthritis Prediction and Progression using Multi-Modal Deep Learning.

1. Introduction

1.1. Background:

Knee Osteoarthritis (KOA) is a chronic joint disease caused by the gradual breakdown of cartilage in the knee. As the cartilage deteriorates, bones begin to rub against each other, leading to pain, stiffness, swelling, and reduced knee movement. This condition affects daily activities and significantly reduces the quality of life of affected individuals, especially older adults.

The prevalence of KOA is increasing due to aging populations, obesity, lack of physical activity, and knee injuries. KOA places a considerable burden on healthcare systems through long-term treatment, frequent hospital visits, and costly surgical procedures. Early detection and continuous monitoring are important to manage the disease effectively and prevent severe complications.

Currently, KOA diagnosis relies on clinical examinations, X-rays, MRI scans, and patient medical history. Although these methods are effective, they are expensive, time-consuming, and mainly available in hospital settings. In addition, they provide only single-time assessments and often rely on one type of data. Advances in data science and intelligent technologies enable the development of automated and integrated approaches for more accurate and accessible KOA prediction and monitoring.

1.2. Research Problem:

Knee Osteoarthritis (KOA) is a common chronic joint disease that causes pain, stiffness, and reduced mobility. Although widely prevalent, current diagnostic and monitoring methods rely on hospital-based clinical visits and expensive imaging techniques, making regular assessment difficult for elderly patients and those in rural or low-resource areas.

Most existing approaches use only a single type of data, such as medical images or clinical records. However, KOA is influenced by multiple factors including joint structure, symptoms, and patient characteristics. Using isolated data sources leads to incomplete assessment and lower prediction accuracy.

In addition, current methods provide only one-time evaluations and do not support continuous monitoring of disease progression. Since KOA progresses gradually, important changes may go unnoticed. Therefore, an integrated and cost-effective solution is needed to improve KOA prediction and monitoring.

Key Problems

- Dependence on expensive, hospital-based diagnostic methods
- Use of single data sources for KOA assessment
- Lack of continuous monitoring of disease progression
- Limited accessibility for elderly and rural patients

1.3. Objectives:

It Number	Objective	Objective number
IT22582942	To develop a machine learning–based model for predicting the presence of Knee Osteoarthritis using demographic, clinical, and biomarker data through data preprocessing, feature engineering, and supervised learning techniques.	1
IT22223708	To develop a deep learning-based model for detecting the presence of Knee Osteoarthritis using X-ray and MRI images by applying image preprocessing and convolutional neural networks.	2
IT22606792	To develop automatically grading the severity of knee osteoarthritis by combining X-ray images with clinical and biomarker data. Deep learning is used to analyze X-ray images, while machine learning models process clinical information to improve accuracy. The predicted severity level is then used to support appropriate treatment recommendations.	3
IT22188472	To develop an IoT-based wearable knee health monitoring component using Vibroarthrography (VAG) signals and biomechanical sensor data, combined with machine learning techniques, to support continuous monitoring and early detection of Knee Osteoarthritis outside clinical environments.	4

2. Data Exploration

2.1. Data Collection

Publicly Available Data : The dataset also includes publicly available data from online platforms such as Kaggle and NDA.

Hospital Collected Data :The dataset includes data physically collected from Knee Osteoarthritis patients at the National Hospital of Sri Lanka

2.2. Dataset Description:

Data source	Description	Resource	Size	Key attributes
Collected demographic, clinical & biomarker dataset	It consists of patient demographic information, clinical symptoms, and laboratory biomarker measurements related to knee osteoarthritis. This real-world dataset was manually collected following ethical guidelines and is used for machine-learning.	Colombo National Hospital, Sri Lanka (in-person data collection)	~400 data records were collected. ~600 synthetic records.	Data Type: Structured tabular data Patient Information: Age, gender, height, weight, BMI Clinical Symptoms: Knee pain, stiffness, swelling, difficulty in movement, history of knee injury Medical History: Obesity, diabetes, hypertension, family history of OA Biomarker Data: Blood test values such as FBS, WBC, platelets, cholesterol, CRP, ESR, RF, and FBC Application: Machine learning–based KOA severity prediction using demographic, clinical and biomarker features

Kaggle – Osteoarthritis Prediction	Knee X-ray image dataset for binary classification to detect whether osteoarthritis is present or not	Kaggle dataset page	3836 X-ray images	Image modality: knee X-ray images, Labels: Normal vs Osteoarthritis (presence/absence), Dataset structure: Train / Test / Valid folders
KOA Dataset with Severity Grading	The dataset consists of knee X-ray images labeled accordi-ng to the Kellgren–Lawrence grading system, ranging from Grade 0 to 4. It is sourced from the OA Initiativ-e & provided via Mendeley Data, making it suitable for automated KOA severity classification.	Kaggle Severity Level	9786 X-ray images	Severity Levels: KL Grade 0–4 Clinical Focus: Joint space narrowing, osteophytes, and sclerosis Application: Automated KOA severity grading using deep learning
Kaggle – Knee Health Dataset Using VAG Signals for AI and IoT	It consists of Vibroarthrography (VAG) signal–based features extracted from knee joint vibration data. The dataset is labeled with knee health conditions and severity levels, making it suitable for AI- and IoT-based knee health monitoring and early detection of Knee Osteoarthritis.	Kaggle dataset	2,500 records	Data Type: Structured tabular data Labels: knee condition, severity level, treatment advised VAG Signal Features: rms_amplitude, peak_frequency, spectral_entropy, zero_crossing_rate, mean_frequency temperature Application: Machine learning–based knee health assessment and continuous monitoring using VAG signals and IoT devices

2.3. Suitability Analysis

2.3.1. Relevance to Individual Research Objectives:

	1	2	3	4
Data s-ce 1	This research aims to predict Knee Osteoarthritis (KOA) using machine learning based on demographic, clinical, and biomarker data. The dataset collected from Colombo National Hospital is highly relevant, as it contains real patient information including age, gender, BMI, key symptoms, and medical history directly associated with KOA. In addition, biomarker data such as FBS, CRP, ESR, cholesterol, RF, and FBC support analysis of inflammatory and metabolic factors, enabling accurate and data-driven KOA risk prediction		The clinical and biomarker dataset strongly supports the objective of predicting KOA severity using non-imaging data, as it includes demographic information, clinical symptoms, & laboratory biomarkers relevant to osteoarthritis progression. This dataset enables machine learning–based analysis of risk factors & complements the X-ray dataset by providing additional clinical context for severity prediction.	

	1	2	3	4
Data source 2		The selected Kaggle Osteoarthritis Prediction dataset directly supports the individual research objective of detecting Knee Osteoarthritis using X-ray images. The labeled knee X-ray images enable the training and evaluation of convolutional neural network models to automatically learn features related to structural joint changes associated with KOA. This makes the dataset well suited for developing an accurate image-based KOA prediction component and for integration into the proposed multi-modal assessment system. Overall, the dataset aligns well with the research objective of building an automated, accurate, and scalable image-based KOA prediction system, making it suitable for integration into the proposed multi-modal Knee Osteoarthritis assessment framework.		
Data source 3			The X-ray dataset aligns well with the research objective of automated knee osteoarthritis severity prediction, as it contains radiographs labeled according to the Kellgren–Lawrence grading system. The visual features present in the images, such as joint space narrowing and osteophyte formation, directly support deep learning–based severity classification & evaluation of model performance.	

	1	2	3	4
Data Source 4				This dataset directly supports the objective of developing an IoT-based knee health monitoring system using Vibroarthrography (VAG) signals and machine learning. The labeled VAG features capture subtle knee joint vibrations associated with osteoarthritis, enabling non-invasive and continuous knee health assessment. Overall, the dataset is well aligned with real-time, wearable, and AI-driven KOA monitoring outside clinical environments.

3. Methodology

3.1. Data Preprocessing:

Objective 1 - IT22582942

	Data Cleaning	Data Normalization	Data Encoding	Feature Engineering	Data Mapping	Manual Correction
Data source 1	x	x	x	x	x	

Objective 2 - IT22223708

	Transformation technique						
	Data Cleaning	Data Normalization	Data Type Conversion	Data Scaling	Feature Engineering	Data Mapping	Data Type Conversion
Data source 2	x	x	x	x	x	x	x

Objective 3- IT22606792

	Transformation technique							
	Data Cleaning	Handling Missing Data	Data Encoding	Data Standardization	Feature Engineering	Manual Correction	Data Normalization	Data Mapping
Data source 1	x	x	x	x	x	x		
Data source 3	x		x		x	x	x	x

Objective 4 - IT22188472

	Transformation technique							
	Data Cleaning	Handling Missing Data	Data Encoding	Data Standardization	Feature Engineering	Manual Correction	Data Normalization	Data Mapping
Data source 4	x	x	x	x		x		x

3.2. Scalability

Objective 1 – IT22582942

The dataset used in this research component initially comprised a limited number of real patient records collected from Colombo National Hospital, containing demographic, clinical, and biomarker attributes relevant to Knee Osteoarthritis (KOA) prediction. To ensure sufficient data volume for reliable machine learning training and evaluation, the dataset was augmented using realistically generated synthetic records, increasing the total size to approximately 1,000 samples. Synthetic data generation was carried out by preserving observed clinical patterns, value ranges, and statistical relationships, ensuring medical plausibility. This scalable dataset structure supports robust model development and can be further expanded in the future through the integration of additional real patient data, enabling continuous improvement of model performance and generalization.

Objective 2 – IT22223708

The selected Kaggle Osteoarthritis Prediction dataset contains 3,836 knee X-ray images, which is sufficient for training, validation, and evaluation of deep learning–based image classification models for Knee Osteoarthritis detection. The dataset size supports effective feature learning using convolutional neural networks, especially when combined with data augmentation techniques to improve generalization.

The dataset structure allows easy expansion by incorporating additional knee X-ray images from other public repositories or hospital sources. New data can be added without modifying the existing preprocessing or model architecture, making the dataset scalable for future experiments and performance improvements

Objective 3 – IT22606792

Data Source 1:

The clinical and biomarker dataset initially consisted of around 400 real patient records collected with demographic, clinical, and laboratory attributes. To improve data sufficiency and model generalization, synthetic data was generated to extend the dataset to approximately 1000 records, while strictly maintaining medical validity by defining realistic value boundaries, constraints, and attribute distributions for each feature. This approach improves scalability for machine learning model training while preserving data consistency, and the dataset can be further expanded in the future by incorporating additional real patient records.

Data Source 3:

The Kaggle based knee X-ray dataset contains 9786 samples across multiple Kellgren Lawrence severity grades to support deep learning model training and evaluation. The dataset is inherently scalable, as additional X-ray images from public repositories or hospital sources can be incorporated using the same preprocessing pipeline and directory-based labeling structure. Furthermore, the use of transfer learning architectures such as EfficientNet enables efficient scaling, allowing new data to be added without retraining models from scratch.

Objective 4 – IT22188472

Data Source 4:

The VAG signal dataset used for this research component consists of 2,500 labeled samples containing vibration-based features extracted from knee joint signals. This dataset size is sufficient for training, validating, and evaluating supervised machine learning models for knee health classification and abnormality detection.

The dataset is inherently scalable, as VAG signals can be continuously collected using IoT-based wearable devices during daily activities. New sensor data can be added without modifying the existing preprocessing pipeline or model architecture. As more users and recording sessions are introduced, the dataset can grow incrementally, supporting long-term monitoring and improved model generalization.

Overall, the dataset structure and IoT-based data acquisition approach enable scalable and sustainable knee health analysis for future expansion and real-world deployment.

3.3. Feature extraction

Objective 1– IT22582942

For the demographic, clinical and biomarker dataset, feature evaluation involves identifying and preparing patient attributes that are clinically relevant to Knee Osteoarthritis (KOA) prediction. Key features were selected from demographic variables (such as age, gender, and BMI), clinical symptoms (including knee pain frequency, stiffness after resting, swelling, and functional difficulty), medical history indicators (such as obesity, diabetes, hypertension, and family history), and laboratory biomarker values (including FBS, CRP, ESR, cholesterol, RF, and FBC).

Derived features such as Body Mass Index (BMI) were calculated using height and weight measurements to better represent obesity-related risk. Data preprocessing steps such as encoding categorical variables, handling missing values, and standardizing numeric attributes were applied to ensure consistency across features.

Tree-based machine learning models, including Random Forest, Gradient Boosting, and XGBoost, were used to implicitly evaluate feature importance during training. These models identified the most influential clinical and biomarker attributes contributing to KOA prediction, enabling a data-driven assessment of feature relevance. This approach supported improved model interpretability and ensured that the most meaningful features were utilized.

Objective 2– IT22223708

This research component focuses on extracting meaningful and discriminative visual features from knee X-ray images to support automated Knee Osteoarthritis (KOA) detection. Instead of relying on handcrafted feature engineering techniques, feature extraction was performed implicitly using deep learning–based models, allowing the system to automatically learn relevant visual patterns directly from the image data.

Prior to feature extraction, image enhancement techniques were applied to improve feature quality. Contrast Limited Adaptive Histogram Equalization (CLAHE) was used to enhance local contrast within knee joint regions. This preprocessing step improved the visibility of important anatomical structures, such as joint margins and bone textures, thereby supporting more effective feature learning by the deep learning models.

Following convolutional feature extraction, Global Average Pooling layers were employed to aggregate spatial feature maps into compact global feature vectors. This approach reduced model complexity, minimized the risk of overfitting, and ensured that only the most salient and discriminative features were retained for classification.

Deep convolutional feature learning was carried out using several pre-trained Convolutional Neural Network (CNN) architectures, including EfficientNet, ResNet50, DenseNet, Inception (GoogLeNet), AlexNet, and YOLO-based classification models.

These CNNs are capable of learning hierarchical representations from raw pixel data, starting from low-level features such as edges and contours, and progressing to high-level semantic features such as bone texture variations, joint space narrowing, and structural irregularities that are indicative of knee osteoarthritis.

Transfer learning played a key role in the feature extraction process. All CNN models were initialized with ImageNet pre-trained weights, enabling the reuse of robust and general-purpose visual features learned from large-scale datasets. During the initial stage of training, the convolutional backbone layers were frozen, allowing the networks to function purely as feature extractors while training only the newly added classification layers. This approach helped reduce training time and mitigated overfitting due to limited medical imaging data.

To further enhance feature relevance, a two-stage training strategy was adopted. In the second stage, selective fine-tuning was applied by unfreezing approximately the last 60% of the convolutional layers in each CNN architecture. This allowed the models to adapt and refine the extracted features specifically for knee osteoarthritis characteristics, improving their sensitivity to subtle pathological patterns present in X-ray images.

The quality and effectiveness of the extracted features were evaluated indirectly through model performance metrics rather than explicit feature selection methods. Metrics such as classification accuracy, confusion matrices, and Receiver Operating Characteristic – Area Under the Curve (ROC-AUC) were used to assess how well the learned features supported KOA detection. Additionally, comparative evaluation across multiple CNN architectures enabled the identification of feature representations that contributed most effectively to accurate and reliable knee osteoarthritis prediction.

Objective 3– IT22606792

Data Source 1:

For the clinical and biomarker dataset, feature extraction involved selecting and transforming relevant patient attributes into a format suitable for machine learning models. Key features included demographic variables, clinical symptoms, medical history indicators, and laboratory biomarker values. Derived features such as Body Mass Index (BMI) were calculated to enhance predictive capability.

Tree-based models such as XGBoost inherently performed feature importance analysis during training, allowing the model to identify the most influential clinical and biomarker attributes. This data-driven feature extraction helped improve model interpretability and prediction accuracy.

Data Source 3:
Feature extraction for X-ray images was performed automatically using deep learning models, without manual hand-crafted feature design. Convolutional Neural Networks (CNNs), particularly EfficientNet and ResNet-based architectures, were used to learn hierarchical features such as edges, textures, joint space patterns, and osteophyte structures directly from knee radiographs. The convolutional layers extracted spatial and structural features that are critical for differentiating between Kellgren–Lawrence severity grades.

Pre-trained models were fine-tuned on the knee X-ray dataset to adapt generic image features to domain-specific osteoarthritis characteristics. The extracted deep features were then passed through fully connected layers for severity classification. This approach ensured robust and discriminative feature learning while reducing the need for manual feature engineering.

Objective 4– IT22188472
For this research component, feature extraction was performed using Vibroarthrography (VAG) signal–based attributes provided in the dataset. The dataset contains pre-extracted numerical features derived from knee joint vibration signals, which are clinically relevant for knee health assessment.

Key VAG features used include RMS amplitude, peak frequency, mean frequency, spectral entropy, and zero-crossing rate. These features capture vibration intensity, frequency distribution, signal irregularity, and joint roughness, which are indicative of abnormal knee joint behavior and osteoarthritis-related changes.

Prior to model training, the extracted features were evaluated for consistency and relevance through exploratory data analysis and scaling. Standardization was applied to numerical features to ensure equal contribution during model learning. The selected VAG features were then used as inputs to supervise machine learning models to classify knee conditions and support continuous knee health monitoring in an IoT-based environment.

4. Modelling and Results

4.1. Key Insights:

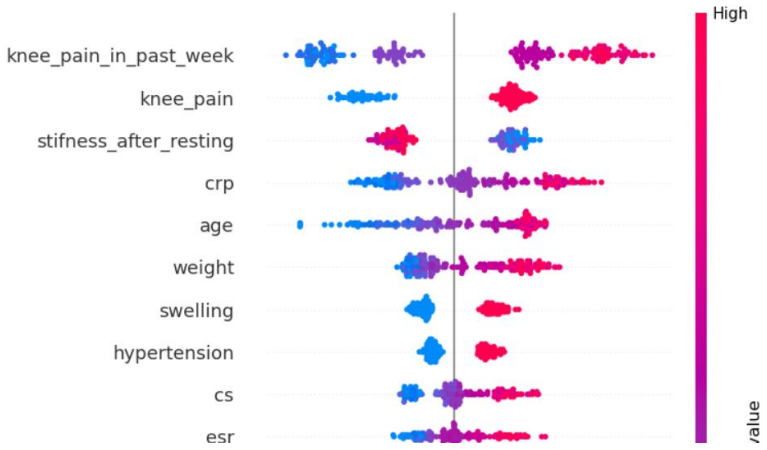
Objective 1 – IT22582942

Machine learning models including Random Forest, Gradient Boosting, and XGBoost were trained using demographic, clinical and biomarker data to predict Knee Osteoarthritis (KOA). Exploratory data analysis using visualizations revealed strong associations between KOA and factors such as knee pain frequency, age, crp, stiffness after resting, and weight. Elevated inflammatory biomarkers, including CRP and ESR, were more common among KOA-positive patients, indicating a link between inflammation and disease presence.

Tree-based models effectively captured non-linear relationships between demographic, clinical, and biomarker features. Feature importance analysis identified BMI, pain-related indicators, and inflammatory markers as key contributors to prediction. Consistent training and validation performance confirmed good model generalization and reliable KOA risk prediction

```
gb_importance = pd.DataFrame({
    "feature": X_train.columns,
    "importance": best_gb.feature_importances_
}).sort_values("importance", ascending=False)

print("Top 15 GB feature_importances_:")
print(gb_importance.head(15))
```

```
Top 15 GB feature_importances_:
                      feature  importance
7     knee_pain_in_past_week    0.285487
6                  knee_pain    0.277516
0                        age    0.112829
23                       crp    0.097165
8     stifness_after_resting    0.051956
3                     weight    0.036852
10                  swelling    0.035779
29                       BMI    0.017126
24                       esr    0.017073
21                        cs    0.012865
18                       fbs    0.012552
25                        rf    0.012445
22               cholesterol    0.006338
11  difficulty_in_performing    0.004684
15              hypertension    0.004203
```



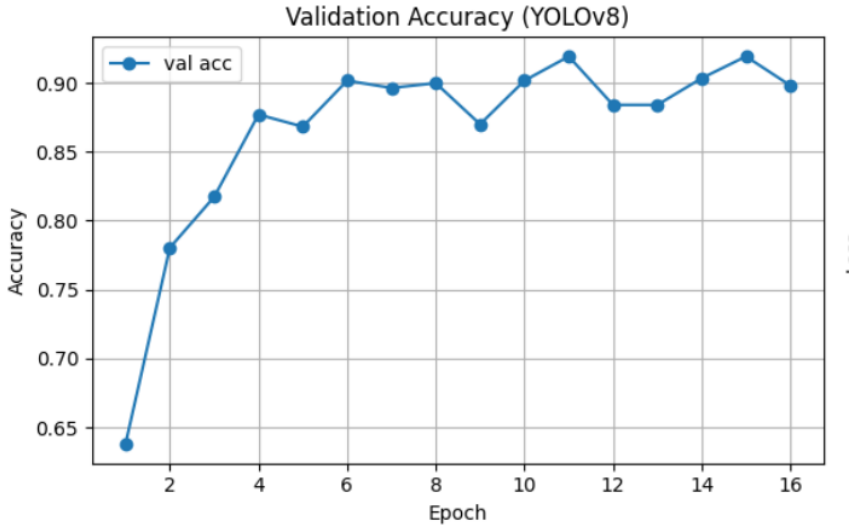
Objective 2 – IT22223708

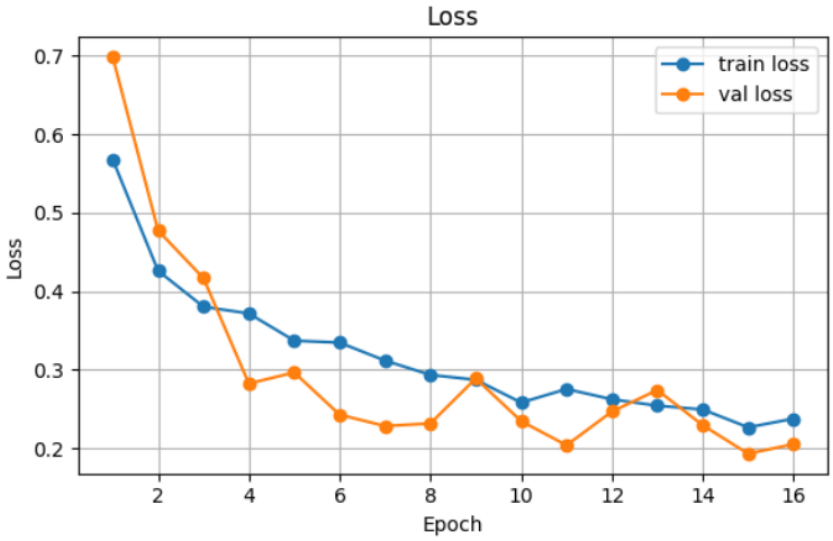
This section summarizes the key outcomes obtained from applying deep learning models for automated Knee Osteoarthritis (KOA) detection using X-ray images. The objective was to evaluate the effectiveness of AI-based image analysis in identifying osteoarthritic patterns and reducing reliance on manual diagnosis.

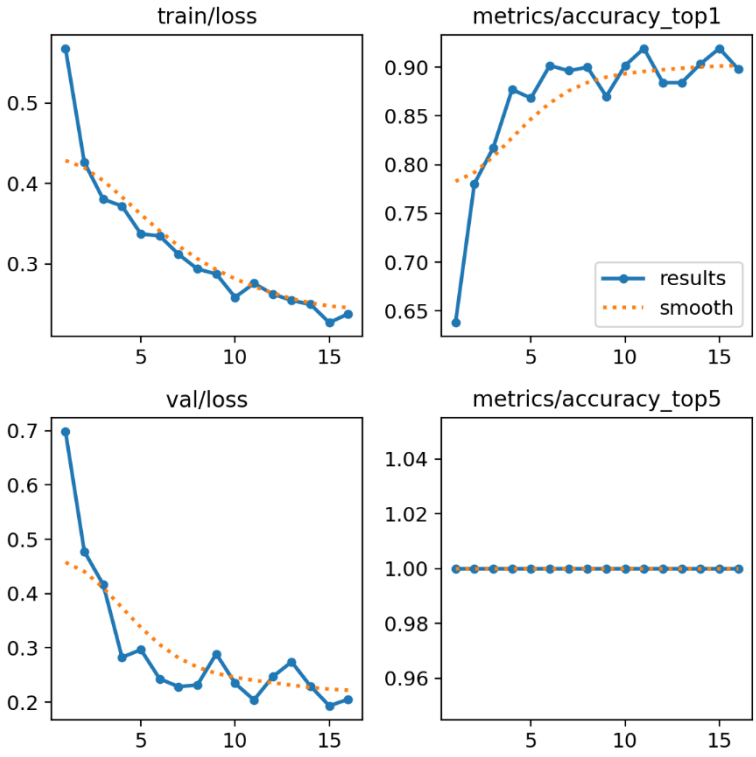
Multiple deep learning architectures were trained and evaluated using a consistent experimental pipeline, including stratified data splitting, image preprocessing, two-stage training, and performance evaluation using standard metrics.

A comparative analysis was conducted by training multiple CNN architectures, including ResNet, DenseNet, EfficientNet, AlexNet, Inception (GoogLeNet), and YOLO-based classifiers. This table presents the training, validation, and testing accuracies obtained for each model, enabling a systematic comparison of model performance and generalization capability.

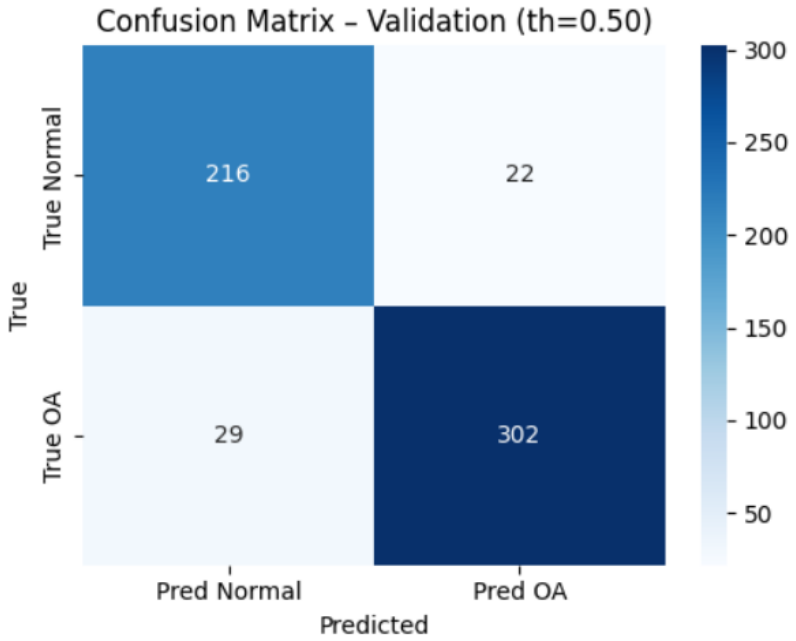
Model	Training Accuracy	Validation Accuracy	Testing Accuracy
DenseNet201	88.26%	90.86%	87.48%
EfficientNetB3	91.30%	99.06%	55.98%
Resnet50	88.22%	89.10%	86.77%
AlexNet	74.15%	80.67%	79.72%
LeNet-5	49.86%	49.38%	52.20%
GoogLeNet	89.62%	87.87%	88.01%
MobileNetV3	88.85%	85.06%	84.30%
YOLOv8	-	91.04%	90.83%
YOLOv11	-	91.39%	91.01%


Validation Accuracy (YOLOv8)
val acc
Accuracy
Epoch
0.65
0.70
0.75
0.80
0.85
0.90
2
4
6
8
10
12
14
16

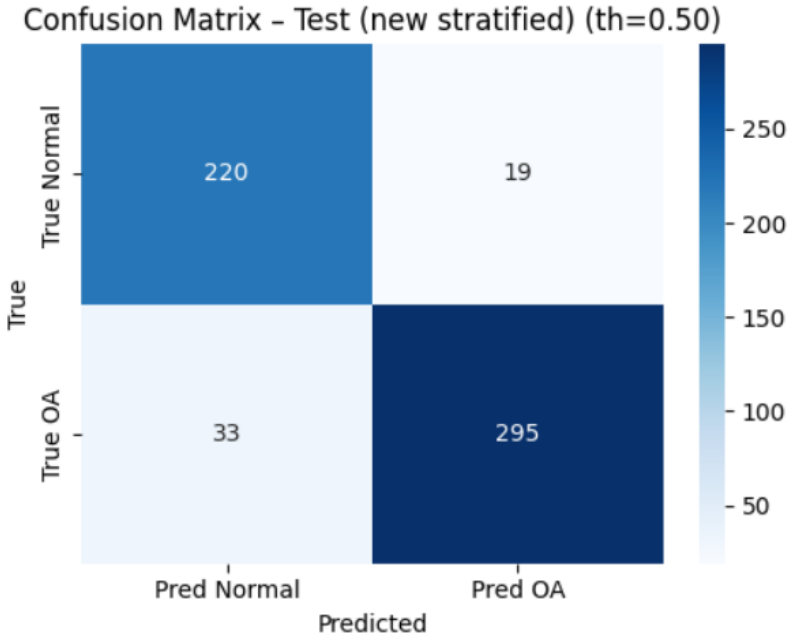

Loss
train loss
val loss
Loss
Epoch
0.2
0.3
0.4
0.5
0.6
0.7
2
4
6
8
10
12
14
16




```
=== Validation (YOLOv8, th=0.50) ===
              precision    recall  f1-score   support

      Normal       0.88      0.91      0.89       238
Osteoarthritis     0.93      0.91      0.92       331

    accuracy                           0.91       569
   macro avg       0.91      0.91      0.91       569
weighted avg       0.91      0.91      0.91       569
```




```
=== Test (new stratified) (YOLOv8, th=0.50) ===
                precision    recall  f1-score   support

        Normal       0.87      0.92      0.89       239
Osteoarthritis       0.94      0.90      0.92       328

      accuracy                           0.91       567
     macro avg       0.90      0.91      0.91       567
  weighted avg       0.91      0.91      0.91       567
```




```
from PIL import Image

img_path = "C:\\Users\\Anuradha\\Downloads\\KOA(X-rays)\\test\\test\\Normal\\9061666R.png"

# Predict
pred = best_model.predict(source=img_path, imgsz=IMG_SIZE, verbose=False)[0]
probs = pred.probs.data.cpu().numpy()
pred_idx = int(probs.argmax())
pred_label = best_model.names[pred_idx]
pred_conf = float(probs[pred_idx])

# Display
img = Image.open(img_path)
plt.figure(figsize=(5,5))
plt.imshow(img, cmap="gray")
plt.axis("off")
plt.title(f"Prediction: {pred_label} ({pred_conf*100:.2f}%)")
plt.show()
```

Objective 3 – IT22606792

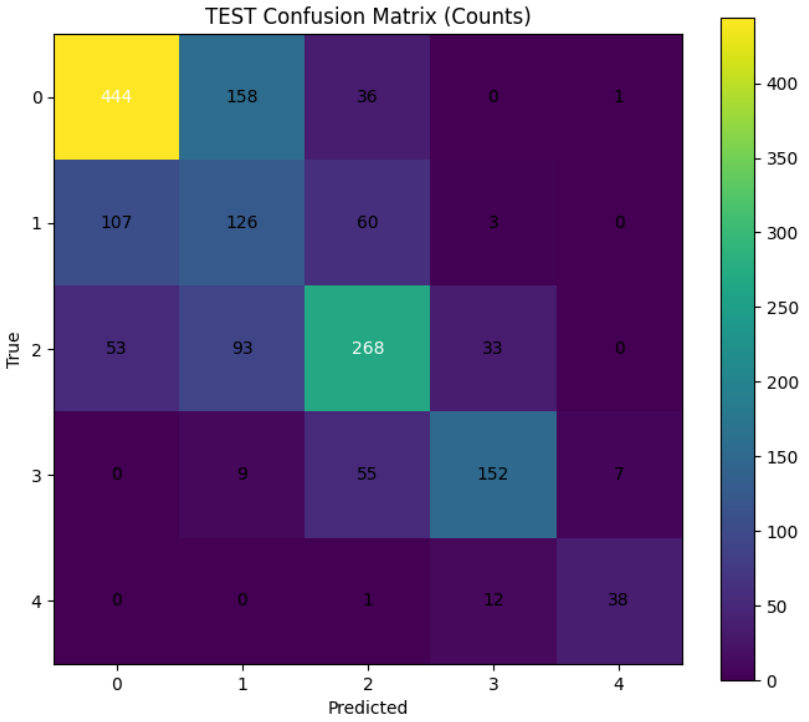
For the X-ray–based knee osteoarthritis severity prediction, deep learning models based on convolutional neural networks were trained using transfer learning. Architectures such as EfficientNetB0, DenseNet121 and ResNet50 demonstrated strong capability in learning discriminative visual features related to joint space narrowing, osteophyte formation, and sclerosis. The trained models achieved stable training and validation performance, indicating effective generalization across multiple severity grades.

In the clinical and biomarker-based prediction component, machine learning models such as XGBoost and Random Forest were trained on structured patient data. Tree-based models performed effectively in capturing non-linear relationships between demographic factors, clinical symptoms, and biomarker values. Feature importance analysis revealed that attributes such as age, BMI, inflammatory markers, and pain-related indicators contributed significantly to severity prediction.

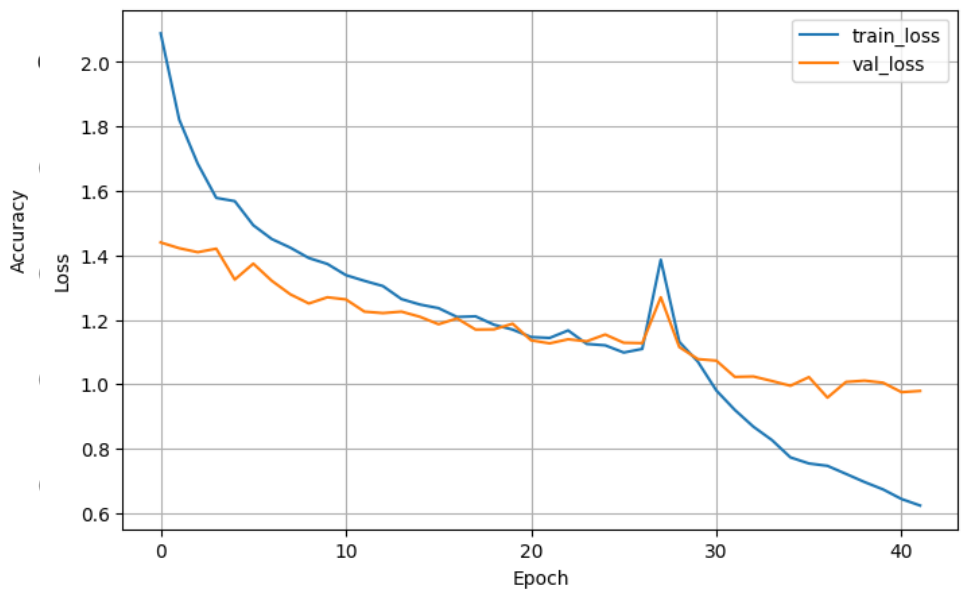
EfficientNetB0 - Classification Report

```
=== TEST Classification Report ===
              precision    recall  f1-score   support

           0     0.7351    0.6948    0.7144       639
           1     0.3264    0.4257    0.3695       296
           2     0.6381    0.5996    0.6182       447
           3     0.7600    0.6816    0.7187       223
           4     0.8261    0.7451    0.7835        51

    accuracy                         0.6208      1656
   macro avg     0.6571    0.6294    0.6409      1656
weighted avg     0.6420    0.6208    0.6295      1656
```

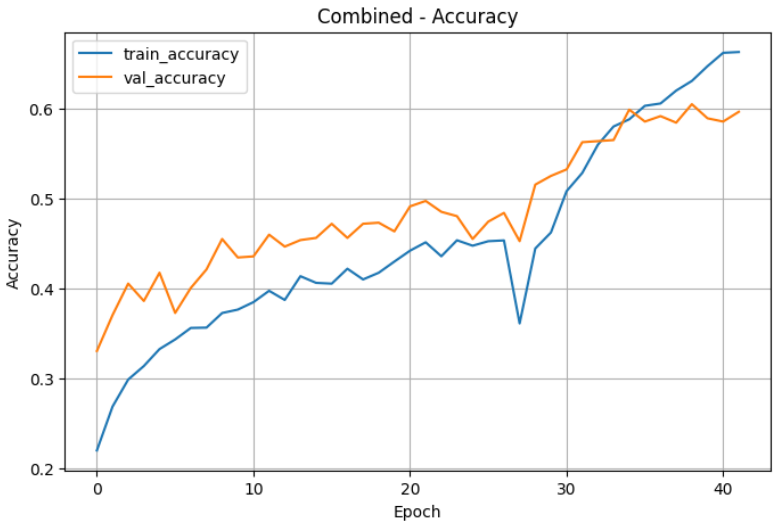
EfficientNetB0 – Confusion Matrix



EfficientNetB0 – Training and Validation Loss Curve



EfficientNetB0 – Training and Validation Accuracy Curve

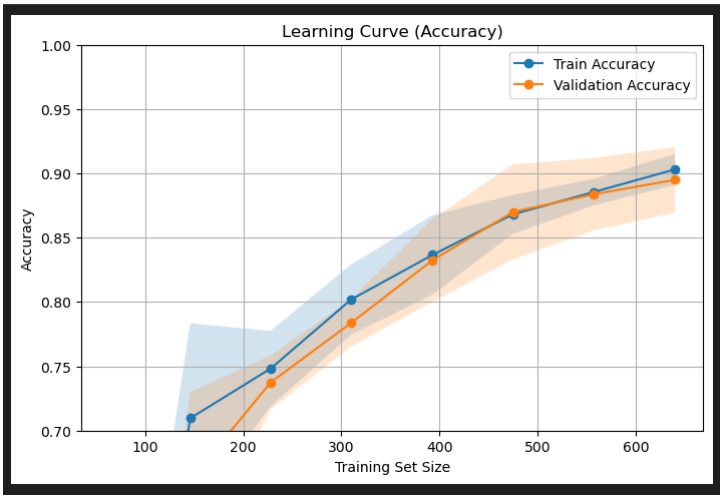


Feature Importance - XGBoost

Total features with permutation importance: 104

	feature	perm_importance	perm_std
3	pain_score	0.5195	0.033945
7	cs	0.0085	0.007089
6	platelets	0.0080	0.002449
11	rf	0.0075	0.004610
9	crp	0.0035	0.002291
4	fbs	0.0035	0.004500
10	esr	0.0035	0.002291
2	weight	0.0030	0.004583
5	wbc	0.0025	0.002500
17	physical_activity_level_Moderate	0.0020	0.003317

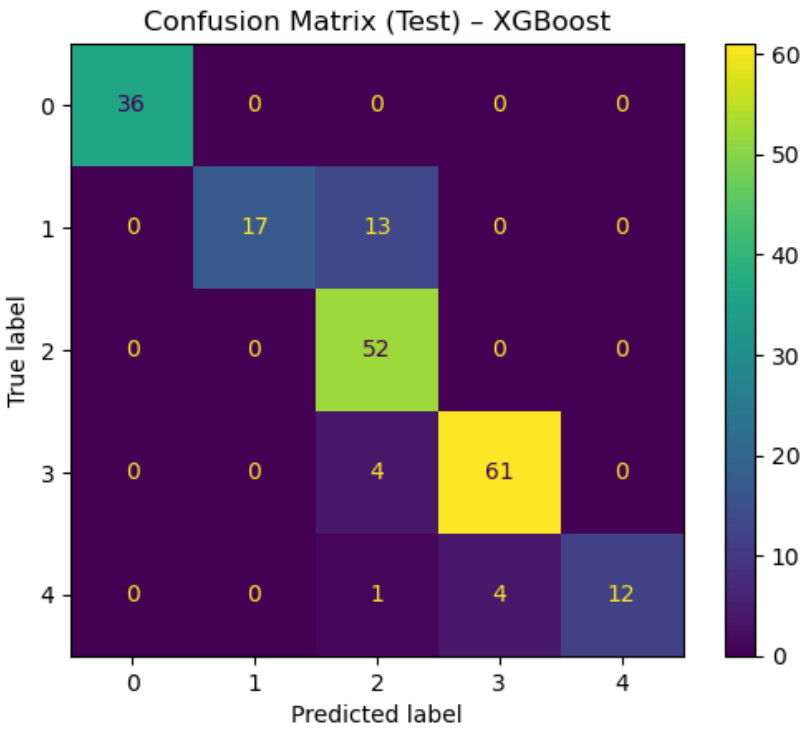
XGBoost - Learning Curve (Accuracy)



XGBoost – Classification Report

```
========================
Train/Val/Test Accuracy
========================
Train Accuracy: 0.8844
Val   Accuracy: 0.8688
Test  Accuracy: 0.8900
Overfitting gap (Train - Val): 0.0156

========================
Classification Report (TEST)
========================
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        36
           1       1.00      0.57      0.72        30
           2       0.74      1.00      0.85        52
           3       0.94      0.94      0.94        65
...
 [ 0 17 13  0  0]
 [ 0  0 52  0  0]
 [ 0  0  4 61  0]
 [ 0  0  1  4 12]]
```

XGBoost – Confusion Matrix



Objective 4 – IT22188472

Several supervised machine learning models were evaluated using the Vibroarthrography (VAG) signal dataset to classify knee health conditions. Based on comparative performance analysis, the Support Vector Machine (SVM) model was selected as the best-performing model for this research component.

Exploratory analysis showed that VAG features such as RMS amplitude, spectral entropy, and peak frequency exhibited clear separations between normal and osteoarthritic knee conditions. These patterns were visualized using distribution plots and correlation analysis, highlighting strong relationships between vibration irregularity and knee abnormalities.

The SVM model demonstrated strong generalization ability, achieving stable accuracy across training and validation datasets. Confusion matrix analysis showed balanced classification performance, while decision boundary visualization confirmed the effectiveness of SVM in separating complex, non-linear patterns in VAG feature space.

Overall, the results confirm that SVM is well suited for vibration-based knee health classification and supports its use in AI-driven, IoT-based continuous knee monitoring systems.

```
Fitting 5 folds for each of 50 candidates, totalling 250 fits
Best Hyperparameters: {'C': 50, 'class_weight': None, 'gamma': 1}
==================================================
Best SVM MODEL (RBF KERNEL) After Hyperparameter Tuning
Training Accuracy : 88.17 %
Validation Accuracy: 77.20 %
Test Accuracy      : 80.40 %

Classification Report ----------------------------
              precision    recall  f1-score   support

        Mild       0.55      0.78      0.65        27
    Moderate       0.70      0.68      0.69        71
      Normal       0.94      0.90      0.92        80
      Severe       0.79      0.72      0.75        72

    accuracy                           0.77       250
   macro avg       0.74      0.77      0.75       250
weighted avg       0.78      0.77      0.78       250

==================================================
```

Figure 1 Classification Report

4.2. Challenges Faced During Data Analysis:

Objective 1 – IT22582942

A key challenge in the clinical and biomarker–based KOA prediction component was the limited availability of real patient data, which restricted initial dataset size. Missing and inconsistent biomarker values required extensive preprocessing to ensure data quality. Additionally, variability in clinical symptoms and laboratory measurements introduced noise, making it difficult to achieve highly separable classes. While synthetic data was used to improve dataset size, maintaining medical realism and minimizing bias remained an ongoing challenge.

Objective 2 – IT22223708

Several challenges were encountered during the data analysis and model development process of this research component. One major challenge was class imbalance within the knee X-ray dataset, where osteoarthritis and normal cases were not equally distributed. This affected model generalization, particularly during testing, and required the use of stratified dataset splitting, class weighting, and threshold tuning to mitigate biased predictions.

Another challenge involved data quality and variability in medical images. Knee X-ray images were collected from different sources and exhibited variations in resolution, contrast, orientation, and imaging conditions. These inconsistencies made it difficult for models to learn stable patterns, necessitating the application of image enhancement techniques such as CLAHE and controlled image augmentation to improve feature visibility and robustness.

Overfitting was also observed during training, where models achieved high training and validation accuracy but performed poorly on unseen test data. This required careful use of two-stage training, regularization techniques, dropout layers, and early stopping to improve generalization performance.

In addition, computational limitations posed a significant challenge. Training deep learning models such as EfficientNet, Inception, and YOLO-based classifiers on a CPU-based system resulted in long training times and limited experimentation. Model complexity, image resolution, and batch size had to be carefully balanced to ensure feasible training within hardware constraints.

Finally, managing model training continuity in Jupyter Notebook environments was challenging, as interruptions led to loss of in-memory variables. This highlighted the importance of periodic model checkpointing and weight saving to avoid retraining and ensure reproducibility.

Objective 3 – IT22606792

One major challenge encountered during image-based analysis was class imbalance across Kellgren–Lawrence severity grades, particularly for higher severity levels. Additionally, variations in image quality, contrast, and acquisition conditions introduced noise, making it difficult to consistently distinguish between adjacent severity classes. The subtle visual differences between early and moderate stages of osteoarthritis further increased classification complexity.

For the clinical and biomarker dataset, limited availability of real patient records posed a challenge for training robust machine learning models. Missing values, inconsistent data entry, and variability in biomarker measurements required careful preprocessing and validation. Although synthetic data generation was used to improve dataset size, maintaining medical realism and avoiding bias remained a key challenge.

Another challenge involved integrating heterogeneous data types from imaging and clinical sources. Aligning predictions from deep learning models with outputs from traditional machine learning models required careful design of the fusion strategy to ensure consistency and reliability.

Objective 4 – IT22188472

One major challenge during data analysis was the overlap of vibration patterns between adjacent severity levels, particularly between mild and moderate Knee Osteoarthritis. This made class separation difficult, as early-stage conditions produce subtle signal differences that are hard to distinguish using vibration-based features alone.

Another challenge was class imbalance, where certain severity levels had fewer samples compared to others. This affected model learning and required careful evaluation using validation accuracy and confusion matrices rather than relying solely on overall accuracy.

Additionally, noise and variability in VAG signals caused by sensor placement, movement speed, and individual gait differences introduced inconsistencies in the data. Extensive preprocessing and feature standardization were required to ensure stable model performance.

Finally, selecting an optimal model and hyperparameters while avoiding overfitting was challenging due to the non-linear nature of sensor data. Hyperparameter tuning using cross-validation was necessary to achieve a balance between model complexity and generalization performance.

5. References

https://www.kaggle.com/datasets/shashwatwork/knee-osteoarthritis-dataset-with-severity

https://www.kaggle.com/datasets/farjanakabirsamanta/osteoarthritis-prediction

https://www.kaggle.com/datasets/ziya07/knee-health-dataset-using-vag-signals-for-ai-and-iot