Hybrid AI Approaches for Clinical Diagnostics: A Case Study on Meningitis

Ibtasam Ur Rehman[1], Abdulraqeb Alhammadi[2], and Jibran K. Yousafzai[3]

[1] [1] Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam
ribtrasam.sdh231@hcmut.edu.vn
[2] [2] Faculty of Artificial Intelligence, University of Technology Malaysia, Kuala Lumpur, Malaysia
abdulraqeb.alhammadi@utm.my
[3] [3] College of Engineering and Applied Sciences, American University of Kuwait, Kuwait
jyousafzai@auk.edu.kw

Abstract. This paper presents a comprehensive comparative analysis of machine learning and deep learning approaches for accurate meningitis classification. Through extensive evaluation of five algorithms including Multi-Layer Perceptron (MLP), XGBoost, Support Vector Machine (SVM), Random Forest, and AdaBoost, we assessed performance using clinical biomarkers from a dataset of patient records. Experimental results demonstrate that MLP Classifier achieved the highest accuracy (93.83%) while XGBoost attained superior ROC AUC (0.991). The study provides detailed analysis of clinical metrics including sensitivity, specificity, and computational efficiency, offering valuable insights into algorithm performance trade-offs for meningitis diagnosis. This research contributes to the understanding of machine learning applications in clinical diagnostics and supports the development of effective decision support tools for meningitis differential diagnosis.

Keywords: Meningitis Classification · Machine Learning · Deep Learning · Clinical Decision Support · Healthcare AI

1 Introduction

1.1 Background and Motivation

Meningitis represents a critical medical emergency where rapid diagnosis is essential, as delayed treatment can lead to severe complications or mortality within hours. Studies show that nearly 90% of pediatric meningococcal deaths occur within 24 hours of diagnosis [1], highlighting the urgent need for faster diagnostic tools. Traditional methods relying on laboratory tests and clinical assessment often cause dangerous treatment delays, particularly in emergency settings. The diagnostic challenge is further complicated by the diverse etiological landscape of

meningitis. Bacterial meningitis remains particularly severe, with incidence ranging from 0.05 to 0.33 per 100,000 annually in developed settings and case fatality rates of 3-4% despite appropriate antimicrobial therapy [2]. While vaccine implementation has reduced the burden from specific serogroups, emerging variants like MenW are associated with increased severity and mortality rates up to 25% [2]. Recent advances in machine learning have demonstrated significant potential for improving meningitis diagnosis. Choi et al. [3] developed an AI model using initial 24-hour patient data to classify meningitis causes into four categories, achieving 89% accuracy and outperforming experienced neurologists in early diagnosis. Similarly, Ghaddaripouri et al. [4] reviewed 16 ML studies involving over 39,000 patients, finding that algorithms like Artificial Neural Networks (ANN) achieved up to 90% diagnostic accuracy using clinical features. Amanullah et al. [5] developed a stacked ensemble of Random Forest and ANN that achieved 92.20% accuracy in classifying meningitis types using CSF biomarkers, while Caragheorgheopol et al. [6] used Random Forest on cerebrospinal fluid cytokine profiles to achieve 91% sensitivity and 93% specificity in distinguishing bacterial from viral meningitis.

1.2 Challenges and Research Contributions

Traditional meningitis diagnosis is hindered by significant delays, as conventional lab methods like cerebrospinal fluid culture take 24-48 hours, creating dangerous treatment gaps. This problem is compounded by overlapping symptoms between bacterial and viral meningitis, variable test sensitivity, and limited access to advanced diagnostics in resource-constrained settings. Previous research has explored various computational approaches to address these limitations. Messai et al. [7] trained a model on 934 patient cases that achieved 88% accuracy for meningococcal meningitis diagnosis while identifying key clinical features. Priya et al. [8] developed an SVM classifier optimized with the Adagrad algorithm using CSF analysis data to distinguish between presence and absence of meningitis, showing improved performance over standard SVM. Lelis et al. [9] created a non-invasive Clinical Decision Support System for early meningitis diagnosis using decision tree models, achieving 94.3% classification accuracy on 26,228 patient records.

2 Methodology

This study presents comparative analysis of ML and DL approaches for classifying meningitis types using clinical biomarkers. This methodology consist of data preprocessing, model development and comprehensive evaluation. The research flow is depicted in the Methodological Framework for Meningitis Classification (Figure 1). Five diverse algorithms were implemented and evaluated, including Multi-Layer Perceptron (MLP), XGBoost, Support Vector Machine (SVM), Random Forest and AdaBoost enabling robust performance comparison across different algorithmic approaches.

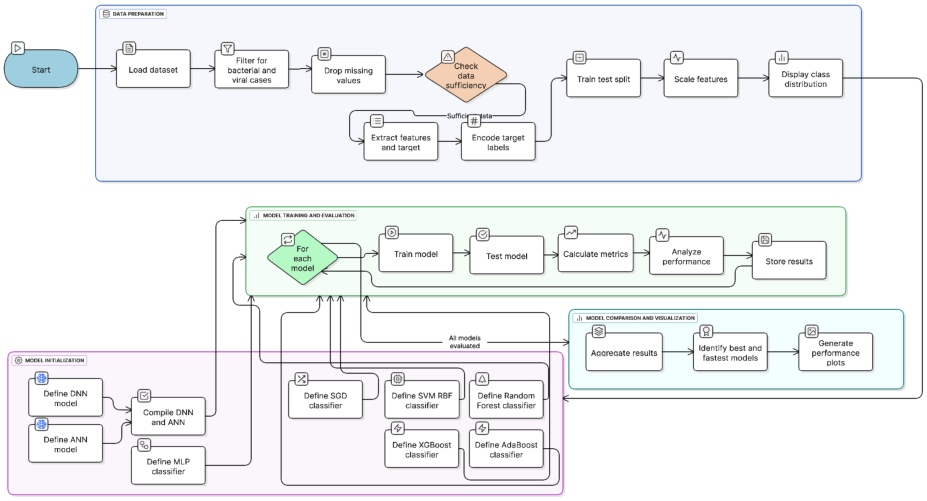

Fig. 1. Methodological Framework for Meningitis Classification using Machine Learning and Deep Learning Approaches

2.1 Dataset Description and Preprocessing

Data Collection and Characteristics The study leveraged the publicly available **Meningitis Classification dataset** [10] from Kaggle, containing 1,200 anonymized patient records. These electronic health records, collected across multiple facilities, feature eight key biomarkers: Age, Gender, WBC Count, Protein Level, Glucose Level, Hemoglobin, WBC Blood Count, Platelets, and CRP Level. The dataset is highly complete with no missing values and shows a **balanced representation** between bacterial and viral meningitis cases (Figure **??**) and comprehensive statistical details are presented in Table 1.

Feature Correlations and Multicollinearity Assessment Correlation analysis revealed clinically relevant relationships between biomarkers (Figure **??**), with strong positive correlations between WBC Count and Protein Level (r=0.70), and strong negative correlations between Hemoglobin and WBC Blood Count (r=-0.85). These patterns align with established pathophysiology, such as co-elevation of inflammatory markers in bacterial infections and anemia in systemic inflammation. While multicollinearity was identified, all features were retained due to their established diagnostic significance in meningitis.

Data Cleaning and Validation Rigorous data cleaning and validation protocols were implemented to ensure dataset integrity. The validation process confirmed complete data quality with no missing values across all 1,200 records and 13 clinical parameters. All laboratory values were verified against established clinical ranges, with statistical outliers identified using interquartile range

Table 1. Dataset Characteristics and Clinical Parameters

Parameter	Value	
	Mean ± SD	**Range**
Demographics		
Age (years)	47.2 ± 26.8	0-118
Gender (Male/Female)	604/529	-
CSF Analysis		
WBC Count (cells/μL)	12,847 ± 5,892	2,008-24,974
Protein (mg/dL)	125.3 ± 78.4	2-299
Glucose (mg/dL)	52.6 ± 39.2	0-148
Hematological Parameters		
Hemoglobin (g/dL)	9.8 ± 5.2	1-18
WBC Blood (cells/μL)	12,156 ± 5,347	4,032-19,998
Platelets ($\times 10^3/\mu$L)	238,451 ± 87,326	100,088-399,479
CRP (mg/L)	33.7 ± 26.9	0-99
Diagnosis		
Bacterial Meningitis	595 (52.5%)	-
Viral Meningitis	538 (47.5%)	-

methods and assessed for clinical relevance. Comprehensive checks ensured no duplicate patient records, consistent data types across parameters, and standardized measurement units. Cross-parameter validation confirmed logical consistency between cerebrospinal fluid profiles and clinical diagnoses, providing a reliable foundation for model development without requiring imputation or extensive preprocessing.

2.2 Hybrid Deep Learning and Machine Learning Models

Neural Network Architectures Three neural network architectures were implemented with varying complexity. The Deep Neural Network (DNN) employed a 128-64-32 layer structure with batch normalization and progressive dropout rates (40%, 30%, 20%) for regularization. The Artificial Neural Network (ANN) used a simplified 64-32 hidden layer design with reduced dropout, while the MLP Classifier implemented a (100, 50) hidden layer architecture. All neural models utilized ReLU activation functions, Adam optimization with learning rate 0.001, and early stopping to prevent overfitting, with the DNN and ANN incorporating sigmoid output activation for binary classification.

Ensemble Methods The ensemble approaches included XGBoost with 100 estimators, maximum depth of 6, and automated missing value handling; Random Forest employing 100 decision trees with feature randomization; and AdaBoost using 100 decision tree stumps with adaptive instance weighting. These methods provided robust performance through different ensemble strategies: gradient boosting in XGBoost, bootstrap aggregation in Random Forest, and sequential reweighting in AdaBoost.

Efficient Linear Models For computational efficiency, SGD Classifier implemented stochastic gradient descent with L2 regularization, offering rapid convergence and minimal memory footprint. The SVM RBF Classifier provided non-linear separation capability through radial basis function kernel with automatic gamma scaling and probability estimation, balancing model complexity with classification performance.

2.3 Experimental Setup

Training and Validation Strategy All models were trained using carefully optimized hyperparameters for fair comparison. Neural networks employed the Adam optimizer (learning rate: 0.001) with binary cross-entropy loss:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

Early stopping (patience: 15 epochs) prevented overfitting. Evaluation used stratified 80-20 train-test split (random state: 42) preserving class distribution between sets.

Evaluation Metrics Comprehensive assessment employed multiple metrics with clinical emphasis on sensitivity for bacterial detection:

- **Accuracy**: $\frac{TP+TN}{TP+TN+FP+FN}$, Overall correctness of the classification model.
- **Sensitivity**: $\frac{TP}{TP+FN}$, Ability to correctly identify true bacterial meningitis cases.
- **Specificity**: $\frac{TN}{TN+FP}$, Ability to correctly identify true viral meningitis cases.
- **F1-Score**: $2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision}+\text{Recall}}$ Balanced measure of precision and recall.
- **ROC AUC**: Area under ROC curve, Overall diagnostic ability across all classification thresholds.
- **Speed**: Inference time efficiency, Practical measure of model deployment feasibility.

3 Results and Analysis

3.1 Overall Performance Comparison

Accuracy Analysis Across Models The comprehensive evaluation of eight machine learning models revealed significant performance variations in meningi-

tis classification. The MLP Classifier achieved the highest accuracy (93.83%), followed closely by XGBoost (93.39%) and DNN (93.39%) as shown in Table 2. All neural network-based approaches consistently outperformed traditional machine learning models, demonstrating the advantage of deep learning architectures in capturing complex patterns from clinical biomarkers. The performance hierarchy established MLP as the optimal choice for accurate meningitis differentiation, while maintaining balanced performance across other clinical metrics.

Training Dynamics, Generalization, and Computational Efficiency Analysis of training dynamics revealed the ANN had the strongest generalization with a minimal train-test gap (1.39%), while tree-based models like XGBoost and Random Forest showed moderate overfitting despite perfect training scores. Neural networks (DNN and ANN) proved to be the most reliable for clinical deployment with stable validation accuracies of 91.76% and 92.31%, respectively. Computational efficiency varied significantly, with the SGD Classifier being the fastest (669.88 speed metric, Table 2), making it suitable for real-time use. Furthermore, feature importance analysis consistently identified CSF WBC count, protein levels, and CRP as the most critical diagnostic features across all models.

Sensitivity, Specificity, Precision, and Recall XGBoost, MLP, and Random Forest achieved highest sensitivity (94.44%), crucial for bacterial meningitis detection. DNN and AdaBoost led specificity (94.96%, 94.12%), excelling at viral identification. MLP provided the optimal clinical balance with high sensitivity (94.44%) and specificity (93.28%), making it most reliable for comprehensive meningitis differentiation.

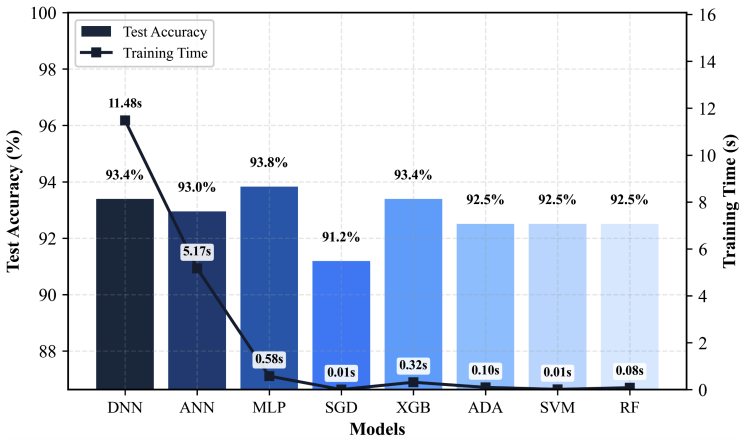

Fig. 2. Model Performance: Accuracy vs. Training Time. MLP achieved highest accuracy (93.83%) with fastest training (0.58s), demonstrating optimal balance for clinical deployment.

Table 2. Comprehensive Model Performance Analysis and Clinical Evaluation Metrics

Model	PERF MET			CLINICAL MET					ERR ANA			COM EFF	
	Acc (%)	AUC	Time (s)	Sen (%)	Spe (%)	Pre (%)	F1 (%)	Spd	Err (%)	FN (%)	FP (%)	Tr Eff	Inf Eff
MLP	93.83	0.987	0.58	94.44	93.28	92.73	93.58	102.61	6.17	5.56	6.72	1.692	107.29
XGB	93.39	0.991	0.32	94.44	92.44	91.89	93.16	340.13	6.61	5.56	7.56	3.162	481.40
SVM	92.51	0.990	0.01	91.67	93.28	92.52	92.09	312.42	7.49	8.33	6.72	74.260	420.34
RF	92.51	0.990	0.08	94.44	90.76	90.27	92.33	148.17	7.49	5.56	9.24	11.944	160.92
ADA	92.51	0.986	0.10	90.74	94.12	93.33	92.02	120.27	7.49	9.26	5.88	9.911	126.47

Abbreviations:
PERF MET: Performance Metrics, CLINICAL MET: Clinical Metrics, ERR ANA: Error Analysis, COM EFF: Computational Efficiency. **Metrics:** Acc: Accuracy, AUC: Area Under ROC Curve, Sen: Sensitivity, Spe: Specificity, Pre: Precision, Spd: Speed (1/time), Err: Error Rate, FN: False Negative Rate, FP: False Positive Rate. **Efficiency:** Tr Eff: Training Efficiency (accuracy/s), Inf Eff: Inference Efficiency (accuracy/s). **Models:** MLP: Multi-Layer Perceptron, XGB: XGBoost, SVM: Support Vector Machine, RF: Random Forest, ADA: AdaBoost.

3.2 Comprehensive Model Evaluation

MLP Classifier demonstrated the most balanced performance with only 14 misclassifications out of 227 samples. Neural networks maintained balanced error distributions, while ensemble methods showed specialized strengths: ANN and AdaBoost minimized false positives to reduce unnecessary treatments, while XGBoost and Random Forest excelled at minimizing false negatives crucial for preventing missed diagnoses. ROC analysis revealed exceptional discriminatory power with XGBoost achieving the highest AUC (0.9907), and all models maintained AUC scores above 0.94, confirming strong classification capability across decision thresholds.

3.3 Conclusion and Future Work

This study presents a comprehensive comparative analysis of machine learning models for meningitis classification. The MLP Classifier emerged as the top performer, achieving the highest accuracy (**93.83%**) and demonstrating the most balanced clinical performance across all evaluation metrics. Neural network architectures consistently showed superior diagnostic capability, while ensemble methods like XGBoost and Random Forest provided strong performance with optimal computational efficiency. The results demonstrate that machine learning approaches can achieve high classification accuracy using standard clinical biomarkers, with XGBoost attaining the highest ROC AUC (0.991) and MLP providing the best overall balance between sensitivity (94.44%) and specificity (93.28%). Future work will focus on enhancing model interpretability through Explainable AI (XAI) techniques, particularly SHAP and LIME, to provide transparent feature importance and individual prediction explanations. This integration is crucial for building clinical trust and facilitating physician acceptance of AI-assisted diagnostics. Additional research directions include validating the models on larger multi-center datasets, exploring temporal patterns in

patient data for progressive diagnosis, and investigating transfer learning approaches to adapt the models for different healthcare settings and patient populations.

References

1. M. R. Foundation, "90% of children and teenagers who die of meningococcal meningitis die within 24 hours," *Meningitis Research Foundation*, January 2020, accessed: 2024. [Online]. Available: https://www.meningitis.org/research-into-how-fast-meningitis-kills
2. T. M. van Soest, N. Chekrouni, N. M. van Sorge, M. W. Bijlsma, M. C. Brouwer, and D. van de Beek, "Epidemiology, clinical features and outcome of adults with meningococcal meningitis: a 15-year prospective nationwide cohort study," *The Lancet Regional Health–Europe*, vol. 30, 2023.
3. B. K. Choi, Y. J. Choi, M. Sung, W. Ha, M. K. Chu, W.-J. Kim, K. Heo, K. M. Kim, and Y. R. Park, "Development and validation of an artificial intelligence model for the early classification of the aetiology of meningitis and encephalitis: a retrospective observational study," *EClinicalMedicine*, vol. 61, 2023.
4. K. Ghaddaripouri, M. Ghaddaripouri, A. S. Mousavi, S. F. Mousavi Baigi, M. Rezaei Sarsari, F. Dahmardeh Kemmak, and M. R. Mazaheri Habibi, "The effect of machine learning algorithms in the prediction, and diagnosis of meningitis: A systematic review," *Health Science Reports*, vol. 7, no. 2, p. e1893, 2024.
5. M. Amanullah, P. Devi *et al.*, "High-performance meningitis etiology prediction: Leveraging csf biomarkers via a stacked ensemble of random forest and artificial neural networks," in *2025 9th International Conference on Inventive Systems and Control (ICISC)*. IEEE, 2025, pp. 1226–1232.
6. R. Caragheorgheopol, C. Ţucureanu, V. Lazăr, S. A. Florescu, D. S. Lazăr, and I. Caraş, "Cerebrospinal fluid cytokines and chemokines exhibit distinct profiles in bacterial meningitis and viral meningitis," *Experimental and Therapeutic Medicine*, vol. 25, no. 5, p. 204, 2023.
7. A. Messai, A. Drif, A. Ouyahia, M. Guechi, M. Rais, L. Kaderali, and H. Cherifi, "Transparent ai models for meningococcal meningitis diagnosis: Evaluating interpretability and performance metrics," in *2024 IEEE 12th International Conference on Intelligent Systems (IS)*. IEEE, 2024, pp. 1–8.
8. S. S. Priya, M. Amanullah, and T. Kujani, "Enhancing meningitis type diagnosis through a novel hybrid ensemble of bagging and gradient boosting," in *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*. IEEE, 2024, pp. 1260–1264.
9. V. M. Lelis, E. Guzman, and M.-V. Belmonte, "Non-invasive meningitis diagnosis using decision trees," *IEEE Access*, vol. 8, pp. 18 394–18 407, 2020.
10. Kaggle Community, "Meningitis classification dataset," 2024, accessed: 2024-10-20. [Online]. Available: https://www.kaggle.com/datasets/chanteep/meningitis-classification