

**PERBANDINGAN EFEKTIVITAS ALGORITMA GENETIKA
DAN RFECV UNTUK MENGOPTIMALKAN RANDOM
FOREST DALAM KLASIFIKASI KESUKSESAN VIDEO
GAME**

TUGAS AKHIR



**OLEH:
ZULVA FARABI
NIM. 1910817110002**

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, JUNI 2026**

**PERBANDINGAN EFEKTIVITAS ALGORITMA GENETIKA
DAN RFECV UNTUK MENGOPTIMALKAN RANDOM
FOREST DALAM KLASIFIKASI KESUKSESAN VIDEO
GAME**

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Sarjana Strata-1 Teknologi Informasi



Oleh:

ZULVA FARABI

NIM. 1910817110002

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS LAMBUNG MANGKURAT
BANJARMASIN, JUNI 2026**

PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Zulva Farabi

NIM : 1910817110002

Fakultas : Teknik

Prodi : Teknologi Informasi

Judul : PERBANDINGAN EFEKTIVITAS ALGORITMA
GENETIKA DAN RFECV UNTUK MENGOPTIMALKAN RANDOM FOREST
DALAM KLASIFIKASI KESUKSESAN VIDEO GAME

Pembimbing Utama : Ir. Eka Setya Wijaya, S.T., M.Kom.

Dengan ini saya menyatakan bahwa dalam Skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya, juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar rujukan.

Banjarmasin, 30 Juni 2026



Zulva Farabi

NIM. 1910817110002

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN
SKRIPSI PROGRAM STUDI S-1 TEKNOLOGI INFORMASI
Perbandingan Efektivitas Algoritma Genetika dan RFECV untuk
Mengoptimalkan Random Forest dalam Klasifikasi Kesuksesan Video Game
Layanan Fakultas Teknik Universitas Lambung Mangkurat
Oleh
Zulva Farabi (1910817110002)

Telah dipertahankan di depan Tim Penguji pada 25 Juni 2026 dan dinyatakan
LULUS

Komite Penguji :	
Ketua :	Muti'a Maulida, S.Kom., M.T.I NIP. 198810272019032013
Anggota 1 :	Muhammad Fajrian Noor, S.Kom., M.Kom. NIP. 199611092023211009
Anggota 2 :	Ir. Muhammad Alkaff, S.Kom., M.Kom., Ph.D. NIP. 198606132015041001
Pembimbing :	Ir. Eka Setya Wijaya, S.T., M.Kom. NIP. 198205082008011010
Utama	

Banjarbaru, 30 Juni 2026

Diketahui dan disahkan oleh:

Wakil Dekan Bidang Akademik
Fakultas Teknik ULM,

Koordinator Program Studi
S-1 Teknologi Informasi,



Dr. Mahamud, S.T., M.T.
NIP. 197401071998021001

Andrean Rizky Baskara, S.Kom., M.Kom.
NIP. 199307032019031011

LEMBAR PERSETUJUAN

LEMBAR PERSETUJUAN

PERBANDINGAN EFEKTIVITAS ALGORITMA GENETIKA DAN RFECV UNTUK
MENGOPTIMALKAN RANDOM FOREST DALAM KLASIFIKASI KESUKSESAN
VIDEO GAME

OLEH
ZULVA FARABI
NIM.1910817110002

Telah diperiksa dan terpenuhi semua persyaratan akademik, administrasi, dan
disetujui untuk dipertahankan di hadapan dewan penguji

Banjarmasin, 24 Juni 2026

Pembimbing Utama,



Ir. Eka Setya Wijaya, S.T., M.Kom.
NIP. 198205082008011010

ABSTRAK

Kesuksesan sebuah video game di platform Steam dapat diukur secara kuantitatif melalui pendapatan (revenue) yang dihasilkan. Penelitian ini bertujuan membandingkan efektivitas dua metode seleksi fitur, yaitu Algoritma Genetika (GA) dan Recursive Feature Elimination with Cross-Validation (RFECV), dalam mengoptimalkan model Random Forest untuk mengklasifikasikan kesuksesan video game. Dataset yang digunakan adalah 1.500 game dengan revenue tertinggi di Steam tahun 2024, dengan kelas "sukses" dan "tidak sukses" ditentukan berdasarkan nilai median revenue dataset. Penelitian dilakukan melalui tiga skenario, yaitu model Random Forest tanpa seleksi fitur (baseline), Random Forest dengan seleksi fitur GA, dan Random Forest dengan seleksi fitur RFECV. Hasil penelitian menunjukkan model baseline dengan 7 fitur menghasilkan akurasi 95,33%. Model RF + GA menghasilkan 4 fitur optimal dengan akurasi 96,00% namun memerlukan waktu komputasi 2.639,08 detik. Model RF + RFECV menghasilkan 3 fitur optimal dengan akurasi yang sama yaitu 96,00%, namun hanya membutuhkan waktu komputasi 19,04 detik. Kedua metode konvergen pada tiga fitur inti yang sama, yaitu `copiesSold`, `price`, dan `avgPlaytime`. Berdasarkan hasil tersebut, RFECV terbukti lebih unggul dibandingkan GA karena mampu mencapai performa klasifikasi yang setara dengan jumlah fitur lebih sedikit dan efisiensi waktu komputasi 138 kali lebih cepat, sehingga menjadi pilihan yang lebih praktis pada dataset berukuran kecil hingga menengah dengan ruang fitur yang terbatas.

Kata kunci: Algoritma Genetika, , klasifikasi kesuksesan video game, Random Forest, RFECV, seleksi fitur

ABSTRACT

The success of a video game on the Steam platform can be quantitatively measured through the revenue it generates. This study aims to compare the effectiveness of two feature selection methods, namely Genetic Algorithm (GA) and Recursive Feature Elimination with Cross-Validation (RFECV), in optimizing a Random Forest model to classify video game success. The dataset used consists of the 1,500 highest-revenue games on Steam in 2024, with the "successful" and "unsuccessful" classes determined based on the median revenue value of the dataset. The research was conducted through three scenarios: a Random Forest model without feature selection (baseline), Random Forest with GA-based feature selection, and Random Forest with RFECV-based feature selection. The results show that the baseline model with 7 features achieved an accuracy of 95.33%. The RF + GA model produced 4 optimal features with an accuracy of 96.00%, but required a computation time of 2,639.08 seconds. The RF + RFECV model produced 3 optimal features with the same accuracy of 96.00%, requiring only 19.04 seconds of computation time. Both methods converged on the same three core features, namely copiesSold, price, and avgPlaytime. Based on these results, RFECV proved superior to GA as it achieved equivalent classification performance with fewer features and was 138 times more computationally efficient, making it a more practical choice for small-to-medium-sized datasets with a limited feature space.

Keywords: feature selection, Genetic Algorithm, Random Forest, RFECV, video game success classification

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadiran Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul "Perbandingan Efektivitas Algoritma Genetika dan RFECV untuk Mengoptimalkan Random Forest dalam Klasifikasi Kesuksesan Video Game". Tugas Akhir ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata-1 pada Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Lambung Mangkurat.

Dalam proses penyusunan Tugas Akhir ini, penulis banyak mendapatkan bimbingan, arahan, serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Kedua orang tua serta keluarga penulis yang senantiasa memberikan doa, dukungan moral, maupun material tanpa henti selama penulis menempuh pendidikan.
2. Bapak Ir. Eka Setya Wijaya, S.T., M.Kom. selaku Dosen Pembimbing Utama yang telah meluangkan waktu, tenaga, dan pikiran untuk membimbing, memberikan arahan, serta masukan yang sangat berharga selama proses penyusunan Tugas Akhir ini.
3. Ibu Muti'a Maulida, S.Kom., M.T.I. selaku Ketua/Penguji I, Bapak Muhammad Fajrian Noor, S.Kom., M.Kom. selaku Penguji II, dan Bapak Ir. Muhammad Alkaff, S.Kom., M.Kom., Ph.D. selaku Penguji III, yang telah memberikan kritik dan saran yang membangun demi penyempurnaan Tugas Akhir ini.
4. Bapak Andreyan Rizky Baskara, S.Kom., M.Kom. selaku Koordinator Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Lambung Mangkurat.
5. Seluruh dosen dan staf Program Studi Teknologi Informasi Universitas Lambung Mangkurat yang telah memberikan ilmu dan bimbingan selama masa perkuliahan.

6. Teman-teman seperjuangan Program Studi Teknologi Informasi angkatan 2019 yang selalu memberikan semangat, bantuan, dan diskusi selama proses pengerjaan Tugas Akhir ini.
7. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah membantu baik secara langsung maupun tidak langsung dalam penyelesaian Tugas Akhir ini.

Tugas Akhir ini terdiri dari beberapa bab dengan sistematika penulisan sebagai berikut:

- BAB I: Pendahuluan, yang mencakup latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, dan manfaat penelitian.
- BAB II: Tinjauan Pustaka, berisi penelitian terkait, landasan teori, dan kerangka pemikiran yang mendasari penelitian.
- BAB III: Metode Penelitian, meliputi alat penelitian, bahan penelitian, dan prosedur penelitian yang digunakan.
- BAB IV: Hasil dan Pembahasan, yang merupakan hasil penelitian serta analisisnya.
- BAB V: Kesimpulan dan Saran, yang berisi kesimpulan dari hasil penelitian serta saran untuk pengembangan lebih lanjut.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari kesempurnaan dan masih terdapat kekurangan baik dari segi penulisan maupun substansi. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga Tugas Akhir ini dapat memberikan manfaat bagi penulis khususnya, dan bagi pembaca serta pengembangan ilmu pengetahuan pada umumnya.

Banjarmasin, 30 Juni 2026

Penulis,



ZULVA FARABI

DAFTAR ISI

SAMPUL DALAM.....	i
PERNYATAAN.....	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR PERSETUJUAN.....	iv
ABSTRAK.....	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN.....	xvi
RIWAYAT HIDUP.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terkait.....	6
2.1.1 <i>A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers</i>	6

2.1.2 Penerapan Sentimen Analisis Acara Televisi Pada Twitter Menggunakan <i>Support Vector Machine</i> dan Algoritma Genetika sebagai Metode Seleksi Fitur.....	7
2.1.3 <i>The Effect of Recursive Feature Elimination with Cross-Validation Method on Classification Performance with Different Sizes of Datasets</i>	7
2.1.4 Genetika Optimalisasi Fitur Seleksi Pada Kismis Di Turki.....	8
2.1.5 Implementasi Algoritma Genetika untuk Seleksi Fitur pada Klasifikasi Genre Musik Menggunakan Metode Random Forest....	9
2.2 Landasan Teori.....	11
2.2.1 Video game.....	11
2.2.2 Steam.....	11
2.2.3 Random Forest.....	13
2.2.4 Seleksi Fitur.....	13
2.2.5 Algoritma Genetika.....	15
2.2.6 Recursive Feature Elimination with Cross-Validation (RFECV). .	19
2.2.7 Evaluasi Model.....	21
2.2.8 Definisi Kesuksesan Game dan Justifikasi Threshold Revenue....	21
2.2 Kerangka Pemikiran.....	22
BAB III METODE PENELITIAN.....	23
3.1 Alat Penelitian.....	23
3.2 Bahan Penelitian.....	24
3.3 Prosedur Penelitian.....	25
BAB IV HASIL DAN PEMBAHASAN.....	31
4.1 Pengumpulan Data.....	31
4.2 Hasil Preprocessing.....	32

4.2.1 Analisis Distribusi Revenue dan Pelabelan Kelas.....	32
4.2.2 Hasil Encoding Fitur Kategorikal.....	34
4.2.3 Hasil Normalisasi Min-Max Scaling.....	34
4.2.4 Pemeriksaan Missing Values Setelah Preprocessing.....	35
4.2.5 Hasil Pembagian Data dan Pemeriksaan Class Balance.....	36
4.3 Random Forest Tanpa Seleksi Fitur (Baseline).....	37
4.3.1 Konfigurasi dan Hasil Pelatihan.....	37
4.3.2 Hasil Evaluasi Model.....	37
4.3.3 Analisis Confusion Matrix.....	38
4.3.4 Feature Importance.....	38
4.4 Random Forest dengan Seleksi Fitur RFECV.....	40
4.4.1 Proses Seleksi Fitur.....	40
4.4.2 Penentuan Jumlah Fitur Optimal.....	40
4.4.3 Hasil Evaluasi Model.....	41
4.4.4 Analisis Confusion Matrix.....	42
4.5 Random Forest dengan Seleksi Fitur Algoritma Genetika.....	43
4.5.1 Konfigurasi dan Proses Evolusi.....	43
4.5.2 Analisis Konvergensi.....	44
4.5.3 Hasil Evaluasi Model.....	44
4.5.4 Analisis Confusion Matrix.....	45
4.6 Perbandingan Hasil Ketiga Skenario.....	46
4.6.1 Tabel Rekapitulasi.....	46
4.6.2 Perbandingan Akurasi dan F1-Score.....	46
4.6.3 Perbandingan Multi-Metrik (Radar Chart).....	47
4.6.4 Perbandingan Jumlah Fitur dan Efisiensi.....	47

4.6.5 Perbandingan Confusion Matrix.....	48
4.7 Implementasi Sistem.....	49
4.7.1 Menu Utama Sistem.....	49
4.7.2 Implementasi Load dan Eksplorasi Data.....	49
4.7.3 Implementasi Preprocessing.....	50
4.7.4 Implementasi Jalankan Skenario.....	50
4.7.5 Implementasi Lihat Hasil dan Bandingkan.....	51
4.7.6 Implementasi Export Hasil.....	51
4.8 Pembahasan.....	52
4.8.1 Analisis Performa Model Baseline.....	52
4.8.2 Analisis Performa Model RF + GA.....	53
4.8.3 Analisis Performa Model RF + RFECV.....	53
4.8.4 Perbandingan GA vs RFECV.....	54
4.8.5 Keterbatasan Penelitian.....	55
BAB V KESIMPULAN DAN SARAN.....	56
5.1 Kesimpulan.....	56
5.2 Saran.....	57
DAFTAR PUSTAKA.....	59
LAMPIRAN.....	63

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait.....	10
Tabel 2. 2 Rumusan Confusion Matrix.....	21
Tabel 3. 1 Daftar alat yang digunakan.....	23
Tabel 3. 2 Daftar library python yang digunakan.....	24
Tabel 3. 3 Uraian dataset yang digunakan.....	25
Tabel 3. 4 Implementasi RFECV.....	28
Tabel 3. 5 Implementasi Algoritma Genetika.....	29
Tabel 4. 1 Evaluasi model baseline Random Forest.....	37
Tabel 4. 2 Evaluasi model RF + RFECV.....	41
Tabel 4. 3 Evaluasi model RF + GA.....	44
Tabel 4. 4 Hasil rekapitulasi ketiga skenario.....	46

DAFTAR GAMBAR

Gambar 2. 1 Logo Steam.....	11
Gambar 2. 2 Ilustrasi Random Forest [24].....	13
Gambar 2. 3 Ilustrasi Algoritma Genetika [24].....	15
Gambar 2. 4 Cara kerja RFECV [31].....	20
Gambar 2. 5 Kerangka Pemikiran.....	22
Gambar 3. 1 Alur Penelitian.....	26
Gambar 4. 1 Sampel dataset yang digunakan.....	31
Gambar 4. 2 Pemeriksaan tipe data.....	31
Gambar 4. 3 Pemeriksaan missing value dataset.....	32
Gambar 4. 4 Statistik deskriptif dataset.....	32
Gambar 4. 5 Distribusi revenue yang bersifat right-skewed.....	33
Gambar 4. 6 Mean dan median revenue dataset.....	33
Gambar 4. 7 Encoding fitur kategorikal.....	34
Gambar 4. 8 Hasil normalisasi dataset.....	35
Gambar 4. 9 Missing value setelah preprocessing.....	35
Gambar 4. 10 Pembagian data training dan testing.....	36
Gambar 4. 11 Confusion matrix model Random Forest Baseline.....	38
Gambar 4. 12 Kepentingan fitur model baseline.....	39
Gambar 4. 13 Chart kepentingan fitur baseline.....	39
Gambar 4. 14 Seleksi fitur RFECV.....	40
Gambar 4. 15 Grafik akurasi cross-validation vs jumlah fitur (RFECV).....	40
Gambar 4. 16 Confusion matrix model RF + RFECV.....	42
Gambar 4. 17 Parameter Algoritma Genetika.....	43

Gambar 4. 18 Pemilihan fitur terbaik Algoritma Genetika.....	43
Gambar 4. 19 Grafik konvergensi fitness Algoritma Genetika selama 100 generasi.....	44
Gambar 4. 20 Confusion matrix model RF + GA.....	45
Gambar 4. 21 Chart perbandingan performa ketiga skenario.....	46
Gambar 4. 22 Radar chart perbandingan multi metrik.....	47
Gambar 4. 23 Chart perbandingan jumlah fitur ketiga skenario.....	47
Gambar 4. 24 Perbandingan Confusion Matrix ketiga skenario.....	48
Gambar 4. 25 Tampilan menu utama sistem	46
Gambar 4. 26 Implementasi Load & Eksplorasi Data	46
Gambar 4. 27 Implementasi Preprocessing	47
Gambar 4. 28 Implementasi Jalankan Skenario	47
Gambar 4. 29 Implementasi Lihat Hasil & Bandingkan	48
Gambar 4. 30 Implementasi Export Hasil	48

DAFTAR LAMPIRAN

Lampiran 1. Lembar Konsultasi.....	63
------------------------------------	----