

**ANALISIS EVOLUSI TOPIK RISET SKRIPSI INFORMATIKA  
MENGUNAKAN *NEURAL TOPIC MODELING* (BERTOPIC) BERBASIS  
INDOBERT UNTUK *DYNAMIC TOPIC ANALYSIS***

**MUHAMAD GALIH  
H1D022052**

**ABSTRAK**

Repositori skripsi universitas menyimpan data historis yang berharga untuk memetakan perkembangan fokus penelitian mahasiswa, namun analisis manual terhadap ratusan dokumen tidak memungkinkan dilakukan. Penelitian ini mengimplementasikan BERTopic sebuah arsitektur *Neural Topic Modeling* yang memanfaatkan model *embedding* IndoSBERT (denaya/indoSBERT-large), yaitu varian arsitektur *Sentence-BERT* yang dibangun di atas fondasi IndoBERT dan dilatih khusus untuk menghasilkan representasi vektor kalimat Bahasa Indonesia berdimensi 256 untuk mengidentifikasi topik riset dan menganalisis evolusi temporalnya pada 473 dokumen skripsi Informatika Universitas Jenderal Soedirman periode 2018–2026. Kinerja BERTopic divalidasi terhadap *baseline* LDA melalui 48 skenario tuning menggunakan *composite score* gabungan. Hasil evaluasi menunjukkan BERTopic mengungguli LDA secara konsisten: *Topic Coherence* 0,5182 vs 0,3583 (+44,6%), *Topic Diversity* 0,8455 vs 0,6500 (+30,1%), dan *composite score* 0,8366 vs 0,5006 (+67,1%). BERTopic menghasilkan 11 topik koheren tanpa *outlier*, sementara LDA hanya menghasilkan 8 topik dengan satu topik generik mendominasi 28,8% korpus. *Dynamic Topic Analysis* menggunakan klasifikasi *relative slope* ( $\pm 0,10$ ) mengungkap 5 topik *emerging* (termasuk *Deep Learning* dan *Machine Learning*), 3 topik *stable*, serta 3 topik *declining* (termasuk Sistem Informasi Desa/UMKM). Temuan ini mengkonfirmasi pergeseran paradigma riset dari pengembangan sistem informasi konvensional menuju kecerdasan buatan pada periode 2022–2024. Hasil analisis disajikan melalui *dashboard* interaktif berbasis Laravel dan FastAPI yang terintegrasi dengan Google Gemini API.

**Kata kunci:** BERTopic, IndoBERT, IndoSBERT, *Dynamic Topic Analysis*, pemodelan topik, LDA, analisis tren riset

**EVOLUTION ANALYSIS OF INFORMATICS THESIS RESEARCH TOPICS  
USING INDOBERT-BASED NEURAL TOPIC MODELING (BERTOPIC)  
FOR DYNAMIC TOPIC ANALYSIS**

**MUHAMAD GALIH  
H1D022052**

**ABSTRACT**

*University thesis repositories store valuable historical data for mapping student research trends, yet manual analysis of hundreds of documents is infeasible. This study implements BERTopic a Neural Topic Modeling framework utilizing IndoSBERT (denaya/indoSBERT-large), a Sentence-BERT variant built upon the IndoBERT foundation and specifically trained to produce 256-dimensional Indonesian sentence embeddings to identify research topics and analyze their temporal evolution across 473 Informatics thesis documents from Universitas Jenderal Soedirman spanning 2018–2026. BERTopic was validated against an LDA baseline through 48 tuning scenarios using a composite score. Results show BERTopic consistently outperforms LDA: Topic Coherence 0.5182 vs 0.3583 (+44.6%), Topic Diversity 0.8455 vs 0.6500 (+30.1%), and composite score 0.8366 vs 0.5006 (+67.1%). BERTopic identified 11 coherent topics with zero outliers, while LDA produced only 8 topics with one generic topic absorbing 28.8% of the corpus. Dynamic Topic Analysis using relative slope classification ( $\pm 0.10$ ) revealed 5 emerging topics (including Deep Learning and Machine Learning), 3 stable topics, and 3 declining topics (including Village/SME Information Systems). These findings confirm a research paradigm shift from conventional information system development toward artificial intelligence-based approaches during 2022–2024. All results are presented through an interactive dashboard built with Laravel and FastAPI, integrated with Google Gemini API.*

**Keywords:** BERTopic, IndoBERT, IndoSBERT, Dynamic Topic Analysis, topic modeling, LDA, research trend analysis