PROPOSAL PENELITIAN AI

# BAB I - PENDAHULUAN

## 1.1 Latar Belakang

Perkembangan teknologi Artificial Intelligence (AI) telah memberikan dampak signifikan di berbagai bidang, termasuk ketenagakerjaan. Dalam konteks aplikasi ketenagakerjaan, modul Job Matching memegang peranan penting untuk menghubungkan pencari kerja dengan posisi yang sesuai berdasarkan data yang dimiliki. Penggabungan AI dengan modul ini memungkinkan peningkatan akurasi dan relevansi pencocokan pekerjaan, memanfaatkan data identitas pencari kerja, pendidikan, keahlian, serta kebiasaan melamar. Penelitian ini fokus pada pengembangan dan integrasi model AI untuk meningkatkan performa dan efektivitas modul Job Matching.

## 1.2 Rumusan Masalah

- Bagaimana mengembangkan model AI yang efektif untuk modul Job Matching dalam aplikasi ketenagakerjaan?

- Bagaimana mengoptimalkan penggunaan data identitas pencaker, pendidikan, keahlian, dan kebiasaan melamar untuk meningkatkan akurasi pencocokan pekerjaan?

- Metode machine learning apa yang paling sesuai untuk data kombinasi tersebut dalam konteks job matching?

- Bagaimana cara mengukur dan meningkatkan performa model AI dalam merekomendasikan pekerjaan yang relevan?

- Apa tantangan dalam pengumpulan, integrasi data, dan pengelolaan privasi data pencari kerja?

## 1.3 Tujuan Penelitian

- Mengembangkan model AI yang mampu merekomendasikan pekerjaan secara akurat pada modul Job Matching.

- Mengoptimalkan penggunaan data identitas, pendidikan, keahlian, dan kebiasaan melamar untuk meningkatkan kualitas pencocokan kerja.

- Mengevaluasi performa berbagai metode machine learning dalam konteks job matching.

## 1.4 Manfaat Penelitian

- Memberikan kontribusi pada pengembangan teknologi AI dalam ketenagakerjaan.

- Membantu pencari kerja mendapatkan rekomendasi pekerjaan yang lebih relevan dan sesuai dengan profil.

- Menjadi referensi bagi pengembang aplikasi ketenagakerjaan dalam mengimplementasikan AI.

# BAB II - TINJAUAN PUSTAKA

## 2.1 Penggunaan Machine Learning untuk Job Matching

Berbagai penelitian telah memanfaatkan algoritma machine learning seperti Random Forest, Gradient Boosting, dan Neural Networks untuk mengolah data perfil pencari kerja dan memberikan rekomendasi pekerjaan yang sesuai. Algoritma ini mampu menangani data numerik dan kategorikal yang kompleks untuk memprediksi kecocokan pekerjaan.

## 2.2 Feature Engineering pada Data Kependidikan, Keahlian, dan Perilaku Aplikasi

Teknik ekstraksi fitur yang efektif sangat penting dalam meningkatkan performa model AI. Pengolahan fitur seperti encoding data kategorikal (identitas, pendidikan, keahlian) dan analisis kebiasaan melamar (frekuensi, jenis posisi) dapat memberikan informasi penting untuk model.

## 2.3 Metode Rekomendasi Berbasis AI untuk Job Matching

Penggunaan embedding dan model rekomendasi adaptif membantu meningkatkan relevansi pencocokan dengan memahami relasi antara pencari kerja dan tipe pekerjaan secara lebih mendalam. Penelitian ini juga menyinggung tantangan dalam personalisasi dan skalabilitas sistem rekomendasi.

# BAB III - METODE PENELITIAN

## 3.1 Pengumpulan Data

- Sumber Data: Dataset berasal dari aplikasi ketenagakerjaan internal dan/atau dataset publik yang relevan.

- Data yang Dikumpulkan: Data identitas pencari kerja (umur, jenis kelamin, lokasi), pendidikan (tingkat, jurusan), keahlian (sertifikasi, skill), serta kebiasaan melamar pekerjaan (frekuensi, jenis posisi).

- Metode Pengumpulan: Pengambilan data langsung dari aplikasi, scraping data publik jika dibutuhkan, dan penggabungan beberapa dataset dengan memperhatikan izin dan privasi.

## 3.2 Deskripsi Dataset

|  |  |
| --- | --- |
| Fitur Dataset | Deskripsi |
| Identitas Pencaker | Informasi demografis: umur, jenis kelamin, lokasi |
| Pendidikan | Tingkat pendidikan terakhir, jurusan/bidang studi |
| Keahlian | Daftar keahlian, sertifikasi, atau keunggulan khusus |
| Kebiasaan Melamar | Frekuensi melamar, riwayat lamaran, jenis posisi lamaran |
| Label Target | Kelas pekerjaan untuk klasifikasi dan rekomendasi |

## 3.3 Metode/Teknik Analisis Data

- Algoritma yang Digunakan:

- Random Forest

- Gradient Boosting (termasuk XGBoost)

- Neural Network dengan embedding untuk fitur kategorikal dan numerik

- Parameter Konfigurasi:

- Random Forest: jumlah pohon, kedalaman pohon

- XGBoost: learning rate, max\_depth, jumlah estimator

- Neural Network: jumlah epoch, batch size, optimizer, arsitektur layer

- Desain Eksperimen:

- Pembagian data: 70% training, 15% validation, 15% testing

- Melakukan cross-validation

- Evaluasi dengan metrik: akurasi, precision, recall, F1-score, dan AUC (jika diperlukan)

## 3.4 Tools dan Platform

- Bahasa Pemrograman: Python

- Library: scikit-learn, XGBoost, TensorFlow/Keras, Pandas, NumPy

- Visualisasi: Matplotlib, Seaborn

- Platform Pengembangan: Jupyter Notebook, Google Colab