

BAB I PENDAHULUAN

I.1 Latar Belakang

CV BeData Technology Indonesia merupakan perusahaan yang bergerak di bidang teknologi data, didirikan pada Oktober tahun 2022. Meskipun masih tergolong baru, CV BeData Technology Indonesia telah menunjukkan perkembangannya, salah satunya melalui program magang bernama BeIntern. Program ini bertujuan untuk memberikan kesempatan kepada mahasiswa dalam mendapatkan pengalaman bekerja secara langsung dalam perusahaan, sekaligus menjadi sarana bagi perusahaan untuk menemukan calon karyawan yang memiliki potensi dan dapat berkontribusi pada kemajuan perusahaan.

Dalam dua *batch* program magang yang telah diselenggarakan, BeIntern telah menarik minat ratusan pendaftar, dengan jumlah peserta yang terus meningkat. Pada *batch* terakhir, terdapat sekitar 1.700 pendaftar, dengan 47% pendaftar datang melalui *Google Form* yang disebarluaskan melalui Instagram, sementara 53% lainnya berasal dari laman *job seeker* Glints. Namun, dengan semakin populernya program magang ini, muncul tantangan baru dalam proses penyaringan CV (*Curriculum vitae*). Proses seleksi secara manual, di mana CV (*Curriculum vitae*) harus ditinjau satu per satu, memakan waktu dan tenaga yang sangat besar.

Proses seleksi CV (*Curriculum vitae*) yang dilakukan secara manual mengharuskan setiap dokumen ditinjau satu per satu. Pendekatan ini tidak hanya memakan waktu dan tenaga, tetapi juga meningkatkan risiko kesalahan dalam menentukan kandidat yang sesuai untuk posisi magang yang dibuka. Sebagai perusahaan yang masih dalam tahap pengembangan, BeData Technology Indonesia juga belum memiliki sistem otomatis yang dapat membantu menyelaraskan keterampilan kandidat dengan kebutuhan posisi yang dibutuhkan.

Untuk mengatasi masalah ini, pengembangan model klasifikasi keterampilan berbasis *machine learning* dapat menjadi solusi. Model ini

dirancang untuk mengidentifikasi keterampilan yang terdapat dalam CV (*Curriculum vitae*) kandidat, lalu memberikan rekomendasi kandidat yang paling sesuai dengan posisi magang yang dibuka. Model klasifikasi ini tidak hanya mempercepat proses seleksi, tetapi juga meningkatkan akurasi dalam mencocokkan keterampilan kandidat dengan kebutuhan posisi.

Dalam perancangan model klasifikasi ini, digunakan metode TF-IDF (*Term Frequency-Inverse Document Frequency*) untuk memproses teks dalam CV dan mengubahnya menjadi fitur numerik. TF-IDF dipilih karena dapat mengidentifikasi seberapa penting sebuah kata (misalnya, keterampilan tertentu) dalam satu dokumen relatif terhadap seluruh dokumen yang ada dalam dataset. Ini sangat cocok untuk mengolah data teks yang beragam seperti CV (*Curriculum vitae*), dimana setiap kata atau frasa mewakili keterampilan dan pengalaman yang dimiliki kandidat.

Selain itu, metode *K-Nearest Neighbors* (KNN) digunakan sebagai algoritma klasifikasi utama. KNN dipilih karena kesederhanaannya dan kemampuannya untuk mengklasifikasikan data baru dengan membandingkan kemiripannya dengan data pelatihan yang ada. Pada konteks ini, KNN membantu mengelompokkan kandidat berdasarkan keterampilan yang terdeteksi, sehingga bisa direkomendasikan untuk posisi yang sesuai. Model ini juga menggunakan pendekatan *One-vs-Rest* (OvR) untuk menangani klasifikasi multi-kelas, dimana setiap kategori pekerjaan diperlakukan sebagai kelas yang berbeda.

Melalui perancangan model klasifikasi keterampilan ini, perusahaan dapat meningkatkan efisiensi dalam menyeleksi kandidat magang dan mengurangi beban kerja manual. Selain itu, pengembangan model berbasis teknologi *machine learning* akan memperkuat komitmen perusahaan dalam menerapkan teknologi inovatif untuk mendukung proses rekrutmen kandidat magang.

I.2 Rumusan Masalah

1. Bagaimana merancang model klasifikasi untuk mengidentifikasi *skill* dari CV kandidat magang di CV BeData Technology Indonesia

2. Bagaimana merancang model klasifikasi untuk mengidentifikasi *skill* dari CV kandidat magang di CV BeData Technology Indonesia

I.3 Batasan Masalah

Berdasarkan rumusan masalah di atas, maka batasan masalah dalam pengembangan model klasifikasi adalah sebagai berikut :

1. Sistem rekomendasi kandidat yang dirancang tidak mencakup pengembangan sistem secara keseluruhan (*frontend* dan *backend*), hanya berfokus pada perancangan model klasifikasi *skill*.
2. Dataset yang digunakan merupakan data yang tersedia dari sumber eksternal (Kaggle).
3. Algoritma *machine learning* yang digunakan untuk proses klasifikasi adalah *K-Nearest Neighbors* (KNN) dengan pendekatan *One-vs-Rest* (OvR) tanpa eksplorasi algoritma lain.

I.4 Tujuan

Tujuan dilakukannya praktik kerja lapangan adalah sebagai berikut :

1. Mengembangkan model klasifikasi untuk mengidentifikasi *skill* dari CV kandidat magang.
2. Mengembangkan sistem rekomendasi yang dapat memberikan saran kandidat magang yang sesuai dengan kebutuhan posisi di CV BeData Technology Indonesia.

I.5 Manfaat

Manfaat dilakukannya praktik kerja lapangan adalah sebagai berikut :

- a. Manfaat Bagi Mahasiswa
 1. Mahasiswa akan mendapatkan pengalaman praktis dalam menerapkan teori dan konsep *machine learning* yang telah dipelajari di perkuliahan ke dalam proyek nyata.
 2. Melalui proyek ini, mahasiswa akan meningkatkan keterampilan teknis dalam pengembangan model klasifikasi dan sistem rekomendasi, serta kemampuan analisis data.

3. Kesempatan untuk bekerja dalam lingkungan profesional memberikan peluang bagi mahasiswa untuk membangun jaringan profesional dengan para profesional di bidang teknologi dan data.
- b. Manfaat Bagi Fakultas
 1. Fakultas dapat menjalin kerjasama yang lebih erat dengan perusahaan dan mengembangkan proyek-proyek kolaboratif lainnya untuk kepentingan pengembangan teknologi dan inovasi.
 2. Pengalaman praktis yang didapat oleh mahasiswa dapat menjadi bahan masukan bagi fakultas dalam memperbarui kurikulum.
 - c. Manfaat Bagi Perusahaan
 1. Sistem rekomendasi yang dikembangkan akan meningkatkan efisiensi dan efektivitas proses seleksi kandidat magang, menghemat waktu dan tenaga yang sebelumnya digunakan untuk screening manual.
 2. Implementasi sistem berbasis *machine learning* menunjukkan komitmen perusahaan terhadap inovasi teknologi, yang dapat meningkatkan citra perusahaan di mata calon kandidat dan mitra bisnis.