International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7434
***Article Received**: 16 May 2023 **Revised:** 04 July 2023 **Accepted:** 31 July 2023*

Recommendations in Social Media Applications to Ensure Personification and Safety using Machine Learning

Prema Gawade[1], Sarang Joshi[2]
[1]Department of Computer Engineering,
SCTR'S Pune Institute of Computer Technology, SPPU
Pune, India
prema.gawade12@gmail.com
[2]Department of Computer Engineering
SCTR's Pune Institute of Computer Technology, SPPU
Pune, India
sajoshi@pict.edu

Abstract—Myriads of social media utilization lead to various issues like personalization hacks, data security problems, and safety. A recommendation is of paramount importance to alleviate this problem when there is a huge amount of data and the number of participants on the platform is increasing exponentially. Unfortunately, modern social media research has enhanced the performance and personalization of recommendations in many fields, yet largely underutilizes the power of artificial intelligence to enable personalized recommendations system for social media platforms like WhatsApp, Facebook, Twitter, etc. With advancement inside the global of technology every hour and every day new features are delivered to the list. In a manner, social platforms are merging into our actual existence, and to achieve personification and related safety, users can get any one safety factor from all 6 classes with this approach. This factor provides the basis for personification and the implementation of safety precautions.

This research proposes recommendations for personification in social media applications. The proposed Modified Inception Resnet V4 Convolutional Neural Network (MInReCNN) outperforms embedded media persona analysis and classification through text, image, and video data. Using these prediction classes better decisions can be made in given social media domain.

Keywords- Machine Learning; Social Media; Personification; Safety; Recommendation.

I. INTRODUCTION

There are many ways where user information is used for different purposes without the permission of the user. Like, Law enforcement agencies are extracting proof to remedy the crime cases that happen due to information breaches, political events are using social platforms as a platform to promote their works and gain popularity to win the approaching elections, and businesses with the assistance of digital marketing they may be focusing onto sell their products and services via developing social campaigns and the employers utilizes the usage of the popular social platforms Facebook, Twitter to evaluate the social interest, LinkedIn to analysis the professional interest of the candidates [1].

This paper proposes a unified framework and evaluation for customized advice, including various context and area knowledge, building the personal model, and analyzing particular characteristics. This is the primary work that goals to have a look at personalized advice on social media subjects and offers a set of research studies and technology to benefit users in social media applications.

Prior art approaches in recommendations permit the categorization of user profiles and not the personalization of user data. Such traditional deterministic approaches are not useful as they can hamper the data quality, leading to critical privacy leakage. Besides, with the help of other approaches like traditional machine learning-based models, to build social relationships, it is very much necessary that the user share his traits, which might also result in extra privacy leakage. So, every user can keep his or her data private with the advent of a personification system. Supporting social media communication randomly with the least privateness leakage using machine learning techniques is important to address, as this will ensure the capabilities of ML can be used to improve personification and related safety.

II. LITERATURE SURVEY

Wherever The personification of users in embedded media can be seen in a variety of past art studies. Personification can be utilized for a variety of goals, including emotion recognition, marketing, and intent prediction [2]. Many researchers are motivated to gather pertinent data so that it can be examined and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7434
***Article Received**: 16 May 2023 **Revised:** 04 July 2023 **Accepted:** 31 July 2023*

applied to individual behaviour. In social media applications, there is a strong correlation between the persona traits and the online behavior, and personification of a user [3]. On the same line, Chang et. al. employed internet activity and location data to infer the persona [4], whereas to forecast the persona identification for certain languages, construct an order of sentences, and blocks of words using current machine learning and deep learning techniques [5]. Analysis of professional social networking sites revealed no significant relationship between personality traits and user profiles [6]. This in part raises the possibility that job experience does not improve one's ability to perceive a person's personality qualities, which primarily influence other social behaviour Su et. al. used deep learning architecture [7]. With the help of user syntactic representation, deliberately defined features, and grammatical annotations that were linked to dictionary information, the author constructed the features from social behaviour.

Understanding the importance of data posted in various forms and finding relationships with context is the purpose of employing a deep learning-based approach. The researcher used a value- and sequence association-based strategy to accomplish the same goal. In order to determine to mean, the value-based deep learning approach uses the frequentist approach, whereas the association-based approach recognizes the temporal relationships between the tokens. The research proposed a convolutional neural network-based approach to fetch embeddings of fixed length from the user data, embedding vectors include latent features to hide the identity representation of data but at the same time, it approximately derives the actual data [8]. According to the literature, determining word embeddings is a time-consuming operation because it necessitates the creation of an embedded feature vector store if the environment does not correspond to an environment of study.

An association-based technique is used to understand the tokens' temporal linkages. This research suggested a convolutional neural network-based approach for obtaining fixed-length embeddings from user input; embedding vectors are latent features that disguise the data's identity representation while also approximating the actual data. Machine learning models are well learnable with the knowledge which is provided to them through embedded media datasets and it supports the claims of subject matter experts in personification.

III. NATURE AND SCOPE OF WORK DONE

This study uses association linkages, text, image, and video data to assure personification and security in embedded media applications. The performance of the learning model was improved by patterns created by feature detection and analysis.

For safety-related judgments, it is required to characterize personification patterns within a specific domain area.

This design uses items, text data, and speech characteristics found in the data flow to analyze a person's persona. Additionally, the data embedding is put through more processing utilizing the feature engineering pipeline before being classified into one of the classes. Only objects that belong to the same category are eligible for connection requests and further communication.

Across a sample of people chosen at random, the evaluation of better personification is conducted.

Based on targeted persona research carried out using automated embedded media data analytics, the solution offers increased personification and safety when onboarding new employees and was highly recommended.

IV. METHODOLOGY

In view of every user profile, past chat history, and previous communications, recommendations can be provided for further communication [9][10]. For instance, Facebook.com is utilizing recommender frameworks for showing friends lists, groups, likes, comments, etc. At the point when a user signs in to the framework, it recommends friends like if they belong to the same community, organization, etc. And further to communicate with known or unknown persons, sending friends requests, accepting the request, sending messages, or giving comments should be from the same class.

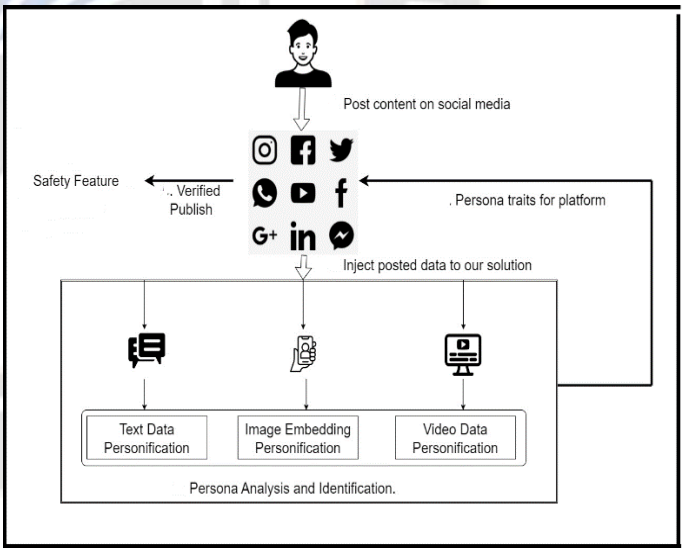

Fig. 1 Safety Recommendation System

For personification and safety decisions in social media applications, it is necessary to classify personification patterns and to decide which class a person belongs to. A strong belief is calculated using language descriptors and persona traits. These characteristics are then used in machine learning algorithms to understand underlying relationships by calculating the overlap of feature space. If both communicators belong to the same class,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7434
***Article Received**: 16 May 2023 **Revised:** 04 July 2023 **Accepted:** 31 July 2023*

then the safety factors are positive 1,2, and 3 otherwise negative 1,2, or 3 based on content filtering.

The steps as shown in Fig.1 are i) Post content on social media ii) Inject posted data into our solution iii) Persona analysis and identification based on text data, image embedding, or video data. iv) Forming of persona traits to check which class persona belongs to. v) Assigned to the specific class. vi) Personification based on content filtering or safety factor.

A. Feature selection through association linkage

For personification on social media, there are two significant issues to consider: association, which adds more personification, and integration of said association-based prediction with traditional attribute values. To evaluate the strength of personas in class there are four measures 1. Syntactical similarity 2. Semantic similarity 3. Linkage similarity 4. Hybrid Similarity [11].

$$HS(u, v) = SS(u, v) + CS(u, v) + IM(u, v) \quad (1)$$

A hybrid is retrieved to determine the relationship's strength and performance levels. Whenever top high variance associative latent characteristics are evaluated, features are selected and the effectiveness of the machine learning algorithm outperforms.

B. Feature selection through content

As embedded media data like text, image, and video data does not follow i.i.d. and contextual data is associated with data entities this may introduce a number of features. With respect to specific case studies and to achieve personification and safety, there is a need of focusing on context features. Association features are helpful in the reduction of the curse of dimensionality.

For content and text data depending on the case study and context, top *k* features from each cluster are extracted by applying a clustering algorithm to form eight different clusters. For images, data features are extracted using stagewise cascading and CNN architectures. Feature extraction is iteratively improved using weight adjustments. As object features are centralized in a specific area, the next epochs are performed in the same area. After feature extraction and using image descriptors personified feature vector is fed to the machine learning algorithm to form clusters.

The labels for clusters using top *k* features are assigned and tried to understand the differences. Labels are Cluster 0: Extroverts, Cluster 1: Introverts, Cluster 2: Sensors, Cluster 3: Intuitive, Cluster 4: Thinkers, Cluster 5: Feelers, Cluster 6: Judgers, Cluster 7: Perceivers [12][13].

To decide on communication six categories are employed, ranging from -3 to 3, with 0 being the lowest level. The -3 category describes the highest likelihood of ignorance or blockage, while 3 describes the likelihood of increased contact on social media platforms.

The probabilities of each sample prediction are sent to the function and then again packetized into six categories to enable these categories one of the first methods shows how previous communication history is connected to past communication history.

V. MATHEMATICAL MODEL

Let S be the solution perspective of the class for personification and safety factors such that,

$S = \{s, e, X, Y, Fme, DD, NDD, F_{friend}, CPU_{CoreCnt}\}$

were,

s: registration, communication in social media

e: safety factor or action function after feature selection and extraction.

X: text data, image data, or video data.

Y: safety factor +1 to +3 with positive and -1 to -3 with negative classes.

Fme:

1. Feature Identification through Association (FIA):
   To identify features in the context of personification research extracts an object's persona traits and association links [11].
2. Social Media Safety Perspective (SMSP):
   Analysis and classification of social media text data done for safety decisions using this research [12].
3. MInReCNN Modified Inception ResNetV4 for personification and safety.
   To achieve personification and safety in social media applications, persona classification is done using speech and image data [13].

DD: known personification features, class distribution, training data.

NDD: Personification features where if video data is there then for specific object functions like text, association features may NULL or may be extracted from image embeddings.

Ffriend: CNN architectures, Logistic Regression, KNN, Adaboost, etc. Machine Learning Algorithms.

CPUCoreCnt: Google Colab environment which has an Intel at 2.4GHz Corei7 CPU, 2 GB GPU, and 25 GB of RAM.

Success: desired personification and related safety.

VI. EXPERIMENT

Python 3.7.1 is used to implement all machine learning models, and the Jupyter Notebook Integrated Development Environment (IDE) is used to run them. It is the most well-known IDE for model development and supports the sci-kit-learn, TensorFlow,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7434
***Article Received**: 16 May 2023 **Revised:** 04 July 2023 **Accepted:** 31 July 2023*

and Keras libraries and packages for machine and deep learning. With 12 GB of RAM and a 2.4GHz Intel Core i7 processor, this data science workbench is installed on an Ubuntu computer.

The dataset of Deepfake Detection Challenge (DFDC) Casual Conversations is used to demonstrate and compare the effectiveness of several machine-learning methods. A total of 138 video files were used to train the model from which 3 sample files are shown in Fig. 2.

Fig. 2 Screenshot of sample video file

Text data is extracted from the video to form a feature vector and is combined with object, emotion, and action feature vectors. The machine learning algorithm is applied to these feature vectors to form different classes and get safety factors. The assessment of performance is made based on metrics such as contingency table, precision, recall, and accuracy.

VII. RESULT AND DISCUSSION

For better personification and safety Feature selection through association, social media application safety factors, and MInReCNN for personification and safety are tested for 138 videos of DFDC Casual Conversations. And videos are placed in clusters considering top *k* features extracted from videos and using machine learning algorithms and the following are accuracies observed.

TABLE 1. ALGORITHM AND ACCURACY

Algorithm Used	Accuracy
K-Means	92.65
DBSCAN	91.02
BIRCH	89.65

.

Key points as shown in Fig. 3 and top terms per cluster it is observed that clusters 1,2,3, have more users and this cluster includes personas that are more closely associated with professionalism.

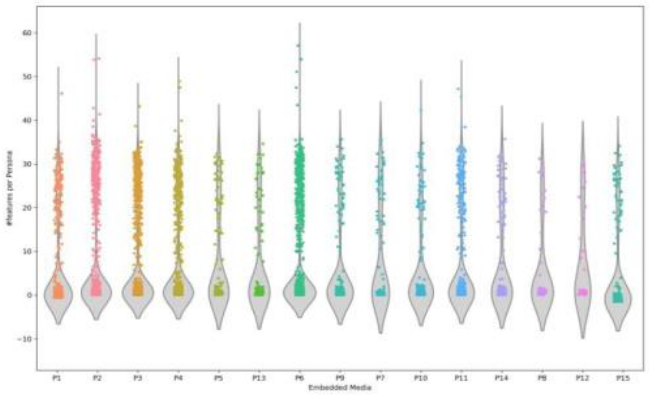

Fig. 3 Feature extraction for specific personas

Fig. 3 shows features extracted for personas from the video are centralized across the specific area and using those clusters are formed to detect for a new social media persona belongs to the same cluster or not.

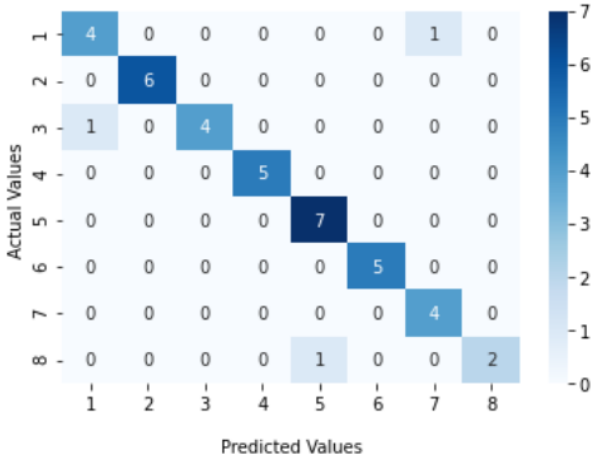

Fig. 4 Confusion Matrix for 40 test videos

Feature vectors of text data, object detection, emotion detection, and action detection are trained for 138 videos and then tested for 40 videos for eight different clusters.

Clusters sequentially from No.1 and 8 are Extroverts, Introverts, Sensors, Intuitive, Thinkers, Feelers, Judgers, and Perceivers, and the result from the above Fig. 4 shows 92 percent accuracy.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7434
***Article Received**: 16 May 2023 **Revised:** 04 July 2023 **Accepted:** 31 July 2023*

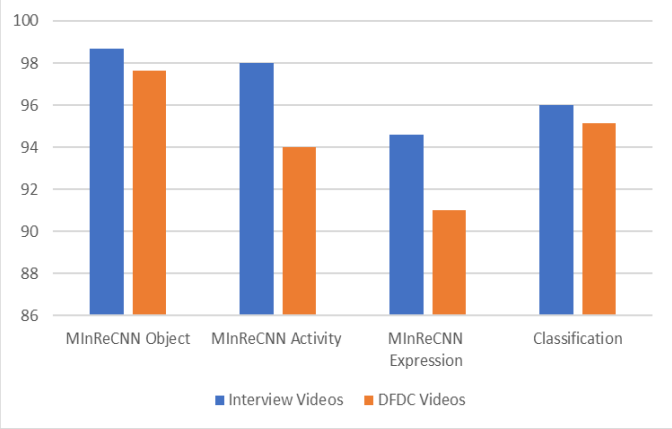

Fig. 5 Comparative analysis of results with different datasets

MInReCNN Modified Inception ResNetV4 is applied to both bunch of videos where stride is selected at the localized area of the object and this is given to action and expression recognition. Along with this audio is converted to text using NLP and classification is applied. For Interview videos and DFDC videos using MInReCNN gives better results and are as shown above in Fig. 5 where it is observed for both types of videos algorithm giving an accuracy of more than 95 percent and classification with different classes where if the object belongs to the same class can be said as positive for communication and if it belongs to different classes can be given with negative class meaning is not safer to have communication and helps in achieving better personification and safety.

VIII.CONCLUSION

This study of social media users with identifying the behavior of the persona based on embedded media context is designed to achieve safety in communication. Research plots a few difficulties and proposed an idea to address them. MInReCNN (Modified Inception Resnet V4 CNN) performs well for object detection, emotion detection, and action detection as compared to other Machine Learning algorithms. With 92 percent of accuracy, clusters are formed for different users where the safety factor either +1 to +3 or -1 to -3 is decided for better personification.

For future work, research exploration objectives are to execute this novel idea, use it to moreover to customize proposals, and assess it in a few spaces with an attention on the effect on user involvement with correlation with different strategies.

REFERENCES

[1] Koosha Zarei, Reza Farahbakhsh, and Noel Crespi.,"Deep dive on politician impersonating accounts in social media.", IEEE Symposium on Computers and Communications (ISCC), Barcelona, Spain, June 2019, PP:1-6.

[2] V. A. Fitri, R. Andreswari, M. A. Hasibuan., "Sentiment analysis of social media twitter with case of antilgbt campaign in Indonesia using naive bayes, decision tree, and random forest algorithm." Procedia CS, 2019, PP: 765-772.

[3] Wen Wu, Li Chen, and Liang He.,"Using personality to adjust diversity in recommender systems.", In Proceedings of the 24th ACM conference on hypertext and social media. ACM," 2013, PP: 225-229.

[4] Prof. Madhuri Zambre. (2016). Analysis and Modeling of Physical Stratum for Power Line Communication. International Journal of New Practices in Management and Engineering, 5(01), 08 - 13. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/42

[5] L. R, W. S and C. KC,"Multiple location profiling for users, relationships from social network and content, Proceedings of the VLDB Endowment, Iss. 5, vol.11," 2011, PP: 1603-1614.

[6] Oren B., Avi C., et.al," Attentive Item2vec: Neural Attentive User Representations". IEEE Proceedings, International Conf. on Acoustics, Speech, Signal Processing (ICASSP)"2020, PP: 3377-3381.

[7] Niels van de Ven, Aniek Bogaert, Alec Serlie, Mark J. Brandt, Jaap J.A. Denisen, "Personality perception based on LinkedIn profiles", Journal of Managerial Psychology, Vol. 32 Issue: 6, 2017, PP: 418-429.

[8] Majumder N., Poria S., et.al. "Deep Learning-Based Document Modeling for Personality Detection from Text.", IEEE Intell. Syst. 2017, 32, PP: 74-79.

[9] Denis Osipychev et.al.," Human intention-based collision avoidance for autonomous cars", IEEE American Control Conference (ACC), July 2017 ISSN: 2378-5861, PP: 2974-2979.

[10] Aravind Kota Gopalakrishna et.al.,"Evaluating machine learning algorithms for applications with humans in the loop" IEEE 14th International Conf. on Networking, Sensing and Control (ICNSC)DOI: 10.1109/ICNSC 2017.8000136. Aug. 2017, PP:1-9.

[11] Kumar, A. ., Reddy, M. P. ., Sushama, C. ., Limkar, S. ., Singh Kaswan, K. ., & Rathod, M. . (2023). Markov Clustering in Leaf Disease Detection Based on Classification using Probabilistic Naïve Bayes Regression for Deep Learning Architecture. International Journal of Intelligent Systems and Applications in Engineering, 11(3s), 184–188. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2559

[12] Prema Gawade,"A Survey on Data Analytics for Personification using Machine Learning", Vol.-7, IJCSE, E-ISSN: 2347-2693, March 2019, PP.18-21.

[13] Prema Gawade, Sarang Joshi., "Feature Selection for Embedded Media in the Context of Personification", ICIRCA proceedings in IEEE, ISBN: 978-1-7281-5375-9, July-2020, PP: 590-594.

[14] Prema Gawade, Sarang Joshi, "Personification and Safety during pandemic of COVID19 using Machine Learning," Nov-2020, ICECA, Coimbatore, proceedings in IEEE, doi:10.1109/ICECA49313.2020.9297555., Dec-2020, PP: 1582-1587

[15] Rossi, G., Nowak, K., Nielsen, M., García, A., & Silva, J. Enhancing Collaborative Learning in Engineering Education

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9s
DOI: https://doi.org/10.17762/ijritcc.v11i9s.7434
***Article Received**: 16 May 2023 **Revised:** 04 July 2023 **Accepted:** 31 July 2023*

with Machine Learning. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/119

[16] Prema Gawade, Sarang Joshi, "Persona Identification based on Social Media Profile Images for Personification and Safety", Webology, Volume 19, Number 1, January, 2022,DOI: 10.14704/WEB/V19I1/WEB19087, PP: 1297-1314.

[17] A. Alsavat, H. EI-Sayed, "Social media analysis using optimized K-Means clustering",June 2016, IEEE Int. Conf. SERA, doi: 10.1109/SERA.2016.7516129, PP: 1-6.

[18] P. Sinha, L. Dey, et.al., "A Hierarchical Clustering Algorithm for Characterizing Social Media Users", In Companion Proceedings of the Web Conference 2020 (WWW '20). Association for Computing Machinery, USA, PP: 353–362.