

LAPORAN SKRIPSI

**DETEKSI TOPIK *FRAUD* MENGGUNAKAN *LATENT*
DIRICHLET ALLOCATION (LDA): STUDI KASUS
PADA DATA APLIKASI X**



FIRSTIAWAN FADHIL THOBIBUDDIN

22.G4.0018

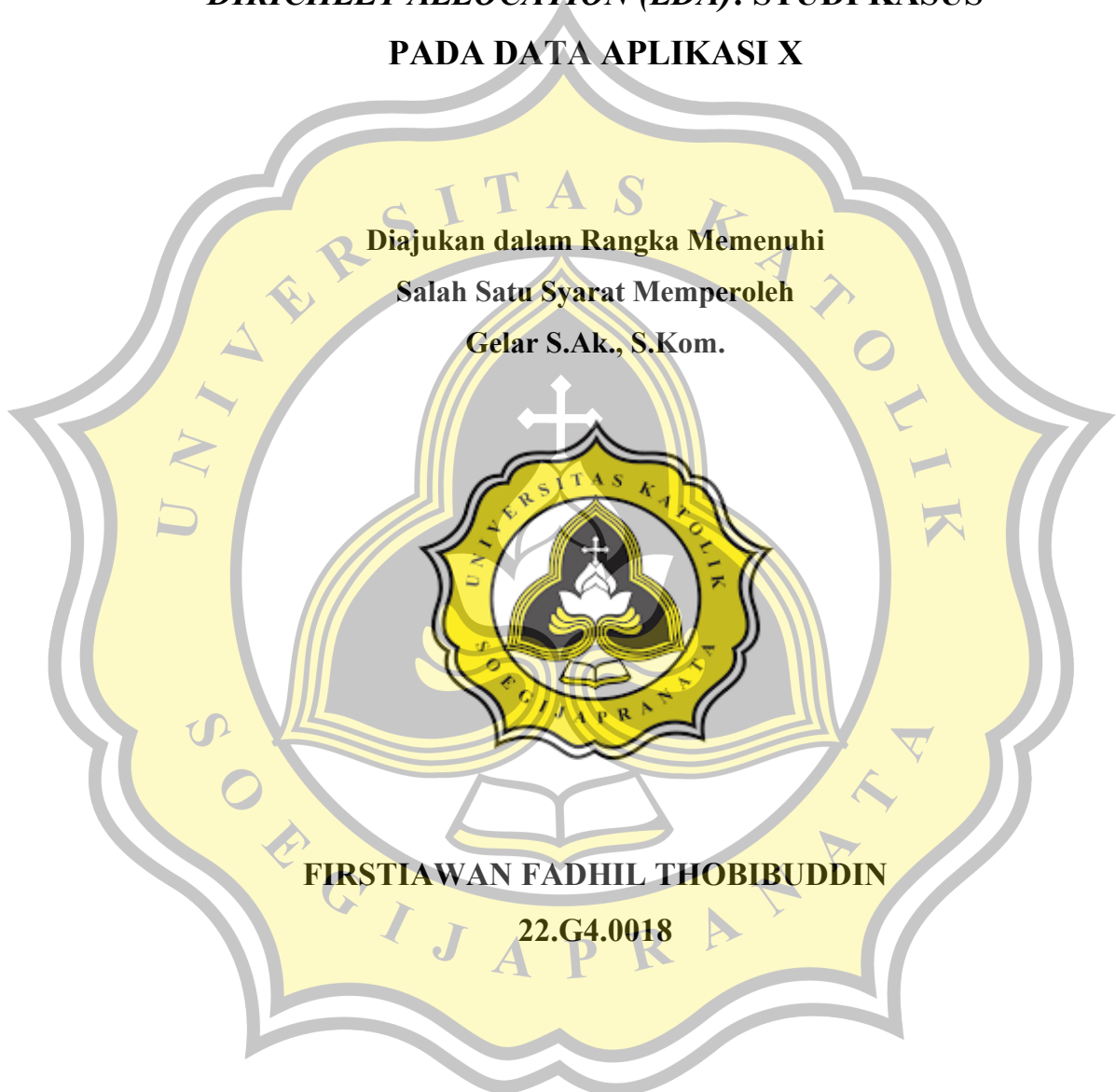
**PROGRAM STUDI AKUNTANSI DAN SISTEM INFORMASI
FAKULTAS EKONOMI DAN BISNIS
UNIVERSITAS KATOLIK SOEGIJAPRANATA
SEMARANG**

2026

LAPORAN SKRIPSI

**DETEKSI TOPIK *FRAUD* MENGGUNAKAN *LATENT*
DIRICHLET ALLOCATION (LDA): STUDI KASUS
PADA DATA APLIKASI X**

**Diajukan dalam Rangka Memenuhi
Salah Satu Syarat Memperoleh
Gelar S.Ak., S.Kom.**



FIRSTIAWAN FADHIL THOBIBUDDIN

22.G4.0018

**PROGRAM STUDI AKUNTANSI DAN SISTEM INFORMASI
FAKULTAS EKONOMI DAN BISNIS
UNIVERSITAS KATOLIK SOEGIJAPRANATA
SEMARANG**

2026

ABSTRAK

Maraknya kasus kecurangan (*fraud*) di Indonesia banyaknya informasi pada media sosial mendorong perlunya pendekatan berbasis data untuk mendeteksi indikasi *fraud* secara lebih cepat dan sistematis. Penelitian ini bertujuan untuk mengidentifikasi pola topik yang berkaitan dengan *fraud* melalui analisis teks pada platform X menggunakan metode *Latent Dirichlet Allocation (LDA)*. Penelitian menggunakan pendekatan *mix methods* dengan teknik web scraping terhadap unggahan publik berbahasa Indonesia periode Januari–November 2025 yang mengandung kata kunci terkait *fraud*. Data diproses melalui tahapan *preprocessing* meliputi *case folding*, *tokenization*, *normalisasi*, *stopword filtering*, serta pembentukan *bigram* dan *trigram*, kemudian dilakukan pemodelan topik dan evaluasi menggunakan C_v (C_v) untuk menentukan jumlah topik optimal. Hasil penelitian menunjukkan bahwa model LDA mampu mengelompokkan percakapan publik ke dalam beberapa topik utama seperti korupsi, manipulasi laporan keuangan, penyalahgunaan aset, dan praktik suap, dengan tingkat koherensi yang menunjukkan kualitas pemodelan yang baik. Implementasi model dalam aplikasi berbasis Streamlit juga membuktikan bahwa sistem dapat digunakan sebagai alat bantu screening awal indikasi *fraud*. Penelitian ini menyimpulkan bahwa LDA efektif sebagai metode eksploratif untuk mendeteksi pola *fraud* pada media sosial, namun hasilnya bersifat indikatif dan memerlukan verifikasi lebih lanjut.

Kata Kunci: *Fraud, Latent Dirichlet Allocation, Topic Modeling, Natural Language Processing, Media Sosial*