

ABSTRACT

TOPIC MODELING OF APP REVIEW IN GOOGLE PLAY BASED ON LATENT DIRICHLET ALLOCATION

Akbar Nafisa Ja'far
14/360047/PA/15758

In recent years, apps on Google Play are growing rapidly. It also help app developers to get feedback from their app users through reviews. Review on apps is different from other product reviews. Review on popular applications exceeds the human ability to identify key information therein. Furthermore, each app has several updates on their features, so reviews can vary over time. This requires to identify topics from app reviews to find what information is discussed in the reviews.

In this study, the author tries to analyze the topic of the app reviews on Google Play with Latent Dirichlet Allocation (LDA) topic modeling. Moreover, the authors also analyze the results of the topic of LDA using Jensen-Shannon divergence and Sliding Window to help understand what users complain from time to time. This study use the review data from the Facebook app on Google Play from December 3, 2017 to March 13, 2018.

Based on this research, the LDA topic modeling has proved to identify complaints, bugs or errors in the application. The results of LDA have been tested for its ease to be interpreted by humans through the Word Intrusion Task. The conclusion is the topic produced by LDA can be interpreted human quite well. Moreover, the results of topic analysis using Jensen-Shannon divergence and Sliding Window successfully identified long-term issues, short-term issues and topic shifts.

Keywords : Latent Dirichlet Allocation, topic modeling, topic chain, Jensen-Shannon Divergence, sliding window, Facebook, Google Play

ABSTRAKSI

PEMODELAN TOPIK DARI *REVIEW APP* DALAM *GOOGLE PLAY* BERBASIS *LATENT DIRICHLET ALLOCATION*

Akbar Nafisa Ja'far
14/360047/PA/15758

Beberapa tahun terakhir, aplikasi di *Google Play* berkembang dengan pesat. Hal ini membantu pengembang aplikasi untuk mendapatkan *feedback* dari pengguna aplikasi mereka melalui *review*. *Review* pada aplikasi berbeda dari *review* produk lainnya. *Review* pada aplikasi yang populer melebihi kemampuan manusia untuk menyimpulkan sebuah informasi utama di dalamnya. Terlebih itu, setiap aplikasi memiliki beberapa pembaharuan pada fitur mereka, sehingga *review* dapat bervariasi dari waktu ke waktu. Untuk itu dibutuhkan identifikasi topik dari *review* aplikasi untuk menemukan informasi apa saja yang sering dibahas dalam *review* pengguna.

Dalam penelitian ini, penulis mencoba menganalisis topik dari *review* pengguna aplikasi pada *Google Play* dengan pemodelan topik *Latent Dirichlet Allocation* (LDA). Selain itu penulis juga menganalisis hasil topik dari LDA menggunakan *Jensen-Shannon divergence* dan *Sliding Window* untuk membantu memahami apa yang dikeluhkan pengguna dari waktu ke waktu. Penelitian ini menggunakan data *review* dari aplikasi *Facebook* pada *Google Play* mulai dari tanggal 3 Desember 2017 sampai 13 Maret 2018.

Berdasarkan penelitian tersebut, pemodelan topik LDA terbukti mampu untuk mengidentifikasi *komplain*, *bug* atau kesalahan pada aplikasi. Hasil eksperimen LDA telah diuji kemudahannya untuk diinterpretasi oleh manusia melalui *Word Intrusion Task*. Kesimpulan yang didapatkan adalah topik yang dihasilkan LDA dapat diinterpretasi manusia dengan cukup baik. Selain itu, hasil dari analisis topik menggunakan *Jensen-Shannon divergence* dan *Sliding Window* berhasil mengidentifikasi masalah jangka panjang, masalah jangka pendek dan pergeseran topik.

Kata kunci : *Latent Dirichlet Allocation*, pemodelan topik, *topic chain*, *Jensen-Shannon Divergence*, *sliding window*, *Facebook*, *Google Play*