

Identifikasi dan reorganisasi intent pada chatbot menggunakan pemodelan topik = Identification and reorganization of chatbot intent using topic modeling.

Darell Hendry, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20517209&lokasi=lokal>

Abstrak

Chatbot sebagai asisten virtual yang digunakan oleh suatu instansi dapat memberikan manfaat bagi penggunanya. Dengan adanya chatbot, pengguna dapat berbicara langsung kepada chatbot melalui pesan singkat, yang kemudian sistem secara spontan mengidentifikasi intent pesan tersebut dan merespons dengan tindakan yang relevan. Sayangnya, cakupan pengetahuan chatbot terbatas dalam menangani pesan oleh pengguna yang semakin bervariasi. Dampak utama dari adanya variasi tersebut adalah adanya perubahan pada komposisi label intent. Untuk itu, penelitian ini berfokus pada dua hal. Pertama, pemodelan topik untuk menemukan intent dari pesan pengguna yang belum teridentifikasi intent-nya. Kedua, pemodelan topik digunakan untuk mengorganisasi intent yang sudah ada dengan menganalisis hasil keluaran model topik. Setelah dianalisis, terdapat dua kemungkinan fenomena perubahan komposisi intent yaitu: penggabungan dan pemecahan intent, dikarenakan terdapat noise saat proses anotasi dataset orisinal. Pemodelan topik yang digunakan terdiri dari Latent Dirichlet Allocation (LDA) sebagai model baseline dan dengan model state-of-the-art Top2Vec dan BERTopic. Penelitian dilakukan terhadap dataset salah satu e-commerce di Indonesia dan empat dataset publik. Untuk mengevaluasi model topik digunakan metrik evaluasi coherence, topic diversity dan topic quality. Hasil penelitian menunjukkan model topik BERTopic dan Top2Vec menghasilkan nilai topic quality 0.036 yang lebih baik dibandingkan model topik LDA yaitu -0.014. Terdapat pula pemecahan intent dan penggabungan intent yang ditemukan dengan analisis threshold proporsi.

.....Chatbot, as a virtual assistant used by an institution, can provide benefits for its users. With a chatbot, users can speak directly to the chatbot via a short message, which then the system spontaneously identifies the intent of the message and responds with the relevant action. Unfortunately, the scope of chatbot knowledge is limited in handling messages by an increasingly varied user. The main impact of this variation is a change in the composition of the intent label. For this reason, this research focuses on two things. First, topic modeling to find intents from user messages whose intents have not been identified. Second, topic modeling is used to organize existing intents by analyzing the output of the topic model. After being analyzed, there are two possible phenomena of changing intent composition: merging and splitting intents because there is noise during the annotation process of the original dataset. The topic modeling used consists of Latent Dirichlet Allocation (LDA) as the baseline model and the state-of-the-art Top2Vec and BERTopic models. The research was conducted on one dataset of e-commerce in Indonesia and four public datasets. The evaluation metrics of coherence, topic diversity, and topic quality were used to evaluate the topic model. The results showed that the BERTopic and Top2Vec topic models produced a topic quality value of 0.036, better than the LDA topic model, which was -0.014. There are also intent splitting and intent merging found by proportion threshold analysis.