

UNDERGRADUATE THESIS

**FINDING SIMILARITY WITHIN DOCUMENTS USING TF-IDF, SENTENCE
TRANSFORMER, AND DISTILBERT AS VECTORIZATION AND COSINE
SIMILARITY AS CALCULATION METHOD**

**MENEMUKAN SIMILARITAS DALAM DOKUMEN MENGGUNAKAN TF-IDF,
SENTENCE TRANSFORMER, DAN DISTILBERT SEBAGAI VEKTORISASI DAN
SIMILARITAS KOSINUS SEBAGAI METODE PERHITUNGAN**

Diajukan untuk memenuhi salah satu syarat memperoleh derajat Sarjana Ilmu Komputer



RAPHAEL OCTAVIAN BONG

18/425532/PA/18424

**COMPUTER SCIENCE STUDY PROGRAMME
DEPARTMENT OF COMPUTER SCIENCE AND ELECTRONICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
UNIVERSITAS GADJAH MADA
YOGYAKARTA**

2023

SKRIPSI (RISET AKADEMIK)

**S1-IUP CREATING PLAGIARISM DETECTION TOOLS USING TF-IDF AND
BERT FOR VECTORIZATION WITH JACCARD AND COSINE AS
METHODS OF CALCULATION**


Telah dipersiapkan dan disusun oleh

Raphael Octavian Bong


18/425532/PA/18424

Telah dipertahankan di depan Tim Penguji
pada tanggal 13 Desember 2023


Susunan Tim Penguji



Drs. Edi Winarko, M.Sc., Ph.D
Ketua Penguji



Arif Nurwidyantoro, M.Cs., Ph.D
Anggota Penguji



Dr. Azhari, MT
Pembimbing

Mengetahui,
a.n. Dekan FMIPA UGM
Wakil Dekan Bidang Pendidikan, Pengajaran
dan Kemahasiswaan



Prof. Drs. Roto, M.Eng., Ph.D
NIP. 196711171993031020