

ABSTRAK

PEMODELAN TOPIK PADA PORTAL BERITA *ONLINE* MENGGUNAKAN *LATENT DIRICHLET ALLOCATION (LDA)*

MOHAMMAD REZZA FAHLEVVI

16/403698/PPA/05215

Portal berita online merupakan salah satu jenis media masa yang populer. Portal berita online memiliki peran antara lain memberikan informasi yang menarik, cepat dan aktual bagi pembacanya. Banyaknya berita yang disajikan pada portal berita online seringkali belum menampilkan topik apa saja yang sedang dibicarakan padahal berita tersebut tidak hanya sekedar untuk konsumsi bacaan saja namun dapat dimanfaatkan menjadi suatu bahan analisis. Dari berita tersebut dapat dicari topik utama atau tren yang sedang dibicarakan. Suatu metode yang cepat dan efisien dibutuhkan untuk mengetahui topik-topik apa saja yang sedang menjadi tren berita tersebut. Salah satu metode yang dapat digunakan untuk mengatasi masalah tersebut adalah pemodelan topik.

Pemodelan topik dilakukan untuk memberikan deskripsi topik bersifat informatif yang dapat langsung diterima oleh pengguna. Pemodelan topik dibutuhkan agar dapat membantu pengguna dengan mudah dan cepat dalam memahami perkembangan isu terkini. Salah satu algoritma dalam pemodelan topik adalah *Latent Dirichlet Allocation (LDA)*. Tahapan penelitian ini dimulai dengan melakukan pengumpulan data, *preprocessing*, pembentukan n-gram, representasi kamus, pembobotan, validasi model topik, pembentukan model topik dan hasil pemodelan topik.

Hasil pemodelan topik LDA pada portal berita *online* rentang waktu tertentu berhasil membentuk topik dengan informasi atau deskripsi topik. Deskripsi topik untuk setiap berita menunjukkan topik-topik tersebut sedang banyak dibicarakan atau diberitakan dalam rentang waktu tertentu oleh pengguna. Berdasarkan hasil evaluasi topik didapatkan nilai terbaik pemodelan topik menggunakan *coherence* terkait jumlah *passes* dan jumlah topik menghasilkan 20 *passes*, 5 topik dengan nilai 0,53 *coherence value* dan dapat dikatakan cukup stabil berdasarkan nilai standar *coherence value*.

Kata kunci : Portal Berita, Pemodelan Topik, *Latent Dirichlet Allocation*, *Coherence Value*

ABSTRACT

TOPIC MODELING ON ONLINE NEWS PORTAL USING LATENT DIRICHLET ALLOCATION

MOHAMMAD REZZA FAHLEVVI
16/403698/PPA/05215

Online news portals are a popular type of mass media. Online news portals have the role of providing interesting, fast and actual information to readers. The amount of news that is presented on online news portals often does not display any topic that is being discussed even though the news is not just for reading consumption but can be used as an analytical material. From the news, you can find the main topics or trends that are being discussed. A fast and efficient method is needed to find out what topics are trending the news. One method that can be used to overcome these problems is topic modeling.

Topic modeling is done to provide informative topic descriptions that can be directly accepted by the user. Topic modeling is needed in order to help users easily and quickly understand the latest development issues. One algorithm in modeling the topic is Latent Dirichlet Allocation (LDA). The stages of this research began with data collection, preprocessing, n-gram formation, dictionary representation, weighting, topic model validation, topic model formation and topic modeling results.

The results of LDA topic modeling on a certain time span online news portal succeed in forming topics with information or topic descriptions. Topic descriptions for each story indicate that the topics are being talked about or preached within a certain time frame by the user. Based on the results of topic evaluation, the best value of topic modeling using coherence related to the number of passes and the number of topics produced 20 passes, 5 topics with a value of 0.53 coherence value and can be said to be quite stable based on the standard value of coherence value.

Keyword : *News Portal, Topic Modelling, Latent Dirichlet Allocation, Coherence Value*