

BAB I

PENDAHULUAN

1.1. Latar Belakang

Berdasarkan UUD 1945 pasal 31 Undang-Undang Dasar 1945, bahwa setiap warga Negara berhak memperoleh dan mengikuti pendidikan dasar dan pemerintah wajib untuk membiayai (Indonesia, 2022). Pentingnya sebuah pendidikan menjadi tanggung jawab pemerintah dalam membangun pendidikan untuk warga Negara agar menjadi manusia yang memiliki karakter serta berkehidupan sosial yang sehat. Spesifikasi singkatnya bisa dikatakan pendidikan itu adalah kunci membentuk sumber daya manusia yang berkualitas (Muhamad Yunus, 2021). Pendidikan selalu berubah dan berkembang secara progresif menjadikan pemerintah harus lebih meningkatkan kualitas dan kuantitas pendidikan yang ada di Indonesia. Dalam hal pengajaran, fasilitas, dan tenaga pengajarpun lebih di tingkatkan, yang berupaya untuk kemajuan Negara. Namun pada kenyataannya masih ada persoalan-persoalan di dunia pendidikan yang ada di Negara ini.

Intruksi Presiden Nomor 7 Tahun 2014 diantaranya mengamanatkan tentang Program Indonesia Pintar (PIP) kepada Kementerian Pendidikan dan Kebudayaan Untuk menyiapkan Kartu Indonesia Pintar (KIP) dan menyalurkan dana program Indonesia Pintar (PIP) kepada siswa yang orang tuanya tidak mampu membiayai pendidikannya (Indonesia, 2022). Program Indonesia Pintar (PIP) melalui Kartu Indonesia Pintar (KIP) merupakan kelanjutan dari program Bantuan Siswa Miskin (BSM) yang mencakup siswa dari jenjang pendidikan SD/MI, SMP/MTs, SMA/SMK/MA, dan siswa/warga yang menempuh pendidikan di Pusat Kegiatan Belajar Mengajar (PKBM) (Muhamad Yunus, 2021). Dalam hal ini pemberian

bantuan Kartu Indonesia Pintar harus di berikan kepada siswa/ warga yang layak untuk mendapatkan bantuan tersebut. Dari pemaparan diatas dapat diketahui permasalahan yang sering dihadapi dalam menentukan kelayakan penerima KIP agar sesuai dengan program pemerintah yang diadakan adalah menentukan siapa yang benar layak mendapatkan KIP, agar program yang di jalankan oleh pemerintah sesuai dengan sasaran dan kriteria dalam pemberiannya, banyak sekali anak-anak yang seharusnya masih menempuh pendidikan tidak bisa mendapatkan pendidikan yang layak yaitu wajib belajar 9 tahun, padahal anak-anak tersebut yang nantinya akan menjadi penerus bangsa ini. Anak-anak yang seharusnya tumbuh menjadi anak-anak yang berkualitas agar ini bisa berkembang kearah yang lebih baik. Karena, raut wajah suatu negeri dapat dilihat dari setinggi apa pendidikan yang dienyam oleh warganya.

Salah satu penyebab rendahnya tingkat pendidikan di Indonesia adalah kemiskinan, ada banyak penelitian yang menjelaskan tentang masalah kemiskinan salah satunya adalah penerapan aplikasi berbasis dataset e-commerce untuk prediksi kemiskinan menggunakan algoritma *naïve bayes*, *xgboost* dan *similarity based feature selection*, dalam penelitian ini diketahui bahwa kemiskinan menginterpretasikan salah suatu keadaan seseorang tidak mampu untuk memenuhi kebutuhan dasar mereka seperti halnya sandang, papan, pangan, kesehatan, dalam menuntut ilmu. Badan Pusat Statistik atau lebih dikenal dengan sebutan BPS menggunakan konsep kemampuan untuk dapat memenuhi kebutuhan (*basic needs approach*) guna mengukur tingkat kemiskinan di Indonesia (Yualinda, 2020).

Dengan menggunakan konsep ini, pengeluaran menjadi tolak ukur dari

kemiskinan yang dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan pangan dan non pangan, sehingga penduduk yang tidak mampu (miskin) adalah penduduk yang memiliki pengeluaran perkapita perbulan dibawah garis kemiskinan. Dalam percobaan yang telah dilakukan, nilainya cukup relevan antara fitur dan nilai asli. Banyaknya fitur yang terlalu sedikit tidak selalu menghasilkan nilai akurasi yang juga kecil, demikian juga sebaliknya, dimana penggunaan sejumlah besar fitur tidak selalu menghasilkan akurasi yang tinggi. Namun, secara keseluruhan metode yang digunakan dapat diterapkan pada data masyarakat untuk memprediksi masyarakat yang tergolong miskin dan yang tidak miskin (Yualinda, 2020).

Dari pemaparan penelitian sebelumnya, belum adanya penelitian yang membahas tentang penerimaan Kartu Indonesia Pintar (KIP) selain pada penelitian pemilihan calon penerima KIP dengan memanfaatkan penggunaan data mining dengan penerapan *C4.5*. Dalam penelitian ini dinyatakan bahwa penerimaan bantuan KIP selalu berkonflik dengan masalah kemiskinan hal ini berarti sejalan dengan penelitian sebelumnya tentang prediksi masyarakat kurang mampu namun dalam penelitian ini berfokus pada prestasi siswa dengan memperhatikan tingkat ekonomi dari suatu keluarga, sehingga diperlukannya evaluasi setiap akhir semester untuk mengetahui hasil belajar yang telah dicapai. Bantuan KIP dapat berupa uang sekolah, bantuan belajar siswa dan penghargaan bagi siswa yang berprestasi atau bagi siswa yang kurang mampu (Muhamad Yunus, 2021).

Maka penelitian ini digunakan untuk memudahkan sekolah dalam menyeleksi agar tepat sasaran, guna mengoptimalkan sekolah untuk mendapatkan calon

penerima KIP. Sehingga bantuan KIP tersebut mampu meringankan tanggungan orang tua dalam permasalahan pembiayaan sekolah dan juga dapat meningkatkan keinginan belajar bagi siswa. Metode penelitian yang digunakan adalah metode klasifikasi *Decision Tree* data mining dengan pemanfaatan algoritma *C4.5* dengan menggunakan variabel berupa pendapatan orangtua, nilai siswa, dan tanggungan orangtua. Berdasarkan hasil uji coba dan pembahasan yang telah dilakukan untuk pemilihan penerima KIP dapat diselesaikan. Data siswa yang diperoleh dari SMPN 35 Jakarta dapat diprediksi dengan akurasi yang cukup tinggi, hasil ini diperoleh dengan akurasi mencapai 90,24% dengan errorate 9,75% sehingga dengan hasil tersebut algoritma *C4.5* sangat baik dalam menentukan pemilihan penerima KIP (Muhamad Yunus, 2021).

Dari kedua penelitian tersebut dapat diketahui bahwa data mining dengan penerapan teknik klasifikasi dapat menjadi solusi untuk permasalahan pemilihan calon penerima KIP namun, agar terhindar dari penelitian dengan metode yang sama untuk penerimaan bantuan KIP maka, akan digunakan metode XGboost yang telah teruji pada penelitian prediksi masyarakat miskin mampu untuk melakukan prediksi dengan memanfaatkan fitur *boosting*. Hal ini juga diperkuat dengan penelitian sebelumnya tentang prediksi *churn* penggunaan algoritma XGboost dikomparasi dengan metode Random Forest untuk mengidentifikasi customer churn sejak awal dalam membantu sektor CRM (*Customer Relationship Management*) agar dapat mempertahankan *customer*, sehingga mengurangi potensi kerugian yang dialami perusahaan. Hasil analisis klasifikasi bahwa algoritma Extreme Gradient Boosting (XGboost) lebih unggul dibandingkan algoritma

Random Forest dilihat dari tingkat akurasi dan nilai AUC. Algoritma Extreme Gradient Boosting (XGboost) mendapatkan nilai akurasi dan AUC sebesar 95.6% dan 0.876, sedangkan algoritma Random Forest mampu menghasilkan nilai akurasi dan AUC sebesar 93.5% dan 0.799.

Penelitian ini bertujuan untuk mengetahui siswa yang layak menerima bantuan KIP dengan menerapkan algoritma XGboost, dataset yang digunakan dibagi menjadi dua kategori, yakni atribut penjelas dan kelas target. Atribut penjelas merupakan kriteria penentu yang menjadi dasar dalam pengklasifikasian dalam menentukan kelas target yaitu dalam kasus penelitian ini yang mencakup usia ayah, usia ibu, pendidikan ayah, pendidikan ibu, penghasilan ayah, penghasilan ibu, jumlah tanggungan keluarga, nilai rata-rata siswa. Sedangkan kelas target merupakan representasi dari hasil prediksi yang terdiri dari dua status, yakni layak atau tidak layak.

Tahapan yang dilakukan dalam penelitian ini dimulai dengan melakukan pengumpulan data dengan melakukan observasi dan wawancara untuk mendapatkan data primer berupa data siswa pada SMPN 02 Kebun Tebu, kemudian melakukan tahapan *preprocessing* yakni tahapan pembersihan data dan transformasi data menggunakan nilai bobot dikarenakan data siswa masih dalam bentuk data kategorikal, lalu dimodelkan menggunakan pemodelan klasifikasi algoritma XGboost, digunakan bahasa pemrograman *python* dengan pemanfaatan *tool google colab* dengan menggunakan library dari *scikit-skelearn*.

Untuk menilai hasil model klasifikasi, dilakukan evaluasi yang mencakup empat aspek, yakni akurasi, *precision*, *recall*, dan *f-1 score*. Harapannya, penelitian

yang dihasilkan dapat dijadikan sebagai model klasifikasi siswa yang layak menerima bantuan sehingga dapat mengantisipasi adanya kesalahan dalam penerimaan bantuan yang tidak tepat sasaran. Dengan menerapkan model klasifikasi XGboost diharapkan akan menghasilkan tingkat akurasi yang lebih baik dengan berfokus pada pengklasifikasian siswa yang layak menerima bantuan KIP secara tepat sasaran.

1.2. Rumusan Masalah

Berdasarkan latar belakang masalah di atas, rumusan masalah dalam penelitian ini adalah:

- a. Bagaimana menerapkan algoritma *XGboost* dalam mengklasifikasikan siswa calon penerima KIP?
- b. Bagaimana hasil analisis dalam hal tingkat akurasi dari kinerja algoritma *XGboost*?

1.3. Batasan Masalah

Untuk memudahkan dalam melaksanakan penelitian dan pengumpulan data, maka batasan masalah yang akan dibahas hanya pada:

- a. Atribut yang akan digunakan berjumlah 8 atribut penjelas dan 1 atribut yang menjadi target yang meliputi usia ayah, usia ibu, pendidikan ayah, pendidikan ibu, penghasilan ayah, penghasilan ibu, jumlah tanggungan keluarga, nilai rata-rata siswa dan untuk kelas target adalah atribut hasil.
- b. Data yang digunakan merupakan data siswa calon penerima KIP tahun 2021 untuk menentukan pola tahun 2022.

- c. Model yang digunakan adalah klasifikasi dengan penggunaan algoritma *XGboost*.
- d. *Tool* yang akan digunakan dalam memprediksi data adalah *Python* dengan menggunakan *Google Colab*, *library scikit-skelearn*, dan validasi menggunakan *Confussion Matrix*.

1.4. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

- a. Membantu pihak SMPN 02 Kebun Tebu dalam pengambilan keputusan bagi siswa yang akan menerima KIP.
- b. Mengetahui tingkat akurasi dari penggunaan algoritma *XGboost* untuk klasifikasi data siswa penerimaan KIP.

1.5. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah:

- a. Dapat membantu pihak SMPN 02 Kebun Tebu dalam pengambilan keputusan dan mengevaluasi terkait penerimaan bantuan KIP untuk siswa yang layak berdasarkan hasil klasifikasi dari 20% dan 30% data uji.
- b. Sebagai bahan acuan bagi peneliti selanjutnya dalam mengembangkan penelitian yang lebih lanjut dengan menggunakan algoritma yang lain dengan tingkat keakuratan yang lebih tinggi.