



**ENHANCING SOFTWARE DEFECT PREDICTION: HHO-BASED  
WRAPPER FEATURE SELECTION WITH ENSEMBLE METHODS**

**SKRIPSI**

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

**Oleh**

**ACHMAD FAUZAN LUTHFI**

**NIM 2111016210004**

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

**MEI 2025**



**ENHANCING SOFTWARE DEFECT PREDICTION: HHO-BASED  
WRAPPER FEATURE SELECTION WITH ENSEMBLE METHODS**

**SKRIPSI**

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

**Oleh**

**ACHMAD FAUZAN LUTHFI**

**NIM 2111016210004**

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

**MEI 2025**

# SKRIPSI

## ENHANCING SOFTWARE DEFECT PREDICTION: HHO-BASED WRAPPER FEATURE SELECTION WITH ENSEMBLE METHODS

Oleh:

**ACHMAD FAUZAN LUTHFI**

**NIM. 2111016210004**

Telah dipertahankan di depan Dosen Penguji pada tanggal 8 Mei 2025.

Susunan Dosen Penguji:

**Pembimbing I**



Rudy Herteno, S.Kom., M.Kom.

NIP. 198809252022031003

**Dosen Penguji I**



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

NIP. 198212042008011006

**Pembimbing II**



Friska Abadi, S.Kom., M.Kom.

NIP. 198809132023211010

**Dosen Penguji II**



Muhammad Itqan Mazdadi, S.Kom., M.Kom.

NIP. 199006122019031013

Banjarbaru, 9 Mei 2025

**Koordinator Program Studi Ilmu Komputer**



[Signature]

NIP. 198704212012122003

## PRAKATA

Puji dan syukur penulis panjatkan ke hadirat Allah SWT, karena atas rahmat dan karunia-Nya, penulisan skripsi yang berjudul **“ENHANCING SOFTWARE DEFECT PREDICTION: HHO-BASED WRAPPER FEATURE SELECTION WITH ENSEMBLE METHODS”** ini dapat diselesaikan dengan baik. Skripsi ini disusun untuk memenuhi syarat untuk memperoleh gelar sarjana di program studi Strata-1 Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lambung Mangkurat.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa keberhasilan ini tidak terlepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak, Oleh karena itu dengan rasa hormat, penulis ingin menyampaikan terima kasih sebesar-besarnya kepada:

1. Allah SWT atas segala rahmat, karunia, hidayah, dan ridho-Nya yang telah memungkinkan penulis menyelesaikan skripsi ini.
2. Ayah, ibu, serta kedua saudara penulis yang selalu memberikan doa, kasih sayang, dan dukungan tanpa henti, sehingga penulis memiliki semangat untuk menyelesaikan skripsi ini.
3. Bapak Rudy Herteno, S.Kom., M.Kom, dan Bapak Friska Abadi, S.Kom., M.Kom, selaku dosen pembimbing utama dan pendamping, yang telah memberikan bimbingan, arahan, dan masukan yang sangat berharga selama proses penyusunan skripsi ini.
4. Ibu Dwi Kartini, S.Kom., M.Kom, selaku ketua program studi Ilmu Komputer, beserta seluruh dosen dan staf Fakultas MIPA Universitas Lambung Mangkurat atas bantuan yang diberikan selama penulis mengikuti studi.
5. Adinda Dyandra Wardhani, seseorang yang selalu menemani dalam keadaan suka maupun duka, yang selalu mendengarkan keluh kesah penulis, dan selalu memberikan dukungan terhadap penulis.

6. Teman-teman IOT; Alif, Tajali, Michael, Nazar, Raidra, Rizky, Yosia dan Gina yang selalu mendukung dalam berbagai hal.
7. Teman-teman seperjuangan di Ilmu Komputer angkatan 2021 serta sahabat-sahabat yang selalu memberikan semangat dan dukungan selama masa perkuliahan.
8. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini, yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis membuka diri terhadap kritik dan saran yang membangun demi penyempurnaan karya ini. Akhir kata, semoga skripsi ini dapat bermanfaat bagi pembaca dan menjadi kontribusi kecil dalam pengembangan ilmu pengetahuan.

Banjarbaru, 2 Mei 2025



**Achmad Fauzan Luthfi**

**NIM. 2111016210004**

## **PERNYATAAN**

Dengan ini saya menyatakan bahwa didalam skripsi 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 didalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 2 Mei 2025

Yang Menyatakan,



**Achmad Fauzan Luthfi**

**NIM. 2111016210004**

## ABSTRAK

### **Enhancing Software Defect Prediction: HHO-Based Wrapper Feature Selection with Ensemble Methods**

(Oleh: Achmad Fauzan Luthfi; Pembimbing: Rudy Herteno, S.Kom., M.Kom., dan Friska Abadi, S.Kom., M.Kom; 2025; 54 halaman)

Kompleksitas data yang semakin meningkat di berbagai bidang menekankan perlunya model klasifikasi yang efektif yang dapat memprediksi hasil dengan akurat sambil mengatasi tantangan seperti ketidakseimbangan kelas dan redundansi fitur. Dataset *MDP NASA* menghadirkan tantangan unik karena karakteristiknya yang beragam dan kelas yang tidak seimbang, yang dapat berdampak negatif pada kinerja model. Penelitian ini bertujuan untuk mengembangkan model klasifikasi yang kuat yang menggunakan teknik pra-pemrosesan canggih, metode pemilihan fitur, dan berbagai algoritma klasifikasi untuk meningkatkan akurasi prediktif. Metodologi yang digunakan melibatkan beberapa langkah kunci, termasuk standarisasi z-score untuk normalisasi data, skala robust untuk meminimalkan dampak dari outlier, dan penerapan teknik ADASYN untuk mengatasi ketidakseimbangan kelas. Pemilihan fitur dilakukan menggunakan metode *Binary Harris Hawk Optimization* (BHHO), dengan KNN sebagai evaluator, untuk mengidentifikasi fitur yang paling relevan untuk tugas klasifikasi. Model kemudian diuji menggunakan beberapa algoritma, termasuk *Random Forest* (RF), *Support Vector Machine* (SVM), dan *Stacking*, dengan kinerja dievaluasi melalui metrik seperti akurasi, AUC, presisi, recall, dan F1-measure. Hasil penelitian menunjukkan bahwa model *Stacking* mencapai akurasi tertinggi sebesar 0,998 dan AUC sebesar 1,000 pada dataset MC1, yang menunjukkan kinerja prediktif yang luar biasa. Selain itu, model RF secara konsisten menunjukkan kinerja yang baik di berbagai dataset, sementara SVM mengalami kesulitan dengan skor yang lebih rendah, terutama pada dataset seperti JM1 dan KC1. Temuan ini menunjukkan bahwa kombinasi pra-pemrosesan yang canggih secara signifikan meningkatkan kinerja model. Selain itu, melakukan *exploratory data analysis* (EDA) yang lebih mendalam dapat membantu mengidentifikasi pola dan hubungan yang mendasari dalam data yang dapat lebih meningkatkan pemilihan fitur dan pelatihan model. Secara keseluruhan, penelitian ini menyoroti potensi model klasifikasi yang dioptimalkan dalam menangani dataset yang kompleks secara efektif, membuka jalan bagi kemajuan di masa depan dalam analitik prediktif..

**Kata kunci:** *Harris Hawk Optimization; Feature Selection; Ensemble Methods; Software Defect Prediction; Wrapper*

## ABSTRACT

### **Enhancing Software Defect Prediction: HHO-Based Wrapper Feature Selection with Ensemble Methods**

(By: Achmad Fauzan Luthfi; Pembimbing: Rudy Herteno, S.Kom., M.Kom., dan Friska Abadi, S.Kom., M.Kom; 2025; 54 halaman)

*The growing complexity of data across various fields underscores the necessity for effective classification models that can accurately predict outcomes while addressing challenges such as class imbalance and feature redundancy. The NASA MDP dataset presents a unique challenge due to its diverse characteristics and imbalanced classes, which can negatively affect model performance. This research aimed to develop a robust classification model that employs advanced preprocessing techniques, feature selection methods, and various classification algorithms to enhance predictive accuracy. The methodology involved several key steps, including z-score standardization for data normalization, robust scaling to minimize the impact of outliers, and the application of the ADASYN technique to address class imbalance. Feature selection was performed using a binary Harris Hawk Optimization (BHHO) method, with KNN as the evaluator, to identify the most relevant features for the classification task. The model was then tested using multiple algorithms, including Random Forest (RF), Support Vector Machine (SVM), and Stacking, with performance evaluated through metrics such as accuracy, AUC, precision, recall, and F1-measure. The results demonstrated that the Stacking model achieved the highest accuracy of 0.998 and an AUC of 1.000 on the MC1 dataset, indicating exceptional predictive performance. Additionally, the RF model consistently performed well across various datasets, while SVM struggled with lower scores, particularly in datasets like JM1 and KC1. The findings suggest that the combination of advanced preprocessing significantly enhances model performance. Furthermore, conducting a more in-depth exploratory data analysis (EDA) could help identify underlying patterns and relationships within the data that may further improve feature selection and model training. Overall, this research highlights the potential of optimized classification models in effectively handling complex datasets, paving the way for future advancements in predictive analytics.*

**Kata kunci:** *Harris Hawk Optimization; Feature Selection; Ensemble Methods; Software Defect Prediction; Wrapper*