

Abstrak

Member pada forum komunitas online sering kali mengajukan pertanyaan yang sama dengan pertanyaan yang sudah diajukan oleh member lain. Sering kali pertanyaan itu mendapatkan tanggapan dari member lain. Akibatnya, terdapat duplikasi pertanyaan dan jawaban yang sama pada forum komunitas online. Akibat selanjutnya tentu akan membebankan sistem, karena sistem memerlukan penyimpanan data yang lebih banyak. Dibutuhkan sebuah sistem untuk mengidentifikasi pertanyaan sama agar member lain tidak perlu menjawab lagi pertanyaan yang sama tersebut. Dalam penelitian ini, sistem identifikasi pertanyaan dibuat menggunakan metode Convolutional Neural Network dengan ekstraksi fitur menggunakan word2vec. Terdapat 4 skenario pada penelitian ini dengan hasil terbaik pada skenario 4 yang menggunakan stemmer dan stopwords. Hasil dari skenario tersebut mendapatkan nilai *precision* 0,834, *recall* 0,761, *f1-score* 0,796, dan akurasi 80,42%. Hasil skenario tersebut mendapatkan hasil yang lebih baik dibandingkan dengan skenario pengujian tanpa menggunakan stemmer dan stopwords, yaitu mendapatkan *precision* 0,810, *recall* 0,735, *f1-score* 0,771, dan akurasi 78,11%.

Kata kunci: kesamaan pertanyaan, convolutional neural network, forum komunitas online, word2vec

