

SKRIPSI



PERBANDINGAN KINERJA *RANDOM FOREST* DAN *SMOTE RANDOM FOREST* DALAM MENDETEKSI DAN MENGUKUR TINGKAT STRES PADA MAHASISWA TINGKAT AKHIR

VIONITA OKTAVIANI

NIM. 2010511027

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
JANUARI 2024**

PERNYATAAN ORISINALITAS

Tugas akhir ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Vionita Oktaviani

NIM : 2010511027

Tanggal : 15 Januari 2024

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 15 Januari 2024

Yang menyatakan,



(Vionita Oktaviani)

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK
KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini :

Nama : Vionita Oktaviani

NIM : 2010511027

Fakultas : Ilmu Komputer

Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :

PERBANDINGAN KINERJA RANDOM FOREST DAN SMOTE RANDOM FOREST DALAM MENDETEKSI DAN MENGUKUR TINGKAT STRES PADA MAHASISWA TINGKAT AKHIR.

Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian Pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 15 Januari 2024

Yang menyatakan,



(Vionita Oktaviani)

LEMBAR PENGESAHAN

Skripsi diajukan oleh :

Nama : Vionita Oktaviani

NIM : 2010511027

Program Studi : S1 Informatika

Judul Skripsi : Perbandingan Kinerja *Random Forest* dan *SMOTE Random Forest* dalam Mendeteksi dan Mengukur Tingkat Stres pada Mahasiswa Tingkat Akhir
Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Yuni Widiastiwi, S.Kom, M.Si.

Penguji 1



Novi Trisman Hadi, S.Pd., M.Kom.

Penguji 2



Nur Rosmawarni, S.Kom., M.Kom.

Pembimbing 1



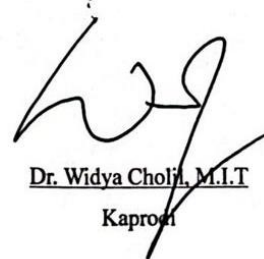
M. Panji Muslim, S.Pd., M.Kom.

Pembimbing 2



Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM

Dekan



Dr. Widya Cholij, M.I.T

Kaprodi

Ditetapkan di : Jakarta

Tanggal Ujian : 12 Januari 2024

KATA PENGANTAR

Puji syukur atas kehadiran Tuhan Yang Maha Esa yang senantiasa melimpahkan rahmat serta hidayah-Nya berupa kesehatan, kesempatan dan kemudahan sehingga penulis dapat menyelesaikan Skripsi dengan judul “Perbandingan Kinerja *Random Forest* dan *SMOTE Random Forest* dalam Mendeteksi dan Mengukur Tingkat Stres pada Mahasiswa Tingkat Akhir”. Adapun tujuan penulisan laporan ini sebagai syarat kelulusan dan memperoleh gelar sarjana prodi Informatika di Universitas Pembangunan Nasional Veteran Jakarta.

Dalam menyusun skripsi ini, tentu saja ada banyak bantuan, dukungan, motivasi, dan doa yang diterima dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Dua orang paling berjasa dalam hidup penulis, Bapak Muhamad Apriantoro dan Ibu Siti Mundaryati, terima kasih atas perjuangannya baik materil dan non materil yang hingga sampai saat ini membuat penulis berada dititik sekarang ini. Mereka memang tidak sempat untuk mengecap bangku perkuliahan namun tidak dapat disangkal bahwasannya mereka mampu untuk mendidik, memotivasi, serta memberikan dukungan sehingga penulis mampu untuk menyelesaikan studinya sampai sarjana. Semoga Allah senantiasa membalas jasa kalian dan menjaga kalian dalam kebaikan dan kemudahan aamiin.
2. Keluarga besar penulis yang selalu mendukung serta mendoakan agar penelitian ini dapat berjalan mudah dan lancar tanpa hambatan dari pihak manapun.
3. *My lover*, Muhamad Naufan Sheva Ramadhan selaku *support system* penulis yang senantiasa telah memberikan dukungan, motivasi serta selalu siap sedia memberikan sandaran serta telinganya untuk mendengarkan keluh kesah yang penulis hadapi dalam mengerjakan tugas akhir dan lainnya. *Trust me, since I know you, I once believed love would be burning red, but I know it's golden, like daylight.*
4. Seluruh Civitas Akademika Fakultas Ilmu Komputer UPN Veteran Jakarta terutama Dosen Pembimbing Tugas Akhir penulis, Ibu Neny Rosmawarni, M.Kom.
5. Rekan yang telah kebersamai penulis sejak semester satu perkuliahan, Febriani Jati Nawangsih dan teman teman S1 Informatika angkatan 2020 yang telah menemani dan berjuang bersama dalam perjalanan perkuliahan ini.

6. Ivon Virginia Patala, Adinda Khairani, Azzura Ferliani Ramadhani, Aisyah Nadhira Salma Maritza, I Gusti Agung Ngurah Fajar Dharmawangsa serta teman – teman magang BPJS Ketenagakerjaan yang selalu memberikan dukungan kepada penulis saat mengerjakan skripsi.
7. Taylor Swift dengan semua lantunan lagunya yang selalu membuat penulis bersemangat dalam mengerjakan skripsi.
8. Serta pihak lain yang tidak dapat penulis sebutkan satu-persatu namun tidak mengurangi rasa syukur dan terima kasih penulis dalam penyelesaian skripsi ini.
9. Terakhir, ribuan terima kasih penulis ucapkan pada diri ini, Vionita Oktaviani yang senantiasa selalu ingin untuk terus berprogres dan berusaha dalam menyelesaikan Tugas Akhirnya. Terima kasih untuk selalu mampu untuk mengendalikan diri meskipun banyak tekanan yang selalu disimpan dan mengobatinya sendiri, terima kasih untuk tidak pernah menyerah sesulit apapun keadaannya dan terima kasih karena telah menjadi diri sendiri setiap saat. Ingatlah bahwa dirimu selalu membanggakan untuk sekecil apapun tiap progressnya. *“If you never bleed, you’re never gonna grow”* -Taylor Swift.

Semoga semua yang telah memberikan bantuan dan dukungan kepada penulis mendapatkan berkah yang melimpah dari Allah SWT.

Penulis menyadari bahwa tulisan ini jauh dari kata sempurna dikarenakan terbatasnya pengetahuan dan pengalaman yang dimiliki. Oleh karena itu, penulis sangat mengharapkan segala bentuk kritik dan saran dari berbagai pihak yang nantinya dapat memotivasi untuk menuju ke arah yang lebih baik. Selain itu, penulis juga berharap bahwa penelitian ini akan memberikan manfaat di kemudian hari.

Jakarta, Januari 2024

Vionita Oktaviani

NIM 2010511027

PERBANDINGAN KINERJA *RANDOM FOREST* DAN *SMOTE RANDOM FOREST* DALAM MENDETEKSI DAN MENGUKUR TINGKAT STRES PADA MAHASISWA TINGKAT AKHIR

VIONITA OKTAVIANI

ABSTRAK

Dalam kehidupan sehari-hari manusia, stres merupakan masalah nyata sehingga menjadi bagian yang tidak terpisahkan. Ketidaksiapan individu dalam menghadapi tuntutan akademis dapat mengakibatkan stres sebagai salah satu gangguan psikologis. Dalam hal ini, stres akademik merupakan stres yang dialami oleh mahasiswa, terutama mahasiswa tingkat akhir. Adanya banyak tekanan baik dari masalah ekonomi, akademik maupun kondisi sosial dapat menjadi pemicu stres bagi mahasiswa tingkat akhir. Penelitian ini bertujuan untuk mengklasifikasikan diagnosa tingkat stress mahasiswa tingkat akhir dengan membandingkan kinerja yang terbaik antara *Random Forest* dengan *SMOTE Random Forest*. Data yang diolah dalam penelitian ini merupakan data yang dihasilkan oleh kuesioner yang berisi 14 pertanyaan yang ditujukan pada mahasiswa tingkat akhir yang sedang melaksanakan skripsi. Adapun hasil dari penelitian ini, disimpulkan bahwasannya metode *Random Forest* dengan menggunakan *SMOTE* mampu mempengaruhi dan meningkatkan evaluasi klasifikasi studi kasus klasifikasi diagnosa mahasiswa tingkat akhir dengan akurasi sebesar 71%, precision sebesar 72% dan recall sebesar 71% pada pembagian 80% data latih, 20% data uji dengan nilai $K=5$.

Kata Kunci: Klasifikasi, *Random Forest*, *Oversampling SMOTE*, Stress.

PERBANDINGAN KINERJA *RANDOM FOREST* DAN *SMOTE RANDOM FOREST* DALAM MENDETEKSI DAN MENGUKUR TINGKAT STRES PADA MAHASISWA TINGKAT AKHIR

VIONITA OKTAVIANI

ABSTRACT

In people's daily lives, stress is a real problem so it becomes an inseparable part. Individual unpreparedness in facing academic demands can result in stress as a psychological disorder. In this case, academic stress is stress experienced by students, especially final year students. The presence of a lot of pressure from economic, academic and social conditions can trigger stress for final year students. This research aims to classify the stress level diagnosis of final year students by comparing the best performance between Random Forest and SMOTE Random Forest. The data processed in this research is data produced by a questionnaire containing 14 questions aimed at final year students who are carrying out their thesis. As for the results of this research, it was concluded that the Random Forest method using SMOTE was able to influence and improve the evaluation of case study classification for final year student diagnosis with an accuracy of 71%, precision of 72% and recall of 71% on a 80% distribution of training data, 20 % test data and K value = 5.

Keywords: *Classification, Random Forest, SMOTE Oversampling, Stress.*

DAFTAR ISI

PERNYATAAN ORISINALITAS	i
LEMBAR PENGESAHAN	iii
KATA PENGANTAR	iv
ABSTRAK.....	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah.....	4
BAB II LANDASAN TEORI	5
2.1 Machine Learning	5
2.2 Random Forest	5
2.3 Data Preprocessing.....	6
2.4 Evaluasi	6
2.5 Imbalance Data.....	8
2.6 SMOTE (Synthetic Minority Oversampling Technique)	8
2.7 Stres.....	8
2.8 Penelitian Terdahulu.....	8
BAB III 12	
METODOLOGI PENELITIAN.....	12
3.1 Tempat dan Waktu Peneltian	12
3.2 Alat Pendukung Penelitian	12
3.3 Diagram Alir Penelitian.....	12
3.4 Jadwal Pelaksanaan	16
BAB IV HASIL DAN PEMBAHASAN	17
4.1 Data	17
4.2 Preprocessing Data.....	19

4.3 Pembagian Data.....	21
4.4 Pembentukan Model Algoritma	23
4.5 Evaluasi	30
4.6 Analisis dan Kesimpulan.....	41
BAB V SIMPULAN DAN SARAN.....	45
3.5 Kesimpulan.....	45
5.2 Saran.....	45
DAFTAR PUSTAKA.....	47
LAMPIRAN.....	49
1. Lampiran hasil cek Turnitin	49

DAFTAR GAMBAR

Gambar 2.1	Diagram alir <i>Random Forest</i> (Khushaktov, 2023)	6
Gambar 3.1	Diagram Alir Penelitian.....	13
Gambar 4.1	Pembagian data dengan <i>train test split</i>	22
Gambar 4.2	Pembentukan model algoritma <i>Random Forest</i>	23
Gambar 4.3	Pohon keputusan <i>Random Forest</i>	25
Gambar 4.4	Pembentukan model algoritma.....	27
Gambar 4.5	Pohon Keputusan SMOTE <i>Random Forest</i>	28
Gambar 4.6	Evaluasi model dan pengujian performa.....	30
Gambar 4.7	Code Confusion matrix data training	31
Gambar 4.8	Confussion Matrix data training 80%	31
Gambar 4.9	Code Confusion matrix data testing.....	32
Gambar 4.10	Confusion Matrix data uji 20%.....	33
Gambar 4.11	Code Confusion matrix data training	36
Gambar 4.12	<i>Confussion Matrix</i> data training 80%.....	36
Gambar 4.13	Code Confusion matrix data testing.....	37
Gambar 4.14	Confusion Matrix rasio data uji 20%	38
Gambar 4.15	Perbandingan akurasi data latih dan data uji.....	43
Gambar 4.16	Perbandingan akurasi data latih dan data uji.....	44

DAFTAR TABEL

Tabel 2.1	Tabel <i>Confussion Matrix</i>	7
Tabel 2.2	Perhitungan Evaluasi Model	7
Tabel 2.3	Derajat <i>Imbalanced Data</i>	8
Tabel 2.4	Penelitian Terdahulu.....	8
Tabel 3.1	Jadwal Pelaksanaan	16
Tabel 4.1	Daftar peubah penelitian	17
Tabel 4.2	Data sebelum transformasi	20
Tabel 4.3	Data setelah transformasi	21
Tabel 4.4	Tabel pengukuran	21
Tabel 4.5	Pembagian data training.....	22
Tabel 4.6	Pembagian data testing.....	22
Tabel 4. 7	Variabel Data Input dan Data Output	23
Tabel 4. 8	Distribusi data pada tiap kelas.....	26
Tabel 4. 9	Distribusi data pada tiap kelas setelah SMOTE	28
Tabel 4. 10	Nilai Confusion Matrix data latih 80%	31
Tabel 4. 11	Nilai Confusion Matrix data uji 20%	34
Tabel 4. 12	Nilai <i>Precision, recall</i> dan <i>F1 – Score</i>	35
Tabel 4. 13	Nilai Confusion Matrix data latih 80%	37
Tabel 4. 14	Nilai Confusion Matrix data uji 20%	39
Tabel 4. 15	Nilai <i>Precision, recall</i> dan <i>F1 – Score</i>	40