



PREDIKSI JUMLAH MAHASISWA UJIAN SKRIPSI DENGAN METODE LEAST SQUARE

Amiruddin Bengnga¹, Rezqiwati Ishak²

^{1,2} Teknik Informatika, Universitas Ichsan Gorontalo

Article Info:

Dikirim: 29 Mei 2021

Direvisi: 20 Juni 2021

Diterima: 29 Juni 2021

Tersedia Online: 30 Juni 2021

Penulis Korespondensi:

Amiruddin

Universitas Ichsan, Gorontalo,
Indonesia

Email: amier.76@gmail.com

Abstrak: Pencapaian target jumlah mahasiswa yang ujian skripsi di setiap semester pada setiap Program Studi yang sudah ditentukan di awal semester tentunya ada yang mencapai target dan tidak mencapai target. Jika tidak mencapai target maka hal ini akan menjadi penilaian kinerja Program Studi menurun, agar hal ini tidak terjadi, maka salah satu solusinya adalah melakukan teknik prediksi dengan menggunakan data yang relevan di periode semester sebelumnya. Metode prediksi yang digunakan adalah metode Least Square karena metode ini cocok digunakan untuk memprediksi data dalam bentuk time series. Data yang digunakan dalam penelitian ini diambil dari 5 Program Studi di Universitas Ichsan Gorontalo. Berdasarkan hasil penelitian perhitungan tingkat kesalahan dengan menggunakan MAPE (Mean Absolute Percentage Error) antara data aktual dan hasil prediksi didapatkan rata-rata error sebesar 16.11% atau tingkat akurasi sebesar 83.89%.

Kata kunci: least square; prediksi; ujian skripsi; MAPE.

Abstract: Achieving the target number of students who have their thesis examination in each semester in each Study Program that has been determined at the beginning of the semester, of course there are those who reach the target and do not reach the target. If it does not reach the target, this will decrease the Study Program's performance assessment, so that this does not happen, then one solution is to carry out prediction techniques using relevant data in the previous semester period. The prediction method used is the Least Square method because this method is suitable for predicting data in the form of time series. The data used in this study were taken from 5 Study Programs at Ichsan University, Gorontalo. Based on the research results, the calculation of the error rate using MAPE (Mean Absolute Percentage Error) between the actual data and the prediction results obtained an average error of 16.11% or an accuracy rate of 83.89%

Keywords: least square; prediction; thesis examination; MAPE.

1. PENDAHULUAN

Keberhasilan suatu Program Studi salah satu tolak ukurnya adalah banyaknya mahasiswa yang ujian skripsi tepat pada waktunya yaitu pada semester delapan, hal ini bisa dicapai tentunya banyak faktor yang harus diperhatikan diantaranya struktur kurikulum, sistem perkuliahan, Dosen pembimbing dan faktor lainnya. Setiap Program Studi setiap semesternya sudah menentukan target jumlah mahasiswa yang akan ujian skripsi akan tetapi jumlah mahasiswa yang berhasil ujian skripsi nanti bisa diketahui setelah berakhirnya semester, tentu kondisi ini sudah terlambat untuk mengambil suatu keputusan bagaimana agar target tadi bisa tercapai, oleh karena itu diperlukan sebuah sistem yang dapat melakukan prediksi atau peramalan berapa jumlah mahasiswa yang akan ujian skripsi untuk semester tertentu jauh-jauh hari sebelumnya, hal ini bisa dilakukan dengan memanfaatkan Data Mining dengan teknik prediksi.

Prediksi atau peramalan merupakan suatu usaha untuk meramalkan atau memperkirakan keadaan di masa yang akan datang melalui pengujian keadaan di masa lalu. Proses prediksi suatu data harus berdasarkan data yang sama dengan data sebelumnya dan dianalisis secara ilmiah dengan memakai metode statistika sehingga mampu memperbaiki keadaan yang akan terjadi pada periode/waktu yang akan datang[1].

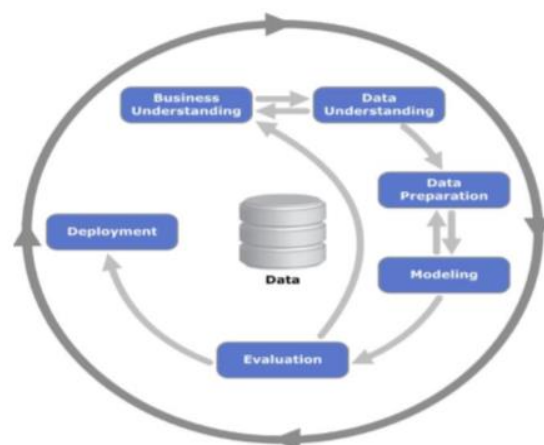
Sesuai kegunaan Data Mining yaitu gudang basis data yang dapat diolah untuk mendapat informasi yang berguna untuk pengambilan suatu keputusan[2], untuk itu pemanfaatan Data Mining terkait data jumlah mahasiswa ujian skripsi pada periode semester sebelumnya dapat diolah atau digunakan untuk memprediksi jumlah mahasiswa yang akan ujian skripsi di semester berikutnya dengan menggunakan metode prediksi. Salah satu metode prediksi yang banyak digunakan adalah metode *Least Square*, karena metode ini dapat diterapkan dalam penentuan persamaan trend data khususnya pada data *time series* dan perhitungannya lebih teliti[1].

Beberapa hasil penelitian sebelumnya yang sudah menerapkan metode *Least Square* diteliti oleh [3] sistem prediksi jumlah kebutuhan obat menggunakan metode *Least Square* berbasis website. Hasil penelitian ini didapatkan persentase nilai akurasi keberhasilan sistem prediksi jumlah kebutuhan obat menggunakan metode *Least Square* adalah 93,49%. Penelitian yang dilakukan oleh [4] tentang memprediksi pendaftar ujian kompetensi laboratorium dengan masukan berupa data jumlah peserta ujian kompetensi dengan tingkat akurasi sebesar 90,01% dan penelitian yang dilakukan oleh [5] metode *Least Square* untuk prediksi penjualan sari kedelai dapat memprediksi penjualan pada periode selanjutnya dengan nilai korelasi sebesar 0,88. Berdasarkan uraian di atas ini membuktikan bahwa metode *Least Square* dapat diterapkan pada penelitian ini yaitu memprediksi jumlah mahasiswa ujian skripsi pada semester berikutnya.

2. METODOLOGI PENELITIAN

2.1 Metode Analisis

Adapun metode analisa yang penulis lakukan adalah menggunakan metode deskriptif, dimana metode ini mengumpulkan data-data yang akan digunakan dan dibuatkan suatu rumusan sehingga didapatkan suatu kesimpulan. Model yang digunakan dalam penyelesaian penelitian ini yaitu menerapkan model CRISP-DM (*Cross Industry Standard Process for Data Mining*), Model CRISP-DM merupakan standar proses dalam pengolahan data mining yang biasa digunakan dalam pemecahan masalah secara umum[6]. Adapaun tahapan dalam standar prosesnya dapat digambarkan pada *life cycle* yang terbagi dalam enam fase sebagai berikut :



Gambar 1. Model CRISP-DM[6]

- a. Pemahaman Bisnis (*Business Understanding*)
Ketepatan waktu dalam penyelesaian studi merupakan target utama setiap Program Studi pada suatu perguruan tinggi dan begitu juga dengan setiap mahasiswa, untuk itu tugas utama dari Program Studi harus bisa memastikan

mahasiswa bisa selesai tepat pada waktunya sehingga target mahasiswa yang ujian skripsi setiap semesternya bisa dilakukan prediksi sebelum terjadi agar bisa dilakukan upaya-upaya, jika hasil prediksi tidak sesuai target yang sudah ditentukan.

b. Pemahaman Data (*Data Understanding*)

Dalam penelitian ini data yang digunakan adalah data jumlah mahasiswa yang sudah ujian skripsi dari 5 Program Studi yang memiliki mahasiswa terbanyak mulai dari periode 2014-2 s/d 2019-2 melalui sistem informasi akademik UNISAN Gorontalo[7].

c. Pengolahan Data (*Data Preparation*)

Pada tahapan ini dilakukan pemisahan data berdasarkan program studi dan posisi semesternya yaitu semester ganjil atau semester genap. Hal ini dilakukan karena data yang ditemukan ada perbedaan nilai trend data untuk periode semester ganjil dan semester genap, dimana pada periode semester ganjil rata-rata setiap prodi jumlah lulusan lebih sedikit dibanding dengan periode semester genap.

d. Pemodelan (*Modelling*)

Metode yang dipilih dalam penelitian ini adalah metode *Least Square*. Untuk melakukan evaluasi tingkat error prediksi atau pengukuran akurasi menggunakan MAPE (*Mean Absolute Percentage Error*), dan untuk membangun aplikasinya menggunakan tools bahasa pemrograman “Microsoft Visual Studio 2010” dengan database MySQL.

e. Evaluasi (*Evaluation*)

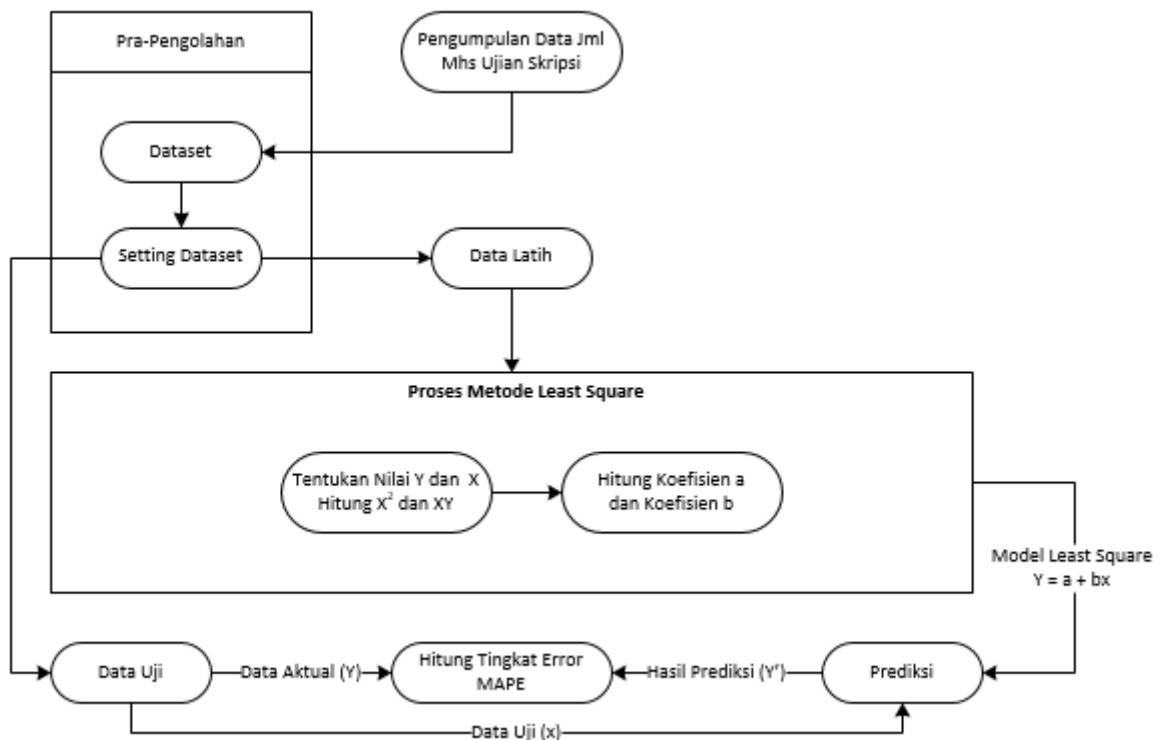
Setelah Model diterapkan, maka perlu diketahui tingkat kesalahan dalam memprediksi dengan menerapkan metode MAPE (*Mean Absolute Percentage Error*).

f. Penyebaran (*Deployment*)

Penelitian ini menghasilkan suatu prediksi yang bisa diterapkan pada *Decision Support System* (DSS), yang bisa digunakan oleh pihak manajemen Universitas Ichsan Gorontalo sebagai dasar dalam menentukan langkah selanjutnya dalam mengatasi permasalahan target jumlah mahasiswa yang akan ujian skripsi setiap semesternya.

2.2 Pemodelan

Tahapan pemodelan ditunjukkan pada gambar 2.



Gambar 2. Pemodelan[8]

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Dataset

Pengumpulan dataset jumlah Mahasiswa ujian Skripsi didapatkan dari laman sistem informasi akademik UNISAN (siakun). Dataset yang diambil hanya 5 sampel Program Studi yang memiliki mahasiswa terbanyak dari 12 Program Studi yang ada di Universitas Ichsan Gorontalo. Dataset yang dikumpulkan mulai dari Periode 2014-2 samapai dengan periode 2019-2, sebagaimana pada tabel berikut :

Tabel 1. Dataset jumlah mahasiswa ujian skripsi[7]

No	Periode	Program Studi (SI)				
		Teknik Informatika	Manajemen	Akuntansi	Ilmu Pemerintahan	Ilmu Hukum
1	20142	187	93	72	130	122
2	20151	79	52	40	26	67
3	20152	160	62	41	92	95
4	20161	104	94	66	35	58
5	20162	171	61	54	98	132
6	20171	51	88	40	32	48
7	20172	165	72	99	104	130
8	20181	49	57	40	22	50
9	20182	125	76	83	101	140
10	20191	83	51	29	18	69
11	20192	102	111	86	91	125

3.2 Pra-Pengolahan Data

Agar dataset di atas dapat diolah maka perlu dilakukan pra pengolahan data yaitu pemisahan data berdasarkan Program Studi dan Periode Ganjil-Genap. Sebagai sampel diambil untuk Program Studi Teknik Informatika

Tabel 2. Dataset prodi teknik informatika periode ganjil[7]

No.	Periode	Jml Mhs Skripsi
1	20151	79
2	20161	104
3	20171	51
4	20181	49
5	20191	83

Tabel 3. Dataset prodi teknik informatika periode genap[7]

No	Periode	Jml Mhs Skripsi
1	20142	187
2	20152	160
3	20162	171
4	20172	165
5	20182	125
6	20192	102

Untuk program studi yang lain dilakukan pra-pengolahan data seperti tahapan di atas. Pada pra-pengolahan data ini ditentukan juga pembagian antara data latih dan data uji, karena datanya sedikit maka data yang digunakan adalah data yang sama antara data latih dan data uji.

3.3 *Pemodelan*

Metode *Least Square* merupakan suatu metode yang dapat diterapkan dalam menghitung persamaan nilai trend data pada data yang berbentuk *time series* dengan menggunakan pendekatan dua kasus data yaitu data bernilai genap dan data bernilai ganjil[1]. Penentuan nilai untuk variabel waktu (*X*) menggunakan aturan tersendiri sehingga didapat nilai nol untuk jumlah nilai variabel waktu. Variabel waktu (*X*) diberikan nilai atau angka yang berbeda pada jumlah waktu yang bernilai ganjil dan waktu yang bernilai genap. Contoh penentuan nilai *X*, untuk jumlah data bernilai ganjil, maka nilai *X*nya adalah : "... , -3, -2, -1, 0, 1, 2, 3, ...". Sedangkan untuk jumlah data bernilai genap, maka nilai *X*nya adalah : "... , -5, -3, -1, 1, 3, 5, ...".

Persamaan trend data pada metode *Least Square* dapat dihitung sebagai berikut :

$$Y = a + bx \tag{1}$$

Keterangan:

Y = Jumlah data berkala

a = Nilai trend data pada periode

b = Nilai trend data pada periode

x = waktu tertentu dalam bentuk kode berdasarkan data yang akan diolah

Koefisien a dan b dihitung dengan persamaan berikut:

$$a = \frac{\sum Y}{n} \tag{2}$$

$$b = \frac{\sum XY}{\sum X^2} \tag{3}$$

3.4 *Penerapan Metode*

Berikut penerapan metode *Least Square* dengan menggunakan dataset pada tabel 2 di atas.

Tabel 4. Tabel koefisien teknik informatika periode ganjil[8]

<i>Periode</i>	<i>Jml Mhs (Y)</i>	<i>Nilai (X)</i>	<i>X.Y</i>	<i>X²</i>
20151	79	-2	-158	4
20161	104	-1	-104	1
20171	51	0	0	0
20181	49	1	49	1
20191	83	2	166	4
Total	366	0	-47	10

$$a = \frac{366}{5} = 73.2$$

$$b = \frac{-47}{10} = -4.7$$

Sehingga didapat persamaan $Y = 73.2 + (-4.7)x$, untuk memprediksi jumlah Mahasiswa yang akan ujian skripsi pada periode 2020-1, dimana nilai *x* nya = 3, maka dapat dihitung sebagai berikut :

$$Y = 73.2 + (-4.7) 3 = 59$$

Berikut penerapan metode *Least Square* dengan menggunakan dataset pada tabel 3 di atas.

Tabel 5. Tabel koefisien teknik informatika periode genap[8]

<i>Periode</i>	<i>Jml Mhs (Y)</i>	<i>Nilai (X)</i>	<i>X.Y</i>	<i>X²</i>
20142	187	-5	-935	25
20152	160	-3	-480	9
20162	171	-1	-171	1
20172	165	1	165	1
20182	125	3	375	9
20192	102	5	510	25
Total	910	0	-536	70

$$a = \frac{910}{6} = 151.6667$$

$$b = \frac{-536}{70} = -7.657143$$

Sehingga didapat persamaan $Y = 151.6667 + (-7.657143)x$, untuk memprediksi jumlah Mahasiswa yang akan ujian skripsi pada periode 2020-2, dimana nilai x nya = 7, maka dapat dihitung sebagai berikut :

$$Y = 151.6667 + (-7.657143)7 = 98$$

Perhitungan pemodelan untuk Program Studi yang lain dilakukan seperti tahapan di atas.

3.5 Evaluasi

Pada tahapan evaluasi model untuk mengetahui tingkat kesalahan perhitungan prediksi dapat digunakan metode MAPE (*Mean Absolute Percentage Error*). Kelebihan penerapan MAPE adalah memberikan informasi kesalahan hasil prediksi terhadap prediksi aktual berdasarkan periode tertentu dalam bentuk nilai presentase kesalahan. Apabila nilai MAPE kurang dari 10%, maka hasil prediksi sangat baik dan apabila nilai MAPE kurang dari 20%, maka hasil prediksi kategori baik[9]. Persamaan MAPE dapat dihitung sebagai berikut :

$$MAPE = \frac{\sum \frac{|y-y'|}{y} \times 100\%}{n} \quad (4)$$

Dimana :

- y' : Data Hasil prediksi
- y : Data aktual/awal
- n : Banyaknya data

Evaluasi model dilakukan berdasarkan program studi, karena datasetnya terbatas, maka data uji sama datanya dengan data latih, berikut data evaluasi untuk data uji Prodi Teknik Informatika.

Tabel 6. Tingkat error mape prodi teknik informatika[8]

<i>Periode</i>	<i>Data Aktual (y)</i>	<i>Data Prediksi (y')</i>	<i>Error MAPE (%)</i>
20142	187	189	1.07
20151	79	82	3.8
20152	160	174	8.75
20161	104	77	25.96
20162	171	159	7.02

<i>Periode</i>	<i>Data Aktual (y)</i>	<i>Data Prediksi (y')</i>	<i>Error MAPE (%)</i>
20171	51	73	43.14
20172	165	144	12.73
20181	49	68	38.78
20182	125	128	2.4
20191	83	63	24.1
20192	102	113	10.78
Total			178.53

$$MAPE = \frac{178.53}{11} = 16.23\%$$

Tingkat kesalahan prediksi mahasiswa ujian skripsi untuk prodi Teknik Informatika berdasarkan perhitungan di atas didapat hasil 16.23% atau tingkat akurasi sebesar 83.77%. Perhitungan evaluasi model untuk data uji program studi yang lain dilakukan dengan tahapan yang sama seperti di atas, adapun rekap hasil evaluasi untuk semua program studi sebagai berikut :

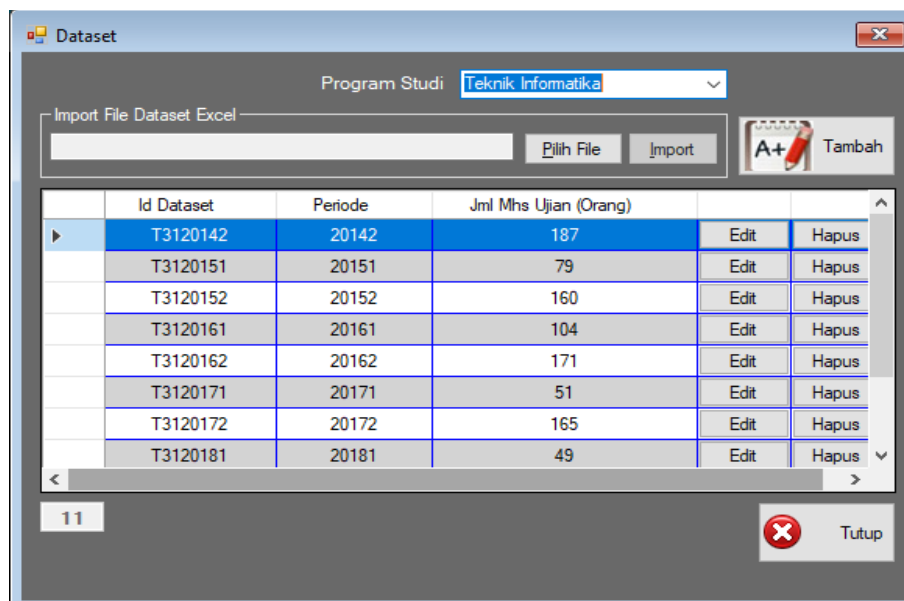
Tabel 7. Rata-rata perhitungan tingkat error MAPE[8]

<i>No</i>	<i>Program Studi</i>	<i>Error MAPE (%)</i>
1	Teknik Informatika	16.23
2	Ilmu Hukum	12.02
3	Ilmu Pemerintahan	11.03
4	Manajemen	21.56
5	Akuntansi	19.69
	Rata-rata	16.11

Setelah dilakukan pengujian tingkat error ke 5 Program Studi didapatkan nilai rata-rata error 16.11% atau tingkat akurasi sebesar 83.89%

3.6 Deployment

Setelah dilakukan pengujian Model dengan hasil akurasi yang baik dimana tingkat errornya <20%, maka bisa dibuatkan suatu sistem prediksi dengan hasil layout sebagai berikut :



Gambar 3. Form Entry Dataset[8]

Form di atas digunakan untuk mengimport dataset yang sudah diolah untuk dimasukkan ke dalam sistem yang sudah dibuat untuk diproses lebih lanjut.

	Id. Dataset	Periode	Nilai (X)	Prediksi Mhs Ujian (Y)
▶	T3120151	20151	-2	82
	T3120161	20161	-1	77
	T3120171	20171	0	73
	T3120181	20181	1	68
	T3120191	20191	2	63
	T3120201	20201	3	59

Gambar 4. Form Prediksi Jumlah Mahasiswa Ujian Skripsi[8]

Form di atas digunakan untuk melakukan prediksi mahasiswa yang akan ujian skripsi pada periode semester yang akan datang dan proses prediksi dilakukan berdasarkan Program Studi.

4. KESIMPULAN

Hasil penelitian ini dapat disimpulkan sebagai berikut :

- 1) Hasil prediksi jumlah mahasiswa ujian skripsi dari 5 Program Studi yang dijadikan sebagai obyek didapatkan hasil pengujian tingkat kesalahan rata-rata sebesar 16.11% atau tingkat akurasi sebesar 83.89%.
- 2) Metode *Least Square* dapat diterapkan untuk memprediksi jumlah mahasiswa ujian skripsi secara baik dan aplikasi yang sudah dibuat dapat diterapkan karena didapatkan tingkat kesalahan kurang dari 20%.
- 3) Disarankan bahwa penelitian ini dapat dilakukan penelitian kembali dengan menerapkan metode prediksi yang lain guna mendapatkan tingkat akurasi yang lebih besar.

DAFTAR PUSTAKA

- [1] Lestari, Sri, I. P., et al. (2019). *Peramalan Stok Spare Part Menggunakan Metode Least Square*, Medan: CV. Sefa Bumi Persada.
- [2] Eko, P. (2013). *Data Mining : Konsep Dan Aplikasi Menggunakan Matlab*. Yogyakarta: CV. Andi Offset.
- [3] Danianty, M. D., Suhery, C., and Hidayati, R. (2020). "Prediksi Jumlah Kebutuhan Obat Menggunakan Metode Least Square Berbasis Website." *Coding*. 08, 02, 33-42.
- [4] Gunadi, Santoni, S. and Na'am, J. (2018). "Tingkat Prediksi Pendaftar Ujian Kompetensi Laboratorium Menggunakan Metode Least Square." *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*. 2, 3, 746-752.
- [5] Hariri, F., R. (2016). "Metode Least Square Untuk Prediksi Penjualan Sari Kedelai Rosi." *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.* 7, 2, 731-736.
- [6] Kusri and Luthfi, E. T. (2019). *Algoritma Data Mining*, Yogyakarta: Andi Publishing.
- [7] BAAK. (2020). "Data Mahasiswa Ujian Skripsi." Gorontalo, [Online]. Available: <https://siakun.unisan.ac.id>.
- [8] Amiruddin. (2021). "Data Primer Penelitian." Gorontalo.
- [9] Berretti, S. and Thampi, S. M. (2016). *Intelligent Systems Technologies and Applications*. Switzerland: Springer.