

ALGORITMA REGRESI LINEAR SEDERHANA DALAM MEMPREDIKSI STATUS MAHASISWA STIKOM TUNAS BANGSA PEMATANG SIANTAR

Selli Oktaviani¹, Solikhun², M.Safii³

^{1,2,3} STIKOM Tunas Bangsa Pematang Siantar, Sumatera Utara, Indonesia
e-mail: ¹Sellyoktaviani585@gmail.com, ²solikhun@amiktunasbangsa.ac.id ,
³m.safii@amiktunasbangsa.ac.id

Abstrak

STIKOM Tunas Bangsa Pematangsiantar merupakan salah satu perguruan tinggi IT yang terkemuka di wilayah Pematang Siantar. Pada saat ini, memiliki jumlah 590 mahasiswa aktif. Jumlah ini meningkat dibandingkan dengan tahun ajaran sebelumnya yang hanya berjumlah 537 mahasiswa. Ini yang membuktikan bahwa STIKOM Tunas Bangsa Pematang Siantar semakin diminati oleh calon mahasiswa. Dari jumlah mahasiswa yang begitu besar, STIKOM Tunas Bangsa Pematang Siantar masih mengalami kesulitan dalam memprediksi status mahasiswa yang terdiri dari mahasiswa aktif dan mahasiswa yang tidak aktif. Dari ke dua status tersebut dibutuhkan metode yang dapat menggali informasi dari data tersebut. Data mining dapat melakukan proses analisa data dan memprediksi data untuk menemukan hasilnya dari data mahasiswa tersebut. Tujuan dari penelitian ini yaitu penulis mencoba menerapkan metode algoritma regresi linear sederhana dalam memprediksi status mahasiswa untuk tahun ajaran selanjutnya yang mungkin saja menjadi bahan pertimbangan solusi kepada pihak akademik Pendidikan di STIKOM Tunas Bangsa kota Pematang Siantar. Dalam kasus ini analisa data mining dilakukan dengan menggunakan algoritma regresi linear sederhana. Solusi yang diusulkan untuk menyelesaikan masalah yang ada saat ini adalah dengan menerapkan algoritma regresi linear sederhana dalam memprediksi status mahasiswa STIKOM Tunas Bangsa Pematang Siantar untuk tahun ajaran yang akan datang. Hasil yang didapatkan yaitu pada sample data tahun 2023 jumlah mahasiswa yang aktif 532 dan prediksi mahasiswa yang non aktif ada 42 mahasiswa, pada sample data tahun 2024 jumlah mahasiswa yang aktif 533 dan prediksi mahasiswa yang non aktif ada 42 mahasiswa, pada sample data tahun 2025 jumlah mahasiswa yang aktif 571 dan prediksi mahasiswa yang non aktif ada 45 mahasiswa, pada sample data tahun 2026 jumlah mahasiswa yang aktif 523 dan prediksi mahasiswa yang non aktif ada 41 mahasiswa, pada sample data tahun 2027 jumlah mahasiswa yang aktif 498 dan prediksi mahasiswa yang non aktif ada 40 mahasiswa.

Kata kunci: Data Mining, Regresi Linear Sederhana, Prediksi, Status Mahasiswa.

Abstract

STIKOM Tunas Bangsa Pematangsiantar is one of the leading IT universities in the Pematangsiantar area. Currently, it has 590 active students. This number has increased compared to the previous school year which only amounted to 537 students. This proves that STIKOM Tunas Bangsa Pematang Siantar is increasingly in demand by prospective students. From such a large number of students, STIKOM Tunas Bangsa Pematang Siantar still experiences difficulties in predicting the status of students consisting of active students and inactive students. From these two statuses, a method is needed that can extract information from the data. Data mining can carry out the process of analyzing data and predicting data to find results from student data. The aim of this research is that the author tries to apply a simple linear regression method in predicting student status for the next academic year which might be

a material consideration for solutions for educational academics at STIKOM Tunas Bangsa, Pematang Siantar city. In this case, data mining analysis was carried out using a regression algorithm. simple linear. The proposed solution to solve the current problem is to apply a simple linear regression algorithm in predicting the status of STIKOM Tunas Bangsa Pematang Siantar students for the next academic year. The results obtained are that in the 2023 data sample the number of active students is 532 and the predicted non-active students are 42 students, in the 2024 data sample the number of active students is 533 and the predicted non-active students are 42 students, in the 2025 data sample the number there are 571 active students and the predicted number of non-active students is 45 students, in the 2026 data sample the number of active students is 523 and the predicted non-active students are 41 students, in the 2027 data sample the number of active students is 498 and the predicted non-active students there are 40 students.

Keywords: Data Mining, Simple Linear Regression, Prediction, Student Status.

1. Pendahuluan

Perkembangan dunia pendidikan selalu disertai dengan berkembangnya teknologi informasi. Kebutuhan akan sumber daya manusia yang memiliki kompetensi menjadi hal prioritas bagi dunia pendidikan sebagai sumber keluaran atau output bagi para peserta didik. Bagi perguruan tinggi swasta, jumlah mahasiswa dan kualitas pendidikan adalah hal yang sangat penting mengingat jumlah Perguruan Tinggi Swasta di Indonesia termasuk yang paling banyak di Asia. Di Sumatera Utara jumlah Perguruan Tinggi Swasta sebanyak 362 dibawah naungan KOPERTIS (Koordinasi Perguruan Tinggi Swasta) Wilayah I SUMUT-Aceh sumber (dikti.go.id). Terdapat dua point penting tersebut tidak boleh dilupakan bagi manajemen pendidikan dalam merancang strategi pendidikan swasta. Khususnya pada kampus STIKOM Tunas Bangsa Pematang Siantar dalam memberikan pelayanan yang baik dan benar bagi mahasiswa. Jumlah mahasiswa dan kualitas pendidikan merupakan dua unsur yang penting bagi keberlangsungan di Perguruan Tinggi Swasta.

STIKOM Tunas Bangsa Pematangsiantar merupakan salah satu perguruan tinggi IT yang terkemuka di wilayah Pematang Siantar. STIKOM Tunas Bangsa Pematang Siantar didirikan pada tahun 2014 oleh Yayasan Muhammad Nasir. Pada saat ini, memiliki jumlah 590 mahasiswa aktif. Jumlah ini meningkat dibandingkan dengan tahun ajaran sebelumnya yang hanya berjumlah 537 mahasiswa. Ini yang membuktikan bahwa STIKOM Tunas Bangsa Pematang Siantar semakin diminati oleh calon mahasiswa. Akan tetapi, dari jumlah 590 mahasiswa tersebut tidak seluruhnya menyelesaikan sampai tahap akhir perkuliahan. Hal ini menjadi kendala bagi perguruan tinggi swasta di dalam melaksanakan proses belajar mengajar.

Dari jumlah mahasiswa yang begitu besar, STIKOM Tunas Bangsa Pematang Siantar masih mengalami kesulitan dalam memprediksi status mahasiswa yang terdiri dari mahasiswa aktif dan mahasiswa yang tidak aktif. Dari ke dua status tersebut dibutuhkan metode yang dapat menggali informasi dari data tersebut. Data mining dapat melakukan proses analisa data dan memprediksi data untuk menemukan hasilnya dari data mahasiswa tersebut. Dalam kasus ini analisa data mining dilakukan dengan menggunakan algoritma regresi linear sederhana.

Solusi yang diusulkan untuk menyelesaikan masalah yang ada saat ini adalah dengan menerapkan algoritma regresi linear sederhana dalam memprediksi status mahasiswa STIKOM Tunas Bangsa Pematang Siantar untuk tahun ajaran yang akan datang. Berdasarkan latar belakang yang telah dijabarkan, maka diambil judul : **“Algoritma Regresi Linear Sederhana Dalam Memprediksi Status Mahasiswa Stikom Tunas Bangsa Pematang Siantar”**.

Istilah Data Mining dan *knowledge discovery in database* (KDD) sering kali digunakan untuk mencari informasi yang tersembunyi pada sebuah database yang besar. KDD akan membantu untuk mengidentifikasi dan mengamati hal-hal yang ingin diketahui, dan mengolah data menjadi informasi dan berkaitan satu sama lain. Proses KDD salah satunya adalah Data Mining [1].

Data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Data mining merupakan salah satu langkah dalam *knowledge discovery in database* (KDD) [1]. Beberapa penelitian terkait mengenai data mining yang telah dicoba maupun dibuat yaitu akan dijabarkan pada penelitian ini.

Menurut hasil pada penelitian Selly Anastassia Amellia Kharis dan Arman Haqqi Anna Zili yang berjudul “*Learning Analytics dan Educational Data Mining pada Data Pendidikan*” yaitu *Learning Analytics dan Educational Data Mining* memiliki potensi untuk dikembangkan pada data pendidikan di Indonesia. *Learning Analytics dan Educational Data Mining* dapat memberikan informasi kepada institusi, pendidik, dan peserta didik sehingga mendukung analisis prediksi dan pada akhirnya dapat meningkatkan motivasi, kinerja dan hasil dari suatu proses pembelajaran. Dengan melacak, menggabungkan, dan menganalisis digital footprint peserta didik, jalur baru untuk kebijakan pendidikan hadir secara lebih terbuka sesuai dengan pertumbuhan database peserta didik [2].

Hasil pada penelitian Sanni Ucha Putri, Eka Irawan dan Fitri Rizky yang berjudul “Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5” yaitu Hasil dari penelitian ini akan dijadikan sebagai acuan untuk dapat melihat apakah seseorang beresiko terkena penyakit diabetes atau tidak berdasarkan atribut yang telah ditetapkan. Berdasarkan pengolahan data menggunakan software RapidMiner didapat nilai akurasi sistem sebesar 90,00%, artinya bahwa rule yang dihasilkan tingkat kebenaran mendekati 100%. Dimana Jumlah benar prediksi positif sebanyak 31 data, salah prediksi positif sebanyak 3 data, jumlah benar prediksi negatif sebanyak 2 data sedangkan salah prediksi negatif sebanyak 13 data. Prediksi positif dari seluruh data menghasilkan 90,00% seseorang terkena penyakit diabetes [3]

Hasil pada penelitian Fidya Arie Pratama, Riri Narasati dan Dita Rizki Amalia yang berjudul “Pengaruh Kata Cashback Terhadap Peningkatan Penjualan Menggunakan Data Mining” yaitu Peneliti menggunakan Algoritma K-Means sebagai metode penelitian dan menggunakan aplikasi Rapid Minner untuk menganalisis data. Hasil penelitian menunjukkan kata cashback meningkatkan tingkat penjualan ganda untuk perusahaan produk kecantikan [4].

Data mining sendiri memiliki beberapa teknik yang sering digunakan antara lain yaitu sebagai berikut :

a. *Clustering*

Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan *cluster*. Objek yang di dalam cluster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan cluster yang lain. Partisi tidak dilakukan secara maual melainkan dengan suatu algoritma *clustering*. Oleh karena itu, *clustering* sangat berguna dan bisa menemukan group atau kelompok yang tidak dikenal dalam data. *Clustering* banyak digunakan dalam berbagai aplikasi seperti misalnya pada *business intelligence*, pengenalan pola citra, *web search*, bidang ilmu biologi, dan untuk keamanan (*security*) [5].

b. *Association Rule*

Association Rule adalah teknik data mining yang digunakan untuk menemukan aturan *associative* antara suatu kombinasi item. Salah satu contoh penerapan *Association Rule* adalah *Market Basket Analysis* [6].

c. *Classification*

Classification merupakan suatu proses untuk menemukan model atau fungsi untuk menggambarkan *class* atau konsep dari suatu data. Proses yang digunakan untuk mendeskripsikan data yang penting serta dapat meramalkan kecenderungan data pada masa depan [7].

d. *Description*

Description salah satu kegunaan data mining untuk menggali dan mengumpulkan data yang banyak. Metode ini berguna untuk kateristik atau menggambarkan suatu kondisi, sejarah, objek [8].

e. Prediction

Prediksi adalah sama dengan ramalan atau perkiraan. Sedangkan menurut Kamus Besar Bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan. Prediksi bisa berdasarkan metode ilmiah ataupun subjektif belaka. Berbeda dengan pengertian prediksi secara bahasa yang berarti lamaran atau perkiraan yang sudah menjadi pengertian yang baku [9].

f. Estimation

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi [10].

2. Metode Penelitian

2.1 Metode Pengumpulan Data

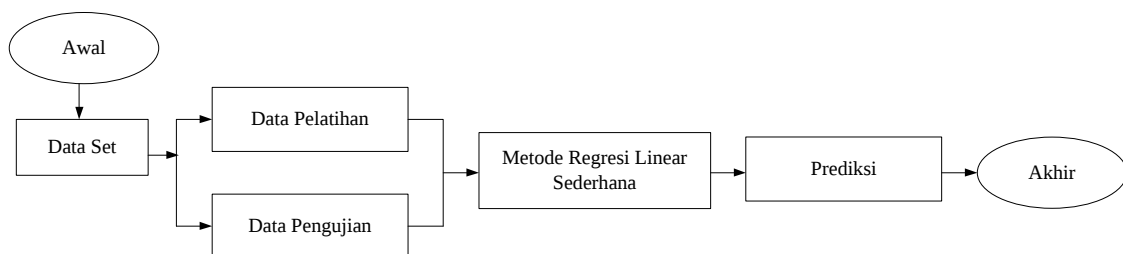
Metode pengumpulan beberapa data yang dilakukan oleh peneliti pada penelitian ini adalah diambil di kantor bagian akademik pendidikan STIKOM Tunas Bangsa Pematang Siantar.

2.2 Sumber Data

Sumber data yang digunakan pada penelitian ini yaitu diambil dari kantor bagian akademik pendidikan yang ada di STIKOM Tunas Bangsa Pematang Siantar. Data ini berisi data mahasiswa yang aktif dan yang tidak aktif. Data ini digunakan untuk memprediksi status mahasiswa untuk tahun ajaran yang akan datang. Data yang dipergunakan peneliti untuk *training* yaitu dari tahun 2015 sampai dengan 2022 dan data yang dipergunakan untuk *testing* yaitu sampel data tahun 2023 sampai dengan 2027.

2.3 Kerangka Kerja Penelitian

Dalam penyelesaian penelitian ini, peneliti menyusun kerangka kerja penelitian seperti pada gambar bagan dibawah ini:



Gambar 1. Kerangka Kerja Penelitian

Yaitu dimulai dari tahap awal, mengumpulkan data set, membagi data dengan data pelatihan dan data pengujian, menemukan metode yaitu metode regresi linear sederhana, memprediksinya sampai dengan tahap akhir.

3. Hasil dan Pembahasan

Pada penelitian ini akan membahas teknik estimasi yang diterapkan untuk memprediksi status mahasiswa yang aktif dan *non* aktif terutama pada kampus STIKOM Tunas Bangsa kota Pematang Siantar. Teknik estimasi yang akan digunakan adalah Regresi Linear Sederhana. Metode ini dipilih karena mampu memprediksi status mahasiswa yang aktif dan yang *non* aktif dengan memanfaatkan data data yang ada pada ruangan akademik pendidikan STIKOM Tunas Bangsa kota Pematang Siantar.

Beberapa penelitian terkait mengenai Regresi Linear Sederhana akan dilampirkan pada penelitian ini. Menurut hasil penelitian yang berjudul “Peramalan Penerimaan Mahasiswa Baru Universitas Samudra Menggunakan Metode Regresi Linear Sederhana” oleh [11] Hasil yang didapat pada penelitian ini adalah penggunaan metode regresi linear sederhana dapat

dipertimbangkan karena jumlah error yang didapat pada hasil prediksi penerimaan mahasiswa baru 1 tahun kedepan tidak terlalu besar yaitu sebesar 21 dari 1428 mahasiswa hasil prediksi, dimana data ini di uji pada data tahun 2018 jika diprediksi pada tahun tersebut, banyak mahasiswa yang masuk pada tahun tersebut adalah 1449.

Menurut hasil penelitian yang berjudul “Peramalan Jumlah Siswa-Siswi Melalui Pendekatan Metode Regresi Linear” oleh [12] Berdasarkan pengujian yang telah dilakukan, dapat disimpulkan bahwa fungsional dari aplikasi telah berjalan sebagaimana mestinya dan aplikasi mampu menampilkan hasil peramalan berdasarkan data trend yang diinputkan.

Menurut hasil penelitian yang berjudul “Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear” oleh [13] Hasil prediksi pada 3 periode mendatang setiap jenis produk menghasilkan nilai yang sama antara hasil sistem dengan hasil perhitungan manual regresi linear. Pengujian tingkat kesalahan terhadap hasil prediksi 3 periode yaitu bulan Januari sampai dengan Maret 2022 setiap produk Axe Deodorant, Bango Kecap, Buavita, Citra Lotion, Sabun Citra, Shampo Clear, Sariwangi, Sunsilk Conditioner, Vixal dan Wall’s Ice Cream tergolong kategori hasil peramalan sangat akurat. Dengan nilai MAPE terkecil pada produk Sunsilk Conditioner sebesar 1%. Dengan demikian, metode regresi linear sangat akurat untuk prediksi penjualan barang jenis unilever.

Regresi linear sederhana adalah sebuah metode pendekatan untuk pemodelan hubungan antara satu variabel dependen dan satu variabel independen. Dalam regresi, variabel independen menerangkan variabel dependennya. Dalam analisis regresi sederhana, hubungan antar variabel bersifat linear, dimana perubahan pada variabel X akan diikuti oleh perubahan Variabel Y secara tetap. Sementara pada hubungan non linear, perubahan variabel X tidak diikuti variabel Y secara proposional [14].

Bentuk umum persamaan regresi linear sederhana yaitu :

$$Y = a + bX$$

Keterangan :

Y = Variabel Dependen

X = Variabel Independen

a = Konstanta

b = Gradien Garis

Persamaan regresi linear yaitu :

$$a = \frac{\sum_{i=1}^n y_i}{n} - b \frac{\sum_{i=1}^n x_i}{n}$$

$$b = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}$$

Gambar 2. Rumus Persamaan Regresi Linear

3.1 Perhitungan Manual Regresi Linear Sederhana

Sebelum menghitung manual regresi sederhana, sebaiknya menyiapkan data set yang akan dihitung. Berikut ini adalah tabel data set jumlah mahasiswa aktif dan non aktif di STIKOM Tunas Bangsa Pematang Siantar.

Tabel 1. Data Set Mahasiswa

Tahun	Mahasiswa aktif	Mahasiswa non aktif
2015	198	14
2016	232	25

Tahun	Mahasiswa aktif	Mahasiswa non aktif
2017	376	32
2018	437	41
2019	508	37
2020	548	54
2021	537	42
2022	590	35

Dibawah ini adalah tabel untuk perhitungan manual pada regresi linear sederhana. Dimulai dari data mahasiswa aktif sebagai variabel independen (x) dan data mahasiswa non aktif sebagai variabel dependen (y). Setelah itu, menghitung seperti yang tertera pada tabel yang dibawah ini.

Tabel 2. Tabel Perhitungan Manual Regresi Linear Sederhana

Tahun	Mahasiswa aktif (x)	Mahasiswa non aktif (y)	xy	x ²	y ²
2015	198	14	2772	39204	196
2016	232	25	5800	53824	625
2017	376	32	12032	141376	1024
2018	437	41	17917	190969	1681
2019	508	37	18796	258064	1369
2020	548	54	29592	300304	2916
2021	537	42	22554	288369	1764
2022	590	35	20650	348100	1225
Σ	3426	280	130113	1620210	10800

Tahapan atau proses yang harus dilakukan untuk melakukan perhitungan manual regresi linear sederhana yaitu hal yang pertama dilakukan adalah dengan cara menentukan nilai b atau gradien garis, yang kedua adalah menentukan nilai a atau nilai konstanta.

1. Menentukan nilai b (gradien garis):

$$b = \frac{n \sum x_i y_i - \left(\sum x_i \right) \left(\sum y_i \right)}{n \sum x_i^2 - \left(\sum x_i \right)^2}$$

$$b = \frac{(\text{jumlah index dikali } \Sigma xy) - (\Sigma \text{variabel independen}(x) \text{ dikali } \Sigma \text{variabel dependen}(y))}{(\text{jumlah index dikali } \Sigma x^2) - (\Sigma \text{variabel independen}(x))^2}$$

$$b = \frac{(8 \times 130113) - (3426 \times 280)}{(8 \times 1620210) - (3426)^2}$$

$$b = \frac{(1040904) - (959280)}{(12961680) - (11737476)}$$

$$b = \frac{81624}{1224204}$$

$$b = 0,0066772446$$

2. Menentukan nilai a (konstanta) :

$$a = \frac{\sum y_i}{n} - b \frac{\sum x_i}{n}$$

$$a = \frac{280}{8} - ((0,0066772446) * \frac{3426}{8})$$

$$a = 35 - ((0,0066772446) * 428,25)$$

$$a = 35 - 28,55363812$$

$$a = 6,446361881$$

Setelah diperoleh nilai a dan b, maka dilakukan perhitungan persamaan regresi linear:

$$Y = a + bX$$

$$Y = 6,446361881 + 0,066675162 X$$

Setelah persamaan regresi linear diperoleh, maka hasil prediksi status mahasiswa yang non aktif dapat diperoleh dengan memasukkan nilai X pada periode tahun 2023 dengan sampel mahasiswa yaitu nilai X sebanyak 532 mahasiswa. Maka, perhitungan prediksi status mahasiswa adalah dengan menggunakan persamaan regresi linear sederhana seperti cara yang diatas, maka:

$$Y = a + bX$$

$$Y = 6,446361881 + 0,066675162 (532)$$

$$Y = 6,446361881 + 35,47119$$

$$Y = 42$$

Maka hasil dari prediksi jika jumlah mahasiswa aktif 2023 adalah 532 maka jumlah prediksi mahasiswa non aktif adalah 42 mahasiswa. Berikut ini hasil prediksi status mahasiswa yang ada di STIKOM Tunas Bangsa Pematang Siantar untuk tahun 2023 sampai dengan 2027.

Tabel 3. Hasil Prediksi Perhitungan Manual Status Mahasiswa

Tahun	Mahasiswa aktif	Mahasiswa non aktif
2023	532	42
2024	533	42
2025	571	45
2026	523	41
2027	498	40

Sumber: sampel data prediksi status mahasiswa 5 tahun yang akan datang

3.2 Perhitungan Menggunakan Rapid Miner Regresi Linear Sederhana

Untuk menghitung prediksi di aplikasi rapid miner, ada beberapa hal yang harus dilakukan.

1. Menentukan data training

Data training diperoleh dari data set mahasiswa aktif dan non aktif yang ada di STIKOM Tunas Bangsa pada tahun 2015, 2016, 2017, 2018, 2019, 2020, 2021 dan 2022.

Tabel 4. Data Training

Tahun	Mahasiswa aktif (x)	Mahasiswa non aktif (y)
2015	198	14
2016	232	25
2017	376	32
2018	437	41
2019	508	37
2020	548	54
2021	537	42
2022	590	35

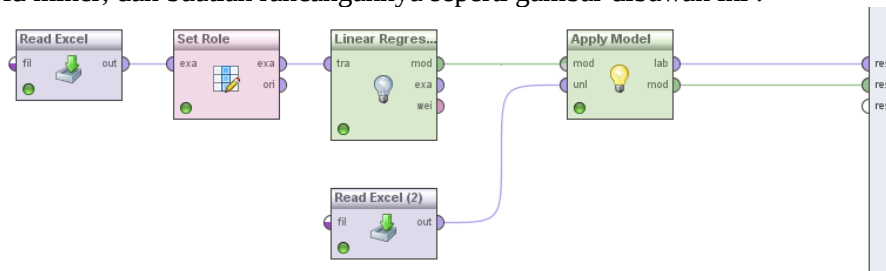
Data testing diperoleh dari data sample yang akan di prediksi untuk tahun ajaran 5 tahun yang akan datang. Jadi, dari data ini kita akan mengetahui labelnya atau hasil prediksinya per 5 tahun yang akan datang.

Tabel 5. Data Testing

Tahun	Mahasiswa aktif (x)	Mahasiswa non aktif (y)
2023	532	?
2024	533	?
2025	571	?
2026	523	?
2027	498	?

2. Rancangan Rapid Miner

Rancangan rapid miner ini digunakan untuk mengatur prediksi yang akan dibuat ataupun dilakukan. Cara membuat model rancangan rapid miner nya yaitu dengan cara buka aplikasi rapid miner, dan buatlah rancangannya seperti gambar dibawah ini :



Gambar 3. Bagan Perhitungan Rapid Miner

Bagan untuk perhitungan di rapid miner yaitu terdiri dari :

1. Read excel 1 untuk data training
2. Read excel 2 untuk data testing
3. Set Role
4. Linear Regression
5. Apply Model

Hasil prediksi yang akan muncul setelah di jalankan pada aplikasi rapid miner versi 5 ini yaitu seperti gambar dibawah ini :

☒ Data View ☐ Meta Data View ☐ Plot View ☐ Advanced Charts ☐ Annotations

ExampleSet (5 examples, 2 special attributes, 2 regular attributes)

Row No.	non aktif	prediction(n...	tahun	aktif
1	?	41.918	2023	532
2	?	41.984	2024	533
3	?	44.518	2025	571
4	?	41.317	2026	523
5	?	39.651	2027	498

Gambar 4. Rumus Perhitungan Rapid Miner

Berdasarkan gambar 4. dapat dilihat hasil dari prediksi status mahasiswa yang ada di STIKOM Tunas Bangsa Pematang Siantar tahun 2023 sampai dengan 2027. Rumus untuk menghitung prediksi status mahasiswa STIKOM Tunas Bangsa Pematang Siantar dapat dilihat pada gambar dibawah ini.

☐ Table View ☒ Text View ☐ Annotations

LinearRegression

$$0.067 * \text{aktif} + 6.446$$

Gambar 5. Rumus Perhitungan Rapid Miner

Berdasarkan gambar 5. yaitu rumus perhitungan prediksi rapid miner, maka akan dilakukannya perhitungan prediksi. Setelah mengetahui rumus untuk menghitung prediksi yang ada di rapid miner, maka hasil yang di dapat yaitu seperti pada tabel dibawah ini. Hasil yang didapat untuk prediksi 5 tahun ajaran yang akan datang.

Tabel 6. Hasil Perhitungan Dengan Menggunakan Rumus di Rapid Miner

Tahun	Mahasiswa Aktif (x)	Mahasiswa Non Aktif (y)	Prediksi
2023	532	?	42
2024	533	?	42
2025	571	?	45

Tahun	Mahasiswa Aktif (x)	Mahasiswa Non Aktif (y)	Prediksi
2026	523	?	41
2027	498	?	40

4. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dijabarkan atau dideskripsikan, maka dapat diambil kesimpulan bahwa regresi linear sederhana dapat digunakan untuk memprediksi status mahasiswa/i di STIKOM Tunas Bangsa kota Pematang Siantar. Hasil yang didapatkan yaitu pada sample data tahun 2023 jumlah mahasiswa yang aktif 532 dan prediksi mahasiswa yang non aktif ada 42 mahasiswa, pada sample data tahun 2024 jumlah mahasiswa yang aktif 533 dan prediksi mahasiswa yang non aktif ada 42 mahasiswa, pada sample data tahun 2025 jumlah mahasiswa yang aktif 571 dan prediksi mahasiswa yang non aktif ada 45 mahasiswa, pada sample data tahun 2026 jumlah mahasiswa yang aktif 523 dan prediksi mahasiswa yang non aktif ada 41 mahasiswa, pada sample data tahun 2027 jumlah mahasiswa yang aktif 498 dan prediksi mahasiswa yang non aktif ada 40 mahasiswa. Dan juga dapat disimpulkan baik perhitungan manual maupun menggunakan rapid miner, hasil yang didapatkan sama.

Beberapa saran yang dapat penulis berikan yang nantinya dapat digunakan untuk mengembangkan penelitian adalah sebagai berikut :

1. Berdasarkan hasil penelitian di STIKOM Tunas Bangsa Pematang Siantar penulis memberikan saran agar diadakan penelitian lebih lanjut mengenai prediksi status mahasiswa menggunakan metode pembandingan lain atau juga bisa menggunakan Sistem Informasi.
2. Untuk penelitian selanjutnya diharapkan menggunakan data yang lebih spesifik lagi agar dapat memaparkan variabel yang lebih jelas.

Daftar Pustaka

- [1] D. Winarti *et al.*, “Penerapan Data Mining untuk Analisa Tingkat Kriminalitas Dengan Algoritma Association Rule Metode FP-Growth,” *J. SIMTIKA*, vol. 4, no. 3, 2021.
- [2] S. Anastassia, A. Kharis, A. Haqqi, and A. Zili, “Learning Analytics dan Educational Data Mining pada Data Pendidikan,” vol. 6, pp. 12–20, 2022.
- [3] F. R. Sanni Ucha Putri, Eka Irawan, “Implementasi Data Mining Untuk Prediksi Penyakit Diabetes,” vol. 2, no. 1, pp. 39–46, 2021.
- [4] F. A. Pratama, R. Narasati, and D. R. Amalia, “Pengaruh Kata Cashback Terhadap Peningkatan Penjualan Menggunakan Data Mining,” vol. 3, no. 2, pp. 1–5, 2019.
- [5] K. Annisa, B. S. Ginting, and M. A. Syari, “Penerapan Data Mining Pengelompokan Data Pengguna Air Bersih Berdasarkan Keluhannya Menggunakan Metode Clustering Pada PDAM Langkat,” vol. 6341, no. April, 2022.
- [6] M. Safii, “Implementasi Data Mining Dengan Metode Pohon Keputusan Algoritma Id3 Untuk Menentukan Status Mahasiswa,” vol. 2, no. 1, pp. 82–86, 2018.
- [7] M. F. D. Nurul Indah Prabawati, Widodo, “Kinerja Algoritma Classification and Regression Tree (Cart) dalam Mengklasifikasikan Lama Masa Studi Mahasiswa yang Mengikuti Organisasi di Universitas Negeri Jakarta Available at : Available at :,” vol. 3, no. 2, pp. 139–145, 2019.
- [8] R. Purba, N. A. Hasibuan, and E. Hatmi, “Implementasi Data Mining Menggunakan Metode Deskripsi Untuk Mengetahui Pola Penentuan Penerima Bantuan Siswa Miskin (BSM) Di Smp N 3 Doloksanggul,” vol. 3, pp. 493–498, 2019, doi: 10.30865/komik.v3i1.1632.
- [9] M. D. Wahyudi, “Penerapan data mining dengan algoritma c4.5 dalam prediksi penjualan buku,” vol. 1, no. 1, pp. 1–6, 2023.
- [10] M. Idris, “Implementasi Data Mining Dengan Algoritma Naïve Bayes Untuk Memprediksi Angka Kelahiran,” vol. 7, pp. 421–428, 2019.
- [11] A. Yordan, T. N. Putri, and D. H. Lamkaruna, “Peramalan Penerimaan Mahasiswa Baru Universitas Samudra Menggunakan Metode Regresi Linear Sederhana,” *J. Tek. Inform.*,

- vol. 2, no. 1, pp. 21–27, 2019, doi: 10.52046/j-tifa.v2i1.237.
- [12] D. A. Trianggana, “a Peramalan Jumlah Siswa-Siswi Melalui Pendekatan Metode Regresi Linear,” *J. Media Infotama*, vol. 16, no. 2, pp. 115–120, 2020, doi: 10.37676/jmi.v16i2.1149.
- [13] A. Anggrawan and N. Azmi, “Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear Sales Prediction of Unilever Products using the Linear Regression Method,” vol. 4, no. 2, pp. 123–132, 2022, doi: 10.30812/bite.v4i2.2416.
- [14] P. S. Ramadhan and N. Safitri, “Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang,” vol. 18, no. 1, pp. 55–61, 2019.