

**ANALISIS KINERJA RANDOM FOREST DAN DECISION
TREE DENGAN OPTIMASI SIMULATED ANNEALING
UNTUK MEMPREDIKSI NILAI ASPD**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
MEYLANI EKA PERTIWI
21.11.4393

Kepada
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025

**ANALISIS KINERJA RANDOM FOREST DAN DECISION
TREE DENGAN OPTIMASI SIMULATED ANNEALING
UNTUK MEMPREDIKSI NILAI ASPD**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
MEYLANI EKA PERTIWI
21.11.4393

Kepada
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS KINERJA RANDOM FOREST DAN DECISION
TREE DENGAN OPTIMASI SIMULATED ANNEALING
UNTUK MEMPREDIKSI NILAI ASPD**

yang disusun dan diajukan oleh

Meylani Eka Pertiwi

21.11.4393

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 30 Juli 2025

Dosen Pembimbing,



Windha Mega Pradnya Duhita, S.Kom., M.Kom.
NIK. 190302185

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS KINERJA RANDOM FOREST DAN DECISION TREE
DENGAN OPTIMASI SIMULATED ANNEALING UNTUK
MEMPREDIKSI NILAI ASPD**

yang disusun dan diajukan oleh

Meylani Eka Pertiwi

21.11.4393

Telah dipertahankan di depan Dewan Penguji
pada tanggal 30 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Subektiningsih, S.Kom., M.Kom.
NIK. 190302413



Anna Baita, S.Kom., M.Kom.
NIK. 190302290



Windha Mega Pradnya Dhuhiya, S.Kom., M.Kom.
NIK. 190302185



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 30 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Meylani Eka Pertiwi
NIM : 21.11.4393

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Kinerja Random Forest Dan Decision Tree Dengan Optimasi
Simulated Annealing Untuk Memprediksi Nilai Aspd**

Dosen Pembimbing : Windha Mega Pradnya Dhuhiita, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 30 Juli 2025

Yang Menyatakan,



Meylani Eka Pertiwi

HALAMAN PERSEMBAHAN

Dengan memanjatkan puji syukur ke hadirat Allah SWT, serta berkat dukungan dan doa dari orang-orang tercinta, saya akhirnya dapat menyelesaikan skripsi ini. Saya menyadari bahwa pencapaian ini tidak lepas dari bantuan, doa, dan dukungan banyak pihak. Oleh karena itu, dengan penuh rasa syukur dan hormat, saya mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT, karena hanya atas izin dan karunia-Nya lah maka skripsi ini dapat dibuat dan selesai pada waktunya.
2. Bapak saya Timbul Raharja dan ibu saya Zunati, yang telah memberikan dukungan moril maupun material serta do'a yang tiada henti untuk kesuksesan saya, karena tiada kata seindah lanjutan do'a dan tiada do'a yang paling khusus selain do'a yang terucap dari orang tua.
3. Adik saya Arga Bagus Surya Atmaja yang selalu memberi do'a dan semangat untuk menyelesaikan skripsi ini.
4. Ibu Windha Mega Pradnya Duhita, S.Kom., M.Kom. selaku dosen pembimbing saya yang telah membimbing dan memberikan arahan selama proses penulisan ini.
5. Seluruh keluarga besar saya yang telah memberikan do'a dan harapan kepada saya untuk lebih termotivasi menjadi lebih baik dan berkembang.
6. Pijar Pahlawan Kalbu sebagai patner saya dalam menghadapi rintangan dan selalu memberikan dukungan penuh untuk kelancaran penulisan ini.
7. Teman-teman seperjuangan, yang telah memberi semangat dan bantuan selama masa kuliah.

KATA PENGANTAR

Puji Syukur saya panjatkan ke hadirat Allah SWT atas Rahmat dan karunia-Nya sehingga skripsi yang berjudul “Analisis Kinerja Radom Forest dan Decision Tree dengan Optimasi Simulated Annealing untuk Memprediksi Nilai ASPD” ini dapat siselesaikan dengan baik.

Penulisan skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana di Program Studi Informatika Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini, penulis telah banyak menerima bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis mengucapkan terima kasih kepada :

1. Kedua orang tua saya, atas doa dan dukungan yang tiada henti.
2. Ibu dosen pembimbing, yang telah memberikan waktu, bimbingan, dan arahan selama proses penyusunan.
3. Almamater tercinta, Universitas Amikom Yogyakarta, yang telah memberikan kesempatan dan dukungan beasiswa selama masa studi saya.
4. Teman-teman seperjuangan, yang telah meberi semangat dan bantuan selama masa kuliah.
5. Semua pihak yang tidak dapat disebutkan satu per satu, yang turut membantu dalam proses penelitian ini.

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan skripsi ini. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa mendatang. Akhir kata, semoga skripsi ini dapat memberikan manfaat dan menjadi referensi bagi penelitian selanjutnya.

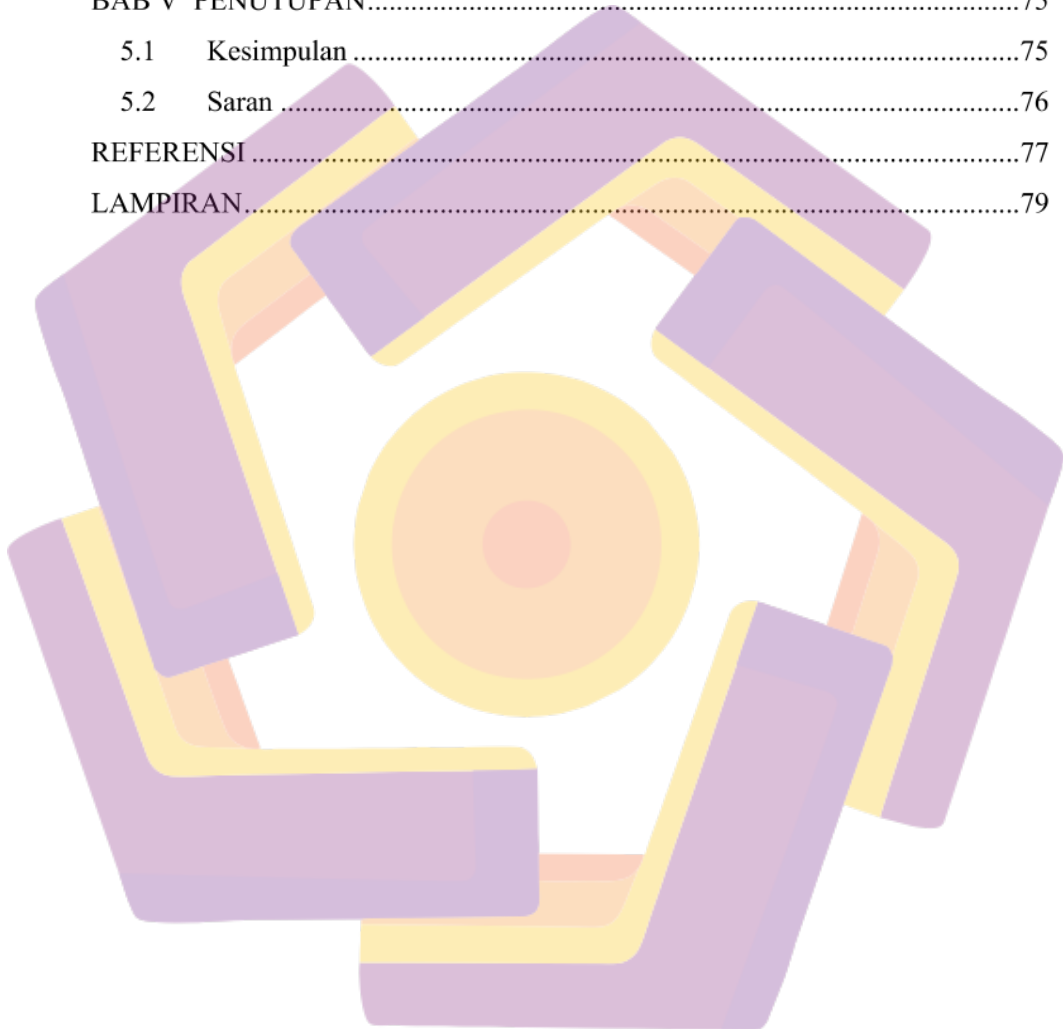
Yogyakarta, 30 Juli 2025

Penulis

DAFTAR ISI

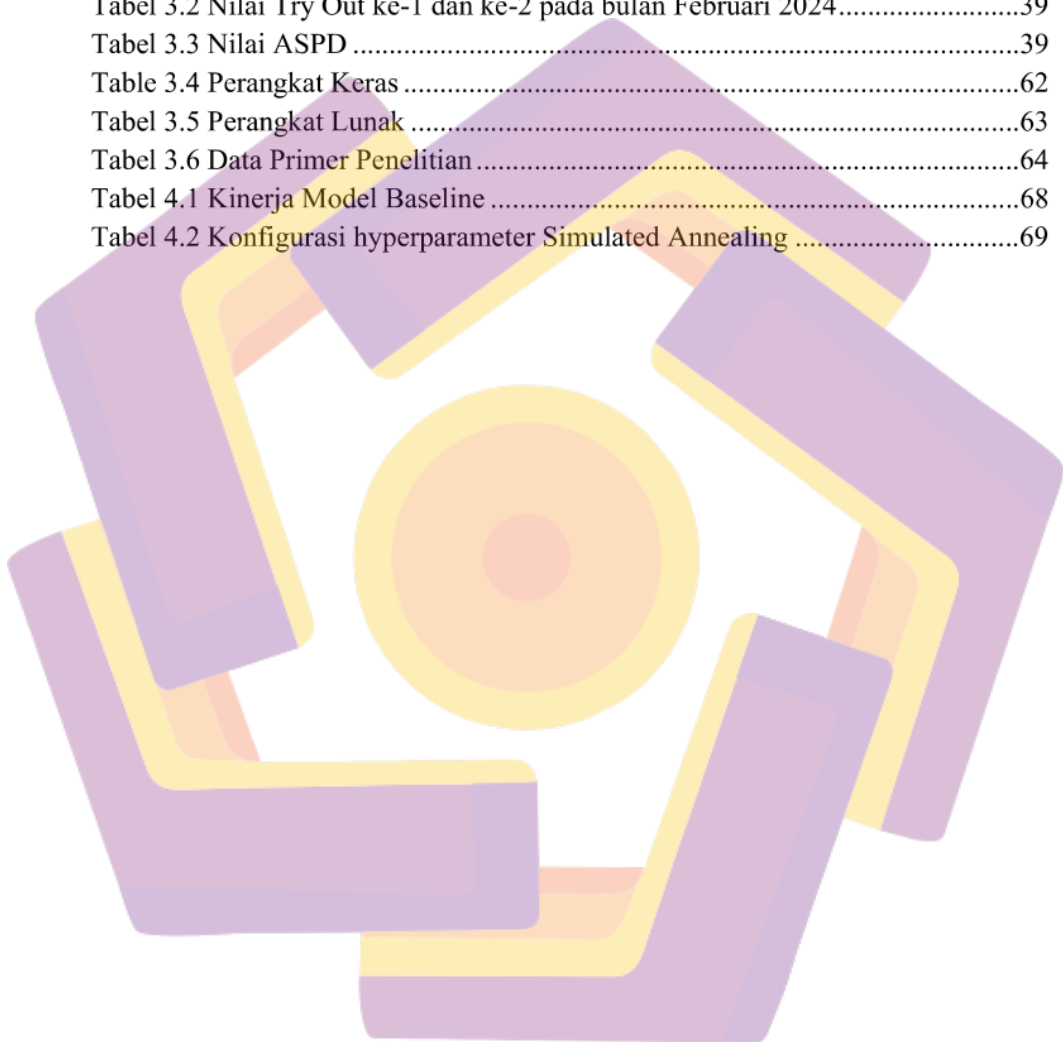
HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMBANG DAN SINGKATAN	xi
INTISARI	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	24
BAB III METODE PENELITIAN	35
3.1 Objek Penelitian.....	35
3.2 Alur Penelitian	36
3.3 Alat dan Bahan.....	62
BAB IV PEMBAHASAN.....	66
4.1 Analisis Data	66

4.2	Pengujian Model <i>Baseline</i>	67
4.3	Optimasi <i>Hyperparameter</i> dengan <i>Simulated Annealing</i>	69
4.4	Perbandingan Kinerja Model Setelah Optimasi.....	70
4.5	Diskusi Temuan	73
BAB V PENUTUPAN.....		75
5.1	Kesimpulan	75
5.2	Saran	76
REFERENSI		77
LAMPIRAN.....		79



DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	10
Tabel 3.1 Nilai Try Out ke-3 pada tanggal 5 – 8 Maret 2024.....	39
Tabel 3.2 Nilai Try Out ke-1 dan ke-2 pada bulan Februari 2024.....	39
Tabel 3.3 Nilai ASPD	39
Table 3.4 Perangkat Keras	62
Tabel 3.5 Perangkat Lunak	63
Tabel 3.6 Data Primer Penelitian	64
Tabel 4.1 Kinerja Model Baseline	68
Tabel 4.2 Konfigurasi hyperparameter Simulated Annealing	69



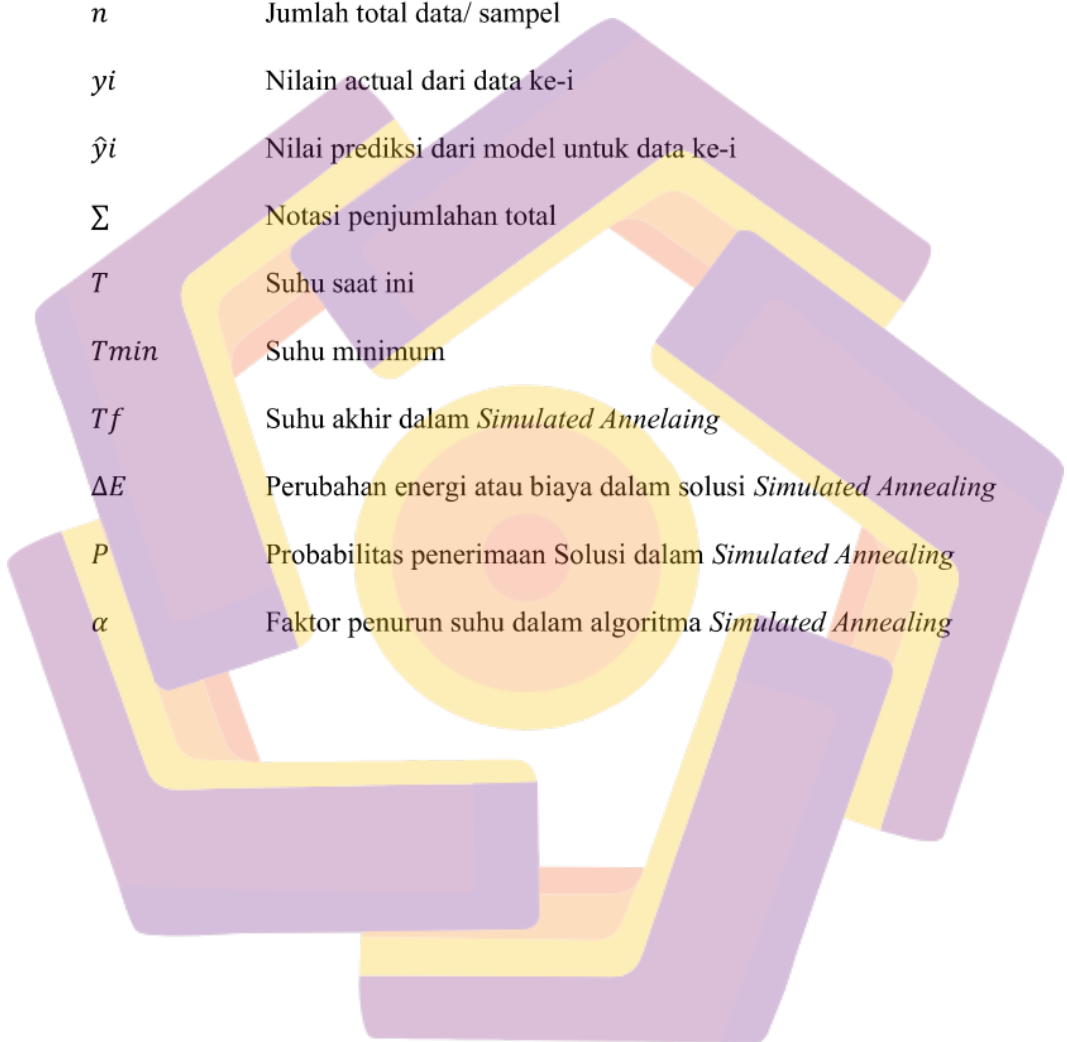
DAFTAR GAMBAR

Gambar 2.1 Cross-Industry Standard Process for Data Mining.....	24
Gambar 2.2 Arsitektur hasil algoritma Random Forest	25
Gambar 2.3 Arsitektur hasil algoritma Decision Tree	27
Gambar 2.4 Simulated Annealing Algorithm	29
Gambar 3.1 Alur Penelitian	37
Gambar 3.2 Dataset yang sudah dimerge.....	41
Gambar 3.3 Menampilkan data yang mengandung Missing Value	42
Gambar 3.4 Data setelah direname dan proses fitur engineering	43
Gambar 3.5 Data yang sudah ditambahkan kolom hasil Feature Engineering	45
Gambar 3.6 visualisasi Korelasi antar data.....	47
Gambar 3.7 Visualisasi fitur importance Random Forest MDI-oob.....	48
Gambar 3.8 kolom yang terpilih berdasarkan metode SelectKBest	49
Gambar 3.9 Parameter default baseline model	53
Gambar 4.1 Distribusi Target ASPD terhadap semua fitur.	67
Gambar 4.2 Perbandingan Matrik Evaluasi Random Forest dan Decision Tree ...	68
Gambar 4.3 Visualisasi perbandingan metrik MAE	70
Gambar 4.4 Visualisai perbandingan metrik MSE	71
Gambar 4.5 Visualisasi perbandingan metrik RMSE	71
Gambar 4.6 Visualisasi perbandingan metrik R^2 Score.....	72

DAFTAR LAMBANG DAN SINGKATAN



PISA	<i>Programme for International Student Assessment</i> (program penilaian siswa internasional).
OECD	<i>Organisation for Economic Co-operation and Development</i> (organisasi kerja sama ekonomi dan pembangunan).
ASPD	Asesmen Standardisasi Pendidikan Daerah (evaluasi kompetensi siswa di DIY).
SMP	Sekolah Menengah Pertama
MTs	Madrasah Tsanawiyah (jenjang pendidikan)
KNN	<i>K-Nearest Neighbor</i> (algoritma pembelajaran mesin).
SA	<i>Simulated Annealing</i> (algoritma optimasi).
TO1	Try Out 1 (nilai uji coba pertama).
LMS	<i>Learning Management System</i> (sistem manajemen pembelajaran).
XGBoost	<i>Extreme Gradient Boosting</i> (algoritma pembelajaran mesin).
LightGBM	<i>Light Gradient Boosting Machine</i> (algoritma pembelajaran mesin).
MAE	<i>Mean Absolute Error</i> (Rata-rata Kesalahan Absolut).
MSE	<i>Mean Squared Error</i> (Rata-rata Kuadrat Kesalahan).
RMSE	<i>Root Mean Squared Error</i> (Akar Kuadrat Rata-rata Kesalahan).
R^2	<i>R-squared</i> (Koefisien Determinasi).
ANN	<i>Artificial Neural Network</i>
MLP	<i>Multi-Layer Perceptron</i>
RF	<i>Random Forest</i>



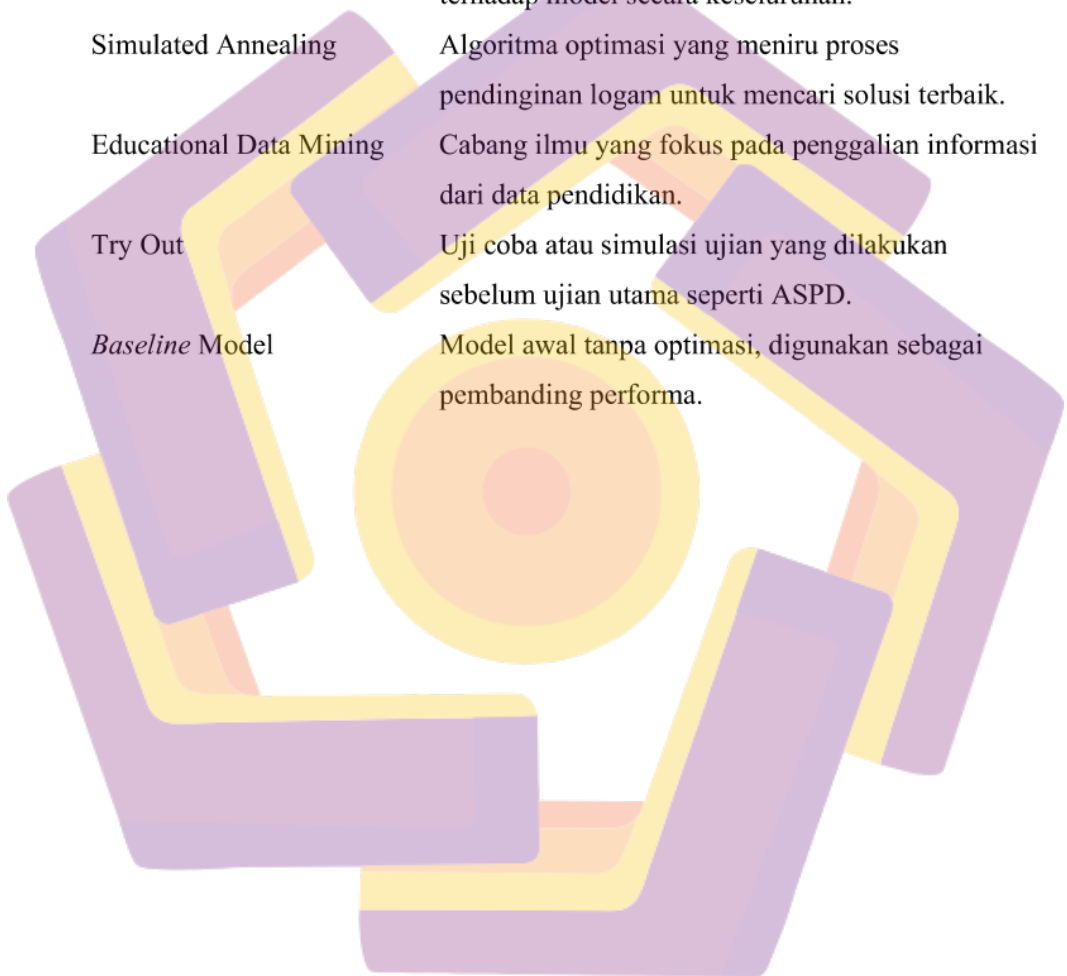
DT	<i>Decision Tree</i>
MDI-OOB	<i>Mean Decrease in Impurity – out of bag</i>
EDM	<i>Educational Data Mining</i>
n	Jumlah total data/ sampel
y_i	Nilain actual dari data ke-i
$\hat{y}_i$	Nilai prediksi dari model untuk data ke-i
Σ	Notasi penjumlahan total
T	Suhu saat ini
T_{min}	Suhu minimum
T_f	Suhu akhir dalam <i>Simulated Annelaing</i>
ΔE	Perubahan energi atau biaya dalam solusi <i>Simulated Annealing</i>
P	Probabilitas penerimaan Solusi dalam <i>Simulated Annealing</i>
α	Faktor penurun suhu dalam algoritma <i>Simulated Annealing</i>

DAFTAR ISTILAH



Literasi Membaca	Kemampuan memahami dan menggunakan teks tertulis
Literasi Sains	Kemampuan menggunakan pengetahuan ilmiah untuk memahami alam.
Numerasi	Kemampuan menggunakan angka dan matematika dasar untuk memecahkan masalah.
Intervensi Pembelajaran	Strategi untuk membantu atau meningkatkan hasil belajar siswa.
Data Science	Bidang interdisipliner yang mengekstrak wawasan dari data.
Machine Learning	Sistem yang belajar dari data untuk membuat keputusan.
Supervised Learning	Pembelajaran mesin dari data berlabel.
Decision Tree	Model prediksi yang menggunakan struktur pohon untuk keputusan.
Ensemble	Menggabungkan beberapa model untuk hasil prediksi yang lebih kuat.
<i>Overfitting</i>	Model terlalu cocok dengan data pelatihan, tidak baik untuk data baru.
Local Optimum	Titik solusi terbaik di area terbatas, bukan keseluruhan.
Metaheuristic	Metode optimasi tingkat tinggi untuk menemukan solusi yang baik.
Out of bag	Sampel data yang tidak digunakan untuk melatih pohon tertentu dalam <i>Random Forest</i> , bisa dipakai untuk evaluasi.
<i>Hyperparameter</i>	Parameter yang ditentukan sebelum proses pelatihan model machine learning.

Feature Selection	Proses memilih subset fitur yang paling relevan untuk model.
<i>Univariate</i> Selection	Pemilihan fitur satu per satu berdasarkan nilai statistik individual.
<i>Multivariate</i> Selection	Pemilihan fitur berdasarkan kontribusi kolektif terhadap model secara keseluruhan.
Simulated Annealing	Algoritma optimasi yang meniru proses pendinginan logam untuk mencari solusi terbaik.
Educational Data Mining	Cabang ilmu yang fokus pada penggalian informasi dari data pendidikan.
Try Out	Uji coba atau simulasi ujian yang dilakukan sebelum ujian utama seperti ASPD.
Baseline Model	Model awal tanpa optimasi, digunakan sebagai pembanding performa.



INTISARI

Penelitian ini bertujuan untuk mengembangkan dan memprediksi nilai Asesmen Standardisasi Pendidikan Daerah (ASPD) menggunakan pendekatan *machine learning* dengan algoritma *Random Forest* dan *Decision Tree*, serta mengoptimalkan performa model menggunakan metode *Simulated Annealing*. Data yang digunakan berasal dari 119 siswa SMP Negeri 2 Mlati pada tahun ajaran 2023/2024, mencakup nilai literasi membaca, literasi sains, dan numerasi dari hasil Try Out serta ASPD (Asesmen Standardisasi Pendidikan Daerah). Data diproses melalui tahap pembersihan, konversi tipe data, penanganan *missing value*, dan pemilihan fitur. Pemilihan fitur dilakukan dengan kombinasi metode *SelectKBest* (*Univariate*) dan *Importance Random Forest* (*multivariat*) untuk mengidentifikasi 6 fitur paling relevan: bahasa indonesia1, rata-rata Literasi1, ASPD Literasi, bahasa indonesia2, ASPD Numerasi, dan ASPD Sains. Model diuji dengan membagi data menjadi 80% data latih dan 20% data uji. Hasil penelitian menunjukkan bahwa model *Random Forest* dan *Decision Tree* yang dioptimalkan dengan *Simulated Annealing* mampu memprediksi nilai ASPD dengan akurasi sangat tinggi, yaitu *Random Forest* 94.55% dan *Decision Tree* 82.70%. Pendekatan ini diharapkan dapat mendukung peningkatan kualitas pendidikan dan mutu pembelajaran di tingkat SMP.

Kata Kunci: *Random Forest*, *Decision Tree*, *Simulated Annealing*, Prediksi Nilai

ABSTRACT

*This research aims to develop and predict the values of the Regional Education Standardization Assessment (ASPD) using a machine learning approach with Random Forest and Decision Tree algorithms, and to optimize model performance using the Simulated Annealing method. The data used was collected from 119 students of SMP Negeri 2 Mlati during the 2023/2024 academic year, including reading literacy, science literacy, and numeracy scores from Try Out results and ASPD. The data processing involved cleaning, data type conversion, handling missing values, and feature selection. Feature selection was performed using a combination of the *SelectKBest* (Univariate) and Importance Random Forest (multivariate) methods to identify the 6 most relevant features: Bahasa Indonesia 1, average Literacy 1, ASPD Literacy, Bahasa Indonesia 2, ASPD Numeracy, and ASPD Science. The model was tested by splitting the data into 80% training data and 20% testing data. The results show that the Random Forest and Decision Tree models, optimized with Simulated Annealing, were able to predict ASPD scores with high accuracy: Random Forest achieving 94.55% and Decision Tree 82.70%. This approach is expected to support the improvement of education quality and learning outcomes at the junior high school level.*

Keywords: *Random Forest, Decision Tree, Simulated Annealing, Value Prediction*