

ANALISIS DAN PREDIKSI PRESENSI AKADEMIK SISWA MENGUNAKAN MACHINE LEARNING

**Syaiful Rahmat¹, Muhammad Rizki Julyadin², Muhamad Ardiansyah Prasetya³, Ihsan Fadhillah⁴,
Perani Rosyani⁵**

¹⁻⁵Universitas Pamulang; Jl. Surya Kencana No.1, Pamulang Barat, (021) 741-2566 atau 7470 9855

¹⁻⁵Jurusan Teknik Informatika, Fakultas Teknik, Universitas Pamulang

e-mail: *¹syaifulrahmat056@gmail.com, ²rjulyadin@gmail.com, ³muhamadardiansyahtkj@gmail.com
⁴icancinangka@gmail.com, ⁵dosen00837@unpam.ac.id

Abstrak

Menganalisis kinerja akademik siswa sangat penting untuk akademik lembaga dan pendidik, untuk mengetahui cara-cara meningkatkan individu siswa kinerja. Proyek ini menganalisis hasil masa lalu siswa termasuk individu mereka atribut termasuk usia, distribusi demografis, latar belakang keluarga dan sikap untuk belajar dan menguji data ini menggunakan alat pembelajaran mesin. Tiga model yaitu; Regresi linier untuk pembelajaran terawasi, regresi linier dengan pembelajaran mendalam dan jaringan saraf diuji menggunakan data uji dan latihan dengan regresi Linier untuk pembelajaran terawasi memiliki rata-rata terbaik rata-rata kesalahan (MAE)

Kata kunci : Analisis, Machine Learning

I. PENDAHULUAN

Kemajuan teknologi merupakan hal yang patut untuk kita syukuri, sebab dengan hadirnya teknologi berbagai pemenuhan kebutuhan hidup manusia menjadi lebih mudah. Pada dasarnya teknologi memberikan implikasi positif dalam sejarah kehidupan manusia, bahkan kemajuan teknologi menjadi bukti perkembangan kemampuan manusia untuk menggunakan nalar dan pikirannya dalam mengelola alam dan potensi diri manusia itu sendiri. Suatu Institusi pendidikan dikatakan sukses jika mampu membawa nama institusi pendidikan tersebut berkembang dan dapat dikenal oleh masyarakat luas yang nantinya institusi pendidikan tersebut benar-benar sebagai sarana atau wadah menciptakan sumber daya manusia yang unggul yang siap mengembangkan ilmu yang didapat dari bangku kuliah dan diterapkan dalam dunia kerja, untuk mencapai pendidikan yang unggul diperlukan sarana dan prasarana salah satunya laboratorium dengan fasilitas yang lengkap, sistem

yang baik serta sumber daya manusia yang hebat, untuk mencapai ini semua institusi pendidikan perlu dikelola dengan baik yang baik serta sumber daya manusia laboratorium yang lengkap memerlukan teknologi dan informasi sebagai media pendukung. Dengan teknologi dan sistem informasi yang baik diharapkan dapat bersaing dan dapat dipercaya oleh masyarakat yang memerlukan informasi pendidikan tersebut, selain itu sistem informasi yang baik yang dibuat akan membantu manajemen dalam pengambilan keputusan.[1]

Untuk memajukan suatu institusi pendidikan diharapkan institusi dapat menentukan sistem yang mana paling sesuai dengan kebutuhan, sistem yang melibatkan media computer atau sistem secara manual karena masing-masing sistem punya kekurangan dan kelebihan. Dimana untuk sistem yang melibatkan media komputer jelas lebih cepat dibanding sistem manual yang tercatat dan menggunakan tangan. Karena tidaklah mungkin suatu sistem yang Komplek dapat berfungsi tanpa melibatkan elemen komputer. Jika laporan dilakukan

secara manual maka akan timbul masalah, hal ini terjadi karena adanya keterbatasan otak manusia. Adapun masalah yang sering terjadi adalah suatu kesalahan dalam pencatatan, penghitungan, lemahnya *database* serta suatu data yang tidak terjamin keamanannya. Dan serta akan terdapat kurangnya kecepatan, dan ketepatan dalam penyajian informasi atau laporan. Setelah adanya sistem yang akan melibatkan media computer tersebut masalah-masalah yang akan timbul dan akan cepatteratasi.[2]

Kinerja siswa menjadi perhatian yang mendalam di lembaga pendidikan di mana berbagai faktor dapat mempengaruhi kinerja siswa. Memprediksi kinerja akademik siswa telah lama menjadi topik penelitian penting di banyak disiplinnya ilmu [3][4]. Untuk itu berdasarkan hasil model prediktif, instruktur dapat mengambil langkah-langkah dalam proaktif dan untuk meningkatkan pembelajaran siswa, terutama bagi siswa yang berkinerja rendah. [5][6]

Ilmu data dan Pembelajaran mesin selama bertahun-tahun telah terbukti sangat efisien dan menentukan di banyak sektor termasuk pendidikan. Proyek ini adalah tentang analisis kinerja beberapa siswa untuk semester tertentu dengan menggunakan pertimbangan berbagai faktor mulai dari psikologis, pribadi, dan lingkungan, dll. Dan prediksi kinerja mereka untuk semester lain menggunakan bahasa pemrograman Python untuk ilmu data dan pembelajaran Mesin.[6] Ilmu data adalah bidang multidisiplin yang berurusan dengan data terstruktur dan tidak terstruktur, menggunakan metode, proses, algoritme, dan sistem ilmiah untuk mengekstrak pengetahuan dan wawasan dari data. Ilmu data terdiri dari bidang data mining dan big data.

Pembelajaran mesin adalah aspek Kecerdasan Buatan (AI) di mana sistem komputasi dapat belajar dari data dan membuat keputusan. Pembelajaran mesin ini telah menjadi salah satu bidang yang terpenting dalam organisasi pengembangan yang mencari cara inovatif untuk memahami aset data guna membantu bisnis mencapai tingkat pemahaman baru. Aplikasi pembelajaran mesin termasuk tetapi tidak terbatas pada deteksi penipuan, prediksi kegagalan peralatan, pengenalan pola dan gambar, dan lain lain[7][8].

Tujuan dari makalah ini adalah untuk menguji kinerja model prediktif yang berbeda pada kinerja siswa dan untuk secara akurat memprediksi nilai kursus siswa.

II. METODE PENELITIAN

Mengusulkan model prediktif untuk kinerja siswa dengan mengklasifikasikan siswa ke dalam kelas biner

(berhasil / tidak berhasil). Model yang diusulkan dibangun di bawah pendekatan penelitian CRISP-DM (Cross Industry Standard Process for Data Mining). Algoritma klasifikasi (OneR, J48, MLP dan IBK) diterapkan pada dataset yang diberikan. Hasil penelitian menunjukkan bahwa akurasi tertinggi dicapai oleh model MPL (73,59%) untuk identifikasi yang berhasil sementara tiga model lainnya berkinerja lebih baik untuk identifikasi siswa yang tidak berhasil.

Membangun model untuk memprediksi keberhasilan akademik siswa dalam suatu mata kuliah dengan mengurangi masalah dimensi data. Berbagai pengklasifikasi pembelajaran mesin seperti NB, MLP dan j48 dievaluasi dalam penelitian ini. Hasilnya menunjukkan bahwa Naïve Bayes memperoleh akurasi tertinggi 76,65%. Model yang diusulkan tidak menangani masalah ketidakseimbangan kelas.

Membahas model prediksi kegagalan siswa berdasarkan teknik pembelajaran mesin untuk menyelesaikan masalah ketidakseimbangan kelas dan dimensi data. Sepuluh pengklasifikasi diterapkan pada dataset. Pengklasifikasi ICRM mencapai akurasi tertinggi 92,7% di antara yang lainnya. Karena karakteristik siswa yang berbeda-beda di setiap tingkat pendidikan, kinerja model yang diusulkan tidak diuji untuk tingkat pendidikan lainnya

Tantangan EDM lainnya adalah memprediksi siswa yang putus sekolah dari mata kuliahnya. Empat metode data mining dengan enam kombinasi atribut diikutsertakan dalam penelitian ini. Hasil penelitian menunjukkan bahwa model support vector machine dengan kombinasi variabel prediktor lebih akurat dalam mengklasifikasikan sebuah data. yang Pencantuman atribut, nilai yang diperoleh dari mata kuliah prasyarat dalam kumpulan data adalah batasan studi ini karena mungkin saja selama studi mata kuliah apa pun, siswa mungkin akan dapat meningkatkan pengetahuannya tentang prasyarat mata kuliah. kursus ini.

Melakukan studi tentang prediksi kinerja siswa. Kontribusi utama dari penelitian ini adalah untuk memperkenalkan faktor sosial baru yang disebut "CAT" yang menggambarkan bahwa pada masa awal orang India dibagi menjadi empat jenis kelompok berdasarkan status sosial mereka dll, yang memiliki efek langsung pada siswa pendidikan. Empat classifiers oneR, MLP, J48, dan IB1 diterapkan pada kumpulan data. Hasil penelitian menunjukkan bahwa model IBI memiliki akurasi tertinggi (82%) yang dicapai.

Bangun versi model ID3 yang lebih baik, yang memprediksi kinerja akademik siswa. Kelemahan model ID3 adalah keinginannya untuk memilih

atribut-atribut tersebut sebagai node yang memiliki nilai lebih. Akibatnya pohon yang dihasilkan tidak efisien. Model yang diusulkan mengatasi masalah tersebut. Dua kelas keluaran dihasilkan oleh model ini (Lulus dan Gagal). Pengklasifikasi termasuk J48, wID3 dan Naïve Bayes diterapkan dan hasilnya dibandingkan. WID3 mencapai akurasi tinggi 93%

Metode

Pemilihan Model

Tiga (3) algoritma atau model regresi yang akan digunakan untuk tujuan proyek ini tercantum di bawah ini;

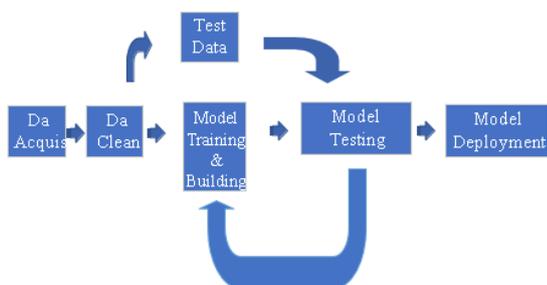
Model 1 - Regresi linier untuk pembelajaran terawasi yang dilaksanakan dengan scikit-learn.

Model 2 - Regresi linier untuk pembelajaran mendalam yang diimplementasikan dengan TensorFlow dan Keras.

Model 3 - Jaringan saraf untuk pembelajaran mendalam yang diimplementasikan dengan TensorFlow dan Keras.

Pelatihan dan Pembuatan Model

Sebelum kita melakukan pememilih algoritma, kita dapat menerapkan pendekatan praktis yang dapat digunakan untuk menyelesaikan sebagian besar dari masalah pembelajaran mesin yang menghasilkan suatu proses pembelajaran mesin seperti yang ditunjukkan pada Gambar 1 di bawah ini..



Gambar 1. Proses Pembelajaran Mesin

Analisis Data : Pada langkah ini dilakukan statistik deskriptif dengan bantuan visualisasi dan plot. Pembersihan.

Data : termasuk pra-pemrosesan, pembuatan profil, pembersihan.

Transformasi Data: melibatkan transformasi data dari keadaan mentah ke keadaan yang cocok untuk pemodelan adalah tempat rekayasa fitur cocok.

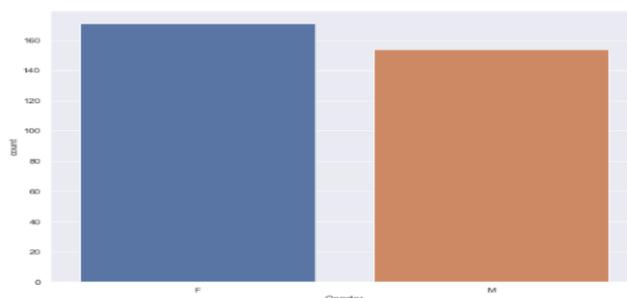
Pemilihan Model: Setelah mengkategorikan masalah dan memahami data, tonggak berikutnya adalah mengidentifikasi algoritma yang dapat diterapkan dan praktis untuk diimplementasikan dalam waktu yang wajar.

Implement the Model: Siapkan model machine learning yang menganalisis performa setiap algoritme pada kumpulan data menggunakan serangkaian kriteria evaluasi yang dipilih dengan cermat.

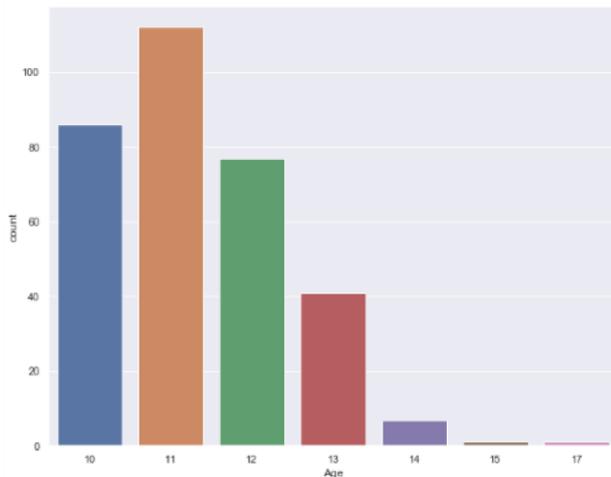
Pengumpulan Data dan Analisis Deskriptif

Data untuk proyek ini dikumpulkan dari Kaggle dengan 648 kumpulan data dan 22 atribut. Data dibagi menjadi 2 bagian dengan 325 set data untuk pelatihan berjudul Train.CSV dan 323 sebagai data uji bernama Test.CSV. Data ini menangkap kinerja siswa yang dipilih secara acak. Dataset kinerja siswa berisi dua puluh dua (22) faktor mulai dari psikologis, pribadi dan lingkungan. Faktor-faktor tersebut antara lain tingkat kehadiran siswa, jarak dari rumah ke sekolah, jam membaca, dukungan pendidikan, status kesehatan, tingkat pendidikan ayah dan ibu dan banyak lagi.

Gambar 2 dan 3 menunjukkan hasil analisis deskriptif yang dilakukan pada kumpulan data uji. Gambar 2 menyajikan plot distribusi jenis kelamin dari jumlah total sedangkan Gambar 3 menunjukkan distribusi usia siswa antara usia 10 sampai 17 dengan frekuensi tertinggi adalah 11 tahun.



Gambar 2. Plot Frekuensi untuk Jenis Kelamin Siswa



Gambar 3. Plot Frekuensi untuk Usia Siswa

Setelah analisis data kita selesai, langkah selanjutnya adalah Transformasi Data. Dari Gambar 4 dan 5 masing-masing menunjukkan sampel dataset kereta asli dan dataset kereta yang diubah. Gambar 4 menunjukkan dataset kereta untuk proyek ini yang menunjukkan semua parameter dari S/N hingga SCORE. Model akan dibangun dengan dataset ini, kemudian digunakan untuk membuat prediksi pada dataset uji sedangkan Gambar 5 adalah data latih yang ditransformasi mengubah semua nilai objek menjadi bilangan bulat.

S/N	Gender	Age	Location	Service	Status	Medu	Fedu	travel	study	recreation	school	business	paid	activities	nursery	higher	internet	laptop	freetime	health	absences	Score			
1	F	13	U	GTS	A	4	4	2	2	0	yes	no	no	no	yes	yes	no	4	3	3	4	22			
2	F	12	U	GTS	T	1	1	1	2	0	no	yes	no	no	no	yes	yes	5	3	3	2	31			
3	F	10	U	LEB	T	1	1	1	2	0	yes	no	no	no	yes	yes	yes	4	3	3	6	37			
4	F	10	U	GTS	T	4	2	1	3	0	no	yes	no	yes	yes	yes	yes	3	2	5	0	42			
5	F	11	U	GTS	T	3	3	1	2	0	no	yes	no	no	yes	yes	no	4	3	5	0	37			
6	M	11	U	LEB	T	4	3	1	2	0	no	yes	no	yes	yes	yes	yes	5	4	5	6	37			
7	M	11	U	LEB	T	2	2	1	1	0	no	no	no	no	yes	yes	yes	4	4	3	0	38			
8	F	12	U	GTS	A	4	4	2	2	0	yes	yes	no	no	yes	yes	no	4	1	1	2	36			
9	M	10	U	LEB	A	3	2	1	2	0	no	yes	no	no	yes	yes	yes	4	2	1	0	48			
10	M	10	U	GTS	T	3	4	1	1	0	no	yes	no	yes	yes	yes	yes	5	5	5	0	37			
11	F	11	U	GTS	T	4	4	1	2	0	no	yes	no	no	yes	yes	yes	5	5	5	0	42			
12	F	10	U	GTS	T	4	4	1	2	0	no	yes	no	no	yes	yes	yes	3	2	2	2	42			
13	F	10	U	GTS	T	2	1	3	3	0	no	yes	no	yes	yes	yes	yes	5	2	4	0	35			
14	M	10	U	GTS	T	4	4	1	1	0	no	yes	no	yes	yes	yes	yes	4	3	5	0	37			
15	M	10	U	LEB	T	4	4	1	1	0	no	yes	no	yes	yes	yes	yes	4	3	5	0	37			
16	M	10	U	GTS	A	2	2	1	3	0	no	yes	no	no	yes	yes	yes	4	5	3	0	43			
17	F	11	U	GTS	T	4	4	1	1	0	no	yes	no	no	yes	yes	yes	4	4	2	6	51			
18	F	11	U	GTS	T	4	4	1	3	0	no	yes	no	no	yes	yes	yes	1	1	1	3	2	100	40	
19	F	11	U	GTS	T	3	3	3	2	0	no	yes	no	no	yes	yes	yes	5	3	4	2	41			
20	F	12	U	GTS	T	3	2	1	1	3	0	no	yes	no	no	yes	yes	1	1	1	5	5	2	23	
21	M	11	U	GTS	T	4	3	1	1	0	no	no	no	no	yes	yes	yes	1	1	1	3	1	5	6	36
22	M	10	U	GTS	T	4	3	1	2	0	no	yes	no	no	yes	yes	yes	4	4	1	0	39			
23	M	10	U	GTS	A	4	4	1	1	0	no	yes	no	no	yes	yes	yes	1	1	1	5	4	5	0	35
24	M	11	U	GTS	T	4	2	1	2	0	no	no	no	no	yes	yes	yes	4	5	5	0	39			
25	M	11	U	LEB	T	2	2	2	2	0	no	yes	no	yes	yes	yes	yes	5	4	5	2	30			

Gambar 4. Contoh Kumpulan Data Kereta Asli

J	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
1	Gender	Age	rural	urban	Status	Medu	Fedu	travel	study	recreation	school	business	paid	activities	nursery	higher	internet	laptop	freetime	health	absences	Score		
2	1	0	13	0	1	0	4	4	2	2	0	1	0	0	0	1	1	1	0	4	3	3	4	22
3	2	0	12	0	1	1	1	1	1	2	0	0	1	0	0	0	1	1	1	5	3	3	2	31
4	3	0	10	0	1	1	1	1	1	2	0	1	0	0	1	1	1	1	1	4	3	3	6	37
5	4	0	10	0	1	1	4	2	1	3	0	0	1	0	1	1	1	1	1	3	2	5	0	42
6	5	0	11	0	1	1	3	3	1	2	0	0	1	0	0	1	1	1	0	4	3	5	0	37
7	6	1	11	0	1	1	4	3	1	2	0	0	1	0	1	1	1	1	1	5	4	5	6	37
8	7	1	11	0	1	1	2	2	1	2	0	0	0	0	0	0	1	1	1	4	4	3	0	38
9	8	0	12	0	1	0	4	4	2	2	0	1	1	0	0	1	1	1	0	4	1	1	2	36
10	9	1	10	0	1	0	3	2	1	2	0	0	1	0	0	1	1	1	1	4	2	1	0	48
11	10	1	10	0	1	1	3	4	1	2	0	0	1	0	1	1	1	1	1	5	5	5	0	37
12	11	0	10	0	1	1	4	4	1	2	0	0	1	0	0	1	1	1	1	3	3	2	2	42
13	12	0	10	0	1	1	2	1	3	3	0	0	1	0	1	1	1	1	1	5	2	4	0	35
14	13	1	10	0	1	1	4	4	1	1	0	0	1	0	1	1	1	1	1	4	3	5	0	37
15	14	1	10	0	1	1	4	3	2	2	0	0	0	0	1	1	1	1	1	5	4	3	0	37
16	15	1	10	0	1	0	2	2	1	3	0	0	1	0	0	1	1	1	1	4	5	3	0	43
17	16	0	11	0	1	1	4	4	1	1	0	0	1	0	1	1	1	1	1	4	4	2	6	51
18	17	0	11	0	1	1	4	4	1	3	0	0	1	0	1	1	1	1	1	3	2	2	100	40
19	18	0	11	0	1	1	3	3	3	2	0	1	1	0	1	1	1	1	0	5	3	4	2	41
20	19	1	12	0	1	1	3	2	1	1	3	0	1	1	1	1	1	1	1	5	5	5	2	23
21	20	1	11	0	1	1	4	3	1	1	0	0	0	0	1	1	1	1	1	3	1	5	6	36
22	21	1	10	0	1	1	4	3	1	2	0	0	0	0	0	1	1	1	1	4	4	1	0	39
23	22	1	10	0	1	1	4	4	1	1	0	0	1	1	0	1	1	1	1	5	4	5	0	35
24	23	1	11	0	1	1	4	2	1	2	0	0	0	0	1	1	1	1	1	4	5	5	0	39
25	24	1	11	0	1	1	2	2	2	2	0	0	1	0	1	1	1	1	1	5	4	5	2	30

Gambar 5. Contoh Kumpulan Data Kereta yang Ditransformasikan

Analisis regresi adalah salah satu bidang terpenting dalam statistik dan pembelajaran mesin. Masalah regresi biasanya memiliki satu variabel dependen yang kontinu dan tidak terbatas.

Regresi linier merupakan salah satu teknik regresi yang paling penting dan banyak digunakan karena kemudahannya dalam menginterpretasikan hasil. Ada hubungan linier antara beberapa variabel dependen pada himpunan variabel independen $x = (x_1, \dots, x_r)$, di mana adalah jumlah prediktor seperti yang ditunjukkan pada persamaan 1: $y = \beta_0 + \beta_1 x_1 + \dots + \beta_r x_r + \epsilon$

Metrik yang digunakan untuk pengukuran dan penentuan kinerja model yang digunakan adalah Mean Absolute Error (MAE), yang menggambarkan perbedaan rata-rata antara dua (2) variabel atas sampel uji di mana semua perbedaan individu memiliki bobot yang sama seperti yang dijelaskan secara matematis dalam persamaan 3.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_i - x|$$

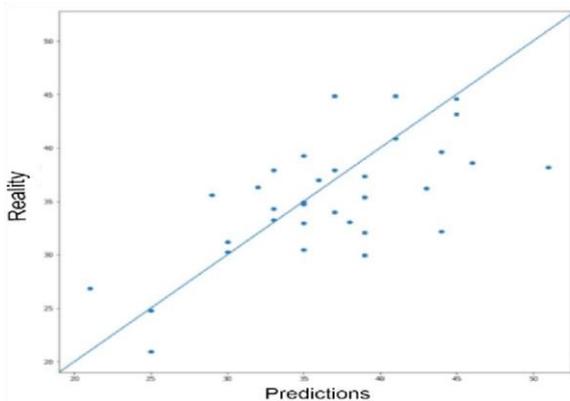
III. HASIL DAN PEMBAHASAN

a) Model 1 - Regresi linier untuk pembelajaran terawasi dengan scikit-belajar.

Semua 21 atribut dengan pengecualian "Skor" digunakan sebagai input (variabel independen) ke sistem sedangkan output (variabel dependen) adalah "Skor" untuk mendapatkan koefisien Beta dari dataset pelatihan. Setelah itu, koefisien beta digunakan

b) Model 2 - Model Regresi Linier untuk

Pembelajaran pada Mendalam dengan TensorFlow dan Keras



Gambar 6. Output kinerja regresi dan grafik regresi linier untuk Model 2

Gambar 6 menampilkan grafik regresi linier untuk model kedua yang memplot skor aktual terhadap skor prediksi dengan metrik: MAE = 4,61. Model tidak berkinerja lebih baik dari model pertama karena peningkatan MAE.

c) Model 3 - Model Jaringan Saraf Tiruan dengan Pembelajaran Mendalam

Untuk mengubah skala data kami, kami akan menggunakan fungsi MinMaxScaler dari Scikit-learn. Dengan `tf.contrib.learn` sangat mudah untuk mengimplementasikan Deep Neural Network. 5 lapisan tersembunyi dengan 200, 100, 50, 25 dan 12 unit masing-masing dan fungsi aktivasi Relu dengan tingkat pembelajaran 0,01 telah diterapkan. Pengoptimal yang digunakan adalah pengoptimal AdaDelta.

IV. SIMPULAN

Untuk mencapai prediksi kinerja siswa dengan kumpulan data berlabel baik yang menjadikannya masalah regresi dan data pembelajaran terawasi, beberapa algoritma regresi terbaik digunakan untuk memecahkan masalah. Regresi linier untuk pembelajaran terawasi dengan pembelajaran Sci-kit memberi kami model yang paling disukai dengan Mean Absolute Error.

Semua ini dicapai dengan dataset kami dari entri; namun, hasilnya bisa lebih baik jika kumpulan datanya jauh lebih besar, katakanlah memiliki ribuan entri lebih banyak atau lebih. Algoritme pembelajaran mesin mencoba menemukan pola dalam data yang membuatnya efisien dalam penggunaan kumpulan data

besar. Ini tidak berarti pembelajaran mesin tidak efektif pada kumpulan data kecil. Untuk penelitian lebih lanjut tentang kinerja siswa, saya merekomendasikan penggunaan data yang sangat besar, setidaknya 5.000 entri dataset. Data besar memberi ruang untuk lebih banyak dari data Kereta yang membantu algoritme menemukan pola dan membuat prediksi yang akurat.

DAFTAR PUSTAKA

- Rosyani, P., & Yunita, D. (2018). Pengembangan Aplikasi Bahan Ajar Kalkulus Berbasis Android. *Jurnal Informatika Universitas Pamulang*, 3(3), 118. <https://doi.org/10.32493/informatika.v3i3.2125>
- Yunita, D., Rosyani, P., & Amalia, R. (2018). Analisa Prestasi Siswa Berdasarkan Kedisiplinan, Nilai Hasil Belajar, Sosial Ekonomi dan Aktivitas Organisasi Menggunakan Algoritma Naïve Bayes. *Jurnal Informatika Universitas Pamulang*, 3(4), 209. <https://doi.org/10.32493/informatika.v3i4.2032>
- L. Cohen, L. Manion dan K. Morrison, *Metode Penelitian dalam Pendidikan*, edisi ke-6., Routledge: Oxon, Inggris, 2007.
- W. B. Ware dan J. P. Galassi, "Menggunakan Data Korelasi dan Prediksi untuk Meningkatkan Prestasi Siswa di Sekolah K-12: Aplikasi Praktis untuk Konselor Sekolah," *Konseling Sekolah Profesional*, vol. 9, tidak. 5, hlm. 344356, 2006.
- S. K. Yadav dan S. Pal, "Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification," *World of Computer Science and Information Technology Journal*, vol. 2, tidak. 2, hlm. 51-56, 2012.
- P. Cortez dan A. Silva, "Menggunakan Data Mining untuk Memprediksi Kinerja Siswa Sekolah Menengah".
- M. Pandey dan V. K. Sharma, "Algoritma Pohon Keputusan Berkaitan dengan Analisis dan Prediksi Kinerja Siswa," *Jurnal Internasional Aplikasi Komputer*, vol. 61, tidak. 13, hlm. 1-5, 2013.
- D. Kabakchieva, "Memprediksi Kinerja Siswa dengan Menggunakan Metode Data Mining untuk Klasifikasi," *Sibernetika dan Teknologi Informasi*, vol. 13, tidak. 1, hlm. 61-72, 2013.

- B. A. Kalejaye, O. Folorunsho dan O. L. Usman, "Memprediksi Nilai Nilai Siswa menggunakan Fungsi Pelatihan Jaringan Syaraf Tiruan," *Journal of Natural Science, Engineering and Technology*, vol. 14, tidak. 1, hlm. 25-42, 2015.
- S. Huang dan N. Fang, "Memprediksi kinerja akademik siswa dalam kursus dinamika teknik: Perbandingan empat jenis model matematika prediktif," *Komputer & Pendidikan*, vol. 61, hlm.