

**OPTIMASI DATA UNTUK MENANGANI KETIDAKSEIMBANGAN
DENGAN MENGGUNAKAN METODE
K- NEAREST NEIGHBOR DAN *Z-TRANSFORMATION*
PADA *SOFTWARE DEFECT PREDICTION***



SKRIPSI

**diajukan untuk memenuhi salah
satu syarat memperoleh gelar
sarjana komputer**

Oleh:

Anjar Maghfiruddin

201569040015

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS YUDHARTA PASURUAN**

2019

PERNYATAAN PENULIS

JUDUL : OPTIMASI DATA UNTUK MENANGANI
KETIDAKSEIMBANGAN DENGAN MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR* DAN *Z-TRANSFORMATION*
PADA *SOFTWARE DEFECT PREDICTION*

NAMA : ANJAR MAGHFIRUDDIN

NIM : 2015 69 04 0015

"Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk di batal kan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut"

Pasuruan, ... Agustus 2019



Anjar Maghfiruddin

Penulis

PERSETUJUAN SKRIPSI

JUDUL : OPTIMASI DATA UNTUK MENANGANI
KETIDAKSEIMBANGAN DENGAN MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR* DAN *Z-TRANSFORMATION*
PADA *SOFTWARE DEFECT PREDICTION*

NAMA : ANJAR MAGHIFIRUDDIN



NIM : 2015.69.04.0015

Skripsi ini telah diperiksa dan disetujui

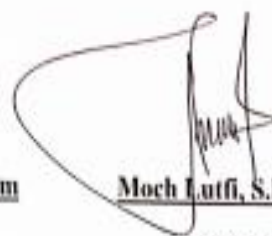
Pasuruan, Agustus 2019

Kaprodi

Pembimbing



Muhammad Imron R. S.Kom, M.Kom
NIK. Y. 0690213121



Moch Lutfi, S.Kom, M.Kom.
NIK.Y 0690201004

PENGESAHAN SKRIPSI

JUDUL : OPTIMASI DATA UNTUK MENANGANI
KETIDAKSEIMBANGAN DENGAN MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR* DAN *Z-TRANSFORMATION*
PADA *SOFTWARE DEFECT PREDICTION*

NAMA : ANJAR MAGHFIRUDDIN

NIM : 2015.69.04.0015

Skripsi ini telah diujikan dan dipertahankan didepan Dewan Penguji pada
Sidang Skripsi tanggal ... Juli 2019. Menurut pandangan kami, Skripsi ini
memadai dari segi kualitas untuk tujuan penganugerahan gelar
Sarjana Komputer (S.Kom)

Pasuruan, Agustus 2019

KetuaPenguji



Rahmad Zainul Abidin, M.Kom

NIK. Y 069.15.07.141

Anggota



Arif Faizin, S.Kom, M.Kom

NIK. Y 069.17.07.002

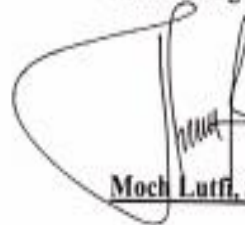
DekanFakultasTeknik



MISBAHUL MUNIR, ST

NIK. Y 069.02.01.015

Pembimbing



Moch Lutfi, S.Kom, M.Kom.

NIK. Y 0690201004

HALAMAN PERSEMBAHAN

Skripsi ini saya persembahkan untuk orang-orang yang saya sayangi dan saya cintai, terutama untuk kedua orangtua saya yang senantiasa mendoakan dan selalu mendukung saya meskipun kedua orangtua saya yaitu SAIHU ROFI'I dan LAILATUL QOMORIYAH yang lagi ada di Bekasi dan saya di Pasuruan semoga kedua orang tua sehat selalu.

Dan untuk adik-adik saya yaitu Firman Maulana dan Riski Perdana tunggu saya pulang ke Bekasi ya adik-adikku dan kita bisa berkumpul kembali.

Tak lupa juga untuk teman-teman IT2015 yang selalu menemaniku baik dalam keadaan suka maupun duka, semoga kita bisa lulus bareng di tahun ini

Untuk kekasihku, Riska Wilfiah yang aku cintai dan aku sayangi

Skripsi ini saya persembahkan untuk kalian semua atas kasih sayang dan bimbingan selama ini sehingga saya dapat menyelesaikan skripsi ini dengan baik.

Selama 4 tahun telah menempuh pendidikan telah tiba saatnya saya akan membuktikan kepada kalian dan dunia bahwa saya telah siap untuk membuka lembaran baru sebagai seseorang yang bertanggung jawab dan berkarya bagi kepentingan banyak orang. Semoga niat dan perbuatan saya kedepan dapat meyakinkan kalian bahwa saya mampu untuk berbagi kebaikan.

Terima kasih kembali yang sebesar-besarnya kepada ayah dan mamah Tercinta,

Aku sayang kalian.

Untuk yang aku hormati, dosenku, pembimbingku dan kaprodi Informatika terimakasih atas bimbingan kalian selama ini, maaf jika selama ini sudah banyak merepotkan kalian.

Dengan segala ketulusan hati

Anjar Maghfiruddin

ABSTRACT

Software that is not found defective during inspection and testing is high-quality software that is able to help the company's business processes effectively and efficiently. Decrease in software quality and high repair costs can be caused by errors or defects in software. Z-Transformasi is a classification model that is suitable for data with numeric attributes or nominal value attributes that are categorical in that each value cannot be added or subtracted, datasets used in software defect prediction research are generally unbalanced, which can reduce model performance defect prediction software. To deal with class imbalances, it is generally divided into two approaches, namely the data level approach and the algorithm level approach. The data level approach is intended to improve class balance, while the algorithm level approach is intended to make classifiers more conducive towards minority classes by combining single classifiers to make them better. K-Nearest Neighbor Algorithm is one algorithm that has been shown to improve the performance of classifiers. So in this study the data level approach is used using Z-Transformation normalization, and for the algorithm level approach using K-Nearest Neighbor. After conducting experiments on the K-Nearest Neighbor and Z-Transformation models, the accuracy is 86.55%, while if only with the K-Nearest Neighbor, the accuracy is 85.73%.

ABSTRAK

Software yang tidak ditemukan cacat selama pemeriksaan dan pengujian merupakan software berkualitas tinggi yang mampu membantu proses bisnis perusahaan secara efektif dan efisien. Penurunan kualitas software dan biaya perbaikan yang tinggi dapat diakibatkan kesalahan atau cacat pada software. *Z-Transformasi* merupakan model klasifikasi yang cocok untuk data dengan atribut yang bersifat *numerik* ataupun atribut yang bernilai nominal yaitu bersifat kategorik dimana tiap nilai tidak bisa dijumlahkan atau dikurangkan, Dataset yang digunakan pada penelitian *software defect prediction* umumnya bersifat tidak seimbang, sehingga dapat menurunkan kinerja model *software defect prediction*. Untuk menangani ketidakseimbangan kelas secara umum dibagi menjadi dua pendekatan, yaitu pendekatan level data dan pendekatan level algoritma. Pendekatan level data ditujukan untuk memperbaiki keseimbangan kelas, Sedangkan pendekatan level algoritma ditujukan untuk membuat pengklasifikasi lebih konduktif terhadap kelas minoritas dengan menggabungkan pengklasifikasi tunggal agar menjadi lebih baik. Algoritma *K-Nearest Neighbor* merupakan salah satu algoritma yang telah menunjukkan dapat memperbaiki kinerja pengklasifikasi, Maka pada penelitian ini digunakan pendekatan level data menggunakan normalisasi *Z-Transformasi*, dan untuk pendekatan level algoritma menggunakan *K-Nearest Neighbor*. Setelah dilakukan percobaan pada model *K-Nearest Neighbor* dan *Z-Transformasi* dapatkan akurasi 86.55% sedangkan jika hanya dengan *K-Nearest Neighbor* saja di dapatkan akurasi 85.73%.

KATA PENGANTAR

Dengan mengucapkan syukur alhamdulillah, penulis memanjatkan puji syukur kehadirat Allah SWT, atas rahmat dan karunia-Nya, penulis dapat menyelesaikan skripsi ini dengan judul **“OPTIMASI DATA UNTUK MENANGANI KETIDAKSEIMBANGAN DENGAN MENGGUNAKAN METODE *K-NEAREST NEIGHBOR* DAN *Z-TRANSFORMATION* PADA *SOFTWARE DEFECT PREDICTION*”** sebagai salah satu syarat yang harus dipenuhi untuk memperoleh gelar sarjana Komputer di Universitas Yudharta Pasuruan.

Sholawat dan salam terlimpah selalu kepada junjungan kita Nabi Muhammad SAW sebagai tauladan (*Uswatun Hasanah*) yang telah menebarkan risalah kebenaran dimuka bumi ini.

Dalam penulisan skripsi ini, Penulis menyadari bahwa banyak hambatan dan kesulitan yang penulis hadapi, alhamdulillah berkat rahmat dan pertolongan Allah SWT serta bantuan dari berbagai pihak akhirnya penulis dapat menyelesaikan skripsi ini. Untuk itu penulis mengucapkan banyak terima kasih kepada semua pihak yang telah membantu serta memberikan dukungan moril maupun materil baik secara langsung ataupun tidak langsung dalam penyusunan skripsi ini. Dalam hal ini penulis mengucapkan terima kasih kepada:

1. Romo KH. M. Sholeh Bahruddin selaku pengasuh Yayasan Darut Taqwa, dimana Universitas Yudharta Pasuruan bernaung dan atas nasehat-nasehat beliau dalam hidup berdampingan dengan masyarakat tanpa membeda-bedakan semua kalangan
2. Teristimewa kepada keluarga tercinta Bapak dan Ibu, yang senantiasa mendoakan dan selalu tak henti-hentinya member dorongan semangat dengan penuh kesabaran, penuh kasih dan ketulusan serta perhatiannya telah memberikan support baik materi, moril serta doa yang tidak pernah putus sehingga penulis dapat menyelesaikan skripsi ini.
3. Bapak Dr. Saifulloh, M.HI, Selaku Rektor Universitas Yudharta Pasuruan.
4. Bapak Misbach Munir, ST, MT, Selaku Dekan Fakultas Teknik.

5. Bapak M. Imron Rosadi, S.Kom, M.Kom selaku prodi Teknik Informatika Universitas Yudharta Pasuruan
6. Bapak Moch. Lutfhi, S.Kom, M.Kom, dosen pembimbing yang telah sudi membantu dalam penyelesaian skripsi ini.
7. Seluruh Bapak dan Ibu Dosen Program Studi Teknik Informatika yang telah memberikan ilmu dan bimbingannya selama masa kuliah.
8. Rekan-rekan Himpunan Mahasiswa Teknik Informatika Universitas Yudharta Pasuruan.
9. Dan semua pihak yang telah membantu penulis dalam penulisan skripsi ini, yang tidak bias penulis sebutkan satu per satu, terimakasih banyak atas semangat dan dukungannya.

MOTTO

Dalam perjalanan sebuah ranting, ia tidak akan sungkan menggugurkan apa yang membuatnya berat dan terhalang.

Tak ragu menggugurkan seluruh daunnya hanya untuk berkembang dan tumbuh lebih baik agar bias bertahan hidup.

Karna dengan meninggalkan apa yang dimiliki saat ini, maka allah akan memberikan ganti dengan kemuliaan dan keagungan yang lebih lagi.

Karena, sesuatu yang menjadi milikmu tidak akan allah biarkan pergi melainkan allah akan menggantinya dengan yang lebih baik.

DAFTAR ISI

PERNYATAAN SKRIPSI	ii
PERSETUJUAN SKRIPSI.....	iii
PENGESAHAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
<i>ABSTRACT</i>	vi
ABSTRAK.....	vii
KATA PENGANTAR	viii
MOTTO	x
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
DAFTAR RUMUS	xvi
DAFTAR LAMPIRAN.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan.....	3
1.4 Batasan Masalah.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematis Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Penelitian Terdahulu.....	5
2.1.1 Integrasi Metode <i>Resampling</i> Cacat Software	5

2.1.2	Klasifikasi Data Tidakseimbang	6
2.1.3	Penerapan Metode <i>Esamble</i> Untuk Klasifikasi.....	6
2.1.4	Penerapan Metode <i>NF</i> Dan <i>KNN</i>	7
2.1.5	Algoritma <i>RU</i> Untuk Ketidak Seimbangan	8
2.2	Landasan Teori	9
2.2.1	Prediksi Cacat Software	9
2.2.2	Ketidakseimbangan Kelas.....	10
2.2.3	Normalisasi Data.....	11
2.2.4	<i>Z-Transformation</i>	11
2.2.5	<i>kNN (k-Nearest Neighbor)</i>	12
2.2.6	Evaluasi Hasil.....	12
2.2.7	<i>Confution Metrix</i>	13
2.2.8	Kerangka Pemikiran.....	15
BAB III METODE PENELITIAN		18
3.1	Analisa kebutuhan	18
3.2	Metode Penelitian.....	18
3.3	Pengumpulan Data	19
3.4	Pengolahan Awal Data	20
3.5	Model/metode Yang Diusulkan.....	21
3.6	Eksperimen Dan Pengujian Model.....	21
3.7	Evaluasi dan Validasi Hasil.....	21
3.8	Metode Pengumpulan Data	21
3.9	Metode Analisis Data	23
3.10	Transformasi Data	24
3.11	Pengukuran Kinerja.....	24

3.12 Perancangan Penelitian.....	24
3.13 Model yang Diusulkan	24
BAB IV HASIL DAN PEMBAHASAN	26
4.1 Hasil.....	26
4.2 Penghitungan Manual.....	26
4.3 Perhitungan Nilai Parameter	26
4.4 Skema Perhitungan.....	29
4.5 Perhitungan <i>Z-Transformasi</i>	29
4.6 Penghitungan <i>K-Nearest Neighbor</i>	37
4.7 Evaluasi <i>K-Nearest Neighbor</i> dengan <i>Z-Transformasi</i>	42
4.8 Perbandingan <i>KNN</i> dengan <i>KNN + Z-Transformasi</i>	43
BAB V KESIMPULAN DAN SARAN	45
5.1 Kesimpulan.....	45
5.2 Saran	45
DAFTAR PUSTAKA	46

DAFTAR TABEL

Tabel 2.1 <i>Related Research</i>	9
Tabel 2.2 <i>Confusion Matrix</i>	15
Tabel 3.1 Dataset.....	22
Tabel 3.2 Dataset <i>Training</i>	27
Tabel 3.3 Dataset <i>Testing</i>	28
Tabel 4.1 Dataset Asli.....	30
Tabel 4.2 Hasil Penghitungan <i>Mean dan Stdev</i>	30
Tabel 4.3 Hasil Normalisasi.....	31
Tabel 4.4 data <i>Training</i> dan <i>Testing</i>	38
Tabel 4.5 Hasil Penghitungan Jarak.....	40
Tabel 4.6 Hasil <i>Confution Matrix</i>	42
Tabel 4.7 Perbandingan Akurasi.....	44

DAFTAR GAMBAR

Gambar 2.1 Rumus <i>Confution Matrix</i>	16
Gambar 2.2 Rumus <i>Confution Matrix</i>	16
Gambar 3.1 Tahapan Penelitian	19
Gambar 3.2 Kerangka Pemikiran.....	25
Gambar 4.2 Skema penghitungan	29
Gambar 4.2 Skema penghitungan <i>KNN & Z-Transformasi</i>	44

DAFTAR RUMUS

Rumus 2.1 <i>Z-Transformasi</i>	12
Rumus 2.2 <i>k-Nearest Neighbor</i>	13
Rumus 4.1 Normalisasi	31
Rumus 4.2 <i>K-Nearest Neighbor</i>	37
Rumus 4.3 Akurasi	43
Rumus 4.3 Akurasi	43
Rumus 4.3 Akurasi	43

DAFTAR LAMPIRAN

Dokumentasi

Daftar Riwayat Hidup

Kartu Seminar

Kartu Bimbingan Skripsi