

**EX-BOOST: KERANGKA KERJA TERINTEGRASI  
*EXPLORATORY DATA ANALYSIS* DAN *EXTREME  
GRADIENT BOOSTING* UNTUK PERAMALAN  
PELUANG KERJA DI MASA DEPAN**

**LAPORAN NON-REGULER**

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



Disusun oleh:

**NOUVAL AMMAR IBRAHIM**

**22.11.5176**

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

**EX-BOOST: KERANGKA KERJA TERINTEGRASI  
*EXPLORATORY DATA ANALYSIS* DAN *EXTREME  
GRADIENT BOOSTING* UNTUK PERAMALAN  
PELUANG KERJA DI MASA DEPAN**

**LAPORAN NON-REGULER**

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



Disusun oleh :

**NOUVAL AMMAR IBRAHIM**

**22.11.5176**

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

HALAMAN PERSETUJUAN

JALUR NON-REGULER

**EX-BOOST: KERANGKA KERJA TERINTEGRASI *EXPLORATORY*  
*DATA ANALYSIS* DAN *EXTREME GRADIENT BOOSTING* UNTUK  
PERAMALAN PELUANG KERJA DI MASA DEPAN**

yang disusun dan diajukan oleh

**Nouval Ammar Ibrahim**

**22.11.5176**

telah disetujui oleh Dosen Pembimbing  
pada tanggal <23 Februari 2026>

**Dosen Pembimbing,**



**Kusnawi, S.Kom, M.Eng**  
**NIK. 190302112**

HALAMAN PENGESAHAN

JALUR NON-REGULER

EX-BOOST: KERANGKA KERJA TERINTEGRASI *EXPLORATORY*  
DATA ANALYSIS DAN *EXTREME GRADIENT BOOSTING* UNTUK  
PERAMALAN PELUANG KERJA DI MASA DEPAN

yang disusun dan diajukan oleh

**Nouval Ammar Ibrahim**

22.11.5176

Telah dipertahankan di depan Dewan Penguji  
pada tanggal <23 Februari 2026>

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

Eli Pujiastuti, S.Kom., M.Kom.

NIK. 190302227

Subektiningsih, S.Kom., M.Kom.

NIK. 190302413

Kusnawi, S.Kom., M.Eng.

NIK. 190302112

Laporan ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal < 23 Februari 2026 >

**DEKAN FAKULTAS ILMU KOMPUTER**



Prof. Dr. Kusnini, M.Kom.

NIK. 190302106

## HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Nouval Ammar Ibrahim

NIM : 22.11.5176

Menyatakan bahwa Laporan dengan judul berikut:

**EX-Boost: Kerangka Kerja Terintegrasi *Exploratory Data Analysis* dan *Extreme Gradient Boosting* untuk Peramalan Peluang Kerja di Masa Depan**

Dosen Pembimbing : Kusnawi, S. Kom , M.Eng.

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 kegiatan **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 ketidak-benaran 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, 23 Februari 2026

  
Nouval Ammar Ibrahim

## HALAMAN PERSEMBAHAN

Dengan segala kerendahan dan ketulusan hati penulis mempersembahkan karya ini kepada:

1. Program Studi Informatika, Universitas Amikom Yogyakarta, yang telah memberikan kesempatan bagi saya untuk menempuh dan memperoleh ilmu.
2. Ayah dan ibuku tercinta yang selalu memberikan doanya dan semangat.
3. Kakak-kakakku tercinta yang selalu mendukung, dan memberikan semangat dan doanya.
4. Semua pihak yang telah memberikan dukungan, membantu penulis dan tidak dapat disebutkan satu persatu.

Semoga Tuhan Yang Maha Esa senantiasa memberikan limpahan rahmat dan karunia atas segala kebbaikannya. Akhirnya penulis berharap semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan segenap pembaca.

Yogyakarta, 23 Februari 2026

Nouval Ammar Ibrahim

## KATA PENGANTAR

Puji syukur kita panjatkan kehadiran Allah Swt. yang telah memberikan rahmat dan hidayah-Nya sehingga saya dapat menyelesaikan proposal skripsi yang berjudul **"EX-Boost: Kerangka Kerja Terintegrasi *Exploratory Data Analysis* dan *Extreme Gradient Boosting* untuk Peramalan Peluang Kerja di Masa Depan"**. Dengan segala kerendahan dan ketulusan hati penulis menyampaikan penghargaan dan ucapan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Prof. Dr.Kusrini, M.Kom., selaku Dekan Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.
3. Bapak Uyock Anggoro Saputro, S.Kom., M.Kom., selaku Ketua Program Studi Informatika, Universitas Amikom Yogyakarta, yang telah memberikan bimbingan dan arahan selama penulis menempuh Pendidikan.
4. Bapak Kusnawi, S. Kom , M.Eng., yang telah memberikan bimbingan dan arahan selama penulis menyelesaikan Tugas Akhir.

Semoga Tuhan Yang Maha Esa senantiasa memberikan limpahan rahmat dan karunia atas segala kebaikannya. Untuk mengembangkan karya ini, dengan segala kerendahan hati penulis mengharapkan kritik dan saran untuk hasil yang lebih baik. Akhirnya penulis berharap semoga tulisan ini dapat bermanfaat bagi ilmu pengetahuan dan segenap pembaca.

Yogyakarta, 23 Februari 2026

Nouval Ammar Ibrahim

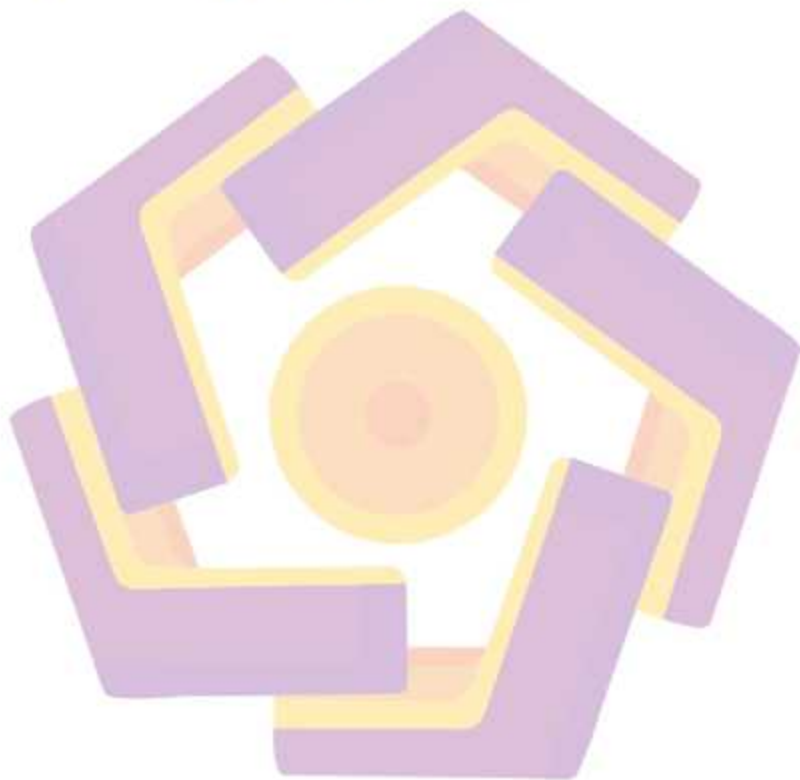
## DAFTAR ISI

|                                        |      |
|----------------------------------------|------|
| Halaman Judul.....                     | i    |
| Halaman Persetujuan.....               | ii   |
| Halaman Pengesahan .....               | iii  |
| Halaman Pernyataan Keaslian Karya..... | iv   |
| Halaman Persembahan .....              | v    |
| Kata Pengantar .....                   | vi   |
| Daftar Isi.....                        | vii  |
| Daftar Tabel.....                      | ix   |
| Daftar Gambar.....                     | x    |
| Daftar Lampiran.....                   | xi   |
| Daftar Lambang dan Singkatan.....      | xii  |
| Daftar Istilah.....                    | xiv  |
| Intisari .....                         | xvi  |
| <i>Abstract</i> .....                  | xvii |
| Bab I Pendahuluan .....                | 1    |
| 1.1. Gambaran Umum.....                | 1    |
| 1.2. Rumusan Masalah .....             | 2    |
| 1.3. Batasan Masalah.....              | 2    |
| 1.4. Tujuan.....                       | 3    |
| Bab II Tinjauan Pustaka .....          | 4    |
| 2.1. Studi Literatur .....             | 4    |
| 2.2. Landasan Teori .....              | 6    |
| 2.2.1 Data Collection: .....           | 6    |
| 2.2.2 Data preprocessing and EDA.....  | 6    |
| 2.2.3 Feature Engineering .....        | 7    |
| 2.2.4 XGBoost.....                     | 7    |
| 2.2.5 Feature Importance.....          | 9    |
| 2.2.6 Interpretation and Insight ..... | 10   |

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| BAB III Metode Penelitian .....                                       | 11 |
| 3.1. Metode.....                                                      | 11 |
| BAB IV Pembahasan .....                                               | 16 |
| 4.1. Pembahasan Hasil dan Analisis Eksperimental.....                 | 16 |
| BAB V Kesimpulan .....                                                | 31 |
| 5.1. Kesimpulan .....                                                 | 31 |
| 5.2. Saran.....                                                       | 31 |
| Referensi .....                                                       | 33 |
| Curriculum Vitae .....                                                | 37 |
| Lampiran dan Bukti Pendukung.....                                     | 38 |
| a. Letter of Acceptance (LOA).....                                    | 38 |
| b. Lembar Review .....                                                | 39 |
| c. Bukti Terbit/Terindex [wajib, sertakan tampilan bukti terbit]..... | 39 |
| d. Sertifikat sebagai Penyaji.....                                    | 46 |
| e. Bukti pembayaran .....                                             | 46 |

## DAFTAR TABEL

|                                                                                                                       |    |
|-----------------------------------------------------------------------------------------------------------------------|----|
| Tabel 4.1. Model accuracy results.....                                                                                | 19 |
| Tabel 4.2. 10 Jobs with the Highest Probability of Surviving probability (year 1 – year 5) according to EX-Boost..... | 20 |
| Tabel 4.3. Most Important Future Skills according to EX-Boost .....                                                   | 21 |

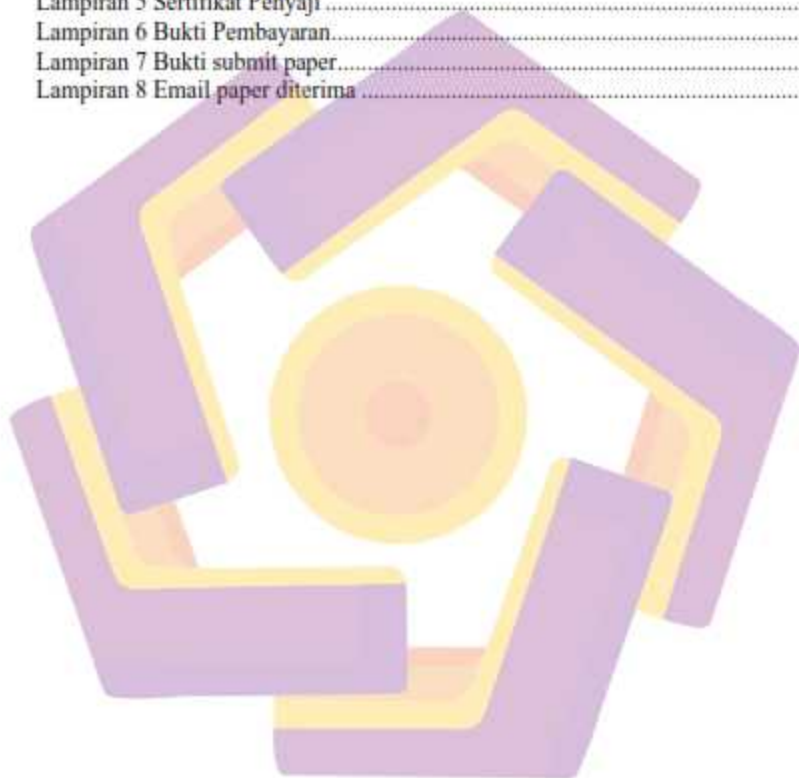


## DAFTAR GAMBAR

|                                                                                                                     |    |
|---------------------------------------------------------------------------------------------------------------------|----|
| Gambar 3.1 Kerangka kerja EX-Boost .....                                                                            | 11 |
| Gambar 4.1 Hasil pelabelan keterampilan masa depan .....                                                            | 16 |
| Gambar 4.2 Hasil pemeringkatan fitur (keterampilan) berdasarkan pengaruhnya terhadap label yang dipertahankan ..... | 16 |
| Gambar 4.3 Visualisasi 10 Keterampilan Masa Depan Terpenting .....                                                  | 17 |
| Gambar 4.4 Visualisasi Keterampilan Paling Dominan pada Pekerjaan yang Tersisa .....                                | 17 |
| Gambar 4.5 Perbandingan Hasil Klasifikasi .....                                                                     | 18 |
| Gambar 4.6 Probabilitas Kelangsungan Hidup Rata-Rata (Tahun 1 - Tahun 5) ...                                        | 22 |
| Gambar 4.7 Jumlah dan persentase missing value pada setiap kolom dataset.....                                       | 23 |
| Gambar 4.8 Menghilangkan missing value pada dataset .....                                                           | 24 |
| Gambar 4.9 Proses labeling pada data .....                                                                          | 24 |
| Gambar 4.10 Proses pemeringkatan fitur keterampilan .....                                                           | 25 |
| Gambar 4.11 Proses pelatihan dan evaluasi dengan Random Forest .....                                                | 26 |
| Gambar 4.12 Proses pelatihan dan evaluasi dengan Logistic Regression .....                                          | 27 |
| Gambar 4.13 Proses pelatihan dan evaluasi dengan XGBoost .....                                                      | 28 |
| Gambar 4.14 Hasil Cross-Validation .....                                                                            | 29 |
| Gambar 4.15 Ringkasan perbandingan model .....                                                                      | 29 |

## DAFTAR LAMPIRAN

|                                             |    |
|---------------------------------------------|----|
| Lampiran 1 Letter of Acceptance (LOA) ..... | 38 |
| Lampiran 2 Lembar Review .....              | 39 |
| Lampiran 3 Publisher .....                  | 39 |
| Lampiran 4 Conference Paper .....           | 40 |
| Lampiran 5 Sertifikat Penyaji .....         | 46 |
| Lampiran 6 Bukti Pembayaran .....           | 46 |
| Lampiran 7 Bukti submit paper .....         | 47 |
| Lampiran 8 Email paper diterima .....       | 47 |

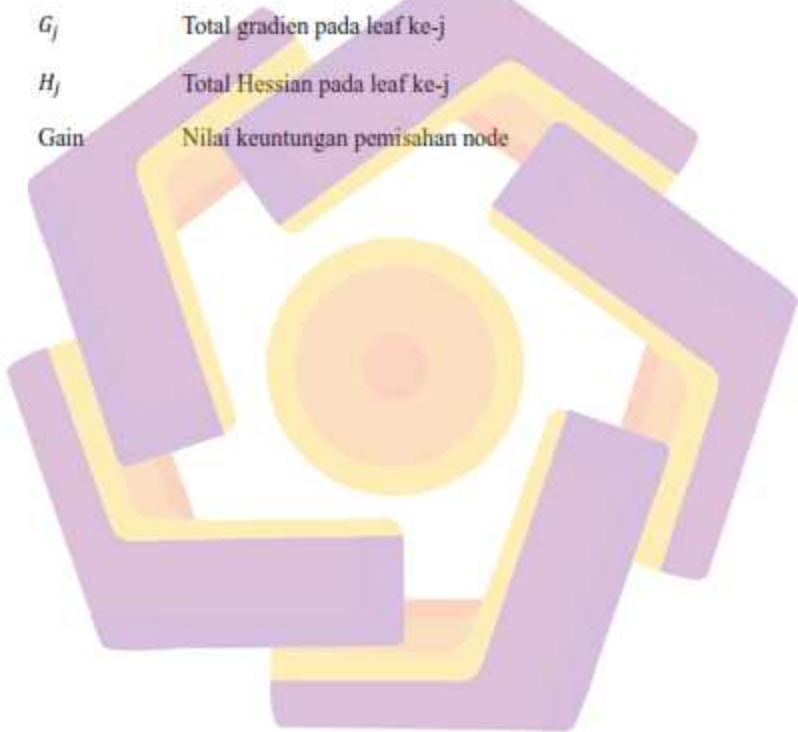


## DAFTAR LAMBANG DAN SINGKATAN



|             |                                                      |
|-------------|------------------------------------------------------|
| SVM         | <i>Support Vector Machines</i>                       |
| AI          | <i>Artificial Intelligence</i>                       |
| CV          | <i>Cross-Validation</i>                              |
| EDA         | <i>Exploratory Data Analysis</i>                     |
| EX-Boost    | <i>Exploratory Data Analysis + XGBoost Framework</i> |
| HR          | <i>Human Resources</i>                               |
| IT          | <i>Information Technology</i>                        |
| MAE         | <i>Mean Absolute Error</i>                           |
| MAPE        | <i>Mean Absolute Percentage Error</i>                |
| RF          | <i>Random Forest</i>                                 |
| SVM         | <i>Support Vector Machine</i>                        |
| XGBoost     | <i>Extreme Gradient Boosting</i>                     |
| $x_i$       | Data ke- $i$                                         |
| $y_i$       | Nilai aktual data ke- $i$                            |
| $\hat{y}_i$ | Nilai prediksi data ke- $i$                          |
| $n$         | Jumlah data                                          |
| $L$         | Fungsi loss                                          |
| $f_k(x_i)$  | Pohon regresi ke- $k$                                |
| $K$         | Jumlah pohon (trees)                                 |
| $\Omega(f)$ | Fungsi regularisasi                                  |
| $\gamma$    | Parameter penalti jumlah leaf                        |

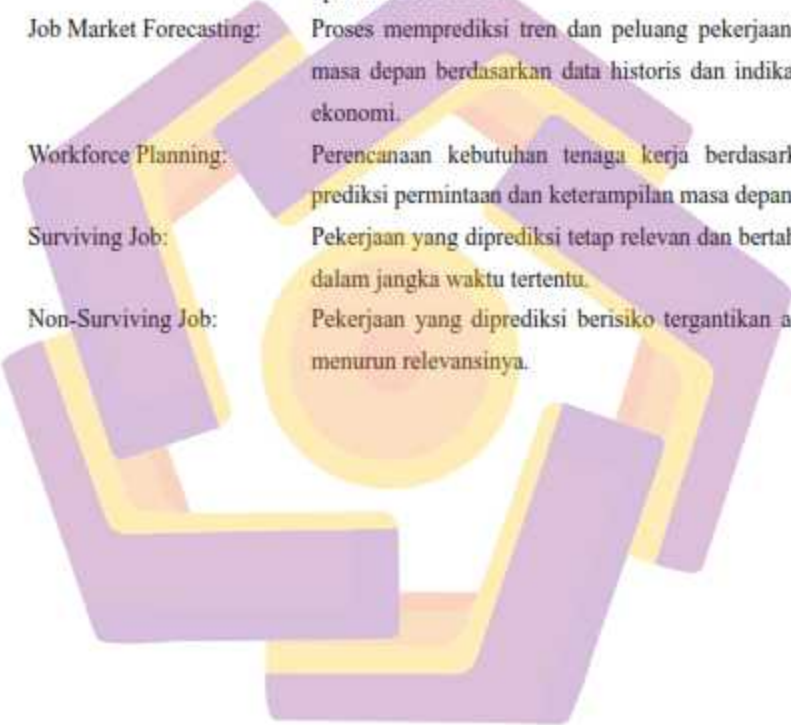
|           |                                 |
|-----------|---------------------------------|
| $\lambda$ | Parameter regularisasi L2       |
| $T$       | Jumlah leaf pada satu pohon     |
| $w_j$     | Bobot leaf ke-j                 |
| $g_i$     | Gradien orde pertama            |
| $h_i$     | Gradien orde kedua (Hessian)    |
| $G_j$     | Total gradien pada leaf ke-j    |
| $H_j$     | Total Hessian pada leaf ke-j    |
| Gain      | Nilai keuntungan pemisahan node |



## DAFTAR ISTILAH



|                                             |                                                                                                                                                              |
|---------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Artificial Intelligence (AI):</i>        | Bidang ilmu komputer yang berfokus pada pengembangan sistem yang mampu meniru kecerdasan manusia dalam pengambilan keputusan.                                |
| <i>Exploratory Data Analysis (EDA):</i>     | Proses analisis awal data untuk memahami pola, tren, anomali, dan karakteristik data sebelum pemodelan.                                                      |
| <i>Extreme Gradient Boosting (XGBoost):</i> | Algoritma machine learning berbasis ensemble boosting yang menggunakan pohon keputusan dan regularisasi untuk meningkatkan akurasi dan mencegah overfitting. |
| <i>EX-Boost:</i>                            | Kerangka kerja hibrida yang mengintegrasikan EDA dan XGBoost untuk peramalan peluang kerja masa depan.                                                       |
| <i>Overfitting:</i>                         | Kondisi ketika model terlalu menyesuaikan data latih sehingga kinerjanya menurun pada data baru.                                                             |
| <i>Cross-Validation (CV):</i>               | Teknik evaluasi model dengan membagi data ke beberapa subset untuk menguji stabilitas dan generalisasi model.                                                |
| <i>Feature Engineering:</i>                 | Proses menciptakan atau memodifikasi fitur dari data mentah untuk meningkatkan performa model.                                                               |
| <i>Feature Importance:</i>                  | Ukuran kontribusi setiap fitur terhadap hasil prediksi model.                                                                                                |
| <i>Ensemble Learning:</i>                   | Pendekatan pembelajaran mesin yang menggabungkan beberapa model untuk memperoleh hasil prediksi yang lebih baik.                                             |
| <i>Regularization:</i>                      | Teknik untuk membatasi kompleksitas model agar tidak overfitting.                                                                                            |



|                         |                                                                                                              |
|-------------------------|--------------------------------------------------------------------------------------------------------------|
| Gain Function:          | Fungsi dalam XGBoost yang digunakan untuk menentukan pemisahan terbaik pada node pohon keputusan.            |
| Gradient:               | Turunan pertama dari fungsi loss yang digunakan untuk memperbaiki kesalahan prediksi.                        |
| Hessian:                | Turunan kedua dari fungsi loss yang digunakan dalam optimasi XGBoost.                                        |
| Job Market Forecasting: | Proses memprediksi tren dan peluang pekerjaan di masa depan berdasarkan data historis dan indikator ekonomi. |
| Workforce Planning:     | Perencanaan kebutuhan tenaga kerja berdasarkan prediksi permintaan dan keterampilan masa depan.              |
| Surviving Job:          | Pekerjaan yang diprediksi tetap relevan dan bertahan dalam jangka waktu tertentu.                            |
| Non-Surviving Job:      | Pekerjaan yang diprediksi berisiko tergantikan atau menurun relevansinya.                                    |

## INTISARI

Perubahan tren ekonomi global telah secara signifikan mengubah dinamika pasar tenaga kerja, mengubah struktur ketenagakerjaan, dan mempercepat perubahan dalam persyaratan keterampilan. Masalah ini menyoroti kebutuhan mendesak akan sistem peramalan yang akurat dan andal untuk memprediksi peluang kerja di masa depan dan mendukung perencanaan tenaga kerja berbasis bukti. Metode peramalan tradisional, yang utama mengandalkan model linier atau statistik, seringkali tidak memadai untuk menangkap hubungan nonlinier dan pola kompleks dalam data multidimensi dan multisumber. Studi ini bertujuan untuk mengatasi batasan-batasan tersebut dengan memperkenalkan EX-Boost, kerangka kerja prediktif terintegrasi yang menggabungkan Exploratory Data Analysis (EDA) dengan algoritma *Extreme Gradient Boosting* (XGBoost). Kerangka kerja ini memanfaatkan data tenaga kerja multi-sumber, termasuk statistik historis, pengumuman lowongan kerja online, dan indikator makroekonomi yang dikompilasi dari O\*NET Online dan laporan tenaga kerja resmi. Model yang diusulkan menerapkan EDA untuk mendeteksi anomali, mengidentifikasi tren, dan memilih fitur berdampak tinggi, sementara XGBoost berfungsi sebagai mesin prediksi karena kemampuannya mengelola data berdimensi tinggi dan interaksi nonlinier. Pentingnya, EX-Boost mengintegrasikan penyesuaian hiperparameter sistematis dan strategi regularisasi yang tangguh untuk memastikan *overfitting* teratasi secara efektif, menghasilkan model yang stabil dan *generalizable*. Hasil menunjukkan bahwa EX-Boost mencapai akurasi uji 92% dan rata-rata validasi silang 89%, mengungguli Regresi Logistik dan menunjukkan stabilitas yang lebih tinggi daripada Hutan Acak meskipun skor uji yang sedikit lebih tinggi. Temuan ini mengonfirmasi ketahanan dan skalabilitas EX-Boost untuk peramalan pasar tenaga kerja. Penelitian selanjutnya akan berfokus pada integrasi aliran data pasar tenaga kerja real-time dan peningkatan interpretabilitas untuk menghasilkan wawasan yang dapat ditindaklanjuti bagi pembuat kebijakan, pendidik, dan pemimpin industri.

**Kata kunci:** *Exploratory Data Analysis*, *Extreme Gradient Boosting* (XGBoost), Peramalan Pasar Tenaga Kerja, Analisis Prediktif, Perencanaan Tenaga Kerja

## ABSTRACT

Shifting global economic trends have significantly reshaped labor market dynamics, transforming employment structures and accelerating changes in skill requirements. This problem underscores the urgent need for accurate and reliable forecasting systems to anticipate future job opportunities and support evidence-based workforce planning. Traditional forecasting methods, which mainly rely on linear or statistical models, are often insufficient to capture nonlinear relationships and complex patterns in highdimensional, multi-source data. This study aims to overcome these limitations by introducing EX-Boost, an integrated predictive framework that combines Exploratory Data Analysis (EDA) with the Extreme Gradient Boosting (XGBoost) algorithm. The framework leverages multi-source labor data including historical statistics, online job postings, and macroeconomic indicators compiled from O\*NET Online and official labor reports. The proposed model applies EDA to detect anomalies, identify trends, and select high-impact features, while XGBoost serves as the predictive engine due to its ability to manage high-dimensional data and nonlinear interactions. Importantly, EX-Boost incorporates systematic hyperparameter tuning and robust regularization strategies to ensure that overfitting is effectively mitigated, resulting in a stable and generalizable model. The results demonstrate that EX-Boost achieved a test accuracy of 92% and a crossvalidation mean of 89%, outperforming Logistic Regression and exhibiting greater stability than Random Forest despite its slightly higher test score. These findings confirm the robustness and scalability of EX-Boost for labor market forecasting. Future work will focus on integrating real-time labor market data streams and enhancing interpretability to generate actionable insights for policymakers, educators, and industry leaders.

**Keyword:** Exploratory Data Analysis, Extreme Gradient Boosting (XGBoost), Job Market Forecasting, Predictive Analytics, Workforce Planning