Social networks have become the most influential digital platforms for human interaction, marketing, and information exchange. Understanding and predicting user behavior in such networks is essential for personalized recommendations, advertising strategies, and detecting malicious activity. This paper presents an analytical framework for Behavioral Prediction using Social Graph Analytics, which combines graph-based relationship data and user activity features to predict future actions within a social network. The proposed model integrates Graph Neural Network (GNN) representations with temporal activity patterns to learn behavioral dependencies among users. Experiments conducted on a Twitter dataset demonstrate an improvement of 10–12% in prediction accuracy compared to traditional machine learning models. The system provides interpretable insights into community influence, helping organizations identify key behavior drivers and influential nodes within a network.

Index Terms— Behavioral Analytics, Data Mining, Graph Neural Networks, Social Network Analysis, User Behavior Prediction.

I. INTRODUCTION

Social networking platforms such as Twitter, Facebook, and Instagram generate massive volumes of user interaction data daily. Each interaction — a post, comment, or follow — represents a social relationship that can reveal valuable behavioral patterns. Predicting how users behave in such environments helps in applications such as content recommendation, sentiment forecasting, market segmentation, and misinformation detection.

Conventional behavior prediction models rely mainly on textual or temporal data, ignoring the network structure among users. However, user behavior in social networks is strongly influenced by social connections and community interactions. Social Graph Analytics captures these structural relationships and enables modeling of how social influence affects user decisions.

This research focuses on developing an analytical system that predicts user behavior using graph-based features combined with machine learning techniques. By transforming the social network into a graph structure and applying Graph Neural Networks (GNNs), we extract both direct and indirect relational patterns among users.

The main contributions of this paper are as follows:

1. A data-driven framework that integrates user activity and graph features for behavioral prediction.

2. Implementation of a Graph Neural Network model to learn social influence patterns.

3. Evaluation of prediction accuracy using real-world social network data.

4. Analysis of how community structures and relationships affect user actions.

II. RELATED WORK

Early research on user behavior prediction primarily used probabilistic or regression-based models to analyze time-series user activities. These models achieved moderate success but failed to capture relational dependencies. Later, data mining approaches such as clustering and decision trees were used to categorize users based on activity levels.

Recent advancements in Graph Neural Networks (GNNs) have enabled learning from graph-structured data, allowing models to encode both node features and edge relationships. Studies such as [1] and [2] demonstrated the effectiveness of GNNs for link prediction, influence estimation, and community detection. However, most of these works focused on static networks and lacked behavioral forecasting capabilities.

Our work differs by combining GNN-based feature learning with dynamic behavioral prediction, allowing both temporal and relational dependencies to contribute to the model's decision-making process.

III. METHODOLOGY

A. System Architecture

The proposed architecture (Fig. 1) includes five main modules:

1. Data Collection – Extract user data from social networking platforms.

2. Data Preprocessing – Clean text, anonymize users, and create an interaction graph.

3. Feature Extraction – Compute network metrics and user activity vectors.

4. Graph Model Training – Train the GNN on extracted graph features.

5. Behavior Prediction – Predict user actions such as sharing, liking, or commenting.

Fig. 1: System Architecture for Behavioral Prediction using Social Graph Analytics (Data Input → Graph Construction → Feature Extraction → GNN Model → Output Prediction)

B. Dataset and Preprocessing

The dataset used for implementation was collected from the Twitter API containing 10,000 users, 100,000 tweets, and 200,000 relationships. The raw data was preprocessed by:

Removing inactive users and bots.

Tokenizing and normalizing text content.

Constructing a directed graph , where V represents users and E represents follower/following relationships.

Each user node contains attributes such as:

Average posting frequency

Sentiment polarity of tweets

Engagement rate (likes + retweets)

Graph metrics (degree, centrality, clustering coefficient)

C. Feature Extraction

User behavior is influenced by personal attributes and social context. We extracted:

1. Structural Features: Degree centrality, betweenness centrality, PageRank score.

2. Temporal Features: Time gap between consecutive posts, activity burstiness.

3. Content Features: Sentiment and keyword embeddings from user posts.

4. Community Features: Detected using Louvain modularity-based clustering.

D. Model Design

The core predictive model is a Graph Neural Network (GNN) implemented using PyTorch Geometric.
The model updates each node’s representation by aggregating information from its neighbors:

$$h_v^{(k+1)} = \sigma \left(W^{(k)} \cdot \text{AGGREGATE} \left(\{ h_u^{(k)} : u \in N(v) \} \right) \right)$$

where is the set of neighbors of node , is a learnable weight matrix, and is a nonlinear activation function.
The final output is a probability score representing the likelihood of a user performing a future behavior (e.g., sharing a post).

E. Implementation Details

Language & Tools: Python, NetworkX, PyTorch Geometric, Pandas.

Hardware: Intel i7 CPU, 16 GB RAM, NVIDIA GTX 1660 GPU.

Training: 80:20 train-test split, learning rate = 0.001, 50 epochs.

Baseline models: Logistic Regression and Random Forest.

IV. IMPLEMENTATION AND RESULTS

A. Evaluation Metrics

The system performance was evaluated using:

Accuracy

Precision

Recall

F1-Score

Area Under Curve (AUC)

B. Quantitative Results

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	78.2%	75.6%	73.9%	74.7%
Random Forest	82.4%	80.1%	79.3%	79.7%
Proposed GNN Model	90.6%	89.3%	88.8%	89.0%

The GNN-based model significantly outperformed traditional classifiers, showing the effectiveness of incorporating graph structure and community dynamics.

C. Discussion

The results indicate that network topology features (e.g., centrality, community membership) contribute substantially to behavioral prediction. Users connected to highly active communities showed a higher likelihood of performing social actions.
The GNN model was also able to highlight influential users whose actions affected group-level behavior.

V. CONCLUSION AND FUTURE WORK

This research demonstrates that integrating graph-based analytics with machine learning can effectively predict user behavior in social networks. The proposed system achieved a prediction accuracy of over 90% on real-world data.

In future work, we plan to:

Extend the model to handle temporal graphs that evolve over time.

Implement explainable AI (XAI) modules to interpret GNN decisions.

Apply federated learning for privacy-preserving social analytics.

Integrate real-time data streaming for live behavioral prediction.

The findings of this work can support applications in recommendation systems, digital marketing, and social influence modeling.