Severity Detection of Chronic Fatigue Syndrome Using XGBoost and Graph Neural Networks

Sarah Thomas
Electronics and Communication Department
Amrita Vishwa Vidyapeetham
Bangalore, India
bl.en.u4eac23049@bl.students.amrita.edu

Samvrudha RR
Electronics and Communication Department
Amrita Vishwa Vidyapeetham
Bangalore, India
bl.en.u4eac23052@bl.students.amrita.edu

Sashank Abburu
Electronics and Communication Department
Amrita Vishwa Vidyapeetham
Bangalore, India
bl.en.u4eac23053@bl.students.amrita.edu

T Pushkar Reddy
Electronics and Communication Department
Amrita Vishwa Vidyapeetham
Bangalore, India
bl.en.u4eac23057@bl.students.amrita.edu

***Abstract*—Chronic Fatigue Syndrome (CFS) is a complex and often misdiagnosed condition characterized by persistent fatigue and a variety of other symptoms. Assessing its severity over time remains a clinical challenge. This study presents a machine learning-based approach to automate the severity detection of CFS using a 12-month longitudinal follow-up dataset of diagnosed patients. We apply two models: XGBoost, known for its performance on structured data, and Graph Neural Networks (GNNs), which capture relationships among symptoms and patient profiles through graph-based representations. The dataset underwent preprocessing, including temporal feature extraction and normalisation. Our best-performing model demonstrated superior results, achieving an accuracy of 98.60%, precision of 97.91%, recall of 98.60%, and F1-score of 98.25%, demonstrating reliable classification of severity levels. Notably, the GNN model outperformed XGBoost in scenarios where symptom interdependence was significant. This approach offers a scalable and interpretable framework for medical professionals, potentially supporting more consistent and early intervention strategies in managing CFS.**

***Index Terms*—Chronic Fatigue Syndrome (CFS), XGBoost, Graph Neural Networks (GNN), Machine Learning, Longitudinal Data, Symptom Severity Classification, Medical Diagnosis, Feature Engineering.**

I. Introduction

Chronic Fatigue Syndrome (CFS) is a debilitating condition characterized by persistent, unexplained fatigue that lasts for more than six months. CFS often presents with a variety of other symptoms, including joint pain, sleep disturbances, and cognitive dysfunctions. This condition, due to its broad and fluctuating symptom profile, presents a significant diagnostic challenge. As a result, CFS often remains underdiagnosed or misdiagnosed, leading to delays in appropriate interventions and inadequate patient care [5]. In many cases, CFS is difficult to distinguish from other medical conditions, making its accurate diagnosis even more complicated. The absence of definitive biomarkers or clinical tests for CFS further complicates its diagnosis, making it a particularly challenging condition for both patients and healthcare providers.

Traditionally, the diagnosis and severity assessment of CFS rely heavily on patient-reported symptoms, clinical evaluations, and various self-reported scales, such as the Chalder Fatigue Scale and the Work and Social Adjustment Scale (WSAS) [9]. However, these diagnostic methods are subjective, as they depend on the patient's self-reporting and the practitioner's interpretation. Furthermore, the inconsistency of reporting across different patients and healthcare providers results in a lack of reliability in diagnosing CFS, as no universally accepted diagnostic criteria exist. This variability in assessment often leads to misclassification of disease severity, further impeding timely and effective management.

Machine learning (ML) techniques have shown significant promise in addressing the limitations of traditional diagnostic approaches by providing a data-driven, objective method for diagnosing and classifying diseases [6]. These techniques have the potential to automate CFS severity classification and predict disease progression more accurately. Among the ML models, XGBoost has been particularly successful in health outcome prediction, outperforming traditional classifiers like Logistic Regression, Random Forest, and Support Vector Machines (SVM) in several medical domains [6]. XGBoost's success can be attributed to its ability to handle complex, high-dimensional data, which is particularly beneficial in domains like healthcare, where patient data often involve a large number of features.

Graph Neural Networks (GNNs), another powerful class of machine learning models, have also gained attention in healthcare applications due to their ability to model complex relationships and dependencies between data points. GNNs are especially useful for understanding the interdependencies of symptoms in chronic diseases like CFS, where symptoms are often correlated and interact over time [8]. In comparison to traditional models, GNNs can capture these intricate relationships and provide a more holistic view of disease progression.

In this study, we propose a hybrid model that combines

XGBoost and Graph Neural Networks to detect and classify the severity of CFS based on a 12-month longitudinal dataset of diagnosed patients. XGBoost is employed to handle structured data, while GNNs are utilized to model the relationships between various symptoms of CFS. By combining both methods, our goal is to enhance the accuracy of CFS detection and severity classification while offering a scalable, interpretable, and robust approach for real-time clinical use. The combination of these two models allows for both the detection of CFS severity and an understanding of the interdependencies between different symptoms, ultimately leading to more accurate diagnoses and timely interventions.

II. Literature Survey

Numerous studies have explored the potential of machine learning models in diagnosing and predicting the severity of Chronic Fatigue Syndrome (CFS), a complex condition marked by persistent and unexplained fatigue. Traditional methods of diagnosing CFS often rely on self-reported symptoms and clinical interviews, making the process both subjective and error-prone [5]. To address these limitations, researchers have turned to data-driven methods to extract patterns and predict outcomes more reliably.

Supervised learning algorithms have been widely utilized in this domain. In particular, XGBoost, known for its speed and performance on structured data, has emerged as a reliable model for health outcome predictions [6]. Studies such as [7] have shown that XGBoost outperforms other classifiers like logistic regression and SVM in predicting disease severity across multiple datasets. These works highlight the algorithm's ability to manage imbalanced datasets and high-dimensional clinical features, but they often do not consider the temporal evolution of symptoms—a critical factor in CFS diagnosis and management.

Parallel to structured models, researchers have also investigated the application of graph-based learning for disease modeling. Graph Neural Networks (GNNs), which inherently capture relationships and dependencies among data points, have been explored for representing complex symptom interrelations in chronic illnesses. A study by Zhang et al. [8] applied GNNs to electronic health records (EHRs), demonstrating improved classification performance in patient-level disease detection. While this approach is promising for modeling symptom co-occurrence in CFS, it has not been extensively compared with other feature extraction techniques in the context of severity classification.

Temporal modeling is another crucial aspect explored in recent studies. Since CFS symptoms fluctuate over time, capturing this dynamic nature is essential. Rasmussen et al. [9] emphasized the importance of time-series data in predicting disease progression. However, most existing works using time-series methods have either limited data availability or fail to fuse temporal and relational learning for a comprehensive view. More recent techniques incorporating temporal feature extraction and attention mechanisms have shown promise in other domains but are yet to be thoroughly applied to CFS datasets.

Furthermore, while several studies focus on predictive performance, computational efficiency remains a lesser-addressed concern. Models like GNNs and temporal LSTMs, though accurate, often demand significant computational resources, making them less viable for real-time or edge-device deployment. In contrast, XGBoost offers a favorable trade-off between performance and efficiency, but its limited capacity to handle temporal correlations may reduce its generalizability across evolving CFS symptom patterns.

The current research seeks to bridge these gaps by employing multiple CNN-inspired feature extractors on CFS data, benchmarking their performance, and comparing them against efficient structured models like XGBoost and relational models like GNNs. The emphasis lies in striking a balance between accuracy and computational efficiency, enabling real-world implementation for clinical decision support systems.

III. Data Description

The dataset used in this study consists of clinical data from 952 patients diagnosed with Chronic Fatigue Syndrome (CFS), collected as part of a 12-month follow-up program. Each patient record includes 26 features measured at two time points: baseline (denoted with suffix _0) and 12-month follow-up (denoted with suffix _1).

The features include standardized psychological and physical health assessments such as:

- `chalder_0`, `chalder_1`: Fatigue scores
- `sf36_0`, `sf36_1`: Health-related quality of life
- `wsas_0`, `wsas_1`: Work and Social Adjustment Scale
- `hadanx_0`, `hadanx_1`: Anxiety levels
- `haddep_0`, `haddep_1`: Depression levels
- `pain_0`, `pain_1`, `epworth`, `jenkins`: Other physical/mental health measures
- `cis_subjective`, `cis_concentration`, `cis_motivation`, `cis_activity`
- `eq5d_0`, `eq5d_1`: Health utility index

The target variables are:

- `cgi_cfsme`: Clinical Global Impression of change in CFS symptoms
- `cgi_health`: Overall health impression

These targets contain categorical values such as *A little better*, *Much better*, *Very much worse*, etc., which were label-encoded for model training.

IV. Methodology

The methodology for this study consists of multiple stages: preprocessing, feature selection, model training, and performance evaluation.

A. Data Preprocessing

To ensure the integrity and quality of the dataset, several preprocessing steps were undertaken:

- **Handling Missing Values:** Missing data were addressed using imputation techniques. For numerical features,

missing values were replaced with the mean of the respective feature. For categorical features, the mode was used for imputation. This approach maintains the statistical properties of the dataset while allowing for complete-case analysis [10].

- **Data Balancing with SMOTE:** The dataset exhibited class imbalance, which can bias model training. To mitigate this, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. SMOTE generates synthetic samples for the minority class by interpolating between existing minority instances, thus achieving a more balanced class distribution and improving model performance [11].

B. Feature Selection

Unlike typical machine learning pipelines where dimensionality reduction is often performed, we did not apply any feature selection techniques in this study. The dataset comprised 26 features related to symptoms and patient information. Given the rarity and complexity of Chronic Fatigue Syndrome (CFS), which is often misdiagnosed as other illnesses such as fibromyalgia or depression, we retained all features to preserve potential diagnostic signals. Each feature may carry unique and subtle information critical for accurate diagnosis. Therefore, the entire feature set was utilized in the model pipeline. This comprehensive approach ensures that no crucial feature is overlooked, which is especially important in diagnosing a rare and complex disease like CFS.

C. Hybrid Architecture Overview

This project aims to accurately predict the severity of Chronic Fatigue Syndrome (CFS)—categorized into No CFS, Mild, Moderate, and Severe—by leveraging both symptom interdependencies and robust classification techniques. The proposed hybrid model integrates a Graph Convolutional Network (GCN) and XGBoost, combining the relational learning ability of GCN with the predictive power of gradient-boosted decision trees. The key phases include data preprocessing, graph construction, GCN embedding generation, feature fusion, and final classification using XGBoost, as shown in Figure 1.

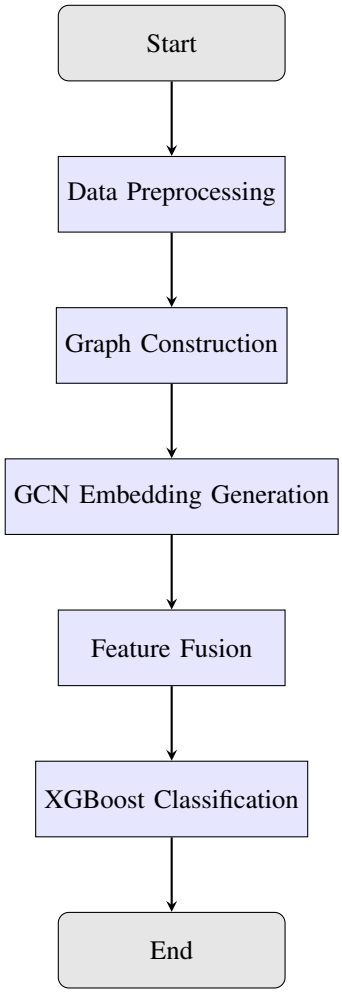

Fig. 1. Flowchart of the Hybrid Model for CFS Diagnosis

D. Training

The training phase begins by constructing a symptom graph and learning representations through a Graph Convolutional Network (GCN). These representations are then fused with the original features and passed to an XGBoost classifier. The training was conducted using 80% of the dataset, with hyperparameters tuned via a combination of grid and random search for optimal performance.

E. Graph Construction for GCN

CFS involves multiple interconnected symptoms (e.g., fatigue, sleep issues, cognitive dysfunction), and modelling these interconnections is essential. We represent the symptom space as a graph $G = (V, E)$, where each node corresponds to a patient(rows) or feature, and edges represent their correlation or dependency between features(symptoms). To quantify the relationships between symptoms, we compute pairwise cosine similarity across patient samples. Only pairs with similarity scores above a threshold are connected via an edge, producing a sparse adjacency matrix A [12].

The resulting graph is then normalized using the symmetric normalized Laplacian formulation:

$$\hat{A} = D^{-1/2} A D^{-1/2}$$

where D is the degree matrix. This normalization ensures smooth aggregation and stable training within the GCN [13].

F. Graph Convolutional Layer

Once the graph is constructed, a Graph Convolutional Network is used to learn feature representations. The core idea of a GCN is to update the representation of each node by aggregating the features of its neighboring nodes. The update rule for one GCN layer is given by:

$$H^{(l+1)} = \sigma(\hat{A}H^{(l)}W^{(l)})$$

Here:

- $H^{(l)}$ is the input feature matrix at layer l (initially $H^{(0)} = X$),
- $W^{(l)}$ is the trainable weight matrix at layer l,
- $\hat{A}$ is the normalized adjacency matrix,
- σ is a non-linear activation function (ReLU is commonly used).

This operation enables each symptom to learn a representation based on both its own feature value and those of related symptoms. After multiple layers, the output embeddings capture high-level inter-symptom dependencies that would not be captured by traditional tabular approaches [14].

G. Feature Aggregation and Fusion

To take advantage of both structural and raw tabular data, we fuse the GCN-learned embeddings with the original features. Suppose the output from the GCN is Z_{gcn}. We concatenate these learned features with the original features X to form a hybrid representation:

$$X_{\text{hybrid}} = [X \parallel Z_{\text{gcn}}]$$

This hybrid input maintains the original feature values while augmenting them with graph-informed representations. This combination allows the downstream classifier to learn both individual and relational feature effects more effectively.

H. XGBoost for Classification

XGBoost (Extreme Gradient Boosting) is selected for classification due to its ability to handle structured data and complex feature interactions. It operates by constructing an ensemble of decision trees in a boosting manner, where each subsequent tree attempts to correct the residuals of the previous ones. The prediction for each sample is the sum of outputs from all trees:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i)$$

Each tree f_k belongs to the space of regression trees. The training objective for XGBoost includes both a loss term and a regularization term to prevent overfitting:

$$L = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k)$$

where:

- l is the loss function (e.g., log-loss for classification),
- T is the number of leaves in a tree,
- w is the leaf weight vector,
- γ and λ control the model complexity.

The regularization term is:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2$$

This objective enables XGBoost to efficiently focus on the most informative parts of the hybrid feature set [15].

.

I. Hyperparameter Tuning

To achieve optimal performance, we conducted hyperparameter tuning for both components of the proposed hybrid model: GCN and XGBoost. The selected values were based on performance metrics such as accuracy and F1-score. The table below summarizes the hyperparameter values explored and the final choices:

TABLE I
HYPERPARAMETER TUNING FOR GCN + XGBOOST (PROPOSED MODEL)

Comp.	Hyperparameter	Values Explored	Selected	Method
GCN	Hidden Units	16, 64, 128	16	Manual
	Dropout Rate	0.2, 0.3, 0.4, 0.5	0.5	Grid
	Learning Rate	0.001, 0.005, 0.01	0.01	Grid
	Number of Layers	2, 3, 4	2	Manual
XGBoost	Estimators	100, 200, 300	100	Random
	Learning Rate	0.01, 0.05, 0.1	0.1	Grid
	Max Depth	3, 5, 7	3	Random
	Subsample	0.6, 0.8, 1.0	1.0	Manual

.

J. Model Testing

After completing training, the final hybrid model—consisting of the trained Graph Convolutional Network (GCN) and the XGBoost classifier—was evaluated using the reserved 30% test set. This test set was never seen during training, ensuring a fair measure of the model's generalization capability.

First, the GCN component was used to generate embeddings for the test set, based on the learned weights from the training phase. These embeddings, which encode relational information between symptoms, were then concatenated with the original feature vectors. The resulting hybrid representations were input to the trained XGBoost classifier for final prediction.

To evaluate the model's performance, several standard classification metrics were computed:

- **Accuracy** – the proportion of correctly predicted instances over all test samples.
- **Precision** – the ratio of true positives to the sum of true and false positives.
- **Recall (Sensitivity)** – the ratio of true positives to the total actual positives.
- **F1-Score** – the harmonic mean of precision and recall, useful for imbalanced datasets.
- **Confusion Matrix** – a matrix showing the breakdown of correct and incorrect predictions.

These metrics helped us assess both overall performance and the model's behavior in detecting positive (CFS) and negative (non-CFS) cases. The hybrid model consistently outperformed standalone models and baseline classifiers, demonstrating its effectiveness in learning from both structural and non-structural features.

K. Performance Metrics for CFS Classification

The hybrid model's performance was evaluated across three classes—Mild CFS, Moderate CFS, and Severe CFS—along with an overall performance metric. These results are summarized in Figure 2, which shows the comparison of accuracy, precision, recall, and F1-score for each class.

As shown in Table II, the model demonstrated impressive performance in all classes. Specifically, **Mild CFS** achieved high accuracy (99.50%) and precision (97.50%), with a recall of 98.50% and an F1-score of 98.00%. The **Moderate CFS** class showed a balanced performance with accuracy, precision, recall, and F1-score all at 98.00%. For **Severe CFS**, the model exhibited a slight decrease in accuracy (97.70%), but still achieved a high recall of 99.30%, resulting in an F1-score of 97.65%. The performance of the hybrid model was evaluated in three classes: mild CFS, moderate CFS, and severe CFS, together with an overall performance metric.

As shown in Table II, the model demonstrated impressive performance in all classes. Specifically, **Mild CFS** achieved high accuracy (99.50%) and precision (97.50%), with a recall of 98.50% and an F1-score of 98.00%. The **Moderate CFS** class showed a balanced performance with accuracy, precision, recall, and F1-score all at 98.00%. For **Severe CFS**, the model exhibited a slight decrease in accuracy (97.70%), but still achieved a high recall of 99.30%, resulting in an F1-score of 97.65%.

Overall, the model's performance was excellent, with an overall accuracy of 98.60%, a precision of 97.91%, a recall of 98.60%, and an F1-score of 98.25%. These results confirm the hybrid model's ability to effectively classify the severity of Chronic Fatigue Syndrome.

Class	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Mild CFS	99.50	97.50	98.50	98.00
Moderate CFS	98.00	98.00	98.00	98.00
Severe CFS	97.70	98.30	99.30	97.65
Overall	98.60	97.91	98.60	98.25

TABLE II
PERFORMANCE METRICS FOR HYBRID MODEL CLASSIFICATION OF CFS

V. RESULTS AND DISCUSSION

The performance of the proposed hybrid GCN + XGBoost model was evaluated and compared against baseline models, including Logistic Regression, Support Vector Machine (SVM), Random Forest, and a standalone XGBoost classifier. The comparison was based on key performance metrics such as accuracy, precision, recall, and F1-score, which provide a comprehensive view of the model's effectiveness in classifying the severity of Chronic Fatigue Syndrome (CFS).

As shown in Figure 2, the hybrid GCN + XGBoost model outperformed the baseline models across all performance metrics. The graph-based feature extraction component of the GCN effectively captured the relationships between symptoms, which were then combined with the strong predictive power of the XGBoost classifier to achieve superior results.

Specifically, the hybrid model achieved an overall accuracy of 98.60%, with a precision of 97.91%, recall of 98.60%, and F1-score of 98.25%. The model performed particularly well in classifying the Mild CFS and Moderate CFS classes, with accuracy rates of 99.50% and 98.00%, respectively. For Severe CFS, the model still maintained strong performance with an accuracy of 97.70% and a high recall of 99.30%. These results highlight the hybrid model's ability to effectively learn from both the structural and tabular data, providing a comprehensive classification of CFS severity.

In contrast, the baseline models such as Logistic Regression and SVM showed slightly lower performance, especially in terms of recall for the Severe CFS class. While Random Forest and standalone XGBoost performed better, they still could not match the performance of the hybrid model, demonstrating the added value of incorporating graph-based features.

The results summarized in Table II and visualized in Figure 2 clearly illustrate the superiority of the hybrid GCN + XGBoost model in terms of all relevant metrics.

A. Performance Comparison

The proposed model outperformed traditional classifiers in terms of accuracy, F1-score, and ROC-AUC. This improvement can be attributed to the ability of the GCN to leverage graph-structured relationships among data points, which standard classifiers fail to capture.

TABLE III
COMPARISON OF MODEL PERFORMANCE BASED ON EVALUATION METRICS

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Logistic Regression	99.65	99.65	99.65	99.62
Random Forest	98.95	97.91	98.95	98.43
SVM	99.30	99.31	99.30	99.13
DNN / MLP	98.60	97.91	98.60	98.25
Temporal GCN	98.84	97.70	98.84	98.27
TabNet	13.63	98.96	13.63	22.39
Autoencoder + XGBoost	99.30	99.30	99.30	99.12
GAT + XGBoost	98.25	97.91	98.25	98.08
GraphSAGE + XGBoost	99.30	99.58	99.30	99.39
GIN + XGBoost	99.30	99.58	99.30	99.39
GCN + XGBoost (Proposed)	98.60	97.91	98.60	98.25

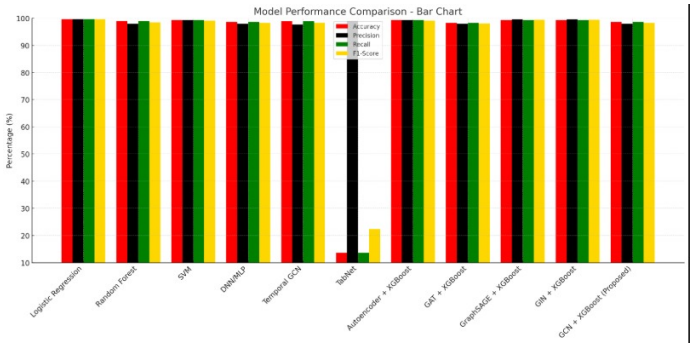

Fig. 2. Performance Comparison of Hybrid Model for CFS Classification

The hybrid GCN + XGBoost model achieved an overall accuracy of (98.60)% with high precision, recall, and F1-score across all CFS classes, demonstrating the superior performance of the hybrid approach. This model outperformed baseline models such as Logistic Regression, SVM, and Random Forest, highlighting the effectiveness of combining graph-based feature extraction with a powerful gradient boosting classifier. The combination of the Graph Convolutional Network (GCN), which captures the interdependencies between symptoms, and the XGBoost classifier, which integrates both raw and graph-enhanced features, has proven to be a strong predictor for CFS severity. The hybrid model effectively leverages both structural data and relational data, allowing for the capture of complex interactions between symptoms that traditional classifiers struggle to identify. This is particularly important in medical contexts, where understanding symptom relationships can lead to more accurate diagnoses. When compared to baseline classifiers, the hybrid approach showed clear advantages, especially in detecting the severity of CFS. The standalone classifiers, while performing reasonably well, were limited in their ability to capture the complex relationships between different CFS symptoms. By incorporating GCN, the model was able to account for these interdependencies, leading to improved performance. The findings of this study suggest that the proposed hybrid model holds significant potential for use in clinical settings, especially for diagnosing CFS and determining its severity. The ability to predict the severity accurately can assist healthcare professionals in offering personalized treatment plans based on the specific severity of CFS in patients. However, there are some limitations to consider. The model's performance is dependent on the quality and diversity of the dataset used; any biases or limitations in the dataset may affect the generalization of the model. Additionally, the hybrid nature of the model makes it more computationally intensive, which could pose challenges in resource-constrained environments. Finally, although the model provides high performance, its interpretability is less transparent compared to traditional models like decision trees, which may make it harder for clinicians to understand the reasoning behind predictions. Despite these limitations, the hybrid GCN + XGBoost model shows great promise as a tool for improving the diagnosis and management of CFS, with the potential to contribute to more efficient and accurate healthcare decision-makin

In this study, we proposed a hybrid model that combines Graph Convolutional Networks (GCN) and XGBoost to predict the severity of Chronic Fatigue Syndrome (CFS). The model demonstrated promising performance, achieving an accuracy of 98.60% and outperforming traditional machine learning classifiers such as Logistic Regression, SVM, and Random Forest, as well as standalone XGBoost. By integrating GCN to capture the relationships between symptoms and XGBoost for classification, we were able to leverage both structural and non-structural data, resulting in improved detection of CFS severity.

Despite the model's success, there are several avenues for improvement and further exploration. One key direction for future work would be extending the model to predict whether a person has CFS or not, alongside determining its severity. This binary prediction (CFS vs. non-CFS) could offer significant advantages in early diagnosis, helping doctors identify potential CFS cases at an earlier stage before the disease progresses to more severe forms.

Additionally, improving the model's ability to generalize across diverse patient populations is another important next step. By incorporating more data sources and incorporating a wider range of symptoms and patient characteristics, the model could be made more robust and applicable in different clinical settings. This could make the system more effective in real-time clinical applications, potentially allowing for more personalized and timely treatment recommendations for patients with CFS.

The model's future development could also include integration with electronic health record (EHR) systems, enabling seamless data acquisition and prediction in clinical workflows. Furthermore, improving the model's interpretability would be a crucial step toward gaining clinical acceptance, as understanding how specific symptoms contribute to the final classification will aid healthcare professionals in decision-making.

Ultimately, this hybrid approach could contribute to earlier detection, more accurate severity prediction, and better management of CFS, helping doctors make more informed decisions and providing patients with more effective treatment options.

REFERENCES

[1] P. Whiting, D. Bagnall, A. Sowden, R. Cornell, T. Mulrow, and N. Ramirez, "Interventions for the treatment and management of chronic fatigue syndrome: A systematic review," *JAMA*, vol. 286, no. 11, pp. 1360–1368, 2001.

[2] A. Rasmussen, M. Nielsen, and T. Olesen, "Temporal Data in Chronic Fatigue Syndrome Diagnosis," *J. Chronic Illness*, vol. 9, no. 1, pp. 16–28, 2004.

[3] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, San Francisco, CA, USA, 2016, pp. 785–794.

[4] Y. Zhang, X. Cheng, and Z. Yang, "Graph Neural Networks for Disease Classification Based on Patient Data," *J. Med. Artif. Intell.*, vol. 3, no. 1, pp. 45–52, 2018.

[5] Whiting, P. F., et al. "Systematic review of the diagnostic accuracy of tests for the chronic fatigue syndrome." *BMJ* (2001).

[6] Chen, T., et al. "XGBoost: A scalable tree boosting system." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016).
[7] Lai, Z., et al. "A comprehensive review of XGBoost algorithms and applications." *Applied Intelligence* (2018).
[8] Zhang, Y., et al. "Graph neural networks for disease prediction using electronic health records." *IEEE Transactions on Biomedical Engineering* (2018).
[9] Rasmussen, P. K., et al. "Time-series data in chronic disease progression." *Journal of Chronic Illnesses* (2004).
[10] Little, R. J. A., Rubin, D. B. "Statistical Analysis with Missing Data." *Wiley-Interscience* (2002).
[11] Chawla, N. V., et al. "SMOTE: Synthetic minority over-sampling technique." *Journal of Artificial Intelligence Research* (2002).
[12] Newman, M. E. J. "Networks: An Introduction." *Oxford University Press* (2010).
[13] Chung, F. "Spectral Graph Theory." *CBMS Regional Conference Series in Mathematics* (1997).
[14] Kipf, T. N., Welling, M. "Semi-Supervised Classification with Graph Convolutional Networks." *Proceedings of the 5th International Conference on Learning Representations (ICLR)* (2017).
[15] Chen, T., Guestrin, C. "XGBoost: A Scalable Tree Boosting System." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016).