IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

A Novel Deep Learning Gut Microbial Analysis for Biomarker Identification and Diagnosis of IBS

Anav Gupta[1], Nirupma Singh[2]

1) The Shri Ram School Moulsari, V37 Moulsari Avenue, Gurugram, Haryana, 122002, India.
2) Department of Biological Sciences and Engineering, Netaji Subhas University of Technology (Formerly NSIT), Dwarka, Delhi, India.; nirupmajadaun21@gmail.com

ABSTRACT: Irritable Bowel Syndrome (IBS) is a widespread functional disorder of the gastrointestinal tract (GI), characterized by abdominal pain and altered bowel habits. Traditional IBS diagnosis is done through self-reported questionnaires characterizing frequencies and severities of associated symptoms. The gut microbiome has shown a significant correlation with various IBS symptoms, and treatments targeting it have been effective in mitigating its symptoms. Previous studies have correlated the gut microbiome with a single label of IBS, whereas it has multiple subtypes. In this study, machine learning (ML) and deep learning (DL) models have been developed to correlate gut microbiome with five IBS symptoms: constipation, acidity, diarrhea, bloating, and burping. Using data from 1100 patients, Graph Neural Networks (GNNs) outperformed traditional ML models by ~15%, while Feed Forward Neural Networks (FFNNs) showed ~20% improvement, achieving 85-90% accuracy for symptom severities and 80-85% for frequencies. Permutation importance was used to compute the important features according to the model, along with Pearson's correlation analysis to identify the direction in which the features varied with the output. Taxon 1737404 (*Murdochiella vaginalis*), 768507 (*Runella slithyformis*), and 1760 (*Actinomyces israelii*) had the highest permutation scores with a positive correlation, while 1236 (*Escherichia coli*) and 186826 (*Enterococcus faecium*) had the strongest permutation scores with a negative correlation. Two web applications were developed for the model, one of which allows other clinicians to upload their datasets, and the other, which returns predictions based on uploaded gut tests. Thus, this study demonstrates the potential of deep learning to leverage gut microbiome data for the accurate prediction of IBS symptoms, along with identifying essential biomarkers.

KEYWORDS: Computational Biology and Bioinformatics, Computational Biomodelling, Irritable Bowel Syndrome, Gut Microbiome Biomarker identification and diagnosis.

Introduction

IBS is a collection of long-term digestive conditions that are prevalent in the GI tract.[1] One in every eleven people globally and one in every fourteen people in India suffer from a set of conditions known as IBS.[2] IBS significantly impacts quality of life, causing chronic abdominal pain, irregular bowel movements, and psychological distress that can affect work productivity and social relationships. Existing literature has shown an imbalance in the gut microbiome, or "gut dysbiosis", to have a strong correlation with IBS.[3] Many therapies that tailor to the gut microbiome have shown a significant improvement in overall IBS Symptoms.[4] The current diagnosis of IBS is based on the Rome IV criteria, which involves clinicians studying self-reported severities and frequencies of various IBS-related symptoms from patients and using that data to diagnose them with IBS.[5] Some of these symptoms include bloating, acidity, constipation, diarrhea, and burping. As per the literature review, there are no reliable biomarkers for IBS.

Research has shown an exponential growth in microbial data, outlining the need for ML.[6] ML models have been developed by correlating the gut microbiome with a single IBS output, showing significant correlations.[7] ML and DL have also been used extensively in microbiome research, with studies showing DL to be more powerful due to its ability to capture microbial dynamics.[8] This research study aims to develop a software and a platform to perform IBS diagnosis based on information gained from the gut microbiome, using symptom-based classification of symptoms. Traditional IBS models, although rare, are trained on a binary IBS output column. This study aims to widen the scope of IBS diagnosis by training the model using the various IBS symptoms (their self-reported severities and frequencies). Testing for severities and frequencies could aid the clinical landscape for IBS, as clinicians can identify targeted biomarkers for each symptom, and subjectivity in patient responses can be eliminated.

Previous microbiome-IBS studies have trained models based on Caucasian and American datasets (such as the European Nucleotide Archive and the American Gut Project). [9, 10] Since the gut microbiome is influenced by the environment, a South Asian/ Indian microbial profile is likely very different from an American/ Caucasian profile. Thus, in this study the relative abundance data of the taxon in the gut microbiome for 1089 Indian patients was gathered with a questionnaire containing how badly (on a scale of 1-10), symptoms such as bloating, acidity, constipation, diarrhea, and burping affected their daily lives and how frequently (per week) these symptoms affected them). Several ML models, including Random Forest (RF) and Adaptive Boosting (AB), were run to build the models.

DOI: 10.36838/v7i8.10

A Graph Convolutional Network (GCN) was also developed, followed by FFNNs, which combine the GNN's ability to effectively capture the taxa in the gut with a simpler architecture, delivering a significantly higher accuracy (of around 85-90% for severities and 80-85% for frequencies). Feature importances were also identified. Overall, this work represents a significant step towards objective, microbiome-based IBS diagnosis and targeted treatments while addressing the critical gap in population-specific research for the Indian subcontinent.

Methods

Data Collection:

1089 Gut Microbiome Tests (GMTs) were sourced from SOVA Health. Microbiome data were derived from 16S rRNA sequencing and provided as relative abundance profiles. These GMTs were collected anonymously, with a unique identifier for each patient, adhering to the ethics of data privacy. Each of these GMTs contained the relative abundances of 31258 taxon and a self-reported questionnaire by patients based on the Rome IV criteria of IBS regarding the severities and frequencies of various IBS symptoms. All the data was anonymized and collected in compliance with data privacy regulations.

Model Building:

Ten models of each type were trained for different symptom severities and frequencies. RF and AB ML models were developed and compared. For the deep learning approach, a GNN was built using Torch Geometric. The network architecture consisted of three Graph Convolutional Network (GCN) layers followed by global pooling and fully connected layers. The algorithm constructs a weighted graph representing relationships between microbiome features based on their co-occurrence patterns in the dataset. Initially, it computes individual feature support values by calculating the proportion of samples where each feature is present. The pairwise relationships between features are then quantified using two complementary metrics: lift and Jaccard similarity. The lift metric measures how much more likely features are to occur together compared to random chance, while the Jaccard similarity captures the overlap between feature occurrences relative to their union. These metrics are combined using a weighted average (70% lift, 30% Jaccard) to produce a final co-occurrence score for each feature pair. The algorithm then creates graph edges for feature pairs whose combined score exceeds a specified threshold (default 0.3). To ensure the graph remains fully connected and to stabilize subsequent graph-based computations, self-loops with maximum weight (1.0) were added for each feature. The resulting graph structure was represented using two Torch tensors: an edge index tensor encoding the connectivity pattern and an edge weight tensor containing the corresponding co-occurrence scores, as shown in Figure 1. This graph representation captured both direct and indirect relationships between microbiome features, enabling the model to leverage community structure information during learning.

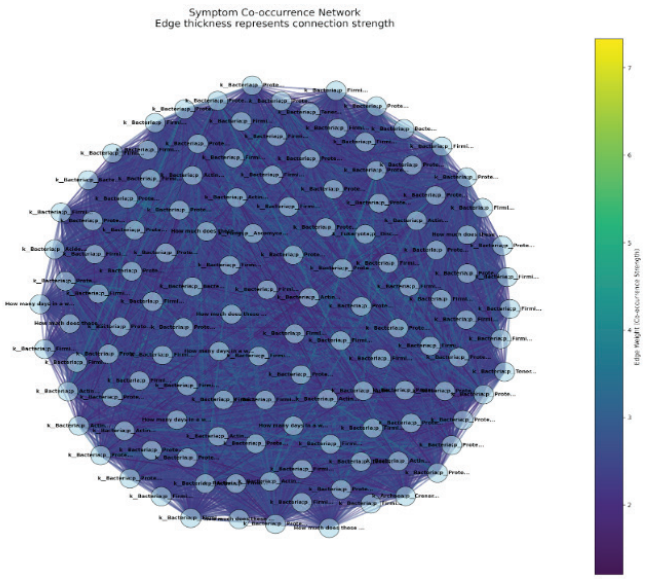


Figure 1: A graph visualisation, created for an individual sample. Round, light blue nodes represent taxon, and edge connections represent co-occurrences between taxon, ranging from violet to light blue: violet having the least weight and yellow having the most weight.

For activation functions, ReLU was used after each GCN layer, followed by batch normalization and a dropout rate of 0.3 to prevent overfitting. The model combined both mean and sum pooling operations to capture different aspects of the graph structure. The final classification layers used ReLU activation with a softmax output layer for multi-class prediction. The GNN implemented a graph convolutional network architecture designed for symptom classification, as shown in Figure 2. It employed three sequential GCN layers (GCNConv) that transformed the input features through message passing operations across the graph structure. Each GCN layer mapped the input to a hidden dimension space, maintaining the same hidden dimensionality across layers. The architecture incorporated batch normalization after each convolution to stabilize training and accelerate convergence. Following the graph convolutions, the model combined global mean and sum pooling operations to aggregate node-level features into graph-level representations. These pooled features were concatenated and processed through two fully connected layers for final classification. Dropout (0.3) is applied throughout to prevent overfitting. The model uses ReLU activation functions between layers to introduce non-linearity. This architecture enables the network to learn both local structural patterns through convolutions and global graph properties through pooling operations.

The following were the hyperparameters that were used for tuning.

Learning rate: [0.01, 0.001, 0.0001]
Hidden Layer Dimensions: [64,128,256]
Dropout rate: [0.2,0.3,0.4]
Co-occurrence threshold: [0.2, 0.3, 0.4]

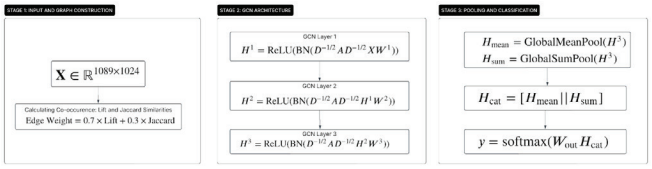


Figure 2: The architecture of the GNN. This figure shows a flowchart of its components, and mathematical functions are utilized to construct the graphs and train the model.

A specialized FFNN architecture, as shown in Figure 3, was developed using PyTorch Geometric, comprising three interconnected components. The proposed FFNN architecture consists of three parallel processing pathways that capture different aspects of the microbiome data. The first pathway implements a feature attention mechanism using two fully connected layers with ReLU activation, which learns to assign importance weights to different taxonomic features. The second pathway, termed the abundance network, processes the raw abundance values through a series of transformations, including linear layers, ReLU activation, batch normalization for stable training, and dropout (p=0.3) for regularization. Similarly, the third pathway, the interaction network, maintains an identical structure to the abundance network but processes the data independently to capture potential inter-species interactions. The outputs from the abundance and interaction networks are concatenated and processed through final layers that gradually reduce dimensionality while maintaining the same regularization techniques (batch normalization and dropout). All hidden layers maintain a consistent dimensionality, while the final output layer produces predictions appropriate for the target variable. This architecture enables the model to simultaneously consider both direct abundance effects and potential ecological interactions while automatically learning which features are most relevant for the prediction task.

Figure 3: The FFNN architecture. A flowchart of its components and mathematical functions is utilized to construct the graphs and train the model.

Evaluation Metrics:

To evaluate the performance of the model, various metrics were utilised to assess the performance of the model. The metrics were accuracy, precision, recall, F1 Score, confusion matrices, specificity, and sensitivity. All metrics were calculated using weighted averages to account for potential class imbalance in the dataset. Permutation importance was computed separately for each symptom model (severity and frequency), allowing symptom-specific feature relevance to be evaluated.

Feature Importance Analysis:

For feature importance analysis, the following methods were employed:

1. Permutation Importance: Feature importance was measured by randomly shuffling feature values and observing the decrease in the model's performance. A larger decrease in performance indicates higher feature importance, as it suggests the model heavily relies on that feature for accurate predictions.

2. Feature Correlation: Spearman's correlation between each microbial abundance was applied, providing a measure of the statistical relationship between individual features and the target variable, along with a direction of that measure.

Permutation importance and feature correlation were computed separately for each symptom model (severity and frequency), allowing symptom-specific feature relevance to be evaluated.

Cross Validation:

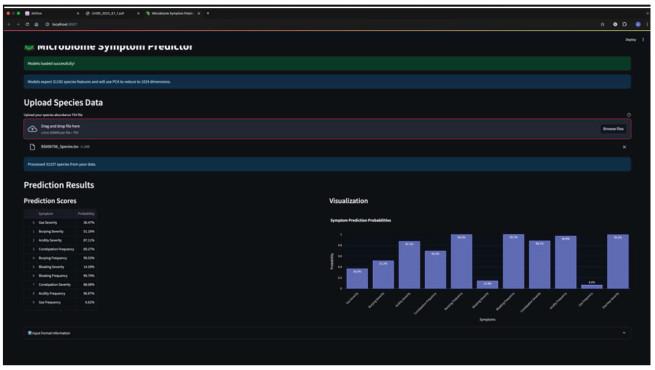
To ensure robust model evaluation, k-fold cross-validation was implemented on the best model (k=5), where the dataset was partitioned into k equal-sized segments. This approach iteratively used k-1 segments for training while reserving one segment for validation, rotating through all possible combinations. This methodology provided a more reliable estimate of the model's generalization performance compared to a single train-test split. The dataset was also tested on 5 external GMTs, and it was able to deliver predictions with an accuracy of ~0.8.

Web Application:

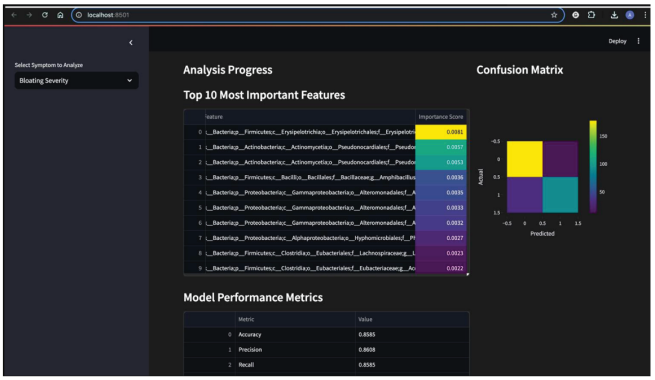
The web applications were then created and deployed. The web applications were built using the Hugging Face platform and are hosted on a HIPAA-compliant cloud server with secure, encrypted data handling. Two web apps were built to include two functionalities: the users can upload their GMT, and the model diagnoses them with IBS symptoms, and clinicians can upload entire datasets on which the model is run, returning validation metrics as well as potential treatment targets (**Figure 4**). Five practicing gastroenterologists reviewed and provided feedback on the web application interfaces and functionality prior to deployment. The web applications are hosted on a HIPAA-compliant cloud with encrypted data transmission. The links of the two web applications are as follows:

1. https://huggingface.co/spaces/anavgupta/ibs-predictions
2. https://huggingface.co/spaces/anavgupta/ibs-dataset

DOI: 10.36838/v7i8.10



(A)



(B)

Figure 4: A layout of web applications. (A). This shows the web application interface developed for Clinicians (B). This shows the web application interface developed for patients.

Result and Discussion

This section presents the results obtained from machine learning and deep learning models applied to the gut microbiome for predicting IBS. The evaluation metrics achieved for FFNNs, RF, AB, and GNNs are listed in Tables 1, 2, 3, and 4, respectively. After cleaning and preprocessing, 1089 samples were retained for analysis. Principal Component Analysis (PCA) reduced dimensionality from 31,259 to 1024. First, RF and AB models were trained, delivering accuracies of 65-75%. After this, GNNs were trained, delivering a 10-15% improvement with accuracies ranging from 80-85%. However, due to its complexity and overfitting on the training set, a lighter FFNN architecture was trained, which not only eliminated complexity but also improved accuracy by ~5%.

Table 1: Evaluation metrics for FFNN models for different symptoms. All symptom severities and frequencies were tested for accuracy, precision, recall, F1-score, specificity, and sensitivity.

FEED FORWARD NEURAL NETWORKS												
Symptom	SEVERITY						FREQUENCY					
	Accuracy	Precision	Recall	F1-Score	Specificity	Sensitivity	Accuracy	Precision	Recall	F1-Score	Specificity	Sensitivity
Bloating	0.88	0.88	0.88	0.88	0.87	0.87	0.86	0.86	0.86	0.86	0.85	0.85
Constipation	0.85	0.85	0.85	0.85	0.76	0.76	0.80	0.80	0.80	0.80	0.74	0.74
Acidity	0.84	0.84	0.84	0.83	0.78	0.78	0.81	0.81	0.8	0.81	0.71	0.71
Diarrhea	0.84	0.83	0.84	0.83	0.77	0.77	0.81	0.82	0.82	0.81	0.82	0.82
Burping	0.90	0.90	0.90	0.90	0.86	0.86	0.84	0.84	0.84	0.84	0.80	0.80

Table 2: Evaluation metrics for RF models for different symptoms. All symptom severities and frequencies were tested for accuracy, precision, recall, F1-score, specificity, and sensitivity.

RANDOM FOREST												
Symptom	SEVERITY						FREQUENCY					
	Accuracy	Precision	Recall	F1-Score	Specificity	Sensitivity	Accuracy	Precision	Recall	F1-Score	Specificity	Sensitivity
Bloating	0.75	0.69	0.59	0.59	0.84	0.63	0.64	0.66	0.68	0.67	0.82	0.66
Constipation	0.78	0.63	0.65	0.64	0.64	0.70	0.61	0.63	0.70	0.67	0.66	0.71
Acidity	0.70	0.66	0.70	0.68	0.65	0.67	0.53	0.50	0.51	0.52	0.65	0.67
Diarrhea	0.68	0.63	0.64	0.67	0.71	0.68	0.58	0.60	0.53	0.52	0.75	0.68
Burping	0.74	0.63	0.62	0.65	0.68	0.72	0.50	0.60	0.54	0.56	0.69	0.72

Table 3: Evaluation metrics for RF models for different symptoms. All symptom severities and frequencies were tested for accuracy, precision, recall, F1-score, specificity, and sensitivity.

ADAPTIVE BOOSTING												
Symptom	SEVERITY						FREQUENCY					
	Accuracy	Precision	Recall	F1-Score	Specificity	Sensitivity	Accuracy	Precision	Recall	F1-Score	Specificity	Sensitivity
Bloating	0.74	0.63	0.60	0.54	0.60	0.66	0.66	0.66	0.68	0.67	0.61	0.66
Constipation	0.74	0.62	0.65	0.63	0.66	0.63	0.67	0.63	0.7	0.67	0.65	0.64
Acidity	0.71	0.64	0.70	0.64	0.65	0.67	0.54	0.50	0.51	0.52	0.65	0.63
Diarrhea	0.61	0.63	0.66	0.67	0.64	0.68	0.55	0.60	0.53	0.52	0.66	0.66
Burping	0.64	0.61	0.68	0.66	0.65	0.63	0.56	0.60	0.54	0.56	0.65	0.63

Table 4: Evaluation metrics for GNN models for different symptoms. All symptom severities and frequencies were tested for accuracy, precision, recall, F1-score, specificity, and sensitivity.

GRAPH NEURAL NETWORKS												
Symptom	SEVERITY						FREQUENCY					
	Accuracy	Precision	Recall	F1-Score	Specificity	Sensitivity	Accuracy	Precision	Recall	F1-Score	Specificity	Sensitivity
Bloating	0.83	0.82	0.81	0.83	0.80	0.76	0.81	0.81	0.81	0.81	0.81	0.80
Constipation	0.81	0.82	0.82	0.82	0.76	0.76	0.75	0.75	0.75	0.75	0.74	0.74
Acidity	0.78	0.79	0.79	0.83	0.74	0.73	0.81	0.76	0.76	0.78	0.72	0.72
Diarrhea	0.84	0.83	0.84	0.83	0.77	0.77	0.81	0.73	0.73	0.73	0.73	0.73
Burping	0.87	0.87	0.86	0.86	0.76	0.76	0.82	0.82	0.82	0.82	0.80	0.80

FFNNs displayed the highest accuracy of 80-85% for symptom frequency and 85-90% for frequencies. Amongst the models built on the severity of different symptoms of IBS, the FFNN model for the severity of burping achieved the highest severity accuracy of 0.90, and amongst the models built on the frequency of different symptoms of IBS, the FFNN model for the frequency of bloating had the highest frequency accuracy of 0.86. ROC-AUC curves were computed for the FFNN, as shown in Figure 5.

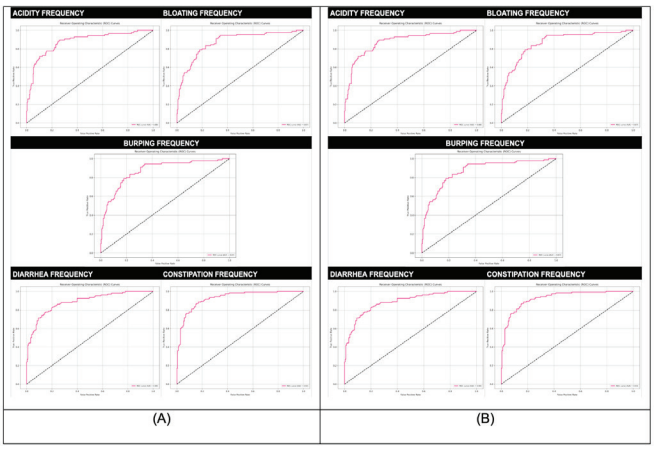


Figure 5: ROC-AUC curves generated for FFNNs using matplotlib. (A) This figure shows the ROC-AUC curves for the models generated for severities of all the symptoms. (B) This figure shows the ROC-AUC curves for the models generated for frequencies of all the symptoms.

To evaluate the model's performance, confusion matrices were built, which gave the frequency of true positives, true neg-

atives, false positives, and false negatives. Confusion matrices give insights into the biases of an ML model and check if it is correctly able to predict high and low outputs. The confusion matrices for different FFNN models built on the frequency and severity of each symptom were computed as listed in Table 5.

Table 5: Confusion Matrices for each symptom severity and frequency. Accuracies for different models lie between 80% and 90%.

FREQUENCIES			SEVERITIES		
ACIDITY			ACIDITY		
	Predicted 0	Predicted 1		Predicted 0	Predicted 1
Actual 0	227	25	Actual 0	213	15
Actual 1	47	44	Actual 1	43	65
BLOATING			BLOATING		
	Predicted 0	Predicted 1		Predicted 0	Predicted 1
Actual 0	165	19	Actual 0	172	17
Actual 1	36	88	Actual 1	25	97
BURPING			BURPING		
	Predicted 0	Predicted 1		Predicted 0	Predicted 1
Actual 0	121	30	Actual 0	145	24
Actual 1	36	105	Actual 1	34	98
CONSTIPATION			CONSTIPATION		
	Predicted 0	Predicted 1		Predicted 0	Predicted 1
Actual 0	207	22	Actual 0	232	20
Actual 1	35	69	Actual 1	41	49
DIARRHEA			DIARRHEA		
	Predicted 0	Predicted 1		Predicted 0	Predicted 1
Actual 0	202	20	Actual 0	263	11
Actual 1	35	71	Actual 1	47	48

A validation was performed on all the FFNN models, after which the model displayed strong validation results with a 1-2% difference in frequencies and a 2-4% difference in severities, as shown in Table 6. The model demonstrated consistent performance across all folds, with minimal variance between different data splits.

Table 6: Fivefold validation results. The metrics tested were accuracy, precision, recall, and F1 score.

CROSS VALIDATION								
Symptom	SEVERITY				FREQUENCY			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
Bloating	± 0.0211	± 0.0215	± 0.0233	± 0.0211	± 0.0159	± 0.0156	± 0.0159	± 0.0166
Constipation	± 0.0241	± 0.0244	± 0.0241	± 0.0241	± 0.0171	± 0.0198	± 0.0171	± 0.0185
Acidity	± 0.0164	± 0.0128	± 0.0141	± 0.0128	± 0.0191	± 0.0213	± 0.0191	± 0.0207
Diarrhea	± 0.0408	± 0.0379	± 0.0334	± 0.0379	± 0.0089	± 0.0091	± 0.0089	± 0.0085
Burping	± 0.0285	± 0.0290	± 0.0303	± 0.0290	± 0.0186	± 0.0193	± 0.0186	± 0.0195

Permutation importance and feature correlation computed separately for each symptom model allowed symptom-specific feature relevance to be evaluated. This approach provided complementary insights into both the predictive power of each taxon and the direction of the relationship between the taxon and the output. Through this, a clinician will be able to identify whether a particular taxon is negatively or positively correlated with the output and how strongly it impacts the model's performance. The top 10 taxa for each symptom, namely diarrhoea, constipation, burping, acidity, and bloating, along with their importance and correlation to IBS, are shown in Tables 7, 8, 9, 10, and 11, respectively.

Table 7: The top ten most important taxa for diarrhea. Arranged from most influential to least influential according to their permutation scores. Correlation coefficients are also included.

DIARRHEA FREQUENCY			DIARRHEA SEVERITY		
TAXON	IMPORTANCE	CORRELATION	TAXON	IMPORTANCE	CORRELATION
Amycolatopsis granulosa	0.0065	-0.0145	*Amycolatopsis granulosa*	0.0087	0.0176
[Clostridium] cocleatum	0.0059	0.0278	*Aeromonas veronii*	0.004	0.0042
Acidaminococcus fermentans	0.0054	-0.0432	*Amniculibacterium sp. G2-70*	0.0037	-0.0185
Acinetobacter bereziniae	0.0051	-0.0089	*Alteromonas sp. RKMC-009*	0.0036	0.0396
Anaerobutyricum soehngenii	0.0047	-0.0256	*[Clostridium] cocleatum*	0.0036	-0.0042
[Enterobacter] lignolyticus	0.0045	0.0183	*[Clostridium] populeti*	0.0036	-0.0122
Alcanivorax profundi	0.0043	-0.0521	*Acinetobacter junii*	0.0036	0.0079
Acetivibrio clariflavus	0.0041	-0.0167	*Acholeplasma oculi*	0.0035	0.0384
Anaerococcus provencensis	0.004	0.0092	*Acetivibrio mesophilus*	0.0032	0.0272
Amniculibacterium sp. G2-70	0.0038	-0.0394	*Acidaminococcus fermentans*	0.0032	-0.0335

Table 8: The top ten most important taxa for constipation. Arranged from most influential to least influential according to their permutation scores. Correlation coefficients are also included.

CONSTIPATION FREQUENCY			CONSTIPATION SEVERITY		
TAXON	IMPORTANCE	CORRELATION	TAXON	IMPORTANCE	CORRELATION
Amycolatopsis circi	0.0064	0.0315	*Amycolatopsis granulosa*	0.0082	0.036
Aminicella lysinilytica	0.0041	-0.0123	*Acetitomaculum ruminis*	0.0053	0.2145
Acidaminococcus fermentans	0.0027	-0.0457	*Anaerophilus nitritogenes*	0.0043	-0.0497
Amphritea atlantica	0.0026	0.0115	*Acetivibrio saccincola*	0.0042	0.0202
Aminomonas paucivorans	0.0025	0.034	*Amycolatopsis sp. CA-128772*	0.0034	-0.0072
Lachnoclostridium aminophilum	0.0023	-0.0122	*Alkalicoccus urumqiensis*	0.0032	0.0071
Eubacterium cellulosolvens	0.0023	-0.0181	*Acinetobacter guillouiae*	0.0029	-0.0356
Aliarcobacter cryaerophilus	0.0023	0.021	*[Clostridium] dakarense*	0.0023	0.0685
Achromobacter sp. GG226	0.0022	0.0465	*Acholeplasma oculi*	0.0023	-0.064
Akkermansia sp. BIOML-A61	0.002	0.0112	*Acinetobacter seiferti*	0.0023	0.0661

Table 9: The top ten most important taxa for burping. Arranged from most influential to least influential according to their permutation scores. Correlation coefficients are also included.

BURPING FREQUENCY			BURPING SEVERITY		
TAXON	IMPORTANCE	CORRELATION	TAXON	IMPORTANCE	CORRELATION
Achromobacter spanius	0.0053	0.0458	*Amycolatopsis granulosa*	0.0075	-0.0126
Amycolatopsis taiwanensis	0.0053	-0.0141	*Amniculibacterium sp. G2-70*	0.0033	-0.0325
Anaerofilum sp. BX8	0.0049	-0.0199	*[Clostridium] cocleatum*	0.0031	0.0299
Aliarcobacter lanthieri	0.004	-0.0375	*[Enterobacter] lignolyticus*	-0.0028	-0.0298
Aeromicrobium sp. zg-Y50	0.0037	0.0132	*Acetivibrio clariflavus*	-0.0027	0.0436
Amycolatopsis granulosa	0.0034	0.0045	*Eubacterium cellulosolvens*	0.0027	-0.163
[Clostridium] populeti	0.0033	-0.0274	*[Clostridium] polysaccharolyticum*	0.0025	0.0261
Alcaligenes sp. SORT26	0.0033	-0.0235	*Acidovorax radicis*	0.0024	0.0032
Acidaminococcus fermentans	0.0029	-0.0307	*Alcanivorax profundi*	0.0023	0.0019
Acinetobacter seiferti	0.0028	0.3497	*Agathobaculum sp. NSJ-28*	0.0022	-0.0081

DOI: 10.36838/v7i8.10

Table 10: The top ten most important taxa for acidity. Arranged from most influential to least influential according to their permutation scores. Correlation coefficients are also included.

ACIDITY FREQUENCY			ACIDITY SEVERITY		
TAXON	IMPORTANCE	CORRELATION	TAXON	IMPORTANCE	CORRELATION
[Enterobacter] lignolyticus	-0.0026	-0.0359	*Amycolatopsis circi*	0.0068	-0.0009
[Flexibacter] sp. ATCC 35103	-0.0021	0.0043	*Anaerobranca californiensis*	0.0057	-0.0198
Acinetobacter bereziniae	0.002	0.0446	*Amphritea atlantica*	0.0053	-0.0117
Acetivibrio ethanolgignens	-0.002	0.0028	*Acetobacter pasteurianus*	0.0044	0.0207
Acetitomaculum ruminis	-0.0019	-0.0118	*Amniculibacterium sp. G2-70*	0.0043	-0.0362
Anaerobutyricum soehngenii	-0.0019	-0.0189	*Romboutsia dakarense*	0.0041	-0.0094
Acetanaerobacterium sp. MSJ-12	-0.0018	-0.0333	*Acinetobacter sp. 105-3*	0.0041	-0.0624
[Clostridium] methoxybenzovorans	-0.0018	-0.041	*Aminicella lysinilytica*	0.0041	-0.0012
Acidaminococcus massiliensis	-0.0018	0.0158	*Alcanivorax gelatiniphagus*	0.0038	-0.0023
Alkalihalobacterium elongatum	-0.0018	0.0014	*Amycolatopsis sp. AA4*	0.0038	-0.0568

Table 11: The top ten most important taxa for bloating. Arranged from most influential to least influential according to their permutation scores. Correlation coefficients are also included.

BLOATING FREQUENCY			BLOATING SEVERITY		
TAXON	IMPORTANCE	CORRELATION	TAXON	IMPORTANCE	CORRELATION
Anabaena lutea	0.0117	-0.0012	*Amycolatopsis granulosa*	0.0089	0.0226
[Ruminococcus] lactaris	0.0092	-0.0569	*Amniculibacterium sp. G2-70*	0.005	-0.0175
Anaerococcus provencensis	0.0065	-0.0793	*Megasphaera vaginalis*	0.0042	0.0336
Acanthamoeba castellanii	0.0057	0.0401	*Amphritea atlantica*	0.0039	0.0188
Achromobacter spanius	0.0051	0.057	*Amycolatopsis circi*	0.0039	-0.0316
Acetivibrio clariflavus	0.0051	-0.0246	*Erysipelatoclostridium cocleatum*	0.0037	0.0345
Anaerococcus sp. mt242	0.0049	-0.0039	*Anaerococcus lactolyticus*	0.0033	-0.0041
[Eubacterium] cellulosolvens	0.0047	-0.0191	*Aminicella lysinilytica*	0.0032	-0.0033
Alteromonas sp. RKMC-009	0.0047	0.0232	*Amycolatopsis niigatensis*	0.0032	0.0624
Abyssibacter profundi	0.0044	-0.0353	*Amycolatopsis sp. AA4*	0.0032	-0.0344

Conclusion

This study developed and validated a novel deep learning approach for IBS diagnosis using gut microbiome data from 1,089 Indian patients. The FFNN achieved 85-90% accuracy for symptom severities and 80-85% for frequencies, outperforming existing models by 15-20%. 71 unique taxonomic biomarkers were identified, with Amycolatopsis, Acinetobacter, Clostridium, and Acetivibrio species occurring repeatedly across symptoms and having the highest permutation scores. The model's performance was validated through 5-fold cross-validation and external testing, demonstrating robust generalization. Two web applications were developed for clinical use, enabling both individual diagnosis and dataset-wide analysis. This work represents a significant advancement in objective IBS diagnosis, particularly for the understudied South Asian population, while providing specific microbial targets for therapeutic intervention. Future work should incorporate longitudinal data and additional health metrics to further improve diagnostic accuracy. The identified biomarkers require experimental validation to confirm their biological relevance in IBS pathophysiology.

Acknowledgments

AG would like to thank Max Kushnir (Co-Founder and CSO, Sova Health) for providing the dataset for the research project. AG would also like to thank Mr. Pankaj Aggarwal (Data Scientist, Duke University) for his valuable guidance and feedback. AG would like to acknowledge Aashna Saraf, Founder of CreatED, for providing valuable feedback and guidance throughout the project.

References

1. MayoClinic *Irritable Bowel Syndrome*; USA, 2024.
2. Oka, P.; Parr, H.; Barberio, B.; Black, C. J.; Savarino, E. V.; Ford, A. C., Global prevalence of irritable bowel syndrome according to Rome III or IV criteria: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* **2020**, *5* (10), 908-917.
3. Menees, S.; Chey, W., The gut microbiome and irritable bowel syndrome. *F1000Res* **2018**, *7*.
4. Zhao, Y.; Zhu, S.; Dong, Y.; Xie, T.; Chai, Z.; Gao, X.; Dai, Y.; Wang, X., The Role of Gut Microbiome in Irritable Bowel Syndrome: Implications for Clinical Therapeutics. *Biomolecules* **2024**, *14* (12), 1643.
5. Camilleri, M., Diagnosis and Treatment of Irritable Bowel Syndrome: A Review. *Jama* **2021**, *325* (9), 865-877.
6. Auslander, N.; Gussow, A. B.; Koonin, E. V., Incorporating Machine Learning into Established Bioinformatics Frameworks. *Int J Mol Sci* **2021**, *22* (6).
7. Fukui, H.; Nishida, A.; Matsuda, S.; Kira, F.; Watanabe, S.; Kuriyama, M.; Kawakami, K.; Aikawa, Y.; Oda, N.; Arai, K.; Matsunaga, A.; Nonaka, M.; Nakai, K.; Shinmura, W.; Matsumoto, M.; Morishita, S.; Takeda, A. K.; Miwa, H., Usefulness of Machine Learning-Based Gut Microbiome Analysis for Identifying Patients with Irritable Bowels Syndrome. *J Clin Med* **2020**, *9* (8).
8. Hernández Medina, R.; Kutuzova, S.; Nielsen, K. N.; Johansen, J.; Hansen, L. H.; Nielsen, M.; Rasmussen, S., Machine learning and deep learning applications in microbiome research. *ISME Communications* **2022**, *2* (1), 98.
9. McDonald, D.; Hyde, E.; Debelius, J. W.; Morton, J. T.; Gonzalez, A.; Ackermann, G.; Aksenov, A. A.; Behsaz, B.; Brennan, C.; Chen, Y.; DeRight Goldasich, L.; Dorrestein, P. C.; Dunn, R. R.; Fahimipour, A. K.; Gaffney, J.; Gilbert, J. A.; Gogul, G.; Green, J. L.; Hugenholtz, P.; Humphrey, G.; Huttenhower, C.; Jackson, M. A.; Janssen, S.; Jeste, D. V.; Jiang, L.; Kelley, S. T.; Knights, D.; Kosciolek, T.; Ladau, J.; Leach, J.; Marotz, C.; Meleshko, D.; Melnik, A. V.; Metcalf, J. L.; Mohimani, H.; Montassier, E.; Navas-Molina, J.; Nguyen, T. T.; Peddada, S.; Pevzner, P.; Pollard, K. S.; Rahnavard, G.; Robbins-Pianka, A.; Sangwan, N.; Shorenstein, J.; Smarr, L.; Song, S. J.; Spector, T.; Swafford, A. D.; Thackray, V. G.; Thompson, L. R.; Tripathi, A.; Vázquez-Baeza, Y.; Vrbanac, A.; Wischmeyer, P.; Wolfe, E.; Zhu, Q.; Knight, R., American Gut: an Open Platform for Citizen Science Microbiome Research. *mSystems* **2018**, *3* (3).
10. Kanz, C.; Aldebert, P.; Althorpe, N.; Baker, W.; Baldwin, A.; Bates, K.; Browne, P.; van den Broek, A.; Castro, M.; Cochrane, G.; Duggan, K.; Eberhardt, R.; Faruque, N.; Gamble, J.; Diez, F. G.; Harte, N.; Kulikova, T.; Lin, Q.; Lombard, V.; Lopez, R.; Mancuso, R.; McHale, M.; Nardone, F.; Silventoinen, V.; Sobhany, S.; Stoehr, P.; Tuli, M. A.; Tzouvara, K.; Vaughan, R.; Wu, D.; Zhu, W.; Apweiler, R., The EMBL Nucleotide Sequence Database. *Nucleic Acids Res* **2005**, *33* (Database issue), D29-33.
11. Kim, J. H.; Lin, E.; Pimentel, M., Biomarkers of Irritable Bowel Syndrome. *J Neurogastroenterol Motil* **2017**, *23* (1), 20-26.
12. Ford, A. C.; Bercik, P.; Morgan, D. G.; Bolino, C.; Pintos-Sanchez, M. I.; Moayyedi, P., Validation of the Rome III criteria for the diagnosis of irritable bowel syndrome in secondary care. *Gastroenterology* **2013**, *145* (6), 1262-70.e1.

Authors

Anav Gupta is presently a senior at The Shri Ram School, Moulsari, Gurugram, Haryana, India. He is keenly interested in technology, Computer Science, and Math. He is a passionate photographer and serves as the editor-in-chief for The Shri Ram School magazine. He has won several prestigious STEM Awards, including being selected as a finalist in the IRIS National Science Fair, and presented this work in the Computational Biology and Bioinformatics Category.

Nirupma Singh is a Bioinformatics Scientist with a doctorate in Biotechnology and Bioinformatics from the University of Delhi, with six years of hands-on research and development experience. Her journey is marked by a robust foundation in Machine Learning and Python, with five years of expertise. Proficient in Linux and cloud servers like AWS. She excels in structural biology, systems biology, protein/gene network analysis, data mining, and computational genomics.

DOI: 10.36838/v7i8.10