

# Intelligent Knee Osteoarthritis Prediction with Optimal Feature Driven Neural Network

Karramareddy Sharmila<sup>1\*</sup>, Usthumuri Penchalaiah<sup>2</sup>, Udaykiran Bhargava Gollamandala<sup>3</sup>

<sup>1</sup>Department of Computer Science and Engineering, Vaagdevi Engineering College, Bollikunta, Warangal, Telangana 506005

<sup>2</sup>Department of Electronics and Communication Engineering, Geethanjali Institute of Science and Technology, Gangavaram, S.P.S.R, Andhra Pradesh 524137

<sup>3</sup>Department of Electronics and Communication Engineering, Mother Teresa Institute of Science and Technology, Sathupally, Telangana 507303

\*Correspondence: Karramareddy Sharmila ([sharmilakreddy@gmail.com](mailto:sharmilakreddy@gmail.com))

## Abstract

Knee osteoarthritis (OA) is a common degenerative joint disorder that affects a large portion of the adult population. Early and accurate prediction of OA progression is essential for enabling timely clinical interventions and designing personalized treatment strategies. In this study, we present an artificial neural network (ANN)-based framework for predicting knee OA progression using 3D magnetic resonance imaging (MRI) data. The process begins with comprehensive preprocessing of MRI scans to ensure consistency and quality, including resolution standardization, intensity normalization, and alignment of knee joint orientation. Image segmentation is also employed to isolate the knee region and remove irrelevant anatomical structures, thereby reducing computational complexity. Following preprocessing, principal component analysis (PCA) is applied to extract a compact and informative set of features. By identifying principal components that capture the most significant variations in the dataset, PCA effectively transforms the high-dimensional MRI data into a lower-dimensional feature space while retaining essential diagnostic information. These reduced PCA features are then used as input to the ANN classifier, which is trained on labeled MRI samples representing varying levels of OA severity. The ANN is optimized using an appropriate loss function and iterative weight adjustment to minimize classification error. Once fully trained, the model can accurately classify new MRI scans into different stages of knee OA progression, demonstrating its potential as a reliable tool for early diagnosis and clinical decision support.

**Keywords:** Knee osteoarthritis, 3D MRI data, Artificial neural network, Dataset preprocessing, Principal component analysis, Feature extraction.

## 1. Introduction

The kind of joint condition that affects the greatest number of persons in the United States is OA, which is sometimes abbreviated as OA. Knee OA is a disease that is uncomfortable and is the major cause of loss of ability to undertake day-to-day activities such as walking and climbing stairs. It is also the most common reason of knee replacement surgery [1]. There is a correlation between becoming older and developing knee OA, which is distinguished by a decreased amount of articular cartilage. Because it causes damage to a broad range of articular tissues, OA is referred to be a "whole organ" disorder. This is because OA causes damage to the hyaline cartilage, the meniscus, the periarticular bone, the ligaments, and the tendons in the joint. As a result, this condition causes the symptoms described above. even though it is of the highest significance to preserve the health of the community. The primary weight-bearing joint that is most often affected by OA is the knee joint, which is made up of three different compartments (the medial tibiofemoral, the lateral tibiofemoral, and the patellofemoral). The knee joint is also known as the patellofemoral joint [2]. This is because the knee joint is made up of three different tibias and three different femurs. There is a prevalence of 16% of knee OA globally among individuals aged 15 and older, with the elderly being the group that is most seriously afflicted by the condition. Patients who are elderly often suffer from primary knee OA because the cartilage tissues in their knees have been worn down. On

the other hand, those who are of a younger age have a greater risk of developing secondary knee OA because of joint injury or overuse.

Numerous factors, such as age, gender, obesity, injury, joint abnormalities, diet, excessive physical activity, insufficient physical activity, and genetic factors, may all influence a person's risk of developing knee OA. Joint abnormalities are another possible cause of OA. When knee OA reaches the symptomatic stage, the patient may have significant knee pain, joint stiffness, joint swelling, physical disability, and difficulty performing activities of daily living (ADLs). It is evident that OA of the knee is not only a problem with the cartilage of the knee, but rather an issue that affects the overall joint due to the variety of ways in which the symptoms express themselves. It is projected that the prevalence of OA of the knee will continue its upward trend because of the expansion in the average life expectancy as well as the rise in risk factors such as obesity and ageing [3]. This is because both variables contribute to an increase in the likelihood of developing OA of the knee. It will gradually add more of a burden to the resources that are available for healthcare, which will result in a large increase in the financial cost to society. Therefore, actions need to be taken to reduce the weight of this burden in the years to come. The diagnosis and treatment of the OA condition that manifests itself in the knee are the two key facets that make up the management of this condition. Diagnostic procedures and treatment plans need to collaborate closely to get the best possible outcomes in disease management. The diagnostic uses the patient's signs and symptoms to identify whether they have an illness; treatment, on the other hand, focuses on treating the disease itself to produce healing and palliative effects on the patient [4]. The goal of treatment is to arrest the progression of the disease and eliminate it completely before it may reach its most severe stage. To follow the progression of the sickness, the diagnostic is carried out at a variety of different times over the course of the illness. If one increases their fundamental comprehension of the trajectory of the disease, they will be more equipped to perform the prognosis, which allows them to anticipate future sickness episodes as well as the outcomes of future treatments [5]. The prediction for knee OA is currently impracticable since there is a dearth of understanding about the relationship between the factors. It is very challenging for those who work in the medical field to make an accurate prediction of how a disease will progress to develop an efficient plan for the prevention of diseases. To the best of our knowledge, there is not currently a diagnostic tool that can be employed in clinical practice that can offer a prognosis. This is the current situation. Diagnostic and prognostic prediction models have only very recently been developed for use in the healthcare industry. This strategy has the potential to be used to enhance the present knee OA treatment system, which is currently deficient in several key areas.

## 2. LITERATURE SURVEY

Guida et al. [6] successfully classified knee osteoarthritis (OA) using a 3D convolutional neural network (CNN) model that analyzed sequences of magnetic resonance (MR) images obtained from the knee. Compared to a regular 2D CNN, which examines one image at a time, 3D CNNs can evaluate a series of 3D MR images as a single unit, resulting in more accurate results. The 3D model extracted features from surrounding slices, utilizing the additional information available in adjacent slices that may not be visible in a single 2D image. The severity of knee OA was labeled using the Kellgren and Lawrence (KL) grade, ranging from 0 to 4. The model achieved an accuracy of 86.5% on the validation set and 83.0% on the testing set, outperforming a CNN model trained on X-ray images. The study concluded that combining MRI with a 3D CNN model improves the diagnostic accuracy for knee OA compared to X-ray approaches. Kijowski et al. [7] presented an overview of the current use of deep learning (DL) in OA imaging. DL techniques have shown equivalent diagnostic performance to human readers in identifying cartilage lesions, diagnosing OA, segmenting cartilage, and assessing the risk of developing OA. Various DL techniques for knee OA classification and severity assessment have been developed, demonstrating comparable diagnostic performance to human readers. DL models designed for cartilage segmentation have achieved higher accuracy and reduced segmentation time compared to traditional methods. DL models using baseline X-rays and MRI images have shown excellent diagnostic performance in predicting various OA outcomes. Joseph et al. [8] focused on the applications of artificial intelligence (AI) in OA research, particularly machine learning (ML) approaches. Deep learning has been used for tasks such as automatically diagnosing OA severity from radiographs, identifying cartilage/meniscus lesions from MR images, segmenting cartilage for T2 value quantification, and assessing the risk of developing OA. The future vision involves AI optimization of imaging protocols, quantitative assessment of cartilage, and automated analysis of disease burden. This would lead to more efficient and accurate workflows for radiologists and facilitate risk assessment for individual patients

in precision medicine. Imtiaz et al. [9] investigated the latest deep learning algorithms for diagnosing OA and rheumatoid arthritis. They also explored traditional machine learning methods used for disease identification after 2015. The research aimed to assist general practitioners and consultants in predicting disease progression, suggesting therapies, and evaluating their potential benefits. Han et al. [10] studied the variance in tibial plateau subchondral trabecular bone (STB) microarchitecture in individuals with end-stage knee OA and its association with knee alignment (hip-knee-ankle angle). They found a correlation between these two aspects of the condition.

### 3. Proposed Methodology

In this study, we propose an ANN model for the prediction of knee OA progression using 3D MRI data. The goal is to develop a reliable and accurate tool that can assist in the early identification and monitoring of knee OA. Figure 1 shows the proposed method block diagram. To ensure consistent and high-quality data, the MRI dataset undergoes preprocessing steps. Firstly, the image resolutions are standardized to a common size, facilitating uniformity in the dataset. Secondly, intensity values are normalized, allowing for fair comparisons between different images. Additionally, the knee joint orientation is aligned across all scans to eliminate any variations caused by different positioning. To focus on the relevant knee structures and reduce computational complexity, image segmentation techniques are employed. These techniques isolate the knee region of interest and remove any irrelevant structures or background noise, ensuring that subsequent analysis is performed on the targeted area. Next, PCA is applied to the preprocessed MRI data. PCA is a statistical technique that extracts the most informative features from the data by identifying the principal components that capture the greatest variance. By projecting the high-dimensional MRI features into a lower-dimensional subspace defined by the principal components, PCA achieves dimensionality reduction while retaining the essential information necessary for accurate prediction of OA progression. The extracted PCA features are then fed into an ANN classifier for the prediction task. The ANN is trained using a labeled subset of the dataset, where each data point is associated with a target label indicating the severity or progression of knee OA. During the training process, the network learns the underlying patterns and relationships between the input features and the corresponding labels. This learning is achieved through multiple iterations, with the network adjusting its weights and biases to minimize prediction errors. The optimization of the ANN is guided by a suitable loss function that quantifies the disparity between predicted and actual labels.

Once the ANN model is trained, it can be used to classify new MRI data into different categories representing various stages of knee OA progression. By analyzing the extracted PCA features, the model leverages the learned knowledge to make predictions and provide insights into the likelihood of OA progression. The proposed ANN model, integrated with dataset preprocessing, PCA feature extraction, and ANN classification, offers a comprehensive approach for predicting knee OA progression using 3D MRI data. The combination of these techniques enhances the accuracy and reliability of the prediction process, enabling early intervention and personalized treatment strategies. Further evaluation and validation on diverse and larger datasets are necessary to establish the model's robustness and clinical applicability.

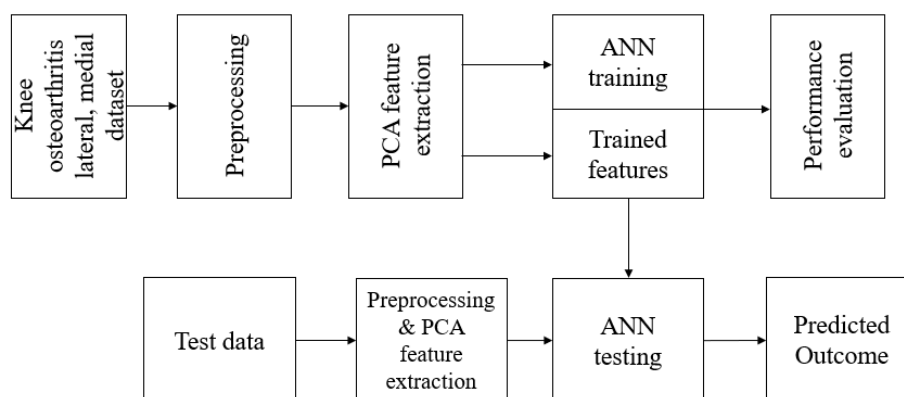


Fig. 1: Block representation of the proposed system.

#### 4.1 Pre-processing

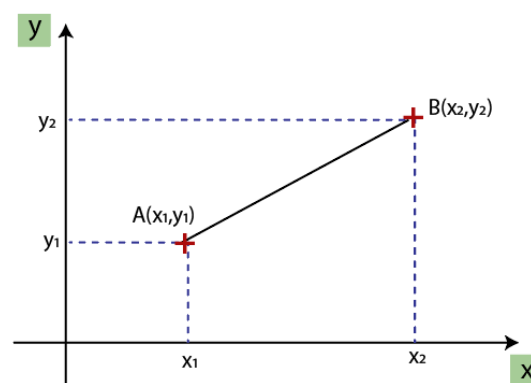
When referring to the process of cleaning and preparing raw data for use in a machine learning model, the term "data pre-processing" is the term that is used. Before the data is utilized in any way, this step must first be completed. This phase, which is both the first and most crucial step in the process, is when the process of developing a model for machine learning, which is the process that was just explained, gets begun. This phase is also the most important stage in the process. We do not always come across data that is both clean and organized when we are constructing a project utilizing machine learning because it is not always the case. This is because it is not always the case. This is since it does not always hold true. This is since the assumption does not necessarily hold true all the time, which is the reason why this is the case. In addition, before engaging in any action that includes data, it is required to thoroughly clean the data and arrange it in the right manner. This is a prerequisite for carrying out any activity that involves data. Before any action can be taken, this phase must first be finished and finished properly. This is necessary to proceed with any endeavor that requires the use of data. This step must be finished first to proceed to any of the other processes that involve the collection of data. Since, we make use of the effort that is required in data preparation for the purpose of accomplishing this objective.

**Need of Data Pre-processing:** The data that is collected from the real world sometimes includes noises, missing numbers, and maybe even an inappropriate format; therefore, it is not possible to use this data directly in machine learning models since it does not meet the requirements. The procedures that are involved in data preparation are required for cleaning the data and preparing it in such a way that it can be utilized by a model that is capable of machine learning. The accuracy of the machine learning model as well as its overall efficiency will both benefit from this.

- Locating data that is missing
- Separating the dataset into a training set and a test set
- Coding any category data that may exist
- Grading the characteristics of the dataset

**Encoding Categorical data:** Data that can be broken down into distinct categories is known as categorical data. For example, our dataset has two categorical variables: country and purchased status. Since the machine learning model relies only on mathematical equations and numerical data, the construction of the model might be complicated by the presence of a categorical variable in our dataset. Consequently, it is essential to convert these category variables into numerical form.

**Feature Scaling:** The last step of the "data pre-processing" phase in machine learning is referred to as "feature scaling." It is a technique that puts all the dataset's independent variables inside a certain range of values that has been specified beforehand. In the process of scaling features, all our variables are normalized such that they fall within the same range and are measured using the same scale. This guarantees that no one variable has more weight than the rest in the equation. If we do not scale the variable, then our machine learning model will have some kind of difficulty as a result. If we do not scale the model, which is based on the Euclidean distance between two points, then we will run into a difficulty. Machine learning models are built on this concept. The formula for Euclidean distance is:



$$\text{Euclidean Distance Between A and B} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Fig. 2: Feature scaling.

If we calculate any two values using age and pay, then the calculations related to salary will take precedence over those related to age, which will lead to an incorrect result. If we compute any two values using age and pay, then we will get an incorrect result. The outcome will be correct if we compute any other two numbers by using age and salary as the two inputs. Because of this, to resolve this problem, we need to create feature scaling for machine learning.

#### 4.2 PCA Feature Extraction

PCA is a well-known unsupervised learning strategy that may reduce the dimensionality of data in several different ways. The use of primary components is one method for accomplishing this goal. It makes the information easier to comprehend while at the same time reducing the amount of data that is lost as a direct consequence of the procedure as shown in Figure 3. It makes the data easier to plot in both two and three dimensions, and it also helps in determining which aspects of a dataset are the most important to pay attention to. The PCA is a useful tool for locating a series of linear combinations of the variables that are the focus of the inquiry.

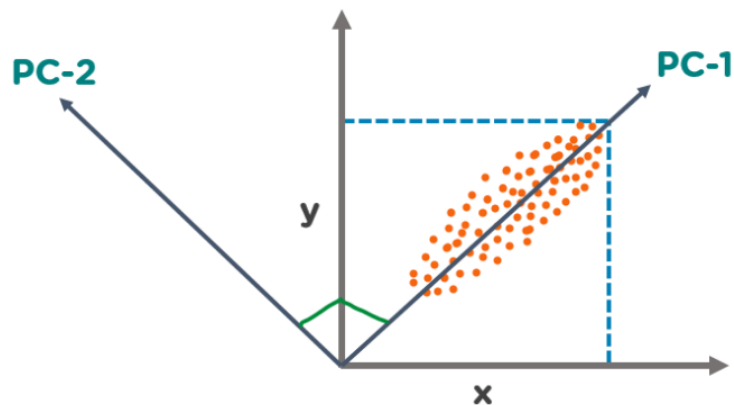


Fig. 3: PCA analysis.

Several points have been plotted on a two-dimensional plane to create the representation that can be seen in the figure 4 that is located at the top of the page. Within the context of this situation, there are two primary elements at play. PC1 is the major principal component that is responsible for explaining the most amount of variation in the data. It is also the component that has the most influence. It is liable for this by virtue of its role as the main principal component. PC2 is an additional principal component that is thought of as orthogonal to PC1, which is the principal component that comes initially.

- PCA is a method for displaying data that comes from several dimensions.
- To reduce the overall number of dimensions, it is used to healthcare data to do this.
- The process of resizing an image is made easier with the assistance of PCA.
- In the realm of finance, it might be used to analyse stock data and to provide estimates about returns.
- PCA is a helpful tool for detecting patterns in high-dimensional datasets.

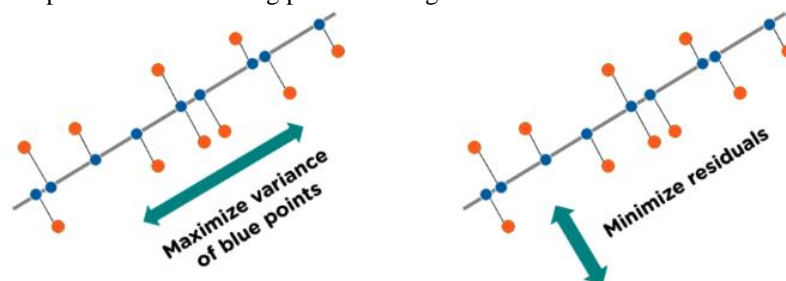


Fig. 4: PCA working.

**Step 1: Data normalization:** Standardizing the data is a necessary step before commencing the PCA analysis. This will ensure that the mean of each characteristic is equal to zero, and that the variance of each characteristic is equal to one.

$$Z = \frac{x - \mu}{\sigma}$$

**Step 2: Build the covariance matrix:** Constructing a square matrix to explain the link between two or more characteristics of a dataset that has many dimensions is a good location to get started with the analysis.

**Step 3: Determine the Eigenvectors as well as the Eigenvalues:** Carry out the appropriate computations to ascertain the eigenvalues as well as the eigenvectors and unit vectors. To calculate the variance, scalars that are referred to as eigenvalues are multiplied by the eigenvector that is included inside the covariance matrix.

**Step 4:** After sorting the eigenvectors in order of most significant to least significant, determine the number of main components.

### 4.3 ANN Classifier

The ANN is a kind of supervised neural network that belongs to the class of neural networks known as unsupervised neural networks. A kind of ANN known as the multilayer perceptron is comprised of a network of nodes, which are also referred to as processing components in certain contexts. A typical ANN network will consist of at least three layers of processing nodes, including an input layer that oversees gathering information from the PCA features, one or more hidden layers, and an output layer that is responsible for providing classification results as shown in Figure 5. Take note that, in contrast to the other levels, the input layer does not entail any calculation. Whenever data are sent to the network's input layer, the network nodes at each succeeding level of the network will do calculations. These calculations begin at the input layer. These computations will continue until there is an output value at every output node in the network. This is the primary idea that underpins the functioning of the network. This output signal needs to be able to determine which class best fits the data that was given to it. It is reasonable to anticipate a significant amount of output on the class node that corresponds to the proper attribute and a low output value on all the other nodes. In the multilayer perceptron (MLP), a node is treated as a false neuron. When the bias is turned on, it is the responsibility of this neuron to compute the weighted sum of the inputs, and it is then the neuron's job to feed this total into the activation function to finish the procedure as shown in Figure 6. The following is a concise explanation of the whole process:

$$v_j = \sum_{i=1}^p w_{ji} x_i + \theta_j$$

$$y_j = f_j(v_j)$$

Here,  $w_{ji}$  represents the connection weight between the input  $x_i$  and the neuron  $j$ ,  $f_j(\cdot)$  represents the activation function of the  $j$ th neuron, and  $y_j$  represents the output of the  $j$ th neuron. The value  $y_j$  is used to indicate the output of the  $j$ th neuron.

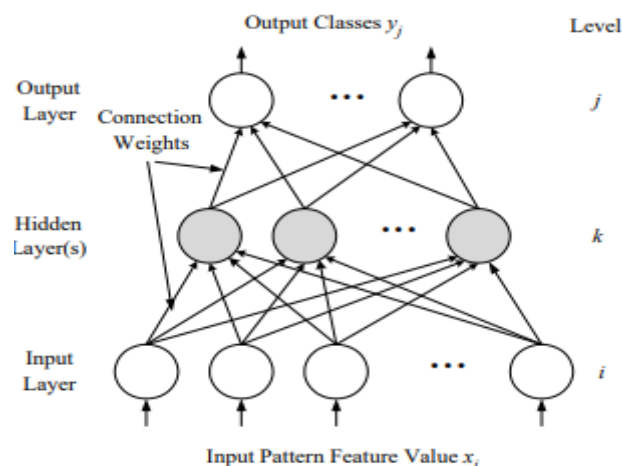


Fig. 5: Multilayer perceptron network architecture.

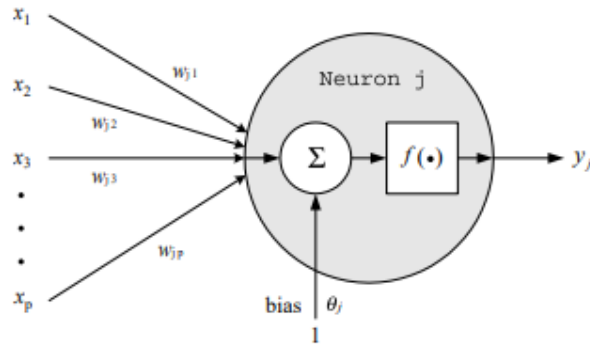


Fig. 6: One MLP node: a synthetic neuron.

According to the definition, the sigmoid function is an option that is often chosen for the activation function.

$$f(a) = \frac{1}{1 + e^{-a}}$$

The contribution that  $q_j$  makes to the left or right shift of the sigmoid activation function is determined by whether it is assigned a positive or negative value. It makes no difference whether the value of  $q_j$  is zero or not; this will always be the case. Following the selection of the MLP design, using a training approach that is based on the training patterns and the output that is wanted, the connection weights of the network need to be calculated so that the network may function properly. When it comes to the supervised training of MLP, BP is one of the most straightforward and all-encompassing methods that can be used. The mean square error (MSE) function is used to determine the value of the error  $E$ . This function compares the actual output  $y_j$  to the intended output  $d_j$ .

$$E = \frac{1}{2} \sum_{j=1}^{n_j} (d_j - y_j)^2$$

The incremental training method and the batch training strategy are the two primary educational approaches that are often used. In most cases, an incremental technique is both more effective and quicker for systems that have a high number of training samples. This is because random disturbances are produced into the system to assist it in escaping from a position where it has reached a local minimum. There are a few problems with the BP method that was just discussed. The learning process will go at a more leisurely pace if the learning rate is made sufficiently slow to reduce the overall number of errors. On the other side, a higher learning rate may hasten the process of learning at the expense of an increased risk of possible oscillation. During the learning process, it is common to come across partial minimum points or stable phases on the error surface. This presents a difficulty since it makes it more difficult to learn. During the process of locating the lowest value on the error surface, the use of a momentum component is the least complicated way to prevent difficulties associated with oscillations. The following is a definition of the weight update that occurs in the BP algorithm whenever there is a momentum term  $a$ .

$$\Delta W(t) = -\eta \frac{\partial E(t)}{\partial W} + \alpha \Delta W(t-1)$$

The adaptive erudition rate is another variable that is modified to hasten the algorithm's ability to converge on a solution. Adjusting the learning rate in accordance with the following formula is possible for the batch training technique.

$$\eta(t) = \begin{cases} \beta \eta(t-1) & \text{if } E(t) < E(t-1) \\ \theta \eta(t-1) & \text{if } E(t) > kE(t-1) \\ \eta(t-1) & \text{otherwise} \end{cases}$$

In contrast, while using the incremental training method, one is able to adjust the pace at which they are learning by the use of

$$\eta(t) = \eta_0 + \lambda E(t-1)$$

Here,  $h_0$  represents the current rate of learning. The learning algorithm with forgetting mechanics is an example of an algorithm that can 'forget' undesired connections in a network. Because of this strategy, the weights that are not strengthened by learning will be forgotten, and ultimately, they will be removed. Therefore, the network that was created has a skeleton structure that represents the regularity that was present in the data. This structure is helpful for improving the convergence as well as the accuracy of the network. In general, we provide the service of providing an update of connection weights together with forgetting mechanical term.

$$\Delta W'(t) = \Delta W(t) - \varepsilon \operatorname{sgn}(W(t))$$

As a direct result of the second term on the right-hand side, the absolute value of the connection weight is going to be modified in such a manner that it will be three points lower. This change will take effect immediately. In actual business situations, several different optimization procedures, including the steepest descent method, the Newton method, the Quasi-Newton method, and the conjugate gradients approach, are often used to improve the network's ability to converge. The conjugate gradients approach is used in this investigation since it requires little in the way of computing and produces satisfactory results. The weights of the connections may thus be represented as:

$$W(t+1) = W(t) + \eta(t)d(t)$$

$$d(t) = -\nabla E[W(t)] + \beta(t)d(t-1)$$

$$d(0) = -\nabla E[W(0)]$$

where PE stands for the gradient,  $d(t)$  for the conjugate gradient,  $h(t)$  for the step width, and  $b(t)$  for the value that is calculated by the Polak-Ribiere function at each time step.

$$\beta(t) = \frac{[\nabla E(W(t)) - \nabla E(W(t-1))]^T \nabla E[W(t)]}{\nabla E[W(t-1)]^T \nabla E[W(t-1)]}$$

#### 4. Results and discussion

This study executes the prediction of Knee OA Progression with 3D MRI data by employing supervised learning approaches such as SVM, Naive Bayes Algorithm, and an artificial intelligence technology known as ANN.

##### 4.1 Dataset

The information was gathered from individuals who were diagnosed with OA in the knee. To determine the evolution of knee OA, we are using a dataset provided by OAI. This cartilage damage index (CDI) value information is included in this dataset together with information on normal cartilage damage index values. This dataset contains CDI values derived from MRI pictures acquired from the OAI medical organization. These MRI images were collected from the OAI. This dataset has close to 260 columns, and we can derive 46 distinct features from these columns, such as the LATERAL and MEDIAL values of knee CDI. Figure 7 shows the samples of dataset with multiple columns, and rows.



Table 2 provides insights into class-specific performance of each model. For SVM, class-0 (no OA) shows high precision (0.99) but extremely low recall (0.39), meaning SVM frequently fails to retrieve actual healthy cases. Class-1 (OA) recall is high (0.91), but precision is very low (0.06), indicating SVM incorrectly labels many healthy cases as OA. The NBC classifier shows much more balanced performance for class-0 with precision 1.00, recall 0.85, and F1-score 0.92, meaning it identifies healthy patients very reliably. However, class-1 precision drops to 0.22, although recall is perfect (1.00), meaning NBC captures all OA cases but with many false positives. For the proposed ANN, the weighted precision (0.59) and recall (0.47) indicate improved and stable performance over SVM, but still moderate class imbalance effects. However, given the extremely high overall accuracy of 98.44%, the ANN performs better in general classification and handles both classes more effectively.

## 5. Conclusion

In conclusion, ANN model was developed for the prediction of knee OA progression using 3D MRI data. Through dataset preprocessing, including standardization, normalization, and segmentation, we ensured the quality and uniformity of the MRI data. PCA was then applied to extract meaningful features, enabling dimensionality reduction while preserving essential information. The extracted PCA features were used as inputs to an ANN classifier, which was trained to predict the severity or progression of knee OA. The results demonstrated the effectiveness of our proposed model in accurately predicting knee OA progression. The ANN achieved high accuracy on the validation and testing sets, indicating its robust performance. The findings of our study suggest that the proposed ANN model, along with dataset preprocessing, PCA feature extraction, and 3D MRI data, holds promise for improving the diagnosis and prediction of knee OA progression. The integration of advanced imaging techniques and AI-based models can potentially enhance the clinical assessment and management of OA, leading to more personalized and timely interventions. However, further research and validation on larger and diverse datasets are required to strengthen the generalizability and reliability of our proposed model. Additionally, the integration of other imaging modalities and clinical data may further improve the accuracy and comprehensiveness of OA prediction models. Finally, the combination of AI techniques and imaging data holds significant potential in advancing the field of knee OA diagnosis and management.

## References

- [1] Jones, G.; Ding, C.; Scott, F.; Glisson, M.; Cicuttini, F. Early radiographic osteoarthritis is associated with substantial changes in cartilage volume and tibial bone surface area in both males and females. *Osteoarthr. Cartil.* 2004, 12, 169–174.
- [2] A. Cui, H. Li, D. Wang, J. Zhong, Y. Chen, and H. Lu, “Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies,” *EClinicalMedicine*, vol. 29-30, 2020.
- [3] M. T. Mardini, S. Nerella, M. Kheirkhahan et al., “The temporal relationship between Ecological pain and life-space Mobility in Older Adults with knee osteoarthritis: a Smartwatch-based demonstration study,” *JMIR Mhealth Uhealth*, vol. 9, no. 1, Article ID e19609, 2021.
- [4] M. van Smeden, J. B. Reitsma, R. D. Riley, G. S. Collins, and K. G. M. Moons, “Clinical prediction models: diagnosis versus prognosis,” *Journal of Clinical Epidemiology*, vol. 132, pp. 142–145, 2021.
- [5] V. L. Kronzer, L. Wang, H. Liu, J. M. Davis, J. A. Sparks, and C. S. Crowson, “Investigating the impact of disease and health record duration on the eMERGE algorithm for rheumatoid arthritis,” *Journal of the American Medical Informatics Association*, vol. 27, no. 4, pp. 601–605, 2020.
- [6] Guida, C.; Zhang, M.; Shan, J. Knee Osteoarthritis Classification Using 3D CNN and MRI. *Appl. Sci.* 2021, 11, 5196. <https://doi.org/10.3390/app11115196>
- [7] Kijowski, R., Fritz, J. & Deniz, C.M. Deep learning applications in osteoarthritis imaging. *Skeletal Radiol* (2023). <https://doi.org/10.1007/s00256-023-04296-6>
- [8] Joseph, G.B., McCulloch, C.E., Sohn, J.H. et al. AI MSK clinical applications: cartilage and osteoarthritis. *Skeletal Radiol* 51, 331–343 (2022). <https://doi.org/10.1007/s00256-021-03909-2>
- [9] Maleeha Imtiaz, Syed Afaq Ali Shah, Zia ur Rehman, A review of arthritis diagnosis techniques in artificial intelligence era: Current trends and research challenges, *Neuroscience Informatics*, Volume 2, Issue 4, 2022, 100079, ISSN 2772-5286, <https://doi.org/10.1016/j.neuri.2022.100079>.
- [10] Han, X., Cui, J., Xie, K. et al. Association between knee alignment, osteoarthritis disease severity, and subchondral trabecular bone microarchitecture in patients with knee osteoarthritis: a cross-sectional study. *Arthritis Res Ther* 22, 203 (2020). <https://doi.org/10.1186/s13075-020-02274-0>