

KLASIFIKASI MINAT SISWA DALAM KEGIATAN EKSTRAKURIKULER DI SEKOLAH DASAR MENGGUNAKAN ALGORITMA C4.5 (STUDI KASUS: SD MUHAMMADIYAH BANGSRI)

Farikha Fatimah Tuzzahro^a, Fida Maisa Hana^b, Soma Setiawan PN^c, Andi Kurniawan^d

^{abcd}Universitas Muhammadiyah Kudus

E-mail: fidamaisa@umkudus.ac.id

Abstrak

Kegiatan ekstrakurikuler adalah kegiatan tambahan diluar mata pelajaran untuk membantu pengembangan peserta didik sesuai dengan kebutuhan, potensi, bakat dan minat mereka melalui kegiatan yang secara khusus diselenggarakan oleh pendidik atau tenaga kependidikan yang berkemampuan dan berkewenangan di sekolah. Penelitian ini untuk mengklasifikasikan minat siswa dalam mengikuti kegiatan ekstrakurikuler yang ada di SD Muhammadiyah Bangsri. Metode yang digunakan adalah deskriptif kuantitatif, sampel dalam penelitian ini adalah siswa kelas 4 sampai 6 di SD Muhammadiyah Bangsri. Teknik pengambilan sampel ini menggunakan teknik total sampling. Teknik pengumpulan data menggunakan observasi dan kuesioner. Tujuan dari penelitian ini adalah untuk mengklasifikasikan minat siswa dalam mengikuti kegiatan ekstrakurikuler dengan menggunakan algoritma C4.5 yang diuji menggunakan software pengolah data otomatis dan akurat menggunakan aplikasi Rapidminer 10.0.000. Hasil dari penelitian ini menunjukkan bahwa bidang ilmu data mining dengan algoritma C4.5 dapat diimplementasikan pada klasifikasi minat siswa dalam kegiatan ekstrakurikuler, setelah menguji data set dengan rapid miner dengan hasil akurasi yang baik. Kesimpulannya adalah Penerapan Algoritma C4.5 pada klasifikasi minat siswa dalam kegiatan ekstrakurikuler di sekolah dasar dapat menghasilkan accuracy yang mencapai 92.86% sampai 85.19%.

Kata Kunci: Klasifikasi, Data Mining, Algoritma C4.5, Rapidminer

Abstract

Extracurricular activities are additional activities outside of the subject matter to help develop students according to their needs, potential, talents and interests through activities specifically organized by educators or educational staff who are capable and authorized at school. This research is to classify students' interest in participating in extracurricular activities at SD Muhammadiyah Bangsri. The method used is descriptive quantitative, the sample in this study were students of grades 4 to 6 at SD Muhammadiyah Bangsri. This sampling technique uses total sampling technique. Data collection techniques using observation and questionnaires. The purpose of this study was to classify students' interest in participating in extracurricular activities using the C4.5 algorithm which was tested using automatic and accurate data processing software using the Rapidminer 10.0.000 application. The results of this study indicate that the field of data mining with the C4.5 algorithm can be implemented in the classification of students' interest in extracurricular activities, after testing the data set with a rapid miner with good accuracy results. The conclusion is that the application of the C4.5 Algorithm to the classification of students' interest in extracurricular activities in elementary schools can produce accuracy that reaches 92.86% to 85.19%.

Keywords: Classification, Data Mining, C4.5 Algorithm, Rapidminer

I. PENDAHULUAN

Kegiatan ekstrakurikuler adalah kegiatan tambahan diluar mata pelajaran untuk membantu pengembangan peserta didik sesuai dengan kebutuhan, potensi, bakat dan minat mereka melalui kegiatan yang secara khusus diselenggarakan oleh pendidik atau

tenaga kependidikan yang berkemampuan dan berkewenangan di sekolah. Ekstrakurikuler yang ada di SD Muhammadiyah Bangsri diantaranya Tapak Suci, English Club, Tilawah Qur'an, Angklung, Tahfidz Qur'an, Hizbul Wathan dan Drumband. Setelah dilakukannya

observasi di dapatkan beberapa hal, salah satunya, ekstrakurikuler dibagi menjadi 2 yaitu ekstrakurikuler wajib seperti Hizbul Wathan dan ekstrakurikuler pilihan seperti Tapak Suci, English Club, Tilawah Qur'an, Angklung, Tahfidz Qur'an dan Drumband. SD Muhammadiyah Bangsri mengharuskan peserta didiknya memilih minimal 1 untuk mengikuti ekstrakurikuler pilihan, dikarenakan sudah tercantum dalam administrasi keuangan di sekolah. Berbeda dengan sekolah dasar lain dimana ekstrakurikuler di sekolah tidaklah wajib dan siswa boleh memilih untuk tidak mengikutinya. Faktor yang menjadi acuan dalam memilih ekstrakurikuler adalah peserta didik memilih berdasarkan minat serta bakat yang peserta didik miliki dan ingin di asah lagi.

Hal yang melatar belakangi pembuatan tugas akhir ini adalah supaya pihak sekolah memiliki data atau klasifikasi minat peserta didik dalam mengikuti ekstrakurikuler secara otomatis dan akurat. Sehingga peneliti ingin meneliti tentang klasifikasi minat siswa dalam mengikuti kegiatan ekstrakurikuler untuk dapat mengetahui klasifikasi minat peserta didik dalam mengikuti kegiatan ekstrakurikuler di SD Muhammadiyah Bangsri secara otomatis dan akurat menggunakan software pengolah data otomatis dan akurat. Disini peneliti hanya meneliti ekstrakurikuler pilihan yang aktif yaitu : Tapak Suci, English Club, Tilawah Qur'an, Angklung, dan Tahfidz Qur'an kepada siswa kelas 4-6 di SD Muhammadiyah Bangsri.

Pada penelitian ini penulis melakukan penelitian dengan proses data mining menggunakan metode Algoritma C4.5. Menurut (Suntoro, 2019) data mining adalah proses untuk mendapatkan informasi yang berguna dari basis data yang besar dan perlu diekstraksi agar menjadi informasi baru dan dapat membantu dalam pengambilan keputusan, manfaat data mining adalah membantu tahapan perencanaan dan memberikan informasi tepat untuk membuat prediksi berdasarkan tren masa lalu dan kondisi saat ini. Algoritma C4.5 merupakan kelompok dari Algoritma Decision Tree. Menurut David Hartanto Kamagi & Seng Hansun (2014), bahwa algoritma C4.5 dapat

diimplementasikan untuk memprediksi tingkat kelulusan mahasiswa dengan empat kategori yaitu lulus cepat, lulus tepat, lulus terlambat dan drop out. Atribut yang paling berpengaruh dalam hasil prediksi adalah IPS semester enam. Dari metode tersebut penulis ingin mengklasifikasikannya kedalam suatu sistem agar lebih mudah proses klasifikasi tersebut. Dengan adanya suatu sistem ini maka proses tersebut akan berjalan secara otomatis ketika memasukkan data yang ada dan menghasilkan data yang akurat dan optimal. Dalam penelitian ini penulis menggunakan aplikasi RapidMiner untuk penghitungan secara otomatis dan akurat, RapidMiner sendiri adalah perangkat lunak open source yang juga merupakan solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi.

Dari penelitian ini, pihak sekolah dapat mengetahui sejauh mana minat ekstrakurikuler pada peserta didik dan sekolah dapat meningkatkan kegiatan ekstrakurikuler serta peserta didik dapat mengembangkan minatnya terhadap ekstrakurikuler sesuai keinginannya. Dan pihak sekolah juga dapat mengetahui minat peserta didik secara otomatis dan akurat.

Tujuan dari penelitian ini adalah untuk menyajikan atau mengklasifikasikan data minat siswa SD Muhammadiyah Bangsri yang mengikuti ekstrakurikuler menggunakan metode algoritma C4.5 secara otomatis menggunakan RapidMiner.

II. LANDASAN TEORI

A. Klasifikasi

Klasifikasi adalah suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan (Aprilla C dkk, 2013).

Proses klasifikasi didasarkan pada empat komponen (Gorunescu, 2011: 15):

1. Kelas

Variabel dependen berupa kategori yang merepresentasikan "label" yang terdapat

pada objek. Contohnya: risiko penyakit jantung, risiko kredit, dan jenis gempa

2. Predictor

Variabel independen yang direpresentasikan oleh karakteristik (atribut) data. Contohnya: merokok atau tidak, minum alkohol atau tidak, besar tekanan darah, jumlah tabungan, jumlah aset, jumlah gaji.

3. Training dataset

Satu set data yang berisi nilai dari kedua komponen di atas yang digunakan untuk menentukan kelas yang cocok berdasarkan predictor.

4. Testing dataset

Berisi data baru yang akan diklasifikasikan oleh model yang telah dibuat dan akurasi klasifikasi dievaluasi.

B. Minat

Minat adalah keinginan yang didorong oleh suatu keinginan setelah melihat, mengamati dan membandingkan serta mempertimbangkan dengan kebutuhan yang diinginkan (Widyastuti, 2008, www.pikiran-rakyat.com/aticle).

Minat sangat berpengaruh dalam belajar ekstrakurikuler karena minat akan dapat melahirkan gairah belajar. Untuk itu minat belajar merupakan hal yang penting dalam pelaksanaan pendidikan, sehingga belajar akan lebih mudah jika minat sudah terbina dalam diri peserta didik. (Almasawi, 2013).

Dalam hal ini faktor yang menjadi indikator minat siswa terhadap

ekstrakurikuler adalah, faktor dari dalam (Intrinsik) :

1. Perhatian

Definisi perhatian menurut Sumadi Suryabrata (2002: 14), kalau diambil dari intinya saja maka definisi perhatian ada dua macam yaitu:

a. Perhatian adalah pemusatan tenaga psikis tertuju kepada suatu objek

b. Perhatian adalah banyak sedikitnya kesadaran yang menyertai suatu aktivitas yang dilakukan.

Perhatian menurut Sumadi Suryabrata (2002: 14) ialah pemusatan tenaga psikis tertuju pada suatu objek. Kemudian menurut

Bimo Walgito (2010: 56) “Perhatian merupakan pemusatan atau konsentrasi dari seluruh aktivitas individu yang ditunjukkan kepada sesuatu atau sekelompok objek”.

2. Perasaan Senang

Perasaan senang terhadap sesuatu objek, baik orang atau benda dapat menimbulkan minat pada diri seseorang, orang merasa tertarik kemudian pada perasaannya timbul keinginan yang dikehendaki agar objek tersebut menjadi miliknya. Dengan demikian, maka individu tersebut dapat mempertahankan objek tersebut. (M Reynaldi Rafiq, 2022)

3. Aktivitas

Aktivitas merupakan kegiatan yang dilakukan secara berkelanjutan akan membentuk sebuah kebiasaan yang akhirnya yang akhirnya dapat menumbuhkan rasa senang dan tertarik (M Reynaldi Rafiq, 2022). Menurut Singgeh D. Gunarsa (1978: 97) “Aktivitas adalah gerakan – gerakan yang timbul menyertai adanya kebutuhan”.

C. Ekstrakurikuler

Dalam sebuah keputusan Dirjen Dikdasmen No. 226/C/Kep/1992, di mana dalam lampiran keputusan itu menjelaskan bahwa ekstrakurikuler adalah kegiatan di luar jam pelajaran biasa dan pada waktu libur sekolah yang dilakukan baik di sekolah maupun di luar sekolah. Tujuan ekstrakurikuler adalah untuk memperdalam dan memperluas pengetahuan siswa, mengenal hubungan antara berbagai pelajar, menyalurkan berbagai bakat dan minat, serta melengkapi upaya pembentukan dan pembinaan manusia seutuhnya.

D. Data Mining

Teknik yang relatif cepat dan mudah untuk menemukan pengetahuan, pola dan atau relasi antar data, secara otomatis. Dengan menggabungkan empat disiplin ilmu komputer seperti pada definisi di atas, pengetahuan bisa diketemukan dalam lima proses berurutan: seleksi, pra pemrosesan, transformasi, data mining, dan interpretasi/evaluasi (Suyanto, 2017).

Ada beberapa faktor yang menjadi alasan utama mengapa menggunakan data mining:

1. Banyaknya data yang terkumpul sehingga memerlukan waktu yang sangat lama dan tenaga ahli yang cukup banyak untuk menganalisisnya.

2. Komputer menjadi salah satu pilihan utama karena kemampuannya dalam kecepatan, ketepatan, tidak pernah lelah dan mudah dioperasikan.

3. Tekanan dari kompetisi bisnis yang terus menguat sehingga menjadikan informasi menjadi sangat penting dan harus segera dimilliki.

4. Mampu menemukan suatu pola yang tidak terpikirkan sama sekali.

E. Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang sangat *popular* yang digunakan oleh banyak peneliti di dunia, hal ini dijelaskan oleh Algoritmaa C4.5 merupakan pengembangan dari algoritma ID3 yang diciptakan oleh J. Rose Quinlan (Mardi, 2017).

Secara Umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar
2. Buat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada.

F. Confusion Matrix

Confusion matrix merupakan suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep *data mining* (Rosandy, 2016).

Melalui *Confusion Matrix*, dapat diketahui keakuratan, tingkat kesalahan, ketepatan dan nilai penarikan. Ada 2 nilai dalam kuantitas *Confusion Matriks* yaitu, akurasi dan laju *error*. Diketahui jumlah data yang diklasifikasi dengan benar, dapat diketahui akurasi hasil prediksi dan mengetahui jumlah data yang diklasifikasikan secara salah, jadi kita dapat mengetahui laju *error* dari prediksi yang dilakukan.

Tabel 2.2. *Confusion Matrix* (Sumber : (Pratiwi & Silvia, 2020))

	Observed	
	TRUE	FALSE
Predicted TRUE	TP	FP
Predicted FALSE	FN	TN

Keterangan :

TP : *True* Positif adalah jumlah data positif yang terklasifikasi dengan benar oleh sistem.

TN : *True* Negatif adalah jumlah data negatif yang terklasifikasi dengan benar oleh sistem.

FN : *False* Negatif adalah jumlah data negatif namun terklasifikasi salah oleh sistem.

FP : *False* Positif adalah jumlah data positif namun terklasifikasi salah oleh sistem.

1. *Accuracy* merupakan perbandingan kasus yang terklasifikasi benar dengan jumlah seluruh kasus.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \times 100\%$$

2. *Recall* menunjukkan berapa persen data kategori positif yang terklasifikasi dengan benar oleh sistem.

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

3. *Precision* mengevaluasi kemampuan sistem untuk menemukan peringkat yang paling relevan dan didefinisikan sebagai *presentase* dokumen yang di *retrieve* dan benar-benar relevan terhadap *query*.

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

4. *Error rate* merupakan kasus yang teridentifikasi salah dalam sejumlah data, sehingga dapat dilihat seberapa besar tingkat kesalahan pada sistem yang digunakan.

$$Error = \frac{FP}{TP} \times 100\%$$

G. Decision Tree

Decision tree adalah sebuah diagram alir yang berbentuk seperti struktur pohon yang mana setiap *internal node* menyatakan pengujian terhadap suatu *atribut*, setiap cabang menyatakan *output* dari pegujian tersebut dan *leaf node* menyatakan kelas-kelas atau distribusi kelas. *Node* yang paling atas disebut sebagai *root node* atau *node* akar.

Sebuah *root node* akan memiliki beberapa *edge* keluar tetapi tidak memiliki *edge* masuk, *internal node* akan memiliki satu *edge* masuk dan beberapa *edge* keluar, sedangkan *leaf node* hanya akan memiliki satu *edge* masuk tanpa memiliki *edge* keluar.

Decision tree digunakan untuk mengklasifikasikan suatu sampel data yang belum diketahui kelasnya ke dalam kelas-kelas yang sudah ada. Jalur pengujian data adalah pertama melalui *root node* dan terakhir adalah melalui *leaf node* yang akan menyimpulkan prediksi kelas bagi data tersebut. *Atribut* data harus berupa data kategori, bila *kontinu* maka *atribut* harus didiskretisasi terlebih dahulu.

H. Supervised Learning dan Unsupervised Learning

Teknik yang digunakan oleh *Supervised Learning* adalah metode klasifikasi di mana kumpulan data sepenuhnya diberikan label untuk mengklasifikasikan kelas yang tidak dikenal. Sedangkan teknik *Unsupervised Learning* sering disebut *cluster* dikarenakan tidak ada kebutuhan untuk pemberian label dalam kumpulan data dan hasilnya tidak mengidentifikasi contoh di kelas yang telah ditentukan (Thupae, Isong, Gasela, & AbuMahfouz, 2018).

I. Microsoft Excel

Microsoft Excel adalah aplikasi untuk mengolah data secara otomatis yang dapat berupa perhitungan dasar, rumus, pengolahan data dan *table*, pembuatan grafik dan manajemen data. *Microsoft excel* memiliki fungsi untuk membuat grafik persamaan matematika, membuat daftar nilai sekolah maupun universitas, melakukan perhitungan statistika, dan lain-lain (Sudarsana et al., 2019).

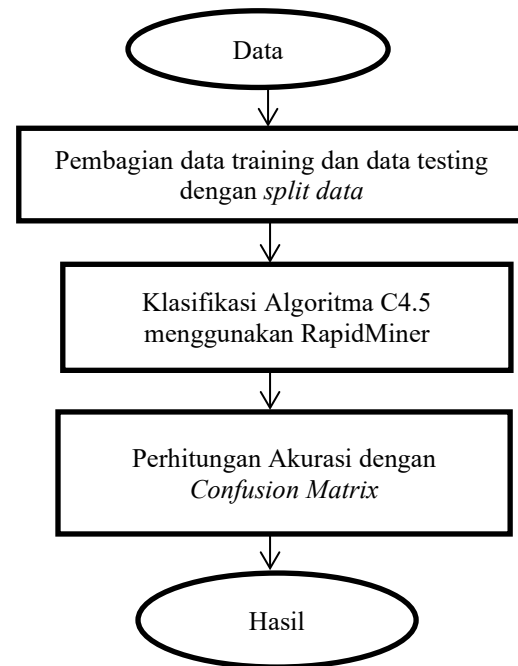
J. RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). *RapidMiner* adalah sebuah solusi untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. *RapidMiner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. *RapidMiner* memiliki kurang lebih 500 operator *data*

mining, termasuk *operator* untuk *input*, *output*, *data preprocessing* dan visualisasi. *RapidMiner* ditulis dengan menggunakan bahasa *java* sehingga dapat bekerja di semua sistem operasi.

K. Kerangka Berpikir

Dibawah ini adalah kerangka berpikir klasifikasi minat siswa dalam mengikuti kegiatan ekstrakurikuler menggunakan metode C4.5. Berikut pada Gambar 2.2 Kerangka Berpikir



Gambar 2.3 Kerangka Berpikir

Dari kerangka berpikir diatas menunjukkan bahwa data minat siswa dalam mengikuti kegiatan ekstrakurikuler yang merupakan sumber data utama dilakukan pembagian data dengan operator *split data* yang ada di *RapidMiner* tahap ini bertujuan untuk membagi data training dan data testing yaitu 0.7 dan 0.3,, selanjutnya melakukan klasifikasi dengan algoritma C4.5 menggunakan *software* pengolah data *RapidMiner* untuk penghitungan secara otomatis. Tahap terakhir adalah melakukan penghitungan akurasi menggunakan *confusion matrix*.

III. METODE PENELITIAN

A. Jenis Penelitian

Penelitian ini menggunakan metode penelitian deskriptif kuantitatif. Metode penelitian deskriptif kuantitatif adalah suatu metode yang bertujuan untuk membuat gambar atau deskriptif tentang suatu keadaan secara objektif yang menggunakan angka, mulai dari pengumpulan data, penafsiran terhadap data tersebut serta penampilan dan hasilnya (Arikunto, 2006).

Metode penelitian ini menggunakan metode angket. Adapun teknik pengambilan datanya menggunakan angket atau kuesioner. Menurut Sugiyono (2017:142) angket atau kuesioner merupakan teknik pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan atau pernyataan tertulis kepada responden untuk dijawab.

B. Populasi Penelitian

Populasi merupakan keseluruhan sumber data yang yang diperlukan dalam waktu penelitian (Saryono, 2011). Populasi target dalam penelitian ini adalah klasifikasi minat siswa di SD Muhammadiyah Bangsri yang dilakukan oleh peneliti kepada siswa SD Muhammadiyah Bangsri kelas 4-6. Dengan total populasinya adalah kelas 4 = 27 siswa, kelas 5 = 42 siswa dan kelas 6 = 23 siswa, jadi total keseluruhan siswa adalah 92 siswa.

C. Sampel Penelitian

Sampel adalah sebagian atau wakil populasi yang diteliti (Arikunto, 2010). Sampel adalah sebagian untuk diambil dari keseluruhan obyek yang diteliti dan dianggap mewakili seluruh populasi (Notoatmodjo, 2010). Teknik sampel dalam penelitian ini adalah total sampling. Sugiyono (2018) menjelaskan, teknik total sampling adalah teknik pengambilan sampel dimana seluruh anggota populasi dijadikan sampel. Alasan pemilihan sampel ini adalah di dasarkan pada jumlah populasi yang relatif kecil, yaitu kurang dari 100 orang. Sampel dalam penelitian ini adalah siswa SD Muhammadiyah Bangsri kelas 4-6 dengan total data yang digunakan adalah 92 siswa.

D. Detail Teknik Pemilihan Sampel

Teknik *sampling* dibagi menjadi dua kelompok yaitu *probability sampling* dan *non probability sampling*. Teknik pemilihan sampel dalam penelitian ini adalah menggunakan *non probability sampling*. Menurut Sugiyono (2017) *probability sampling* adalah teknik pengambilan sampel yang tidak memberikan peluang atau kesempatan yang sama bagi setiap unsur atau anggota populasi untuk dipilih menjadi sampel. Pada penelitian ini, peneliti menggunakan *total sampling*, menurut Sugiyono (2017) teknik *total sampling* adalah teknik pengambilan sampel dimana seluruh anggota populasi dijadikan sampel.

Pada penelitian ini menggunakan teknik *total sampling* dengan rumus: :

$$n = \frac{N}{1 + Ne^2}$$

$$n = \frac{92}{1 + 95(0,05)^2}$$

$$n = \frac{92}{1 + 92(0,025)}$$

$$= \frac{92}{1 + 2,3}$$

$$= \frac{92}{1 + 3,3}$$

$$= 27,8$$

Ket :

n : Jumlah sampel

N : Besar populasi

e : batas toleransi

E. Tempat penelitian

Penelitian akan dilakukan di SD Muhammadiyah Bangsri. SD Muhammadiyah Bangsri adalah sekolah dasar swasta yang dibawah oleh lembaga Muhammadiyah.

F. Instrumen Penelitian

Penelitian ini menggunakan 1 instrumen penelitian, yaitu:

1. Kuesioner

Menurut Sugiyono (2017:142) angket atau kuesioner merupakan teknik pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan atau pernyataan tertulis kepada responden untuk dijawab.

Instrumen yang digunakan penulis dalam mengambil kuesioner adalah dengan menggunakan kertas yang didalamnya terdapat beberapa pertanyaan yang harus di jawab.

Alasan menggunakan metode kuesioner karena dengan metode tersebut dapat menghemat waktu. Serta pengisiannya dilakukan langsung oleh siswa dengan menggunakan jenis kuesioner. Karena jawaban sudah tersedia, siswa memilih jawaban yang sudah ada di kuesioner tersebut. Variabel dalam penelitian ini adalah minat siswa kelas 4 – 6 SD Muhammadiyah Bangsri dalam kegiatan ekstrakurikuler.

G. Teknik Analisis Data

Analisis data menurut Sugiyono (2018:482) adalah proses mencari dan menyusun secara sistematis data yang diperoleh dari hasil wawancara, catatan lapangan dan dokumentasi, dengan cara mengorganisasikan data ke dalam kategori, menjabarkan ke dalam unit-unit, melakukan sintesa, menyusun ke dalam pola, memilih mana yang penting dan yang akan dipelajari, dan membuat kesimpulan sehingga mudah dipahami oleh diri sendiri maupun orang lain.

Pada penelitian ini, peneliti menggunakan teknik data mining dan klasifikasi menggunakan Algoritma C4.5 untuk teknik analisis datanya. Data mining sendiri adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (pramudiono, 2006). Berdasarkan fungsionalitasnya, tugas-tugas data mining bisa dikelompokkan kedalam delapan kelompok (Kotu et al, 2015): Classification, Regression, Clustering, Assosiation, Anomaly Detection, Forecasting, Text Mining amf Feature Selection. Dan metode data mining yang digunakan oleh penulis adalah Klasifikasi, Klasifikasi adalah suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah

didefinisikan. Teknik ini dapat memberikan klasifikasi pada data baru dengan memanipulasi data yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan (Aprilla C dkk, 2013). Dalam klasifikasi terdapat beberapa algoritma yang digunakan salah satunya adalah Algoritma C4.5 yang akan digunakan oleh penulis pada penelitian ini, algoritma C4.5 adalah Algoritma C4.5 merupakan algoritma yang sangat populer yang digunakan oleh banyak peneliti di dunia, hal ini dijelaskan oleh Algoritmaa C4.5 merupakan pengembangan dari algoritma ID3 yang diciptakan oleh J. Rose Quinlan (Mardi, 2017).

Alasan penulis menggunakan data mining klasifikasi dengan menggunakan algoritma C4.5 karena pada penelitian sebelumnya teknik yang penulis gunakan ini memiliki tingkat akurasi yang baik.

Metode penelitian berisi jenis penelitian, pendekatan, populasi, sampel teknik sampling, waktu dan tempat penelitian, instrumen penelitian, teknik pengambilan data dan analisis data.

IV. HASIL DAN PEMBAHASAN

A. Hasil Penelitian

Pada tahap ini merupakan hasil dari implementasi algoritma C4.5 dalam membuat pohon keputusan dan menghitung akurasi menggunakan aplikasi pengolah data RapidMiner studio 10.

Sebelum melakukan pengujian menggunakan aplikasi rapidminer. Penulis melakukan beberapa tahapan dalam pengolahan datanya, sesuai dengan kerangka berpikir pada bab 2. Yaitu Data hasil penelitian kepada siswa kelas 4 sampai 6 SD Muhammadiyah Bangsri yang dimasukkan terlebih dahulu kedalam microsoft excel, selanjutnya membagi data dengan split data kemudian melakukan klasifikasi algoritma C4.5 menggunakan aplikasi RapidMiner studio 10, pengolahan data ini menggunakan algoritma C4.5 yang secara otomatis akan terhitung pengujiannya menggunakan RapidMiner dan tahap terakhir adalah perhitungan akurasi secara otomatis

menggunakan RapidMiner yang menghasilkan olah data di RapidMiner studio 10.

B. Data

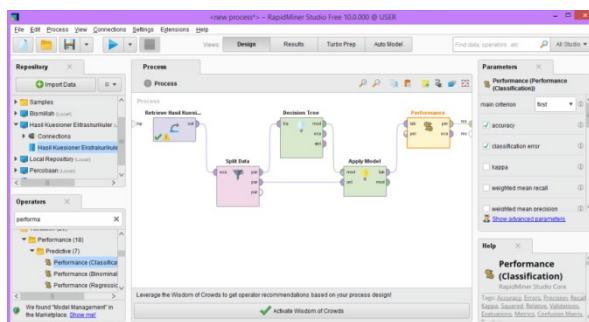
Pada tahap ini penulis memindahkan data hasil kuesioner yang telah di bagikan ke siswa kelas 4 sampai 6 SD Muhammadiyah Bangsri ke dalam Microsoft Excel. Kuesioner akan terlampir di halaman lampiran. Setelah semua hasil penelitian dimasukkan kedalam Microsoft excel, selanjutnya adalah pengolahan data menggunakan RapidMiner Studio 10 untuk mendapatkan hasil pohon keputusan dan akurasi yang otomatis dan akurat.

C. Pembagian data training dan data testing dengan split data

Pada tahap ini data dibagi menjadi 2 yaitu data training dan data testing dengan operator split data yang ada di software pengolah data otomatis RapidMiner, yang akan dijelaskan dan diaplikasikan pada tahap selanjutnya yaitu klasifikasi algoritma C4.5 dengan menggunakan software pengolah data otomatis RapidMiner.

D. Klasifikasi Algoritma C4.5 menggunakan RapidMiner

Tahap selanjutnya adalah Klasifikasi Algoritma C4.5 menggunakan RapidMiner atau melakukan pengujian menggunakan aplikasi RapidMiner. Dalam tahap ini penulis melakukan pengolahan data hasil penelitian menggunakan kuesioner kedalam software pengolah data RapidMiner untuk penghitungan dan pengujian secara otomatis. Berikut adalah langkah-langkah penggunaan RapidMiner studio 10.

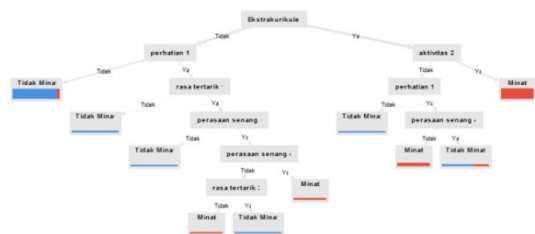


Gambar 4.1. RapidMiner

Drop data penelitian ke proses kemudian cari operator split data, yang berguna untuk untuk membagi data menjadi data training dan data testing, lalu drop ke dalam process. Setelah itu sambungkan output retrieve_hasil_kuesioner ke input split data. 7.

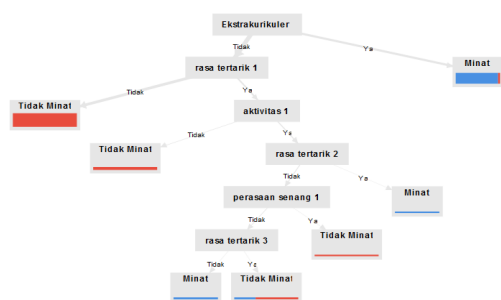
Setelah menghubungkan output retrieve ke input split data, ubah parameters atau klik edit numeric yang ada pada split data partitions dengan rasion 0.7 untuk data training dan 0.3 untuk data testing. Langkah selanjutnya pada operators drop decision tree ke process. Decision tree berguna untuk menghitung pohon keputusan menggunakan algoritma C4.5. Kemudian drop apply model ke proses Apply model berguna untuk mengaplikasikan hasil tree pada training ke data testing. Tahap terakhir adalah performance. Drop performance ke dalam process. Performance digunakan untuk mengevaluasi kinerja model yang memberikan daftar nilai kriteria kinerja secara otomatis sesuai dengan tugas yang diberikan. Disini untuk menguji accuracy. Kemudian klik run dan akan menghasilkan pohon keputusan dan akurasi sebagai berikut:

1. Pohon Keputusan Ekstra Tapak Suci



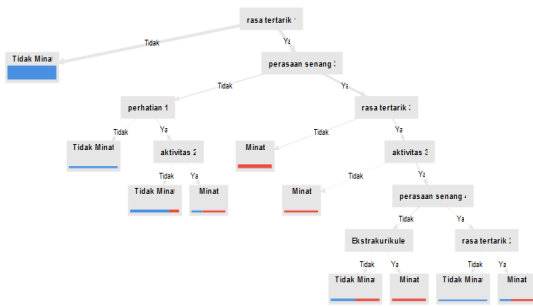
Gambar 4.2. Pohon Keputusan Ekstra Tapak Suci

2. Pohon Keputusan Ekstra English Club

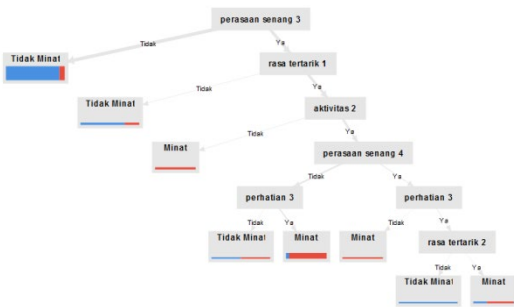


Gambar 4.3. Pohon Keputusan Ekstra English Club

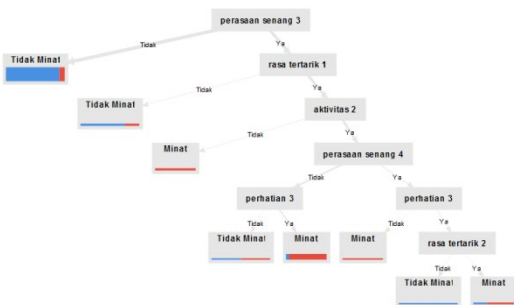
3. Pohon Keputusan Ekstra Tilawah Qur'an



Gambar 4.4. Pohon Keputusan Ekstra Tilawah Qur'an
4. Pohon Keputusan Ekstra Tahfidzh Qur'an



Gambar 4.5. Pohon Keputusan Ekstra Tahfidzh Qur'an
5. Pohon Keputusan Ekstra Angklung



Gambar 4.6. 1. Pohon Keputusan Ekstra Angklung

Setelah mendapatkan hasil pohon keputusan, berikut adalah hasil akurasi seperti yang terdapat di bab 2 yaitu tahap terakhir adalah perhitungan akurasi menggunakan confusion matriks, perhitungannya sudah dihitung secara otomatis menggunakan aplikasi RapidMiner studio 10.

E. Perhitungan akurasi dengan Confusion Matrix

1. Akurasi Tapak Suci

accuracy: 92.86%			
	true Tidak Minat	true Minat	class precision
pred Tidak Minat	13	2	86.67%
pred Minat	0	13	100.00%
class recall	100.00%	86.67%	

Gambar 4.7. Akurasi Ekstra Tapak Suci

Berdasarkan hasil pengujian diatas di dapatkan bahwa akurasi eskrtrakurikuler Tapak Suci adalah 92.86%, Recall 100% dan Precision 86.67%.

2. Akurasi English Club

accuracy: 85.19%			
	true Minat	true Tidak Minat	class precision
pred Minat	10	3	76.92%
pred Tidak Minat	1	13	92.86%
class recall	90.91%	81.25%	

Gambar 4.8. Akurasi Ekstra English Club

Berdasarkan hasil pengujian diatas di dapatkan bahwa akurasi eskrtrakurikuler English Club adalah 85.19%, Recall 90.91% dan Precision 76.92%.

3. Akurasi Tilawah Qur'an

accuracy: 85.71%			
	true Tidak Minat	true Minat	class precision
pred Tidak Minat	17	2	89.47%
pred Minat	2	7	77.78%
class recall	89.47%	77.78%	

Gambar 4.9. Akurasi Ekstra Tilawah Qur'an

Berdasarkan hasil pengujian diatas di dapatkan bahwa akurasi eskrtrakurikuler Tilawah Qur'an adalah 85.71%, Recall 89.47% dan Precision 89.47%.

4. Akurasi Tahfidzh Qur'an

accuracy: 89.29%			
	true Tidak Minat	true Minat	class precision
pred Tidak Minat	17	3	85.00%
pred Minat	0	8	100.00%
class recall	100.00%	72.73%	

Gambar 4.10. Akurasi Ekstra Tahfidzh Qur'an

Berdasarkan hasil pengujian diatas di dapatkan bahwa akurasi eskrtrakurikuler Tahfidzh Qur'an adalah 89.29%, Recall 100.00% dan Precision 85.00%.

5. Akurasi Angklung

	true Tidak Minat	true Minat	class precision
pred Tidak Minat	19	3	86.36%
pred Minat	0	6	100.00%
class recall	100.00%	66.67%	

Gambar 4.11. Akurasi Ekstra Angklung

Berdasarkan hasil pengujian diatas di dapatkan bahwa akurasi ekstrakurikuler Angklung adalah 89.29%, Recall 100.00%, dan Precision 86.36%.

Dari hasil accuracy ekstrakurikuler yang ada di SD Muhammadiyah Bangsri dengan hasil accuracy 80%-90% adalah baik.

4.2 Pembahasan

Penelitian ini adalah untuk mengklasifikasikan minat siswa dalam mengikuti kegiatan ekstrakurikuler yang ada di SD Muhammadiyah Bangsri. Pengujian ini dilakukan menggunakan algoritma C4.5 yang pengujian atau perhitungannya menggunakan aplikasi pengolah data otomatis yaitu RapidMiner studio 10 untuk mendapatkan hasil yang otomatis dan akurat. Berdasarkan hasil pengujian menggunakan aplikasi RapidMiner dalam penelitian ini, hasilnya menunjukkan bahwa penerapan algoritma C4.5 dalam mengklasifikasikan minat siswa adalah baik atau bisa diterapkan dengan hasil akurasi yaitu 80% sampai dengan 90%. Dengan hasil akurasi sebagai berikut :

Tabel 5.1. Akurasi Ekstrakurikuler

Ekstrakurikuler	Accuracy
Tapak Suci	92.86%
English Club	85.19%
Tilawah Qur'an	85.71%
Tahfidzh Qur'an	89.29%
Angklung	89.29%

Dari data tersebut terlihat masing-masing ekstrakurikuler memiliki nilai akurasi yang tidak sama. Akurasi yang paling tinggi didapatkan oleh ekstrakurikuler tapak suci yaitu sebesar 92.86%. dan akurasi yang paling rendah adalah English club dengan akurasi sebesar 85.19%.

Hasil penelitian menunjukkan bahwa algoritma C4.5 dapat diterapkan dalam

klasifikasi minat siswa dalam mengikuti kegiatan ekstrakurikuler dengan bantuan olah data otomatis RapidMiner dan dapat menampilkan hasil klasifikasi berupa pohon keputusan dan akurasi yang baik, berdasarkan dari table milik Rino Gupitho, didapati bahwa semua ekstrakurikuler dapat dikategorikan dengan klasifikasi yang baik.

Tabel 5.2. Klasifikasi nilai (sumber: Rino Gupitho, 2018)

No	Nilai	Klasifikasi
1	0-55	Rendah
2	56-75	Sedang
3	76-100	Tinggi

Dari table di atas, dapat disimpulkan bahwa klasifikasi nilai untuk klasifikasi minat siswa dalam kegiatan ekstrakurikuler ini diklasifikasi tinggi dengan nilai 76-100 dengan dilakukan pengujian menggunakan aplikasi olah data otomatis RapidMiner studio 10.

V. KESIMPULAN

5.1 Kesimpulan

Berdasarkan uji coba dan evaluasi yang telah dilakukan maka dapat diambil kesimpulan sebagai berikut:

1. Diketahui bahwa bidang ilmu data mining dengan algoritma C4.5 dapat diimplmentasikan pada klasifikasi minat siswa dalam kegiatan ekstrakurikuler, setelah menguji data set dengan rapid miner, maka dihasilkan akurasi sebagai nilai ketentuan seberapa besar keakuratan aplikasi dalam menentukan baik buruknya klasifikasi minat siswa dalam kegiatan ekstrakurikuler.
2. Penerapan Algoritma C4.5 pada klasifikasi minat siswa dalam kegiatan ekstrakurikuler di sekolah dasar dapat menghasilkan accuracy mencapai 92.86% untuk Ekstrakurikuler tapak suci, 85.19% untuk ekstrakurikuler English club, 85.71% untuk ekstrakurikuler tilawah qur'an, 89.29% untuk ekstrakurikuler tahfidzh qur'an dan ekstrakurikuler

angklung dengan menggunakan aplikasi RapidMiner.

5.2 Saran

Dalam penelitian ini, penulis menyadari masih banyak kekurangan-kekurangan pada skripsi ini. Oleh karena itu, penulis memberikan saran kepada pembaca maupun penelitian selanjutnya untuk:

1. Menggunakan algoritma lainnya, seperti algoritma K-nearest Neighbor (K-NN), Support Vector Machine (SVM), Naïve Bayes, dll. Untuk membandingkan tingkat akurasi hasil yang diperoleh.
2. Sebaiknya jumlah data kuesioner ditambah, sehingga dapat diperoleh hasil akurasi fungsi algoritma yang lebih baik.

DAFTAR PUSTAKA

- Aminullah, M. (n.d.). 2021. Perbandingan Performa Klasifikasi Machine Learning dengan Teknik Resampling pada Data Set Tidak Seimbang.
- Andika P, F. (2007). Penambahan aturan asosiasi pada dataset nilai akademik mahasiswa dengan algoritma. 1–137.
- Awliya Wanhari. (2018). Perbandingan Algoritma C4. 5 dan Naive Bayes Untuk Klasifikasi Mustahik. *Repository.Uinjkt.Ac.Id*. http://repository.uinjkt.ac.id/dspace/handle/123456789/53608%0Ahttp://repository.uinjkt.ac.id/dspace/bitstream/123456789/53608/1/AWLIYA_WANHARI-FST.pdf
- Bob Gary Gerardo, S. E. (2019). TEKNIK DATA MINING DALAM PENILAIAN PENGAJARAN GURU BERDASARKAN INDEKS KEPUASAN SISWA.
- Cahyono, N. D. (2017). MINAT SISWA DALAM MENGIKUTI KEGIATAN EKSTRAKURIKULER OLAHRAGA DI SEKOLAH MENENGAH ATAS NEGERI 2 PLAYEN.
- Change, G., Cimino, M., York, N., Alifah, U., Mayssara A. Abo Hassanin Supervised, A., Chinatown, Y., Staff, C., & Change, G. (2021). IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA C4.5 PADA KLASIFIKASI PENJUALAN HIJAB. *Paper Knowledge . Toward a Media History of Documents*, 3(2), 6.
- David Hartanto Kamagi, S. H. (2014). Implementasi Data Mining dengan Algoritma C4.5 untuk Memprediksi Tingkat Kelulusan.
- Di, E., & Mentaya, S. (2022). *K-Nearest Neighbor Ekstrakurikuler Di Smkn 1 Mentaya Hilir Selatan Menggunakan Metode K-Nearest Neighbor*.
- Gupitha. (2018). *SISWA SMK TERPADU LAMPANG SUBANG Pendahuluan Kajian Teori Penelitian Terkait Ekstrakurikuler*. IV(2), 31–37.
- Husna, F. (2021). IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA C4.5.
- Ibrahim, I., Gani, H., Lamusu, R., & Humolungo, Y. (2022). Perbandingan Algoritma Naïve Bayes Dan C4.5 Untuk Klasifikasi Bantuan Rumah Sehat. *Jurnal Ilmu Komputer (JUİK)*, 2(1), 72. <https://doi.org/10.31314/juik.v2i1.1477>
- Kustanti, E., Ba'diah, U., Sinawati, & Muhammad, F. (2018). Klasifikasi Siswa berdasarkan Nilai pada Bidang Ekstrakurikuler Menggunakan Metode K-Nearest Neighbor. *Journal of Big Data Analytic and Artificial Intelligence*, 4(STMIK PPKIA Tarakanita Rahmawati), 1–8.
- Luluk Elvitaria, M. H. (2017). Smk Analisis Kesehatan Abdurrah Menggunakan Algoritma. *(Jurnal Teknologi Dan Sistem Informasi Univrab)*, 2(2), 220–233.
- Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Edik Informatika*, 2(2), 213–219. <https://doi.org/10.22202/ei.2016.v2i2.1465>

- Mochamad Rizki Ilham, P. (2016). Implementasi Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Kepuasan Pelanggan Taksi Kosti. *Simplementasi Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Kepuasan Pelanggan Taksi Kosti*, Vol. 4, No(5), 11.
- Mugi Raharjo, R. J. (2019). Implementasi Metode Decision Tree Klasifikasi Data Mining untuk Prediksi Peminatan Jurusan Robotika oleh Mahasiswa.
- Nadya Nurul Fauziah, L. N. (2022). Pelaksanaan Kegiatan Ekstrakurikuler Gamelan di Sekolah Dasar
- Putri, R. E. (2021). Implementasi data mining untuk prediksi efektivitas pada mesin injection menggunakan algoritma c4. 5 studi kasus: pt. tridaya artaguna santara. *Skripsi*, 5. https://repository.uinjkt.ac.id/dspace/handle/123456789/57846%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/57846/1/RIZKY_EVITA_PUTRI-FST.pdf
- Rafiq, M. R. (2022). *Klasifikasi Minat Siswa Terhadap Kegiatan Ekstrakurikuler Di SMK N 1 Mentaya Hilir Selatan Menggunakan Metode K-NEAREST NEIGHBOR*.
- Rivaldi, M. A., Nurudin, A. A., & Nugraheni, W. (2022). Survei Minat Siswa Dalam Mengikuti Kegiatan Ekstrakurikuler Futsal SMA Negeri Se-Kota Sukabumi. *Jurnal Educatio FKIP UNMA*, 8(2), 498–504. <https://doi.org/10.31949/educatio.v8i2.1963>
- Ryanwar. (2020). Penerapan Metode Algoritma C4.5 Untuk Memprediksi Loyalitas Karyawan Pada Pt.Xyz Berbasis Web. *Fakultas Sains Dan Teknologi Universitas Buddhi Dharma*, 103.
- Saifullah. (2019). *Deteksi Kelayakan Fisik Air Untuk Konsumsi Menggunakan Naive Bayes Classifier*.
- Siti Sundari, A. P. (2022). Penerapan Data Mining Klasifikasi Terhadap Kepuasan Pelanggan Toko Bombay Textile.
- Sugiyono, P. (2018). Metodologi Penelitian.
- Sukmawati, S., Februariyanti, H., Jananto, A., Stikubank Semarang Jl Tri Lomba Juang No, U., & Kode Pos, S. (2022). I N F O R M a T I K a Perbandingan Algoritma C 4.5 Dan Algoritma Naive Bayes Untuk Klasifikasi Pekerja Migran Indonesia. *Jurnal Informatika, Manajemen Dan Komputer*, 14(1), 7–16.
- Telaumbanua, D., & Kurniawati, I. (2022). Penerapan Algoritma C4.5 Untuk Klasifikasi Kepuasan Pelanggan Pada Jasa Layanan Pengiriman. *JoMMiT: Jurnal Multi Media Dan IT*, 6(1). <https://doi.org/10.46961/jommit.v6i1.524>
- Wulandari, R. Y. (2018). Analisa Data Mining Dengan Metode Klasifikasi Untuk Produk Cacat Pada Pt. Shuangying International Indonesia. *Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa*, 14(5), 2–5. <https://ecampus.pelitabangsa.ac.id/pb/AmbilLampiran?ref=22722&jurusan=&jenis=Item&usingId=false&download=false&clazz=ais.database.model.file.LampiranLain>
- Zahedi, B., Nahid-Mobarakeh, B., Pierfederici, S., & Norum, L. E. (2016). A robust active stabilization technique for dc microgrids with tightly controlled loads. *Proceedings - 2016 IEEE International Power Electronics and Motion Control Conference, PEMC 2016*, VI(1), 254–260. <https://doi.org/10.1109/EPEPEMC.2016.7752007>