A PROJECT REPORT ON

**OSTEOARTHRITIS PREDICTION WITH MACHINE LEARNING**

Submitted by [SE\_C\_02]

**Student Name Roll Number**

**Shilpa Patil 24**

**Omkar Tilekar 55**

**Yash Limbhore 66**

**Anuj Sharma 23**

Under The Guidance of

### Prof. Ashwini Shinde

**(Asst. Prof. Dept. of AI&DS Engg.)**

Towards the fulfilment of Second Year Under Graduate Course-Project Based Learning-II in Artificial Intelligence and Data Science Engineering

of

SAVITRIBAI PHULE PUNE UNIVERSITY (2024-25)



Department of Artificial Intelligence and Data Science Engineering Ajeenkya DY Patil School of Engineering

Dr. D.Y. Patil Knowledge City, Charholi (Bk), Lohegaon, Pune - 412105



Department of Artificial Intelligence and Data Science Engineering Ajeenkya DY Patil School of Engineering

Dr. D.Y. Patil Knowledge City, Charholi (Bk), Lohegaon, Pune - 412105

CERTIFICATE

This is to certify that Project Report

on

“**OSTEOARTHRITIS PREDICTION WITH MACHINE LEARNING**”

Submitted by

**Student Name Roll Number**

**Shilpa Patil 24**

**Omkar Tilekar 55**

**Yash Limbhore 66**

**Anuj Sharma 23**

Towards the Fulfilment of Second Year under Graduate Course in Artificial Intelligence and Data Science Engineering

of

SAVITRIBAI PHULE PUNE UNIVERSITY (2024-2025)

#### Prof. Ashwini Shinde Dr. Farook Sayyad

Guide Principal

# ACKNOWLEDGEMENT

It gives me great pleasure and immense satisfaction to present this project report on “**OSTEOARTHRITIS PREDICTION WITH MACHINE LEARNING**”, which is the result of the unwavering support, expert guidance, and focused direction of my guide **Prof. Ashwini Shinde** to whom I express my deep sense of gratitude and humble thanks, for her valuable guidance throughout the presentation work. The success of this project has throughout depended upon an exact blend of hard work and unending cooperation and guidance, extended to me by the superiors at our college.

Furthermore, I am indebted to **Dr. Bhagyashree Dhakulkar,** HoD of Artificial Intelligence and Data Science Engineering and **Dr. Farook Sayyad,** the Principal whose constant encouragement and motivation inspired me to do my best.

Last but not least I sincerely thank to my colleagues, the staff, and all others who directly or indirectly helped us and made numerous suggestions that have surely improved the quality of my work.

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sr. No** | **Contents** | **Page No.** |
| **1** | **ABSTRACT** | **1** |
| **2** | **INTRODUCTION** | **2** |
| **3** | **MOTIVATION** | **3** |
| **4** | **PROBLEM STATEMENT** | **4** |
| **5** | **LITERATURE SURVEY** | **5** |
| **6** | **PROPOSED SYSTEM** | **7** |
| **7** | **METHODOLGY** | **10** |
| **8** | **SOFTWARE AND HARDWARE REQUIREMENTS** | **12** |
| **9** | **IMPLEMENTATION** | **13** |
| **10** | **RESULTS** | **15** |
| **11** | **ADVANTAGES** | **16** |
| **12** | **CONCLUSION** | **17** |
| **13** | **FUTURE SCOPE** | **18** |
| **14** | **REFERENCES** | **19** |
| **15** | **PARTICIPATION/PUBLICATION CERTIFICATES** | **20** |

# List of Figures

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **Figure Number** | **Figure Name** | **Page Number** |
| 1 | **System Architecture** | 14 |
| 2 | **Streamlit interface (Grade 4)** | 18 |
| 3 | **Streamlit interface (Grade 2)** | 19 |
|  |  |  |

## ABSTRACT

## Osteoarthritis (OA) is a long-term, degenerative joint condition that impacts millions of individuals globally, causing discomfort, stiffness, and decreases mobility overtime. Early identification and diagnosis play a vital role in slowing disease progression and enhancing the patient's overall quality of life. Conventional diagnostic methods, such as X-ray imaging and clinical assessments, provide reliable results but can be time-consuming.

## Therefore, using Machine Learning(ML), OA can be predicted before time ans irreversible changes can be avoided. .The model presents Deep Learning-based approach that leverages convolutional neural netwroks(CNNs) for classifying OA using X-ray images. The feature extractor that model uses is EfficientNetB0 and custom dense layers for classification.

## The proposed model demonstrates optimistic results in distinguishing between different OA severity levels that is (Grade 0-4), which offers a scalable and efficient model for clinical usage.

## By integrating deep learning with medical imaging(X-rays) and other patient's metadata, this model aims to automate OA, enhance treatments outcomes, support clinical decisions and enable timely interventions.

## Future work will focus on training model with larger dataset and real time clinical applications.

## INTRODUCTION

Osteoarthritis, the most common form of Arthritis, is a chronic progressive joint disorder that affects millions of people worldwide leading to pain, stiffness, and reduced mobility. OA is caused because of biological, genetic and mechanical aspects combined altogether. Key aspects that cause OA are age, injury, genetics, obesity.It is characterized by the cartilage degradation and inflammation in joints. This condition usually affects the weight bearing joints, for example knees, spine, hips leading to decreased quality of life of patients. There is no permanent cure for arthritis, but several measures can be taken to ease the pain and discomfort it causes.

There are more than 180 million cases of arthritis in India alone. Out of this, a study done in 2024 reported that the number of OA patients rose from approximately 23.46 million in 1990 to about 62.35 million in 2019.Therefore, early diagnosis and severity assessment are very crucial to slow down the disease progression, effective treatment and management, to improve treatment outcomes and enhance patient's quality of life.

OA prediction and classification can be automated with the help of machine learning. Machine Learning is the study of how a machine learns from experience through data and outcomes, which later finds patterns in the dataset and make predictions using algoritms. ML emerged as a transformative tool in medical field and diagnosis. Deep Learning has demonstrated great results in medical image analysis and is proved better than the traditional computer vision methods and techniques. Convolutional Neural Networks (CNNs) particularly have been widely used feature extraction and classifying medical images(X-rays).By understanding the patterns within these images, CNNs can achieve high accuracy in detecting OA and grading its severity.

Our aim is to develop an AI-driven solution for early osteoarthritis (OA) prediction using deep learning-based image analysis. This system leverages medical imaging (X-ray scans) and patient metadata to accurately classify OA severity levels and assist healthcare professionals in making timely treatment decisions.The system is designed to process X-ray images using deep learning models, particularly Convolutional Neural Networks (CNNs), for automated OA detection and severity classification.This process and data can help predict OA efficiently however, using patient metadata such as medical history, age, and weight along with X-rays in the diagnosis can be found helpful. This will improve the accuracy of the predictive model.

## MOTIVATION

Challenges with Traditional OA Diagnosis

1. Late-stage detection: OA is often diagnosed only when significant joint damage has occurred.
2. Time-consuming & expensive: MRI scans and other advanced diagnostics are not always accessible.

Therefore it is necessary to detect osteoarthritis at an early stage to prevent irreversible damages.

1. Early Intervention-Early prediction allows for prompt treatment. This reduces disease progression.
2. Improved Outcomes-Timely interventions improve patient quality of life.
3. Resource Allocation-Prediction helps allocate healthcare resources efficiently.

## PROBLEM STATEMENT

## Develop an AI-driven system using machine learning for early osteoarthritis prediction.

## By combining medical imaging, clinical data, and predictive modelling, this approach enables timely interventions, personalized treatment plans, and better patient outcomes.

## 5. LITERATURE SURVEY

The incorporation of Machine Learning and Deep Learning in the medical field has yielded remarkable advancements in recent years, positively impacting treatment methodologies. Previous studies have investigated the application of machine learning for osteoarthritis (OA) detection, utilizing techniques such as support vector machines (SVM), random forests, and convolutional neural networks (CNNs) to enhance diagnostic accuracy and efficiency.

Among these, deep learning models—especially CNN-based architectures—have shown remarkable effectiveness in medical image classification. Several studies have leveraged pretrained models like ResNet, VGG, and DenseNet to enhance OA prediction accuracy.

Various studies have confirmed that this integration results in automated methods for OA classification and prediction.

In the following study, multiple pre-trained deep learning models, including ResNet101 and DenseNet121, were utilized for the classification of osteoarthritis (OA) severity. These models, known for their advanced feature extraction capabilities, were fine-tuned to analyze X-ray images and differentiate between varying OA severity levels.The outcomes demonstrated improved accuracy with residual neural networks, achieving approximately 89% accuracy in multi-class classification.This improvement highlights the effectiveness of deep learning in detecting intricate patterns in medical imaging, leading to more precise OA severity grading. Additionally, the use of residual connections in ResNet101 helped overcome vanishing gradient issues, enabling deeper networks to learn more efficiently.By leveraging these advanced architectures, the study reinforces the potential of AI-driven diagnostic tools in reducing human errors, enhancing early detection, and supporting clinical decision-making. Future enhancements, such as integrating patient metadata and Explainable AI (XAI), could further refine predictions and ensure wider clinical applicability.[7].

This study incorporates Explainable AI (XAI) techniques, such as Grad-CAM, alongside deep learning models like EfficientNetB7, to enhance the interpretability and reliability of OA severity classification. By utilizing Grad-CAM, the model provides visual explanations of its predictions, helping clinicians understand which regions of the X-ray contribute most to the diagnosis.The study demonstrated that the model could effectively distinguish between severe and normal OA cases, achieving an impressive accuracy of up to 99.13%. This high accuracy underscores the potential of deep learning in automated OA detection, reducing subjectivity in traditional assessments.Moreover, integrating XAI techniques ensures greater transparency in AI-driven medical diagnoses, increasing trust among healthcare professionals. Future advancements could focus on incorporating patient metadata and expanding datasets to further enhance model generalization and real-world applicability.[12].

This study explores various machine learning approaches and their applications in knee osteoarthritis (OA) research, focusing on key areas such as prediction, segmentation, classification, and disease progression. By leveraging advanced ML techniques, researchers aim to develop more accurate and automated diagnostic tools for OA assessment.The study emphasizes the critical role of Explainable AI (XAI) in enhancing the accuracy, reliability, and trustworthiness of these models. XAI ensures that AI-driven decisions are transparent and interpretable, making them more suitable for clinical deployment.Furthermore, the research highlights the potential of integrating multi-modal data, including patient history, imaging, and biomarkers, to improve OA diagnosis and progression tracking.

Future advancements may focus on refining model generalization and ensuring seamless integration into healthcare systems for widespread clinical adoption.[14]

This study provides us an overview of machine learning techniques and algorithms applicable to osteoarthritis (OA) prediction, analysis, and treatment planning. It explores how ML-based models can enhance early detection, severity classification, and personalized treatment strategies, offering a more data-driven approach to OA management.The review emphasizes the importance of diverse and extensive datasets to improve the robustness and generalizability of AI models. It emphasis on the need for continued research and collaboration to develop more effective predictive tools.Additionally, the study suggests that incorporating multi-source data—including medical history, lifestyle factors, and imaging data—can significantly enhance model performance, ultimately leading to better patient outcomes and improved clinical decision-making.[13].

This study provides a detailed review of research articles published between 2006 and 2020, focusing on the applications of machine learning (ML) in osteoarthritis (OA) diagnosis and management. It categorizes ML applications into four key areas:Classification, Segmentation, Predictions or Regression and Optimum Post-Treatment Planning.The study highlights the core characteristics of ML algorithms, comparing their performance, advantages, and limitations in automated OA detection and analysis. It underscores the importance of deep learning, feature engineering, and hybrid AI models to improve diagnostic accuracy and enhance clinical decision-making. Future research directions include integrating Explainable AI (XAI), federated learning for data privacy, and multi-modal data fusion to refine ML-driven OA assessments.[11]

All these studies are inclined towards making OA diagnosis easy and to provide AI-driven solutions to improve the existing traditional methods of OA prediction.

**6. PROPOSED SYSTEM**

The proposed system aims to develop an AI-driven solution for early osteoarthritis (OA) prediction using deep learning-based image analysis.

This system leverages medical imaging (X-ray scans) and patient metadata to accurately classify OA severity levels and assist healthcare professionals in making timely treatment decisions.

System Overview

The system is designed to process X-ray images using deep learning models, particularly Convolutional Neural Networks (CNNs), for automated OA detection and severity classification.

Additionally, patient metadata such as age, weight, and medical history may be incorporated to enhance prediction accuracy.

Key Components:

1.) Data Collection and Preprocessing:

* Acquisition of X-ray images and patient metadata.
* Image enhancement techniques such as contrast adjustment and noise reduction.
* Data augmentation (rotation, flipping, scaling) to improve model generalization.

2.) Deep Learning Model for OA Prediction

* Pretrained CNN Backbone: EfficientNetB0 for feature extraction.
* Fully Connected Layers: Custom dense layers for classification.
* Softmax Output Layer: Multi-class classification (Grades 0-4).

3.) Training and Optimization:

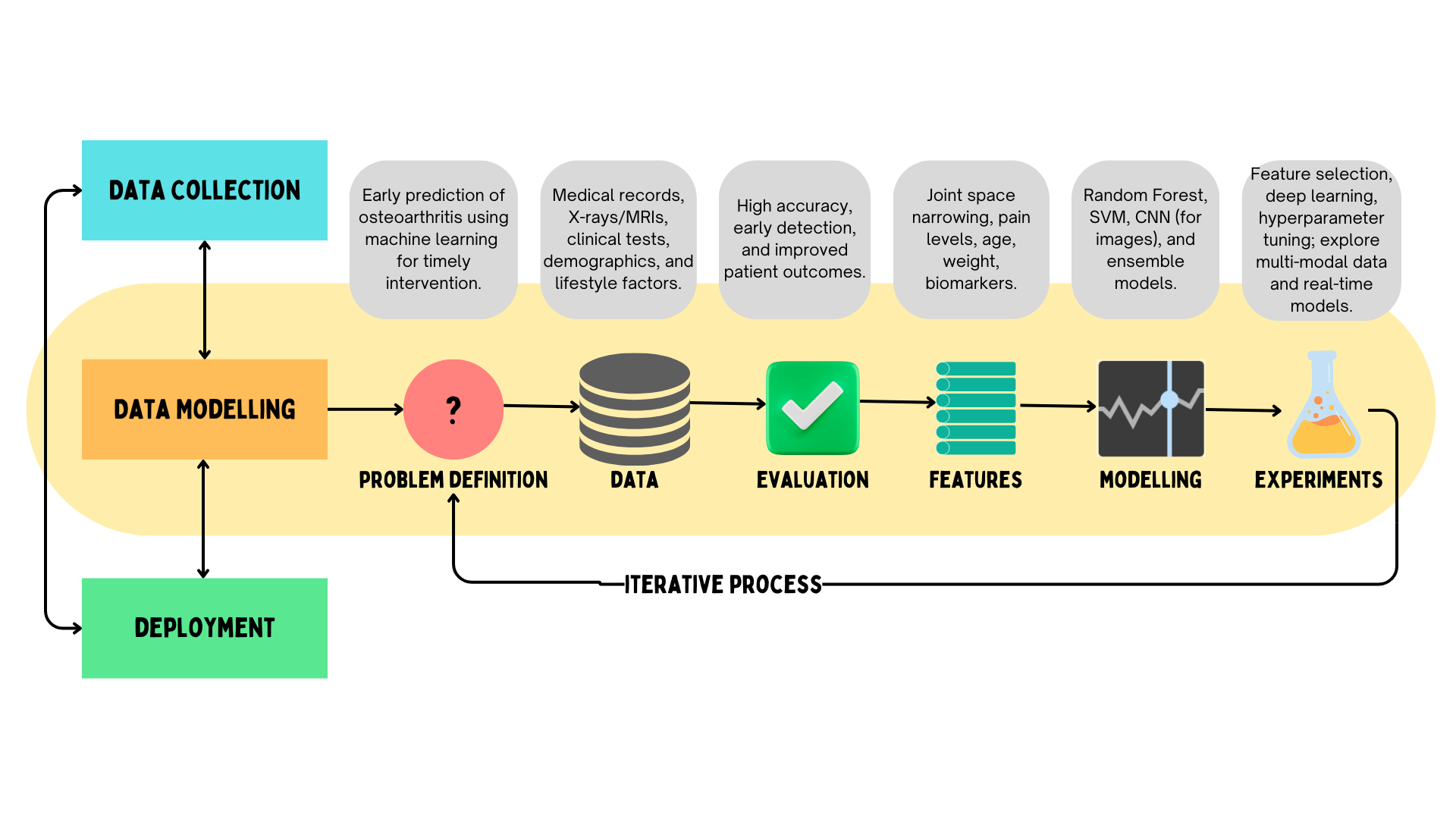
* Loss Function: Categorical cross-entropy for multi-class classification.
* Optimizer: AdamW with weight decay for improved regularization.
* Learning Rate Scheduler: ReduceLROnPlateau to adaptively adjust learning rates.
* Class Imbalance Handling: Weighted loss to prevent bias toward majority classes.

4.) Model Evaluation and Deployment:

* Performance assessment using accuracy, precision, recall, and F1-score.
* Deployment via web-based or cloud-based API for real-time predictions.
* Integration with hospital management systems for clinical applications.

**Expected Outcomes**

* Early OA Detection: Identification of osteoarthritis at an early stage.
* Improved Diagnostic Accuracy: Reduction in human error in radiographic assessments.
* Scalability and Accessibility: AI-driven diagnosis for remote healthcare applications.
* Personalized Treatment Planning: Tailoring interventions based on patient-specific data.

****

**Figure 1. System Architecture**

The above image shows a systematic approach is classifying osteoarthritis using machine learning. Following are the processes:

1. **Data Collection** – Gathering medical records, X-rays, MRIs, clinical tests, demographics, and lifestyle factors.
2. **Problem Definition** – Defining the objective, such as early prediction of osteoarthritis for timely intervention.
3. **Data** – Storing and managing relevant medical data.
4. **Evaluation** – Assessing model performance to ensure high accuracy and improved patient outcomes.
5. **Feature Engineering** – Selecting critical features such as joint space narrowing, pain levels, age, weight, and biomarkers.
6. **Modelling** – Using machine learning techniques like Random Forest, SVM, CNN, and ensemble models.
7. **Experiments** – Refining models through feature selection, deep learning, hyperparameter tuning, and exploring multi-modal data.
8. **Deployment** – Implementing the trained model for real-world usage.

9.**Iterative Process** – Continuous refinement based on evaluation to improve accuracy and effectiveness.

# METHODOLOGY

## A.MODEL ARCHITECTURE

## 1)Pretrained Backbone:

a pretrained DL model is for extracting relevant features from the input images that is xrays.

## - used for feature extraction

## -include\_top=False : to remove original classifier

## -weights='imagenet': to leverage pretrained weights.

## -trainable=False : freezes weights to prevent overfitting.

## 2) Feature Extraction (Global Average Pooling)

## - it converts extracted features into a fixed-size vector.

-also helps to reduce the parameters and enhance efficiency of the model

## 3)Fully Connected Layers (Dense Layers)

## -Multiple Dense layers with decreasing neurons.

## -L2 Regularization (l2(0.01)) to prevent overfitting.

## -Dropout (0.3) for better generalization.

## 4) Final Output Layer

## -Uses softmax activation for multi-class classification(Grades 0-4).

- The output is the probability scores for every class which leads to accurate OA severity classification.

## B.TRAINING DETAILS

1. **Loss function:**

Categorical Crossentropy: Measures the divergence between predicted and actual class probabilities, making it ideal for multi-class classification tasks.

1. **Optimer:**

AdamW: A variant of the Adam optimizer that integrates weight decay (weight\_decay=0.01) to enhance model regularization and minimize overfitting.

1. **Learning Rate Adjustment:**

ReduceLROnPlateau: Automatically lowers the learning rate by a factor of 0.5 and patience=3 when the validation loss (val\_loss) stagnates for three consecutive epochs, facilitating better convergence.

1. **Data Augmentation:**

ImageDataGenerator - Enhances dataset variability by applying transformations such as rotations, flipping, and scaling, improving model generalization.

1. **Handling Class Imbalance:**

compute\_class\_weight: Determines appropriate class weights to ensure balanced learning, preventing underrepresented classes from being overshadowed during training.

## 8. SOFTWARE AND HARDWARE REQUIREMENTS

## SOFTWARE REQUIREMENTS

1. Streamlit Module
2. Tensorflow
3. Scikitlearn
4. Os module
5. Keras
6. Jupyter notebook/Google colab

## HARDWARE REQUIREMENTS

1. Computer/laptop
2. Medical imaging source(X-rays)

## 9. IMPLEMENTATION

## WhatsApp Image 2025-03-25 at 3.27.00 PM

**Figure 2. Streamlit interface (Grade 4)**



**Figure 3. Streamlit interface (Grade 2)**

The image showcases the interface created using streamlit of an osteoarthritis X-ray Classification System, which utilizes

deep learning to assess the severity of arthritis based on X-ray images. The system allows users to upload an X-ray image and automatically predicts the severity grade.

In this instance, a knee joint X-ray has been uploaded, and the system has classified the arthritis severity as Grade 4, indicating severe arthritis.

This classification system can aid in the early diagnosis and assessment of osteoarthritis, assisting medical professionals in decision-making and treatment planning resulting in improving patients quality of life.

## 10. RESULTS

The trained model was evaluated on a test set and its performance metrics were recorded.

The following are the results:

| **Performance metrics** | **Values** |
| --- | --- |
| Classification Reshape Accuracy | 58.50% |
| Classification Reshape Loss | 1.1729 |

-Overall Loss: 1.8535

Currently, the model achieves an accuracy of 58.50%, which is expected to improve as it is trained on a larger and more diverse dataset. Expanding the dataset will enhance the model’s ability to recognize patterns, reduce overfitting, and improve generalization.

Additionally, fine-tuning hyperparameters, optimizing data augmentation techniques, and incorporating patient metadata could further boost performance, making the model more reliable for real-world clinical applications.

## 11.ADVANTAGES

**Early Detection & Intervention** – Machine learning models can identify OA in its initial stages, allowing for timely medical intervention to slow disease progression and improve patient outcomes.

**Enhanced Diagnostic Accuracy** – AI-powered systems minimize human errors and inconsistencies in radiographic assessments, leading to more reliable and standardized diagnoses.

**Cost-Effective Approach** – By utilizing predictive features from alternative imaging techniques, ML models can reduce dependence on expensive diagnostic tools such as MRIs, making OA detection more affordable.

**Scalability & Accessibility** – Once trained, ML models can be deployed across various healthcare facilities, making advanced OA diagnostics accessible to a broader population.

**Potential for Remote Screening** – AI-driven diagnostic tools can be incorporated into telemedicine platforms, enabling OA screening and early detection in remote or underserved areas, improving healthcare accessibility.

**Explainability & Clinical Decision Support** – Explainable AI (XAI) techniques can provide insights into key risk factors, helping clinicians understand model predictions and make informed treatment decisions.

**Personalized Treatment Plans** – AI-driven models can customize treatment strategies based on individual patient risk factors, ensuring a more targeted and effective approach to disease management.

**Multi-Modal Data Integration** – Machine learning can analyze and correlate multiple data sources, including clinical records, imaging scans, genetic information, and biomarkers, for more comprehensive and accurate prediction

## CONCLUSION

### In a nutshell, AI integration in osteoarthritis detection enhances diagnostic accuracy, enables early detection, and improves patient care.

### Advanced algorithms like CNN analyse medical imaging with high precision, paving the way for AI-driven personalized healthcare.

### As technology evolves, machine learning promises to revolutionize OA diagnosis by improving speed and accuracy.

### FUTURE SCOPE

1. **Wearable Sensor Integration** – The combination of IoT and AI can enable real-time gait analysis, allowing early detection of OA symptoms through continuous movement tracking.

**2. Personalized Treatment Plans** – Machine learning models can predict patient-specific responses to various treatment options, enabling more tailored and effective therapeutic strategies.

**3. Federated Learning for Data Privacy** – This approach allows AI models to be trained on datasets from multiple hospitals without sharing sensitive patient information, ensuring privacy while improving accuracy.

**4. Explainable AI (XAI)** – By incorporating interpretability methods, AI models can provide transparent and understandable insights, assisting clinicians in making informed diagnostic and treatment decisions.

**5. Deployment in Healthcare Systems** – AI-powered diagnostic tools can be seamlessly integrated into hospital workflows, enhancing efficiency, reducing diagnostic time, and supporting healthcare professionals in decision-making.

**6.Multi-Modal Data Fusion** – Combining imaging data, patient history, genetic markers, and real-time sensor data can provide a more comprehensive understanding of OA progression and risk factors.

**7. Predictive Analytics for Disease Progression** – AI models can analyze patient data to forecast OA progression, helping doctors intervene early and recommend lifestyle modifications or preventive measures.

**8.Automated Severity Grading** – AI can be trained to automatically classify OA severity levels based on X-ray and MRI scans, reducing the subjectivity in manual grading.

**9.Cloud-Based AI Solutions** – Cloud-based AI platforms can make OA diagnosis more accessible, allowing healthcare providers worldwide to utilize powerful AI models without needing high-end local hardware.

**10.AI-Driven Drug Discovery for OA Treatment** – Machine learning can accelerate drug discovery by identifying potential therapeutic compounds and predicting their effectiveness, leading to faster development of OA treatments.

## REFERENCES

[1]Sabat, S. K., Panda, S., Merli, P. K., & Bhuyan, S. K.. Prevalence and correlates of arthritis in Indian older adults: Findings from the longitudinal aging study of India.International Journal of Health Science (2022)

[2]"Deep Learning for Osteoarthritis Diagnosis: A Review," IEEE, 2023

[3]"Predicting Osteoarthritis Progression Using Machine Learning," Nature Medicine, 2022

[4]"Automated Knee Osteoarthritis Detection via CNNs," Medical Imaging Journal, 2021

[5]Prevalence of the Knee Osteoarthritis Risk Factors Among Young Adult Population - An Observational Study, 2023

[6]Exploring the Burden of Knee Osteoarthritis in Rural South India: Community Prevalence, Risk Factors, and Functional Assessment Among Adults Aged 40 and Above, 2024

[7]A. S. Mohammed, A. A. Hasanaath, G. Latif and A. Bashar, "Knee osteoarthritis detection and severity classification using residual neural networks on preprocessed x-ray images", 2023.

[8]Simone Castagno, Benjamin Gompels, Estelle Strangmark, Eve Robertson-Waters, Mark Birch, Mihaela van der Schaar, Andrew W McCaskie.Understanding the role of machine learning in predicting progression of osteoarthritis. 2024

[9]Taghi Ramazanian, Sunyang Fu, Sunghwan Sohn.Prediction Models for Knee Osteoarthritis: Review of Current Models and Future Directions. 2023

[10]Anran Xuan, Haowei Chen,Zhaohua Zhu. The application of machine learning in early diagnosis of osteoarthritis. .2023

[11]C. Kokkotis, S. Moustakidis, E. Papageorgiou, G. Giakas, D.E. Tsaopoulos .Machine learning in knee osteoarthritis: A review. 2020

[12]Rafique Ahmed, Ali Shariq Imran .Knee Osteoarthritis Analysis Using Deep Learning and XAI on X-Rays. 2024

[13]Marie Binvignat ,Valentina Pedoia, Atul J Butte, Karine Louati, David Klatzmann, Francis Berenbaum, Encarnita Mariotti-Ferrandiz, Jérémie Sellam.Use of machine learning in osteoarthritis research: a systematic literature review. (2022)

[14]Suman Rani, Minakshi Memoria , Tanupriya Choudhury and Ayan Sar. A Comprehensive Review of Machine Learning’s Role within KOA

1. **PARTICIPATION/PUBLICATION CERTIFICATES**