p-ISSN: 2746-9190
e-ISSN: 2746-9240
Volume 1, No. 1, 2022

IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU MENGGUNAKAN *RANDOM FOREST*

Zaskila Nurfadilla[1*], Faisal[2]

Universitas Islam Negeri Alauddin Makassar[1], Universitas Islam Negeri Alauddin Makassar [2]
zaskilanurfadilla27@gmail.com[1], faisal@uin-alauddin.ac.id[2]*

Abstrak

Tingkat ketepatan kelulusan mahasiswa pada perguruan tinggi menjadi salah satu kriteria penilaian akreditasi kampus. Semakin banyak mahasiswa yang lulus tepat waktu maka semakin baik pula kinerja perguruan tinggi tersebut. Tingkat kelulusan mahasiswa sulit di prediksi secara dini, sehingga mengakibatkan keterlambatan kelulusan. Untuk mengurangi tingkat keterlambatan lulus kuliah untuk mahasiswa, perlu di didik secara serius agar dapat lulus tepat waktu. Salah satu metode penyelesaiannya untuk permasalahan tersebut yaitu dengan melakukan prediksi terhadap kelulusan ketepatan mahasiswa dengan menggunakan metode atau penambangan data. Tujuan dari sistem ini yaitu untuk mempermudah para dosen di kampus untuk mengklasifikasikan mahasiswa yang tergolong lulus tepat waktu menggunakan metode Random Forest. Hasil klasifikasi menggunakan Algoritma Random Forest menggunakan 1.351 data, maka didapatkan hasil evaluasi denga nilai akurasi sebesar 90.74% dengan membagi dan testing sebanyak 80:20 Sistem berhasil menampilkan visualisasi data utuk memprediksi kelulusan tepat waktu dengan mengimplementasikan metode .

Kata kunci*: Random Forest, Kelulusan, Prediksi.*

Abstract

IMPLEMENTATION OF DATA MINING TO PREDICATE STUDENT GRADUATION ON TIME USING RANDOM FOREST. *The level of accuracy of student graduation in tertiary institutions is one of the criteria for assessing campus accreditation. The more students who graduate on time, the better the college's performance will be. Students' graduation rates are difficult to predict early, resulting in delays in graduation. To reduce the rate of delay in graduating college for students, it is necessary to be educated seriously in order to graduate on time. One method of solving this problem is by predicting the accuracy of student graduation by using or methods. The purpose of this system is to make it easier for lecturers on campus to classify students who are classified as graduating on time using the Random Forest method. The results of the classification using the Random Forest Algorithm using 1,351 data, then the evaluation results with an accuracy value of 90.74% by dividing the training and testing data as much as 80:20 The system successfully displays data visualization to predict graduation on time by implementing .*

Keywords*: , Random Forest, Graduation, Prediction.*

AGENTS
JOURNAL OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE
p-ISSN: 2746-9190
e-ISSN: 2746-9240
Volume 1, No. 1, 2022

1. PENDAHULUAN

Pendidikan merupakan kebutuhan utama dipenuhi bagi Indonesia, karena pendidikan merupakan salah satu parameter negara dikatakan maju. Setiap kampus yang ada di Indonesia memiliki data mahasiswa yang sangat besar. Salah satunya adalah data mengenai jumlah kelulusan mahasiswa. Banyak hal yang dapat mempengaruhi lama masa studi mahasiswa, baik faktor internal maupun eksternal. Tingkat kelulusan mahasiswa sulit di prediksi secara dini, sehingga mengakibatkan keterlambatan kelulusan. Salah satu metode penyelesaiannya untuk permasalahan tersebut yaitu dengan melakukan prediksi terhadap kelulusan ketepatan mahasiswa dengan menggunakan metode *data mining* atau penambangan data. Prediksi ini bertujuan untuk mengetahui kelulusan mahasiswa tepat waktu atau tidak, yang diharapkan hasilnya dapat memberikan informasi dan masukan bagi pihak perguruan tinggi dalam membuat kebijakan demi perbaikan di masa yang akan datang [1]. Apabila mahasiswa telah diketahui klasifikasinya, maka selanjutnya akan diserahkan ke bagian pimpinan untuk diberikan tritment khusus kepada mahasiswa yang bersangkutan. Berdasarkan permasalahan tersebut di atas, maka dibuatlah sistem yang mampu memprediksi kelulusan mahasiswa tepat waktu yang diharapkan dapat membantu pihak dosen dalam mengklasifikasikan mahasiswa yang lulus tepat waktu. Aspek untuk menentukan apakah mahasiswa itu lulus tepat waktu atau tidak yaitu nilai IPS dari mahasiswa mulai dari semester satu sampai semester empat. Untuk mengurangi tingkat kelulusan mahasiswa yang lambat perlu dilakukan prediksi dini atau pediksi awal yang mampu dijadikan sebagai acuan pertimbangan pengambilan keputusan dalam memberikan tritmen dan nasehat kepada mahasiswa yang bersangkutan, agar mahasiswa mampu memanfaatkan waktunya sebaik mungkin.

2. METODE

Klasifikasi digunakan untuk menempatkan beberapa bagian yang tidak diketahui pada data dalam kelompok yang sudah diketahui. Pengelompokkan atau klasifikasi menggunakan variabel target dengan nilai nominal [2].

Algoritma *Random Forest* adalah metode yang terdiri dari gabungan pohon klasifikasi (CART) yang saling independen. Hasil prediksi yang diperoleh berdasarkan hasil *voting,* atau *vote* yang terbanyak dari kumpulan pohon klasifikasi yang terbentuk [3] Data yang akan digunakan kemudian akan diklasifikasikan, pada penelitian ini data akan dibagi menjadi kelas ltepat waktu atau tidak untuk kategori kelulusan.

Data yang akan digunakan kemudian akan diklasifikasikan, pada penelitian ini data akan dibagi menjadi kelas tepat waktu atau tidak untuk kategori kelulusan. Datasetnya yang digunakan dalam penelitian ini adalah Mobile Price Prediction yang didapat dari situs Kaggle yaitu
https://www.kaggle.com/baladikaalhariri

Proses pembelahan simpul diperoleh dengan melakukan suatu perhitungan yang dikenal dengan *Information Gain* dilakukan untuk mencari variabel yang akan dipilih untuk menumbuhkan pohon.

Untuk menumbuhkan *Information Gain* dilakukan perhitungan *information theory* yaitu *Entropy*. Dengan persamaan sebagai berikut:

$$Entropy\ (Y) = -\sum_i p(c|Y) log_2\ p(c|Y) \quad (1)$$

Dimana Y merupakan himpunan kasus dan $p(c|Y)$ merupakan proporsi nilai Y terhadap kelas c.

Sedangkan untuk mendapatkan *Gainnya* menggunakan persamaan sebagai berikut :

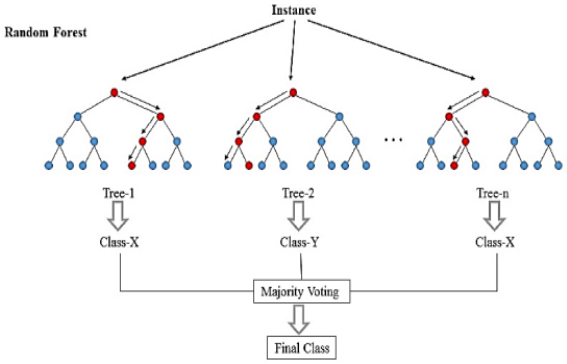
$$Gain\ (Y, a) = Entropy\ (Y) - \sum_{v \in Values(a)} \frac{|Y_v|}{|Y_a|} Entropy\ (Y_v) \quad (2)$$

Dimana nilai (a) merupakan semua yang memungkinkan dalam himpunan kasus *a.*

AGENTS
JOURNAL OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE
p-ISSN: 2746-9190
e-ISSN: 2746-9240
Volume 1, No. 1, 2022

Y_v merupakan subkelas dari Y dengan kelas v yang berhubungan dengan kelas a. Y_a merupakan semua nilai yang sesuai dengan nilai a [4].

Data yang akan digunakan kemudian akan diklasifikasikan, pada penelitian ini data akan dibagi menjadi kelas tepat waktu atau tidak untuk kategori kelulusan. Random forest adalah kombinasi dari masing – masing tree yang baik kemudian dikombinasikan ke dalam satu model [5][6]



Gambar 1. Alur kerja dari metode Random Forest

Adapun proses klasifikasi *Random Forest* yaitu :

1. Data sampel yang ada akan dilakukan pengacakan kedalam decision *tree*.
2. Menentukan variabel dataset secara acak.
3. Setelah pohon terbentuk maka akan dilakukan voting pada setiap kelas dari data sampel.
4. Kemudian hasil dari setiap kelas kemudian diambil *vote* yang paling banyak.

3. HASIL DAN PEMBAHASAN

a. Variabel Penelitian

Variabel yang akan digunakan sebanyak empat. Dataset ini berisi nilai IPS semester 1 sampai semester 4.

b. Perhitungan manual klasifikasi *algoritma Random Forest.*

Perhitungan manual *algoritma Random Forest* diperlukan untuk menganalisa *algoritma* yang digunakan. Berikut contoh perhitungan untuk mendapatkan hasil apakah aksinya tidak atau ya.

Tabel 1. Contoh Dataset

No	kehadiran	lingkungan	prakara	kerjasama	Aksi
1	Kurang	Kurang peduli	Cukup Inspiratif	Tidak mampu	Tidak
2	Rajin	Kurang peduli	inspiratif	Tidak mampu	Tidak
3	Cukup	Peduli	Inspiratif	Tidak mampu	Tidak
4	Rajin	Kurang peduli	inspiratif	mampu	ya
5	Rajin	Kurang peduli	Inspiratif	Mampu	Ya
6	Rajin	Kurang peduli	inspiratif	mampu	Ya
7	Cukup	Peduli	Inspiratif	Mampu	Ya
8	Cukup	Peduli	Inspiratif	Tidak mampu	Tidak
9	Rajin	Peduli	Cukup inspiratif	Mampu	Ya
10	Rajin	Kurang peduli	Cukup inspiratif	Mampu	Tidak
11	Rajin	Peduli	Cukup inspiratif	Mampu	Ya
12	Rajin	Kurang peduli	inspiratif	Mampu	Ya
13	Rajin	Kurang peduli	inspiratif	Mampu	?

Perhitungan nilai *entropy* menggunakan persamaan 1 dan untuk menghitung nilai *gain* menggunakan persamaan 2.

$$Entropy = -\left(\frac{4}{9}log2\left(\frac{4}{9}\right) + \left(\frac{5}{9}log2\left(\frac{5}{9}\right)\right)\right)$$
$$= \mathbf{0.991}$$

$$Gain\ (prakara) = 0.918 - \left(\frac{4}{12}(1) + \frac{8}{12}(0.81)\right)$$
$$= \mathbf{1.125}$$

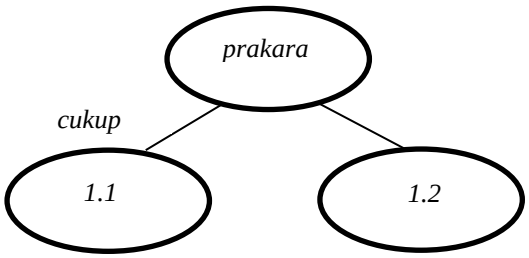
Karena prakara memiliki *gain* tertinggi maka prakara dapat menjadi node 1.

Tabel 2. Perhitungan Nilai Entropy dan Gain.

node 1	atribut		jumlah	tidak	ya	entropi	gain
	total		12	4	8	0.918	
	lingkungan						0.174
		kurang peduli	9	4	5	**0.991**	
		peduli	3	0	3	0	
	prakara						**1.125**
		cukup inspiratif	4	2	2	1	
		inspiratif	8	2	6	0.8112	

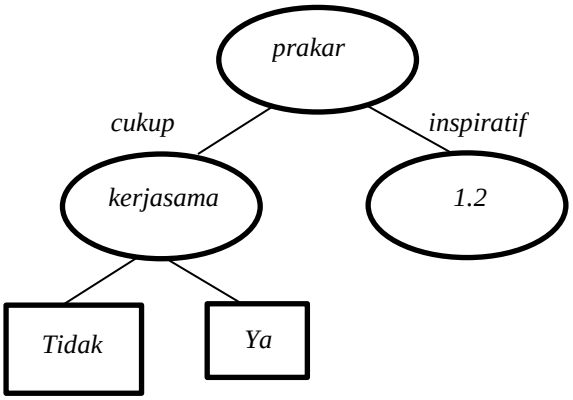
Sehingga didapatkan pohon keputusan sementara yaitu seperti diagram dibawah ini:

AGENTS
JOURNAL OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE
p-ISSN: 2746-9190
e-ISSN: 2746-9240
Volume 1, No. 1, 2022



Gambar 2. Diagram Decision tree node 1

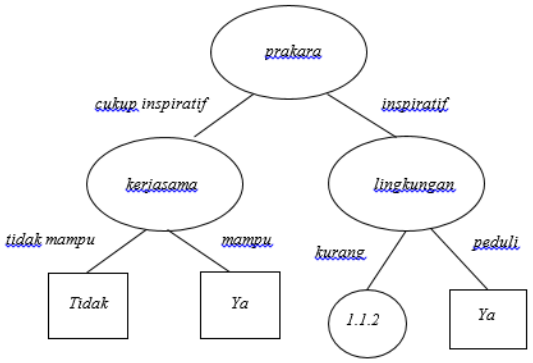
Tahap selanjunya adalah menghitung entropy dan gain node 1.1 dengan kondisi prakara = cukup inspiratif, dan mengambil variable acak yaitu lingkungan dan kerjasama. Karena nilai gain dari kerjasama lebih besar dari lingkungan maka kerjasama berada pada node 1.1. Sehingga menghasilkan diagram seperti dibawah ini.



Gambar 3. Diagram Decision tree node 1.1

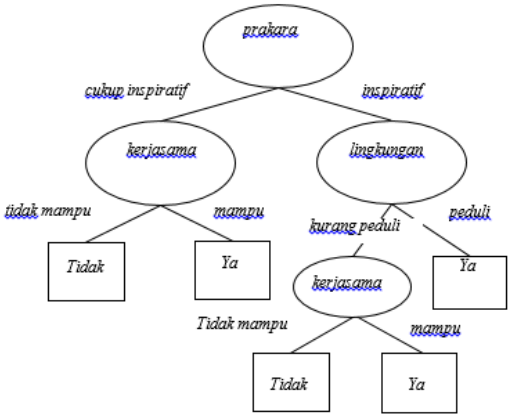
Kemudian menghitung node 1.2 dengan kondisi prakara = inspiratif dengan variabel acak yang dipilih pada node 1.2 adalah kehadiran dan lingkungan. Nilai gain tertinggi adalah lingkungan.

Pada gambar diatas hanya berisi kelas peduli dan kurang peduli, maka simpul kanan sudah homogen, sehingga tidak perlu lagi dilakukan pemecahan. Pemilihan variabel yang akan dijadikan pemilihan berikutnya yaitu dengan mencari niali gain dan nilai entropi. Sehingga menghasilkan diagram node 1.2 yaitu sebagai berikut :



Gambar 4. Diagram Decision tree node 1.2

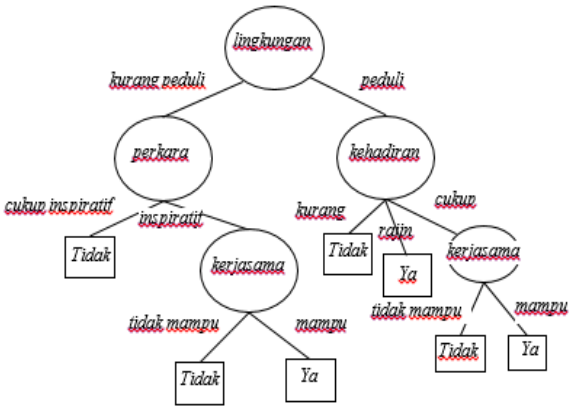
Karena simpul kiri dari cabang lingkungan masih bersifat *heterogen* sehingga dilakukan pemecahan selanjutnya. Tahap selanjutnya yaitu menghitung node 1.1.2 dengan kondisi prakara=inspiratif, lingkungan = kurang peduli dengan variabel acak yang dipilih pada node 1.1.2 adalah kehadiran dan kerjasma. Dengan menghitung nilai gain dan entropi seperti pada perhitungan sebelumnya menggunakan persamaan 1 dan 2. Dan nilai gain terringgi dari variabel yang digunakan adalah variabel kerjasama yang membandingkan antara variabel kondisi prakara dan variabel lingkungan . Hasil dari pohon yang terbentuk selanjutnya Sehingga menghasilkan diagram node 1.1.2.



Gambar 5. Diagram Pohon Keputusan Pertama

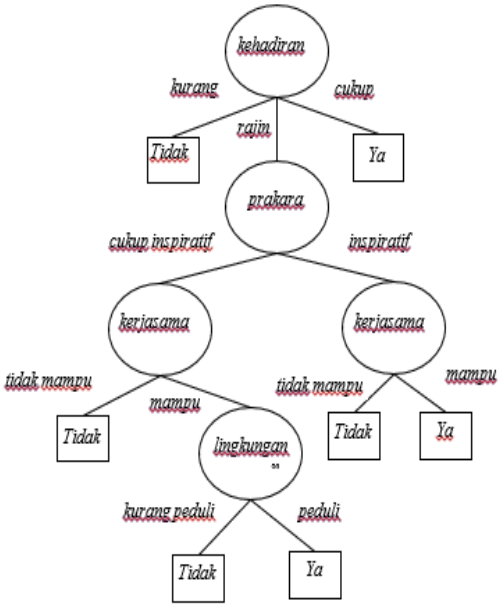
Langkah selanjutnya adalah membuat pohon keputusan dengan mengikuti langkah awal pembuatan pohon keputusan dengan variabel yang berbeda secara acak. Pada gambar dibawah hanya berisi kelas peduli dan kurang peduli, maka simpul kanan sudah homogen, sehingga tidak perlu lagi dilakukan pemecahan. Sehingga didapatkan pohon keputusan yang lain sebagai berikut:

AGENTS
JOURNAL OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE
p-ISSN: 2746-9190
e-ISSN: 2746-9240
Volume 1, No. 1, 2022



Gambar 6. Diagram Pohon Keputusan Kedua

Setelah pohon keputusan pertama terbentuk, maka tahap selanjutnya adalah membuat pohon keputusan lainnya dengan mengambil variabel yang berbeda dengan menggunakan rumus persamaan 1 dan 2 sehingga pohon yang terbentuk pada metode ini menghasilkan bentuk yang berbeda- beda seperti pada gambar pohon keputusan yang terbentuk seperti gambar diagram pohon keputusan pertama dan kedua.



Gambar 7. Diagram Pohon Keputusan Ketiga

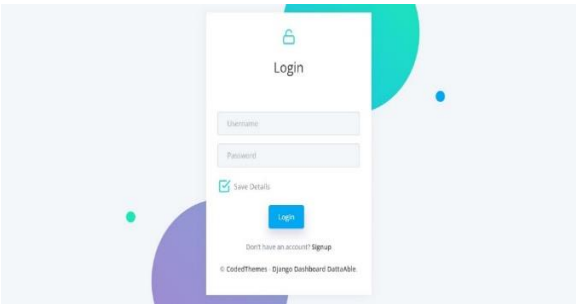
Gambar diatas merupakan salah satu pohon keputusan yang terbentuk dengan mengambil variabel awal dengan membandingkan antara variabel lingkungan prakara dan kehadiran. Hasil dari perbandingan mendapatkan nilai gain yang tertinggi adalah variabel lingkungan. Perhitungan percabangan dilakukan untuk membuat pohon keputusan akhir yang mampu melakukan prediksi awal.

Setelah membuat pohon keputusan selanjutnya yaitu menguji data yang belum diketahui kelasnya. Hasil dari pohon keputusan yang telah dibuat akan digabungkan**,** melalui pemungutan suara model atau rata-rata, menjadi model ansambel tunggal yang pada akhirnya mengungguli keluaran pohon keputusan individu. Pada gambar 5,6,7, menghasilkan kelas "tidak" hasil yang didapatkan setelah menghitung *data output* maksimal yaitu "tidak".

c. Implementasi Sistem

1) *Interface* Halaman Login

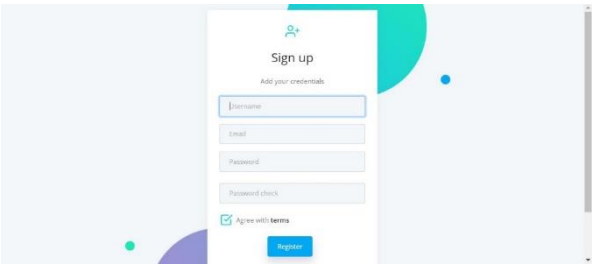
Untuk mengakses halaman utama (Dashboard), pengguna harus melakukan login menggunakan username dan password. Jika belum mempunyai akun, maka bisa melakukan registrasi dengan klik Sign Up.



Gambar 8. Menu Cetak Hasil Prediksi

2) Implementasi Halaman Registrasi

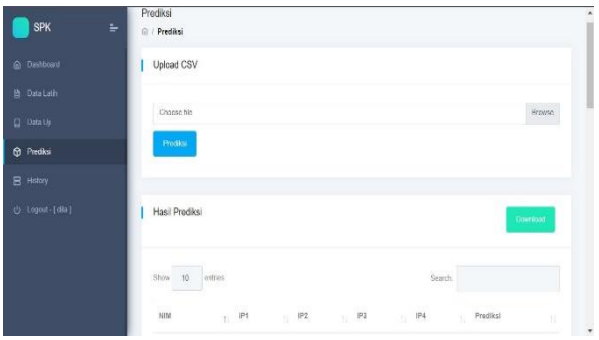
Pengguna dapat melakukan pendaftaran akun terlebiih dahulu sebelum melakukan login, dengan memasukkan Username, Email, Password dan Password check. Kemudian 'klik' centang pada 'agree with terms'.



Gambar 9. Antarmuka Registrasi

AGENTS
JOURNAL OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE
p-ISSN: 2746-9190
e-ISSN: 2746-9240
Volume 1, No. 1, 2022

3) Implementasi Halaman Dashbord

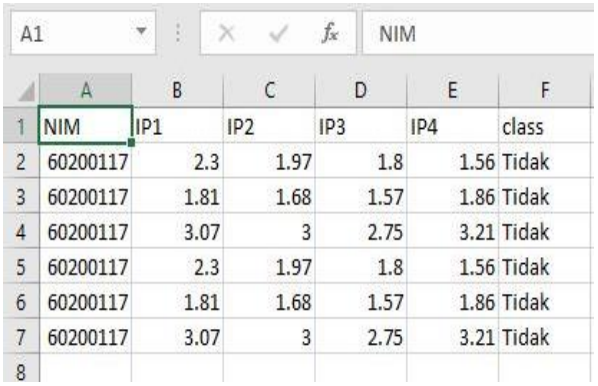


Gambar 10. Halaman Dashbord

Setelah berhasil, maka akan tampil Dashbord dari web Sistem Pendukung Keputusan. Selanjutnya akan tampil menu yang dapat diakses oleh pengguna, yaitu Dataset, Prediksi, Model Predic dan History.

4) Implementasi Halaman Data Latih

Halaman Dataset ini merupakan halaman yang berfungsi untuk mengupload dataset yang akan digunakan dengan menggunakan format file.csv.



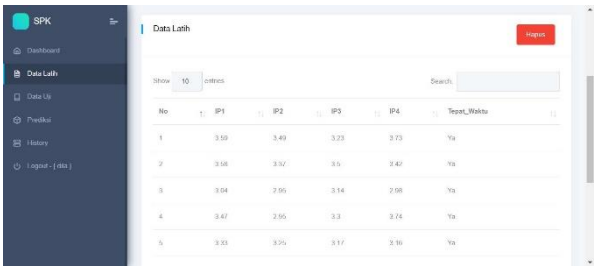
A1 | NIM

	A	B	C	D	E	F
1	NIM	IP1	IP2	IP3	IP4	class
2	60200117	2.3	1.97	1.8	1.56	Tidak
3	60200117	1.81	1.68	1.57	1.86	Tidak
4	60200117	3.07	3	2.75	3.21	Tidak
5	60200117	2.3	1.97	1.8	1.56	Tidak
6	60200117	1.81	1.68	1.57	1.86	Tidak
7	60200117	3.07	3	2.75	3.21	Tidak
8						

Gambar 11. Halaman Data Latih 1

5) Implementasi Halaman Data Latih

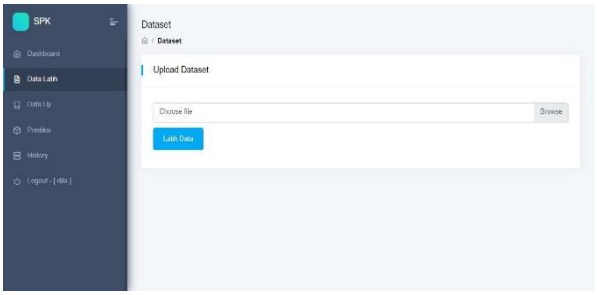
Halaman Dataset ini merupakan halaman yang berisi informasi data nilai IPS semester 1 sampai Semester 4, yang telah diketahui apakah lulus tepat waktu atau tidak. Pada halaman ini user hanya bisa melihat dan hapus dataset yang digunakan untuk membentuk model Random Forest.



Gambar 12. Halaman Data Latih 2

6) Implementasi Antarmuka Prediksi

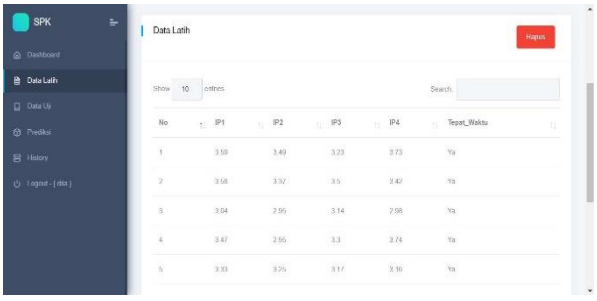
Pada halaman ini user akan melakukan uji prediksi apakah seseorang lulus tepat waktu atau tidak. User memasukkan data sesuai sesuai dengan atribut yang telah disediakan.



Gambar 13. Antar Muka Prediksi

7) Implementasi Upload Data Prediksi

Pada menu prediksi pengguna dapat mengupload file *data testing* dengan menggunakan format .csv yang belum diketahui kelasnya.

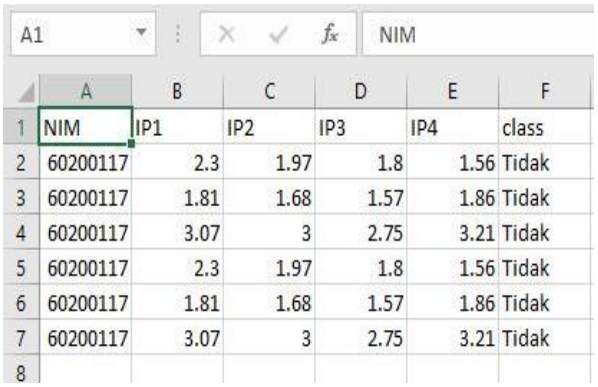


Gambar 14. Upload Data Prediksi

AGENTS
JOURNAL OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE
p-ISSN: 2746-9190
e-ISSN: 2746-9240
Volume 1, No. 1, 2022

8) Implementasi Download Data Prediksi

Pada tombol download user mampu mengambil data. .Pada menu prediksi pengguna dapat mendownload file data yang telah diketahui hasil prediksinya dengan menggunakan format excel.

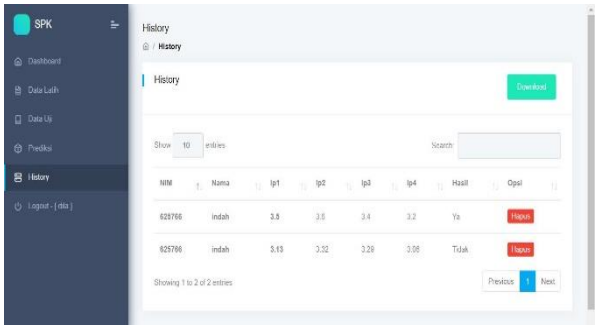


	A	B	C	D	E	F
1	NIM	IP1	IP2	IP3	IP4	class
2	60200117	2.3	1.97	1.8	1.56	Tidak
3	60200117	1.81	1.68	1.57	1.86	Tidak
4	60200117	3.07	3	2.75	3.21	Tidak
5	60200117	2.3	1.97	1.8	1.56	Tidak
6	60200117	1.81	1.68	1.57	1.86	Tidak
7	60200117	3.07	3	2.75	3.21	Tidak
8						

Gambar 15. Download Data Prediksi

9) Implementasi Halaman History

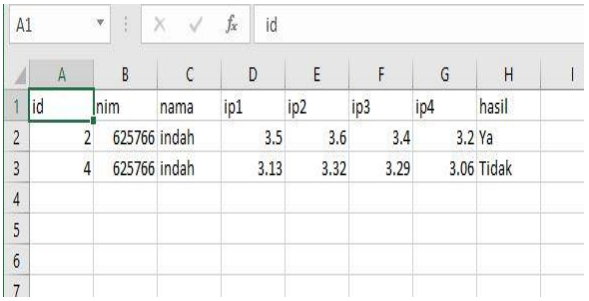
Pada bagian ini semua hasil prediksi yang telah didapatkan hasilnya akan ditampilkan secara keseluruhan.



Gambar 16. Halaman History

10) Implementasi Halaman Download Hasil Prediksi

Pada bagian ini semua hasil prediksi yang telah didapatkan hasilnya akan dicetak dengan format *excel.*



	A	B	C	D	E	F	G	H
1	id	nim	nama	ip1	ip2	ip3	ip4	hasil
2	2	625766	indah	3.5	3.6	3.4	3.2	Ya
3	4	625766	indah	3.13	3.32	3.29	3.06	Tidak
4								
5								
6								
7								

Gambar 17. Download Cetak Hasil Prediksi

d) Pengujian sistem hasil klasisfikasi algoritma *Random Forest.*

Adapun pengujian sistem yang akan dilakukan yaitu pengujian *Confusion Matrix* dan *Cross - validation* dipilih untuk menganalisis algoritma dan mengevaluasi kinerja model algoritma *Random Foresr* (Berrar, 2019:1-5). Pengujian sistem bertujuan untuk melakukan pengeksekusian sistem dengan menguji setiap proses dan kesalahan yang mungkin terjadi. Pengujian sistem yang akan dilakukan oleh penulis yaitu menggunakan Confusion Matrix dengan menghitung nilai accuracy, precission, dan recall. Selanjutnya pemilihan jenis *cross validation* dapat didasarkan pada ukuran datanya. *K-fold cross validation* untuk menguji sistem mendapatkan hasil sebagai berikut:

Tabel 3. Hasil Uji *K-Fold Validation*

K	Akurasi
2	85.01
3	89.04
4	90.74
5	91.19
6	90.93
7	90.75
8	90.69
9	90.87
10	90.87

Dari hasil analisis menggunakan *cross validation* menggunakan bantuan library *python*. Pengujian yang dilakukan menggunakan variasi jumlah *k-fold* 2-10 dan menghasilkan nilai k=optimum adalah 5 dengan nilai performan 91.19% dan

p-ISSN: 2746-9190
e-ISSN: 2746-9240
Volume 1, No. 1, 2022

memiliki persentase akurasi yang sangat baik, dengan demikian dapat disimpulkan penambahan *data training* tidak berpengaruh signifikat terhadap peningkatan performa nilai akurasi pada metode *Random Forest.* Selanjutnya menguji algoritma dengan menganalisis terhadap nilai entropy yang digunakan.

Tabel 4. Analisis Terhadap Nilai Entropy

N_Pohon	Akurasi (%)
200	91.72
400	90.34
600	89.94
800	89.55
1000	89.55
1200	89.55
1400	89.55
1600	91.12
mean	90.16

Pada tabel 4 diatas menjelaskan mengenai analisis terhadap banyaknya pohon yang dibentuk pada metode *Random Forest*. Pengujian yang dilakukan menggunakan variasi jumlah pohon yang terbentuk sebanyak 200, 400, 600, 800, 1000, 1.200, 1400, 1600 Berdasarkan hasil analis yang diperoleh jumlah nilai entropi yang normal sebanyak 800. Hasil nilai akurasi yang dihasilkan berdasarkan varian jumlah nilai entropy tidak mengalami perubahan yang besar sehingga dapat disimpulkan bahwa metode *Random Forest* baik digunakan untuk data yang banyak dan jumlah pohon yang beragam.

4. KESIMPULAN

Berdasarkan hasil penelitian Terhadap Data Kelulusan Tepat Waktu Menggunakan Algoritma *Random Forest*, diperoleh kesimpulan sebagai berikut:

a. Penggunaan dengan metode *Random Forest* dapat diterapkan dalam melakukan prediksi kelulusan mahasiswa tepat waktu.
b. Berdasarkan hasil analisis dari *cross validation*, didapatkan nilai tertinggi dari pengujian banyaknya nilai k adalah 5, artinya pembagian *data training* dan testingnya adalah 80:20, dimana didapatkan akurasi sebesar 91,19% pada saat pengujian *cross validation.*
c. Pembagian pada *data training* dan *data testing* tidak terlalu berpengaruh signifikat terhadap performa akurasi yang didapatkan pada kasus ini.
d. Hasil nilai akurasi yang dihasilkan berdasarkan varian jumlah nilai entropy tidak mengalami perubahan yang besar sehingga dapat disimpulkan bahwa metode *Random Forest* baik digunakan untuk data yang banyak dan jumlah pohon yang beragam.

5. DAFTAR PUSTAKA

[1] Rohmawan, E. P. (2013). *Prediksi Kelulusan Mahasiswa Tepat Waktumenggunakan Metode Desicion Tree*. 21–30.
[2] Linawati, Safitri Dan Nurdiani, S., & Dkk. (2020). *Prediksi Prestasi Akademik Mahasiswa Menggunakan*. *Viii*(1), 47–52.
[3] Wuryani, N., & Agustiani, S. (2021). *Random Forest Classifier Untuk Deteksi Penderita Covid-19 Berbasis Citra Ct Scan*. *7*(2). *https://Doi.Org/10.31294/Jtk.V4i2*
[4] Nugroho, Y. S., & Nova, E. (2017). *Sistem Klasifikasi Variabel Tingkat Penerimaan Konsumen Terhadap Mobil Menggunakan Metode Random Forest*. *9*(October), 24.
[5] *RP Hidayanti., Sari,M.., & Hariani. (2022). Prediksi Harga Batu Bara Menggunakan Regresi Kuadratik. Jurnal JESSI (Journal of Embedded Systems Security and Intelligent Systems.*
[6] *Salman,N., & Sari,M. (2020). Pengaruh Penyetelan Hyperparameter Terhadap Kinerja Prediksi Random Forest Pada Pendeteksian Spam. Jurnal INSTEK (Informatika Sains dan Teknologi), 5(2, 149-158.*