



Article Title: The Significance of artificial intelligence in the early diagnosis of multiple myeloma

The Significance of artificial intelligence in the early diagnosis of multiple myeloma

Anisha S S¹, Shajin Nargunam A²

ABSTRACT

A malignancy known as multiple myeloma develops in a type of white blood cell known as a plasma cell. The plasma cells are responsible for the production of antibodies. In multiple myeloma, healthy blood cells are displaced by malignant plasma cells that build up in the bone marrow. Using machine learning techniques, artificial intelligence has radically changed the world of oncology research. Machine learning is a branch of artificial intelligence that makes use of algorithms to analyse data, draw conclusions from it, and then utilise those conclusions to make decisions in the present. In this study, we explore the potential applications of artificial intelligence in the diagnosis of multiple myeloma and present the most important machine learning and deep learning experiments conducted in the field. One of the most severe haematological cancers worldwide is multiple myeloma.

Keywords: Multiple Myeloma, Artificial Intelligence. Machine Learning, Deep Learning, Diagnosis, Bone Disease

1 Introduction

Artificial intelligence is the simulation of human intelligence processes by machines, especially computers. Numerous tools and algorithms are part of artificial intelligence, which can automate the process of developing hypotheses. Most common subsets of artificial intelligence are machine learning (ML) and deep learning (DL).

Machine Learning focus on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. ML typically uses three categories of algorithms: supervised, unsupervised, and reinforcement (Figure 1).

Deep learning, another subset of AI (DL), layers algorithms to form artificial neural networks (ANNs) that can learn and make wise decisions on their own. In essence, it more closely matches the operation of a genuine "brain," resulting in a "human-like" AI method. DL uses two different categories of algorithms: supervised learning, unsupervised learning (Figure 2).



Article Title: The Significance of artificial intelligence in the early diagnosis of multiple myeloma

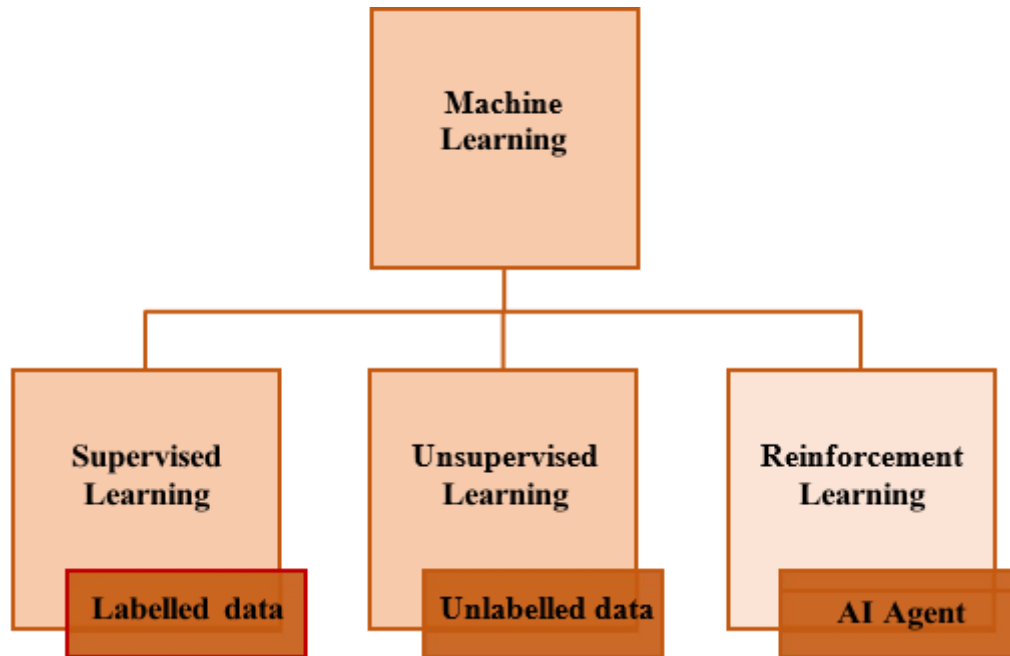


Figure 1: *Classification of ML algorithms*

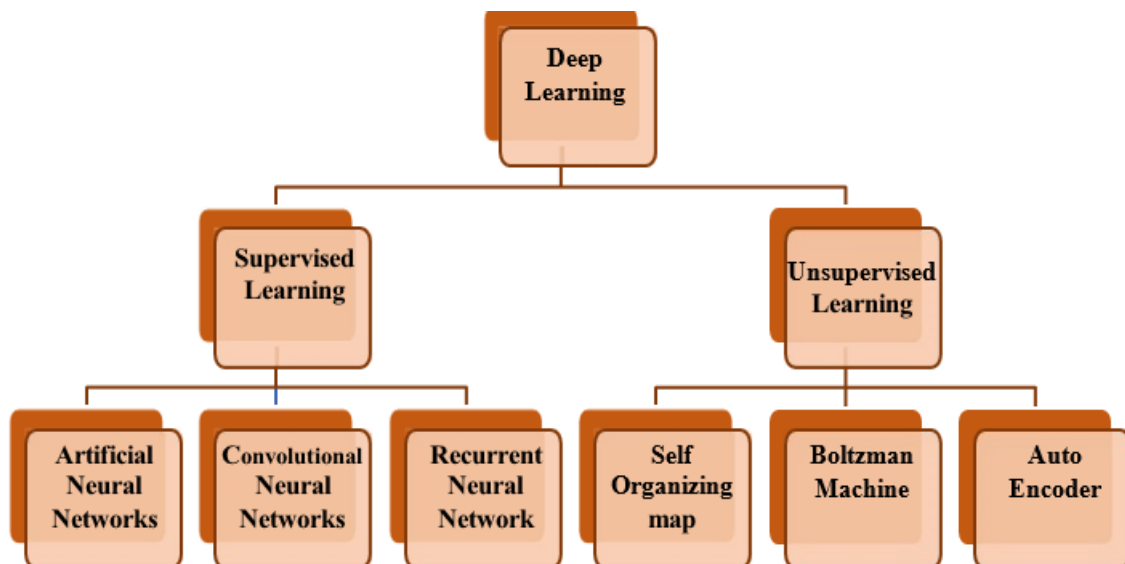


Figure 2: *Classification of Deep Learning algorithms*

Convolutional neural networks are widely employed in image processing. CNN is frequently used in areas such as face recognition, text analysis, tissue localization, and biological image detection or recognition [1].

Multiple myeloma is a haematological malignancy characterized by the uncontrolled proliferation of plasma cells in the bone marrow. Multiple myeloma is typically staged based



Article Title: The Significance of artificial intelligence in the early diagnosis of multiple myeloma

on the International Staging System (ISS), which takes into account levels of specific blood markers (beta-2 microglobulin and albumin) and genetic abnormalities. We discuss the challenges associated with early detection of multiple myeloma and explore recent advancements in screening, imaging, and biomarker-based approaches.

2 Related Works

Various studies focus on AI applications in multiple myeloma diagnoses.

Machine learning methods to identify multiple myeloma consist of the work by Waddellet al. [5] employed support vector machines to forecast the likelihood of multiple myeloma's chance for early development using machine learning.

In the study conducted by Bhattacharyya et al. [3] Machine learning was used to discover potential biomarkers suggestive of bone lesions in patient blood serum samples to detect bone lesions. Random forests and partial least squares discriminant analysis were both shown to be efficient.

Convolutional neural networks are used by Turki et al. [4] with transfer Learning techniques for Multiple Myeloma Diagnosis.

The work by Xu et al. [2] detected bone abnormalities from images using machine learning techniques. The study tested multiple machine learning algorithms for this goal, and convolutional neural networks appeared to yield the greatest accuracy.

In the study by D. Pamela et al. [6] Multiple myeloma detection accuracy has been demonstrated to be improved by combining several detection methods. For instance, it has been demonstrated that combining imaging and flow cytometry improves the precision of identification than using only one of these methods alone.

3 The Learning Approaches

This section describes the machine learning and deep learning techniques used for multiple myeloma diagnosis.

3.1 Artificial Neural Network (ANN)

The ANN is explained in this section. One input layer, three hidden layers, and one output layer make up the architecture. The input layer's size is determined by the size of the input image. The He Normal initialization method [7], which helps prevent overloading and performs well for corrected activations, is used to initialize the weights in the hidden layers. The Adam optimizer [8] is employed due to its acceptance, general effectiveness, and speed in use.

3.2 Convolutional Neural Network (CNN)

The ImageNet's VGG16 model, developed by K. Simonyan et al. [9] is the architecture used. a scaled-down, personalized Initially, CNN was put to the test, the VGG16 network outperformed it in terms of precision as well as speed.



Article Title: The Significance of artificial intelligence in the early diagnosis of multiple myeloma

3.3 Support Vector Machine (SVM)

Support vector machine (SVM) is a well-liked machine learning (ML) technique for problems including classification and regression. Vapnik [10] developed SVM in the second half of the 20th century. SVMs have been widely used in many other fields than disease detection, such as protein fold, distant homology identification, speech detection, and text recognition. While using SVM for data analysis [11] choosing the right kernel and variables is crucial.

3.4 Random Forest (RF)

The Random forest was implemented using the Scikit-Learn toolkit. The number of features to pick at each node is a crucial decision. Internal computations of the dependence, classifier strength, and generalization error are produced for guidance [12]. These are referred to as "out-of-bag estimates. The feature split criteria, an arbitrary state, and the number of trees to be employed are initialized in the RF classifier.

3.5 Decision Tree

Rule-of-thumb for the decision tree (DT) algorithm is divide and conquer. The property in DT models may have several values known as classification trees; leaves denote unique classes, while branches denote the mix of qualities that give rise to those class labels[13].

3.6 K-Nearest Neighbour

Evelyn Fix and Joseph Hodges created the nonparametric classification method known as K-nearest neighbour (KNN) in 1951. Both classification and regression analysis can be done with KNN. Class membership is the result of KNN classification. A voting system is utilized to categorize the samples. The distance between two data samples is calculated using Euclidean distance methods. The average of the KNN values is used to project the value in regression analysis [14].

3.7 Naive Bayes

The naive Bayes (NB) classifier is a probabilistic classifier with a Bayesian foundation. It makes predictions about the likelihood of each class's membership based on a particular record or data point. The most likely class is the one with the highest likelihood. The NB classifier is used to forecast the possibility rather than make predictions [15].

4 Discussion

The artificial intelligence techniques used for the identification of multiple myeloma detailed in the related works are analysed in Table 1. Although identification and diagnosis of multiple myeloma (MM) using the current AI-based detection systems have shown promising results, each scheme has pros and cons.



Article Title: The Significance of artificial intelligence in the early diagnosis of multiple myeloma

Table 1: AI techniques in Multiple Myeloma Detection

AI-Based Multiple Myeloma Detection Schemes	Imaging Modalities	AI Techniques	Pros	Cons
Blood Tests and Biomarkers	Blood samples	SVM, RF, CNN, RCNN, Principal Component Analysis	provide prognostic information, helps to predict disease aggressiveness and patient outcomes, Wide Availability	Limited Specificity, complex to interpret, Biomarker levels can fluctuate over time. Limited Predictive Value
Medical Imaging Analysis	X-ray, CT, MRI, and PET scans	CNN, Transfer learning	Non-Invasive Enhanced Accuracy	Dependency on Imaging Modalities, False Positives, and False Negatives
Radiomics	X-rays, CT, MRI, and PET scans	feature selection methods and classifiers,	Non-Invasive, visualization of lesions aiding in diagnosis, staging, and monitoring treatment response	Limited Soft Tissue Differentiation, Variations in interpretation can lead to discrepancies in diagnosis and treatment decisions
Prognostic Prediction	patient demographic s, genetic profiles, treatment histories, and response data,	SVM, RF, RNN	predict disease progression and patient outcomes based on medical image data and clinical variables	Interpretability. Model Overfitting, Radiation Exposure, requires specialized training and expertise
Decision Support Systems	X-rays, CT, MRI, PET, ultrasound and mammography	decision trees, RF, CNN, RCNN, Bayesian Networks	Timely Diagnosis, Improved Treatment Planning, Integration of Multiple Data Sources	Inaccurate or incomplete data can lead to erroneous recommendations or missed diagnoses, Technological Dependence



Article Title: The Significance of artificial intelligence in the early diagnosis of multiple myeloma

Lesion Segmentation	X-rays, CT, MRI, and PET scans	Deep learning techniques, such as U-Net, CNN, and Mask R-CNN.	Identify and delineate multiple myeloma lesions within medical images.	Interpretability. Model Overfitting
---------------------	--------------------------------	---	--	-------------------------------------

The above-mentioned works identify multiple myeloma only in advanced stages. The factors which are known to increase the risk of developing multiple myeloma include age (most commonly diagnosed in people over 65), gender (slightly more common in males), race (more prevalent in African Americans), family history of the disease, and exposure to certain chemicals or radiation. Multiple myeloma often presents with nonspecific symptoms, which may include bone pain (especially in the back or ribs), fatigue, weakness, recurrent infections, frequent fractures, unexplained weight loss, kidney problems, and anaemia. Here lies the importance of early detection of this disease.

Machine learning algorithms can analyze large-scale genomic, proteomic, or metabolomic datasets to identify biomarkers associated with multiple myeloma. By considering various molecular signatures, these algorithms can potentially detect early signs of the disease or predict its progression.

Performance metrics are used to evaluate the effectiveness of machine learning and deep learning models in multiple myeloma detection. Some commonly used performance metrics are Accuracy, Sensitivity, Specificity, and Area under the Receiver Operating Characteristic Curve.

5 Result

In this study, the scope of AI techniques for detecting multiple myeloma is focussed. Various machine learning and deep learning algorithms were used in the existing works. From the literature, it is evident that myeloma can be diagnosed with the aid of AI. A well-trained AI algorithm on a wide range of data in serology, histology, genetic data, and radiology can detect multiple myeloma in early stage.

6 Conclusion

Researchers are particularly interested in certain diseases, like heart disease, breast cancer, kidney disease, diabetes, Alzheimer's, and Parkinson's diseases, that have been addressed considering machine learning/deep learning-based techniques. The survival rate of multiple myeloma varies based on many factors. These include the stage of cancer, the person's age and general health, and how well the treatment plan works. The disease needs to be diagnosed as early as possible since the stage of cancer determines the survival rate. In this paper, we have presented an overview of various algorithms that are quite promising in the detection of multiple myeloma. Different algorithms are compared to get a clear view of how to choose a



Article Title: The Significance of artificial intelligence in the early diagnosis of multiple myeloma

better algorithm for detection.

Multiple Myeloma detection may be more precise if Machine learning algorithms are trained on a wide range of data types including medical images, genomic data, and clinical data. Future research could focus on the development of personalized treatment plans based on individual patient characteristics and medical history. Further research and validation on larger datasets and in real-world clinical settings are recommended to confirm the findings of this study and explore the generalizability and robustness of the identified AI techniques. Continued advancements in artificial intelligence and machine learning hold great potential for improving multiple myeloma detection, leading to earlier diagnoses, better treatment outcomes, and enhanced patient care. It's important to note that while AI-based detection schemes have shown promise in multiple myeloma diagnoses, they are not intended to replace medical professionals. These schemes aim to augment and assist healthcare providers in the decision-making process, improving accuracy and efficiency. Medical experts should always be involved in the final diagnosis and treatment decisions.

References

1. Moi Hoon Yap; Gerard Pons; Joan Marti; Sergi Ganau; Melcior Sentis; Reyer Zwiggelaar; Adrian K. Davison; Robert Marti, Year: 2017, "Automated breast ultrasound lesions detection using convolutional neural networks", IEEE journal of biomedical and health informatics, Vol: 22, no: 4, pp.1218 – 1226.
2. Michael Waddell; David Page; John Shaughnessy Jr, Year: 2005, "Predicting cancer susceptibility from single-nucleotide polymorphism data: a case study in multiple myeloma", In Proceedings of the 5th international workshop on Bioinformatics, pp. 21 – 28.
3. Sudeepa Bhattacharyya; Joshua Epstein; Larry J. Suva, Year: 2006, "Biomarkers that discriminate multiple myeloma patients with or without skeletal involvement detected using SELDI-TOF mass spectrometry and statistical and machine learning tools", Disease Markers, Vol: 22, no: 4, pp. 245 – 255.
4. Turki Turki; Zhi Wei; Jason TL Wang, Year: 2017, "Transfer learning approaches to improve drug sensitivity prediction in multiple myeloma patients", IEEE Access, Vol: 5, pp. 7381 – 7393.
5. Lina Xu; Giles Tetteh; Jana Lipkova; Yu Zhao; Hongwei Li; Patrick Christ; Marie Piraud; Andreas Buck; Kuangyu Shi; Bjoern H. Menze, Year: 2018, "Automated whole-body bone lesion detection for multiple myeloma on 68 Ga-Pentixafor PET/CT imaging using deep learning methods", Contrast media & molecular imaging 2018.
6. M. M. Shinu; D. Pamela; G. Glan Devadhas, Year: 2022, "A Systematic Review on Recent Techniques for the Detection of Multiple Myeloma", In 2022 Third International Conference on



Article Title: The Significance of artificial intelligence in the early diagnosis of multiple myeloma

Intelligent Computing Instrumentation and Control Technologies (ICICICT), pp. 1573 – 1579.

7. Kaiming He; Xiangyu Zhang; Shaoqing Ren; Jian Sun, Year: 2015, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification”, In Proceedings of the IEEE international conference on computer vision, pp. 1026 – 1034.
8. D. Kingma; B. J. Adam, Year: 2014, “A method for stochastic optimization”, arxiv [cs. lg]. 2014 (2017)
9. Karen Simonyan; Andrew Zisserman, Year: 2014, “Very deep convolutional networks for large-scale image recognition”, arXiv preprint arXiv: 1409.1556.
10. Harris Drucker; Donghui Wu; Vladimir N. Vapnik, Year: 2019, “Support vector machines for spam categorization”, IEEE Transactions on Neural networks, Vol: 10, no: 5, pp.1048 – 1054.
11. Essam H. Houssein; Marwa M. Emam; Abdelmgeid A. Ali; Ponnuthurai Nagarathnam Suganthan, Year: 2021, “Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review”, Expert Systems with Applications, Vol: 167, pp.114161.
12. Leo Breiman, Year: 2001, “Random forests”, Machine learning, Vol: 45, pp. 5 – 32.
13. Rajesh S. Walse; Gajanan D. Kurundkar; Santosh D. Khamitkar; Aniket A. Muley; Parag U. Bhalchandra; Sakham N. Lokhande, Year: 2020, “Effective use of naïve bayes, decision tree, and random forest techniques for analysis of chronic kidney disease”, In Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2020, Vol: 1, pp. 237 – 245.
14. Evelyn Fix; Joseph Lawson Hodges, Year: 1989, “Discriminatory analysis. Nonparametric discrimination: Consistency properties”, International Statistical Review/Revue Internationale de Statistique, Vol: 57, no: 3, pp. 238 – 247.
15. A. Nurrohmah; S. Abdullah; H. Murfi, Year: 2020, “Parkinson’s disease subtype classification: Application of decision tree, logistic regression and logit leaf model”, In AIP Conference Proceedings, Vol: 2242, no: 1. AIP Publishing.