

Sistem Pendukung Keputusan Penentuan Jurusan di SMKN 1 Pungging Menggunakan *Gradient Boosting Tree*

Elok Nur Fauziah¹, Salamun Rohman Nudin²

^{1,2}Teknik Informatika, Jurusan Teknik Informatika, Fakultas Teknik
Universitas Negeri Surabaya

¹elok.17051204051@mhs.unesa.ac.id

²salamunrohman@unesa.ac.id

Abstrak— Saat ini belum ada sistem yang membantu calon siswa yang akan masuk ke SMKN 1 Pungging untuk memilih jurusan yang sesuai dengan minat dan bakat siswa. Sehingga perlu dilakukan analisis terhadap kesesuaian minat dan bakat siswa dengan jurusan yang akan dipilih. Analisis ini akan membentuk sebuah Sistem Pendukung Keputusan (SPK). Sistem Pendukung Keputusan merupakan suatu alat yang dapat membantu proses pengambilan keputusan dengan menggunakan berbagai metode. Metode yang digunakan dalam penelitian ini adalah *Gradient Boosting Tree* (GBT). *Gradient Boosting Tree* merupakan teknik *machine learning* untuk masalah klasifikasi yang menghasilkan prediksi dan banyak digunakan karena efisiensi, akurasi, dan interpretabilitasnya. *Gradient Boosting Tree* merupakan algoritma yang sesuai dengan permasalahan penentuan jurusan karena *Gradient Boosting Tree* akan menghasilkan prediksi dari kriteria yang diberikan dalam bentuk *decision tree* yang akan menghasilkan keputusan berupa jurusan yang sesuai. Tujuan dari penelitian ini adalah untuk mengembangkan sistem pendukung keputusan penentuan jurusan untuk calon siswa di SMKN 1 Pungging dengan menggunakan *Gradient Boosting Tree*. Hasil dari penelitian ini adalah sistem pendukung keputusan penentuan jurusan untuk calon siswa di SMKN 1 Pungging dengan akurasi 100% untuk *training data* dan 96.34% untuk *testing data*. Adanya sistem ini diharapkan dapat membantu calon siswa dalam menentukan keputusan pemilihan jurusan di SMKN 1 Pungging.

Kata Kunci— Pemilihan Jurusan, Sistem Pendukung Keputusan, *Gradient Boosting Tree*, Python

I. PENDAHULUAN

SMK Negeri 1 Pungging merupakan sekolah menengah kejuruan yang memiliki 8 jurusan diantaranya Teknik Kendaraan Ringan, Teknik Pengelasan, Teknik Pemesinan, Multimedia, Teknik Komputer dan Jaringan, Teknik Instalasi Tenaga Listrik, Teknik Program Pertelevisian serta Teknik Ototronik [1]. Pemilihan jurusan merupakan hal yang penting bagi siswa karena pilihan ini dapat mempengaruhi karir masa depan siswa, kehidupan sosial di masa depan, dan kepuasan pribadi siswa [2]. Berdasarkan penelitian yang dilakukan oleh Youthmanual, 92% siswa SMA/SMK sederajat bingung dan tidak tahu akan menjadi apa kedepannya [3]. Dan berdasarkan penelitian yang dilakukan oleh Urzu Rahmad (2018), dari 313 siswa sebanyak 213 siswa atau sebesar 64,4% dari keseluruhan responden dalam pemilihan jurusan didasari oleh faktor

lingkungan alih-alih sesuai dengan minat dan bakat mereka [4].

Dengan adanya permasalahan siswa yang merasa tidak sesuai dengan jurusan mereka setelah diterima dan kesulitan dalam mengikuti materi pembelajaran, maka perlu dilakukan analisis terhadap kesesuaian minat dan bakat siswa dengan jurusan yang akan dipilih sebelum memasuki Sekolah Menengah Kejuruan (SMK). Sehingga ketika siswa diterima pada jurusan yang dipilih, siswa akan mudah dalam menerima pembelajaran atau menyesuaikan diri. Analisis ini akan membentuk sebuah Sistem Pendukung Keputusan (SPK) dengan proses yang kompleks, karena mencakup banyak faktor atau kriteria yang akan mempengaruhi *lifecycle* atau siklus hidup dari sistem [5].

Sistem Pendukung Keputusan (SPK) kesesuaian minat dan bakat siswa dengan jurusan dapat dipengaruhi oleh beberapa faktor diantaranya yaitu penilaian secara akademis siswa. Maka untuk mempermudah analisis ini, penulis membuat sistem yang dapat membantu calon siswa untuk memilih jurusan yang sesuai terutama jurusan di SMK Negeri 1 Pungging.

Algoritma yang digunakan adalah *Gradient Boosting Tree* (GBT). *Gradient Boosting Tree* merupakan algoritma yang sesuai dengan permasalahan penentuan jurusan karena *Gradient Boosting Tree* akan menghasilkan prediksi dari kriteria yang diberikan dalam bentuk *decision tree* yang akan menghasilkan keputusan berupa jurusan yang sesuai. Berdasarkan hal tersebut, maka penulis mengembangkan penelitian yang berjudul “Sistem Pendukung Keputusan Penentuan Jurusan di SMKN 1 Pungging Menggunakan *Gradient Boosting Tree*”. Adanya sistem ini diharapkan dapat membantu calon siswa dalam menentukan keputusan pemilihan jurusan di SMKN 1 Pungging.

A. Sistem Pendukung Keputusan

Dalam sebuah proses pengambilan keputusan, dibutuhkan sebuah alat yang dapat membantu dalam proses pengambilan keputusan tersebut agar keputusan yang diambil adalah keputusan yang benar atau dengan pertimbangan yang baik atau berimbang. Menurut V. Amalia dan S. Hamidani, Sistem Pendukung Keputusan (SPK) adalah sebuah sistem yang dapat memberikan fungsi manajemen data sesuai dengan model tertentu, sehingga memungkinkan pengguna sistem untuk mendapatkan solusi

terbaik [6]. Dengan kata lain, SPK bukanlah alat pengambilan keputusan, tetapi alat untuk mendukung pengambilan sebuah keputusan.

Dalam perkembangan ilmu komputer dan teknologi, Sistem Pendukung Keputusan merupakan suatu alat yang dapat membantu proses pengambilan keputusan dengan menggunakan berbagai metode yang diteliti oleh para ahli [7]. Setiap metode dalam SPK memiliki perhitungan yang berbeda dalam proses pengambilan keputusannya. Dalam penelitian ini metode yang digunakan adalah *Gradient Boosting Tree*.

B. *Gradient Boosting Tree*

Gradient Boosting Tree (GBT) merupakan teknik *machine learning* untuk masalah klasifikasi yang menghasilkan prediksi, biasanya berupa pohon keputusan [8]. Algoritma *Gradient Boosting Tree* diperoleh dengan memasukkan kriteria *Differentiable Loss Function* $L(y_i, F(x))$ [9]. *Gradient Boosting Tree* adalah algoritma *machine learning* yang sangat populer dan telah terbukti berhasil di banyak domain dan merupakan salah satu metode terkemuka untuk memenangkan kompetisi Kaggle [10].

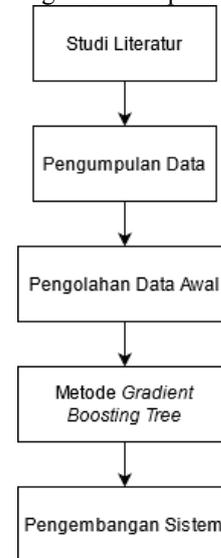
Ada beberapa varian algoritma *boosting* yang berfokus pada masalah klasifikasi. Sepanjang tahun 1990-an banyak pendekatan yang dikembangkan dan yang paling sukses adalah algoritma AdaBoost yang dikembangkan oleh Freund dan Schapire pada tahun 1999. Pada tahun 2000, Friedman menghubungkan AdaBoost dengan konsep statistik (seperti *loss function* dan *additive modeling*), yang memungkinkan untuk menggeneralisasi *boosting framework* ke masalah regresi dan *multiple loss function*. Hal ini mengarah pada model *Gradient Boosting Tree* yang kita gunakan saat ini dan yang dianggap sebagai implementasi paling modern yang pernah dibangun [10]. *Gradient Boosting Tree* banyak digunakan karena efisiensi, akurasi, dan interpretabilitasnya. *Gradient Boosting Tree* mencapai kinerja mutakhir dalam banyak tugas *machine learning*, seperti klasifikasi kelas jamak, prediksi klik, dan *learning* menentukan peringkat [11].

Teknik *boosting* memberi bobot pada setiap observasi dan mengubah bobot setelah training pada *classifier*. Bobot untuk observasi *classifier* yang salah bertambah dan bobot untuk observasi *classifier* yang benar dikurangi. Pengamatan yang bobotnya telah dimodifikasi digunakan untuk melatih pengklasifikasi berikutnya. Akhirnya, pengklasifikasi yang berbeda digabungkan menjadi pengklasifikasi yang lebih akurat. Friedman mengembangkan *Gradient Boosting* (GB). Ini adalah metode langkah demi langkah dengan berfokus pada pengurangan gradien dari fungsi kerugian (*loss function*) di model sebelumnya [12]. Fungsi kerugian dapat diartikan sebagai tingkat kesalahan oleh model. Secara umum, semakin besar fungsi kerugian, semakin besar kemungkinan model akan meleset. Karena tujuannya adalah untuk meminimalkan fungsi kerugian dan tingkat kesalahan,

pendekatan terbaik adalah menurunkan fungsi kerugian ke arah gradien.

II. METODE PENELITIAN

Konsep atau tahapan dari metode penelitian untuk membangun Sistem Pendukung Keputusan Penentuan Jurusan di SMKN 1 Pungging menggunakan *Gradient Boosting Tree* akan digambarkan pada Gbr. 1.



Gbr. 1 Tahapan Penelitian

Dalam penelitian ini terdapat beberapa tahapan yang dilakukan agar penelitian dapat berjalan dengan baik seperti pada Gbr. 1, penjelasan dari masing-masing tahapan akan dijabarkan sebagai berikut:

A. *Studi Literatur*

Tahap studi literatur adalah tahap pertama yang dilakukan penulis dalam melakukan penelitian ini. Studi literatur dilakukan dengan mempelajari penelitian sebelumnya mengenai permasalahan yang diangkat dalam penelitian yaitu mengenai Sistem Pendukung Keputusan pemilihan jurusan serta mengenai metode yang akan diterapkan dalam penelitian ini yaitu *Gradient Boosting Tree*.

B. *Tahap Pengumpulan Data*

Tahapan kedua adalah pengumpulan data. Data yang digunakan dalam penelitian ini diperoleh dari SMK Negeri 1 Pungging yang berada dibawah naungan Pemerintahan Kabupaten Mojokerto, Jawa Timur. Pada tahap ini data yang diperoleh adalah data persyaratan minimum untuk memasuki masing-masing jurusan di SMK Negeri 1 Pungging pada tahun 2020, selain itu juga diperoleh data siswa yang diterima di SMK Negeri 1 Pungging pada tahun 2020. Adapun untuk tahap pengumpulan data dilakukan dengan menghubungi Kepala Sekolah SMK Negeri 1 Pungging, untuk mendapatkan data-data serta informasi yang berhubungan dengan masing-masing jurusan yang ada di SMK Negeri 1 Pungging.

Data yang diperoleh berisi informasi nama siswa dan nilai siswa yang diwakili dalam data yang ditetapkan oleh 4 atribut yang merupakan nilai mata pelajaran siswa yang mempengaruhi diterimanya pada satu jurusan di SMKN 1 Pungging. 4 atribut tersebut berupa data numeric, sedangkan target atau jurusan diterimanya siswa berupa data object. Data siswa yang diterima di SMKN 1 Pungging memiliki 8 kelas yang merupakan jurusan yaitu, Teknik Kendaraan Ringan, Teknik Pengelasan, Teknik Pemesinan, Multimedia, Teknik Komputer dan Jaringan, Teknik Instalasi Tenaga Listrik, Teknik Program Pertelevisionan serta Teknik Ototronik dengan jumlah sampel data untuk masing-masing kelas yaitu 34 sampel.

TABEL I
ATRIBUT DATASET

Atribut	Deskripsi	Tipe Data
Math	Nilai Matematika siswa yang diterima di jurusan tertentu	Numeric
Eng	Nilai Bahasa Inggris siswa yang diterima di jurusan tertentu	Numeric
Ind	Nilai Bahasa Indonesia siswa yang diterima di jurusan tertentu	Numeric
IPA	Nilai Ilmu Pengetahuan Alam siswa yang diterima di jurusan tertentu	Numeric

C. Pengolahan Data Awal

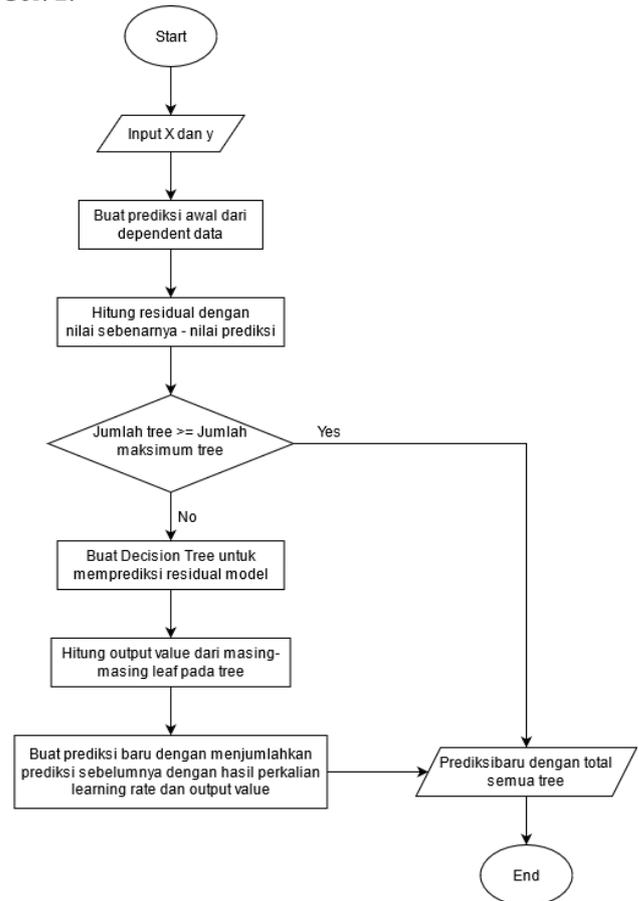
Pengolahan data awal atau *pre-processing data* merupakan langkah lanjutan atau tindak lanjut dari proses pengumpulan data. Data mentah yang diperoleh dari proses pengumpulan data terdiri dari banyak kolom diantaranya kolom Nomor, Nama Siswa, Nilai Matematika, Nilai Bahasa Inggris, Nilai Bahasa Indonesia, Nilai Bahasa Inggris, Nilai IPA, dan Kompetensi Keahlian, untuk itu perlu dilakukan pemilihan data yang dipergunakan dalam penelitian ini yaitu Nilai Matematika, Nilai Bahasa Inggris, Nilai Bahasa Indonesia, Nilai Bahasa Inggris, Nilai IPA, dan Kompetensi Keahlian. Data Nilai Matematika, Nilai Bahasa Inggris, Nilai Bahasa Indonesia, Nilai Bahasa Inggris, dan Nilai IPA adalah data yang berisi nilai yang menentukan ke Kompetensi Keahlian apa siswa akan masuk, sedangkan data Kompetensi Keahlian adalah data jurusan tempat siswa dengan nilai tertentu diterima. Pembersihan data dilakukan terhadap data-data ganda, *null*, dan data yang tidak memiliki tipe yang tidak sesuai.

Pengolahan data awal dilakukan dengan melakukan tranformasi tipe data keseluruhan menjadi numerik. Karena data dependen dari *dataset* yang diperoleh memiliki tipe data *object* yang harus diubah terlebih dahulu ke bentuk numerik agar dapat diproses dengan algoritma *Gradient Boosting Tree*.

D. Metode Gradient Boosting Tree

Metode yang digunakan yaitu menerapkan algoritma *Gradient Boosting Tree* sebagai algoritma pengklasifikasi. Data yang telah diperoleh dari tahap pengumpulan data yang kemudian diolah pada pengolahan data awal, akan dilakukan klasifikasi menggunakan *Gradient Boosting Tree*.

Pembelajaran dengan *Gradient Boosting Tree* akan terjadi dengan optimisasi *loss function* atau selisih dari nilai prediksi dan nilai sebenarnya sehingga nilai prediksi akan semakin mendekati nilai sebenarnya. Diagram alur dari prosedur *Gradient Boosting Tree* akan ditunjukkan pada Gbr. 2.



Gbr. 2 Diagram Alur Algoritma *Gradient Boosting Tree*

Deskripsinya adalah sebagai berikut: Pertama, masukkan data X dan y dimana X adalah independent variable dan y adalah dependent variable dan jumlah sampel adalah n dimana n dimulai dengan 1, $\{(x_i, y_i)\}_{i=1}^n$.

Kemudian, dimulai dengan *leaf* yang mewakili prediksi awal untuk setiap sampel. Untuk klasifikasi, prediksi awal merupakan $\log(\text{odds})$ dari *dependent variable* yang merupakan peluang kemunculan dari kategori dalam kolom target dimana dalam penelitian ini adalah Kompetensi Keahlian. Dan bisa dibilang sebagai rata-rata klasifikasi. Setelah diperoleh prediksi awal, hitung residual dari masing-masing sampel, residual adalah selisih dari nilai prediksi dan nilai sebenarnya dari masing-masing sampel.

Langkah selanjutnya yaitu membangun *Decision Tree* untuk memprediksi residual menggunakan atribut yang sudah ditentukan sebelumnya, dalam penelitian ini merupakan Nilai Matematika, Nilai Bahasa Indonesia, Nilai Bahasa Inggris, dan Nilai IPA. Jika jumlah residual lebih banyak dari jumlah *leaf* dalam *tree*, beberapa residual akan

berakhir di *leaf* yang sama, sehingga perlu dilakukan persamaan

$$\frac{\sum Residual_i}{\sum [Prediksi Sebelumnya_i \times (1 - Prediksi Sebelumnya_i)]} \quad (1)$$

Setelah dihitung seluruh output pada masing-masing *leaf* pada *tree*, buat prediksi baru untuk masing-masing sampel yaitu dengan menjumlahkan prediksi sebelumnya yaitu $F_0(x)$ dengan hasil perkalian *learning rate* dengan output value dari *leaf* yang sesuai dari *tree* sebelumnya sehingga bisa didapatkan Kompetensi Keahlian sebagai prediksi baru.

E. Tahap Pengembangan Sistem

Tahap pengembangan sistem adalah tahap dimana penulis merancang sistem yang akan dibuat sehingga sesuai dengan tujuan pembuatan artikel ilmiah ini sehingga tidak melenceng dari tujuan penelitian. Pada tahap ini akan dilakukan analisis sistem dan desain sistem.

1) Analisis Sistem

a. Identifikasi Masalah

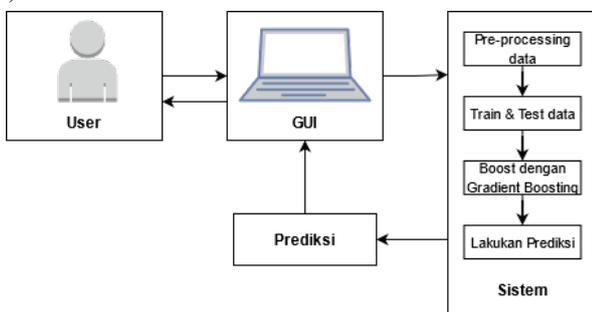
Berdasarkan latar belakang yang telah diuraikan sebelumnya maka permasalahan yang dapat dirumuskan yaitu bagaimana pengembangan Sistem Pendukung Keputusan Penentuan Jurusan di SMKN 1 Pungging Menggunakan *Gradient Boosting Tree*.

b. Analisis Kebutuhan Fungsional

Sistem Pendukung Keputusan Penentuan Jurusan di SMKN 1 Pungging dapat melakukan hal-hal sebagai berikut:

- Pengguna dapat melakukan pencarian data yang akan dijadikan *training data* dan *testing data*
- Sistem dapat melakukan klasifikasi pada *training data* dan *testing data*
- Sistem dapat menampilkan akurasi dari *training data* dan *testing data*
- Pengguna dapat memasukkan data baru
- Sistem dapat melakukan prediksi untuk data baru
- Sistem dapat menampilkan hasil prediksi dari data yang dimasukkan oleh pengguna

2) Desain Sistem



Gbr. 3 Desain Sistem

Gbr. 3 menjelaskan bahwa pertama pengguna akan diarahkan ke GUI atau tampilan dari program, kemudian sistem akan memproses data yang dimasukkan oleh pengguna dan akan menghasilkan prediksi yang akan

ditampilkan kembali kepada pengguna. Namun, sistem yang akan dibuat akan dibagi menjadi 2 proses yaitu proses pengolahan data awal (*pre-processing data*) dan proses klasifikasi menggunakan *Gradient Boosting Tree*, berikut adalah diagram alur atau *flowchart* dari kedua proses tersebut:



Gbr. 4 Flowchart Pre-processing data

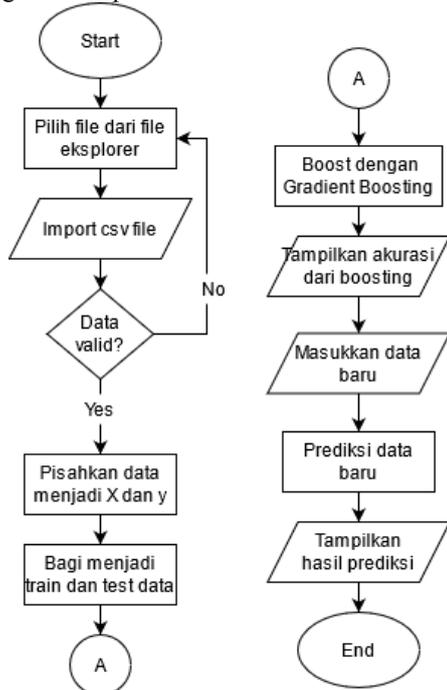
Data mentah yang diperoleh dari tahap pengumpulan data perlu diolah sehingga menjadi data yang dapat diproses oleh algoritma *Gradient Boosting Tree*. Dari Gbr. 4 dapat dilihat bahwa proses diawali dengan melakukan *import dependencies* atau memasukkan *libraries* yang diperlukan untuk menjalankan program. Proses selanjutnya yaitu memasukkan data yang akan diproses, yaitu data yang diperoleh dari SMK Negeri 1 Pungging.

Data tersebut akan dijadikan *dataset*, yaitu bentuk utama penyimpanan data dalam bahasa pemrograman Python. Setelah menjadi *dataset*, data dapat diproses sesuai dengan kebutuhan kita.

Pada data awal yang diperoleh dari tahap pengumpulan data, masih ada beberapa kolom yang tidak diperlukan dalam sistem, sehingga perlu dilakukan penghapusan pada kolom tersebut. Kemudian, dilakukan pula pengubahan nama kolom target yang sebelumnya bernama “Kompetensi Keahlian” menjadi kolom “Decision” untuk mempermudah proses pengolahan selanjutnya.

Kolom Decision yang merupakan kolom target pada data tersebut masih memiliki categorial value dengan tipe data object, sehingga perlu diubah menjadi integer agar dapat

diproses menggunakan algoritma *Gradient Boosting Tree*. Setelah itu *dataset* tersebut disimpan dalam *format file .csv* agar dapat digunakan dalam proses klasifikasi dan prediksi menggunakan algoritma *Gradient Boosting Tree* yang memiliki diagram alur pada Gbr. 5.



Gbr. 5 Flowchart proses klasifikasi dan prediksi menggunakan algoritma *Gradient Boosting Tree*

Gbr. 5 merupakan diagram alur yang menggambarkan proses dari sistem yang dijalankan oleh pengguna. Saat program dijalankan, pengguna dapat melakukan pencarian *file* melalui *file explorer*, *file* yang dimasukkan adalah *file csv* yang merupakan hasil dari proses pada Gbr. 4 dan berisi *dataset* yang siap untuk diproses dengan algoritma *Gradient Boosting Tree*. Kemudian sistem akan memeriksa apakah data yang dimasukkan *valid* atau tidak, jika data *valid* maka akan dilakukan proses selanjutnya, jika tidak maka pengguna harus memasukkan kembali data yang *valid*.

Setelah data dimasukkan, *dataset* di dalamnya akan dipisahkan menjadi *X* dan *y*, dimana *X* adalah *independent variable* dan *y* adalah *dependent variable*. *X* berisi data nilai siswa dan *y* berisi kolom target atau kolom *Decision*.

Dataset yang telah dipisahkan menjadi *X* dan *y* akan dibagi menjadi *training data* dan *testing data* dengan porsi 70% *training data* dan 30% *testing data*. Setelah itu, *training data* akan dilatih dengan algoritma *Gradient Boosting Tree*. Setelah diperoleh model yang optimal, model tersebut akan digunakan untuk menguji *testing data*. Kemudian, setelah dilakukan *training* dan *testing* akan dihitung skor atau akurasi dari masing-masing proses *training* dan *testing*.

Ketika hasil akurasi dari *training data* dan *testing data* telah diperoleh, maka kita dapat melakukan prediksi terhadap data baru atau data nilai yang dimasukkan oleh pengguna. Hasil dari prediksi ini adalah jurusan di SMK

Negeri 1 Pungging berdasarkan nilai yang dimasukkan oleh pengguna.

III. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan Sistem Pendukung Keputusan Penentuan Jurusan untuk Calon Siswa di SMKN 1 Pungging Menggunakan *Gradient Boosting Tree* dimana prosesnya dibagi menjadi 2 yaitu *pre-processing* data dan sistem itu sendiri, berikut ini adalah hasil dari 2 proses tersebut:

A. Hasil Pre-Processing Data

Data yang belum diproses atau data mentah yang diperoleh dari SMKN 1 Pungging yang berisi 272 data nama siswa, nilai siswa, dan kompetensi keahlian masing-masing siswa dimana data tersebut memiliki beberapa kolom yang tidak diperlukan dalam sistem dan memiliki tipe yang belum sesuai dengan kebutuhan sistem seperti sebagai berikut:

TABEL II
DATA YANG BELUM DIPROSES

No	Nama Siswa	MA TH	EN G	IND	IPA	Kompetensi Keahlian
1	Abid Rohman	91	84.9	92.5	76.7	Teknik Pemesinan
2	Achamd Ferry Irawan	88	82.5	89.7	74.8	Teknik Pemesinan
3	Achmad Awwalun Ni'am	89.3	83	94.6	74.6	Teknik Pemesinan
4	Affandi Oky Saputra	89.5	83.9	93.3	75	Teknik Pemesinan
5	Ahmad Reavasta Andrade	88.2	84.9	93.2	79.1	Teknik Pemesinan

Dari Tabel II, dapat dilihat bahwa masih ada data yang tidak diperlukan dan masih ada tipe data yang tidak sesuai sehingga perlu dilakukan pemrosesan data awal dengan menggunakan *Jupyter Notebook* untuk menghapus kolom yang tidak diperlukan dalam proses klasifikasi data seperti data No dan Nama Siswa karena yang diperlukan hanya data Nilai Matematika, nilai Bahasa Inggris, Nilai Bahasa Indonesia, Nilai IPA, dan Kompetensi Keahlian. Langkah selanjutnya yaitu mengubah *categorical values* pada kolom Kompetensi Keahlian yang merupakan kolom target menjadi *numerical values* dengan cara mendeteksi semua kategorinya dan mengonversinya menjadi *numerical values* sehingga hasilnya menjadi sebagai berikut dan bisa diproses dengan menggunakan algoritma *Gradient Boosting Tree*.

TABEL III
HASIL PEMETAAN DATA KOMPETENSI KEAHLIAN

class	encoded
Multimedia	0
Produksi Dan Siaran Program Televisi	1
Teknik Instalasi Tenaga Listrik	2
Teknik Kendaraan Ringan Otomotif	3
Teknik Komputer Jaringan	4
Teknik Ototronik	5
Teknik Pemesinan	6
Teknik Pengelasan	7

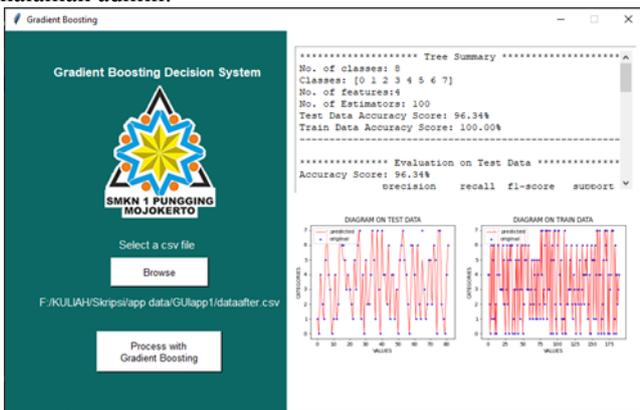
Hasil akhir dari *pre-processing data* adalah data yang siap digunakan untuk sistem disajikan pada Tabel IV.

TABEL IV
DATA SETELAH DIPROSES

MATH	ENG	IND	IPA	Decision
91	84.9	92.5	76.7	6
88	82.5	89.7	74.8	6
89.3	83	94.6	74.6	6
89.5	83.9	93.3	75	6
88.2	84.9	93.2	79.1	6

B. Hasil Pengujian Sistem

Sistem dibuat dengan bahasa pemrograman Python dan dibagi menjadi 2 tampilan yaitu tampilan untuk admin dan tampilan untuk *client* yang merupakan calon siswa di SMKN 1 Pungging, berikut ini adalah tampilan dari halaman admin:



Gbr. 6 Tampilan halaman admin

Pada tampilan halaman admin di Gbr. 6 terdapat tombol Browse yang jika ditekan akan mengarahkan ke direktori untuk mencari file hasil *pre-processing data* yang diperlukan dalam sistem. File tersebut akan diproses menggunakan Gradient Boosting dengan menekan tombol Process. Proses yang dilakukan adalah menjadikan file sebagai dataset dan memisahkannya sebagai training data dan testing data.

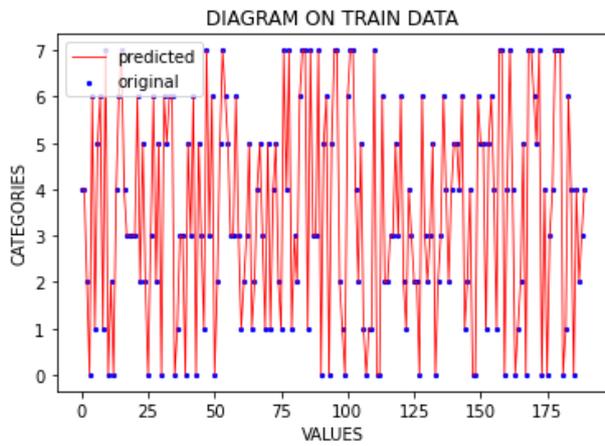
Dalam uji coba ini training data diambil sebanyak 70% dari data keseluruhan dan testing data diambil sebanyak 30% dari data keseluruhan. Hasil akurasi yang diperoleh berdasarkan sistem yang dibuat adalah 100% untuk training data dan 96,34% untuk testing data.

Berdasarkan model yang dibuat, training data mendapatkan akurasi sebesar 100%. Detail akurasi atau kualitas prediksi dari algoritma klasifikasi dengan *Gradient Boosting Tree* dari *training data* dapat dilihat berdasarkan *Classification Report* yang diperoleh. *Classification report* merupakan evaluasi model klasifikasi dengan algoritma *Gradient Boosting Tree* dari *library sklearn* yang berisi *precision* yang merupakan presentase prediksi yang benar, *recall* yang merupakan kemampuan *classifier* untuk menemukan semua *instance* positif, *F1-score* yang merupakan presentase dari prediksi positif yang benar, dan *support* yang merupakan jumlah kemunculan kelas dalam klasifikasi tersebut. Hasil *classification report* dari *training data* dapat dilihat pada Tabel V.

TABEL V
HASIL EVALUASI DATA TRAINING

	precision	recall	F1-score	support
0	1.00	1.00	1.00	26
1	1.00	1.00	1.00	21
2	1.00	1.00	1.00	23
3	1.00	1.00	1.00	27
4	1.00	1.00	1.00	21
5	1.00	1.00	1.00	24
6	1.00	1.00	1.00	24
7	1.00	1.00	1.00	24
accuracy			1.00	190
macro avg	1.00	1.00	1.00	190
weighted avg	1.00	1.00	1.00	190

Dari Tabel V dapat dilihat bahwa total data yang digunakan sebagai *training data* adalah 70% dari keseluruhan data yaitu sebanyak 190 data, dan dari 8 kategori diperoleh total akurasi 100%. Hasil akurasi dari *training data* dapat divisualisasikan dalam diagram pada Gbr. 7.



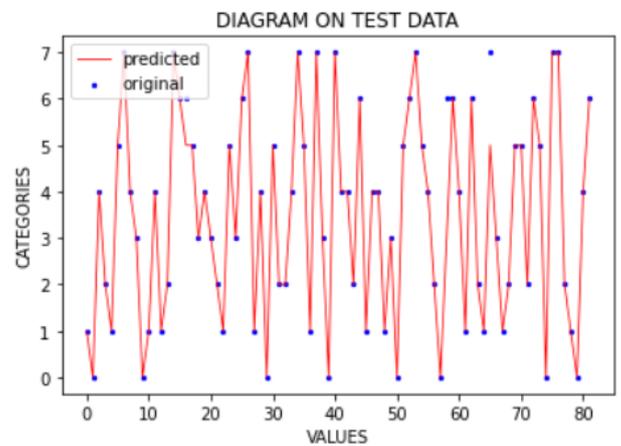
Gbr. 7 Diagram Training Data

Pada diagram *training data* di Gbr. 7 dapat dilihat bahwa dari 190 train data seluruh prediksinya sesuai dengan target yang terdiri dari 8 kategori yaitu nilai 0 sampai 7 yang merupakan nilai numerik dari 8 jurusan di SMKN 1 Pungging seperti yang ditunjukkan pada hasil pemetaan di Tabel III. Setelah melakukan proses pada *training data*, selanjutnya adalah pengujian sistem pada *testing data* yang menghasilkan akurasi sebesar 96.34% dengan hasil *classification report* pada Tabel VI.

TABEL VI
HASIL EVALUASI DATA TESTING

	precision	recall	F1-score	support
0	1.00	1.00	1.00	8
1	1.00	1.00	1.00	13
2	1.00	1.00	1.00	11
3	0.88	1.00	0.93	7
4	1.00	1.00	1.00	13
5	0.83	1.00	0.91	10
6	1.00	0.80	0.89	10
7	1.00	0.90	0.95	10
accuracy			0.96	82
macro avg	0.96	0.96	0.96	82
weighted avg	0.97	0.96	0.96	82

Dari Tabel VI dapat dilihat bahwa total data yang digunakan sebagai *testing data* adalah 30% dari keseluruhan data yaitu sebanyak 82 data. Dari 8 kategori dapat dilihat ada beberapa prediksi yang tidak benar dengan total akurasi 96.34% atau dibulatkan menjadi 96%. hasil akurasi dari *testing data* divisualisasikan dalam Gbr. 8.



Gbr. 8 Diagram Testing Data

Pada diagram test data di Gbr. 8 dapat dilihat bahwa dari 82 *testing data* ada beberapa prediksi tidak sesuai dengan target yang terdiri dari 8 kategori yaitu nilai 0 sampai 7 yang merupakan nilai numerik dari 8 jurusan di SMKN 1 Pungging seperti yang ditunjukkan pada hasil pemetaan di Tabel III.

Setelah didapatkan akurasi sebesar 100% untuk *data training* dan 96.34% untuk *testing data*, model dapat digunakan untuk melakukan prediksi terhadap jurusan yang sesuai dengan nilai calon siswa yang akan masuk di SMKN 1 Pungging. Berdasarkan penelitian sebelumnya yang dilakukan oleh Mutrofin, et al., pada tahun 2020, Gradient Boosting Tree menghasilkan akurasi testing data sebesar 91,71% [13], maka nilai akurasi dari penelitian ini lebih besar dan dapat dikatakan bahwa akan lebih akurat dalam pengambilan keputusan.

Berikut ini adalah halaman client dimana pengguna dapat memasukkan nilai dari masing-masing mata pelajaran dan mendapatkan prediksi berupa jurusan yang sesuai dengan nilai yang telah dimasukkan.



Gbr. 9 Hasil tampilan halaman client

Pada tampilan program di Gbr. 9 dapat ditunjukkan hasil prediksi untuk nilai yang telah dimasukkan adalah kompetensi keahlian Teknik Pemesinan, beberapa hasil pengujian terhadap data nilai yang berbeda ditunjukkan pada Tabel VII.

TABEL VII
HASIL PENGUJIAN TERHADAP DATA BARU

No	Nilai Matematika	Nilai Bahasa Indonesia	Nilai Bahasa Inggris	Nilai IPA	Prediksi Kompetensi Keahlian
1	80.4	72.8	93.1	84.5	Teknik Pengelasan
2	86.7	80.9	72.8	73.4	Teknik Kendaraan Ringan Otomotif
3	73.1	85.7	84.8	87.3	Produksi Dan Siaran Program Televisi
4	92.4	81	83.5	80.8	Teknik Kendaraan Ringan Otomotif
5	70.5	78.9	90.6	88.3	Teknik Pengelasan
6	91.8	82	88.9	76.1	Teknik Pemesinan
7	88.2	92.8	86.2	74.4	Teknik Instalasi Tenaga Listrik
8	87.7	81	92.9	86.4	Teknik Ototronik
9	90	90	77.1	75.3	Teknik Komputer Jaringan
10	81.7	88.6	79	89.3	Multimedia

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan oleh penulis dan didasarkan dari pembahasan sebelumnya dapat ditarik kesimpulan sebagai berikut:

1. Telah berhasil dikembangkan Sistem Pendukung Keputusan Penentuan Jurusan untuk Calon Siswa di SMKN 1 Pungging Menggunakan *Gradient Boosting Tree*.
2. Hasil menunjukkan akurasi dari sistem yang dibuat berdasarkan data yang diperoleh dari SMKN 1 Pungging adalah 100% untuk *training data* dan 96.34% untuk *testing data*.
3. Sistem Pendukung Keputusan Penentuan Jurusan untuk Calon Siswa di SMKN 1 Pungging Menggunakan *Gradient Boosting Tree* dapat membantu calon siswa untuk menentukan jurusan sesuai dengan nilai yang dimiliki.

V. SARAN

Berdasarkan pengujian terhadap Sistem Pendukung Keputusan Penentuan Jurusan untuk Calon Siswa di SMKN 1 Pungging Menggunakan *Gradient Boosting Tree*, penulis dapat memberikan beberapa saran sebagai berikut:

1. Aplikasi dapat dikembangkan lebih lanjut menggunakan bahasa pemrograman lain yang lebih dinamis berbasis Website.
2. Sistem Pendukung Keputusan dapat dikembangkan menggunakan metode lain dengan maksud mendapatkan akurasi yang lebih tinggi dan dapat memberi prediksi lebih dari 1 jurusan.

UCAPAN TERIMA KASIH

Ucapan terima kasih penulis ucapkan kepada Tuhan Yang Maha Esa, karena atas karunianya, penulis dapat menyelesaikan penelitian ini dengan baik. Kemudian ucapan terima kasih juga penulis sampaikan kepada semua pihak yang terlibat dalam terselesaikannya penelitian ini.

REFERENSI

- [1] SMKN 1 Pungging. [Online]. <http://www.smkn1pungging.sch.id/>, tanggal akses: 6 Maret 2021.
- [2] Bahri Kamal and Ghea Dwi Rahmadiane, "Pengaruh Persepsi, Akreditasi Prodi, Dan Promosi Terhadap Keputusan," *Jurnal Inspirasi Bisnis & Manajemen*, vol. I, no. 2, pp. 145-158, 2017.
- [3] Perusahaan Rintisan Inkubasi Kemenristekdikti "Youthmanual" Bantu Siswa Memilih Jurusan Kuliah Sesuai Potensi Secara Online. [Online]. <https://www.ristekbrin.go.id/perusahaan-rintisan-inkubasi-kemenristekdikti-youthmanual-bantu-siswa-memilih-jurusan-kuliah-sesuai-potensi-secara-online-2/>, tanggal akses: 6 Maret 2021.
- [4] Urzu Rahmad, "ANALISIS REFERENSI SISWA SEKOLAH MENENGAH KEJURUAN DALAM MEMILIH PROGRAM STUDI DI PERGURUAN TINGGI," *JMSP (Jurnal Manajemen dan Supervisi Pendidikan)*, vol. II, no. 2, pp. 147-152, 2018.
- [5] Othman Subhi Alshamrani and Adel Alshibani, "Automated decision support system for selecting the envelope and structural systems for educational facilities," *Building and Environment*, no. 181, 2020.
- [6] Veradilla Amalia and Syafi'ul Hamidani, "Sistem Pendukung Keputusan Rekrutmen Anggota KPSS Pemilu dengan Metode Simple Multi Attribute Rating Technique," *Jurnal Teknologi Informasi & Komunikasi*, vol. XI, no. 2, pp. 232-244, 2020.
- [7] Roki Hardianto, Fana Wiza, and Wirdah Choiriah, "SISTEM PENDUKUNG KEPUTUSAN UNIVERSITAS FAKULTAS TERBAIK UNIVERSITAS LANCANG KUNING MENGGUNAKAN METODE SMART DAN MOORA," *Teknologi dan Sistem Informasi Univrab*, vol. VI, no. 1, pp. 33-40, 2021.
- [8] Stefanos Fafalios, Pavlos Charonyktakis, and Ioannis Tsamardinos, "Gradient Boosting Trees," 2020.
- [9] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, *The elements of statistical learning: data mining, inference, and prediction.*: Springer Science & Business Media, 2009.
- [10] Brad Boehmke and Brandon Greenwell, *Hands-on machine learning with R.*: CRC Press, 2019.
- [11] Guolin Ke et al., "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, no. 300, pp. 3146-3154, 2017.
- [12] Yung-Chia Chang, Kuei-Hu Chang, and Guan-Jhih Wu, "Application of eXtreme gradient boosting trees in the construction of credit risk assessment models for financial institutions," *Applied Soft Computing*, vol. 73, no. 914-920,

- 2018.
- [13] Siti Mutfin, M. Mughniy Machfud, Diema Heryka Satyareni, Raden Venantius Hari Ginardi, and Chastine Faticah, "Komparasi Kinerja Algoritma C4. 5, Gradient Boosting Trees, Random Forests, dan Deep Learning pada Kasus Educational Data Mining," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. VII, no. 4, pp. 807-814, 2020.