BAB 1. PENDAHULUAN

1.1 Pendahuluan

Perkembangan teknologi informasi memberikan peluang besar dalam pengolahan data, termasuk dalam memetakan informasi mengenai alumni perguruan tinggi. Salah satu data penting yang dapat dianalisis adalah informasi pekerjaan alumni pada *Tracer Study*. Data *Tracer Study* ini memiliki nilai strategis karena dapat digunakan oleh program studi untuk mengevaluasi relevansi kurikulum, menyesuaikan kompetensi lulusan dengan kebutuhan industri, serta memberikan gambaran bagi calon mahasiswa mengenai prospek karier. Menurut Anggarawati dkk. (2022), pengelompokan data alumni berperan penting dalam penentuan strategi pengembangan mutu lulusan dan penyusunan kurikulum berbasis kebutuhan pasar kerja.

Pada kajian sebelumnya, analisis data pekerjaan alumni dilakukan menggunakan metode K-Means sebagai teknik pengelompokan. Penelitian oleh Diqy dkk. (2019) menunjukkan bahwa K-Means mampu memberikan hasil pengelompokan bidang pekerjaan alumni yang terstruktur. Namun, kajian tersebut masih memiliki keterbatasan, terutama dalam tahap *preprocessing* data yang belum optimal. Data pekerjaan alumni biasanya ditulis secara bebas pada kolom isian *tracer study*, sehingga profesi yang sama dapat ditulis dengan istilah berbeda, seperti "programmer", "software developer", atau "pengembang aplikasi". Akibatnya, hasil pengelompokan menjadi kurang akurat karena variasi penulisan tersebut tidak ditangani dengan baik.

Untuk mengatasi permasalahan tersebut, diperlukan metode *preprocessing* teks yang lebih mendalam agar data dapat diolah dengan lebih seragam dan akurat. Menurut Sari dkk. (2021), *preprocessing* teks mencakup tahapan *case folding, tokenizing, stopword removal, dan stemming*. Namun, tahapan tersebut saja belum cukup untuk merepresentasikan data teks dalam bentuk yang dapat diproses oleh algoritma *clustering*. Oleh karena itu, metode TF-IDF (*Term Frequency-Inverse Document Frequency*) menjadi solusi penting dalam mengubah data teks menjadi

representasi numerik yang mencerminkan tingkat kepentingan setiap kata dalam konteks keseluruhan data. Dengan TF-IDF, istilah-istilah yang sering muncul namun tidak informatif dapat dikurangi bobotnya, sementara istilah spesifik yang relevan akan memiliki bobot lebih tinggi, sehingga meningkatkan kualitas pengelompokan (Laksana dkk. 2025).

Kajian ini berfokus pada pengembangan sistem pekerjaan alumni dari tiga program studi di bidang teknologi informasi, yaitu Program Studi Manajemen Informatika, Teknik Informatika, dan Teknik Komputer di Politeknik Negeri Jember. Ketiga program studi ini dipilih karena memiliki karakteristik lulusan yang sejenis namun dengan kekhasan masing-masing, sehingga analisis komparatif terhadap pola pekerjaan alumni dapat memberikan wawasan yang lebih komprehensif. Sistem yang dikembangkan diharapkan mampu mengotomatisasi pengolahan data alumni dan menghasilkan pengelompokan pekerjaan yang lebih informatif melalui tahapan *preprocessing* teks, ekstraksi fitur menggunakan TF-IDF, dan pengelompokan data menggunakan algoritma K-Means.

Hasil yang diperoleh diharapkan tidak hanya bermanfaat bagi pihak akademik dalam mengevaluasi profil lulusan dan menyusun strategi pengembangan kurikulum, tetapi juga menjadi rujukan bagi mahasiswa dan stakeholder lain dalam memahami pola pekerjaan alumni pada bidang teknologi informasi. Dengan penerapan TF-IDF sebagai metode ekstraksi fitur, penelitian ini memberikan kontribusi dalam peningkatan akurasi pengelompokan data pekerjaan yang bersifat tekstual dan tidak terstruktur.

Ruang lingkup kajian ini terbatas pada data alumni yang berasal dari Program Studi Manajemen Informatika, Teknik Informatika, dan Teknik Komputer dalam rentang waktu tahun 2019 hingga 2025. Data pekerjaan alumni yang menjadi fokus utama diperoleh melalui tracer study maupun sumber administrasi akademik. Analisis yang dilakukan dibatasi pada aspek pengelompokan pekerjaan. Metodologi yang digunakan adalah algoritma K-Means dengan dukungan preprocessing teks dan ekstraksi fitur menggunakan TF-IDF untuk memastikan data pekerjaan dapat dikelompokkan secara lebih akurat dan representatif.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah:

a. Bagaimana cara membangun sistem klaster pekerjaan alumni yang dapat memproses data *tracer study*?
b. Bagaimana penerapan algoritma K-Means dalam pengelompokan pekerjaan alumni yang bersifat *unsupervised*?
c. Bagaimana hasil pengelompokan K-Means dalam hal fleksibilitas dan akurasi pengelompokan data pekerjaan alumni?

1.3 Tujuan

Adapun tujuan dari penelitian ini adalah:

a. Mengimplementasikan algoritma K-Means untuk mengelompokkan data pekerjaan alumni dari Program Studi Manajemen Informatika, Teknik Informatika, dan Teknik Komputer secara unsupervised.
b. Mengevaluasi kinerja algoritma K-Means dalam mengelompokkan data pekerjaan alumni yang bersifat tekstual dan tidak terstruktur menggunakan metrik evaluasi clustering.
c. Menyediakan informasi pengelompokan pekerjaan alumni yang dapat digunakan sebagai acuan evaluasi program studi dan pengembangan kurikulum.

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

a. Memberikan metode pengolahan data alumni yang efisien dan otomatis.
b. Menjadi acuan untuk penelitian lanjutan terkait pengelompokan atau clustering data teks alumni.
c. Menjadi bahan pertimbangan bagi program studi dalam mengevaluasi dan mengembangkan kurikulum berdasarkan pola pekerjaan alumni serta memberikan informasi mengenai prospek pekerjaan lulusan.