

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berdasarkan hasil observasi yang dilakukan di Dinas Sosial Kota Tasikmalaya didapatkan informasi bahwa tingkat kemiskinan di Kota Tasikmalaya masih tinggi. Oleh karena itu, Dinas Sosial Kota Tasikmalaya melaksanakan program bantuan sosial pangan non tunai atau BPNT. BPNT merupakan bantuan sosial pangan yang disalurkan dalam bentuk non tunai dari pemerintah kepada penerima setiap bulannya. Proses dalam menentukan penerima bantuan sosial oleh pihak Dinas Sosial Kota Tasikmalaya masih dilakukan secara manual meskipun banyak data yang diproses. Penerapan sistem yang masih manual menggunakan perhitungan dari *microsoft excel* ini dinilai kurang akurat dalam memperoleh hasil penerima bantuan sosial karena membutuhkan waktu yang relatif lama dalam mengambil keputusan.

Dari permasalahan tersebut, untuk mengatasi masalah ini diperlukan perhitungan yang sistematis dalam menentukan siapa yang berhak menjadi penerima bantuan sosial. Dalam memproses data dibutuhkan model yang dapat menjelaskan data dengan pengaplikasiannya, maka dibuatlah sebuah model *machine learning* yang dapat membantu memproses data tersebut. Dalam proses ini, ada banyak algoritma *machine learning* yang dapat diterapkan. Pemberian klasifikasi penerimaan bantuan sosial pangan non tunai pada penelitian ini menggunakan algoritma *K-Nearest Neighbor* dan *Adaptive Boosting*. Algoritma *K-*

Nearest Neighbor dipilih karena memiliki kemampuan untuk mendeteksi dan menganalisa permasalahan yang sifatnya kompleks dan *non-linear* (Prasetio dan Susanti, 2019). Sedangkan, algoritma *Adaptive Boosting* merupakan salah satu varian dari beberapa algoritma *boosting* yang dapat mengubah model klasifikasi lemah menjadi model klasifikasi kuat (Freund dan Schapire, 1999).

Model *Machine Learning* yang telah dibuat akan dievaluasi menggunakan perhitungan *Confusion Matrix* dan nilai *Area Under The Curve* (AUC) yang dihasilkan dari kurva *Receiver Operating Characteristic* (ROC). Hal ini digunakan untuk melihat kinerja hasil dari masing-masing algoritma dan sejauh mana perbedaan akurasi yang dihasilkan algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* dalam memprediksi penerima bantuan sosial.

Dalam sebuah penelitian yang dilakukan (Firasari dkk., 2020) tentang perbandingan algoritma KNN dan Naïve Bayes pada klasifikasi masyarakat miskin penerima bantuan sosial menghasilkan Algoritma Naive Bayes dengan akurasi 89,04% sedangkan algoritma K-NN menghasilkan akurasi 87,67%. Perhitungan dua algoritma tersebut termasuk dalam kategori klasifikasi yang baik. Penelitian yang sama dilakukan Sumiah dan Mirantika pada tahun 2020 mengenai perbandingan metode *K-Nearest Neighbor* dan *Naive Bayes* dengan studi kasus rekomendasi penentuan mahasiswa penerima beasiswa, dalam penelitian ini algoritma *K-Nearest Neighbor* mempunyai akurasi yang lebih tinggi yaitu 100% dibandingkan algoritma *Naïve Bayes* yang menghasilkan akurasi sebesar 99,89%.

Menurut Yunita, Pada tahun 2017 dalam penelitian mengenai perbandingan dua algoritma yaitu algoritma *K-Nearest Neighbor* dan *Decision Tree* untuk

menentukan risiko kredit kepemilikan mobil. Penelitian ini dilakukan menggunakan aplikasi Rapid Miner dan hasilnya algoritma *K-Nearest Neighbor* memiliki nilai akurasi yang lebih baik. Penelitian mengenai perbandingan algoritma klasifikasi juga dilakukan oleh Prianti dkk., pada tahun 2020 yaitu perbandingan metode *K-Nearest Neighbor* dan *Adaptive Boosting* dengan menggunakan *base learner*-nya yaitu *Classification and Regression Trees (CART)* pada kasus klasifikasi multi kelas. Hasil dari perbandingan tersebut mendapatkan model terbaik dengan akurasi tertinggi yaitu pada metode *Adaboost* dengan banyak nilai *accuracy*, *precision* serta *recall* yang lebih besar dibandingkan dengan nilai pada model algoritma *K-Nearest Neighbor*.

Berdasarkan latar belakang yang diuraikan, penelitian ini akan membandingkan kinerja dari dua algoritma *machine learning* yang menghasilkan akurasi tertinggi dari penelitian-penelitian sebelumnya yaitu algoritma *K-Nearest Neighbor* dan *Adaptive Boosting*. Pada penelitian sebelumnya perbandingan algoritma antara *K-Nearest Neighbor* dan *Adaptive Boosting* pemodelan dilakukan dengan lebih sedikit data dan sedikit atribut yang digunakan dimana hal tersebut dapat berpengaruh pada akurasi yang dihasilkan setiap algoritmanya. Oleh karena itu, pada penelitian ini akan melakukan perbandingan dari kedua algoritma tersebut dengan lebih banyak data dan atribut yang digunakan untuk memperoleh hasil algoritma yang paling maksimal dalam menentukan penerima bantuan sosial pangan non tunai.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat dirumuskan masalah dalam

penelitian ini adalah:

1. Bagaimana mengimplementasikan algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* untuk memprediksi penerima bantuan sosial pangan non tunai?
2. Bagaimana performa yang dihasilkan dari algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* untuk memprediksi penerima bantuan sosial pangan non tunai dengan menggunakan lebih banyak data dan atribut yang digunakan?
3. Sejauh mana perbedaan akurasi yang dihasilkan algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* dalam memprediksi penerima bantuan sosial pangan non tunai?

1.3 Batasan Masalah

Terdapat beberapa batasan-batasan masalah yang digunakan sehingga penelitian dapat dilaksanakan secara spesifik. Batasan masalah pada penelitian ini, sebagai berikut:

1. Data yang digunakan diambil dari data penerima bantuan sosial pangan non tunai pada tahun 2020 dan bersumber dari Data Terpadu Kesejahteraan Sosial (DTKS) Kota Tasikmalaya.
2. Membahas perbandingan dua algoritma yaitu algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* untuk memprediksi penerima bantuan sosial pangan non tunai.
3. Pembuatan model *machine learning* menggunakan sintaks bahasa pemrograman *python* dengan beberapa *library* pada *python*.

4. Menggunakan 47 atribut untuk membuat model dalam memprediksi penerima bantuan sosial pangan non tunai.
5. Menggunakan *ROC Curve* dan *Confussion Matrix* dalam evaluasi model prediksi penerima bantuan sosial pangan non tunai.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Mengimplementasikan algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* untuk memprediksi penerima bantuan sosial pangan non tunai.
2. Menunjukkan performa kinerja yang dihasilkan dari algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* untuk memprediksi penerima bantuan sosial pangan non tunai dengan menggunakan lebih banyak data dan atribut yang digunakan.
3. Mendapatkan hasil perbandingan dari algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* dalam memprediksi penerima bantuan sosial pangan non tunai berupa tingkat akurasi serta menentukan algoritma mana yang menghasilkan akurasi terbaik.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat bermanfaat bagi seluruh pihak yang terkait, diantaranya :

1. Secara Aplikatif
 - a. Diharapkan dari hasil prediksi penerima bantuan sosial pangan non tunai ini dapat membantu pengambilan keputusan pihak Dinas Sosial Kota

Tasikmalaya untuk menentukan warga yang menjadi ‘Penerima’ dan ‘Bukan Penerima’ bantuan agar bantuan tepat sasaran.

2. Secara Akademis
 - a. Mengetahui cara pembuatan model *Machine Learning* dengan algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* untuk memprediksi penerima bantuan sosial pangan non tunai.
 - b. Memberikan wawasan bagaimana performa penggunaan algoritma *K-Nearest Neighbor* dan *Adaptive Boosting* dalam memprediksi penerima bantuan sosial pangan non tunai untuk dapat menghasilkan prediksi yang baik dalam menentukan warga yang menjadi penerima dan bukan penerima bantuan.

1.6 Metodologi Penelitian

Metodologi penelitian berisi mengenai waktu dan tempat penelitian, tahapan penelitian, pendekatan penelitian, jenis penelitian, variabel penelitian, serta objek penelitian. Metode penelitian yang digunakan untuk menyelesaikan penelitian ini diantaranya:

1. Studi Literatur, merupakan pengumpulan data-data serta sumber yang berhubungan dengan penelitian ini. Studi literatur didapatkan dari jurnal dan *e-proceeding* terkait metode dan algoritma yang digunakan, selain itu diperoleh dari buku dan internet serta dokumen yang berkaitan dengan objek penelitian.
2. Pengumpulan dan Analisis Data

Data yang dibutuhkan untuk penelitian ini adalah data penerima bantuan sosial pangan non tunai tahun 2020 yang didapatkan dari Data Terpadu Kesejahteraan Sosial Kota Tasikmalaya yang dapat dijadikan sampel dalam membandingkan kinerja dua Algoritma klasifikasi terbaik yaitu algoritma *K-Nearest Neighbor* dan *Adaptive Boosting (Adaboost)* untuk melihat algoritma mana yang menghasilkan akurasi yang lebih tinggi.

3. Pengolahan Data

Agar data yang didapatkan valid dan berkualitas, maka perlu dilakukan beberapa tahap pengolahan data, yaitu: *Data Validation, Data Integration and Transformation* dan *Data Size Reduction and Discretization*.

4. Pelatihan Model Algoritma *K-Nearest Neighbor* dan *Adaptive Boosting*

Pada tahapan ini data akan digunakan untuk pelatihan model dalam membandingkan kinerja 2 Algoritma klasifikasi terbaik yaitu algoritma *K-Nearest Neighbor* dan *Adaptive Boosting (Adaboost)* untuk melihat algoritma mana yang menghasilkan akurasi yang lebih tinggi. Proses ini dilakukan beberapa skenario dalam pembagian data.

5. Evaluasi Model

Dari skenario-skenario pembagian data yang ada pada setiap pemodelan dua algoritma yaitu *K-nearest Neighbor* dan *Adaptive Boosting* akan menghasilkan masing-masing akurasi dari setiap algoritma dan skenario-skenario pembagian data. Setelah itu akan dibandingkan dan

dipilih suatu model dan skenario terbaik berdasarkan hasil perhitungan *Confusion Matrix* dan nilai *AUC* pada *ROC Curve*.

6. Penarikan Kesimpulan

Tahapan ini merupakan tahapan akhir dari penelitian dimana hasil penelitian tersebut berupa Analisa perbandingan dua algoritma klasifikasi dalam prediksi penerima bantuan sosial dengan mengimplementasikan model *Machine Learning* menggunakan algoritma *K-Nearest Neighbor* dan *Adaptive Boosting (Adaboost)*.

1.7 Sistematika Penulisan

Aturan dan sistematika penulisan yang digunakan dalam penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab pendahuluan ini berisi latar belakang umum yaitu gambaran secara garis besar tentang isi laporan, rumusan masalah, batasan permasalahan pada penelitian, tujuan penelitian, manfaat penelitian yang diperoleh, metode penelitian serta sistematika penulisan laporan tugas akhir.

BAB II LANDASAN TEORI

Bab ini berisi pembahasan teori - teori yang saling berhubungan dengan penelitian seperti konsep serta metode dan algoritma yang terkait dengan penelitian ini. Pada bab ini juga berisi penjelasan dari penelitian sebelumnya yang relevan dengan penelitian ini.

BAB III METODOLOGI

Dalam bab ini berisi uraian metode yang digunakan dalam melakukan penelitian, mulai dari waktu dan tempat penelitian, objek penelitian, variabel penelitian, matriks penelitian serta tahapan-tahapan yang dilakukan pada penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi pemaparan hasil serta pembahasan terhadap perancangan pada bab sebelumnya, yaitu bagaimana alur sebuah data apabila diolah dengan berdasarkan metode *Adaboost* dan algoritma *K-Nearest Neighbor* dan bagaimana suatu sistem dikembangkan berdasarkan algoritma yang digunakan dengan mengimplementasikan model *Machine Learning* menggunakan *Framework Flask* dengan pemrograman *Python*, serta dilakukan pula uji coba sistem untuk mendapatkan hasil yang sesuai dengan tujuan awal penelitian.

BAB V KESIMPULAN DAN SARAN

Bab ini merupakan bab terakhir berisi kesimpulan dan saran dari hasil penelitian, serta merupakan garis besar dari metode penelitian yang telah dilakukan. Kesimpulan adalah hasil akhir dari penelitian yang dilakukan, sedangkan Saran berisi tentang rekomendasi sesuai dengan keterbatasan yang ada pada sistem.