

BAB I

PENDAHULUAN

1.1 Latar Belakang masalah

Perkembangan teknologi informasi dan analisis data di era transformasi digital telah mengubah cara lembaga pemerintah, termasuk Dinas Tenaga Kerja Kabupaten Labuhanbatu, mengelola informasi ketenagakerjaan untuk mendukung kebijakan berbasis data yang akurat dan efisien. Penerapan *machine learning* menjadi salah satu pendekatan yang banyak digunakan dalam analisis data sosial, ekonomi, dan ketenagakerjaan. Kabupaten Labuhanbatu sebagai salah satu daerah yang memiliki dinamika ekonomi cukup tinggi di Provinsi Sumatera Utara, menghadapi tantangan serius dalam mengelola isu pengangguran. Pengangguran tidak hanya menjadi persoalan ekonomi, tetapi juga berdampak pada aspek sosial, kesejahteraan masyarakat, dan stabilitas pembangunan daerah.

Pengelolaan data ketenagakerjaan di Dinas Tenaga Kerja Kabupaten Labuhanbatu hingga saat ini masih dilakukan secara manual dan parsial, sehingga tidak mampu mendukung analisis prediktif berbasis data yang dibutuhkan untuk perumusan kebijakan tenaga kerja yang proaktif. Akibatnya, proses analisis dan pemetaan status tenaga kerja belum mampu menggambarkan kondisi yang sesungguhnya. Oleh karena itu, penelitian ini diarahkan untuk membandingkan performa dua metode *machine learning* yaitu Random Forest dan Support Vector Machine (SVM), guna menentukan model klasifikasi terbaik yang dapat digunakan untuk memetakan status pengangguran dan mendukung kebijakan ketenagakerjaan di Kabupaten Labuhanbatu.

Masalah pengangguran di Kabupaten Labuhanbatu muncul dari berbagai faktor yang saling berkaitan, seperti ketidaksesuaian antara keterampilan tenaga kerja dengan kebutuhan pasar kerja, tingkat pendidikan, pengalaman kerja, dan kondisi sosial ekonomi masyarakat. Data yang dihasilkan dari survei dan administrasi ketenagakerjaan belum diolah secara optimal sehingga kebijakan yang diambil sering bersifat reaktif, bukan

prediktif. Metode statistik konvensional tidak lagi cukup untuk memetakan pola non - linear dan multidimensi dalam data ketenagakerjaan, sehingga diperlukan penerapan metode machine learning yang mampu mengenali interaksi kompleks antar variabel seperti usia, pendidikan, pengalaman kerja, dan kondisi sosial ekonomi. Dinas Tenaga Kerja Kabupaten Labuhanbatu membutuhkan sistem yang mampu melakukan klasifikasi status tenaga kerja baik bekerja, setengah menganggur, maupun menganggur dengan tingkat akurasi tinggi dan dapat digunakan sebagai dasar penyusunan kebijakan yang lebih tepat sasaran. Oleh karena itu, diperlukan suatu pendekatan analitis yang dapat mengelola kompleksitas data dan menghasilkan model prediksi yang akurat.

Salah satu pendekatan yang dapat digunakan untuk menjawab permasalahan tersebut adalah penerapan algoritma *machine learning* berbasis klasifikasi. Penelitian ini memilih dua algoritma utama, yaitu *Random Forest* dan *Support Vector Machine* (SVM), karena keduanya terbukti efektif dalam klasifikasi data sosial - ekonomi dan memiliki kemampuan yang saling melengkapi: *Random Forest* unggul dalam interpretasi variabel penting (*feature importance*), sedangkan SVM unggul dalam kemampuan generalisasi dan pemisahan kelas data yang optimal. Keduanya memiliki keunggulan dalam memproses data berukuran besar dengan variabel yang beragam.

Random Forest merupakan metode yang menggabungkan banyak pohon keputusan untuk mengurangi kesalahan prediksi dan meningkatkan stabilitas hasil. Metode ini juga memiliki kemampuan untuk menunjukkan tingkat kepentingan variabel (*feature importance*), yang dapat membantu Dinas Tenaga Kerja dalam memahami faktor - faktor utama penyebab pengangguran. Di sisi lain, *SVM* bekerja dengan mencari *hyperplane* terbaik yang dapat memisahkan kelas data secara optimal. Meskipun *SVM* dikenal memiliki kemampuan generalisasi yang baik, terutama pada data yang terstruktur dan bersih, metode ini cenderung membutuhkan penyesuaian parameter yang kompleks agar hasilnya optimal.

Melalui penelitian ini, kedua metode tersebut dibandingkan secara sistematis untuk menentukan model terbaik dalam mengklasifikasikan status pengangguran di Labuhanbatu.

Penelitian ini memberikan solusi berbasis teknologi yang konkret terhadap masalah pengangguran di tingkat daerah. Penerapan model machine learning diharapkan mampu mentransformasi sistem pengambilan keputusan di Dinas Tenaga Kerja Kabupaten Labuhanbatu, dari yang bersifat manual dan deskriptif menjadi berbasis analisis prediktif, sehingga kebijakan penanggulangan pengangguran dapat disusun secara lebih tepat sasaran dan efisien.

Hasil klasifikasi yang dihasilkan oleh model dapat digunakan untuk mengidentifikasi kelompok masyarakat yang berisiko tinggi mengalami pengangguran. Informasi ini memungkinkan dinas terkait untuk merancang program pelatihan dan peningkatan keterampilan yang lebih spesifik serta sesuai dengan kebutuhan pasar kerja. Selain itu, penerapan *Random Forest* dan *SVM* memberikan alternatif yang terukur dalam pemanfaatan data ketenagakerjaan daerah. Hasil penelitian ini diharapkan dapat diintegrasikan dalam sistem informasi ketenagakerjaan berbasis *dashboard* interaktif, yang dapat menampilkan data pengangguran secara real - time berdasarkan variabel seperti usia, tingkat pendidikan, atau wilayah administratif. Dengan demikian, penelitian ini bukan hanya menghasilkan model klasifikasi yang efektif, tetapi juga mendukung peningkatan efisiensi dan transparansi dalam perumusan kebijakan publik.

Keterbaruan penelitian ini terletak pada penerapan algoritma machine learning untuk menganalisis data ketenagakerjaan daerah, yang selama ini belum banyak dikaji dalam konteks pemerintahan kabupaten di Indonesia. Penelitian ini berkontribusi pada literatur data - driven policy making dengan memberikan bukti empiris bagaimana model klasifikasi dapat membantu pemerintah daerah dalam merumuskan kebijakan ketenagakerjaan yang adaptif. Sebagian besar penelitian sebelumnya berfokus pada penerapan algoritma *machine learning* di bidang keuangan, kesehatan, dan pendidikan, sementara isu pengangguran di sektor pemerintahan daerah belum banyak

dikaji menggunakan pendekatan ini. Melalui perbandingan antara *Random Forest* dan *SVM*, penelitian ini menghadirkan nilai tambah berupa pemahaman mendalam terhadap karakteristik data ketenagakerjaan di Labuhanbatu serta rekomendasi praktis bagi instansi pemerintah dalam memilih algoritma yang sesuai. Penelitian ini tidak hanya menilai performa teknis kedua metode, tetapi juga menyoroti aspek implementatifnya dalam kebijakan publik berbasis data. Dengan pendekatan yang komprehensif, penelitian ini diharapkan menjadi langkah awal menuju pengembangan sistem prediksi ketenagakerjaan yang berkelanjutan di Kabupaten Labuhanbatu. Integrasi antara teknologi dan kebijakan publik yang dihasilkan dari penelitian ini diharapkan mampu memberikan kontribusi nyata dalam mengurangi tingkat pengangguran dan meningkatkan kualitas tenaga kerja di daerah tersebut.

1.2 Rumusan masalah

Hasil penelitian ini memberikan pemahaman mengenai karakteristik data ketenagakerjaan di Kabupaten Labuhanbatu, serta membandingkan kinerja algoritma *Random Forest* dan *SVM* dalam mengklasifikasikan status pengangguran. Berdasarkan metrik evaluasi, algoritma dengan kinerja terbaik dapat dipilih untuk mendukung kebijakan ketenagakerjaan daerah yang lebih efektif.

1. Bagaimana karakteristik data ketenagakerjaan di Kabupaten Labuhanbatu yang digunakan dalam penerapan algoritma machine learning untuk klasifikasi status pengangguran?
2. Bagaimana penerapan algoritma *Random Forest* dan *Support Vector Machine* (*SVM*) dalam mengklasifikasikan status pengangguran berdasarkan variabel - variabel ketenagakerjaan daerah?
3. Bagaimana hasil perbandingan kinerja kedua algoritma tersebut berdasarkan metrik evaluasi (*accuracy*, *precision*, *recall*, dan *F1-score*) untuk menentukan metode yang paling optimal bagi kebijakan ketenagakerjaan daerah?

1.3 Batasan Masalah

Hasil penelitian ini menjelaskan bahwa penelitian memiliki batasan yang jelas pada aspek data, metode, evaluasi, dan implementasi, yang memungkinkan fokus yang lebih tajam pada penerapan algoritma machine learning untuk klasifikasi status pengangguran di Kabupaten Labuhanbatu.

1. Penelitian ini hanya menggunakan data ketenagakerjaan yang diperoleh dari Dinas Tenaga Kerja Kabupaten Labuhanbatu, yang mencakup variabel utama seperti usia, tingkat pendidikan, pengalaman kerja, jenis kelamin, dan status pekerjaan.
2. Metode yang digunakan dalam penelitian ini dibatasi pada dua algoritma *machine learning* berbasis klasifikasi, yaitu *Random Forest* dan *Support Vector Machine* (SVM), tanpa membandingkan dengan algoritma lain di luar kedua metode tersebut.
3. Pengujian performa model difokuskan pada metrik evaluasi utama, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*, serta tidak mencakup pengujian lanjutan seperti *cross-validation* pada data dari wilayah lain di luar Kabupaten Labuhanbatu.
4. Implementasi hasil penelitian difokuskan pada penyusunan rekomendasi model prediktif yang dapat diadaptasikan oleh Dinas Tenaga Kerja Kabupaten Labuhanbatu, tanpa melakukan integrasi langsung ke dalam sistem informasi ketenagakerjaan secara real-time.

1.4 Tujuan dan Manfaat Penelitian

Berdasarkan hasil penelitian ini terbukti dalam memberikan kontribusi ilmiah di bidang ketenagakerjaan, khususnya dengan penerapan algoritma machine learning untuk klasifikasi status pengangguran. Penelitian juga memberikan acuan metodologis bagi para peneliti dalam memilih algoritma yang sesuai dengan karakteristik data sosial ekonomi daerah. Dengan menyediakan informasi berbasis data yang akurat dalam merancang kebijakan ketenagakerjaan yang lebih baik.

1.4.1 Tujuan Penelitian

1. Menganalisis karakteristik data ketenagakerjaan di Kabupaten Labuhanbatu yang menjadi dasar dalam proses klasifikasi status pengangguran dengan pendekatan *machine learning*.
2. Menerapkan algoritma *Random Forest* dan *Support Vector Machine (SVM)* untuk melakukan klasifikasi status pengangguran berdasarkan variabel - variabel ketenagakerjaan yang tersedia.
3. Membandingkan kinerja kedua algoritma tersebut menggunakan metrik evaluasi utama seperti *accuracy*, *precision*, *recall*, dan *F1 - score* untuk menentukan model yang paling optimal dalam klasifikasi status pengangguran.
4. Menghasilkan model klasifikasi terbaik yang dapat dijadikan rekomendasi bagi Dinas Tenaga Kerja Kabupaten Labuhanbatu dalam menyusun kebijakan berbasis data guna mengurangi tingkat pengangguran di wilayah tersebut.

1.4.2 Manfaat Penelitian

1. Memberikan kontribusi ilmiah dalam pengembangan kajian penerapan algoritma *machine learning* pada bidang ketenagakerjaan, khususnya dalam perbandingan metode *Random Forest* dan *SVM* untuk klasifikasi status pengangguran.
2. Menjadi acuan metodologis bagi peneliti atau akademisi dalam memilih dan menerapkan algoritma klasifikasi yang sesuai dengan karakteristik data sosial ekonomi di tingkat daerah.
3. Menyediakan informasi berbasis data yang akurat bagi Dinas Tenaga Kerja Kabupaten Labuhanbatu untuk memetakan kondisi ketenagakerjaan dan merancang program pelatihan serta penempatan kerja yang lebih efektif.
4. Mendukung pengembangan sistem informasi ketenagakerjaan berbasis prediksi yang dapat diintegrasikan ke dalam kebijakan publik, sehingga membantu pemerintah daerah dalam pengambilan keputusan yang lebih cerdas dan efisien.

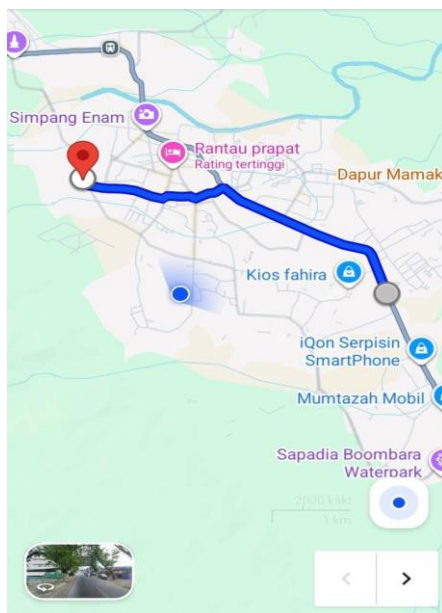
1.5 Tinjauan Umum objek Penelitian

Dinas Tenaga Kerja Kabupaten Labuhanbatu merupakan instansi pemerintah daerah yang memiliki tugas pokok dan fungsi dalam penyelenggaraan urusan pemerintahan di bidang ketenagakerjaan. Lembaga ini berperan penting dalam mengatur, mengawasi, serta meningkatkan kualitas tenaga kerja melalui berbagai program seperti pelatihan keterampilan, penempatan tenaga kerja, dan penanggulangan pengangguran. Kabupaten Labuhanbatu sendiri terletak di Provinsi Sumatera Utara dan dikenal sebagai wilayah dengan aktivitas ekonomi yang cukup beragam, meliputi sektor pertanian, perdagangan, dan industri jasa. Kondisi ini menyebabkan dinamika ketenagakerjaan di wilayah tersebut sangat kompleks, karena tingkat pendidikan, keterampilan, serta ketersediaan lapangan kerja tidak selalu seimbang. Oleh karena itu, pengelolaan data ketenagakerjaan menjadi aspek yang krusial untuk memahami kondisi tenaga kerja secara lebih komprehensif dan menyusun kebijakan yang efektif.

Selama ini, sistem pendataan ketenagakerjaan di Dinas Tenaga Kerja Kabupaten Labuhanbatu masih dilakukan secara manual atau semi - digital, sehingga data yang dihasilkan belum sepenuhnya optimal untuk analisis mendalam. Keterbatasan ini berdampak pada sulitnya melakukan klasifikasi dan prediksi status tenaga kerja secara akurat, baik untuk kategori bekerja, setengah menganggur, maupun menganggur. Melalui penelitian ini, penggunaan teknologi *machine learning* diharapkan dapat membantu dinas dalam meningkatkan akurasi analisis dan mempermudah identifikasi kelompok masyarakat yang berisiko tinggi mengalami pengangguran. Dengan demikian, penerapan metode *Random Forest* dan *Support Vector Machine (SVM)* bukan hanya menjadi inovasi dalam analisis data, tetapi juga langkah strategis untuk memperkuat kapasitas Dinas Tenaga Kerja dalam mewujudkan kebijakan berbasis data yang adaptif terhadap perubahan sosial dan ekonomi di Kabupaten Labuhanbatu.



Dinas Tenaga Kerja Kabupaten Labuhanbatu



Denah Lokasi dari Universitas Labuhanbatu ke Dinas Tenaga Kerja

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini disusun secara terstruktur agar pembahasan dapat tersaji secara logis, runtut, dan mudah dipahami. Setiap bab memiliki keterkaitan yang saling mendukung dalam menjelaskan proses penelitian mulai dari perumusan masalah hingga penarikan kesimpulan. Adapun sistematika penulisan laporan penelitian ini adalah sebagai berikut:

BAB I Pendahuluan Bab ini menjelaskan gambaran umum penelitian yang mencakup latar belakang masalah, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian, tinjauan umum objek penelitian, serta sistematika penulisan. Bagian ini berfungsi sebagai landasan awal untuk memahami arah dan fokus penelitian yang dilakukan.

BAB II Landasan Teori Bab ini berisi uraian teori - teori yang relevan dengan penelitian, meliputi konsep dasar *machine learning*, algoritma *Random Forest* dan *Support Vector Machine (SVM)*, serta kajian literatur terkait penelitian terdahulu. Bab ini bertujuan untuk memberikan dasar teoretis yang memperkuat analisis dan interpretasi hasil penelitian.

BAB III Metode Penelitian Bab ini menjelaskan metode yang digunakan dalam penelitian, termasuk sumber data, variabel penelitian, teknik pengumpulan data, tahapan *preprocessing*, implementasi algoritma, serta metode evaluasi model. Bagian ini juga memaparkan pendekatan perbandingan yang digunakan untuk mengukur performa *Random Forest* dan *SVM* dalam klasifikasi status pengangguran.

BAB IV Implementasi dan Pembahasan Bab ini menyajikan hasil implementasi algoritma *Random Forest* dan *SVM*, analisis hasil pengujian model, serta interpretasi perbandingan kinerja keduanya berdasarkan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Pembahasan juga mencakup analisis faktor-faktor yang memengaruhi performa model dan relevansinya terhadap kebijakan ketenagakerjaan di Kabupaten Labuhanbatu.

BAB V Kesimpulan dan Saran Bab terakhir berisi kesimpulan yang diperoleh dari hasil penelitian serta saran yang diajukan untuk pengembangan penelitian selanjutnya dan penerapan praktis oleh Dinas Tenaga Kerja Kabupaten Labuhanbatu. Kesimpulan dirumuskan berdasarkan hasil analisis yang telah dilakukan, sedangkan saran disusun untuk memberikan arah perbaikan dan pengembangan sistem analisis ketenagakerjaan di masa mendatang.