

INTISARI

ANALISIS TREN TOPIK PENELITIAN PADA ABSTRAK PUBLIKASI ILMIAH MENGGUNAKAN LATENT DIRICHLET ALLOCATION DAN BERTOPIC

Oleh

Jesicha Hety Manika

22/509709/PPA/06464

Seiring dengan pertumbuhan eksponensial komunitas riset, terjadi lonjakan dalam penggunaan berbagai teknologi, metode, eksperimen, dan solusi inovatif. Kemajuan yang sangat cepat dan berbagai literatur yang melimpah membuat peneliti, terutama yang memulai studi dalam bidang baru, seringkali merasa kesulitan untuk menemukan informasi terkini dan relevan dalam disiplin atau topik riset tertentu dengan banyaknya informasi yang tersedia, pemahaman yang mendalam tentang karya-karya sebelumnya dan batasan-batasan penelitian menjadi semakin penting. Pentingnya melacak perkembangan konten riset dan tren riset menjadi jelas. Dalam melakukan analisis tren memerlukan waktu untuk memahami sejauh mana perubahan terjadi, apakah tren tersebut stabil atau berfluktuasi dan, bagaimana tren ini dapat mempengaruhi pengambilan keputusan atau perkiraan di masa depan. Oleh karena itu, diperlukan cara yang efisien untuk mengidentifikasi tren riset berdasarkan topik dan waktu. Melalui pendekatan *Latent Dirichlet Allocation* (LDA) dan *BERTopic*, penelitian ini secara cermat memetakan distribusi topik dalam literatur riset. Penelitian ini menekankan pentingnya pemilihan model yang tepat, khususnya dalam konteks analisis data dengan aspek temporal, guna mendapatkan gambaran yang lebih akurat tentang evolusi topik. Berdasarkan hasil evaluasi dengan menggunakan 1.517 data abstrak, *BERTopic* menunjukkan nilai *coherence* lebih tinggi (0.7204) dibandingkan dengan LDA (0.5635), yang mengindikasikan bahwa *BERTopic* lebih akurat dan relevan dalam analisis topik. Dengan demikian, *BERTopic* lebih efektif dalam menangkap variasi topik yang lebih rinci dan melibatkan berbagai disiplin ilmu. Sedangkan LDA lebih cenderung menghasilkan topik yang bersifat lebih umum.

Kata Kunci: Tren, *Topic Modeling*, *Latent Dirichlet Allocation*, *BERTopic*

ABSTRACT

ANALYSIS OF RESEARCH TOPIC TRENDS IN SCIENTIFIC PUBLICATION ABSTRACTS USING LATENT DIRICHLET ALLOCATION AND BERTOPIC

by

Jesicha Hety Manika

22/509709/PPA/06464

Along with the exponential growth of the research community, there has been a surge in the use of innovative technologies, methods, experiments, and solutions. The rapid progress and abundance of literature means that researchers, especially those starting in a new field, often find it challenging to find current and relevant information in a particular research discipline or topic. With so much information, an in-depth understanding of previous works and research boundaries becomes increasingly important. Tracking the development of research content and research trends becomes clear. Trend analysis takes time to understand the extent of change, whether trends are stable or fluctuating, and how these trends may affect future decision-making or forecasting. Therefore, an efficient way to identify research trends based on topic and time is needed. Through the Latent Dirichlet Allocation (LDA) and BERTopic approaches, this study carefully maps the distribution of topics in the research literature. It enables researchers to track dynamic changes in research trends over some time. This research emphasizes the importance of choosing the right model, especially in the context of data analysis with temporal aspects, to get a more accurate picture of topic evolution. Based on the evaluation results using 1,517 abstract data, BERTopic shows a higher coherence value (0.7204) than LDA (0.5635), indicating that BERTopic is more accurate and relevant in topic analysis. Thus, BERTopic is more effective in capturing more detailed topic variations and involving various disciplines. In comparison, LDA is more likely to produce topics that are more general in nature.

Keywords: Trend, Topic Modeling, Latent Dirichlet Allocation, BERTop