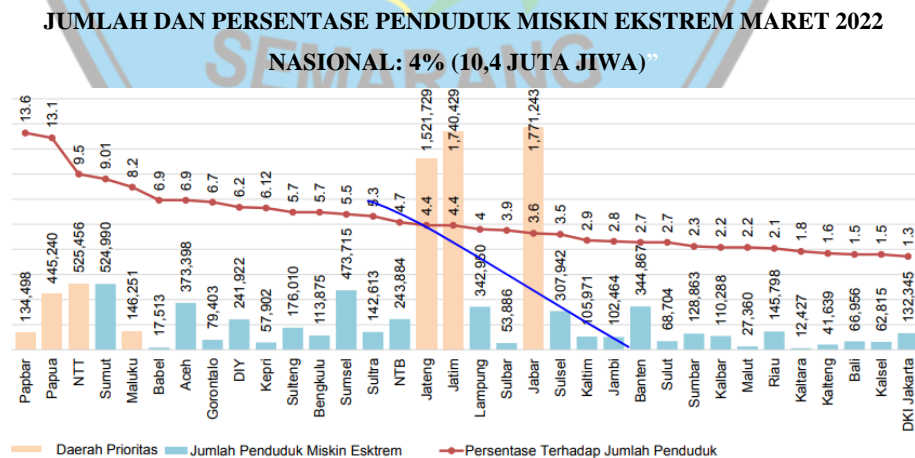


BAB I PENDAHULUAN

1.1 Latar Belakang

Kemiskinan merupakan isu sentral bagi setiap negara didunia, khususnya bagi negara berkembang, pengentasan kemiskinan dan menciptakan kesejahteraan bagi rakyat merupakan tujuan akhir suatu negara termasuk di Indonesia. Bappenas mendefinisikan kemiskinan sebagai keadaan dimana individu atau sekelompok laki-laki dan perempuan tidak mampu mencukupi hak-hak dasarnya untuk mengembangkan dan mempertahankan kehidupan. Hak-hak dasar tersebut meliputi: Pemenuhan kebutuhan pangan, pendidikan, kesehatan, air bersih, pertanian, pekerjaan, perumahan, sumber daya alam dan lingkungan hidup (Bappenas, 2021).

Berdasarkan data publikasi badan pusat statistik (BPS) pada September 2022 bahwa jumlah penduduk miskin di Indonesia sebesar 9,57% atau sebanyak 26,36 juta jiwa. Sementara persentase penduduk miskin ekstrem pada Maret 2022 sebesar 10,4 juta jiwa (TNP2K, 2022). Data jumlah dan persentase penduduk miskin ekstrem bulan Maret 2022 Secara Nasional ditunjuk pada gambar 1.1.



Gambar 1.1 Jumlah dan Persentase Penduduk Miskin Ekstrem

Pemerintah menetapkan sasaran prioritas pengentasan kemiskinan ekstrem sebanyak 212 kabupaten/kota dengan tingkat kemiskinan ekstrem diperkirakan berada di angka 3-3,5 persen, dimana Provinsi Jawa Tengah masuk didalamnya. Terdapat 19 daerah di Jawa Tengah yang masuk kategori miskin ekstrem termasuk Kabupaten Pemalang yaitu dengan presentase kemiskinan ekstrem tertinggi di Jawa Tengah, dengan nilai 9,52% (TNP2K, 2022).

Setiap pemimpin baik pemimpin daerah maupun pusat menjadikan permasalahan penduduk miskin sebagai tujuan utama yang harus dituntaskan. Menurut Suharto dalam (Ortega, 2021) program jaring pengaman sosial melalui bantuan langsung tunai merupakan program pemerintah berupa pengaman sosial yang diberikan kepada kelompok-kelompok yang rentan menyusul adanya dampak-dampak negatif jangka pendek akibat diterapkannya suatu kebijakan (Ortega, 2021).

Bantuan langsung tunai dana desa atau disingkat dengan (BLT-DD) merupakan bantuan keuangan yang bersumber dari dana desa dan diperuntukkan bagi masyarakat miskin yang berdomisili di desa dan rentan yang menghadapi kesulitan dalam memenuhi kebutuhan hidupnya sehari-hari masyarakat yang belum menerima bantuan sosial lain seperti program keluarga harapan (PKH), bantuan sosial tunai (BST), UMKM, kartu prakerja, dan bantuan sosial lainnya (Sasuwuk dkk., 2021)

Permasalahan selama ini, mekanisme penyaluran bantuan sosial dana desa belum mengadopsi metode yang tepat dalam menentukan objek penerima bansos, sehingga menimbulkan subjektivitas yang tinggi dalam menentukan Objek penerima bansos. Banyaknya pengaduan masyarakat tentang bantuan sosial yang tidak tepat sasaran tentu menjadi perhatian khusus untuk masalah ini. Kita membutuhkan sistem pendukung keputusan yang menggunakan metode untuk mendukung keputusan sistem. Karena jumlah data yang diolah relatif besar dan belum adanya pola aturan yang baku, sulit untuk memberikan penilaian

yang tepat dan adil dalam menentukan penerima bansos. Akibatnya pengambilan keputusan seringkali tidak jelas sehingga menimbulkan kekacauan dan kecemburuan sosial di masyarakat.

Perkembangan teknologi informasi pada saat ini berdampak pada kehidupan sehari-hari yang membutuhkan informasi yang lebih akurat. Dalam mengambil keputusan dari data biasanya hanya mengandalkan data operasional saja, padahal ini tidak cukup. Maka diperlukan analisis data untuk menggali pola atau informasi penting dari yang jumlah data yang besar, yang disebut *data mining*. Menurut Larose dalam (Mardi, 2014) "Data mining adalah bidang dari beberapa keilmuan yang menggabungkan teknik dari pembelajaran mesin, pengenalan pola, statistik, database dan visualisasi yang mengatasi permasalahan pengambilan informasi dari database yang besar.

Terdapat banyak teknik pengerjaan *data mining* untuk menemukan pola atau informasi tersembunyi, diantaranya Klasterisasi (*clustering*), Regresi (*regression*), Asosiasi (*association*) dan Klasifikasi (*classification*). Metode klasifikasi yang baik akan menghasilkan sedikit kesalahan klasifikasi atau akan menghasilkan peluang kesalahan alokasi yang kecil" (Johnson dan Wichern, 2007). Dalam mengevaluasi sebuah model Klasifikasi, melihat performa dari sebuah algoritma adalah hal yang penting. Performa klasifikasi dapat dihitung dengan melihat *confusion matrix* dengan melihat tingkat akurasi, *presisi*, *recall*, *specificity*, *f1-score* dan nilai *AUC* dari curva *ROC*.

Penelitian sebelumnya mengenai komparasi/perbandingan metode dalam *data mining* telah banyak dilakukan Dengan jumlah data dan atribut yang berbeda-beda. salah satunya yaitu penelitian yang dilakukan oleh Dinda Rosyada Nuruzzulfa (Nuruzzulfa, 2020). yang mengambil tema Klasifikasi Korban Kecelakaan Lalu Lintas Di Kabupaten Pati. Penelitian dilakukan dengan jumlah data sebanyak 207 data. Hasilnya didapat bahwa nilai akurasi tertinggi diperoleh dengan

menggunakan algoritma *Naïve Bayes* dengan nilai akurasi sebesar 96,13%. Akas Bagus setiawan (2018) sebelumnya juga telah melakukan penelitian mengenai perbandingan klasifikasi *Decision Tree C4.5* dan *Naïve Bayes* untuk menunjang keputusan pinjaman uang (studi kasus di koperasi karyawan PT. Karyamitra budisentosa pandangan pasuruan).

Decision Tree adalah model prediksi terhadap suatu keputusan menggunakan struktur hirarki atau pohon. Setiap pohon memiliki cabang, cabang mewakili suatu atribut yang harus dipenuhi untuk menuju cabang selanjutnya hingga berakhir di daun (Jayanti dkk., 2008). Hasilnya metode *naïve bayes* diperoleh nilai *AUC* rata-rata sebesar 86% berdasarkan 4 model sampling. Sedangkan metode *C4.5* dengan mendapatkan performa dari nilai rata-rata *AUC* sebesar 0,786 berdasarkan 4 model sampling masing-masing metode diuji menggunakan skenario uji *K-fold* dengan *cross validation* sehingga *naïve bayes* merupakan metode terbaik yang direkomendasikan dalam bentuk sebuah *decision suport system* (setiawan, 2018).

Pedoman yang jelas diperlukan untuk mengambil keputusan yang tepat sasaran, sehingga memudahkan tim seleksi untuk mengevaluasi penerima bansos Dana Desa. Dengan memanfaatkan Algoritma *C4.5* dan *naïve bayes*, penerapan *data mining* dapat membantu dalam menemukan data dalam database. Algoritma *C4.5* merupakan pengembangan Algoritma ID3 untuk meningkatkan beberapa karakteristik, yaitu: dapat digunakan untuk menangani atribut kontinu, dapat berurusan dengan data *training* yang nilai atributnya hilang serta dapat digunakan untuk memproses pemangkasan pohon yang dibangun (Mantas dan Abellan, 2014).

Naïve bayes merupakan perhitungan teorema bayes yang paling sederhana, karena dapat mengurangi kompleksitas komputasi menjadi multiplikasi sederhana dari probabilitas. Selain itu, algoritma *naïve bayes* juga mampu menangani set data yang memiliki banyak atribut (Yoo dkk., 2012). Pemilihan penggunaan Algoritma *C4.5* dan *naïve*

bayes pada penelitian ini didasarkan pada beberapa alasan, yaitu: Selain kedua Algoritma tersebut sama-sama mudah diimplementasikan dan sama-sama dapat memberikan hasil yang baik dalam kasus klasifikasi, kedua Algoritma tersebut juga mempunyai beberapa keunggulan masing-masing. *C4.5* merupakan Algoritma klasifikasi pohon keputusan yang efisien dalam menangani atribut bertipe diskret dan numerik (Han et al., 2012). diharapkan dengan menggabungkan algoritma *C4.5* dengan *Naïve bayes*, akan dapat dihasilkan data baru berupa pola aturan yang dapat digunakan sebagai acuan dalam menyeleksi pemohon Bantuan Sosial khususnya Bantuan Langsung Tunai Dana Desa. untuk membantu tim seleksi membuat keputusan.

1.2 Rumusan Masalah

“ Berdasarkan latar belakang di atas, maka rumusan masalah yang dibahas dalam penelitian sebagai berikut :

1. Bagaimana tingkat Performa Algoritma *C4.5* dan *naïve bayes* pada Klasifikasi bantuan sosial dana desa di desa?
2. Bagaimana hasil perbandingan tingkat performa antara Algoritma *C4.5* dan Algoritma *naïve bayes* pada Klasifikasi bantuan sosial dana desa?.”

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan penelitian yang hendak dicapai sebagai berikut :

1. Untuk mendapatkan tingkat performa Algoritma *C4.5* dan *naïve bayes* pada Klasifikasi bantuan sosial dana desa.
2. Untuk mendapatkan hasil perbandingan tingkat performa antara Algoritma *C4.5* dan Algoritma *naïve bayes* pada klasifikasi bantuan sosial dana desa.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari hasil penelitian ini adalah sebagai berikut:

2

1. Manfaat Teoritis

Penelitian ini dapat digunakan sebagai referensi atau literatur ilmiah dalam mengembangkan metode klasifikasi *machine learning* terutama algoritma *C4.5* dan *Naïve bayes*

2. Manfaat Praktis

a. Bagi Peneliti, Mengetahui cara menginterpretasikan Algoritma *C4.5* dan *naïve bayes* pada klasifikasi bansos di desa Asemtoyong, kecamatan Taman, Kabupaten Pamelang.

b. Bagi Masyarakat, sebagai informasi dan tambahan pengetahuan di bidang statistik khususnya penerapan Algoritma *Naïve bayesian* dan *C4.5* dalam klasifikasi bansos dana desa diharapkan pembaca dapat mempelajari lebih lanjut.

c. Bagi lembaga pemerintah, sebagai bahan informasi dan referensi tambahan di bidang statistik serta bisa dijadikan sebagai pedoman bagi pemerintah untuk menyalurkan bantuan sosial agar tepat sasaran serta memprioritaskan masyarakat yang paling membutuhkan.

1.5 Batasan Penelitian

Untuk membatasi pembahasan yang terlalu luas maka peneliti membatasi penelitian ini dengan hal-hal sebagai berikut:

1. Data yang digunakan adalah data yang diperoleh dari dinas sosial Kabupaten Pamelang hasil verifikasi dan validasi data sistem informasi kesejahteraan sosial Jawa Tengah (SIKS-DJ) Pada Bulan November tahun 2022 dengan sampel data sebanyak 337.

2. kasus terfokus pada klasifikasi penerima bantuan sosial tunai dana desa di desa Asemtoyong Kecamatan Taman Kabupaten Pemalang Provinsi Jawa Tengah.
3. Metode yang digunakan adalah Algoritma *C4.5* dan Algoritma *Naïve bayes*. Dengan perhitungan perbandingan data *training* dan data *testing*, sebesar 70%, 75%, dan 80% untuk data *training*, dan data *testing* yaitu sebesar 30%, 25%, dan 20%.
4. Pemilihan model terbaik yaitu dengan melihat performa nilai *accurasi*, *precision*, *recall*, *f1-score* dan Nilai *AUC* tertinggi.

