

Prediksi Kelulusan Mahasiswa Dengan Metode Naïve Bayes dan Artificial Neural Network: Studi Kasus Fakultas Teknik UNIS Tangerang

Ummu Habibah R¹, Achmad Solichin²

¹Department of Informatic, Universitas Tangerang Raya, Indonesia

²Department of Information System, Universitas Budi Luhur, Indonesia

Article Info

Article history:

Received 10 Jan 2022

Revised 20 Feb 2022

Accepted 16 Mar 2022

Keywords:

Predictions

Graduate Students

Naïve Bayes

Artificial Neural Network

ABSTRACT

Sheikh - Yusuf Islamic University (UNIS) is one of the leading Islamic Universities in Tangerang City which has a vision of making Islamic Higher Education superior in the field of science and technology in Southeast Asia by 2036. With the vision of the university, the engineering faculty must produce graduates in the engineering field who are superior, strong, and able to compete at national and international levels. The problems faced by the Faculty of Engineering, UNIS Tangerang, are the low percentage of students who graduate on time compared to students who graduate late. This study aims to assist the Faculty of Engineering in predicting student graduation so that it can be anticipated earlier. This study uses the attribute number of credits, IP from semesters 1 to 4. The methods used in this study are Naïve Bayes and Artificial Neural Network. The data used in this study used 330 student records who graduated from 2012 to 2016. Accuracy results obtained after testing with a system using 20% data testing obtained an accuracy of 63.63%, precision 71.05%, recall 67.5%, and AUC 62.6%.

Copyright © 2022 Universitas Indraprasta PGRI.
All rights reserved.

Corresponding Author:

Ummu Habibah Romlah

Department of Informatic,

Universitas Tangerang Raya, Indonesia

Komp. Perumahan Sudirman Indah, Jl. Ki Mas Laeng Jl. Syeh Mubarak No. 25

Tigaraksa Kab. Tangerang – Banten

Email: uhabibahr@untara.ac.id

1. PENDAHULUAN

Lulus tepat waktu adalah keinginan seluruh mahasiswa, tidak hanya itu lulus tepat waktu adalah keuntungan bagi kedua belah pihak, yaitu mahasiswa dan instansi pendidikan. Dalam mengukur status kelulusan dapat diperoleh dari indeks prestasi mahasiswa setiap semester. Jika indeks prestasi mahasiswa lebih rendah maka mahasiswa akan mengakibatkan lulus tidak tepat waktu [1].

Permasalahan terkait dengan jumlah mahasiswa lulus tidak tepat waktu terbilang sedikit dibanding dengan mahasiswa lulus tepat waktu, namun hal tersebut perlu ditindak lanjuti karena jumlah mahasiswa lulus tidak tepat waktu akan berdampak pada nilai akreditasi fakultas. Banyak faktor penyebab mahasiswa lulus tidak tepat waktu, namun belum diketahui penyebab pasti faktor keterlambatan kelulusan mahasiswa tersebut.

Perlu dilakukan penggalian data yang masih tersembunyi dan pengelolaan data sehingga menjadi pengetahuan serta informasi baru yang dapat dimanfaatkan untuk menindak lanjuti mahasiswa yang bermasalah pada akademik berjalan atau mahasiswa yang masih aktif.

Klasifikasi merupakan proses penemuan model (fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label

kelasnya tidak diketahui. Proses klasifikasi didasarkan pada empat komponen yaitu kelas *predictor*, *training dataset* dan *testing dataset* [2].



Gambar 1. Grafik Perbandingan Kelulusan Mahasiswa

Pada Gambar 1 merupakan grafik perbandingan kelulusan mahasiswa pada Fakultas Teknik dengan tahun masuk angkatan 2012-2016 untuk semua program studi. Pada jumlah mahasiswa yang lulus tidak tepat waktu pada tahun akademik 2012 mengalami kenaikan, dan pada tahun akademik 2016 mahasiswa lulus tidak tepat waktu mengalami kenaikan kembali dibandingkan dengan tahun akademik sebelumnya.

Berdasarkan penelitian sebelumnya yang telah dilakukan oleh penelitian pendahulu, penelitian ini menggunakan metode naïve bayes dan neural network untuk mengatasi masalah yang ada yaitu rendahnya presentase mahasiswa yang lulus tepat waktu dibanding mahasiswa lulus tidak tepat waktu. Peneliti sebelumnya menggunakan algoritma yang sama tetapi data yang digunakan berbeda. Atribut yang digunakan pada penelitian ini serta sudah ditentukan oleh pakar pada penelitian sebelumnya yaitu Asal Sekolah, Total SKS, Total Cuti, Kemampuan Finansial, IPS 1, IPS 2, IPS 3, IPS 4. Penelitian dilakukan di Fakultas Teknik Universitas Islam Syekh Yusuf (UNIS) Tangerang.

Tabel 1. Penelitian Terkait Prediksi Kelulusan Mahasiswa

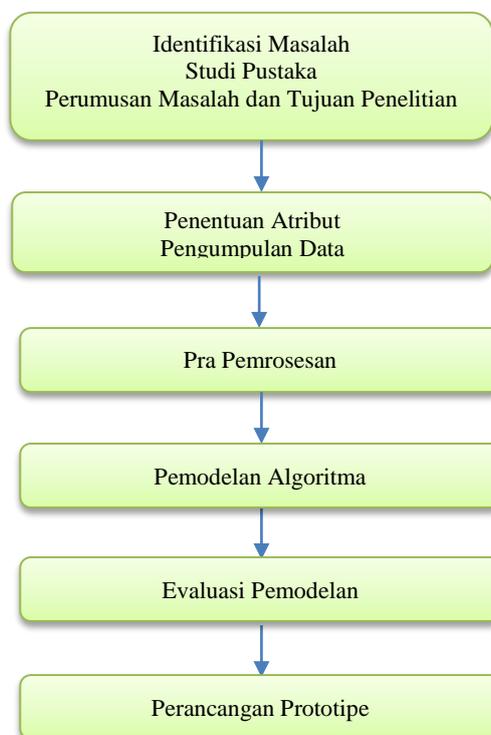
Paper	Tujuan Penelitian	Hasil Penelitian
[3]	Prediksi Kelulusan Mahasiswa	Berdasarkan hasil pengujian metode <i>decision tree</i> memiliki akurasi sebesar 74,51% dan <i>artificial neural network</i> sebesar 79,74%.
[4]	Prediksi Kelulusan Mahasiswa berbasis Neural Network	Hasil dari penelitian ini mendapatkan nilai akurasi hingga 99,9% menggunakan 166 sample data tersebut.
[5]	Perkiraan Waktu Studi Mahasiswa menggunakan Algoritma Naïve Bayes	hasil prediksi dari ketepatan lama studi dari mahasiswa angkatan 2008 adalah sebesar 254 mahasiswa diprediksi "Tepat Waktu" dan sisanya yaitu 4 orang diprediksi "Tidak Tepat Waktu"
[6]	Prediksi Masa Studi Mahasiswa menggunakan Algoritma C4.5	Jumlah data training yang akan digunakan sebanyak 83 dengan tingkat kesalahan 5% baik prediksi tepat waktu dan data prediksi tidak tepat waktu secara acak (<i>simple random sampling</i>).
[7]	Prediksi Lama Studi Mahasiswa dengan Algoritma Neural Network (Backpropagation)	Setelah dilakukan pembelajaran, dan pengujian pola-pola yang dilatih, diperoleh hasil bahwa 99% % test terhadap pola-pola tersebut benar (akurat) sesuai dengan target.

Penelitian ini menggunakan metode Naïve Bayes dan Neural Network berdasarkan data yang telah didapatkan dengan mencari beberapa penelitian terdahulu terkait prediksi kelulusan mahasiswa, dalam penelitian tersebut beberapa menggunakan Naïve Bayes saja dan Neural Network saja, sedangkan dalam penelitian ini ialah membandingkan kedua metode tersebut yaitu Naïve Bayes dan Neural Network yang nantinya akan dipilih salah satu metode terbaik setelah dilakukan testing menggunakan data yang telah didapatkan.

Tujuan yang ingin dicapai dalam penelitian ini ialah menghasilkan model prediksi kelulusan mahasiswa dengan metode Naïve Bayes atau Neural Network tersebut, diharapkan dengan adanya model prediksi ini dapat membantu fakultas untuk melakukan pencegahan dini terhadap mahasiswa yang lulus tidak tepat waktu.

2. METODE

Tujuan penelitian ini ialah untuk memprediksi sejak dini mahasiswa yang Lulus Tepat Waktu dan Tidak Tepat Waktu. Metode yang digunakan pada penelitian ini adalah metode CRISP-DM. CRISP-DM adalah metode yang dibentuk oleh komisi eropa pada tahun 1996 yang menerapkan standar dalam proses *data mining*. Dalam CRISP-DM terdapat 6 (enam) fase yang akan dilakukan pada penelitian pengembangan data mining.



Gambar 2. Tahapan Proses Penelitian

Identifikasi Masalah

Gambar 1 telah menjelaskan masalah yang terdapat pada Fakultas Teknik UNIS Tangerang yaitu jumlah mahasiswa yang lulus tidak tepat waktu tidak sebanding dengan mahasiswa lulus tepat waktu.

Studi Pustaka

Dalam penelitian ini dilakukan studi pustaka sebelumnya, dengan mencari penelitian sebelumnya yang terkait dengan judul penelitian.

Perumusan Masalah dan Tujuan Penelitian

Setelah dilakukan identifikasi masalah maka tujuan dari penelitian ialah menemukan model prediksi kelulusan mahasiswa dengan metode Naïve Bayes dan Artificial Neural Network.

Penentuan Atribut dan Pengumpulan Data

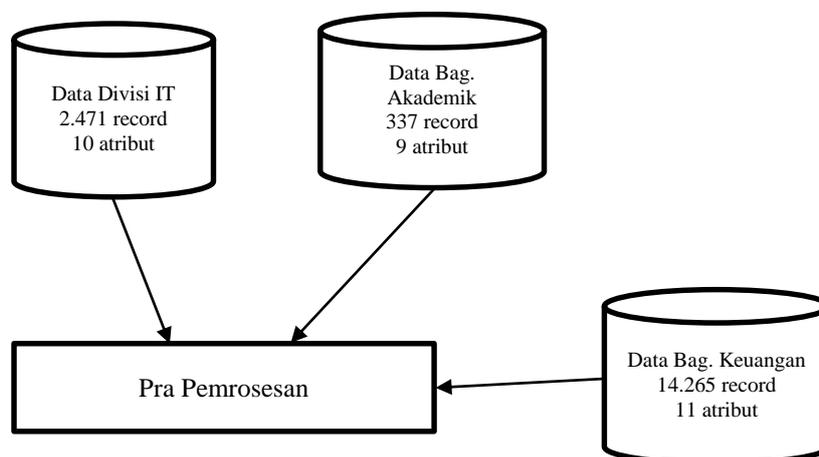
Penentuan atribut pada penelitian ini dilakukan dengan mengambil beberapa atribut pada penelitian sebelumnya dengan tujuan yang sama. Kemudian dilakukan pengumpulan data, dataset yang digunakan pada penelitian ini ialah data mahasiswa yang telah lulus dengan tahun masuk Angkatan 2012 hingga 2016 pada Fakultas Teknik untuk semua program studi yaitu Teknik Informatika, Teknik Industri, Teknik Kimia, dan Teknik Sipil. Data mentah yang dikumpulkan dari divisi IT sejumlah 1.589 *record* dari data ini sudah terdapat atribut yang dibutuhkan tetapi ini hanya data lulusan 2(dua) tahun, sedangkan yang dibutuhkan ialah data lulusan 5(tahun) terakhir. Selanjutnya data didapat dari divisi tata usaha yaitu diberikan buku wisudawan 5(lima) tahun terakhir. Kemudian karena dinilai tidak efisien data didapat kembali dari divisi IT dalam bentuk softcopy sejumlah 2.471 *record* telah ditambahkan dengan data yang sebelumnya didapat dari divisi IT, dalam data tersebut hanya terdapat mahasiswa jurusan teknik informatika dan yang dibutuhkan ialah semua program studi pada fakultas teknik yaitu teknik informatika, teknik sipil, teknik kimia dan teknik industri. Selanjutnya data didapatkan.

Id periode	NIM	Id statusmhs	Ips	Ipk	Sem mhs	Sks smt	Sks total	Ipk lulus	Sks lulus
20121	1204030008	A	3	3	1	11	11	3	11
20122	1204030008	A	2,67	2,88	2	6	17	2,88	17
20151	1204030008	A	3,68	3,35	7	20	97	3,35	97
20152	1204030008	L	0,86	3,37	8	14	104	3,37	104
20151	1504030098	A	3,18	3,18	1	20	20	3,18	20

Tabel 2. Data awal dari Divisi IT

Pra Pemrosesan

Pada fase ini dilakukan pra pemrosesan data yang telah didapatkan, dengan melakukan pembersihan data yaitu menyatukan seluruh data sesuai dengan atribut yang dibutuhkan dan melakukan reduksi terhadap data yang tidak dibutuhkan. Setelah mereduksi data yang tidak dibutuhkan, kemudian dilakukan transformasi data yaitu merubah data yang didapat sesuai dengan data yang dibutuhkan dengan cara membentuk atribut baru sebagai parameter dan generalisasi pada data yang berlebihan dari yang dibutuhkan.



Gambar 3. Komposisi Dataset Awal

Pada gambar 3 merupakan sekumpulan dataset awal yang didapatkan yaitu data mahasiswa yang telah lulus dari divisi IT berjumlah 2.471 record dan data mahasiswa yang telah lulus dari bagian akademik berjumlah 337 record, data keuangan mahasiswa yang telah lulus dari bagian keuangan 14.265 record. Dataset tersebut berisi seluruh data mahasiswa pada tahun 2012-2016, terdapat data mahasiswa yang telah lulus, mahasiswa tidak aktif, dan mahasiswa drop out dan *history* pembayaran mahasiswa yang telah lulus

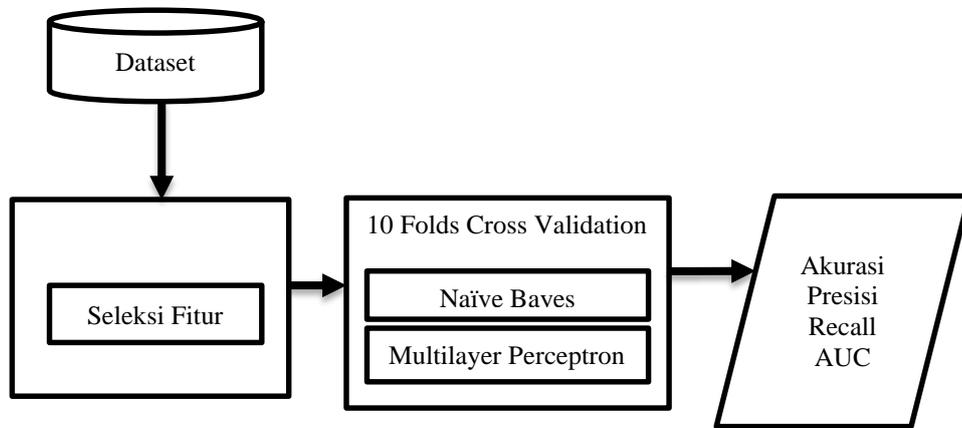
Pemodelan Algoritma

Pada tahap ini dilakukan pemodelan algoritma menggunakan tools yaitu WEKA dengan melakukan eksplorasi dan eksperimen untuk mendapatkan akurasi yang terbaik. Proses pemodelan ini menggunakan algoritma Naïve Bayes dan Multilayer Perceptron.

Evaluasi Pemodelan

Fase ini dilakukan evaluasi pemodelan sebelum merancang prototipe yaitu evaluasi menggunakan *jupyter notebook python* untuk membuktikan ketepatan hasil pengujian yang didapat. Model pembelajaran dengan metode *naïve bayes* dan *neural network* berdasarkan penelitian awal yang telah dilakukan sebelumnya. Metode yang terbaik setelah dilakukan eksplorasi menggunakan aplikasi WEKA pada penelitian ini ialah algoritma *multilayer perceptron* lebih tinggi akurasinya dan untuk algoritma *naïve bayes* pada urutan kedua setelah dilakukan eksperimen menggunakan WEKA untuk mendapatkan hasil perbandingan akurasi terbaik maka tidak bisa menggunakan eksplorasi dengan 1(satu) algoritma saja dan kemudian dilakukan pemodelan juga

menggunakan *jupyter notebook* dengan bahasa pemrograman *python*. Dimulai dengan melakukan *import library* yang akan digunakan, selanjutnya menginstall *library* jika masih ada *library* yang belum tersedia pada *library python* tersebut. Setelah didapatkan kernel terbaik maka selanjutnya mengevaluasi prediksi dengan data testing.



Gambar 4 . Eksplorasi menggunakan WEKA

Perancangan Prototipe

Tahap terakhir ialah perancangan prototipe yaitu menggunakan framework flash dari hasil model yang telah dibuat terlebih dahulu menggunakan aplikasi WEKA dan *Jupyter Notebook*.

Naïve Bayes

Metode naïve bayes merupakan salah satu teknik data mining yang paling populer untuk mengklasifikasikan data dalam jumlah besar dan dapat digunakan untuk probabilitas keanggotaan suatu class.

Algoritma Naïve bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. Naïve bayes merupakan pengklasifikasian metode probabilistik dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema bayes. Teorema tersebut dikombinasikan dengan Naïve dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi Naïve bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri kelas lainnya[8].

Naïve bayes merupakan teknik prediksi berbasis probabilistic sederhana yang berdasar pada penerapan teorema bayes (aturan bayes) dengan asumsi independensi (ketidaktergantungan) yang kuat (naif). Dengan kata lain, dalam naïve bayes model yang digunakan adalah model fitur independent.

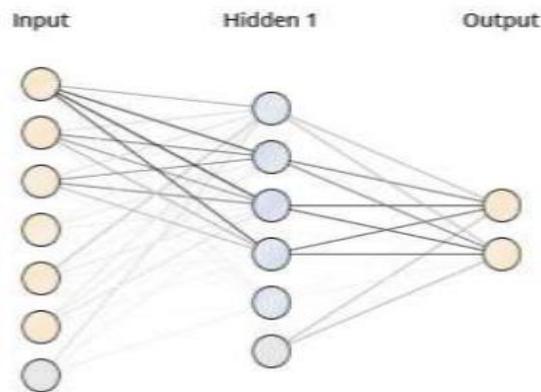
$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \dots \dots \dots (1)$$

Penjelasan dari persamaan (1) adalah sebagai berikut :

- P(H|E) : Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis H terjadi jika diberikan bukti (*evidence*) E terjadi.
- P(E|H) : Probabilitas sebuah bukti E terjadi akan memengaruhi hipotesis H.
- P(H) : Probabilitas awal (priori) hipotesis H terjadi tanpa memandang bukti apapun.
- P(E) : Probabilitas awal (priori) bukti E terjadi tanpa memandang hipotesis Atau bukti yang lain.

Artificial Neural Network

Multilayer perceptron merupakan model yang paling banyak digunakan untuk melakukan proses prediksi. Pada *Multilayer Perceptron* variabel independen maupun dependen dapat memiliki tingkat pengukuran metrik maupun nonmetric. *Multilayer Perceptron* ialah ANN turunan dari *perceptron*, berupa ANN umpan balik (*feedforward*) dengan satu atau lebih layer tersembunyi (*hidden layer*).



Gambar 5. Model Arsitektur Neural Network

Neural network (jaringan saraf tiruan) adalah model non-linear yang rumit dibangun dari komponen yang secara individu berperilaku mirip seperti model regresi. Neural network dapat direpresentasikan sebuah grafik, dan beberapa sub-grafik tampaknya ada integritas yang sama dengan gerbang logika. Struktur dari jaringan neuron atau saraf secara terperinci dirancang terlebih dahulu. *Neural network* adalah model mesin pembelajaran yang meniru aspek pembelajaran dari pengalaman masa lalu untuk memprediksi masa depan [9].

Neural Network adalah model algoritmik yang terinspirasi oleh bagaimana neuron dalam otak manusia bekerja. Setiap neuron di otak manusia saling terhubung dan informasi mengalir dari masing-masing neuron tersebut. Adapun metode penelitian yang digunakan dalam penelitian ini meliputi metode pengumpulan data, dan metode algoritma neural network menggunakan data training, learning, maupun testing [10].

Artificial Neural Network (ANN) sejak diperkenalkan pada sekitar tahun 1940 telah banyak diimplementasikan pada berbagai bidang keilmuan. ANN banyak digunakan untuk melakukan prediksi atau peramalan. *Artificial neural network (ANN)* terinspirasi dari kesadaran atas *complex learning system* pada otak yang terdiri dari set- set neuron yang saling berhubungan secara dekat. Jaringan neuron mampu melakukan tugas yang sangat kompleks seperti klasifikasi dan pemahaman pola. ANN dapat memperkirakan rentang yang cukup luas suatu model statistika dan fleksibel dalam menggambarkan model (linier maupun nonlinier). ANN dapat digunakan untuk permasalahan yang sama dengan permasalahan statistika multivariat seperti *multiple regression*, analisa diskriminan, dan analisa kluster.

Algoritma Back-Propagation

Backpropagation merupakan algoritma pembelajaran yang terawasi dan biasanya digunakan oleh perceptron dengan banyak lapisan untuk mengubah bobot- bobot yang terhubung dengan neuron-neuron yang ada pada lapisan tersembunyinya. Algoritma backpropagation menggunakan error output untuk mengubah nilai bobot-bobotnya dalam arah mundur (*backward*). Untuk mendapatkan error ini, tahap perambatan maju (*feedforward*) harus dikerjakan terlebih dahulu [11].

Masukan dari node input diteruskan ke hidden layer kemudian dilanjutkan ke node output. Setiap hubungan dari unit i ke unit j memiliki bobot w_{ij} yang mengindikasikan kekuatan dari koneksi. Jumlah dari pembobotan, a_j , untuk suatu input x_{ij} dan bobot w_{ij} . Nilai galat, $E_j(n)$, antara output actual $y_j(n)$ dan nilai output dari neuron $d_j(n)$ dihitung dengan persamaan 2.

$$E_j(n) = d_j(n) y_j(n) \dots \dots \dots (2)$$

Multilayer Perceptron

Multilayer Perceptron lebih baik daripada *Single Layer Perceptron (SLP)*. Setiap node h_i pada 1 hidden layer berbobot (**Gambar 2a**) dianalogikan sebagai garis- garis yang membentuk pola **polygon konveks** (**Gambar 2b**). Pola ini akan semakin bertambah seiring dengan bertambahnya jumlah hidden layer. Dengan kemampuan ini, MLP dapat meningkatkan kemampuan belajar (training) dari algoritma JST [12].

3. HASIL DAN PEMBAHASAN

3.1. Eksplorasi menggunakan WEKA

Eksplorasi menggunakan WEKA dilakukan sebelum menggunakan *python*. Dalam eksplorasi ini untuk membandingkan algoritma yang digunakan dengan menggunakan dataset yang sudah selesai dilakukan prapemrosesan atau atribut yang siap digunakan untuk digunakan. Dalam penelitian ini menggunakan dataset sebanyak 330 *record* yang dibagi menjadi dua yaitu data latih dan data uji.

Algoritma (Klasifikasi)	Akurasi	Presisi	Recall	Nilai AUC	Eksplorasi Algoritma
Naïve Bayes	0.735	0.75	0.74	0.75	
Multilayer Perceptron	0.853	0.86	0.85	0.88	

Tabel 3 merupakan hasil eksplorasi perbandingan algoritma menggunakan WEKA, dari eksplorasi tersebut didapatkan hasil akurasi tertinggi dan Nilai AUC tertinggi yaitu *Multilayer Perceptron*. Kemudian penelitian dilanjutkan menggunakan *Multilayer Perceptron*.

3.2. Eksperimen dengan WEKA

Setelah melakukan eksplorasi dengan WEKA maka dilakukan kembali eksperimen. Perbedaan dari eksplorasi dan eksperimen ialah eksplorasi perbandingan menggunakan algoritma satu persatu, jika eksperimen perbandingan semua algoritma menjadi satu.

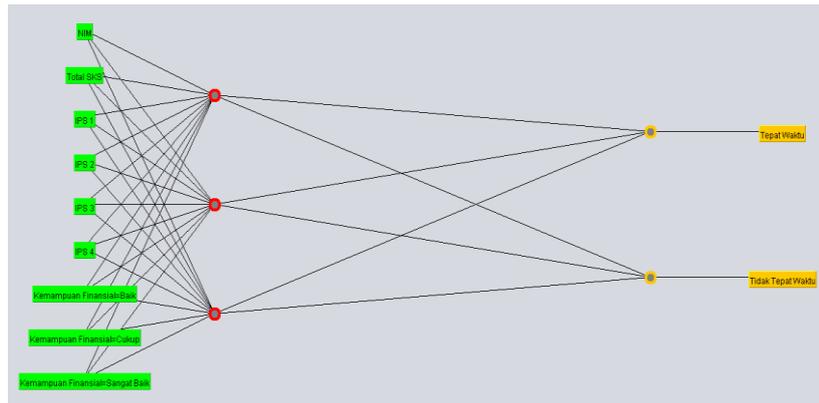
Tabel 4. Hasil eksperimen perbandingan algoritma

Algoritma (Klasifikasi)	Akurasi	Presisi	Recall	Nilai AUC
Naïve Bayes	0.719	0.723	0.720	0.742
Multilayer Perceptron	0.871	0.876	0.871	0.882

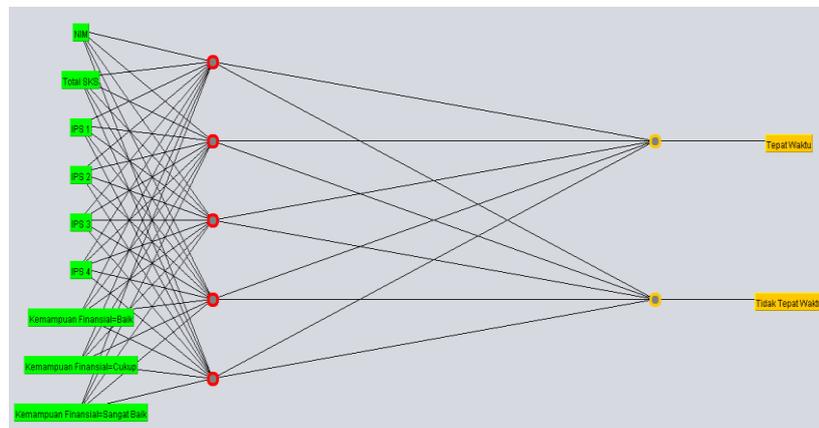
Pada tabel 4 merupakan hasil eksperimen menggunakan WEKA, dilakukan eksperimen tersebut maka algoritma multilayer perceptron tetap unggul dibandingkan dengan algoritma naïve bayes. Berdasarkan hasil eksperimen tersebut maka pengembangan atau perancangan prototipe menggunakan algoritma multilayer perceptron.

3.3. Pemodelan dengan WEKA

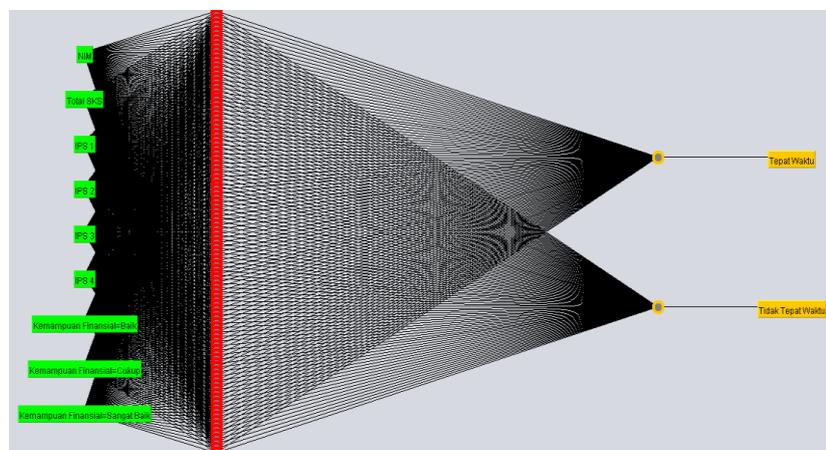
Setelah melakukan eksplorasi dan eksperimen menggunakan WEKA maka dilakukan pula pemodelan menggunakan WEKA sebelum melakukan deployment pada *python*. Dengan menentukan beberapa parameter dalam *Multilayer Perceptron* menggunakan WEKA untuk mendapatkan *error per epoch* terkecil.

Gambar 6. Model *multilayer perceptron*

Pada gambar 6 merupakan model pertama yang digunakan dengan beberapa parameter yaitu 3 (tiga) *hidden layer*, dengan *learning rate* 0.1 dan momentum 0.2, dan kemudian untuk epochnya 500. Maka didapatkan hasil *error per epoch* yaitu 0.0777273.

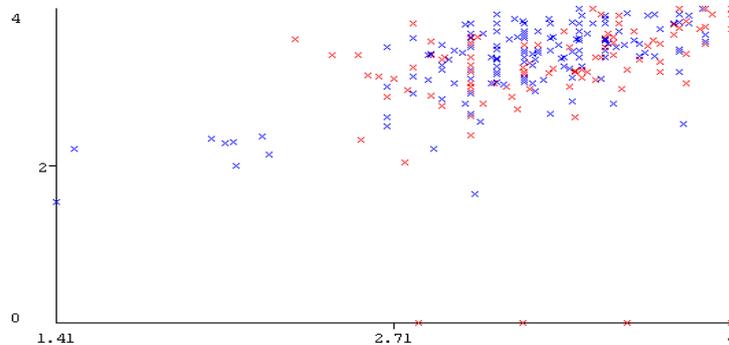
Gambar 7. Model Kedua *multilayer perceptron*

Pada gambar 7 merupakan model kedua menggunakan *multilayer perceptron* menggunakan beberapa parameter yaitu terdapat 5(lima) *hidden layer*, dengan *learning rate* 0.3 dan momentum 0.2, kemudian untuk epochnya menggunakan 700 maka didapatkan hasil *error per epoch* yaitu 0.0702772.

Gambar 8. Model ketiga *multilayer perceptron*

Pada gambar 8 dilakukan eksplorasi kembali dengan *multilayer perceptron* karena pada model kedua yaitu 0.70 sedangkan untuk eksplorasi model ketiga menggunakan beberapa parameter yaitu terdapat 100 (seratus) *hidden layer*, dengan learning rate yang digunakan ialah 0.001 dan momentumnya 0.9, serta menggunakan epoch 500 maka didapatkan hasil *error per epoch* yaitu 0.1864394.

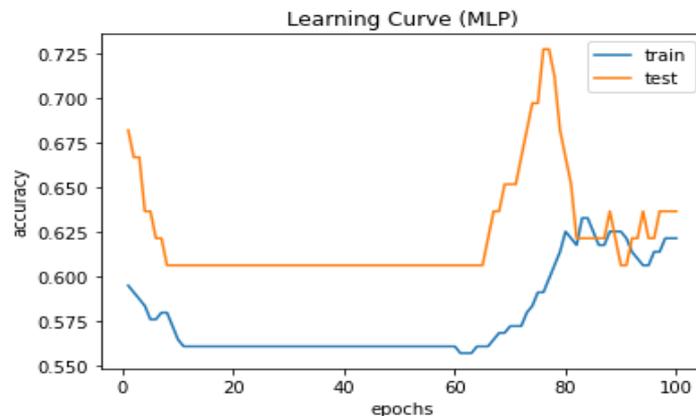
3.4. Visualisasi Data Training



Gambar 9. Plot Matrix Data Mining

Pada gambar 9 merupakan hasil visualisasi dari data training yang digunakan yaitu 80%. Hasil visualisasi menggunakan atribut IP Semester 3, dengan berwarna merah ialah Lulus Tidak Tepat Waktu dan berwarna biru ialah Lulus Tepat Waktu.

3.5. Implementasi Pemodelan



Gambar 10. Grafik akurasi

Gambar 10 merupakan hasil dari pemodelan *multilayer perceptron* menggunakan *Jupyter Notebook python* yaitu grafik akurasi alpha dengan menggunakan parameter alpha 0.0005 dan *learning rate* 0.0001 dengan maksimal epoch yang digunakan ialah 100 epoch.

Setelah dilakukan eksplorasi dan eksperimen menggunakan WEKA dan *Jupyter Notebook* maka implementasi pada penelitian ini ialah menggunakan algoritma *multilayer perceptron*, karena hasil akurasi yang didapatkan lebih tinggi dibandingkan dengan Naïve Bayes dan dilakukan uji coba menggunakan data testing tetap yang didapatkan ialah algoritma *multilayer perceptron* lebih tinggi dibandingkan dengan Naïve Bayes.

```

...
split data train dan testing
test_size=0.2 -> 20% testing, 80% training
random_state=38 -> aga
...
Xtrain, Xtest, ytrain, ytest = train_test_split(X,y, test_size=0.2, random_state=38)

```

Gambar 11. Pembagian data testing dan data training

Xtrain								
	Total SKS	IPS 1	IPS 2	IPS 3	IPS 4	Finansial_Baik	Finansial_Cukup	Finansial_Sangat Baik
198	80	3.03	3.02	3.65	3.39	1	0	0
184	83	3.50	3.45	3.51	3.56	0	0	1
213	80	3.18	3.21	3.25	3.34	1	0	0
161	77	2.42	3.09	3.21	3.16	0	0	1
18	79	3.05	3.15	2.47	3.40	0	1	0
...
23	77	3.21	3.00	3.00	3.21	0	1	0
225	80	2.65	3.05	3.73	3.74	1	0	0
258	80	3.38	3.36	3.55	3.47	1	0	0
316	77	3.03	3.20	3.68	3.50	1	0	0
53	74	2.75	2.68	3.10	3.34	0	1	0

264 rows × 8 columns

```
len(Xtrain), len(Xtest)
```

(264, 66)

Gambar 12. Hasil pembagian data testing dan data training

Pada gambar 11 dan 12 merupakan Implementasi pemodelan *multilayer perceptron* menggunakan *jupyter notebook* setelah sebelumnya dilakukan pula pemodelan menggunakan WEKA. Pada gambar 4.5 merupakan uji coba kembali menggunakan *jupyter notebook* dengan menggunakan data testing 20% dari keseluruhan dataset yang telah dibagi sebelumnya yaitu data testing tersebut terdapat 66 *record*.

4. PENUTUP

Pada penelitian ini menggunakan atribut input yaitu Total SKS, KEMAPUAN FINANSIAL, IP SEMESTER 1, IP SEMESTER 2, IP SEMESTER 3, IP SEMESTER 4. Menggunakan *class* atau *output* yaitu Lulus Tepat Waktu dan Tidak Tepat Waktu. Hasil pemodelan yang didapatkan menggunakan aplikasi WEKA dengan algoritma *Multilayer Perceptron* akurasi yang didapat 0,871, presisi 0,876, recall 0,871 dan AUC 0,882. Setelah dilakukan uji coba dengan sistem menggunakan data testing 20% maka didapatkan akurasi 63,63%, presisi 71,05%, recall 67,5%, dan AUC 62,6%.

Model yang dihasilkan pada penelitian saat ini dapat memberikan peluang masa depan bagi Fakultas Teknik UNIS Tangerang untuk mencegah pencegahan dini mahasiswa Lulus Tidak Tepat Waktu, penelitian selanjutnya dapat dilakukan menggunakan dengan menambahkan fitur lain yang berpengaruh serta membandingkan beberapa metode klasifikasi yang terbaru.

DAFTAR PUSTAKA

- [1] I. Ali dan S. Lana, "Optimasi Parameter Artificial Neural Network Menggunakan Algoritma Genetika Untuk Prediksi Kelulusan Mahasiswa," vol. 18, no. 1, hal. 54–59, 2019.
- [2] D. Salmu, S. dan A. Solichin, "Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naïve Bayes : Studi Kasus UIN Syarif Hidayatullah Jakarta Prediction of Timeliness Graduation of Students Using Naïve Bayes : A Case Study at Islamic State University Syarif Hidayatullah Jakarta," *Pros. Semin. Nas. Multidisiplin Ilmu*, no. April, hal. 701–709, 2017.
- [3] E. P. Rohmawan, "Menggunakan Metode Desicion Tree," *J. Ilm. Matrik Vol.20 No.1, April 2018* 21–30, 2013.
- [4] B. Rahmani dan H. Aprilianto, "Early model of student's graduation prediction based on neural network," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 12, no. 2, hal. 465–474, 2014, doi: 10.12928/TELKOMNIKA.v12i2.1603.
- [5] A. Jananto, "Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa," *Tekno. Inf. Din.*, vol. 18, no. 1, hal. 9–16, 2013.

- [6] S. Haryati, A. Sudarsono, dan E. Suryana, "Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu)," *J. Media Infotama*, vol. 11, no. 2, hal. 130–138, 2015.
- [7] D. Kartini, "Penerapan Data Mining dengan Algoritma Neural Network (Backpropagation) Untuk Prediksi Lama Studi Mahasiswa," *Pros. Semin. Nas. Sisfotek*, vol. 3584, hal. 235–241, 2017, [Daring]. Tersedia pada: www.seminar.iaii.or.id %7C ISSN 2597-3584 (media online).
- [8] S. Salmu dan A. Solichin, "Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naïve Bayes : Studi Kasus UIN Syarif Hidayatullah Jakarta," *Semin. Nas. Multidisiplin Ilmu 2017*, no. April, hal. 701–709, 2017.
- [9] N. Purwati, R. Nurlistiani, dan O. Devinsen, "Data Mining Dengan Algoritma Neural Network Dan Visualisasi Data Untuk Prediksi Kelulusan Mahasiswa," *J. Inform.*, vol. 20, no. 2, hal. 156–163, 2020, doi: 10.30873/ji.v20i2.2273.
- [10] R. Ridwan, H. Lubis, dan P. Kustanto, "Implementasi Algoritma Neural Network dalam Memprediksi Tingkat Kelulusan Mahasiswa," *J. Media Inform. Budidarma*, vol. 4, no. 2, hal. 286, 2020, doi: 10.30865/mib.v4i2.2035.
- [11] M. D. Y. M. F. Amin, "Implementasi Algoritma Backpropagation Untuk Memprediksi Kelulusan Mahasiswa," *Klik - Kumpul. J. Ilmu Komput.*, vol. 5, no. 2, hal. 169, 2018.
- [12] E. Tungadi, I. Thalib, M. Nur, dan Y. Utomo, "Machine Learning Penentuan Penerima Beasiswa Peningkatan Prestasi Akademik (PPA) Menggunakan Metode Jaringan Saraf Tiruan (JST)," no. August 2019, 2018.