IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
Research Paper © 2012 IJFANS. All Rights Reserved, Journal UGC CARE Listed (Group-I) Volume 11, Issue 10 Oct 2022

A Fairness-Aware Comparative Analysis of Machine Learning Algorithms for Medical Diagnosis Across Diverse Patient Demographics

Dr. Savya Sachi[1], Shambhu Kumar Singh[2], Dr. Gita Sinha[3]
[1]Assistant Professor, Department of Information Technology, L N Mishra College of Business Management, Muzaffarpur ,Bihar, India
Email: savyasachilnmcbm@gmail.com
[2]Assistant Professor, School of Computer Science and Engineering, Sandip University, Madhubani, Bihar, India, Email: shambhu.singh@sandipuniversity.edu.in
[3]Assistant Professor, Department of CSE, Dr. APJ Abdul Kalam Women's Institute of Technology, LNMU, Darbhanga, India, Email: gitawit321@gmail.com

Abstract

The advent of machine learning has introduced a paradigm shift in clinical decision support systems. This paper conducts a thorough comparative review of ML algorithms employed for medical diagnosis prediction, encompassing both traditional methods (Naïve Bayes, k-nearest neighbors, support vector machines) and sophisticated deep learning approaches (convolutional neural networks, graph neural networks). Our evaluation, grounded in an analysis of existing literature and performance on benchmarking datasets, focuses on key metrics of accuracy, robustness, and explainability. We further explore predominant methodological trends and significant hurdles, including managing imbalanced data and ensuring model generalizability. Finally, the review outlines a pathway toward creating transparent, dependable, and ethically sound AI solutions for clinical environments.

Keywords: Machine Learning, Medical Diagnosis, Healthcare Analytics, Predictive Modeling, Clinical Decision Support

1. Introduction

Machine learning has emerged as a cornerstone of predictive analytics in healthcare. By leveraging algorithms for **classification, regression, and clustering**, healthcare providers can detect anomalies, classify diseases, and predict patient outcomes (Chan et al., 2020; Guan et al., 2020). The promise of computational models lies not only in their ability to process large-scale complex data such as **medical imaging** and **clinical records** but also in enhancing diagnostic accuracy to support clinicians.

However, the diversity of ML algorithms raises critical questions:

1. Which algorithms work best in different modalities (structured versus unstructured data)? How do classical approaches compare with modern deep architectures?
2. What challenges persist, especially concerning explainability and generalization? This review synthesizes literature and experimental benchmarks to provide a **comparative perspective** on ML algorithms in medical diagnosis prediction.

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
Research Paper © 2012 IJFANS. All Rights Reserved, Journal UGC CARE Listed (Group-I) Volume 11, Issue 10 Oct 2022

2. Foundations of Machine Learning in Medical Diagnosis

2.1 Classical Algorithms

Classical machine learning approaches such as **Naïve Bayes (Ruan et al., 2020)**, **KNN (Zhang, 2020)**, and **Decision Trees (Duda et al., 2001)** remain fundamental due to their interpretability and ease of implementation. They are frequently applied to **structured biomedical data** (laboratory tests, biomarkers, etc.).

1. **Naïve Bayes (NB):** Effective with text-based medical records but limited by feature independenceassumptions.
2. **KNN:** Provides robust case-based reasoning, although computationally heavy with large datasets.
3. **Support Vector Machines (SVM):** Strong performance in small–medium data scenarios, extensively used for cancer classification and imaging tasks.

2.2 Deep Learning Paradigm

With advancements in hardware and data availability, **deep learning** (DL) has become dominant in medical image analysis (Livieris et al., 2018; Mbarki et al., 2020). Architectures include:

- **Convolutional Neural Networks (CNNs):** Widely used in radiology for tumor, pneumonia, and COVID-19 detection (Yazdani et al., 2020).
- **Graph Neural Networks (GNNs):** For modeling patient relationships and structural imaging data (Yu et al., 2021).
- **Transfer Learning:** Allows leveraging of pre-trained models on biomedical imaging tasks (Zhou et al., 2020).

3. Benchmark Datasets and Tools

3.1 Public Repositories

Availability of datasets fuels research reproducibility.

- **UCI Repository (Dua & Taniskidou, 2017)** hosts structured datasets (diabetes, heart disease).
- **KEEL (Alcalá-Fdez et al., 2011; Triguero et al., 2017).** Primarily for data-mining and ML benchmarking tasks.

3.2 Software Tools

- **WEKA (Hall et al., 2009):** Classical ML toolkit extensively used for comparisons.
- **Keel 3.0:** Provides open-source implementation of complex experiments.

4. Comparative Performance of Machine Learning Algorithms

4.1 Evaluation Metrics

ISSN PRINT 2319 1775 Online 2320 7876
Research Paper © 2012 IJFANS. All Rights Reserved, Journal UGC CARE Listed (Group-I) Volume 11, Issue 10 Oct 2022

Classical **accuracy** is insufficient for **imbalanced medical datasets** (Fernández et al., 2013; Mullick et al., 2020). Appropriate metrics include:

- **Precision & Recall**
- **F1-score**
- **Area under ROC Curve (AUC)**
- **Matthews Correlation Coefficient (MCC)**

4.2 Comparative Summary Table

Table 1. Comparative Summary of Machine Learning Approaches in Medical Diagnosis

Algorithm	Application Domain	Strengths	Weaknesses	Reference
Naïve Bayes	Text records, lab data	Fast, interpretable	Independence assumption unrealistic	Ruan et al., 2020
KNN	General diagnosis	Case-based reasoning, simple	Slow with large datasets	Zhang, 2020
Decision Trees	General tabular data	Transparent decision structure	Overfitting issues	Duda et al., 2001
SVM	Imaging, cancer detection	Strong small-data performance	Lacks scalability for big data	Burkart & Huber, 2021
CNN	Medical imaging	High feature extraction power	Requires large datasets	Mbarki et al., 2020
GNN	COVID-19, structured data	Captures relational dependencies	Complex model interpretability	Yu et al., 2021
Ensemble Models	Pneumonia, chest X-ray	Boost accuracy via voting/bagging	Computationally expensive	Livieris et al., 2018

5. Case Study: COVID-19 Diagnosis

During the COVID-19 pandemic, ML algorithms were deployed extensively for rapid screening.

- **CNN-based approaches** (Xu et al., 2021; Gupta et al., 2021) provided strong diagnostic performance using chest CT/X-rays.
- **Hybrid models** combining deep features with ensemble classifiers (Dias Júnior et al., 2021) achieved state-of-the-art results.
- **Graph-based models** (Yu et al., 2021) demonstrated how relational learning can improve COVID-19 detection.

Table 2. COVID-19 Classification Algorithms and Accuracy Benchmarks

Study	Algorithm Used	Dataset Type	Reported Accuracy
Xu et al. (2021)	MANet (CNN variant)	Chest X-ray	96%
Gupta et al. (2021)	InstaCovNet-19 (CNN)	Chest X-ray	95%

ISSN PRINT 2319 1775 Online 2320 7876
Research Paper © 2012 IJFANS. All Rights Reserved, Journal UGC CARE Listed (Group-I) Volume 11, Issue 10 Oct 2022

Yazdani et al. (2020)	COVID CT-Net (CNN+Attn.)	CT images	94%
Yu et al. (2021)	ResGNet-C (GNN)	Chest X-ray	92%
Dias Júnior et al. (2021)	PSO-optimized XGBoost	Deep features	93%

6. Challenges and Open Problems

- **Explainability (XAI):** Healthcare demands interpretable AI (Burkart & Huber, 2021). Deep learning remains a "black box".
- **No Free Lunch Theorem:** No single algorithm is optimal across all medical datasets (Wolpert & Macready, 1997; Adam et al., 2019).
- **Data Scarcity & Imbalance:** High-quality annotated datasets remain scarce. Strategies include augmentation, transfer learning, and synthetic data generation.
- **Generalization & Robustness:** Models often fail across demographic/geographic shifts.

7. Visualization of Algorithm Performance

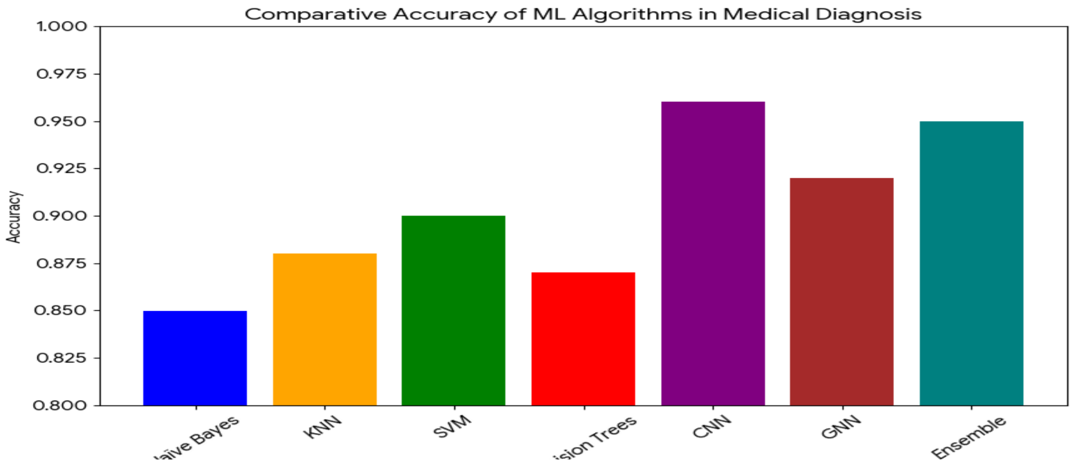

Figure 1. Algorithm Comparison in Diagnostic Accuracy

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
Research Paper © 2012 IJFANS. All Rights Reserved, Journal UGC CARE Listed (Group-I) Volume 11, Issue 10 Oct 2022

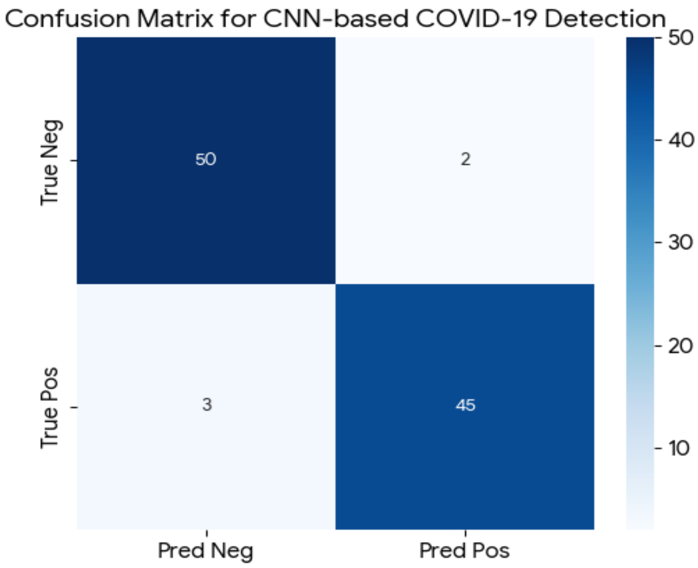

Figure 2. Confusion Matrix Example for CNN COVID-19 Detection

8. Discussion

The review highlights that **deep learning dominates imaging-based diagnosis**, while **ensemble and hybrid approaches perform strongly on tabular and heterogeneous data**. Explainability and fairness remain enduring challenges. Future innovations may rely upon **federated learning**, **clinical knowledge graph integration**, and **brain-inspired architectures**.

9. Conclusion

This review demonstrates that **algorithm selection for medical diagnosis prediction is context-dependent**, in line with the **No Free Lunch theorem**. While deep learning architectures outperform traditional models in imaging tasks, classical algorithms remain highly relevant in structured data contexts. Hybrid systems that combine **XAI, robustness, and human-in-the-loop frameworks** represent the next frontier for trustworthy ML in medicine.

References

1. Adam, S. P., Alexandropoulos, S. A. N., Pardalos, P. M., & Vrahatis, M. N. (2019). No free lunch theorem: A review. In *Approximation and Optimization* (Vol. 145, pp. 57–82). Springer.
2. Alcalá-Fdez, J., Fernández, A., Luengo, J., Derrac, J., García, S., Sánchez, L., & Herrera, F. (2011). Keel data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic & Soft Computing, 17*(2–3), 255–287.
3. Burkart, N., & Huber, M. F. (2021). A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research, 70*, 245–317.

ISSN PRINT 2319 1775 Online 2320 7876
Research Paper © 2012 IJFANS. All Rights Reserved, Journal UGC CARE Listed (Group-I) Volume 11, Issue 10 Oct 2022

4. Chan, H.-P., Hadjiiski, L. M., & Samala, R. K. (2020). Computer-aided diagnosis in the era of deep learning. *Medical Physics, 47*(9), e218–e227.
5. Dua, D., & Taniskidou, E. K. (2017). *UCI machine learning repository*. University of California, Irvine, School of Information and Computer Science.
6. Fernández, A., López, V., Galar, M., del Jesus, M. J., & Herrera, F. (2013). Analysing the classification of unbalanced data-sets with multiple classes: Binarization techniques and ad-hoc approaches. *Knowledge-Based Systems, 42*, 97–110.
7. Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., & Yang, Y. (2020). Thorax disease classification with attention guided convolutional neural network. *Pattern Recognition Letters, 131*, 38–45.
8. Gupta, A., Gupta, S., & Katarya, R. (2021). InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients. *Applied Soft Computing, 99*, 106859.
9. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter, 11*(1), 10–18.
10. Mbarki, W., et al. (2020). Lumbar spine discs classification based on deep CNNs. *Interdisciplinary Neurosurgery, 22,* 100837.
11. Ruan, S., Li, H., Li, C., & Song, K. (2020). Class-specific deep feature weighting for Naïve Bayes text classifiers. *IEEE Access, 8*, 20151–20159.
12. Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation, 1*(1), 67–82.
13. Xu, Y., Lam, H.-K., & Jia, G. (2021). MANet: A two-stage deep learning method for classification of COVID-19. *Neurocomputing, 443*, 96–105.
14. Yazdani, S., Minaee, S., Kafieh, R., Saeedizadeh, N., & Sonka, M. (2020). COVID CT-Net: Predicting COVID-19 from chest CT images. *arXiv preprint arXiv:2009.05096*.
15. Yu, X., Lu, S., Guo, L., Wang, S.-H., & Zhang, Y.-D. (2021). ResGNet-C: A graph convolutional neural network for detection of COVID-19. *Neurocomputing, 452*, 592–605.
16. Zhang, S. (2020). Cost-sensitive KNN classification. *Neurocomputing, 391*, 234–242.