

ABSTRACT

A NEURAL NETWORK-BASED METHOD FOR CLASSIFYING DEPRESSION SEVERITY LEVELS ON INDONESIAN TWITTER USERS

Classifying the depression severity levels of Twitter users could help medical professionals to detect users in the early stage of depression. Due to the vast amount of data generated by Twitter users, it is impractical to analyze them manually. Hence, machine learning has been extensively studied in the literature. BDI is one of the widely used metrics to evaluate the depression severity level of a person. BDI divides a person's depression severity level into six levels, from normal to extreme depression. Hence, predicting the severity level of Twitter users can be considered an ordinal classification problem, which deep learning techniques can solve.

In particular, Convolutional Neural networks (CNN) and Long Short Term Memory (LSTM) are among the most advanced techniques to extract hidden features in tweets. This research compared the performance of these two architectures for classifying depression severity levels based on BDI. A survey involving more than 1,000 Twitter users was conducted to obtain their BDI levels, then their tweets in the past three months were collected and used as the dataset to train the models. Finally, CNN and LSTM were compared using the Mean Absolute Error (MAE).

On average, LSTM was slightly better than CNN, with an average MAE of 1.313 ± 0.09 and 1.382 ± 0.04 , respectively. However, LSTM outperformed CNN when categorizing borderline level depression (Level 3) to extreme level depression (Level 6), where the data on these levels are significantly smaller than those of the lower levels (Levels 1 and 2). This emphasizes the potential of LSTM for detecting early depression in Twitter users.

Keyword: depression, ordinal classification, convolutional neural network, long short term memory, consistent rank logits

INTISARI

METODE BERBASIS *NEURAL NETWORK* UNTUK MENGKLASIFIKASIKAN TINGKAT KEPARAHAN DEPRESI PADA PENGGUNA *TWITTER* INDONESIA

Mengklasifikasikan tingkat keparahan depresi pengguna Twitter dapat membantu profesional medis untuk mendeteksi pengguna pada tahap awal depresi. Karena banyaknya data yang dihasilkan oleh pengguna Twitter, tidak praktis untuk menganalisisnya secara manual. Oleh karena itu, pembelajaran mesin telah dipelajari secara ekstensif dalam literatur. BDI adalah salah satu metrik yang banyak digunakan untuk mengevaluasi tingkat keparahan depresi seseorang. BDI membagi tingkat keparahan depresi seseorang menjadi enam tingkatan, dari depresi normal hingga ekstrim. Oleh karena itu, memprediksi tingkat keparahan pengguna Twitter dapat dianggap sebagai masalah klasifikasi ordinal, yang dapat diselesaikan dengan teknik pembelajaran mendalam.

Secara khusus, Convolutional Neural networks (CNN) dan Long Short Term Memory (LSTM) adalah salah satu teknik paling canggih untuk mengekstrak fitur tersembunyi dalam tweet. Penelitian ini membandingkan kinerja kedua arsitektur tersebut untuk mengklasifikasikan tingkat keparahan depresi berdasarkan BDI. Sebuah survei yang melibatkan lebih dari 1.000 pengguna Twitter dilakukan untuk mendapatkan tingkat BDI mereka, kemudian tweet mereka dalam tiga bulan terakhir dikumpulkan dan digunakan sebagai dataset untuk melatih model. Terakhir, CNN dan LSTM dibandingkan menggunakan *Mean Absolute Error* (MAE).

Rata-rata, LSTM sedikit lebih baik dari CNN, dengan rata-rata MAE $1,313 \pm 0,09$ dan $1,382 \pm 0,04$, masing-masing. Namun, LSTM mengungguli CNN ketika mengkategorikan depresi tingkat batas (Level 3) hingga depresi tingkat ekstrem (Level 6), di mana data pada level ini secara signifikan lebih kecil dibandingkan dengan level yang lebih rendah (Level 1 dan 2). Ini menekankan potensi LSTM untuk mendeteksi depresi dini pada pengguna Twitter.

Keyword: depresi, klasifikasi ordinal, convolutional neural network, long short term memory, consistent rank logits