

BAB V

PENUTUP

5.1. Kesimpulan

Penelitian ini bertujuan menerapkan metode HDBSCAN (*Hierarchical Density-Based Spatial Clustering of Applications with Noise*) Untuk melakukan pengelompokan data ketenagakerjaan provinsi di Indonesia yang memiliki pencila (*outlier*). Metode HDBSCAN memiliki keunggulan dalam mengidentifikasi *outlier* secara otomatis dengan mengeksplorasi berbagai skala kepadatan untuk menemukan kluster yang stabil, sementara titik yang dianggap *outlier* akan dipisahkan menjadi kluster tersendiri.

Hal tersebut berhasil dibuktikan dalam penelitian ini. Secara umum, HDBSCAN mampu mengelompokkan data ketenagakerjaan provinsi di Indonesia yang mengandung *outlier* karena data memiliki variasi kepadatan yang mencerminkan perbedaan kondisi masyarakat dan faktor alam, yang memengaruhi jumlah penduduk usia produktif, tingkat pendidikan, serta komposisi sektor lapangan pekerjaan. Berdasarkan proses analisis dan evaluasi, penelitian ini menghasilkan kesimpulan sebagai berikut:

1. Metode HDBSCAN berhasil diterapkan pada data ketenagakerjaan provinsi di Indonesia yang mengandung *outlier*, dengan pengujian parameter *min_cluster_size* = 2–6 dan metrik = Euclidean serta Manhattan. Parameter terbaik diperoleh pada *min_cluster_size* = 2 dan metrik *Euclidean* dengan nilai *Silhouette Coefficient* sebesar 0,546 serta hasil visualisasi kluster yang optimal, di mana titik-titik berhasil terbagi menjadi 3 kluster utama dan 1 kluster *noise* tanpa mengganggu kluster utama. Hasil kluster tersebut meliputi kluster 0 berjumlah 3 provinsi, kluster 1 berjumlah 25 provinsi, kluster 2 berjumlah 3 provinsi, dan kluster -1 (*noise*) berjumlah 3 provinsi. Terlihat adanya ketimpangan jumlah anggota, di mana kluster 1 memiliki 25 provinsi sementara kluster lainnya hanya 3 provinsi. Hal ini disebabkan HDBSCAN mengelompokkan titik berdasarkan kepadatan, sehingga provinsi pada kluster 1 memiliki kisaran nilai yang sempit dari hasil standarisasi dan dikategorikan

sebagai satu kelompok besar yang menunjukkan karakteristik mayoritas provinsi serupa.

2. *Outlier* dalam data ketenagakerjaan juga dapat digunakan sebagai salah satu indikator evaluasi hasil klasterisasi HDBSCAN. Sebelum proses klasterisasi, dilakukan deteksi *outlier* menggunakan metode *Local Outlier Factor* (LOF). Hasil LOF menunjukkan 6 provinsi sebagai *outlier*, yaitu Jawa Timur, Jawa Tengah, Jawa Barat, Banten, DKI Jakarta, dan Sumatera Utara. Berdasarkan hasil klasterisasi dengan parameter terbaik, HDBSCAN mengelompokkan 3 provinsi *outlier* menjadi klaster *noise*, sementara 3 provinsi *outlier* lainnya berada pada klaster utama tetapi tidak bergabung dengan provinsi yang bukan *outlier*. Hal ini menunjukkan bahwa meskipun termasuk *outlier*, ketiga provinsi tersebut masih memiliki kepadatan atau jarak kemiripan yang cukup dekat.
3. Hasil klasterisasi data ketenagakerjaan provinsi yang diperoleh, digunakan untuk memahami kesiapan daya saing kelompok provinsi di pasar tenaga kerja dengan menganalisis dominasi rata-rata setiap klaster terkait pendidikan, sektor pekerjaan, dan proporsi penduduk usia produktif. Hasil analisis ini diharapkan dapat memberikan gambaran yang jelas mengenai provinsi yang memiliki kesiapan daya saing tenaga kerja lebih tinggi dan provinsi yang masih memerlukan perhatian dalam pengembangan sumber daya manusia.
4. Penelitian ini juga berhasil mengembangkan GUI (*Graphical User Interface*) untuk memvisualisasikan cara kerja HDBSCAN serta menerapkannya langsung pada data ketenagakerjaan yang mengandung *outlier*. GUI ini menyediakan fitur input data sehingga pengguna dapat memasukkan data baru, memahami proses analisis seperti statistika deskriptif dan deteksi *outlier*, melakukan proses klasterisasi, dan memvisualisasikannya melalui peta interaktif provinsi Indonesia. Fitur ini meningkatkan interaktivitas dan kemudahan dalam eksplorasi data, sehingga membantu pemangku kepentingan memahami kelompok provinsi yang sudah siap maupun yang masih memerlukan pengembangan dalam meningkatkan daya saing SDM di pasar tenaga kerja untuk mendukung penyusunan kebijakan yang lebih tepat sasaran.

5.2. Saran Pengembangan

Melihat hasil serta keterbatasan penelitian ini, terdapat beberapa rekomendasi yang dapat dijadikan acuan untuk pengembangan maupun penelitian lanjutan, yaitu:

1. Menambahkan variabel lain terkait data ketenagakerjaan yang relevan dan membantu memperluas serta memperkuat gambaran daya saing seperti tingkat partisipasi angkatan kerja (TPAK), tingkat pengangguran terbuka (TPT), dan pendapatan per kapita.
2. Mengembangkan penelitian dengan menggunakan data panel (time series per provinsi) sehingga dinamika perkembangan tenaga kerja dapat dipantau dari waktu ke waktu.
3. Metode HDBSCAN dapat digunakan dengan kondisi yang sama seperti data ketenagakerjaan yang memiliki distribusi tidak merata berdasarkan kepadatan.
4. Menambahkan metrik evaluasi kualitas kluster, misalnya Davies-Bouldin Index dan Dunn Index agar penilaian performa kluster tidak hanya bergantung pada Silhouette Coefficient dan dapat mencakup perspektif yang lebih luas.
5. Mengembangkan GUI yang tidak hanya menampilkan hasil klusterisasi secara visual, tetapi juga menyediakan fitur interaktif untuk analisis data baru, simulasi skenario kebijakan, dan rekomendasi strategis berbasis kluster sehingga dapat menjadi alat bantu bagi pemangku kebijakan untuk memahami karakteristik kelompok provinsi, mengidentifikasi provinsi yang membutuhkan perhatian khusus, serta merancang kebijakan peningkatan daya saing tenaga kerja secara lebih tepat sasaran dan aplikatif.

Dengan penerapan saran-saran tersebut, penelitian selanjutnya diharapkan dapat menghasilkan analisis kluster ketenagakerjaan yang lebih akurat, representatif, dan aplikatif sehingga dapat mendukung pengambilan keputusan dan perumusan kebijakan yang lebih tepat sasaran dalam meningkatkan daya saing tenaga kerja di Indonesia.