



**IMPLEMENTASI METODE *RANDOM OVER-UNDER SAMPLING*
PADA ALGORITMA *RANDOM FOREST CLASSIFIER* DALAM
PROSES ANALISIS SENTIMEN PADA WEBSITE STEAM (STUDI
KASUS: GAME CYBERPUNK 2077)**

LAPORAN SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**UNIVERSITAS
MERCU BUANA**
LINTANG ANNURYA DEWA
41519010054

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA JAKARTA
2023**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan dibawah ini:

Nama : Lintang Annurya Dewa
NIM : 41519010054
Program Studi : Teknik Informatika
Judul Laporan : Implementasi Metode *Random Over-Under Sampling* Pada Algoritma *Random Forest Classifier* Dalam Proses Analisis Sentimen Pada Website Steam (Studi Kasus: Game Cyberpunk 2077)

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 19 Juni 2023

UNIVERSITAS
MERCU BUANA



Lintang Annurya Dewa




HALAMAN PENGESAHAN

Saya yang bertanda tangan dibawah ini:

Nama : Lintang Annurya Dewa
NIM : 41519010054
Program Studi : Teknik Informatika
Judul Laporan : Implementasi Metode *Random Over-Under Sampling* Pada Algoritma *Random Forest Classifier* Dalam Proses Analisis Sentimen Pada Website Steam (Studi Kasus: Game Cyberpunk 2077)

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Siti Maesaroh, S.kom., M.T.I. 
NIDN : 0913059003
Ketua Penguji : Anis Qerid 
NIDN : 0328127203
Penguji 1 : Achmad Kodar 
NIDN : 0323085801

MERCU BUANA

Jakarta, ... 20 Juli 2023

Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jokonowo, S.Si., M.T.I



Dr. Bagus Priambodo, ST., M.T.I

KATA PENGANTAR

Dengan memanjatkan puji syukur kehadirat Allah SWT, karena atas rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan Laporan Skripsi ini dengan judul **“Implementasi Metode Random Over-Under Sampling Pada Algoritma Random Forest Classifier Dalam Proses Analisis Sentimen Pada Website Steam (Studi Kasus: Game Cyberpunk 2077)”**. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Penulis menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi penulis untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Dr. Bambang Jokonowo, S.Si, MTI selaku Dekan Fakultas Ilmu Komputer Universitas Mercu Buana.
2. Bagus Priambodo, ST, MTI selaku Ketua Program Studi Teknik Informatika.
3. Ibu Siti Maesaroh, S.Kom., MTI selaku Dosen Pembimbing Skripsi yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini.
4. Seluruh staff pengajar Fakultas Ilmu Komputer Universitas Mercu Buana yang telah memberikan ilmu pengetahuan yang tak ternilai selama penulis menempuh Pendidikan di Fakultas Ilmu Komputer Universitas Mercu Buana
5. Ibu dan Ayah tercinta, yang selalu memberikan dukungan, kasih sayang, doa, nasehat, serta atas kesabarannya yang luar biasa dalam setiap langkah hidup penulis. Penulis berharap dapat menjadi anak yang dapat dibanggakan.
6. Adik penulis, terima kasih atas doa dan dukungannya, sampai saat ini penulis masih berusaha agar menjadi kakak yang baik.
7. Seluruh keluarga tersayang yang senantiasa mendoakan dan memberikan semangat dalam penyelesaian skripsi ini.
8. Seluruh teman-teman penulis di Fakultas Ilmu Komputer UMB Warung Buncit angkatan 2019. Zakki, Bayu, Rani, Denise, Ilham, serta beberapa yang tiba-tiba

menghilang entah kemana. Terima kasih atas bantuan, saran, diskusi, serta keseruan disaat jam-jam kosong.

9. Seluruh teman-teman platform Discord. Khususnya server DYWSC, Eja, Ipan, Zakki, Bayu, Iwa, Menace, Denis, CJ, serta seluruh member lainnya yang ikut meramaikan sehingga menciptakan suasana baru. Terima kasih atas keseruan, kesenangan, semangat, canda tawa yang membahagiakan di tengah-tengah terjadinya Pandemi COVID-19. Sehingga penulis terus termotivasi dalam menempuh perkuliahan.

Akhir kata, saya berharap Allah SWT berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.



Jakarta, 20 Juli 2023

UNIVERSITAS
MERCU BUANA

Penulis

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Lintang Annurya Dewa
NIM : 41519010054
Program Studi : Teknik Informatika
Judul Laporan : Implementasi Metode *Random Over-Under Sampling* Pada Algoritma *Random Forest Classifier* Dalam Proses Analisis Sentimen Pada Website Steam (Studi Kasus: Game Cyberpunk 2077)

Demi pengembangan ilmu pengeahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalti-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpang, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 27 Juli 2023

Yang Menyatakan,



Lintang Annurya Dewa

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI.....	ii
HALAMAN PENGESAHAN	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	6
2.1 Penelitian Terkait.....	6
2.2 Teori Pendukung.....	14
BAB III METODE PENELITIAN	25
3.1 Jenis Penelitian	25
3.2 Tahap Penelitian	25
BAB IV PEMBAHASAN	28
4.1 Dataset	28
4.2 Pre-processing.....	29
4.3 Data Modelling	35
4.3 Visualisasi Data	37
4.4 Pengujian	40
4.5 Analisa Hasil.....	46

BAB V KESIMPULAN DAN SARAN	48
5.1 Kesimpulan	48
5.2 Saran	49
DAFTAR PUSTAKA.....	50
LAMPIRAN	53



UNIVERSITAS
MERCU BUANA

DAFTAR TABEL

Tabel 1 Penelitiat Terkait	14
Tabel 2 Contoh Casefolding dan Hyperlink Removal.....	17
Tabel 3 Contoh Penggunaan Stemming.....	18
Tabel 4 Contoh Penggunaan Tokenizing	18
Tabel 5 Contoh Penggunaan Stopwords	19
Tabel 6 Contoh Punctuation Removal	20
Tabel 7 Contoh Kalimat untuk TF-IDF	22
Tabel 8 Tabel TFIDF.....	23
Tabel 9 Proporsi data Random Over-Under Sampling.....	47
Tabel 10 Performa Model	47



UNIVERSITAS
MERCU BUANA

DAFTAR GAMBAR

Gambar 1 Tahapan Random Forest	15
Gambar 2 Cara kerja over-under sampling.....	16
Gambar 3 Tahapan Penelitian.....	25
Gambar 4 Aplikasi Scrapping dengan libary Steam-reviews 0.2.1	28
Gambar 5 Dataset Hasil Scraping.....	28
Gambar 6 Python Library	29
Gambar 7 Proses Load dataset menggunakan Pandas.....	30
Gambar 8 Kode untuk menampilkan info data null	31
Gambar 9 Output dari eksekusi	31
Gambar 10 Membersihkan data null pada label Review dan Vote.....	31
Gambar 11 Re-mapping label True dan False	32
Gambar 12 Hasil Re-mapping	32
Gambar 13 Penggunaan fungsi df.drop_duplicates.....	32
Gambar 14 Pembuatan Fungsi untuk menghapus Hyperlink dan Markup.....	33
Gambar 15 Pembuatan Fungsi untuk menghapus data numeric.....	33
Gambar 16 Pembuatan fungsi untuk menghapus data yang mengandung emoji.....	33
Gambar 17 Pembuatan fungsi untuk menghapus spasi berlebih	33
Gambar 18 Pembuatan fungsi untuk menghapus simbol dan tanda baca	33
Gambar 19 Fungsi untuk aplikasi Stopwords.....	34
Gambar 20 Fungsi untuk melakukan Stemming dan Tokenisasi.....	34
Gambar 21 Fungsi untuk mengaplikasikan dengan meng-combine semua fungsi	34
Gambar 22 Baris Kode untuk menampilkan hasil text cleaning	34
Gambar 23 Hasil Text Cleaning	35
Gambar 24 Pembobotan Kata menggunakan TF-IDF	35
Gambar 25 Pembagian data dengan Rasio 80:20	36
Gambar 26 Proses Klasifikasi Algoritma RF	36
Gambar 27 Fungsi Confusion Matrix.....	36
Gambar 28 Confusion Matrix Rasio 80:20	37
Gambar 29 WordCloud Sentimen Positif.....	38
Gambar 30 WordCloud Sentimen Negatif.....	39
Gambar 31 Barplot	40
Gambar 32 Visualisasi Confusion Matrix 80:20	41
Gambar 33 Tabel Pengujian Confusion Matrix 80:20.....	41
Gambar 34 Visualisasi Confusion Matrix 70:30	42
Gambar 35 Tabel Pengujian Confusion Matrix 70:30.....	43
Gambar 36 Kode Over-Under Sampling	44
Gambar 37 Pembagian dataset 80:20	44
Gambar 38 Pembuatan model dengan Pipeline.....	45
Gambar 39 Visualisas Confusion Matrix dari Over-Under Sampling dengan 80:20	45
Gambar 40 Tabel Confusion Matrix Over-Under Sampling	45

DAFTAR LAMPIRAN

Lampiran 1 Bimbingan.....	53
Lampiran 2 Luaran Tugas Akhir.....	55
Lampiran 3 Bukti Submit / Published Artikel Ilmiah.....	55
Lampiran 4 Naskah Jurnal.....	56
Lampiran 5 Curriculum Vitae.....	63
Lampiran 6 Surat Pernyataan HKI.....	64
Lampiran 7 Sertifikat BNSP/Bukti Daftarnya.....	66



UNIVERSITAS
MERCU BUANA