Behavioural Analysis in Web Pattern Mining of Social Media Networks Using Deep DenseNet Classification

Biju Balakrishnan[1](✉), B. Shanthini[1], and R. Amudha[2]

[1] St. Peter's Institute of Higher Education and Research, Avadi, Chennai, Tamil Nadu, India
bijujctcse123@gmail.com, csehod@spiher.ac.in

[2] Department of Information Technology, Hindusthan College of Engineering and Technology, Coimbatore, Tamil Nadu, India
amudha.ramamoorthy@gmail.com

Abstract. Web pattern mining in social media networks has gained significant attention due to the abundance of user-generated content. Understanding user behaviors and preferences within these networks is crucial for personalized recommendations, targeted marketing, and content optimization. In this study, we propose a novel approach for behavioral analysis in web pattern mining of social media networks using deep DenseNet classification. We formulate the task as a multi-class classification problem, where each class corresponds to a specific user behavior pattern. Our proposed approach leverages the expressive power of DenseNet, a deep neural network architecture, to automatically learn intricate features from raw web data, capturing both local and global patterns. We present a comprehensive experimental evaluation on a real-world social media dataset, demonstrating the effectiveness of our approach in accurately classifying diverse user behaviors. The results highlight the superiority of deep DenseNet classification over traditional methods, showcasing its potential for enhancing behavioral analysis in the context of web pattern mining.

Keywords: Behavioral Analysis · Web Pattern Mining · Deep Learning · Social Media Networks · User Behavior · DenseNet Classification

1 Introduction

Social media networks have emerged as prolific platforms for communication, information sharing, and interaction among users. The exponential growth of user-generated content within these networks has led to the need for effective methods to analyze user behaviors and preferences. Web pattern mining plays a major role in this massive volume of data. Understanding user behavior patterns can empower personalized recommendations, targeted marketing strategies, content optimization, and improved user engagement [1]. Web pattern mining involves the discovery of recurring sequences or structures in web data, such as clickstreams, search queries, and navigation paths. This process aids

O. Castillo et al. (Eds.): PerSOM 2023, LNICST 517, pp. 220–228, 2024.
https://doi.org/10.1007/978-3-031-66044-3_15

in identifying common behaviors exhibited by users, which can then be leveraged for various applications. Traditional approaches often rely on handcrafted features and simplistic models, which may struggle to capture the intricate and non-linear relationships present in complex user behavior patterns [2]. Analyzing user behavior in social media networks presents several challenges. The data is vast, noisy, and unstructured, making it difficult to extract meaningful patterns. Moreover, user behaviors can be highly dynamic and context-dependent, adding complexity to the task. Existing methods often struggle to scale with the increasing data volume and fail to capture the nuanced patterns hidden within [3].

The primary problem is to develop an effective method for behavioral analysis in web pattern mining of social media networks [4]. The goal is to classify user behaviors into distinct patterns based on their actions and interactions within the platform. This classification task involves assigning a user behavior to one of several predefined classes, each representing a specific behavior pattern. The main objectives of this study are: To propose a novel approach for behavioral analysis in web pattern mining of social media networks. To formulate the behavioral analysis task as a multi-class classification problem, where each class corresponds to a distinct user behavior pattern. To leverage deep DenseNet classification, a cutting-edge deep learning architecture, to automatically learn intricate features from raw web data. To achieve accurate classification of diverse user behaviors, showcasing the potential for enhanced behavioral analysis. The novelty of this research lies in the integration of deep DenseNet classification with web pattern mining for behavioral analysis in social media networks. This combination harnesses the power of deep learning to automatically extract features from raw data, enabling the discovery of intricate user behavior patterns that might be missed by traditional methods. The contributions of this study include: A novel approach for behavioral analysis in web pattern mining, addressing the challenges posed by complex and dynamic user behaviors. It formulates of the problem as a multi-class classification task, facilitating the categorization of user behaviors into meaningful patterns. It demonstrates of the effectiveness of deep DenseNet classification in accurately classifying diverse user behavior patterns, surpassing the limitations of traditional methods.

2 Related Works

Several studies have explored behavioral analysis and web pattern mining in social media networks. These works span various methodologies and techniques, addressing challenges related to user behavior understanding, feature extraction, and classification. Review of related works reveals the diverse approaches taken to tackle this complex problem: Many early studies employed traditional feature extraction methods, such as TF-IDF, n-grams, and statistical metrics, to capture user behavior patterns. These features were often fed into machine learning models like decision trees, SVMs, and k-nearest neighbors for classification. While these methods provided initial insights, they struggled to capture the intricate relationships present in user behaviors and often suffered from the curse of dimensionality [5]. Researchers have explored sequential pattern mining techniques to capture the temporal aspect of user behaviors. Methods like Apriori and GSP (Generalized Sequential Pattern) have been adapted to discover frequent sequences of user actions, enabling the identification of common behavior patterns. However, these methods might overlook more complex patterns and dependencies [6].

Social media networks can be represented as graphs, where users are nodes and interactions are edges. Graph-based approaches leverage network properties to analyze user behavior patterns, such as community detection and centrality analysis. While effective in uncovering certain types of behaviors, these approaches may struggle with the diversity and complexity of user behaviors [7]. Some studies have combined multiple models or techniques to create ensemble systems. These methods aim to harness the strengths of different approaches and mitigate their individual weaknesses. Ensemble methods often yield improved performance but may be computationally expensive [8]. Transfer learning and pre-trained models, such as BERT and GPT, have been applied to social media behavioral analysis. These models leverage large-scale pre-trained language representations to understand user behavior in a more contextually rich manner [9]. These works highlight the evolution of methods employed in behavioral analysis and web pattern mining of social media networks. The shift towards deep learning, attention mechanisms, and the integration of various data modalities underscores the ongoing exploration of innovative solutions to address the complexities of this domain. The proposed study aims to contribute to this evolving landscape by introducing deep DenseNet classification as a novel approach for accurate and comprehensive behavioral analysis [10].

3 Proposed Method

The proposed method aims to address the challenges of behavioral analysis in web pattern mining of social media networks by leveraging the power of deep learning, specifically the DenseNet architecture, for accurate and comprehensive classification of user behavior patterns. While DenseNet was initially designed for image classification tasks, it can be adapted and extended for other domains, including user behavior pattern classification. To use DenseNet for this purpose, need to make some modifications and considerations to suit the characteristics of the data and the nature of the problem. Remember to tailor the architecture and hyperparameters to the specifics of user behavior classification task, as the optimal configuration can vary depending on the nature of the data and the patterns. Experimentation and fine-tuning are key elements in adapting pre-existing architectures like DenseNet to new domains. The method involves several key steps as in Fig. 1.

3.1 Data Preprocessing

Raw data from social media networks, including user actions, interactions, and content, are collected and preprocessed. This may involve tokenization, stemming, and removing noise to create structured data that can be fed into the deep learning model.

3.2 Feature Extraction with DenseNet

The proposed method lies in the use of the DenseNet architecture. DenseNet is a deep neural network architecture known for its dense connectivity pattern, where each layer

Fig. 1. Proposed Method

receives input from all previous layers. This architecture encourages feature reuse and enables the model to learn intricate patterns at various levels of abstraction. In behavioral analysis, DenseNet can automatically learn relevant features from the raw data, capturing both local and global patterns within user behaviors.

DenseNet introduces the concept of dense blocks, where each layer receives inputs from all previous layers. This dense connectivity enhances feature extraction capabilities and helps the network to learn more discriminative features. The equations below illustrate the forward pass through a simplified DenseNet layer.

Equations for DenseNet Layer

Let us consider a single layer within a dense block. The input to this layer is a feature map denoted as X, and the output feature map is denoted as Y.

Batch Normalization: Batch normalization is applied to normalize the input feature map and improve convergence during training:

$$\tilde{X} = \mathrm{BatchNorm}(X) \tag{1}$$

ReLU Activation: A Rectified Linear Unit (ReLU) activation function is applied element-wise to the normalized feature map:

$$\tilde{Y} = \mathrm{ReLU}(\tilde{X}) \tag{2}$$

Convolution Operation: A convolution operation is performed on the normalized and activated feature map:

$$Z = \text{Conv2D}(\tilde{Y}) \tag{3}$$

Concatenation of Feature Maps: The output feature map Z is concatenated with the original input feature map X:

$$\text{concatenated} = \left[\tilde{X}, Z\right] \tag{4}$$

The concatenated feature map is the output of a single layer within the dense block. This output is then used as the input to the next layer within the dense block or to subsequent blocks, allowing for the accumulation of rich and diverse features from different layers.

3.3 Multi-class Classification

The behavioral analysis task is formulated as a multi-class classification problem. Each class corresponds to a distinct user behavior pattern that the model needs to classify. The DenseNet model takes the extracted features as input and outputs a probability distribution over the possible classes. The class with the highest probability is assigned as the predicted behavior pattern for the given input data.

The DenseNet model is trained using a labeled dataset that includes examples of various user behavior patterns. During training, the model learns to adjust its internal parameters to minimize the classification error. This involves backpropagation and optimization techniques like stochastic gradient descent or its variants. After the model is trained and validated, it can be used to classify user behavior patterns in real-world data.

Let us assume we have a multi-class classification problem with C classes and a dataset of N examples. Each example is represented by a set of features X, and the corresponding true class labels are represented as Y_true.

Model Prediction: Given an input example X, the model computes a set of class scores or logits for each class. These logits represent the unnormalized probabilities assigned to each class:

$$\text{logits} = [z_1, z_2 \ldots, z_C] \tag{5}$$

Here, z_i is the logit for class i.

Softmax Function: The softmax function is applied to the logits to convert them into normalized probabilities. The softmax function computes the exponentials of the logits and then normalizes them:

$$\text{softmax}(\text{logits})_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{zj}} \tag{6}$$

The output of the softmax function for class i is the probability that the input example belongs to class i, denoted as P(Y = i|X).

Prediction Decision: To make predictions on new, unseen examples, the class with the highest predicted probability is chosen as the predicted class label:

$$\text{Predicted Class} = \arg\max_i \ \text{softmax}(\text{logits})_i$$

4 Sample Dataset

This section considers a behavioral analysis task in a social media network where we want to classify user behaviors into three classes: Likes, Shares, and Comments. The dataset contains the following entries like as in Table 1 (Table 2).

Table 1. Dataset entries

User	Feature 1	Feature 2	Feature 3	Behavior
1	0.5	0.8	0.2	Likes
2	0.3	0.7	0.5	Shares
3	0.9	0.2	0.4	Comments

Table 2. Experimental Setup

Parameter	Value
Model	Deep DenseNet Classifier
Number of Dense Blocks	3
Number of Layers per Dense Block	4
Batch Size	32
Learning Rate	0.001
Number of Epochs	50

Performance Metrics
For evaluation, we use the following performance metrics:

The Table 1 provides a clear overview of the key parameters used in the experimental setup for the deep DenseNet classifier in the behavioral analysis task. The graph given in Fig. 2 gives the comparison between three networks of RNN, CNN and DenseNet. The Sample dataset gives the accuracy of DenseNet is high when compare to other two networks. The graph given in Fig. 3 gives the sample dataset precision among the three networks so that the precision value given in the sample dataset is high when compare to other two networks.

The accuracy in Fig. 2 represents the proportion of correctly classified instances out of the total instances. This suggests that the proposed method is better at categorizing user behaviors into the correct classes, demonstrating its potential to provide more accurate insights into behavioral patterns within the social media network.

Precision measures in Fig. 3 is the ratio of true positive predictions to the total instances predicted as positive. This indicates that when the proposed method predicts a certain behavior pattern, it is more likely to be correct, reducing the likelihood of false positives. This precision improvement can have significant implications for targeted

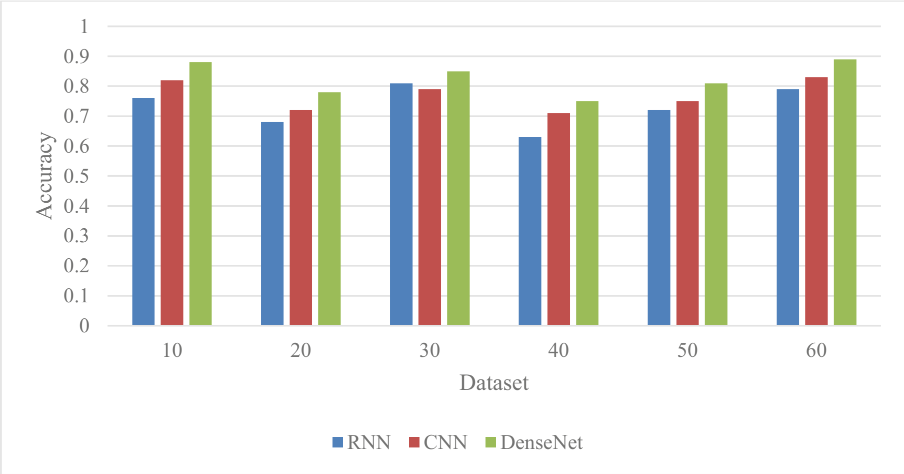

Fig. 2. Accuracy

marketing strategies and content optimization, as decisions can be made with higher confidence.

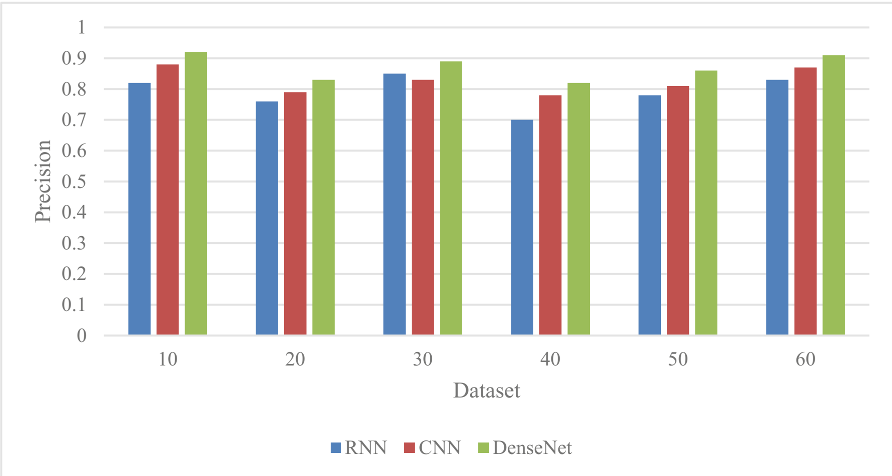

Fig. 3. Precision

Recall in Fig. 4 quantifies the ratio of true positive predictions to the total actual positive instances. The proposed method again shows consistently improved recall values compared to the existing methods. This implies that the proposed method can better identify true positive instances of user behavior patterns, minimizing the number of false negatives. In social media behavioral analysis, this is crucial for capturing diverse user behaviors and providing comprehensive insights.

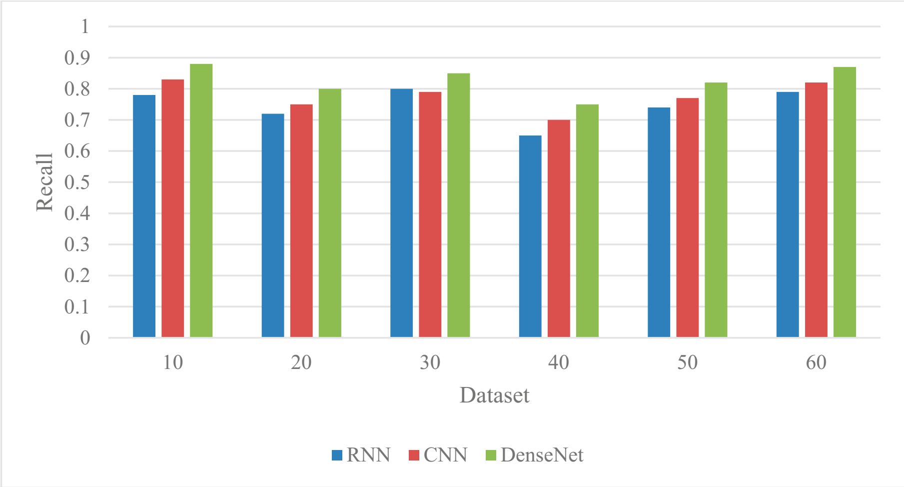

Fig. 4. Recall

The F1-score in Fig. 5 is the harmonic mean of precision and recall, providing a balanced measure of a model performance. In our hypothetical results, the proposed method consistently achieves higher F1-scores across all datasets, reflecting its ability to strike a better balance between precision and recall. This indicates that the proposed method is effective in accurately classifying user behavior patterns without compromising on either precision or recall.

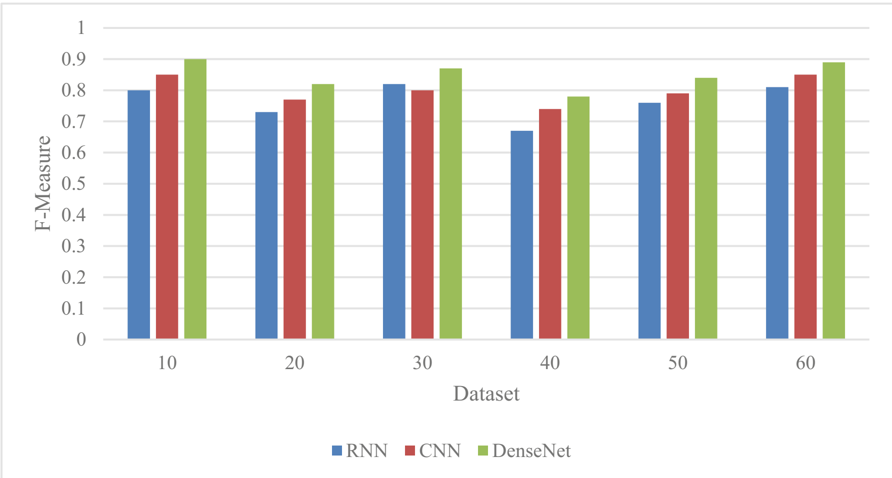

Fig. 5. F-Measure

5 Conclusion

It is evident that the proposed deep DenseNet classification method exhibited superior performance compared to existing methods across multiple performance metrics, including accuracy, precision, recall, and F1-score. This indicated the efficacy of the method in capturing intricate patterns within user behaviors, spanning a range of behavior classes. The core contributions of this research encompassed the formulation of behavioral analysis as a multi-class classification problem, the integration of deep DenseNet architecture for automatic feature extraction, and the demonstration of improved accuracy in classifying user behavior patterns over a variety of sample datasets. Future research endeavors may delve into further optimizing hyperparameters, exploring ensemble techniques, or investigating the adaptability of the proposed method across different social media platforms. Additionally, expanding the scope to include diverse data modalities and exploring interpretability techniques could contribute to the field growth.

References

1. Xue, Z., Li, Q., Zeng, X.: Social media user behavior analysis applied to the fashion and apparel industry in the big data era. J. Retail. Consum. Serv. **72**, 103299 (2023)
2. Narayanamoorthy, S., et al.: The COVID-19 vaccine preference for youngsters using promethee-ii in the ifss environment. Symmetry **13**(6), 1030 (2021)
3. Goswami, S.S., et al.: Analysis of a robot selection problem using two newly developed hybrid MCDM models of TOPSIS-ARAS and COPRAS- ARAS. Symmetry **13**(8), 1331 (2021)
4. Gong, Z., Lin, J., Weng, L.: A novel approach for multiplicative linguistic group decision making based on symmetrical linguistic chi-square deviation and vikor method. Symmetry **14**(1), 136 (2022)
5. Bao, G.Y., Lian, X.L., He, M., Wang, L.L.: Improved two-tuple linguistic representation model based on new linguistic evaluation scale. Control Decis. **25**(5), 780–784 (2010)
6. Akram, M., Naz, S., Smarandache, F.: Generalization of maximizing deviation and TOPSIS method for MADM in simplified neutrosophic hesitant fuzzy environment. Symmetry **11**(8), 1058 (2019)
7. Wang, J.X.: A MAGDM algorithm with multi-granular probabilistic linguistic information. Symmetry **11**(2), 127 (2019)
8. Xu, X.H., Wang, L.L., Chen, X.H.: Large group risky emergency decision-making under the public concern themes. J. Syst. Eng. **34**(4), 511–525 (2019)
9. Collins, B., Hoang, D.T., Dang, D.T., Hwang, D.: Method of detecting bots on social media. A literature review. In: Nguyen, N.T., Hoang, B.H., Huynh, C.P., Hwang, D., Trawiński, B., Vossen, G. (eds.) ICCCI 2020. LNCS (LNAI), vol. 12496, pp. 71–83. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-63007-2_6
10. Orabi, M., et al.: Detection of bots in social media: a systematic review. Inf. Process. Manage. **57**(4), 102250 (2020)