



**ANALISIS PERBANDINGAN VARIASI POLYNOMIAL-FIT-SMOTE
DENGAN KLASIFIKASI BERBASIS POHON PADA PREDIKSI CACAT
PERANGKAT LUNAK**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

WILDAN NUR HIDAYATULLAH

NIM. 2011016210003

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

JUNI 2024



**ANALISIS PERBANDINGAN VARIASI POLYNOMIAL-FIT-SMOTE DENGAN
KLASIFIKASI BERBASIS POHON PADA PREDIKSI CACAT PERANGKAT
LUNAK**

Skripsi

**Untuk Memenuhi Persyaratan
Dalam Menyelesaikan Strata-1 Ilmu Komputer**

Oleh

**WILDAN NUR HIDAYATULLAH
NIM. 2011016210003**

**PROGRAM STUDI S-1 ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMBUNG MANGKURAT
BANJARBARU**

JUNI 2024

SKRIPSI


ANALISIS PERBANDINGAN VARIASI POLYNOMIAL-FIT-SMOTE DENGAN KLASIFIKASI BERBASIS POHON PADA PREDIKSI CACAT PERANGKAT LUNAK

Oleh :


WILDAN NUR HIDAYATULLAH
NIM. 2011016210003

Telah dipertahankan di depan Dosen Penguji pada tanggal 28 Juni 2024

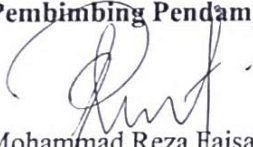
Susunan Dosen Penguji :
Pembimbing Utama


Rudy Hertono S.Kom., M.Kom
NIP. 198809252022031003


Penguji I


Radityo Adi Nugroho, S.T., M.Kom
NIP. 198212042008011006

Pembimbing Pendamping

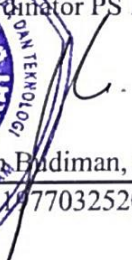

Mohammad Reza Faisal, S.Si., S.T., M.T., Ph.D
NIP. 197612202008121001

Penguji II


Setyo Wahyu Saputro, S.Kom., M.Kom
NIP. 198808072023211027



Koordinator PS Ilmu Komputer,


Widiadiman, S.T., M.Kom.
NIP. 197703252008121001

SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam jurnal ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Pustaka.

Banjarbaru, 28 Juni 2024

Yang Menyatakan,



Wildan Nur Hidayatullah

NIM. 2011016210003

ABSTRAK

ANALISIS PERBANDINGAN VARIASI POLYNOMIAL-FIT-SMOTE DENGAN KLASIFIKASI BERBASIS POHON PADA PREDIKSI CACAT PERANGKAT LUNAK

(Oleh: Wildan Nur Hidayatullah; Pembimbing: Rudy Herteno, S.Kom., M.Kom. dan Mohammad Reza Faisal, S.Si., S.T., M.T., Ph.D.; 2024; halaman)

Cacat perangkat lunak menghadirkan tantangan signifikan terhadap keandalan sistem perangkat lunak, yang sering kali mengakibatkan kerugian ekonomi yang substansial. Studi ini menguji kemanjuran varian Polynomial-fit-SMOTE (pf-SMOTE) yang dikombinasikan dengan klasifikasi berbasis pohon untuk prediksi cacat perangkat lunak, dengan memanfaatkan kumpulan data NASA Metrics Data Program (MDP). Metodologi penelitian melibatkan pembagian kumpulan data menjadi subset pelatihan dan pengujian, penerapan oversampling pf-SMOTE, dan evaluasi kinerja klasifikasi menggunakan Decision Trees, Random Forests, and Extra Trees. Temuan menunjukkan bahwa kombinasi oversampling pf-SMOTE-star dengan klasifikasi Extra Trees mencapai akurasi rata-rata tertinggi (90,91%) dan AUC (95,67%) di 12 kumpulan data MDP NASA. Hal ini menunjukkan potensi varian pf-SMOTE untuk meningkatkan efektivitas klasifikasi. Namun, penting untuk dicatat bahwa kehati-hatian diperlukan terkait potensi bias yang diperkenalkan oleh data sintetis. Temuan-temuan ini merupakan kemajuan signifikan atas upaya-upaya penelitian sebelumnya, yang menggarisbawahi peran penting pemilihan algoritma yang cermat dan karakteristik kumpulan data dalam mengoptimalkan hasil klasifikasi. Implikasi yang penting mencakup kemajuan dalam keandalan perangkat lunak dan dukungan keputusan untuk manajemen proyek perangkat lunak. Penelitian di masa mendatang dapat menyelidiki sinergi antara varian pf-SMOTE dan metode klasifikasi alternatif, serta mengeksplorasi integrasi penyetelan hiperparameter untuk lebih menyempurnakan kinerja klasifikasi.

Kata Kunci: Polynomial-Fit-SMOTE, Decision Tree, Random Forest, Extra Trees, Software Defect Prediction.

ABSTRACT

A COMPARATIVE ANALYSIS OF POLYNOMIAL-FIT-SMOTE VARIATIONS WITH TREE-BASED CLASSIFIERS ON SOFTWARE DEFECT PREDICTION

(By: Wildan Nur Hidayatullah; Pembimbing: Rudy Herteno, S.Kom., M.Kom. dan Mohammad Reza Faisal, S.Si., S.T., M.T., Ph.D.; 2024; page)

Software defects present a significant challenge to the reliability of software systems, often resulting in substantial economic losses. This study examines the efficacy of polynomial-fit SMOTE (pf-SMOTE) variants in combination with tree-based classifiers for software defect prediction, utilising the NASA Metrics Data Program (MDP) dataset. The research methodology involves partitioning the dataset into training and test subsets, applying pf-SMOTE oversampling, and evaluating classification performance using Decision Trees, Random Forests, and Extra Trees. Findings indicate that the combination of pf-SMOTE-star oversampling with Extra Tree classification achieves the highest average accuracy (90.91%) and AUC (95.67%) across 12 NASA MDP datasets. This demonstrates the potential of pf-SMOTE variants to enhance classification effectiveness. However, it is important to note that caution is warranted regarding potential biases introduced by synthetic data. These findings represent a significant advancement over previous research endeavors, underscoring the critical role of meticulous algorithm selection and dataset characteristics in optimizing classification outcomes. Noteworthy implications include advancements in software reliability and decision support for software project management. Future research may delve into synergies between pf-SMOTE variants and alternative classification methods, as well as explore the integration of hyperparameter tuning to further refine classification performance.

Keyword: Polynomial-Fit-SMOTE, Decision Tree, Random Forest, Extra Trees, Software Defect Prediction.

KATA PENGANTAR

Puji syukur penulis panjatkan ke Tuhan kita Yang Maha Esa karena atas berkat rahmat dan karunia-Nya penulis dapat menyelesaikan jurnal yang berjudul “*A Comparative Analysis of Polynomial-fit-SMOTE Variations with Tree-Based Classifiers on Software Defect Prediction*” untuk memenuhi syarat dalam menyelesaikan pendidikan program S1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Pada lembar ini penulis ingin menyampaikan ucapan terimakasih kepada semua pihak yang sangat mendukung penulis dalam pembuatan dan penyusunan jurnal ini, adapun yang dimaksud adalah sebagai berikut:

1. Kedua orang tua dan keluarga besar yang selalu memberikan bantuan, semangat, doa dan dukungan dalam proses penyelesaian jurnal ini.
2. Bapak Rudy Herteno, S.Kom., M.Kom. selaku dosen pembimbing utama, dan Bapak Mohammad Reza Faisal, S.Si., S.T., M.T., Ph.D. selaku dosen pembimbing pendamping yang turut serta membantu dan meluangkan waktu demi kelancaran dalam penyelesaian jurnal ini.
3. Bapak Irwan Budiman, S.T., M. Kom. selaku Ketua Program Studi Ilmu Komputer FMIPA ULM, atas bantuan dan izin beliau jurnal ini dapat diselesaikan.
4. Seluruh Dosen dan staff Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini yang sangat bermanfaat.
5. Teman-teman dan sahabat-sahabat keluarga Ilmu Komputer yang memberikan dukungan dan selalu mengingatkan serta mendoakan dalam proses mengerjakan jurnal.
6. Serta semua pihak yang tidak dapat disebutkan satu persatu yang telah turut membantu dalam penyelesaian jurnal ini.

Akhir kata penulis menyadari sepenuhnya bahwa penulisan ini jauh dari sempurna. Semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan pembaca khususnya serta mendapat keridhaan Allah SWT.

Banjarbaru, 28 Juni 2024



Wildan Nur Hidayatullah