

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Di tengah perkembangan dunia bisnis yang terus berlanjut, perusahaan seringkali menghadapi situasi di mana mereka memiliki banyak karyawan dengan beragam latar belakang demografis. Laporan terbaru yang dikeluarkan oleh Badan Pusat Statistik (BPS) menunjukkan bahwa, berdasarkan data dari Survei Angkatan Kerja Nasional (Sakernas) pada bulan Februari 2023, jumlah angkatan kerja mencapai 146,62 juta orang, mengalami peningkatan sebanyak 2,61 juta orang dibandingkan dengan bulan Februari 2022. Sementara itu, terdapat 7,99 juta orang yang masih menganggur, mengalami penurunan sebesar 0,41 juta orang jika dibandingkan dengan Februari 2022.

Kenaikan jumlah pekerja yang disertai dengan pertumbuhan ekonomi yang meningkat di Indonesia menunjukkan bahwa pertumbuhan ekonomi dapat diperkuat dengan keterlibatan sumber daya manusia (Ramadhani, R., & Firmansyah, Y., 2023). Hal ini menggambarkan bahwa peran karyawan dalam proses pertumbuhan ekonomi menjadi semakin penting. Selain itu, laporan tersebut juga mencerminkan bahwa perusahaan-perusahaan harus lebih berfokus pada manajemen sumber daya manusia yang efektif untuk mengelola karyawan dengan beragam latar belakang demografis agar dapat mengoptimalkan kontribusi mereka dalam menghadapi tantangan dan peluang bisnis yang terus berkembang.

Performa tenaga kerja merupakan elemen utama dalam mempengaruhi tingkat efisiensi dan efektivitas lembaga perusahaan, baik dalam meningkatkannya maupun mengurangnya (Idehlu, HA et al. 2019). Oleh karena itu, penting bagi

lembaga perusahaan untuk memberikan gaji yang pantas dan adil. Salah satu hal yang bisa dilakukan adalah membagi gaji berdasarkan seberapa baik karyawan bekerja dan berapa banyak mereka berkontribusi. Ini bisa menjadi cara untuk mendorong karyawan agar selalu bekerja dengan baik dan akhirnya meningkatkan kinerja keseluruhan. Namun dari hasil analisis yang telah dilakukan oleh Ramadhani, R., & Firmansyah, Y. (2023), keberagaman demografis tenaga kerja, seperti usia, jenis kelamin, pendidikan, dan sebagainya juga memiliki dampak positif yang signifikan pada cara karyawan bekerja. Berdasarkan penelitian di PT XYZ, terdapat sekitar 29,5% dari kinerja karyawan dipengaruhi oleh faktor-faktor demografi ini.

Melihat pengaruh positif yang signifikan dari faktor-faktor demografi terhadap kinerja karyawan, dapat dipertimbangkan untuk menerapkan strategi penggajian yang lebih berfokus pada aspek demografis ini. Salah satu langkah yang bisa diambil adalah menggunakan *clustering* untuk pemodelan data aktual tingkat tinggi seperti gaji (Anggarawati, S. 2022). *Clustering* adalah alat yang paling berguna untuk penggalan data, kompresi, estimasi kepadatan probabilitas, dan banyak tugas penting lainnya (Vadyala, S. R., et. al., 2021), serta merupakan teknik yang esensial dalam mengorganisir data atau objek berdasarkan sejauh mana hubungan antar data diperlukan (Faisal, M., & Zamzami, E. M., 2020). Terdapat berbagai jenis algoritma pengelompokan seperti *K-means*, PAM, DBSCAN, dan lainnya yang tersedia. *K-means* adalah salah satu metode pengelompokan yang paling umum diterapkan serta dapat digunakan dengan mudah (Faisal, M., & Zamzami, E. M., 2020) dan *K-means* tidak diragukan lagi digunakan dalam

konteks ini (Celebi, M. E., Kingravi, H. A., & Vela, P. A. 2013).

Dalam penelitian Martín, I., Mariello, A., Battiti, R., & Hernández, JA (2018) menerapkan pengelompokan *K-means* pada atribut-atribut yang memiliki ketergantungan linier yang signifikan terhadap gaji. Hasil dari penelitian ini menghasilkan 9 kluster yang memisahkan penawaran berdasarkan rentang gaji, meskipun penekanannya adalah pada klasifikasi dengan metode *decision tree* yang mencapai akurasi terbaik sebesar 84%. Di sisi lain, penelitian oleh Syaputri, D., Noprita, P. H., & Romelah, S. (2021) juga memberikan wawasan berharga dengan berfokus pada pengelompokan data distribusi sosial ekonomi masyarakat menggunakan metode *K-means*, menghasilkan 4 kluster yang paling optimal dengan nilai indeks Davies-Bouldin (DBI) sebesar 0,87.

Berdasarkan kerangka ini, peneliti bertujuan untuk melanjutkan penelitian terkait dengan judul “Algoritma *Clustering K-Means* Dalam Analisis *Cluster* Gaji Berdasarkan Data Demografi Dengan *Mechine Learning* Menggunakan Python”. Tujuan utama penelitian ini adalah untuk mengelompokkan Gaji berdasarkan faktor-faktor demografis menggunakan metode *K-means clustering*. Dengan demikian, penelitian ini diharapkan dapat memberikan pandangan yang lebih mendalam tentang cara yang lebih efektif dan efisien dalam mengelola gaji karyawan dengan mempertimbangkan faktor-faktor demografis yang juga berperan penting.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang masalah yang telah dipaparkan maka muncul rumusan masalah penelitian yaitu; Bagaimana menghasilkn k optimal pada *K-means clustering* dalam klasterisasi gaji berdasarkan data demografil?

1.3 Ruang Lingkup

1. Sumber data dalam penelitian ini menggunakan dataset yang diperoleh dari *website Kaggle.com* yaitu *Jog description salary in Indonesia* dengan jumlah record 3000.
2. Data demografi yang akan digunakan mencakup atribut seperti jenis kelamin, pendidikan, gaji, lamanya bekerja, pekerjaan, dan lokasi.
3. Metode *K-means clustering* untuk mengelompokkan gaji terhadap faktor-faktor demografi yang telah disebutkan
4. Penelitian ini membatasi jumlah kelompok (K) yang akan dibentuk dalam *K-means clustering* menggunakan metode *elbow* untuk mendapatkan k optimal.
5. Penelitian ini akan membatasi analisis hanya pada faktor-faktor demografi yang tersedia dalam dataset
6. Menggunakan metrik jarak *euclidean distance* untuk mengukur kesamaan antara data.
7. Penelitian ini akan memfokuskan pada data demografi sebagai pengaruh utama dalam analisis perbandingan gaji.
8. Tidak akan ada penjelasan mendalam tentang faktor-faktor eksternal atau sejarah ekonomi yang mungkin memengaruhi perbandingan gaji.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk mencapai *cluster* optimal dalam mengelompokkan gaji berdasarkan atribut demografis menggunakan metode *K-means clustering*.

1.5 Manfaat Penelitian

1. Dengan memperhitungkan karakteristik demografis dalam pengelompokan gaji, perusahaan dapat lebih memastikan bahwa karyawan dengan profil yang serupa diperlakukan secara adil dalam hal gaji mereka.
2. Memberikan pengetahuan serta perspektif tentang data mining dan merinci langkah-langkah yang terlibat dalam analisis dengan metode *K-means clustering*.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Bab ini membahas mengenai latar belakang, perumusan masalah, ruang lingkup, tujuan penelitian, manfaat penelitian, dan pengaturan isi penelitian.

BAB II TINJAUAN PUSTAKA DAN DASAR TEORI

Pada bab ini akan dibahas mengenai sumber pustaka yang relevan yang digunakan sebagai referensi dalam penelitian serta landasan teori yang mendukung penelitian dalam penyusunan skripsi.

BAB III METODE PENELITIAN

Bab ini membahas tentang prosedur penelitian yang melibatkan bahan, peralatan, serta tahapan kerja dan pengumpulan data.

BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bagian ini membahas implementasi penelitian, seperti hasil analisis yang sesuai dengan pembahasan sebelumnya.

BAB V PENUTUP

Bab ini berisi kesimpulan dan rekomendasi terkait hasil penelitian yang telah dilakukan