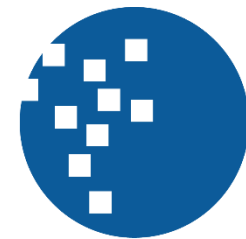


**PREDIKSI BENCHMARK SPEC CPU2017 MENGGUNAKAN
RANDOM FOREST**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

Tugas Akhir

BAYU BASKORO ADJIE

00000061617

**PROGRAM STUDI TEKNIK KOMPUTER
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

**PREDIKSI BENCHMARK SPEC CPU2017 MENGGUNAKAN
RANDOM FOREST**



Tugas Akhir

**Diajukan sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Teknik**

BAYU BASKORO ADJIE

00000061617

**PROGRAM STUDI TEKNIK KOMPUTER
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Bayu Baskoro Adjie

Nomor Induk Mahasiswa : 00000061617

Program Studi : Teknik Komputer

Skripsi dengan judul:

Prediksi Benchmark SPEC CPU2017 Menggunakan Random Forest

Merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 14 Januari 2026



Bayu Baskoro Adjie

HALAMAN PERSETUJUAN

Tugas Akhir dengan judul

PREDIKSI BENCHMARK SPEC CPU2017 MENGGUNAKAN RANDOM
FOREST

Oleh

Nama : Bayu Baskoro Adjie

NIM : 00000061617

Program Studi : Teknik Komputer

Fakultas : Teknik dan Informatika

Dengan susunan pembimbing sebagai berikut:

Pembimbing



Samuel, M.T.I
0304038902

Ketua Program Studi
Teknik Komputer



Samuel, M.T.I
0304038902

HALAMAN PENGESAHAN

Tugas Akhir dengan judul

Prediksi Benchmark SPEC CPU2017 Menggunakan Random Forest

Oleh

Nama : Bayu Baskoro Adjie
NIM : 00000061617
Program Studi : Teknik Komputer
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Rabu, 14 Januari 2026

Pukul 15.00 s.d 17.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang



Nabila Husna Shabrina, S.T., M.T.
0321099301

Penguji



Dareen Kusuma Halim, S.Kom., M.Eng.Sc.
0317129202

Pembimbing



Samuel, M.T.I
0304038902

Ketua Program Studi Teknik Komputer



Samuel, M.T.I

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Bayu Baskoro Adjie
NIM : 00000061617
Program Studi : Teknik Komputer
Jenjang : S1
Judul Karya Ilmiah : Prediksi Benchmark SPEC CPU2017 Menggunakan
Random Forest

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - ☐ Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - ☐ Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 14 Januari 2026



Bayu Baskoro Adjie

KATA PENGANTAR

Puji Syukur atas selesainya penulisan Tugas Akhir ini dengan judul: “Prediksi Benchmark SPEC CPU2017 Menggunakan Random Forest” dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar S1 Jurusan Teknik Komputer Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan tugas akhir ini, sangatlah sulit bagi saya untuk menyelesaikan tugas akhir ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Dr. Andrey Andoko, selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, selaku Dekan Fakultas Universitas Multimedia Nusantara.
3. Samuel Hutagalung, M.T.I, selaku Ketua Program Studi Universitas Multimedia Nusantara.
4. Samuel Hutagalung M.T.I, sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Teman-teman angkatan 2021 yang sedang berjuang bersama selama perkuliahan.

Semoga tugas akhir ini dapat digunakan sebagai informasi untuk pembaca.

Tangerang, 27 Januari 2026



Bayu Baskoro Adjie

PREDIKSI BENCHMARK SPEC CPU2017 MENGGUNAKAN RANDOM FOREST

Bayu Baskoro Adjie

ABSTRAK

Benchmark SPEC CPU2017 digunakan secara luas oleh akademisi, vendor prosesor, dan penyedia layanan komputasi sebagai standar evaluasi performa CPU. Namun, sebagaimana ditunjukkan dalam penelitian Meiliu Lu et al., pendekatan prediksi performa berbasis machine learning masih menghadapi keterbatasan ketika diaplikasikan pada konfigurasi perangkat keras yang tidak sepenuhnya terwakili dalam data pelatihan. Kondisi ini menjadi tantangan bagi pemangku kepentingan industri yang membutuhkan estimasi performa CPU secara cepat tanpa menjalankan benchmark secara langsung. Penelitian ini mengembangkan dan mengevaluasi model prediksi berbasis Random Forest untuk mengestimasi empat metrik utama SPEC CPU2017, yaitu CINT2017, CFP2017, CINT2017rate, dan CFP2017rate, berdasarkan spesifikasi perangkat keras. Model dilatih menggunakan dataset resmi SPEC CPU2017 melalui tahapan pra-pemrosesan, rekayasa fitur, dan seleksi fitur multi-tahap. Evaluasi performa dilakukan melalui validasi internal menggunakan skema hold-out dan cross-validation, serta pengujian kemampuan generalisasi model menggunakan data pengujian yang dipisahkan sejak awal dari data pelatihan. Hasil pengujian menunjukkan bahwa model Random Forest mampu mempertahankan akurasi prediksi yang tinggi pada tahap training dan validasi dengan nilai kesalahan yang relatif rendah. Pada tahap pengujian generalisasi, model masih mampu menghasilkan estimasi performa yang cukup mendekati nilai *benchmark* aktual, meskipun pada beberapa konfigurasi tertentu terjadi peningkatan deviasi prediksi. Temuan ini menunjukkan bahwa distribusi dan keragaman data pelatihan berperan penting dalam menentukan kemampuan generalisasi model prediksi performa CPU berbasis *machine learning*.

Kata kunci: SPEC CPU2017, Random Forest, Prediksi Performa CPU, Evaluasi Model, Generalisasi

BENCHMARK PREDICTION OF SPEC CPU2017 DATA USING RANDOM FOREST

Bayu Baskoro Adjie

ABSTRACT (English)

The SPEC CPU2017 benchmark is widely used by academics, processor vendors, and computing service providers as a standard for evaluating CPU performance. However, as shown in the research by Meiliu Lu et al., machine learning-based performance prediction approaches still face limitations when applied to hardware configurations that are not fully represented in the training data. This condition poses a challenge for industry stakeholders who need quick CPU performance estimates without running benchmarks directly. This study develops and evaluates a Random Forest-based prediction model to estimate four main SPEC CPU2017 metrics, namely CINT2017, CFP2017, CINT2017rate, and CFP2017rate, based on hardware specifications. The model was trained using the official SPEC CPU2017 dataset through pre-processing, feature engineering, and multi-stage feature selection. Performance evaluation was conducted through internal validation using a hold-out and cross-validation scheme, as well as testing the model's generalization ability using test data separated from the training data from the outset. The test results show that the Random Forest model is able to maintain high prediction accuracy at the training and validation stages with a relatively low error rate. During the generalization testing phase, the model was still able to produce performance estimates that were fairly close to the actual benchmark values, although there was an increase in prediction deviation in certain configurations. These findings indicate that the distribution and diversity of training data play an important role in determining the generalization capabilities of machine learning-based CPU performance prediction models.

Keywords: SPEC CPU2017, Random Forest, CPU Performance Prediction, Model Evaluation, Generalization

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	v
ABSTRAK	vii
<i>ABSTRACT (English)</i>	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Pertanyaan Penelitian	6
1.3 Batasan Penelitian	6
1.4 Tujuan Penelitian	7
1.5 Manfaat Penelitian	7
1.6 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	9
2.1 Justifikasi Solusi	11
2.1.1 Dataset (Pengembangan)	11
2.1.2 Jenis Benchmark (Adopsi)	11
2.1.3 Fitur Input (Pengembangan)	12
2.1.4 Fitur Output (Pengembangan)	12
2.1.5 Preprocessing Konversi (Pengembangan)	12
2.1.6 Preprocessing Scaling (Pengembangan)	13
2.1.7 Analisis Fitur (Pengembangan)	13
2.1.8 Split Data Testing: Training (Pengembangan)	13
2.1.9 Model & Struktur (Pengembangan)	14
2.1.10 Algoritma Pelatihan (Pengembangan)	14
2.1.11 Pengujian Model (Adopsi)	14
2.2 Tinjauan Teori	15
2.2.1 Benchmark	15
2.2.2 Random Forest	16
BAB III METODE PENELITIAN	17
3.1 Perancangan Solusi	17
3.1.1. Pemilihan Dataset dan Benchmark	18

3.1.2. Pre-Processing (Konversi, Encoding, Scaling)	19
3.1.3. Pengembangan Model	21
3.1.4. Evaluasi Performa Model pada Tahap Training dan Validasi	24
3.1.5. Pengujian Data	25
3.2 Metode Pengujian	25
3.2.1. Evaluasi Model	25
3.2.2. Lingkungan Pengujian Ground Truth	26
BAB IV IMPLEMENTASI DAN PENGUJIAN	27
4.1. Spesifikasi Sistem	27
4.2. Implementasi Solusi	27
4.2.1. Pengumpulan dan pembersihan data awal	27
4.2.2. Transformasi data dan Feature Engineering	28
4.2.3. Implementasi Seleksi Fitur	29
4.2.4. Implementasi Pelatihan dan Penskalaan Model	31
4.2.5. Evaluasi Performa Model pada Tahap Training dan Validasi	32
4.3. Pengujian dan Analisis Solusi	35
4.3.1. Perbandingan Aktual dengan Prediksi Model	35
4.3.2. Analisis Akurasi Generalisasi (Ground Truth)	41
BAB V SIMPULAN DAN SARAN	44
5.1 Simpulan	44
5.2 Saran	44
DAFTAR PUSTAKA	46
LAMPIRAN	50
Lampiran A - Turnitin (Wajib)	50
Lampiran B - Konsultasi Bimbingan (Wajib)	54
Lampiran C - Daftar Penggunaan AI (Wajib)	56

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR TABEL

Tabel 1.1 Teknologi atau Metode pada Penelitian Meiliu Lu	4
Tabel 2.1 Teknologi atau Metode yang dikembangkan	9
Tabel 3.1 Jumlah dataset SPEC CPU	18
Tabel 3.2 Skema Pembagian Data Training dan Testing	21
Tabel 3.3 Parameter model RF	22
Tabel 4.1 Hasil Evaluasi Validasi pada Model	32
Tabel 4.2 Perbandingan Aktual dan Prediksi Model (CINT2017)	36
Tabel 4.3 Perbandingan Aktual dan Prediksi Model (CFP2017)	37
Tabel 4.4 Perbandingan Aktual dan Prediksi Model (CINT2017rate)	38
Tabel 4.5 Perbandingan Aktual dan Prediksi Model (CFP2017rate)	40



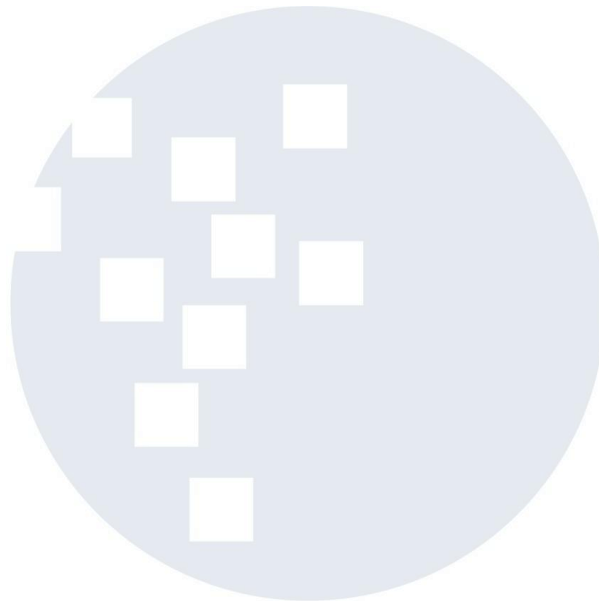
DAFTAR GAMBAR

Gambar 2.1 Diagram RF	16
Gambar 3.1 Alur penelitian	17
Gambar 4.1 Pembersihan kolom dan data	28
Gambar 4.2 Transformasi log1p pada fitur dan target	28
Gambar 4.3 Feature Engineering pada prosesor dan memori	29
Gambar 4.4 Penggunaan Target Mean Encoding	29
Gambar 4.5 Clustering K-Means	30
Gambar 4.6 Seleksi fitur Spearman dan Hierarchical	30
Gambar 4.7 RFECV	30
Gambar 4.8 Penerapan 5-Fold Cross Validation	31
Gambar 4.9 Scaling data dengan StandardScaler	31
Gambar 4.10 Grafik MAPE pada Model	33
Gambar 4.11 Grafik R2 pada Model	34



DAFTAR LAMPIRAN

Lampiran A - Turnitin (Wajib)	50
Lampiran B - Konsultasi Bimbingan (Wajib)	54
Lampiran C - Daftar Penggunaan AI (Wajib)	56



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA