Journal of Advanced Zoology

ISSN: 0253-7214
*Volume **44** Issue **S-6 Year 2023** Page **634:642***

Determining Efficient Machine Learning Techniques for Grading of Knee Osteoarthritis

Bhavana V*, NallapaReddy Pooja, Shalini V R, Sarvamangala D R

Department of Computer Science & Engineering REVA University Bangalore, India
R17CS244@cit.reva.edu.in
R17CS257@cit.reva.edu.in
R17CS299@cit.reva.edu.in
sarvamangala.dr@reva.edu.in

**Corresponding author's E-mail: R17CS244@cit.reva.edu.in*

Article History

Received: 06 June 2023
Revised: 05 Sept 2023
Accepted: 29 Nov 2023

Abstract

Osteoarthritis (OA) of the Knee is a degenerative joint disease mainly caused due to loss of articular cartilages. The paper introduces an approach to quantify knee osteoarthritis (OA) severity using KL grades. This approach combines EDA (Exploratory Data Analysis), Pre-processing and Feature Engineering techniques. The amount of damage to the knee can be graded using KL scale (0-4). The automated detection of Knee Osteoarthritis (KOA) based on KL grades which corresponds to severity stages has been given in the paper. In the study public dataset from Osteoarthritis Initiative (OAI) has been used to evaluate the proposed approach with very promising results. Different accuracy metrices like F1 score, Receiver operating characteristic curve (ROC), Area Under Curve (AUC) and Precision were used to find the best algorithm amongst the classification models in Machine learning. Random forest and Decision trees algorithms were considered efficient giving an accuracy of 96.9% and 91.6% respectively. Our study is an economically better approach when compared to x-rays for OA detection.

Keywords: *Joint space narrowing (JSN), Kellgren-Lawrence grades, Knee Osteoarthritis, Machine learning*

1. Introduction

Knee bears the entire body weight over its synovial hinge joint which has two condylar joints and one saddle joint. With an increasing age there is a wear and tear in these joints over time, and a gradual reduction is Joint space width (JSW).

Osteoarthritis (OA) can be better understood as a degenerative disease whose major result can be knee JSN. It is the most generic chronic disease which has lot of impact on the public health.

KOA is well said as the loss of the knee's articular cartilage which is the main reason to protect the bones from harsh wear and tear and joint frictions. This mainly causes the damage of bones underneath the cartilages which also affects the soft tissues surrounding it.

This is the most rheumatologic problem and also the most frequent joint disease with a prevalence of 10% in male and 13% in female in US. It is mostly seen in people aged above 60years and specially in women than men, but the prevalence increases drastically with age. It causes pain, mobility limitation in people who are affected by it. Currently, no known cure is discovered for KOA, but many biological, medical and environmental risk factors are involved for the progression and development of KOA. So certain precautionary measures have to be taken to prevent this.

Machine learning techniques such as decision trees, k-nearest neighbour, and advanced classification models can be used to predict the stages of KOA. These methods with combined features of selection methods can be used for very large data sets which can significantly contribute in predicting the risk factors of KOA.

For Knee OA severity the Kellgren-Lawrence (KL) grading system is considered to be the gold standard. The joint space narrowing (JSN) and osteophytes formation becomes the basis for knowing

the severity of KOA using KL grades. The newest and a feature specific way for grading of OA is the (OARSI) ATLAS. This allows gradings for major features which directly or indirectly contribute to the final gradings of knee osteoarthritis. Few to mention: femoral osteophytes (FO), the tibial-osteophytes (TO) and the Joint Space Narrowing (JSN) (Refer to Figure 1).

Figure 1: Two grades (KL-0 and KL-3) of KOA.

In the Figure (1) FL, FM, TL and TM represents the femoral lateral, femoral medial, TL and TM compartments respectively. (a)A knee which is not affected by OA i.e, KL grade 0 (b) A knee affected by OA with its severity as grade 3 according to Kellgren -Lawrence. The arrow in (b) indicates very less Joint space narrowing (JSN) which is the reason for it to fall under KL grade 3.

Let us see the different grades of KOA and their observable usual symptoms:

- **Grade 0**: With none of the symptoms. Complete absence of osteoarthritis
- **Grade 1**: Tentative stage of KOA. This includes partial joint space narrowing (JSN) and there is a possibility of tiny osteophytes.
- **Grade 2**: Very minimal symptoms with definite osteophytes and possibility of joint space narrowing (JSN).
- **Grade 3**: A moderate level of disease infection with multiple osteophytes. It also includes certain JSN with some sclerosis. Bone ends deformity can be the later effects of this stage.
- **Grade 4**: This is the final and sever stage of KOA. It can be seen as large osteophytes which marks large JSN, severe stage of sclerosis and confirmed deformity of bones. (Refer Figure 2).

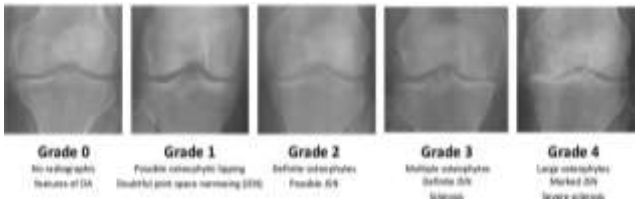

Figure 2: Kellgren-Lawrence grading System

In the study, we have compared the KL grades for knee joint where 0-4 grades define the stages of osteoarthritis. Here we focus on using the KL grades of the samples on different Machine learning algorithms to find for the best one in correctly detecting the severity of OA. We have calculated the initial accuracy, F1 score, precision and recall scores. Receiver operating characteristic curve (ROC) is plotted for TPR and FPR which efficiently shows the performance of different classification models at all thresholds.

The remaining part of the paper is organized as follows. Section II, reviews the literature survey conducted. Section III is the problem definition. Section IV describes the materials and methods used in this research, including OAI database, pre-processing, feature analysis, machine learning methods and evaluation metrics. In Section V describes the methodology used. Section VI is a detail showcase of the analysed results from the experiments. In Section VII gives the conclusion of the work done.

I. LITERATURE SURVEY

In [1] the authors have predicted about the progression of knee osteoarthritis on Magnetic Resonance Images (MRI). To predict Machine learning methods like those of ANN, Random Forest, SVM etc. The above-mentioned methods predict the progression by measuring KL grade, JSM grade, JSL grade.

In [2] authors have done evaluation on effects of laser photo biomodulation as a therapy to reduce the pain and providing better life for patients without suffering because of pains. The authors concluded

PBM is effective by analysing the score of Visual Analog Scale (VAS), Range of Joining Motion (ROM), WOMAC. Calculation is done based on mobility problem, stiffness and difficulty in performing daily activities.

In [3] author obtained knee sound signal by using inventive stethoscope device named as Goniometer. Computer takes the signal as input and process the signal based on obtained data. Then, transformation of signal will be done. Partition index is calculated by obtaining the power spectrum. Vibroarthrographic signals (VAG) are compared with x-ray images, then author proved that VAG signals of knee are best alternative for x-ray examinations.

This [4] introduced a better approach for knee OS by using Radiographic images. Authors have combined pre-processing to this approach CNN and LSTM are used as extraction method and classification method. Osteoarthritis Initiative (OAI) is the data set used for calculation. This approach gives the best results compared to before methods.

This [5] is about analysing the severity of OA by detecting radiographic. In this paper, they compared all the approaches that are used to analyse the knee OA and mentioned which approach gives the accurate results. This paper concentrated on the important issues related to this feature extraction in Osteoarthritis and main methods used by different researchers.

In [6], the development is done on automatic method for Kellegren-Lawrence and OARSI grade predictions using Knee Radio graphs. This method depends on deep learning and fifty layers of networks. They used Transfer learning from imageNet and OAI dataset. This detects the radiographic OA presence, more accurate than the current methods.

This [7] focused on establishing references of Joint Space Width (JSW) which is used in diagnosis and leads to knee cartilage diseases. This was done on 30 degrees knee computed radiographs. The in-built electronic callipers were used to measure the JSW of each knee.

This [8] is about developing MUS procedures and detecting knee OA. MUS scores is compared with plain radiographs and osteoarthritis outcome score (KOOS). They proved MUS score in this study was dependable and is valid in detection knee Osteoarthritis than standard radiographs of the knees.

II. PROBLEM DEFINITION

Determining efficient machine learning techniques for grading of knee osteoarthritis.

2. Materials And Methods

A. DATASET

The complete data and all the CSV files used in our study were selected from the US database which is Osteoarthritis Initiative (OAI). It is a prospective, multicentre and longitudinal observational study of knee osteoarthritis. From the baseline cohort, there were approximately 9000 participants. We selected a good sample of approximately 2600 participants who had complete data (cyst grades, sclerosis grades, osteophytes grades, JSN and KL grades). The datasets can be publicly available for viewing on the OAI website.

TABLE 1. DATASET SHOWING THE MAJOR FEATURES AMONGST ALL 19 FEATURES

Grades	Kellgren and Lawrence (grades 0- 4)	osteophytes (OARSI grades 0- 3) tibia lateral compartment	osteophytes (OARSI grades 0- 3) tibia medial compartment	osteophytes (OARSI grades 0- 3) femur medial compartment	osteophytes (OARSI grades 0- 3) femur lateral compartment
0	321	1524	895	1278	1492
1	346	671	1167	552	600
2	1067	186	291	291	239
3	624	203	231	463	253
4	226				

i. PREPROCESSING OF DATASET

As said earlier, OAI is a multicentre study. Thus, the datasets collected from the baseline cohort where conflicting. Due to this, pre-processing becomes the major step before classification. First, many CSV files which were chosen for the study which were merged together to form a very large dataset. Later, we looked into the basic information which was put forth by the data. An EDA (Exploratory data analysis) was made to decide about the relevant features whose corelations had a big impact on the KL grades. Further, we pre-processed the data by removing the outliers. Label encoders were used for conversions of labels into numeric forms which would be easier for the machine learning algorithms to process. After the final formation of the accurate dataset, we were left with 2584 samples who would be classified for KL grades based on 19 different features.

ii. FEATURE ANALYSIS

Corelations were calculated for each feature corresponding to the target class. The positive corelations were considered eliminating the negative corelations. 19 sets of informative features were selected for which the corresponding target label was the KL grades. These set of 19 features were further classified as two groups. 9 features were from medial compartment (which includes medial femur and medial tibia) and 9 features were from lateral compartment (which again includes lateral femur and lateral tibia). As per many researches, medial OA is more common than lateral OA. That is why we have analysed all the features in these sub groups for classification.

iii. MACHINE LEARNING METHODS USED

In our work, we have explored and used five machine learning methods to learn which works the best for classifying KL grades based on different metrics. The five machine learning methods are Random Forest, Decision trees, Naïve Bayes, KNN and Logistic Regression.

Random forest are powerful classifiers which can handle large data as it can balance the datasets efficiently when a class is more infrequent than other classes the data. It is a combination of multiple decision trees which works in the training phase of dataset. For the final result, it considers all the individual tree's output and combines them to form a single output. Hence, we find this more accurate.

Decision tree does not require much computation time and is very intuitive. It forces to consider all possible outcomes of a decision and also traces every path to an end conclusion. It is easy to explain with simple math and does not involve any complex formulas.

Naive Bayes is seen as a group of simple probabilistic classifiers which will basically be working with Bayes theorem. Naïve independence assumptions between all the features are considered in this method. It is largely used for an automatic medical diagnosis. When we compare other classification fields, Naïve Bayes is not as competitive as other advanced machine learning methods such as Logistic Regression and K-nearest Neighbors Algorithm, but in our work, we found that it has achieved a better performance as compared to the latter ML algorithms.

Logistic Regression and KNN algorithms were used to get a comparative study of how our dataset works when applied on different feature engineering models.

iv. EVALUATION

In this work we have used many different metrics to show up the performance of our all five classifiers:

Accuracy, Precision, Recall (also called as sensitivity measure), F1 score (F score), and the ROC curve. ROC curves show the relation between classification sensitivity and specificity. It is based on the classifier confidence threshold, on how much it increases or decreases. F-measure clearly indicates the complete classification accuracy in terms of weighted average of precision and recall. It is calculated for a particular confidence threshold. Evaluation metrics formulas are given here

Precision =

True Positive / (TP + FP)

Recall = TP / (TP + FN)

F1 Score = 2*(Recall * Precision) / (Recall + Precision)

TP= True Positive, FP= False Positive, FN=False Negative

Confusion matrix also called as an error matrix is a table layout, allows for easy visualization of the performance of various different algorithms. It suits best for supervised learning which we have used in this work to get a better understanding.

III. METHODOLOGY

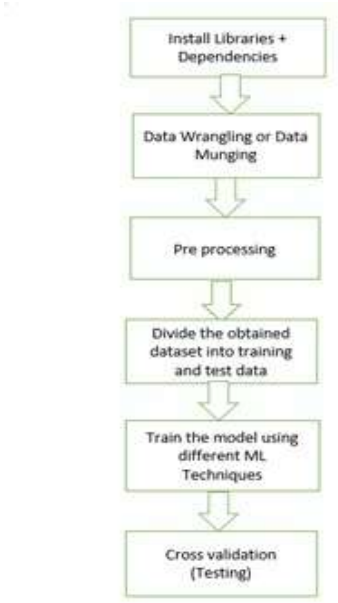

Figure 3: Flowchart of proposed model

Initially, we install libraries required. Further, data wrangling is done to combine data from various files.We make use of multiple CSV files for the collection of datasets. Preprocessing becomes necessary to remove the outliers and to retain the necessary features. The dataset is trained on the training data-set and tested on test-set. Later the model is trained using various Machine Learning classsification algorithms and the obtained results are analyzed. For various performance metrics including accuracy, ROC, AUC and F-measure.

3. Results and Discussion

EXPERIMENTS DONE FOR PREDICTING KL GRADES

i. Predicting Kellgren-Lawrence grade accuracy for Random Forest algorithm using ROC through OVR (one versus rest) scheme - For Random Forest, the best performance was achieved with AUC of value 0.969 and F-measure of value 0.89.

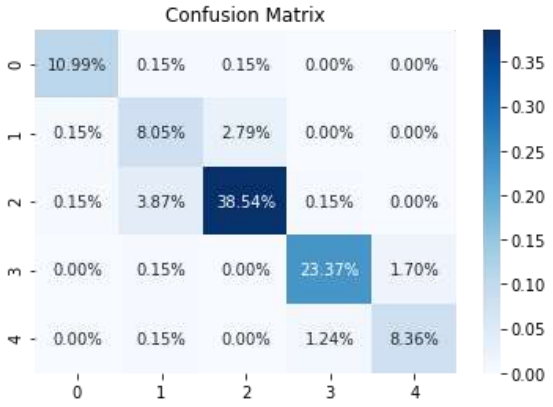

Figure 4: Confusion Matrix (Error Matrix) for Random Forest

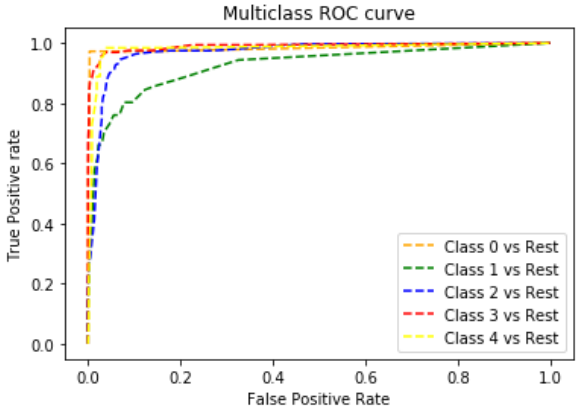

Figure 5: ROC curve for Random Forest

ii. Predicting KL grade accuracy for Decision tree algorithm using ROC through OVR (one versus rest) scheme - For Decision Tree, the best results was seen with AUC of 0.916 followed by F-measure score of 0.88.

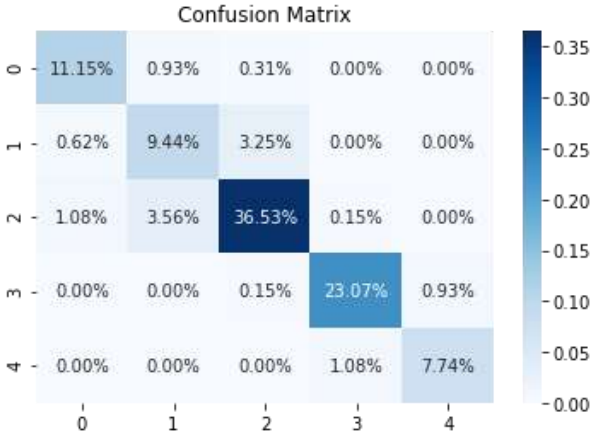

Figure 6: Confusion matrix for Decision Trees

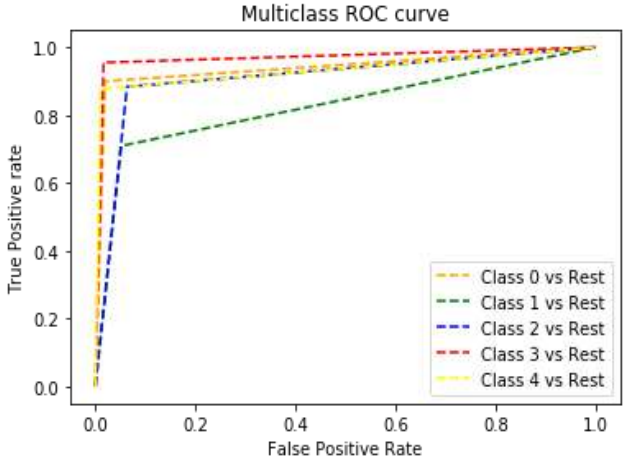

Figure 7: ROC curve for Decision Trees

iii. Predicting KL grade accuracy for Naïve bayes algorithm using ROC through OVR (one versus rest) scheme - For Naïve Bayes, the best results was seen with AUC of 0.622 followed by F-measure score of 0.43.

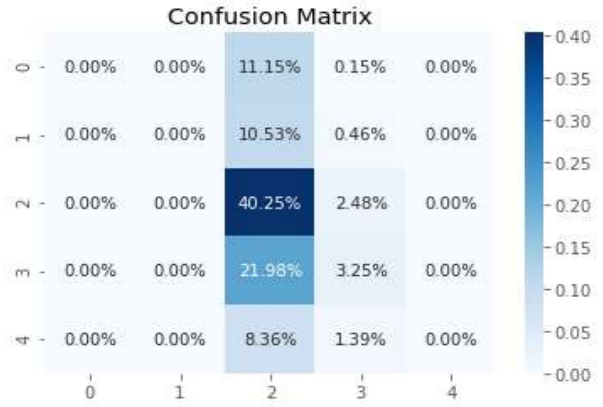

Figure 8: Confusion matrix for Naïve Bayes classifier

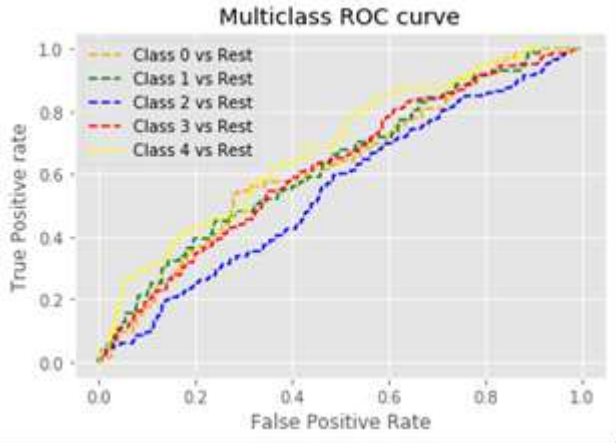

Figure 9: ROC curve for Naïve Bayes classifier

iv. Predicting KL grade accuracy for Logistic Regression algorithm using ROC through OVR (one versus rest) scheme - For Logistic Regression, the best results was seen with AUC of 0.44 followed by F-measure score of 0.44.

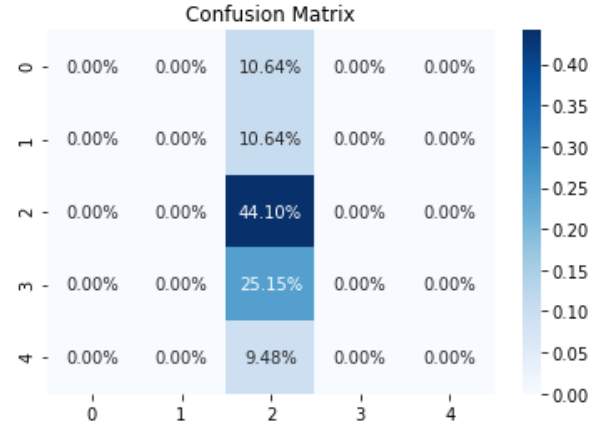

Figure 10: Confusion matrix for Logistic Regression

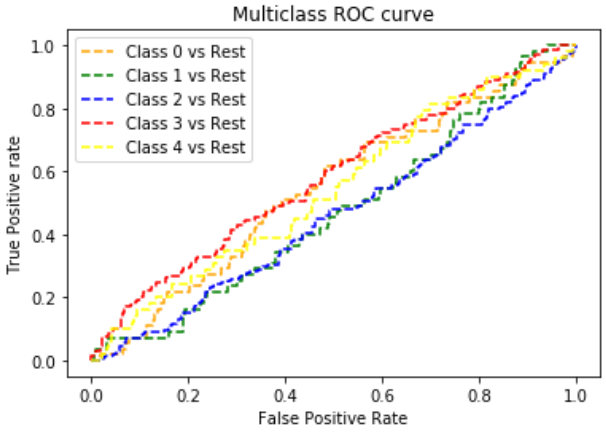

Figure 11: ROC curve for Logistic Regression

v. Predicting KL grade accuracy for KNN algorithm using ROC through OVR (one versus rest) scheme - For KNN, the best results was seen with AUC of 0.502 followed by F-measure score of 0.49.

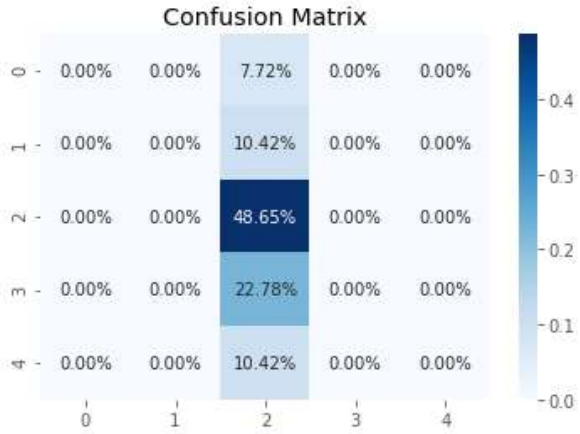

Figure 12: Confusion matrix for KNN

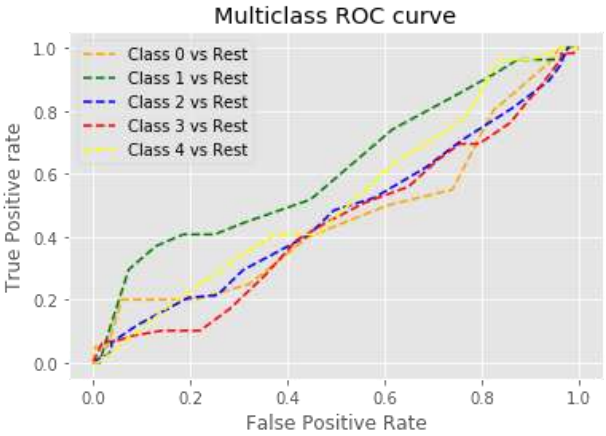

Figure 13: ROC curve for KNN

A. ACCURACY OBTAINED

TABLE 2. DIFFERENT METRICS DEPICTION FOR EACH OF THE FIVE CLASSIFIERS

Classifier	Precision	Recall	F measure	ROC area
Random Forest	**0.876**	**0.889**	**0.877**	**0.969**
Decision Trees	0.866	0.879	0.866	0.916
Naïve Bayes	0.189	0.212	0.434	0.622
Logistic Regression	0.109	0.202	0.444	0.441

KNN	0.101	0.208	0.497	0.502

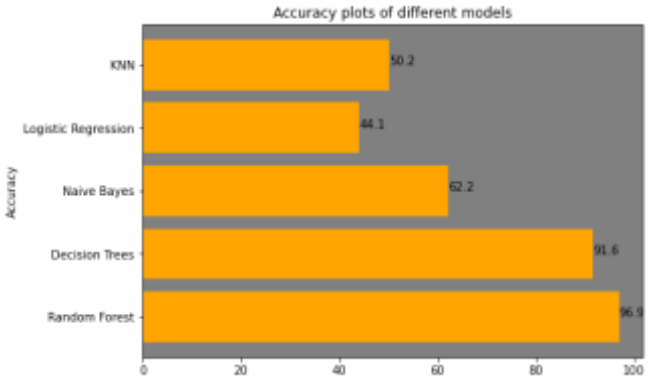

Figure 14: Accuracy plots of the various machine learning models analyzed.

4. Conclusion

The study was conducted for the comparative study of various machine learning classification algorithms for analysing a grading system for KOA. We in our work have proposed an approach to classify severity of knee OA using CSV files based on KL grading system. We have used different machine learning models to find the efficiency of each in correctly classifying the KL grades. Classification based on Random Forest provides good results when compared to other models.

References:

[1] Yaodong Du ; Rania Almajalid ; Juan Shan ; Ming Zhang, "A Novel Method to Predict Knee Osteoarthritis Progression on MRI Using Machine Learning Methods," Volume: 17 , Issue: 3 , July 2018

[2] Ana Paula Pelosi Mota ; Salmo P. Silva ; Alessandra Batista ; Amanda F.F. Barros , "Analysis of the Treatment of Knee Osteoarthritis using Photobiomodulation Performed with a Low Power Laser," 2019 SBFoton International Optics and Photonics Conference (SBFoton IOPC).

[3] Tsair-Fwu Lee ; Wei-Chun Lin ; Li-Fu Wu ; Hung-Yu Wang "Analysis of Vibroarthrographic Signals for Knee Osteoarthritis Diagnosis," 2012 Sixth International Conference on Genetic and Evolutionary Computing

[4] Rima Tri Wahyuningrum ; Lilik Anifah "A New Approach to Classify Knee Osteoarthritis Severity from Radiographic Images based on CNN-LSTM Method" 2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST)

[5] Pooja P. Kawathekar ; Kailash J. Karande "Severity analysis of Osteoarthritis of knee joint from X-ray images" 2014 International Conference on Signal Propagation and Computer Technology (ICSPCT)

[6] Aleksei Tiulpin; Simo Saarakkala;" Automatic Grading of Individual Knee Osteoarthritis Features in Plain Radiographs Using Deep Convolutional Neural Networks" @copyright 2020 by the authors. Licensee MDPI, Basel, Switzerland.

[7] Ismail Anas ; Tabari Abdulkdir Musa ; Isyaku Kabiru ; Abdulkadir Adekunle Yisau ; Idris Sulaiman Kazaure ; Suwaid Muhammad Abba; Saleh Muhammad Kabir;" Digital radiographic measurement of normal knee joint space in adults at Kano, Nigeria" Depatrment of Radiology, Nigeria; Published on 15 April 2013.

[8] B.F. Riecke; R. Christensen; S. Torp-Pedersen; "An Ultrasound score for knee osteoarthritis: a cross-sectional validation study" Department of Radiology, Elsevier, Published on 30 September 2014.