

Computational Linguistics and In-telligent Systems (COLINS-2021), Kharkiv, Ukraine, April 22-23, 2021. – CEUR Workshop Proceedings, 2021, 2870, Volume I, pp. 916-936

2. Оцінювання дистанційного навчання URL - <https://profcenter.com.ua/otsinyuvannya-dystantsijnogo-navchannya/> (дата звернення: 29.11.2023).

3. Моделі дистанційного навчання URL - <https://op.ua/ru/pedclass/nauchnaya-statya/modeli-distanciynogo-navchannya> (дата звернення: 29.11.2023).

RECURRENT NEURAL NETWORK MODEL ARCHITECTURE FOR DETECTING A TENDENCY TO ATYPICAL BEHAVIOR OF INDIVIDUALS BY TEXT POSTS

Olena Sobko

Teacher

olena.sobko.ua@gmail.com

Oleksandr Mazurets

Ph.D in Engineering Science, Associate Professor

exe.chong@gmail.com

Volodymyr Didur

Bachelor student

mamont.m6@gmail.com

Illia Chervonchuk

Bachelor student

illia.chervon4@gmail.com

Computer Science Department

Khmelnytskyi National University, Ukraine

With the growing popularity of social networks, people are giving away more personal information in the online environment. This creates a unique opportunity to apply neural network methods to analyze textual content and identify potentially atypical behavior.

Social media has become not only a platform for communication and information exchange, but also a space where atypical behavior can occur for many reasons. Social media is a platform for atypical behavior, as anonymity and remoteness allow some people to express their thoughts and actions without traditional restrictions. The lack of physical contact and distance from the real world make people less responsible for their actions [1].

Social networks have become a kind of virtual space where people willingly share their inner state, problems and experiences for various reasons. People, especially in today's virtual world, feel the need to interact and communicate, and

social media provides an opportunity to see and feel interaction with others, even if they are far away – sharing their experiences becomes a means of support and mutual assistance. With the help of social media, people can seek recognition and support from their environment. Sharing a personal life profile can be aimed at gaining support, advice, or even just acknowledgment of the importance of one's existence. For many people, sharing their problems and experiences in the virtual space is similar to the process of psychotherapy – getting feedback, advice, or just a loyal listener can have a positive impact on mental health [2].

The task of detecting a tendency to atypical behavior based on a user's text post is important and relevant for several reasons. First, the Internet and social networks have become an important part of modern communication, where people express their thoughts, feelings, and interests. In such a virtual space, situations may arise when users exhibit atypical behavior that can affect their safety and comfort, as well as create a negative impact on others [3].

Automated detection of signs of atypical behavior in text messages is an important task of information technology aimed at ensuring security and ethical communication in the digital environment [4]. Atypical behavior can include aggressive statements, depressive and suicidal thoughts, attempts at provocation, spreading false information, or offensive comments [5]. Therefore, automated detection of atypical behavior is important for providing timely support to people who may be experiencing emotional stress or other difficulties. Providing support can help reduce the risk of conflicts and help improve the emotional state of participants in virtual communication, so it is a relevant area of information technology.

The aim of this work is to develop a recurrent neural network model architecture for detecting a tendency to atypical behavior of individuals by text posts, and its practical implementation.

Recurrent neural network model architecture is used in the method for detecting a tendency to atypical behavior of individuals by text posts. The method for detecting a tendency to atypical behavior of individuals by text posts is dedicated to improving the detection of a tendency to atypical behavior by a user's text post by analyzing it with neural network tools. The input data of the method for detecting a tendency to atypical behavior by text posts are a text post, a trained neural network model for detecting a tendency to atypical behavior, and a trained neural network model for determining the type of mental disorder that affects atypical behavior.

The first step is to load the trained neural network models. The method uses two trained neural network models: to detect a tendency to atypical behavior and to determine the type of mental disorder that affects atypical behavior.

The second step is to translate the text message into English. It takes place in two stages, the first of which is to determine the current language of the post, and then to translate from the currently detected language to English.

The next step is pre-processing the user's text message, which includes removing stop characters and stop words, as well as the tokenization process.

The fourth step is a neural network assessment of the propensity for atypical behavior. This assessment is binary. If the assessment is positive (there is a tendency

to atypical behavior), step five takes place; if the assessment is negative, step five is not performed.

At step five, a neural network assessment of the type of mental disorder that affects atypical behavior is performed. The paper presents ten types of possible behavioral disorders.

The sixth step is the formation of a conclusion for the user, which includes an assessment of the predisposition to atypical behavior, as well as the type of possible psychological disorder that can lead to atypical behavior. The output of the method is a numerical assessment of the propensity for atypical behavior and a certain type of mental disorder.

Thus, the method for detecting a tendency to atypical behavior of individuals by text posts works by converting input data in the form of a text post, a trained neural network model for detecting a tendency to atypical behavior, and a trained neural network model for determining the type of mental disorder that affects atypical behavior into output data in the form of a numerical estimate of the tendency to atypical behavior and the determined type of mental disorder.

The method uses two recurrent neural network models: to detect a tendency to atypical behavior and to determine the type of mental disorder.

The model for detecting a tendency to atypical behavior is a typical RNN model with two outputs. The architecture of the neural network model is shown in Figure 1.

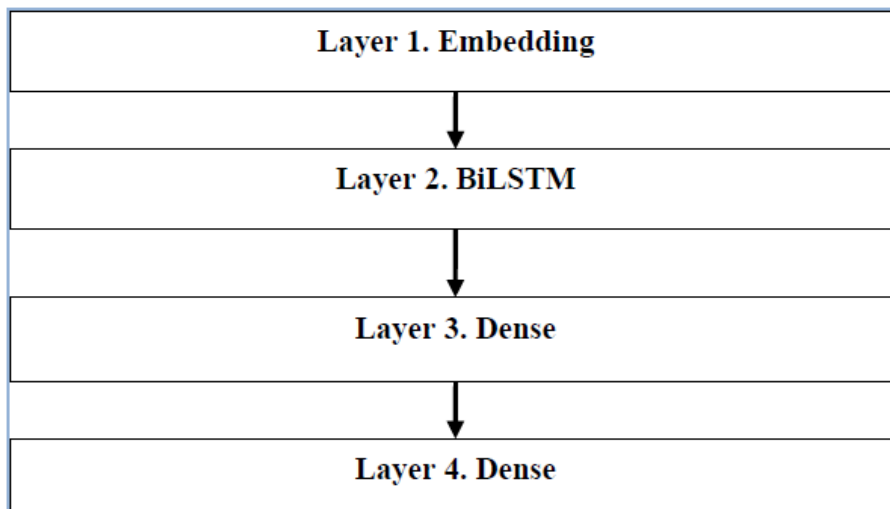


Figure 1. Recurrent neural network model architecture for detecting tendencies to atypical behavior.

The first layer is the Embedding layer, which transforms words into fixed-length vectors to help the model understand the semantic meaning of words and their relationships in context.

The next layer is the BiLSTM, which is a modified version of a conventional LSTM layer that processes input sequences in both directions using two separate LSTM layers, each containing 128 neurons.

In the penultimate layer of Dense, each input is connected to each output, covering all the data from the previous layer. For the purposes of this study, this layer

contains 128 neurons. This means that there will be 128 neurons in the layer with weights that will be adjusted during training.

The final layer of the neural network, like the previous one, is a Dense layer. Since it completes the model, the softmax activation function is used. The number of neurons in this layer corresponds to the number of unique classes in the training dataset, which in this study is 2 (there is a tendency to atypical behavior / no tendency to atypical behavior). The softmax function generates a probability vector for each class, where each value indicates the probability that the input sample belongs to a particular class.

The model for determining the type of mental disorder that may affect the tendency to atypical behavior is also an RNN model, but given the work with ten classes, its architecture is more multilayered and complex. The architecture of the neural network model is shown in Figure 2. Unlike the previous model, here, in addition to the usual layers, there are also dropout layers.

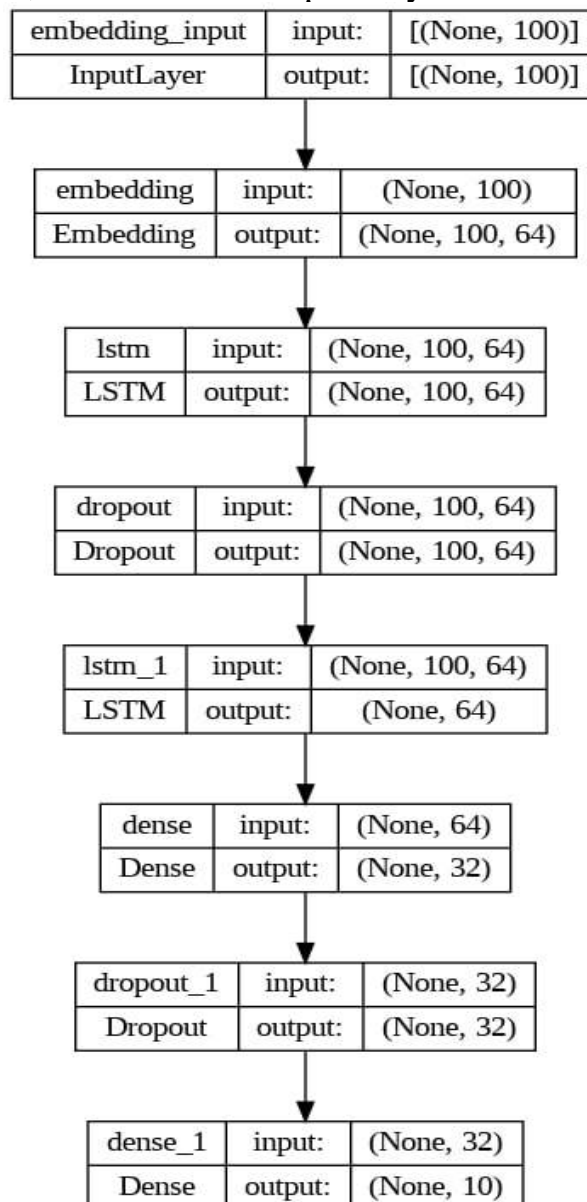


Figure 2. Recurrent neural network model architecture for detecting the type of mental disorder.

The model starts with the Embedding layer, which converts the input words into fixed-length vectors, allowing the model to better understand the semantic meaning of words and their contextual relationships. The input size of the dictionary is 5,000, the output dimension of the vectors is 64, and the length of the input sequence is determined by the `max_length` parameter, which can be experimentally changed.

This is followed by a 64-neuron LSTM layer that processes the data sequence and returns the full sequence of results, allowing subsequent layers to have access to all time steps. To prevent overtraining, a Dropout layer is used, which randomly discards 50% of the neurons.

The next LSTM layer also has 64 neurons, but unlike the previous one, it only returns the last initial state, which represents a general understanding of the sequence. This is followed by the Dense layer with 32 neurons and ReLU activation, which provides nonlinearity and helps the model learn more complex relationships. To further reduce the probability of overlearning, another Dropout layer is used with the same probability of 0.5.

The final Dense layer has as many neurons as there are unique classes in the training data. In this case, their number is determined by the size of the set of class labels `y_train`, which is 10. Using the softmax activation function in this layer allows you to get a probability vector for each class, where each value represents the probability of the input sample belonging to a particular class.

Thus, we present the architectures for two recurrent neural network models that are the input data of the method for detecting a tendency to atypical behavior in a text message. The architecture of the RNN model for detecting a tendency to atypical behavior has 2 outputs responsible for the presence of a tendency to atypical behavior or the absence of a tendency to atypical behavior, and the architecture of the RNN model for determining the type of mental disorder provides 10 outputs, each of which corresponds to the type of mental disorder that may affect the tendency to atypical behavior.

References

1. Living by example. How Is Social Media Changing Human Behavior? URL: <https://www.livingbyexample.org/how-is-social-media-changing-human-behavior/>
2. Sirivisa creative. Sharing Secrets on Social Media. URL: <https://www.sirivisacreative.com/blogs/sharing-secrets-on-social-media>
3. M. M. Rock. Social Media and Behavioral Health: Normal and Abnormal Behaviors. Behavioral Pediatric Healthcare for Nurse Practitioners. 2018. URL: <https://study.com/learn/lesson/abnormal-behavior-examples-criteria.html>
4. Nazarov V., Molchanova M. Information System for Detecting Abusive Speech in Audio Content by Means of Natural Language. Proceedings of V International Scientific and Practical Conference «Modern strategies of global scientific solutions». December 27-29, 2023. Stockholm, Sweden, International Scientific Unity. Pp. 132-135.
5. Zalutskaya O., Molchanova M., Sobko O., Mazurets O., Pasichnyk O., Barmak O., Krak I. Method for Sentiment Analysis of Ukrainian-Language Reviews in E-Commerce Using RoBERTa Neural Network. CEUR Workshop Proceedings, 2023, vol. 3387, pp. 344-356.