

Penerapan Machine Learning menggunakan algoritma C4.5 berbasis PSO dalam Menganalisa Data Siswa Putus Sekolah

Agus Surip¹, Muhamad Aji Pratama², Irfan Ali³, Arif Rinaldi Dikananda⁴, Ade Irma Purnamasari^{5*}

¹Rekayasa Perangkat Lunak; STMIK IKMI CIREBON; Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45131, 0231-490480; e-mail: info.ikmicirebon@gmail.com.

²Rekayasa Perangkat Lunak; STMIK IKMI CIREBON; Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45131, 0231-490480; e-mail: info.ikmicirebon@gmail.com.

³Rekayasa Perangkat Lunak; STMIK IKMI CIREBON; Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45131, 0231-490480; e-mail: info.ikmicirebon@gmail.com.

⁴Rekayasa Perangkat Lunak; STMIK IKMI CIREBON; Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45131, 0231-490480; e-mail: info.ikmicirebon@gmail.com.

⁵Teknik Informatika; STMIK IKMI CIREBON; Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45131, 0231-490480; e-mail: info.ikmicirebon@gmail.com.

* Korespondensi: e-mail: irma2974@yahoo.com

Diterima: 29 Maret 2021; Review: 8 April 2021; Disetujui: 13 April 2021, 147-155

Cara sitasi: Agus Surip, Muhamad Aji Pratama, Irfan Ali, Arif Rinaldi Dikananda, Ade Irma Purnamasari. 2021. Penerapan Machine Learning menggunakan algoritma C4.5 berbasis PSO dalam Menganalisa Data Siswa Putus Sekolah. Informatics for Educators and Professionals. Vol (5) No.2 :Hal: 147-155

Abstrak : Pendidikan sangat memegang peranan penting dalam meningkatkan kualitas Sumber daya manusia yang ada di Indonesia. Untuk itu Pemerintah menggalakkan program wajib belajar 12 tahun. SMK As Salam, adalah salah satu Sekolah Menengah Kejuruan (SMK) swasta yang ada di Kecamatan Gegesik yang memiliki angka putus sekolah yang berubah setiap tahunnya. Ada beberapa penyebab seorang anak putus sekolah, diantaranya faktor ekonomi dimana orang tua tidak sanggup membiayai anaknya untuk sekolah, bisa juga faktor membantu usaha orang tua, faktor siswanya sendiri, atau faktor lingkungan dan teman – teman siswa tersebut. Perlu dilakukan penganalisaan terhadap data siswa agar dapat diklasifikasikan siswa yang berpotensi putus sekolah. Sehingga kepala sekolah dan bagian kesiswaan di SMK As Salam dapat mengambil keputusan agar siswa yang putus sekolah bisa diantisipasi setiap tahunnya. fokus maka penelitian ini adalah penggunaan model decision tree dalam mengkasifikasi data putus sekolah yang dioptimasi menggunakan PSO di SMK As Salam Gegesik. Dari hasil klasifikasi tersebut diketahui nilai akurasi dari model decision tree tersebut adalah 90.86 %. Artinya keakuratan dalam klasifikasi tersebut sudah cukup baik. Sedangkan hasil klasifikasi model decision tree yang menggunakan PSO diketahui memiliki nilai akurasi 92.95 %. Artinya keakuratan dalam klasifikasi tersebut menjadi lebih baik dari decision tree yang tanpa optimasi PSO.

Kata kunci: Sekolah, Decision Tree, PSO

Abstract : Education plays an important role in improving the quality of human resources in Indonesia. For this reason, the Government is promoting a 12 year compulsory education program. SMK As Salam, is one of the private Vocational High Schools (SMK) in Gegesik District which has a dropout rate that changes every year. There are several reasons for a child dropping out of school, including economic factors where parents cannot afford to pay for their children to go to school, it could also be factors helping parents' businesses, students' own

factors, or environmental factors and students' friends. It is necessary to analyze student data so that it can be classified as students who have the potential to drop out of school. So that the principal and the student department at SMK As Salam can make decisions so that students dropping out of school can be anticipated every year. The focus of this research is the use of the decision tree model in classifying the data on dropout which is optimized using PSO at SMK As Salam Gegesik. From the classification results, it is known that the accuracy value of the decision tree model is 90.86%. This means that the accuracy of the classification is good enough. While the classification results of the decision tree model using PSO are known to have an accuracy value of 92.95%. This means that the classification accuracy is better than the decision tree without PSO optimization

Keywords : *Dropping out of school, Decision Tree, PSO*

1. Pendahuluan

Pendidikan sangat memegang peranan penting dalam meningkatkan kualitas SDM. Oleh karena itu, sejak tahun 1994 pemerintah telah menerapkan program wajib belajar pendidikan dasar 9 tahun dengan harapan semakin baik tingkat pendidikan akan semakin baik pula tingkat kesejahteraan. Selanjutnya program wajib belajar ditambah menjadi 12 tahun, hal ini seiring dengan perkembangan teknologi yang menuntut pendidikan dan keterampilan yang semakin tinggi. Sejalan dengan yang digaungkan oleh Pemerintah Kabupaten Cirebon terutama di Kecamatan Gegesik yang memiliki lembaga pendidikan dari TK sampai SMA/SMK.

Ada banyak cara yang dapat digunakan untuk menganalisis permasalahan di dunia pendidikan, salah satunya dengan data mining. Data mining merupakan gabungan sejumlah disiplin ilmu komputer yang didefinisikan sebagai penemuan pola-pola baru dari kumpulan-kumpulan data sangat besar, meliputi metode-metode yang merupakan irisan dari artificial intelligence, machine learning, statistics, dan database systems. Data mining sering juga disebut knowledge discovery in database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan. [1]

Contoh penelitian yang dilakukan dengan model algoritma klasifikasi C4.5 dan Particle Swarm Optimization adalah penelitian untuk membantu pendidik memprediksi siswa bermasalah dengan akurat. Metode penelitian yang digunakan metode eksperimen, dengan mengambil data dari SMAN 2 Banjarbaru. Kesimpulan yang didapat dalam penelitian ini adalah optimasi decision tree C4.5 dan Particle Swarm Optimization menjadi lebih baik dalam memprediksi nilai akurasi dari pada hanya menggunakan metode C4.5 saja. [2]

Penelitian lain yang berkaitan dengan optimasi decision tree menggunakan Particle Swarm Optimization yaitu meneliti tentang data siswa putus sekolah siswa kelas XII pada SMAN 2 Barabai, Hulu Sungai Tengah Banjarmasin. Metode penelitian yang digunakan metode eksperimen dengan tahapan pengumpulan data (data gathering), pengolahan data awal (data pre-processing), desain eksperimen dan pengujian metode. Pada tahap ini eksperimen dilakukan dengan menerapkan ID3 pada data yang sudah melalui preprocessing dan diuji tingkat akurasi klasifikasinya dengan confusion matrix. Kemudian diterapkan pada PSO dan diuji tingkat akurasi untuk dibandingkan dengan hasil sebelumnya yang tanpa menggunakan PSO. Kesimpulan dalam penelitian ini bahwa PSO dapat meningkatkan akurasi dari klasifikasi ID3 yang diterapkan pada data siswa putus sekolah.[3]

Berdasarkan data dari neraca Pendidikan daerah Kab. Cirebon yang diakses di situs <https://npd.kemdikbud.go.id/?> siswa yang putus sekolah di Sekolah Menengah Kejuruan (SMK) setiap tahun ada perubahan, berikut datanya :

Tabel 1 Jumlah Siswa Putus Sekolah Kab. Cirebon 3 tahun terakhir

No	Tahun	Jumlah Siswa Putus Sekolah
1	2017	1.051 siswa (Per bulan Nov 2017)
2	2018	1.017 siswa (Per bulan Nov 2018)
3	2019	145 siswa (Per bulan Juli 2019)

SMK As Salam, adalah salah satu Sekolah Menengah Kejuruan (SMK) swasta yang ada di Kecamatan Gegesik yang memiliki angka putus sekolah yang berubah setiap tahunnya.

Berikut tabel siswa putus sekolah di SMK As Salam dalam 3 tahun terakhir yang bersumber dari informasi Wakil Kesiswaan SMK As Salam Gegesik.

Tabel 2 Jumlah Siswa Putus Sekolah SMK As Salam 3 tahun terakhir

No	Tahun	Jumlah Siswa Putus Sekolah
1	2018	15 Siswa
2	2019	12 Siswa
3	2020	10 Siswa

Ada beberapa faktor penyebab seorang anak putus sekolah, diantaranya faktor ekonomi, faktor orang tua disebabkan oleh latar belakang pendidikan orang tua yang sedang-rendah, broken home, kemudian orang tua meninggal, hingga keputusan orang tua yang salah, faktor internal siswa disebabkan kurangnya minat terhadap sekolah dan kurangnya motivasi bersekolah, faktor lingkungan baik itu lingkungan tempat tinggal maupun lingkungan di sekolah.[4]

Akar masalah dari penelitian ini adalah perlu dilakukan penganalisaan terhadap data siswa sehingga dapat diklasifikasikan siswa yang berpotensi putus sekolah. Adapun data yang klasifikasikan diantaranya Nama Siswa, No Induk Siswa, jenis kelamin, pekerjaan dan penghasilan Orang tua, berkebutuhan Khusus atau tidak, penerima KIP atau bukan, jarak dari rumah ke sekolah dan kehadiran siswa yang meliputi jumlah alpa, jumlah izin dan jumlah sakit. Kemudian dari data tersebut kepala sekolah dan bagian kesiswaan di SMK As Salam dapat memprediksi apakah siswa tersebut berpotensi putus sekolah atau tidak, sehingga pihak sekolah dapat menentukan kebijakan yang harus diambil selanjutnya.

2. Metode Penelitian

Metode penelitian adalah langkah yang dimiliki dan dilakukan oleh peneliti dalam rangka untuk mengumpulkan informasi atau data serta melakukan investigasi pada data yang telah didapatkan tersebut. Metode penelitian kuantitatif merupakan suatu penelitian yang menggunakan data berupa angka-angka yang terkumpul sebagai hasil penelitian, dan dianalisis dengan menggunakan metode statistika [4].

Pada penelitian ini, penulis menggunakan pendekatan penelitian kuantitatif. Metode penelitian kuantitatif dapat diartikan juga sebagai metode penelitian yang berlandaskan pada filsafat positivisme, digunakan untuk meneliti pada populasi atau sampel tertentu, pengumpulan data menggunakan instrumen penelitian, analisis data bersifat kuantitatif/statistik, dengan tujuan untuk menguji hipotesis yang telah ditetapkan. Metode ini disebut sebagai metode positivistik karena berlandaskan pada filsafat positivisme. Metode ini sebagai metode ilmiah karena telah menemui kaidah-kaidah ilmiah yaitu konkrit/empiris, objektif, terukur, rasional dan sistematis. Metode ini disebut metode kuantitatif karena data penelitian berupa angka-angka dan analisis menggunakan statistik.[5]

Tahapan penelitian merupakan suatu proses memperoleh atau mendapatkan suatu pengetahuan atau memecahkan permasalahan yang dihadapi, yang dilakukan secara ilmiah, sistematis dan logis. Tahapan penelitian pada penelitian adalah sebagai berikut :

Teknik Pengumpulan data

Data adalah unit informasi yang direkam media yang dapat dibedakan dengan data lain, dapat dianalisis dan relevan dengan program tertentu. Dalam penelitian ini tahapan teknik pengumpulan data yang penulis lakukan sebagai berikut:

Observasi merupakan metode yang dilakukan penulis dengan cara mendatangi langsung tempat riset yang ingin diteliti penulis. Penulis melakukan pengamatan langsung dengan beberapa acuan yaitu menganalisis beberapa dokumen terkait dengan sekolah tersebut termasuk data jumlah siswa, data siswa putus sekolah, penghasilan orang tua dan data yang diperlukan penulis yang mendukung penelitian ini dalam hal ini di SMK As Salam Gegesik.

Wawancara: metode yang dilakukan penulis dengan cara melakukan tanya jawab secara langsung kepada narasumber. Penulis melakukan tanya jawab dengan kepala sekolah SMK As Salam Gegesik, wakil kepala sekolah bagian kesiswaan SMK As Salam Gegesik.

Studi Pustaka merupakan metode yang digunakan penulis sebagai pendukung dan referensi. Buku, jurnal, dan situs yang berhubungan dengan penelitian dan penulisan.

Seleksi

Pada tahap awal ini penulis menyeleksi data – data yang diperlukan dalam penelitian seperti Nama siswa, Nomor Induk Siswa Nasional, jenis kelamin siswa, pekerjaan Orang tua, penghasilan orang tua, alamat siswa, jarak siswa ke sekolah, jumlah keluarga siswa .

Preprocessing

Pembersihan data (data cleaning) merupakan proses menghilangkan noise dan data yang tidak konsisten. Data yang digunakan dalam penulisan penelitian ini terdiri dari dua sumber data, yaitu data induk siswa (Nama, NISN dan Jenis Kelamin) dan data nilai (orang tua pekerjaan Orang tua, penghasilan orang tua).

Data Mining

Data mining mengolah data mentah yang tersimpan di dalam basis data sehingga menghasilkan informasi yang berguna. Data mentah merupakan data yang disimpan hanya sebagai dokumentasi. Pengolahan data mentah dapat dimanfaatkan untuk pembuatan laporan, mengontrol kebutuhan, menemukan informasi dan pengetahuan baru. Data mining menelusuri data pada database untuk membangun model dan menggunakannya untuk mengenali pola data lain yang tidak tersimpan dalam basis data. Tahap ini memilih Algoritma Particle Swarm Optimization (PSO) yang akan digunakan dalam klasifikasi.

Evaluasi

Evaluasi model dengan dilakukan penghitungan fungsi precision dan recall terhadap model yang dibangun. Dengan membandingkan hasil prediksi positif benar dan hasil positif salah pada suatu kelas klasifikasi. Untuk proses evaluasi dengan confusion matrix maka akan diperoleh nilai precision, recall, dan accuracy yang didapat dari rumus:

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} * 100\%$$

$$Presisi = \frac{TP}{FP+TP} * 100\%$$

$$Recall = \frac{TP}{FN+TP} * 100\%$$

[11]

Keterangan :

TP : jumlah kasus positif yang diklasifikasikan sebagai positif

FP : jumlah kasus negatif yang diklasifikasikan sebagai positif

TN : jumlah kasus negatif yang diklasifikasikan sebagai negatif

FN : jumlah kasus positif yang diklasifikasikan sebagai negative

3. Hasil dan Pembahasan

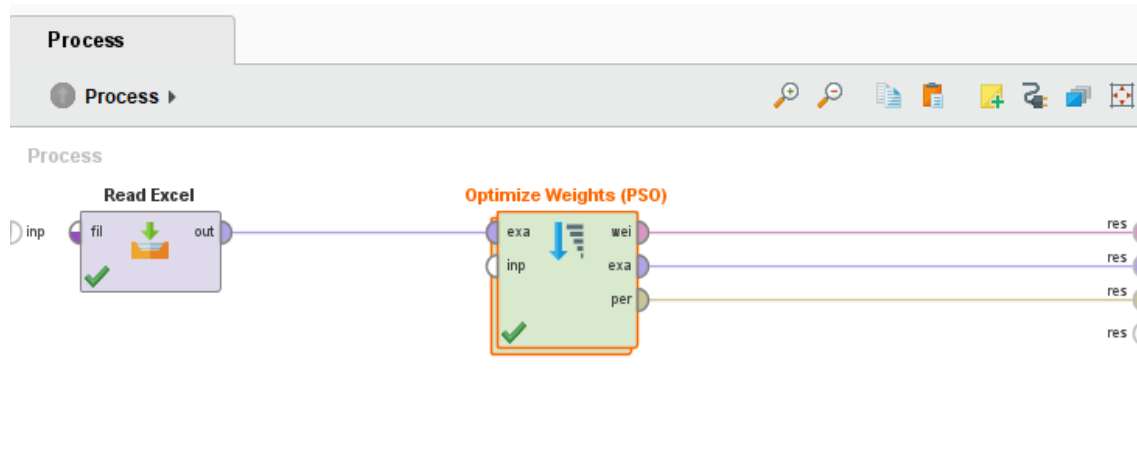
Hasil

Tahap awal dalam penelitian ini adalah menyeleksi data siswa siswa SMK As Salam Kecamatan Gegesik yang berjumlah 440 siswa pada Tahun Ajaran 2020/2021 data yang dibutuhkan adalah Nama, Jenis Kelamin, NISN, Alamat, Pekerjaan, Penghasilan Orang Tua, Jumlah Alpha, Jumlah Sakit, Jumlah Izin, Jarak Rumah Ke Sekolah, dan Potensi Putus Sekolah.

Data yang didapat adalah data mentah yang akan digunakan dalam pemrosesan data mining. Agar data dapat diproses dengan baik data harus dalam keadaan Normal, maka perlu adanya preprocessing. Agar tidak ada lagi data yang mengandung noise atau missing value dan data dinyatakan normal, data terlebih dahulu harus melalui proses KDD. Proses KDD sendiri merupakan proses dimana untuk mencari dan mengidentifikasi pola sebelum data ini dapat diolah lebih lanjut, Sehingga data sudah bisa kita gunakan sebagai dataset.

Karena data tersebut masih belum bisa digunakan untuk pemrosesan data mining, karena terdapat 2 atribut identitas (nomer unik/nomer identitas) dan pada atribut NIPD banyak data yang missing value. Maka langkah preprocessing yang dilakukan adalah dengan membuang atribut NIPD dengan alasan banyak data missing value dan terdapat 2 nomer identitas (NIPD dan NISN). Karena atribut NISN diisi dengan lengkap oleh pihak Tata Usaha SMK As Salam Gegesik, maka atribut NISN memiliki role sebagai ID.

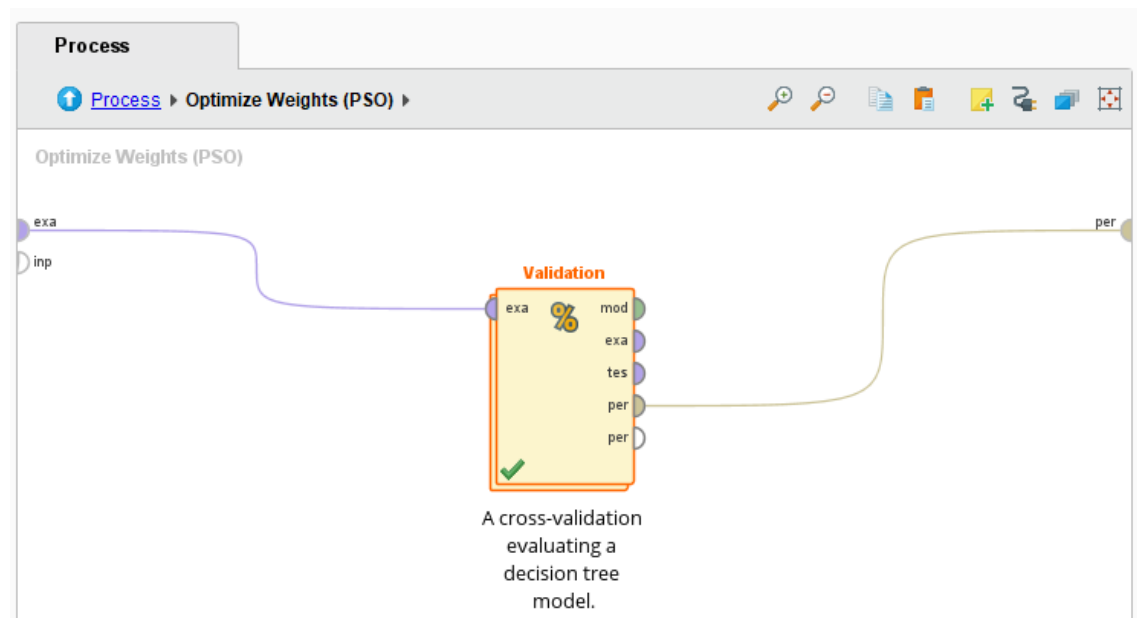
Pada penelitian ini, model proses yang digunakan adalah Decision Tree berbasis PSO, dengan menggunakan tools Rapid Miner. Rapid Miner merupakan software platform untuk mengolah berbagai data mining. Sebelum mengeolah sebuah data mining dengan tools Rapid Miner, langkah pertama yang harus dilakukan yaitu menyusun atau membuat model proses dengan operator yang sesuai dengan teknik data mining yang digunakan dan operator yang telah tersedia pada tools Rapid Miner.



Sumber : Hasil Penelitian (2021)

Gambar 1 Model Proses Secara Global

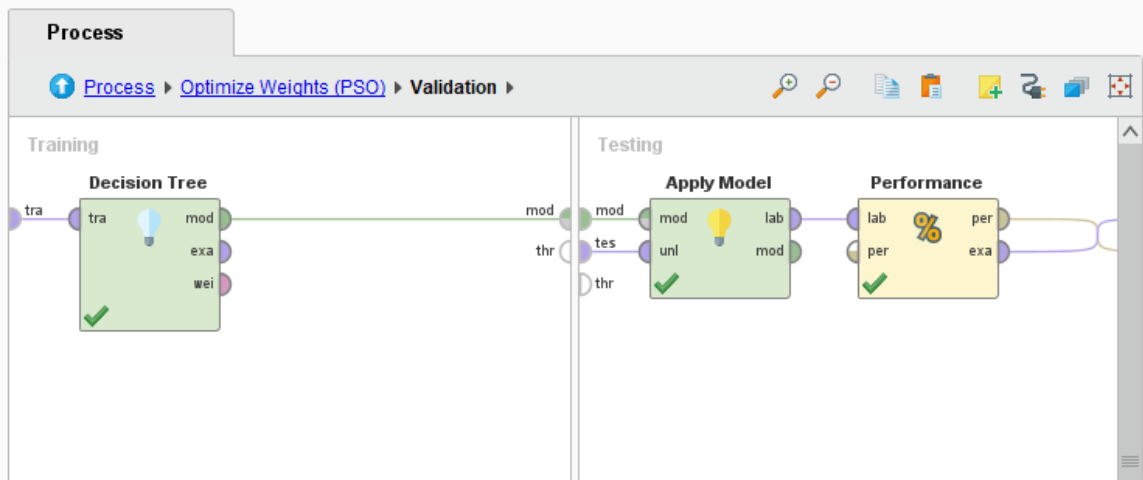
Dalam model tersebut terdapat 2 operator untuk menggambarkan model secara umum, yaitu operator read excel dan operator optimize weights (PSO). Operator read excel berfungsi untuk penyedia data set yang akan digunakan. Sedangkan operator optimize weights PSO digunakan untuk model pemrosesannya. Di dalam operator optimize weights (PSO) terdapat operator lain, berikut gambarnya :



Sumber : Hasil Penelitian (2021)

Gambar 2 Operator Validation di dalam Operator Optimize Weights (PSO)

Operator tersebut digunakan untuk jenis model klasifikasi, dimana di dalam operator terdapat beberapa operator lagi di dalamnya. Operator validation ini berfungsi untuk memvalidasi hasil proses algoritma utama yang digunakan yaitu decision tree.



Sumber : Hasil Penelitian (2021)

Gambar 3 Model Decision Tree

Model tersebut ada di dalam operator validation, model ini adalah model utama yang digunakan dalam memprediksi atau mengklasifikasi data. Langkah selanjutnya adalah menekan tombol start proses dan model proses akan segera dieksekusi. Berikut hasil dari pemrosesan :



Sumber : Hasil Penelitian (2021)

Gambar 4 Graph Decision Tree

Graph hasil pemrosesan tersebut menampilkan model pohon keputusan yang dihasilkan dari proses tersebut. Aturan pertama jika Alpha lebih dari 5,5 maka berpotensi putus sekolah. Jika alpha kurang dari 5,5 maka turun ke aturan kedua yaitu jika jumlah izin lebih dari 3,5 maka siswa berpotensi putus sekolah. Sedangkan jika jumlah izin kurang atau sama dengan 3,5 maka turun ke aturan ketiga yaitu jika jumlah sakit lebih dari 4 maka siswa berpotensi putus sekolah. Sedangkan jika jumlah sakit kurang dari 4 maka siswa tidak berpotensi putus sekolah.

Berdasarkan plot dengan jumlah alpha melebihi 5 memiliki potensi putus sekolah yang dominan. Sedangkan untuk hasil plot lainnya yaitu penghasilan orang tua / wali dengan penghasilan 1 juta lebih tidak memiliki potensi putus sekolah.

Evaluasi yang pertama untuk mengukur kualitas klasifikasi tersebut ialah dengan melihat tingkat akurasi. Berikut adalah performa hasil klasifikasi yang dilakukan :

Tabel 3 Akurasi *Decission tree*

	<i>true tidak</i>	<i>true ya</i>	<i>class precision</i>
pred. tidak	279	39	87.74%
pred. ya	2	120	98.36%
<i>class recall</i>	99.29%	75.47%	

Sumber : Hasil Penelitian (2021)

Berdasarkan tabel diatas dapat diketahui prediksi tidak dengan *true* tidak terdapat 279 item, sedangkan prediksi tidak dengan *true ya* memiliki 39 item, sehingga prediksi tidak memiliki *class precision* sebesar 87.74 %. Sedangkan prediksi dengan *true* tidak memiliki 2 item dan prediksi ya dengan *true ya* memiliki 120 item sehingga prediksi tersebut memiliki *class precision* sebesar 98.36%. Dari hasil klasifikasi tersebut diketahui nilai akurasi dari model tersebut adalah 90.86 %. Artinya keakuratan dalam klasifikasi tersebut sudah cukup baik.

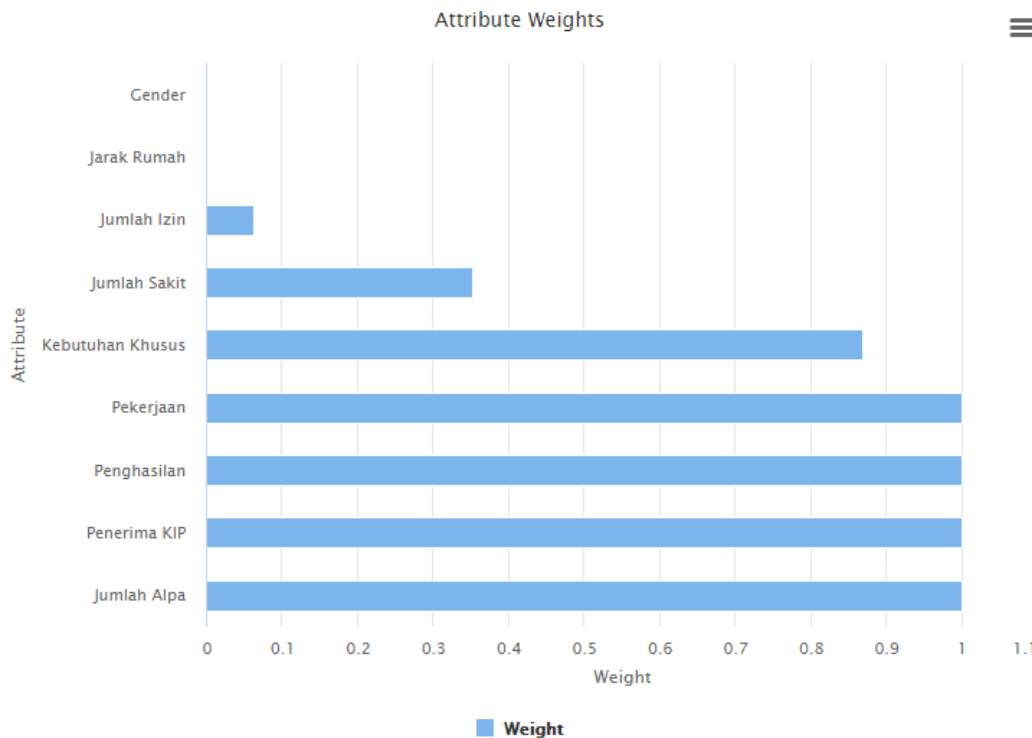
Hasil tersebut adalah proses dengan tanpa optimasi dari PSO. Sedangkan hasil performa *decision tree* yang dioptimasi dengan PSO adalah sebagai berikut :

Tabel 4 Akurasi *decision tree* dengan PSO

	<i>true tidak</i>	<i>true ya</i>	<i>class precision</i>
pred. tidak	276	26	91.39%
pred. ya	5	133	96.38%
<i>class recall</i>	98.22%	83.65%	

Sumber : Hasil Penelitian (2021)

Berdasarkan tabel diatas dapat diketahui prediksi tidak dengan *true* tidak terdapat 276 item, sedangkan prediksi tidak dengan *true ya* memiliki 26 item, sehingga prediksi tidak memiliki *class precision* sebesar 91.74 %. Sedangkan prediksi dengan *true* tidak memiliki 5 item dan prediksi ya dengan *true ya* memiliki 133 item sehingga prediksi tersebut memiliki *class precision* sebesar 96.38 %. Dari hasil klasifikasi tersebut diketahui nilai akurasi dari model tersebut adalah 92.95 %. Artinya keakuratan dalam klasifikasi tersebut menjadi lebih baik dari *decision tree* yang tanpa optimasi PSO. Adapun attribute *weights* hasil dari PSO dapat diketahui sebagai berikut :



Sumber : Hasil Penelitian (2021)

Gambar 5 Attribute Weights

Berdasarkan plot tersebut dapat dilihat bahwa terdapat 2 atribut yang tidak memiliki *weight*. Atribut tersebut adalah gender dan jarak rumah. Dapat disimpulkan bahwa atribut gender dan jarak tidak berpengaruh terhadap hasil prediksi yang dilakukan. Dari hasil klasifikasi tersebut diketahui nilai akurasi dari model tersebut adalah 92.95 %. Artinya keakuratan dalam klasifikasi tersebut menjadi lebih baik dari *decision tree* yang tanpa optimasi PSO

Pembahasan

Berdasarkan hasil pemrosesan tersebut dapat diketahui bahwa diketahui hasil dari *decision tree* biasa memiliki prediksi tidak dengan *true* tidak terdapat 279 item, sedangkan prediksi tidak dengan *true* ya memiliki 39 item, sehingga prediksi tidak memiliki *class precision* sebesar 87.74 %. Sedangkan prediksi dengan *true* tidak memiliki 2 item dan prediksi ya dengan *true* ya memiliki 120 item sehingga prediksi tersebut memiliki *class precision* sebesar 98.36 %. Dari hasil klasifikasi tersebut diketahui nilai akurasi dari model tersebut adalah 90.86 %. Artinya keakuratan dalam klasifikasi tersebut sudah cukup baik.

Sedangkan hasil dari *decision tree* yang dioptimasi menggunakan PSO memiliki prediksi tidak dengan *true* tidak terdapat 276 item, sedangkan prediksi tidak dengan *true* ya memiliki 26 item, sehingga prediksi tidak memiliki *class precision* sebesar 91.74 %. Sedangkan prediksi dengan *true* tidak memiliki 5 item dan prediksi ya dengan *true* ya memiliki 133 item sehingga prediksi tersebut memiliki *class precision* sebesar 96.38 %. Dari hasil klasifikasi tersebut diketahui nilai akurasi dari model tersebut adalah 92.95 %. Artinya keakuratan dalam klasifikasi tersebut menjadi lebih baik dari *decision tree* yang tanpa optimasi PSO.

4. Kesimpulan

Berdasarkan dari hasil penelitian yang dilakukan di SMK As Salam Gegesik, maka kesimpulan yang didapat adalah sebagai berikut, algoritma *decision tree* dapat digunakan untuk mengklasifikasi siswa yang berpotensi putus sekolah. Adapun yang menjadi kriteria adalah Nama siswa, NISN, Gender, Pekerjaan Orang tua dan Penghasilan Orang tua, kebutuhan khusus, penerima KIP, jumlah alpa, jumlah izin, jumlah sakit, jarak rumah.

Algoritma decision tree yang dioptimasi dengan PSO memiliki akurasi lebih baik dibandingkan dengan decision tree yang tanpa dioptimasi. Hal ini didasarkan pada hasil dari decision tree memiliki class precision prediksi tidak sebesar 87.74 % dan prediksi ya sebesar 98.36 %. Sedangkan hasil dari decision tree yang dioptimasi menggunakan PSO memiliki class precision prediksi tidak sebesar 91.74 % dan prediksi ya class precision sebesar 96.38 %.

Dari hasil klasifikasi diketahui nilai akurasi dari model decision tree adalah 90.86 % artinya keakuratan dalam klasifikasi tersebut sudah cukup baik. Sedangkan Dari hasil klasifikasi model decision tree yang dioptimasi menggunakan PSO diketahui nilai akurasi adalah 92.95 %. Artinya keakuratan dalam klasifikasi tersebut menjadi lebih baik dari decision tree yang tanpa optimasi PSO.

Agar selanjutnya penelitian ini dapat dikembangkan lagi kedepannya, berikut adalah saran dari peneliti yang dijadikan bahan renungan penelitian yang akan datang, agar hasil penelitian ini lebih baik lagi, dapat menggunakan jumlah data yang lebih banyak lagi dan melakukan pengujian dengan dataset yang lebih lengkap atributnya sehingga model yang sudah dihasilkan akan lebih teruji lagi kedepannya; Menggunakan algoritma lain dalam klasifikasi data misalnya ANN atau Naïve Bayes; Untuk mendapatkan nilai akurasi yang lebih baik lagi dari penelitian ini, dapat melakukan optimasi dengan teknik optimasi lainnya.

Referensi

- [1]. O. Nurdiawan, F. A. Pratama, D. A. Kurnia, Kaslani, and N. Rahaningsih, "Optimization of Traveling Salesman Problem on Scheduling Tour Packages using Genetic Algorithms," *J. Phys. Conf. Ser.*, vol. 1477, no. 5, pp. 0–6, 2020, doi: 10.1088/1742-6596/1477/5/052037.
- [2]. Febie Elfaladonna, A. R. (2019). Analisa Metode Classification-Decission Tree Dan Algoritma. *Science And Information Technology*, 2(1), 10–17.
- [3]. H. Mustofa and A. A. Mahfudh, "Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes," *Walisongo J. Inf. Technol.*, vol. 1, no. 1, p. 1, 2019, doi: 10.21580/wjit.2019.1.1.3915.
- [4]. O. Nurdiawan, F. A. Pratama, D. A. Kurnia, Kaslani, and N. Rahaningsih, "Optimization of Traveling Salesman Problem on Scheduling Tour Packages using Genetic Algorithms," *J. Phys. Conf. Ser.*, vol. 1477, no. 5, 2020, doi: 10.1088/1742-6596/1477/5/052037.
- [5]. Fitriani, S., Informatika, T., Teknik, F., & Sidoarjo, U. M. (2019). *Penerapan Data Mining Menggunakan Metode Klasifikasi Naïve Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur*. 3(2), 98–104.
- [6]. R. Firmansyah, "Web Klarifikasi Berita untuk Meminimalisir Penyebaran Berita Hoax," in *Jurnal Informatika*, 2017.
- [7]. M. Athaillah, Y. Azhar, and Y. Munarko, "Perbandingan Metode Klasifikasi Berita Hoaks Berbahasa Indonesia Berbasis Pembelajaran Mesin," *J. Repos.*, vol. 2, no. 5, p. 675, 2020, doi: 10.22219/repositor.v2i5.692.
- [8]. O. Nurdiawan, D. A. Kurnia, D. Solihudin, T. Hartati, and T. Suprpti, "Comparison of the K-Nearest Neighbor algorithm and the decision tree on moisture classification," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1088, no. 1, p. 012031, 2021, doi: 10.1088/1757-899x/1088/1/012031.
- [9]. Kominfo, "Jumlah Pengguna Internet 2017 Meningkat, Kominfo Terus Lakukan Percepatan Pembangunan Broadband," *Kementrian Komun. dan Infotmatika*, 2018.
- [10]. O. Nurdiawan, I. Ali, C. L. Rohmat, and A. R. Rinaldi, "InfoTekJar : Jurnal Nasional Informatika dan Sistem Kendali Sensor Tanah Sebagai Pemonitor Tingkat Kelembaban Media Tanam Padi," *Nas. Inform. dan Teknol. Jar.*, vol. 1, pp. 0–4, 2020.
- [11]. Gorunescu, F. (2011). *Data Mining Concepts, Models and Techniques*. Berlin: Springer