
ANALISIS SISTEM PENDUKUNG KEPUTUSAN DENGAN MODEL KLASIFIKASI BERBASIS MACHINE LEARNING DALAM PENENTUAN PENERIMA PROGRAM INDONESIA PINTAR

Andri Nata¹, Suparmadi²

STMIK Royal, Kisaran

e-mail: andrinata0202@gmail.com

***Abstract:** The PIP program is a poor student assistance given to students from underprivileged families to be able to carry out learning activities at school. However, the problem is that it is still found that students who receive this educational assistance come from well-to-do families. On the other hand, there are still students who are less able to get help. It is clear that at the education unit level this program has not been fully targeted. So this study aims to apply Decision Support System Technology in determining the recipients of the Smart Indonesia Program (PIP) at Tamansiswa Sukapeace Junior High School. Researchers will compare machine learning-based classification models such as the k-Nearest Neighbor, Naïve Bayes, Support Vector Maching, Boosting and Decision Tree Random Forest methods. It aims to determine the method in the classification model that has the best accuracy. The method with the best accuracy is then installed in a web-based system.*

***Keywords:** Smart Indonesia Program; Classification Models; Machine Learning.*

Abstrak: Program PIP adalah bantuan siswa miskin yang diberikan kepada siswa dari keluarga kurang mampu untuk dapat melakukan kegiatan belajar di sekolah. Akan tetapi, permasalahannya adalah masih dijumpai siswa penerima bantuan pendidikan ini berasal dari keluarga yang mampu. Sebaliknya masih dijumpai siswa yang kurang mampu justru tidak mendapatkan bantuan. Hal ini jelas bahwa di level satuan pendidikan program ini belum sepenuhnya tepat sasaran. Maka penelitian ini bertujuan untuk menarapkan Teknologi Sistem Pendukung Keputusan dalam penentuan penerima bantuan Program Indonesia Pintar (PIP) di SMP Tamansiswa Sukadamai. Peneliti akan membandingkan model klasifikasi berbasis machine learning seperti metode k-Nearest Neighbor, Naïve Bayes, Support Vector Maching, Boosting dan Decision Tree Random Forest. Hal ini bertujuan untuk mengetahui metode dalam model klasifikasi yang memiliki akurasi terbaik. Metode dengan akurasi terbaik ini kemudian dipasang didalam sistem berbasis web.

Kata kunci: Program Indonesia Pintar; Model Klasifikasi; Mechine Learning.

PENDAHULUAN

Salah satu upaya nyata yang dilakukan dalam program wajib belajar Sembilan tahun yaitu melalui bantuan siswa miskin[1]. Program ini kemudian dilanjutkan dengan nama program Indonesia Pintar (PIP). Program PIP adalah bantuan siswa miskin yang diberikan kepada siswa dari keluarga yang kurang mampu untuk dapat

melakukan kegiatan belajar di sekolah. Sehingga program ini adalah yang memberikan kesempatan yang seluas luasnya kepada peserta didik yang masuk dalam kategori miskin untuk tetap bersekolah atau tidak putus sekolah. Bantuan ini memberikan peluang bagi siswa untuk mengikuti pendidikan di level yang lebih tinggi. Bersumber dari pip.kemendikbud.go.id pada 3 Desember 2021 Jumlah penerima program bantuan

siswa miskin yang sudah dicairkan pada tahun 2021 adalah sebesar 12.908.246 siswa dengan rincian SD: 7.625.596, SMP : 3.598.925, SMA : 710.037 dan SMK : 973.688. Program ini akan berjalan dengan lancar sesuai dengan tujuannya jika yang diberikan bantuan tepat sasaran yaitu siswa yang memenuhi kategori kurang mampu.

Akan tetapi, permasalahannya adalah masih dijumpai siswa penerima bantuan pendidikan ini berasal dari keluarga yang mampu. Sebaliknya masih dijumpai siswa yang kurang mampu justru tidak mendapatkan bantuan. Hal ini jelas bahwa di level satuan pendidikan program ini belum sepenuhnya tepat sasaran. Berdasarkan observasi yang peneliti lakukan salah satu penyebab dari masalah ini adalah pemilihan siswa yang berhak mendapatkan bantuan ini masih dilakukan secara manual dan cenderung subjektif. Misalnya dikarenakan faktor keluarga maupun faktor keterbatasan sumber daya dalam seleksi dilapangan yang menyebabkan nama-nama penerima bantuan ini kurang tepat sasaran. Misalnya di Sekolah Menengah Pertama Tamansiswa Di Kabupaten Asahan Desa Sukadamai. Kekurangan sumberdaya dalam menyeleksi siswa yang layak dan tidak layak dalam program bantuan ini mengalami kesulitan. Sehingga salah satu solusi yang dapat digunakan untuk memudahkan adalah melalui penerapan teknologi komputer.

Berdasarkan permasalahannya di atas, maka penelitian ini bertujuan untuk menerapkan Teknologi Sistem Pendukung Keputusan dalam penentuan penerima bantuan Program Indonesia Pintar (PIP) di SMP Tamansiswa Sukadamai. Model yang digunakan dalam sistem pendukung keputusan ini adalah model klasifikasi berbasis Machine Learning. Model klasifikasi dipilih karena memiliki kesesuaian dalam menentukan status layak dan dilayak berdasarkan kumpulan data.

Penelitian terkait penentuan status kelayakan menggunakan model klasifikasi pernah dilakukakan oleh Hanun (2020) dengan judul “Penerapan Algoritma

Klasifikasi Random Forest Untuk Penentuan Kelayakan Pemberian Kredit Di Koperasi Mitra Sejahtera” Penelitian ini bertujuan untuk memprediksi kelayakan kredit dengan penerapan Algoritma Klasifikasi Random Forest agar dapat memberikan solusi untuk penentuan kelayakan pemberian kredit.

Akan tetapi, model klasifikasi yang digunakan hanya Random Forest saja. Sehingga tidak didapati perbandingan akurasi yang lebih baik karena tidak membandingkan dengan metode-metode lain yang terdapat dalam model klasifikasi. Sedangkan dalam penelitian yang akan dilakukan ini, peneliti akan membandingkan model klasifikasi seperti metode k-Nearest Neighbor, Naive Bayes, Support Vector Maching, Boosting dan Decision Tree Random Forest. python. Hal ini bertujuan ini mengetahui metode dalam model klasifikasi yang memiliki akurasi terbaik. Metode dengan akurasi terbaik ini kemudian akan dipasang didalam sistem berbasis web, sehingga pengguna memiliki sistem yang dapat mudah digunakan dan memiliki akurasi yang paling baik.

Tujuan Penelitian

1. Menentukan model klasifikasi terbaik dalam menganalisa data kelayakan penerima Program Indonesia Pintar.
2. Merancang dan mendeployment Sistem Pendukung Keputusan dengan Model Klasifikasi Berbasis Mechine Learning dalam Penentuan Penerima Program Indonesia Pintar berbasis web.

Study Kelayakan

Didasarkan pada pendekatan Information Economic (IE) penerapan rencana sistem dalam penelitian ini layak untuk dilakukan. Hal ini dikarenakan secara quasi tangible benefit, sistem yang akan dibangun dapat digunakan untuk memperbaiki proses perencanaan, perbaikan pengambilan keputusan dalam satuan pendidikan. Penelitian ini sejalan dengan renstra penelitian institusi peneliti

yaitu Sekolah Tinggi Manajemen Informatika dan Komputer Royal yang salah satu bidang fokusnya adalah Machine Learning dan Kecerdasan Buatan.

Penelitian terkait Kelayakan penerima Program Indonesia pintar pada tahun 2022 dilakukan oleh Pratama dengan judul “Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Penerima Program Indonesia Pintar”. Penelitian ini bertujuan untuk menentukan penerima bantuan melalui penerapan data mining menggunakan algoritma Naïve Bayes yang dapat mengklasifikasikan tingkat kelayakan masyarakat dalam menerima bantuan sehingga diperoleh hasil yang lebih baik dalam menentukan penerima bantuan Program Indonesia Pintar. Akan tetapi, selain metode yang digunakan hanya metode naïve bayes dan akurasi dari klasifikasi dalam penelitian ini juga tidak ditemukan. Sehingga evaluasi terhadap sistem yang dibangun belum terukur dengan baik.

Machine Learning

Machine Learning (ML) atau pembelajaran mesin merupakan pendekatan dalam AI yang banyak digunakan untuk menggantikan atau menirukan perilaku manusia untuk menyelesaikan masalah atau melakukan otomatisasi. Sesuai namanya, ML mencoba menirukan bagaimana proses manusia atau makhluk cerdas belajar dan menggeneralisasi.¹ Setidaknya ada dua aplikasi utama dalam ML yaitu, klasifikasi dan prediksi. Ciri khas dari ML adalah adanya proses pelatihan, pembelajaran, atau training. Oleh karena itu, ML membutuhkan data untuk dipelajari yang disebut sebagai data training. Klasifikasi adalah metode dalam ML yang digunakan oleh mesin untuk memilah atau mengklasifikasikan obyek berdasarkan ciri tertentu sebagaimana manusia mencoba membedakan benda satu dengan yang lain. Sedangkan prediksi atau regresi digunakan oleh mesin untuk menerka keluaran dari suatu

data masukan berdasarkan data yang sudah dipelajari dalam training. Metode ML yang paling populer yaitu Sistem Pengambil Keputusan, Support Vector Machine (SVM) dan Neural Network.

Model Klasifikasi

Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukannya kedalam kelas tertentu dari jumlah kelas yang tersedia. Klasifikasi melakukan pembangunan model berdasarkan data latih yang ada, kemudian menggunakan model tersebut untuk mengklasifikasikan pada data yang baru. Klasifikasi dapat didefinisikan sebagai pekerjaan yang melakukan pelatihan/pembelajaran terhadap fungsi target yang memetakan setiap set atribut (fitur) ke satu jumlah label kelas yang tersedia. Sebuah sistem yang melakukan klasifikasi diharapkan dapat melakukan klasifikasi semua data set dengan benar, tetapi tidak dapat dipungkiri bahwa kinerja sistem tidak bisa 100% benar sehingga sebuah sistem klasifikasi juga harus diukur kinerjanya. Umumnya, pengukuran kinerja klasifikasi dilakukan dengan matriks konfusi. Klasifikasi adalah tipe analisis data yang dapat membantu orang menentukan kelas label dari sampel yang ingin di klasifikasi. Klasifikasi merupakan Metode supervised learning, metode yang mencoba menemukan hubungan antara atribut masukan dan atribut target. Tujuan klasifikasi untuk meningkatkan kehandalan hasil yang diperoleh dari data. Beberapa model yang terdapat dalam klasifikasi antara lain: k-Nearest Neighbor, Naive Bayes, Support Vector Maching, Boosting dan Decision Tree Random Forest. Akurasi Model Klasifikasi dapat dilihat salah satu dari accuracy, Sensitive atau Recall dan Precision yang didapatkan dari Confusion Matrix.

METODE

Metode penelitian ini menggunakan beberapa tahapan yaitu:

1. Penyusunan Instrumen
Pada tahapan ini penyusunan instrumen dilakukan untuk alat pengumpulan data terkait kelayakan penerima bantuan program PIP. Ketua peneliti menyusun draf instrumen yang akan dilakukan pengujian instrumen dan divalidasi oleh pakar. Sedangkan anggota melakukan mendokumentasikan kegiatan validasi oleh pakar.
2. Pengumpulan Data
Pengumpulan data dilakukan terkait data populasi penerima program PIP beserta seluruh criteria seleksi yang digunakan selama ini. Ketua peneliti pada tahapan ini melakukan koordinasi pada satuan pendidikan untuk mendapatkan izin pelaksanaan. Anggota melakukan penyebaran instrumen pengumpulan data dan mendokumentasikan seluruh data dan proses selama pengumpulan data
3. Menganalisis Data dengan Model Klasifikasi
Data yang dikumpulkan akan dianalisis menggunakan model regresi yang ada didalam metode k-Nearest Neighbor, Naive Bayes, Support Vector Maching, Boosting dan Decision Tree Random Forest. Pada tahap ini anggota melakukan kegiatan data cleaning, data preprocessing dan data transformation sebagai persiapan data sebelum dilakukan analisis. Ketua peneliti bertugas untuk menganalisis data menggunakan model klasifikasi menggunakan Python.
4. Perancangan dan Implementasi Sistem
Setelah mendapatkan model klasifikasi dengan performance accuracy, *sensitivity*, dan *precision* terbaik, selanjutnya dilakukan proses perancangan sistem dan implementasi sistem. Bahasan

pemograman yang digunakan adalah python dan HTML. Pada tahapan ini ketua melakukan analisa perbandingan performa klasifikas. Selanjutnya ketua dan anggota bersama-sama dengan anggota melakukan proses perancangan dan implementasi sistem berbasis web.

5. Evaluasi Sistem
Evaluasi sistem dilakukan untuk mengetahui kekurangan sistem yang telah dibangun serta menerima masukan dari pengguna terkait feasibility study produk dilakukan di satuan pendidikan. Anggota mengumpulkan instrument evaluasi sistem. Ketua dan anggota peneliti melakukan diskusi bersama manajemen dan penyesuaian sistem dari masukan dan melakukan perbaikan sistem.

HASIL DAN PEMBAHASAN

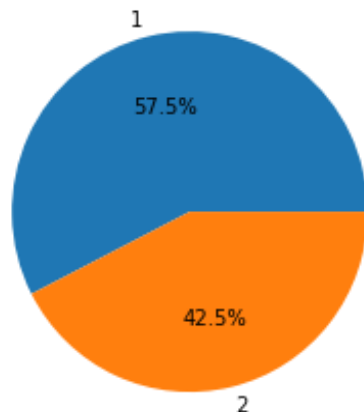
Dataset yang terkumpul sebanyak 73 menjelaskan bahwa dalam seleksi penerima beasiswa PIP di SMA taman siswa sukadamai menggunakan kriteria Pengasilan Orang Tua, Tanggungan, Kepemilikan Rumah, Pembayaran Listrik, Pembayaran Air dan Kelayakan. Data set ini kemudian dilakukan proses pembersihan data hingga data transformasi untuk dapat diolah menggunakan model Machine Learning.

	penghasilanortua	tanggungan	kepemilikrumah	pmbayaranlistrik	pmbayaranair	kelayakan
0	3	2	2	3	2	2
1	4	2	2	2	3	2
2	1	1	1	1	1	1
3	2	2	2	2	2	1
4	1	2	1	2	2	1
...
68	1	1	1	2	2	1
69	3	2	2	1	1	1
70	3	3	2	3	2	2
71	4	2	2	3	3	2
72	3	2	1	2	3	2

73 rows x 6 columns

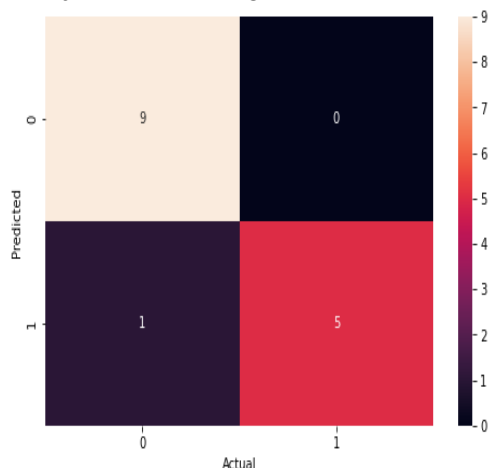
Gambar. Hasil import dataset pada Python

Persentase Setiap Class pada Dataset



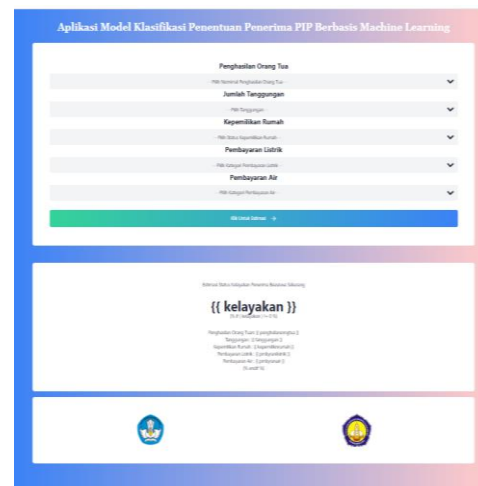
Gambar. Gambar Pembagian Class data Layak (1) dan Tidak Layak (2)

Selanjutnya dilakukan pembagian data training sebenar 80% dan data testig sebesar 20%. Selanjutnya dilakukan pengukuran akurasi dari model klasifikasi ditunjukkan melalui gambar.



Gambar. Pengukuran Akurasi model dengan Seaborn

dimana hasil akurasi menghasilkan : Tingkat Akurasi dengan menggunakan Naive Bayes sebesar 0.9333333333333333. Dikarenakan akurasi yang telah dihasilkan menunjukkan performa yang cukup baik yaitu 93%, maka model siap untuk di deploy secara local menggunakan HTML.



Gambar. Deploy model klasifikasi

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, model klasifikasi naive bayes memiliki kinerja yang baik dengan akurasi sebesar 93,3% dengan rasio data uji sebesar 20% dan data training 80%.

UCAPAN TERIMA KASIH

Ucapan terima kasih kepada Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi atas pendanaan yang diberikan sehingga penelitian ini

DAFTAR PUSTAKA

Pratama, F.D., Zufria, I., & Triase, T. (2022). Implementasi Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Penerima Program Indonesia Pintar. *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, 7(1), 77-84.

Rohaeni, N. E., & Saryono, O. (2018). Implementasi Kebijakan Indonesia Pintar (PIP) Melalui Kartu Indonesia Pintar (KIP) dalam Upaya Pemerataan Pendidikan. *Indonesian Journal of Education Management & Administration Review*, 2(1), 193-204.

- Gunawan, R. D., Oktavia, T., & Borman, R. I. B. I. (2018). Perancangan Sistem Informasi Beasiswa Program Indonesia Pintar (PIP) Berbasis Online (Tudi Kasus: SMA N 1 Kota Bumi). *MIKROTIK: Jurnal Manajemen Informatika*, 8(1), 43-54.
- Yusup, W. B., Ismanto, B., & Wasitohadi, W. (2019). Evaluasi Program Indonesia Pintar dalam peningkatan akses pendidikan di sekolah menengah pertama. *Kelola: Jurnal Manajemen Pendidikan*, 6(1), 44-53.
- Perdana, N. G., & Widodo, T. (2013). Sistem Pendukung Keputusan Pemberian Beasiswa Kepada Peserta Didik Baru Menggunakan Metode TOPSIS. *Semantik* 2013, 3(1), 265-272.
- Limbong, T., Muttaqin, M., Iskandar, A., Windarto, A. P., Simarmata, J., Mesran, M., ...& Wanto, A. (2020). Sistem Pendukung Keputusan: Metode & Implementasi. *Yayasan Kita Menulis*.
- Fauzan, R., Indrasary, Y., & Muthia, N. (2018). Sistem Pendukung Keputusan Penerimaan Beasiswa Bidik Misi di POLIBAN dengan Metode SAW Berbasis Web. *Jurnal Online Informatika*, 2(2), 79-83.
- Shobha, G., & Rangaswamy, S. (2018). Machine learning. In *Handbook of statistics* (Vol. 38, pp. 197-228). Elsevier.
- Ahmad, A. (2017). Mengenal artificial intelligence, machine learning, neural network, dan deep learning. *J. Teknol. Indones.*, no. October, 3.
- Utomo, D. P., & Mesran, M. (2020). Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung. *Jurnal Media Informatika Budidarma*, 4(2), 437-444.
- Hendrian, S. (2018). Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan. *Faktor Exacta*, 11(3).
- Wahyuni, E. D., Arifiyanti, A. A., & Kustyani, M. (2019). Exploratory data analysis dalam konteks klasifikasi data mining. *ReTII*, 263-269.
- Annur, H. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes. *ILKOM Jurnal Ilmiah*, 10(2), 160-165.
- Utomo, D. P., & Purba, B. (2019, September). Penerapan Datamining pada Data Gempa Bumi Terhadap Potensi Tsunami di Indonesia. In *Prosiding Seminar Nasional Riset Information Science (SENARIS)* (Vol. 1, pp. 846-853).
- Mustafa, M. S., Ramadhan, M. R., & Thenata, A. P. (2018). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Creative Information Technology Journal*, 4(2), 151-162..