JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
SICAIET
2024

Research Article
Collection: Artificial Intelligence and Emerging Technologies

A Diagnostic Tool for Rheumatoid Arthritis and Disease Activity Levels using Machine Learning Classifiers

Javeria Batool[1], Ayesha Hakim[1*], and Javeria Jabeen[1]

[1]Institute of Computing, Muhammad Nawaz Shareef University of Agriculture (MNSUA), Multan, Pakistan
[*]Corresponding Author: Ayesha Hakim. Email: ayehsa.hakim@mnsuam.edu.pk

Academic Editor: Salman Qadri **Published:** February 01, 2024

Abstract: Rheumatoid Arthritis (RA) stands out as a severe chronic inflammatory disease potentially leading to disability. It mostly affects the joints in body including small joints of hands and feet such as wrists, finders, and toes and large joints such as knees and shoulders. The symptoms resembles much with other inflammatory arthritis diseases that is a major challenge in its early diagnosis. This paper presents an automated tool for diagnosis of RA and its disease activity levels using machine learning classifiers based on signs and symptoms that RA patients noticed and recorded. The dataset comprises of anonymised clinical data, lab reports, and demographics information of 104 RA patients collected from a local hospital in Multan. This study compares the performance of five supervised machine learning classifers including AdaBoost, Support Vector Machines, Random Forest, Naïve Bayes, and Decision Trees for diagnosis of RA and its activity level. The algorithms selected the clinicaly robust features using the parameter of information gain for classificatio of RA and non-RA patients. Disease activity levels are further categorised using DAS-28 score in terms of severity. The goal of this study is to determine the best suited machine learning algorithm for RA diagnosis that can be used to assist physician in early diagnosis. The performance of AdaBoost classifier was better than other algorithms in terms of precision, recall, F1-score, error rate, and specificity for diagnosis of RA and its various activity levels.

Keywords: Disease activity level; Information Gain; Diagnostic tool; Rheumatoid Arthritis; Classification; DAS-28.

1. Introduction

Rheumatoid Arthritis (RA) stands out as the prevalent chronic inflammatory autoimmune joint disorder, impacting approximately 1-2% of the worldwide population [1]. The condition is marked by inflammation of the joints, leading to swelling and stiffness due to immune-triggered synovial inflammation. This, in turn, causes joint damage, progressive polyarthritis, and persistent disability. While the disorder primarily impacts the joints, it is recognized as a systemic ailment with additional symptoms beyond the joints, including rheumatoid nodules, lung complications, vasculitis, and associated health issues. Consequently, the overall life expectancy of individuals with this condition is diminished in the general population.

Medical practitioners suggest early treatment of the diseases based on an accurate assessment of disease activity level. The stage of RA illness can be determined using a variety of methods, including laboratory reports that include the acute phase response, which is a direct reflection of the core inflammation; and patient-based variables like pain assessment, the number of swollen or tender joints, and an assessment of the overall disease activity [3]. However, most of these symptoms overlap with other types of arthritis. The American Rheumatism Association (ACR) introduced classification guidelines in 1987 (listed in Table 1) to distinguish patients with established RA from those with other inflammatory arthritis[4].

Table 1. Classification criteria developed by the American College of Rheumatology (ACR) are used to distinguish between patients with and without RA [4].

Criterion **The criteria 1-4 must have been existed for a minimum of six weeks.**	Definition
Morning Stiffness	Morning stiffness affecting the joints persists for a minimum of one hour before reaching its maximum relief.
Arthritis of ≥3 joint areas	A physician has clinically witnessed synovitis in at least three joint regions at the same time.
Arthritis of hand joints	At least one part of the wrist, metacarpophalangeal (MCP), or proximal interphalangeal (PIP) joint exhibits swelling.
Symmetric arthritis	The same joint regions are simultaneously affected on both sides of the body.
Rheumatoid nodules	Subcutaneous nodules on extensor surfaces, located on bony prominences, or juxta-articular areas are present.
Serum RF	Rheumatoid factor (RF) test result is positive.
Radiographic changes	Anterior hand and wrist radiographs show radiographic alterations suggestive with RA.
Satisfied 4/7 criteria = RA	

ACR evaluates the disease's initial symptoms in response to therapy; it does not take the patient's degree of disease activity into account. 113 RA patients' samples were used by a group of Dutch researchers in 1990 to establish and validate the Disease Activity Score (DAS-28) [5]. A numerical score called the Disease Activity Score in 28 joints (DAS28) is used to quantify the degree of rheumatoid arthritis disease activity. The tender joint count (TJC), swollen joint count (SJC), erythrocyte sedimentation rate (ESR), and the patient's overall health—which is usually assessed using a visual analogue scale (VAS)—are all considered when calculating the DAS28 calculation. The following is the formula:

$DAS28 = 0.56 \times \sqrt{TJC28} \times \sqrt{(SJC28)} + 0.36 \times \ln(ESR) + 0.014 \times GH$

TJC28 = 28 specific joints' tender joint count

SJC28= The swelling of 28 specific joints

ESR stands for Erythrocyte Sedimentation Rate

GH stands for general health.

ln= Natural log

The score of the DAS28 formula given in Eq. 1 ranges from 0 to 10 and Rheumatoid Arthritis patients are evaluated for disease activity state as listed in Table 2.

Table 2. Disease Activity Level determined by DAS-28 score. This table is extracted from [6].

Disease Activity Score (DAS)	Disease Activity Level
DAS<2.6	REMISSION
DAS28<=3.2	LOW
3.2<DAS28<5.1	MODERATE
DAS28>5.1	HIGH

The identification of diseases and the grading of disease activity levels have shown to be greatly aided by artificial intelligence and image processing techniques. In a recent study [7], researchers classified patients into RA and non-RA patients using an artificial neural network that was trained on data containing six features: age, sex, RF, anti-CCP, 14-3-3η protein, and anti-carbamylated protein (CarP) antibodies. The proposed model achieved 90.6% classification accuracy on a dataset of 670 participants from China. Another study [8] presented an unsupervised model to detect inflammatory changes caused by RA on a dataset of 99 images of wrist MRIs. They obtained a pixel-level nontrivial changes to represent inflammatory

progression. The survey presented in [9] provides summary of work done for diagnosis of RA using computer vision and Artificial Intelligence. The majority of the study [10–14] focuses on diagnosing RA by utilizing imaging data—such as X-rays, MRIs, and CT scans—to identify minute alterations in the joints. The survey also presented the studies that are related to diagnosis of RA based on blood tests and clinical data such as [15-16]. However, most of the studies did not classify RA into disease activity levels. Disease activity level has been scored in [17] on 1342 Doppler ultrasound images of arthritis patients, but it has not been evaluated on RA patients.

2. Materials and Methods

The proposed structure of this study consists of six phases as shown in Fig. 1. The description of each phase is presented as follows:

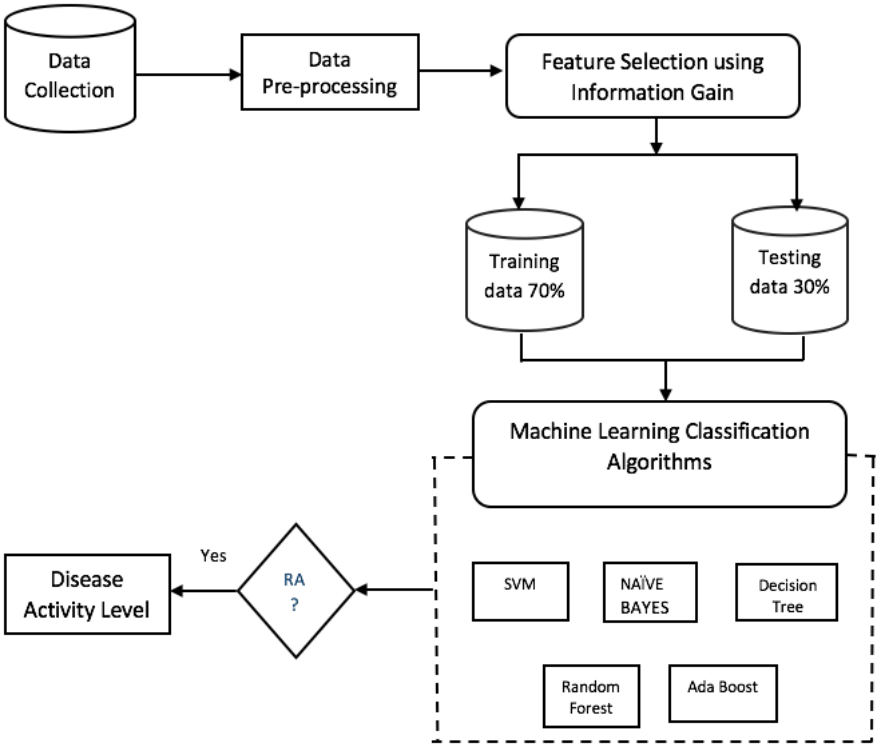

Figure 1. The proposed methodology comprising of six phases: data collection, data pre-processing, feature selection using information gain, percentage split of dataset, machine learning classification, and disease activity level.

3. Data Collection

There are over 100 distinct forms of arthritis [18] each with its own set of symptoms and indications, but body soreness, stiffness, inflammation, and painful joints are the common complaints among all patients. For an accurate diagnosis, it is critical to differentiate between RA from other types of arthritis. To collect an accurate dataset, the authors obtained diagnostic criteria for RA patients from a panel of expert rheumatologists at a private hospital in Multan, Pakistan. Authors found ACR and DAS28 criteria to detect disease activity levels from previous studies. The most accurate diagnostic criterion was selected after discussion with rheumatologists and a questionnaire was designed to collect data from patients.

Table 3. Description of features in the dataset

No.	Features	Value type
Demographic Features		
1	Age	integer
2	Gender	M, F, Not Disclosed
3	Height in inches	integer
4	Weight in kg	integer

5	BMI (Body mass index)	integer
6	Family history(disease)	Yes, No
7	Smoking	Yes, No, Ex-smoker
Symptoms		
8	Morning stiffness	Yes, No
9	Activity reduces pain	Yes, No
10	Low-Grade Fever	Yes, No
11	Swelling in Joints	Yes, No
12	Tenderness in Joints	Yes, No
13	Symmetry	Yes, No
14	Pain in Wrist	Yes (Low, Medium, High), No
15	Pain in Shoulder	Yes (Low, Medium, High), No
16	Pain in Knees	Yes (Low, Medium, High), No
17	Pain in Spine	Yes (Low, Medium, High), No
18	Pain in Ankles	Yes (Low, Medium, High), No
19	Pain in Hip	Yes (Low, Medium, High), No
20	Pain in Neck	Yes (Low, Medium, High), No
21	No. of MTP joints in pain	integer
22	No. of MCP joints in pain	integer
23	No. of PIP joints in pain	integer
24	Swollen Joint Count (SJC)	(0-28)
25	Tender Joint count (TJC)	(0-28)
26	Pain Level	Low, Medium, High
27	G. Health	(0-10)
28	Loss of ROM (Range of Motion)	Yes, No
29	Disease On set time	integer
30	RA Nodules	Yes, No
31	Fatigue	Yes, No
Lab findings		
32	Rheumatoid Factor	integer
33	Anti-CCP	integer
34	ESR	integer
35	CRP	integer
36	HB level	integer
37	Platelets count	integer
38	WBC	integer
39	X-ray	Categorical

40	Prognosis	Not patient, Low level, Medium, High

The authors collected data from 104 patients using stratified random sampling. The group of strata included all patients who visited the rheumatologist in a private hospital in Multan for six months (10, January 2021 to 10, July 2021). We did not record names, addresses, contact details, or any other identifiers to ensure the privacy of patients. All data was recorded after signing a consent form from the patient. The collected dataset is comprised of demographic, clinical symptoms, and lab findings of patients. A total of forty parameters were recorded for each patient as shown in Fig. 2. The description of the value type of each feature is listed in Table 3 and the range of selected features in the entire dataset is given in Table 4.

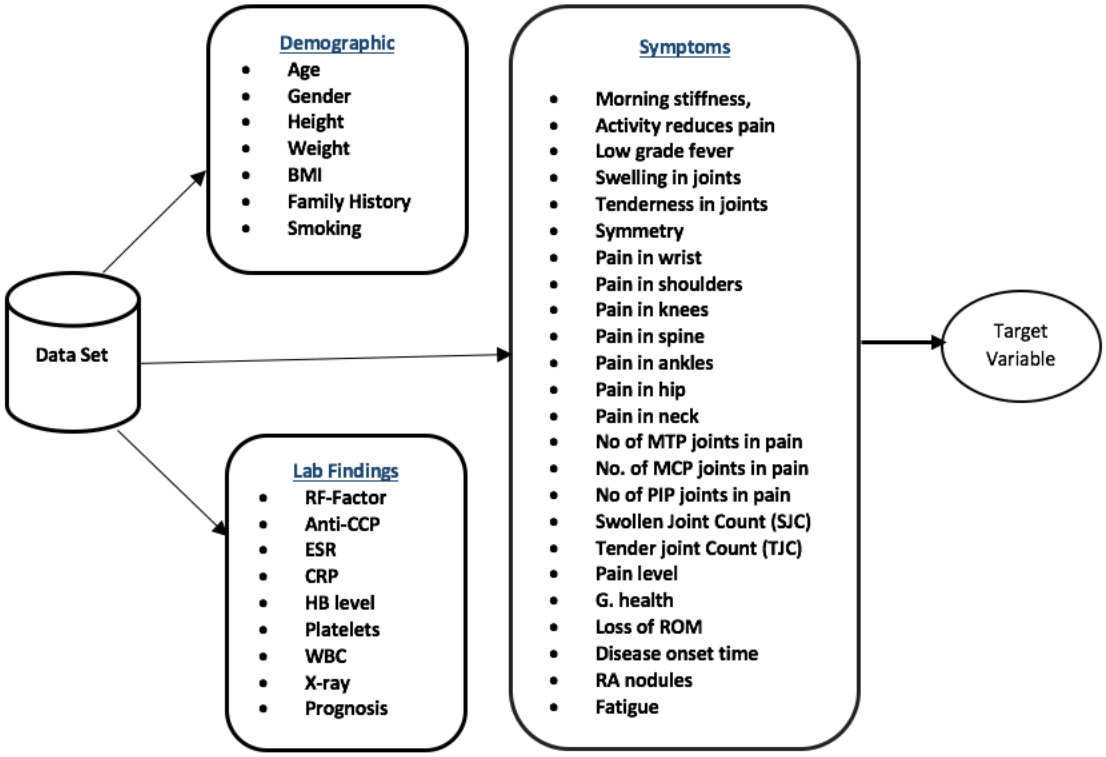

Figure 2. The dataset is developed by collecting forty parameters including anonymized demographic, clinical symptoms, and lab findings of patients visiting private hospital in Multan, Punjab.

3.1 Data pre-processing

As low-quality data can lead to erroneous predictions, data quality is a crucial aspect of disease prediction and diagnosis. During the pre-processing phase, the mean calculation method was applied to eliminate incomplete or missing values by replacing them with median and mode. We performed normalization of parameters to ensure compatibility for comparison. It involves generating z-values based on the mean and standard deviation to modify all integer data to fall inside a limited range. The z-value replaced each value of a parameter in the range [0,1].

Table 4. Range of selected features in the entire dataset

Sr.	Feature	Range
1	Age	24-64 years
2	Height in inches	58-70
3	Weight in KGs	45-90
4	BMI	17.6-34.4
5	Swollen joint count	0-9
6	Tender Joint Count	0-8
7	RF- Factor(IU/ml)	2-105
8	Anti-CCP(u/ml)	4-99
9	ESR(mm/hr)	5-110

10	CRP(mg/L)	0-88
11	HB(g/dl)	8-14.4
12	Platelets per microlitre	207000-755000
13	WBC	4.0-16.9

Feature Selection using Information Gain (IG)

Feature selection is used to select a subset of features to enhance relevancy and minimize redundancy toward the target attribute. We employed the Information Gain approach for the selection of the most pertinent features from the dataset, like that used by [19], to approximate the value of each variable by calculating the difference between post entropy and prior entropy. It constitutes a filtering approach where subsets of features are ranked in descending order, determined by their entropy gain. Information gain investigates each characteristic individually by calculating its information gain, as well as determining the value and relevance of each feature to the target class. Target class entropy is calculated from the full dataset, and conditional entropies are subtracted for all function conceivable values. The frequency count for the target class by feature value is required for entropy computation. The occurrences of each feature value were counted inside those occurrences of each class, and the entropy for that feature value was calculated. This process was repeated for each conceivable function value [20].

3.2 Percentage Split of dataset

We split the data using a random 70-30 percentage split method where 70% of the data will be used for training and 30% will be used for validation. For validation, we used performance metrics of confusion matrix, precision, accuracy, recall, Receiver Operating Characteristic (ROC) curve, and F1-score.

Machine Learning Classification Algorithms

We utilized the open-source toolkit scikit-learn, developed in Python, for executing tasks such as data pre-processing, feature selection, model development, and model evaluation. Using five models (Naïve Bayes, Support Vector Machine, Decision Tree, Random Forest, and AdaBoost), we classified patients into two groups: RA patients and non-RA patients.

3.3 Disease Activity Level

After classification, the selected features of RA patients will be used to measure the RA disease activity level using DAS-28 score calculated using Eq. 1 following the criteria listed in Table 2.

4. Results

We selected useful features for classification based on information gain score. Only those variables were included for processing where Information Gain score was greater than 0.1 in the ranking as shown in Fig. 3. This reduced the number of features to sixteen from a total of forty features for each patient in the dataset. The most useful features (IG score >0.5) for diagnosis of RA patients include Tender Joint Count (TJC) and Swollen Joint Growth (SJC) that account for tenderness and swelling in more than one joint. In patients with rheumatoid arthritis (RA), elevated levels of C-Reactive Protein (CRP) (IG score > 0.4) are an important marker associated with systemic inflammation.

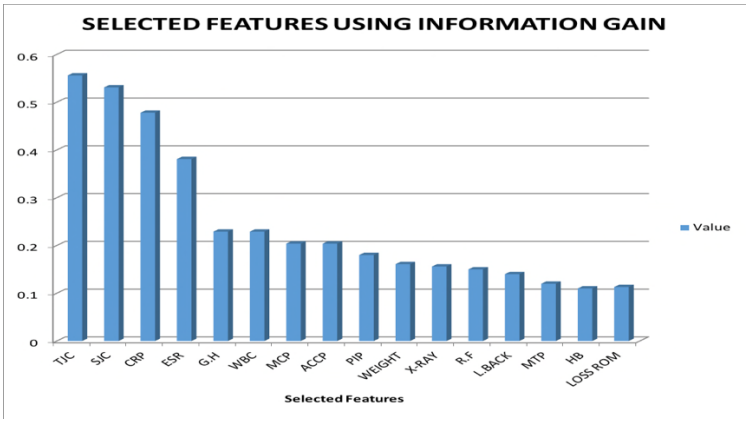

Figure 3. Sixteen selected features out of total forty features, where Information Gain score >0.1. Abbreviations: TJC: Tender Joint Count; SIC: Swollen Joint Growth; CRP: C-Reactive Protein; ESR: Erythrocyte

Sedimentation rate; G. H.: General Health; WBC: White blood cells; ; MCP: Metacarpophalangeal joints; ACCP: Anti-Cyclic Citrullinated Peptide; PIP: Proximal Interphalangeal Phalanges; WEIGHT: Body weight; X-RAY: x-ray imaging; R. F.: Rheumatoid factor; L. BACK: Low Back Pain; MTP: Metatarsophalangeal Joint; HB: Haemoglobin; LOSS ROM: Range of Motion.

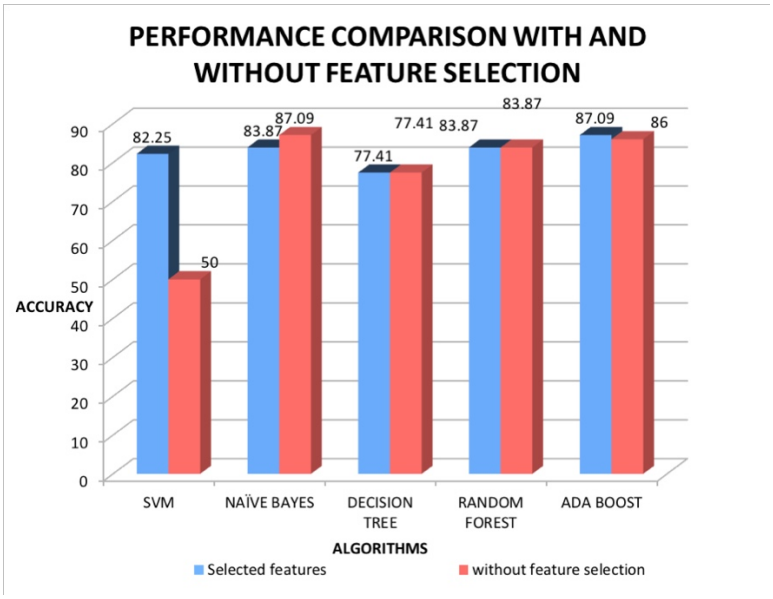

Figure 4. Performance comparison of machine learning classification models before and after feature selection based on information gain score. Feature selection improved accuracy of SVM and Adaboost, while maintain the performance of Decision Tree and Random Forest, without significant loss of information.

A useful indicator of inflammation (IG>0.3), erythrocyte sedimentation rate (ESR), also known as erythrocyte rate, is the rate at which red blood cells settle in a test tube after an hour. White blood cells (WBC) and general health (G. H) are also chosen as significant factors (IG score >0.2) because RA is an autoimmune disease in which the body's defense system mistakenly targets healthy cells, causing inflammation in the afflicted areas. Weight is associated with high chances of RA and around two-third of RA patents are obese that is one of the main reasons of sever activity level of RA disease. Low level of Hemoglobin (HB) is another main symptom since the body creates fewer red blood cells than needed in RA patients.

To determine the importance of selected features, we compared the classification performance of our models before and after feature selection. Fig. 4 shows classification accuracy of the five selected models before feature selection (using all forty features) and after feature selection (using sixteen selected features). The graph show that the feature selection technique improves classification accuracy of SVM and Adaboost, while maintaining the accuracy of Decision Tree and Random Forest. Classification after feature selection reduced the space-time complexity of the algorithms that reduced the training time without major information loss.

Table 5 shows the comparison of the accuracy, precision, recall, and F1-score of the five models used for classification. AdaBoost outperformed the other machine learning models, according to the results.

Table 5. Performance Metrics of Machine Learning Classifiers used for classification into RA and Not-RA patient.

Classification Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Naive Bayes	83.9	84.7	83.9	83.6
SVM	82.2	84	82.2	81
Decision Tree	77.4	78.0	77.0	77.4
Random Forest	83.9	85.6	83.9	83.6

Adaboost	**87.1**	**88.0**	**87.1**	**86.0**

We further take data of all RA patients and classified them according to disease activity level following the DSA-28 formula. Performance of five machine learning models for classification of low-level disease state (DAS<=3.2) is shown in Table 6:

Table 6. Performance of algorithms for classification of Low active disease

Classification Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Naive Bayes	63	87	73	91
SVM	54	100	70	91
Decision Tree	63	63	63	87
Random Forest	63	100	78	93
AdaBoost	63	100	78	93

Performance of five machine learning models for classification of moderate-level disease state (3.2<DAS<5.1) is shown in Table 7:

Table 7. Performance of algorithms for classification of Low active disease

Classification Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Naive Bayes	93	80	86	85
SVM	90	78	84	83
Decision Tree	80	80	80	80
Random Forest	93	77	84	83
AdaBoost	93	82	87	87

Performance of five machine learning models for classification of moderate-level disease state (DAS>5.1) is shown in Table 8:

Table 8. Performance of algorithms for classification of Low active disease

Classification Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Naive Bayes	93	80	86	85
SVM	90	78	84	83
Decision Tree	76	61	68	80
Random Forest	88	71	78	82
AdaBoost	93	82	87	87

Table 6 and 7 shows that AdaBoost performed better in classification of high and moderately active disease, while in case of low activity level Random Forest and AdaBoost performed similarly to one another.

5. Discussion

There are many symptoms of RA that overlap with other types of arthritis or relevant diseases. Along with these symptoms, there are many other features such as age, gender, and weight that are often ignored in the diagnostic process. The goal of this research is to develop a diagnostic tool that classifies the degree of disease activity in addition to helping with the accurate diagnosis of RA. Various treatment approaches are available based on the severity of the disease, as illustrated in Fig. 5.

Artificial Intelligence and image processing techniques proved to be very useful in disease diagnosis and scoring of disease activity level. Five common machine learning models have been used for classification of not-RA and RA patients along with the disease activity level. The models have been evaluated on an anonymized dataset of 104 patients collected from local hospital. The data was collected through questionnaire based on the most accurate criteria of RA diagnosis used by expert rheumatologists. Adaboost performs best among other classifiers achieving 87.1% accuracy and 88% precision on classification of RA

and not-RA patients, 93% accuracy and 82% precision on classification of high and moderately active disease, and 63% accuracy and 100% precision on classification of low active disease. It is hard to classify RA in early stages or low active stages due to large number of overlapping symptoms with other types of arthritis.

The proposed diagnostic tool may reduce the diagnostic time by performing automated data analysis. It may also be used as a second opinion on a medical diagnosis of RA. We are working on improving accuracy of the selected model by increasing the number of patients records in the dataset. The accuracy may be improved by integrating images data into the feature-set. The model may be linked with a mobile application to perform RA diagnosis of patients in remote areas where there is no direct access to the health facilities.

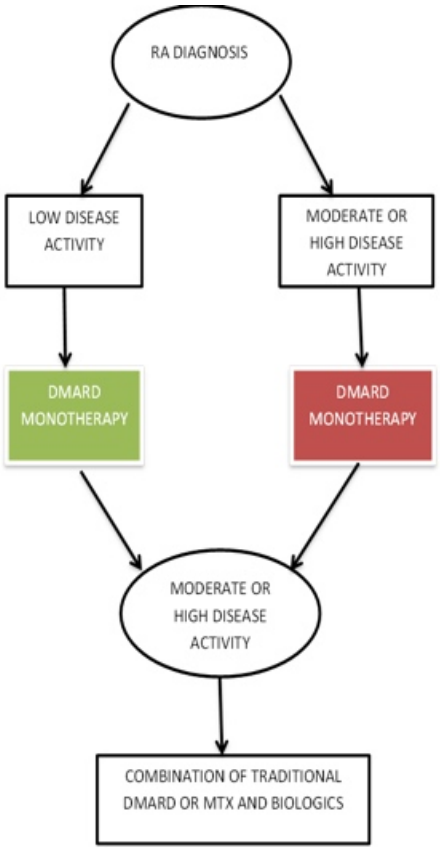

Figure 5. Treatment recommendations based on disease activity level, adapted from [21]. Classification of RA into disease activity level helps rheumatologists in adopting appropriate treatment pathway.

6. Conclusions

In conclusion, this study proposed a machine learning classifier-based diagnosis tool for RA and its disease activity levels. The research addressed the challenge of efficient and precise diagnosis associated with overlapping symptoms that resemble other forms of inflammatory arthritis. The proposed approach seeks to improve the accuracy of RA diagnosis by integrating the Disease Activity Score (DAS-28) and classifying patients according to their degree of disease activity. A dataset consisting of 104 individuals' demographics, clinical symptoms, and test results were used in this study. The performance of five machine learning classifiers: Adaboost, Random Forest, Support Vector Machine, Naïve-Bayes, and Decision Tree was evaluated for RA diagnosis on the collected dataset. An information gain-based feature selection process was used to find relevant characteristics for RA diagnosis. AdaBoost performed better than other classifiers, according to the data, showing improvements in recall, precision, F1-score, error rate, and specificity.

Acknowledgments: We acknowledge Dr. Ghulam Shabbir, Rheumatologist at Medicare, Multan for supporting in data gathering, evaluation, and labelling.

References

1. Almutairi, K, Nossent, J, Preen, D, et al. The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. Rheumatol. Int 2021; 41(5): 863–877
2. Zafar, ZA, Alam, MA, Sarfraz, M, Ahmad, et. al. Frequency of extra-articular manifestations in cohort of Pakistani patients presented with RA at Independent University Hospital Faisalabad. TPMJ 2021; 28(06): 819–827.
3. Smolen, JS, & Aletaha, D. The assessment of disease activity in rheumatoid arthritis. Clin Exp Rheumatol 2010; 28(3): 59-62.
4. Paimela, L. The radiographic criterion in the 1987 revised criteria for rheumatoid arthritis. Reassessment in a prospective study of early disease. Arthritis Rheum 1992; 35(3): 255–258.
5. Heijde, DMFM, Van't Hof, MA, Van Riel, PLCM, et al., LBA. Judging disease activity in clinical practice in rheumatoid arthritis: First step in the development of a disease activity score. ARD 1990; 49(11): 916–920.
6. Barczyńska, TA, Dura, M, Blumfield, E, et al. DAS28 score vs. Ultrasound examination for assessment of rheumatoid arthritis disease activity: Comparison and discussion of pros and cons. Reumatol. 2015; 53(4): 213–218.
7. Bai, L, Zhang, Y, Wang, P, et al. Improved diagnosis of rheumatoid arthritis using an artificial neural network. Scientific Reports 2022; 12(1): 1-7.
8. Hassanzadeh, T., Shamonin, D. P., Li, Y., Krijbolder, D. I., Reijnierse, M., van der Helm-van, A. H., & Stoel, B. C. 2024. A deep learning-based comparative MRI model to detect inflammatory changes in rheumatoid arthritis. Biomedical Signal Processing and Control, 88, 105612.
9. Wang, J., Tian, Y., Zhou, T., Tong, D., Ma, J., & Li, J. 2023. A survey of artificial intelligence in rheumatoid arthritis. Rheumatology and Immunology Research, 4(2), 69-77.
10. Ahalya, R. K., Snekhalatha, U., & Dhanraj, V. 2023. Automated segmentation and classification of hand thermal images in rheumatoid arthritis using machine learning algorithms: A comparison with quantum machine learning technique. Journal of Thermal Biology, 111, 103404.
11. Imtiaz, M., Shah, S. A. A., & ur Rehman, Z. 2022. A review of arthritis diagnosis techniques in artificial intelligence era: Current trends and research challenges. Neuroscience Informatics, 2(4), 100079.
12. Nieuwenhuis, W. P., Krabben, A., Stomp, W., Huizinga, T. W., van der Heijde, D., Bloem, J. L., ... & Reijnierse, M. 2015. Evaluation of magnetic resonance imaging–detected tenosynovitis in the hand and wrist in early arthritis. Arthritis & rheumatology, 67(4), 869-876.
13. Maziarz, K., Krason, A., & Wojna, Z. 2021. Deep learning for rheumatoid arthritis: Joint detection and damage scoring in x-rays. arXiv preprint arXiv:2104.13915.
14. Bai, L., Zhang, Y., Wang, P., Zhu, X., Xiong, J. W., & Cui, L. 2022. Improved diagnosis of rheumatoid arthritis using an artificial neural network. Scientific Reports, 12(1), 9810.
15. Mc Ardle, A., Kwasnik, A., Szentpetery, A., Hernandez, B., Parnell, A., de Jager, W., ... & Pennington, S. R. 2022. Identification and evaluation of serum protein biomarkers that differentiate psoriatic arthritis from rheumatoid arthritis. Arthritis & Rheumatology, 74(1), 81-91.
16. Rychkov D, Neely J, Oskotsky T, et al. Cross-Tissue Transcriptomic Analysis Leveraging Machine Learning Approaches Identifies New Biomarkers for Rheumatoid Arthritis. Front Immunol. 2021;12:638066.
17. Andersen, JKH, Pedersen, JS, Laursen, MS, et al. Neural networks for automatic scoring of arthritis disease activity on ultrasound images. RMD open 2019; 5(1): e000891.
18. Salmeron, JL, Rahimi, SA, Navali, AM, et al. Medical diagnosis of Rheumatoid Arthritis using data driven PSO–FCM with scarce datasets. Neurocomputing 2017; 232: 104–112.
19. Ahmad, W, Ahmad, A, Iqbal, A, et al. Intelligent hepatitis diagnosis using adaptive neuro-fuzzy inference system and information gain method. Soft Comput. 2019; 23(21): 10931–10938.
20. Alhaj, TA, Siraj, MM, Zainal, A, et al. Feature selection using information gain for improved structural-based alert correlation. PLoS one 2016; 11(11): 1–18.

21. Jasvinder, S, Saag, KG, & Bridges, SL. American College of Rheumatology Guideline for the treatment of rheumatoid arthritis. Arthritis Care Res 2015; 67(10): 1335-486.