Nano World Journal

https://doi.org/10.17756/nwj.2023-s4-081

Research Article

Open Access

Enhanced Automatic Recommender System: Leveraging Sentiment Analysis and Deep Learning

Lingam Sunitha*, **Kambam Shreya, Baridula Umamaheshwar Reddy and Godavarthi Yashwitha**

Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, Telangana, India

***Correspondence to:**
Lingam Sunitha
Department of Computer Science and Engineering,
Vardhaman College of Engineering,
Hyderabad, Telangana, India.
E-mail: sunithavvit@gmail.com

Received: September 19, 2023
Accepted: December 04, 2023
Published: December 07, 2023

Citation: Sunitha L, Shreya K, Reddy BU, Yashwitha G. 2023. Enhanced Automatic Recommender System: Leveraging Sentiment Analysis and Deep Learning. *NanoWorld J* 9(S4): S483-S487.

Published by United Scientific Group

Abstract

Online social networks offer helpful information about people's perspectives on a variety of topics. These data are gathered and evaluated by applications such as monitoring and recommendation systems (RS). The advanced wellbeing observation component of the knowledge-based recommendation system (KBRS) described in this work can help detect clients who might be stressed out or depressed. The KBRS, which is activated based on ontologies and sentiment analysis, sends signals of joy, relaxation, soothing, or motivation to people who are experiencing psychological challenges depending on the outcomes of the monitoring. A means for alerting appropriate parties is also part of the solution in case the monitoring system detects a depressed issue. The suggested strategy successfully distinguished between unhappy and focused clients with a precision of 0.89 and 0.90. Using a convolutional brain organization and bidirectional long short-term memory (BLSTM) - recurrent neural networks (RNN), depressing and upsetting sentences are detected. The recommended KBRS scored 94% of consumers who were highly delighted, according to trial results, compared to 69% for a RS that didn't use sentiment metrics or ontologies. Furthermore, results from random tests revealed that the proposed method makes only modest demands on the processing, memory, and energy of modern mobile electronic devices.

Keywords

Knowledge-based recommendation system, Recommendation systems, Deep learning, Recurrent neural networks

Introduction

Integrating an automatic RS based on sentiment analysis, deep learning, and nanotechnology within an online social network can profoundly impact user engagement, content recommendations, and personalized experiences. An automatic RS leveraging sentiment analysis, deep learning, and nanotechnology represents an innovative fusion at the forefront of technology-driven recommendations. The integration of sentiment analysis, deep learning, and nanotechnology in an automatic RS represents a cutting-edge approach. This convergence promises not only more accurate and personalized recommendations but also addresses concerns regarding data security, making it a transformative innovation in recommendation technology.

Users of online social networks (OSNs) have substantially expanded, and by the year 2020, there may be 2.95 billion OSN members globally. The large clientele of OSN is typically attributed to the increase in the number of portable web-connected devices, including mobile phones and tablets. OSN has evolved into a multifaceted and global platform for the expression of thoughts, feelings, unhealthy habits, and good behaviours. Recently, several applications in the field of medical informatics have benefited from the analysis of messages posted on

OSN. Negatively spelled words can convey sorrow, anxiety, or dissatisfaction in phrases. In order to strengthen the role of the citizen as a collaborator rather than a client in the delivery of public services, social media's contribution to public administration is crucial [1]. Similar to this, words with positive meanings might denote joy, happiness, or satisfaction. On the other side, if a person is in a good mood, it can be thought that they are more emotionally steady and self-assured. Stress is a normal, everyday aspect of life, but excessive or persistent stress can be bad for one's physical and mental health. According to published research studies, chronic stress has been linked to a variety of illnesses, including severe depression, sleeplessness, etc. [2]. Customers behave differently on OSN, and if the emotional intensity of the conveyed expression consistently fluctuates between low and high levels, or the opposite, this may be a sign of a serious problem like bitterness or traumatic occurrences. Depression is one of the most prevalent mental illnesses in the world, affecting people of all ages and socioeconomic backgrounds. Unfortunately, not many individuals can still recognize sadness. Most research on health systems uses sensor technologies to spot mental health issues [3]. A RS application may be used to improve an individual's sentimental well-being and mood while they are suffering unpleasant emotional states [4]. In the healthcare sector, ontology-based RS is used to provide precise results from sickness treatment approaches (Figure 1).

The purpose of this project is to propose a RS that uses the KBRS approach. This method covers a wide range for health-related situations that isn't addressed by current proposal frameworks. The main objective of this project is to offer an RS that adheres to the notion of a KBRS. For the planned KBRS, a thorough health monitoring system and an opinion examination approach are also recalled. In order to identify potential users who could be experiencing stress or depression, the monitoring system examines OSN utterances. An optimal approach based on a BLSTM-RNN is taken into consideration to differentiate potential mental diseases from CNN (Convolutional neural network) in order to achieve this objective. Then, messages for cheering up, relaxing, having fun, or rousing these folks are sent utilising a KBRS. Potential mental illnesses are identified using an ideal method based on CNN and the BLSTM-RNN. After that, a KBRS is used to send these people encouraging, cheerful, calming, or loosening

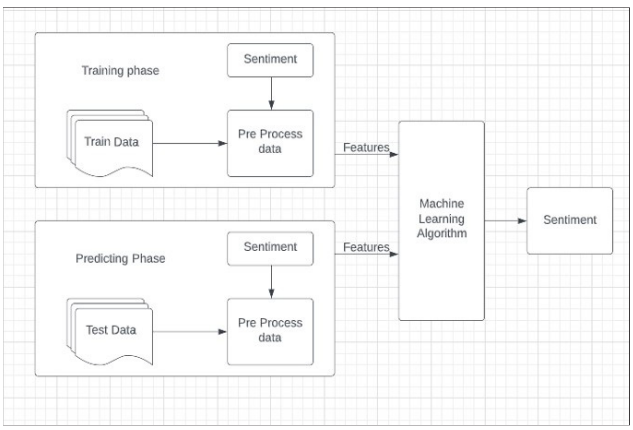

Figure 1: Entity relationship diagram.

up signals. The force degree of these exchanges is determined by the opinion power of the phrases put on an OSN, as determined by a superior feeling investigation metre known as eSM2.

Literature review

Hancock et al. [5] and Liu [6], discovered that people who are depressed tend to write in fragments. Additionally, these individuals have a history of severe sleeplessness and employ the first-person pronoun in their utterances. Thus, the statements posted on OSN may be used to infer information about their conduct. In order to identify users who are very likely to attempt suicide, particular terms in the phrases can be tracked and evaluated. Then, the right intervention can be made.

Using a mix of LSTM (Long short-term memory), and Conditional Random Field, a CNN, Ma and Hovy [7] offer network model to assess sentences meaning through character-level presentations. Zhang et al. [8] achieved an average accuracy of 80% when classifying emotions from microblogs with an emphasis on psychiatric diseases using several machine learning (ML) classifiers. The suggested approach, which relied on data from Twitter activity to detect stress, had a 69% accuracy rate. Using data from OSN, authors investigate the causes of postpartum depression. Studies on mood monitoring systems analyse signals from online networks using ML algorithms, and they achieve an accuracy of 57%.

RNNs and Conditional Random Fields are combined by Rosa et al. [9] to achieve the greatest results on Named Entity Recognition datasets. An improved version of the LSTM called the BLSTM is commonly used for labelling tasks.

Experimentation

Existing system

All currently available applications employ classic ML methods like Random Forest or Support Vector Machine to identify user sentiment.

Disadvantages of existing system

Algorithm accuracy, which is 69%, is not improved. They don't keep track of user personal information like personal profiles to provide inspirational ontologies.

Proposed system

To identify stressed and depressed users, the suggested approach has an accuracy of 0.89 and 0.90, respectively. According to experimental findings, the suggested KBRS received a 94% extremely happy user rating. Maintain all user information, including age, sleep schedule, and personal or professional biography. We employ the algorithms CNN, RNN, and BLSTM [10, 11].

The recommended strategy has an accuracy of 0.88 and 0.91 for identifying stressed and depressed users, respectively. The recommended KBRS obtained a 94% highly satisfied user rating, according to trial results.

Keep track of all user details, such as age, sleeping habits, and a personal or professional biography. The algorithms CNN, RNN, and BLSTM are used.

Deep learning

However, profound learning is a subset of ML whereas artificial intelligence is a subset of man-made awareness. Computerized reasoning is a general term used to describe methods that enable personal computers to mimic human behavior. All of this is made possible by artificial intelligence, a collection of computations that are built using data.

Handling power was a constraint on the complexity of brain networks. Because of improvements in big data analysis, which let computers to monitor, understand, and respond to complicated events more quickly than people, larger, more jumbled brain networks are now possible. Deep learning has benefited discourse recognition, language interpretation, and picture arranging. It allows for the resolution of any example acknowledgment issue without the need for human connection (Figure 2).

Contrarily, the association of the human mind was the driving force behind only a small portion of ML, known as deep learning. In order to achieve conclusions that are equivalent to those made by individuals, deep learning computations continuously examine data involving a predetermined sensible system. Deep learning does this using several types of computations and brain organisations.

The design of the human brain served as the inspiration for the arrangement of the brain. Brain structures may be able to comprehend patterns and organise a variety of facts similarly to how they would ordinarily.

The supporting pieces make up the framework. Data set created using data gathered from OSNs includes client profile and client information. A database contains 360 messages, with 90 of each category (calming, inspiring, upbeat, or tranquil messages) available to the user. Users have previously had the option to choose one or two text styles when they are going through a stressful or sad period. The messages were written by three psychologists, and they were approved by another three psychologists.

To assess whether the sentences collected from OSN were written when sad or under stress, ML was utilised to filter the sentences. The system uses it to monitor emotional health.

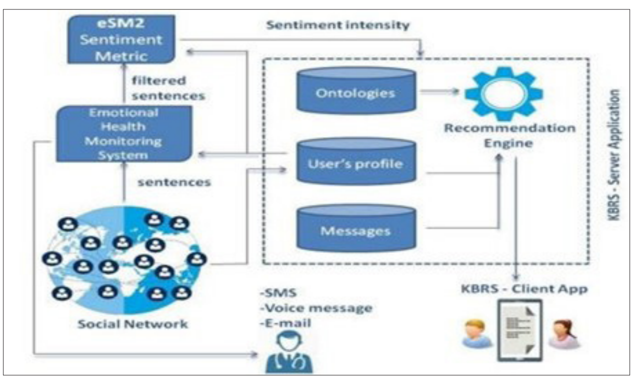

Figure 2: System architecture [10, 11]

eSM2's sentiment analysis filters the phrases and assigns them a sentiment score on a scale of 0 to 5. Earlier studies have looked into and validated this spectrum. Depending on how powerfully the statement conveys emotion, the message's potency will vary. The three message levels are extreme, medium, and lower. If the monitoring system detects that the user is very/deeply frightened, it will send them a highly reassuring message. Examples of exceedingly positive expressions include the adverbs greatly, tremendously, powerfully, among others. Additionally, the message type is consistent with the preference discovered after testing in the first phase [6].

The recommendation motor uses ontologies, which are a set of classes, things, and links. The metaphysics web language is used to convey ontologies. It is possible to extract the data for each class using OSN. Scenarios involving health are utilised with the Nuadu ontology collection [12]. The project made use of the following Nuadu classes.

- Personal ontology: Reflects the private data that is stored about each individual, such as their gender, profession, and preferences.
- Activity ontology: The person's actions are fully described. It's possible from the records that the person's routine has altered.
- Sleep ontology: Explains user routines and behaviours.
- Risk ontology: Information about bad behaviours that have been related to stress and the onset of illnesses, such smoking and drinking alcohol.
- Context ontology: Explains the surrounds of the subject (home, school, work, or travel). This information is crucial because it can explain why activity records were absent for a while or why sleeping habits changed.

A recommendation engine is the procedure that generates a list of recommendations.

Implementation

Because no one is available to produce OSN phrases for this project, I'm directly using the OSN dataset from Twitter to train the BLSTM and Random Forest train models, evaluate their performance, and forecast sentiment from fresh test messages. The creator of this project hired accessors and made them type phrases on OSN networks in order to create his dataset. He then trained the BLSTM and Random Forest algorithms using those phrases.

Modules

The project includes the following modules.

- Upload OSN dataset: The dataset will be uploaded to the application using this module.
- Generate train and test model from OSN dataset: With the use of this module, we will read every message in the dataset and create a train and test model using the characteristics we've discovered.
- Build CNN BLSTM-RNN model using Softmax: This lesson will assist us in building a deep learning BLSTM

model using a dataset, and we'll use test data to evaluate the model's performance.

- Run Random Forest algorithm: This method is also being used to evaluate the accuracy of random forest and BLSTM.
- Upload test message and predict sentiment and stress: When we send test messages using this module, the program will detect stress by utilizing the BLSTM model to test data.
- Accuracy graph: By doing this, a graph of accuracy for BLSTM and random forest will be produced.

Algorithms

The algorithms CNN, RNN, and BLSTM are used in this research.

CNN

Using learnable weights and biases, a CNN may give different objects and elements in an input image a value before being able to distinguish between them. CNNs require significantly less pre-processing than other classification methods. Contrary to basic techniques, which need hand-engineering filters, CNNs can learn these features and filters. The configuration of the visual cortex has an impact on the architecture of a CNN, which is comparable to the connecting network of neurons in the human brain. Individual neurons are only sensitive to inputs in the limited region of the visual field known as the Receptive Field. There are numerous overlapping fields like this that encompass the entire visual field.

RNN

An imaginary brain network known as an intermittent brain organisation sets up connections between hubs into a coordinated or undirected chart along a transitory succession (RNN). As a result, it might briefly exhibit a strong pattern of behavior. Using their internal state, RNNs, which are created from feedforward neural networks, can accommodate input sequences of different lengths (memory). Since recurrent neural networks are Turing complete, they can theoretically run any programme to handle any input sequence. They can therefore be useful in tasks like linked, unsegmented handwriting recognition and speech recognition. RNNs have an endless impulse response, in contrast to CNNs' finite impulse response. Both varieties of networks have temporally varying behavior. Finite impulse recurrent networks are directed acyclic graphs that can be unrolled and swapped out for strictly feedforward neural networks, as opposed to infinite impulse recurrent networks, which are directed cyclic graphs that cannot be unrolled.

BLSTM networks

It's conceivable that the production will also be influenced by subsequent productions. This is especially true for applications like natural language processing, where the traits of the word or phrase we are trying to predict may depend on the context supplied by the entire surrounding sentence, not just the words that came before it. A BLSTM, which is effectively two RNNs placed on top of one another that reads the input from left to right and right to left, can be used to address this problem. Each time step's output will be based on the concealed state of both RNNs. With the use of bidirectional RNNs, the network may provide the start and end of the sequence similar weights, which frequently improves performance.

Results and Discussion

Dataset description

The considered dataset has 504 sample records, and each record is having three specifications namely tweet, emoticon and sentiment. Tweet denotes message, emoticon is the emoji used by the client and the last one is sentiment of him. Of which these columns, tweet and emoticon are independent variables and sentiment is a dependent variable. The possible values for sentiment variable are 0,2,4 which denotes deeply stressed, stressed and happy respectively. This is the considered sentiment metric.

Comparison of algorithms

Table 1, table 2, table 3, and figure 3 present the observed results.

Table 1: Table for comparison of accuracy of Random Forest and BLSTM algorithms.

Algorithm	Accuracy
Random Forest	69%
BLSTM	94%

Table 2: Hyperparameters and their values used in BLSTM algorithm.

Hyperparameter name	Value
Number of units in a dense layer	3
Dropout	20%
Activation function	SoftMax
Loss function	Categorical cross entropy
Number of epochs	7
Batch size	32

Table 3: Hyperparameters and their values used in Random Forest algorithm.

Hyperparameter name	Value
Number of estimators	10
Max. depth	2
Random state	4

Conclusion

The eSM2 was designed to better understand user profile factors, geographic location, and sentence theme in order to estimate a message's sentiment intensity. The sentiment metrics that are currently in use do not account for these two factors. When compared to the results of the eSM2 and eSM measures, the findings of the eSM2 in the perceptual assessment of the RS were superior. This demonstrated how

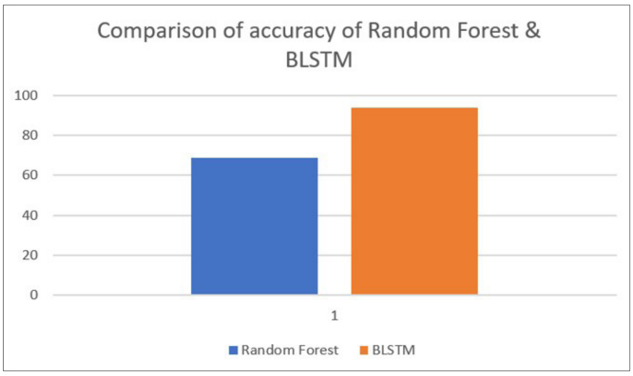

Figure 3: Graph for comparison of accuracy of Random Forest and BLSTM algorithms.

crucial it is to incorporate additional user profile variables in order to improve the sentiment measure. The suggested KBRS also made use of the ontology idea. There aren't many studies out there right now that employ OSN data to pinpoint stress levels. Using the BLSTM-RNN for illness entity detection and the CNN for character level representation, an accuracy of 0.89 and 0.90, respectively, was achieved by the solution for monitoring the sad or stressed state of OSN users. These accuracy levels are significant compared to those from comparable studies. The suggested KBRS was put up against another KBRS whose presenting assessment tests don't take an emotion measure or metaphysics into account. Results show that the proposed KBRS performs better than the RS without opinion measure and philosophy by obtaining 94% and 69% of clients who are extremely satisfied, respectively. Customers guarantee, using more traditional and non-redid data, that an RS that doesn't take philosophy and a feeling metre into account works ineffectively. The KBRS's finest output demonstrates the value of using an ontology and, more particularly, a focused sentiment analysis as opposed to a generic sentiment analysis. The most significant contribution of the study is that the relevant users often experienced an improvement in their emotional state as a result of the recommended messages that were sent to them.

Acknowledgements

None.

Conflict of Interest

None.

References

1. Glavan IR, Mirica A, Firtescu BN. 2016. The use of social media for communication in official statistics at European level. *Roman Stat Rev* 64(4): 37-48.
2. Lin H, Jia J, Qiu J, Zhang Y, Shen G, et al. 2017. Detecting stress based on social interactions in social networks. *IEEE Trans Know Data Eng* 29(9): 1820-1833. https://doi.org/10.1109/TKDE.2017.2686382
3. Weiner IB, Greene RL. 2017. Handbook of Personality Assessment. John Wiley & Sons.
4. Al-Qurishi M, Hossain MS, Alrubaian M, Rahman SMM, Alamri A. 2017. Leveraging analysis of user behavior to identify malicious activities in large-scale social networks. *IEEE Trans Ind Inf* 14(2): 799-813. https://doi.org/10.1109/TII.2017.2753202
5. Hancock JT, Gee K, Ciaccio K, Lin JMH. 2008. I'm sad you're sad: emotional contagion in CMC. In Proceedings of the ACM Conference on Computer Supported Cooperative Work, San Diego, CA, USA.
6. Liu B. 2017. Many Facets of Sentiment Analysis. In Cambria E, Das D, Bandyopadhyay S, Feraco A (eds) A Practical Guide to Sentiment Analysis. Socio-Affective Computing. Springer, Cham, pp 11-39.
7. Ma X, Hovy E. 2016. End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF. *arXiv preprint* 1-10. https://doi.org/10.48550/arXiv.1603.01354
8. Zhang Y, Xu C, Li H, Yang K, Zhou J, et al. 2018. HealthDep: an efficient and secure deduplication scheme for cloud-assisted eHealth systems. *IEEE Trans Ind Inf* 14(9): 4101-4112. https://doi.org/10.1109/TII.2018.2832251
9. Rosa RL, Schwartz GM, Ruggiero WV, Rodríguez DZ. 2018. A knowledge-based recommendation system that includes sentiment analysis and deep learning. *IEEE Trans Ind Inf* 15(4): 2124-2135. https://doi.org/10.1109/TII.2018.2867174
10. Rosa RL, Rodriguez DZ, Bressan G. 2015. Music recommendation system based on user's sentiments extracted from social networks. *IEEE Trans Consumer Electron* 61(3): 359-367. https://doi.org/10.1109/TCE.2015.7298296
11. Rosa RL, Rodríguez DZ, Schwartz GM, de Campos Ribeiro I, Bressan G. 2016. Monitoring system for potential users with depression using sentiment analysis. In IEEE International Conference on Consumer Electronics, Las Vegas, NV, USA.
12. World Health Statistics 2016: Monitoring Health for the SDGs, Sustainable Development goals. [https://www.who.int/publications/i/item/9789241565264] [Accessed December 05, 2023]