

## COMPARISON OF LEAST SQUARE AND QUADRATIC METHODS ON PREDICTION THE NUMBER OF NEW STUDENT APPLICANTS

Atin Hasanah<sup>\*1</sup>, Kusrini<sup>2</sup>, Kusnawi<sup>3</sup>

<sup>1,2,3</sup>PJJ Master's Study Program in Informatics Engineering, Universitas AMIKOM Yogyakarta, Indonesia  
Email: <sup>1</sup>[atin.hasanah92s2@students.amikom.ac.id](mailto:atin.hasanah92s2@students.amikom.ac.id), <sup>2</sup>[kusrini@amikom.ac.id](mailto:kusrini@amikom.ac.id), <sup>3</sup>[khusnawi@amikom.ac.id](mailto:khusnawi@amikom.ac.id)

(Article received: June 13, 2023; Revision: July 04, 2023; published: Desember 23, 2023)

### Abstract

New student registration is held every year with several mechanisms. However, in recent years the number of applicants has decreased even though it had experienced a surge in the previous year. So that, it is necessary to have a prediction to predict the number of applicants in the coming year. In addition, the results of these predictions can be used as material for consideration in determining the quota/ceiling for the number of new student admissions in the following academic year. This research used the Least Square and Quadratic methods to predict the number of new student applicants based on data on the number of applicants from the 2014/2015 to 2022/2023 academic years. Performance testing of the two methods was tested with three (3) testing methods : MAE, MAPE, and MSE. The performance test found that the Quadratic method is more suitable with the MAPE value in the "Good" forecasting accuracy category, which is 11%. For the MAE value, it gets 452,17 and an MSE of 302069,04. While Least Square produces a MAPE value in the "Enough" forecasting accuracy category of 30%, for the MAE value, it gets 996,97 and an MSE of 1494205,36.

**Keywords:** least square, MAE, MAPE, MSE, prediction, quadratic.

## KOMPARASI METODE LEAST SQUARE DAN QUADRATIC PADA PREDIKSI JUMLAH PENDAFTAR MAHASISWA BARU

### Abstrak

Pendaftaran Mahasiswa Baru digelar disetiap tahunnya dengan beberapa jalur yang dibuka. Namun, beberapa tahun terakhir jumlah pendaftar mengalami penurunan padahal di tahun sebelumnya sempat mengalami kelonjakan. Sehingga, perlu adanya suatu prediksi untuk meramalkan jumlah pendaftar di tahun mendatang. Selain itu, hasil dari prediksi tersebut dapat menjadi bahan pertimbangan penentuan kuota/pagu jumlah penerimaan mahasiswa baru di tahun akademik selanjutnya. Penelitian ini merapkan metode *Least Square* dan *Quadratic* untuk memprediksi jumlah pendaftar mahasiswa baru berdasarkan data jumlah pendaftar dari tahun akademik 2014/2015 sampai 2022/2023. Pengujian performance kedua metode tersebut di uji dengan tiga (3) metode pengujian yakni MAE, MAPE dan MSE. Pada pengujian performance tersebut, didapat hasil bahwa metode *Quadratic* lebih cocok dengan nilai MAPE masuk dalam kategori akurasi peramalan "Baik" yakni sebesar 11%, untuk nilai MAE mendapat 452,17 dan MSE sebesar 302069,04. Sedangkan *Least Square* menghasilkan nilai MAPE masuk dalam kategori akurasi peramalan "Cukup" yakni sebesar 30%, untuk nilai MAE mendapat 996,97 dan MSE sebesar 1494205,36.

**Kata kunci:** least square, MAE, MAPE, MSE, prediksi, quadratic.

### 1. PENDAHULUAN

Pendaftaran Mahasiswa Baru digelar setiap tahunnya dengan melalui beberapa jalur yang digelar. Pada pendaftaran mahasiswa baru di tingkat UIN/IAIN/STAIN terdapat 2 jalur yang digelar secara nasional yakni SPAN-PTKIN dan UM-PTKIN. Sedangkan tingkat lokal di IAIN Madura menambahkan dua jalur yaitu Mandiri Reguler dan baru-baru ini menambahkan jalur Mandiri Prestasi. Sehingga di IAIN Madura menggelar empat jalur

masuk yakni SPAN-PTKIN, UM-PTKIN, Mandiri Prestasi dan Mandiri Reguler setiap tahunnya

Beberapa tahun terakhir, jumlah pendaftar mahasiswa baru mengalami penurunan dari tahun sebelumnya yang sempat mengalami kenaikan pada jumlah pendaftarannya. Sehingga, perlu adanya suatu sistem yang dapat memprediksi atau meramalkan jumlah pendaftar di tahun mendatang. Hasil prediksi tersebut yang akan digunakan untuk menjadi bahan pertimbangan dalam penentuan kuota/pagu jumlah

penerimaan mahasiswa untuk tahun akademik selanjutnya. Maka dari itu dengan menggunakan data pendaftar mahasiswa baru dari tahun-tahun sebelumnya tersebut akan dilakukan prediksi tentang jumlah pendaftar. Pada penelitian ini menggunakan metode *Least Square* dan *Quadratic* yang nantinya akan dikomparasi atau dibandingkan mana yang lebih baik untuk dapat diterapkan atau digunakan sebagai bahan pertimbangan penentuan kuota/pagu jumlah penerimaan mahasiswa baru. Sebagai landasan penentuan metode yang akan digunakan, dipaparkan beberapa penelitian terkait metode tersebut walau beberapa penelitian yang dipaparkan ini tidak berkaitan dengan data pendaftaran mahasiswa. Namun, pertimbangan dari hasil penelitian sebelumnya tersebut sehingga metode ini yang akan digunakan dalam memprediksi jumlah pendaftar mahasiswa baru.

Pada penelitian yang dilakukan oleh Ines Saraswati Machfiroh dkk, yang membahas tentang trend penjualan sepeda motor untuk periode mendatang dengan membandingkan hasil dari dua (2) metode yakni *Semi Average Method* dan *Least Square Method*. Data yang digunakan merupakan data penjualan sepeda motor dari tahun 2015-2019 (tahun penjualan dan total unit motor yang terjual). Penelitiannya memperoleh hasil performa sebesar 31,89% pada *Least Square* dan 43,96% pada *Semi Average* yang dimana pengujian performa tersebut menggunakan metode MAPE [1].

Harlen Gilbert Simanullang, dkk menggunakan *Least Square* dalam memprediksi jumlah pasien covid. Dari hasil penelitian tersebut didapat hasil peramalan kasus positif sebanyak 33457 kasus, meninggal 1385 kasus dan sembuh 34662 kasus [2].

Ahmad Ridwan, dkk memprediksi penjualan dengan menggunakan metode *Least Square* yang cocok dalam penerapannya. Hal ini dibuktikan dengan hasil prediksi pada Januari 2018 yang data aslinya 125 unit sedangkan data hasil prediksinya 145 unit yang dimana hal tersebut hanya terpaut 20 unit [3].

Sagung Putri Chandra Astiti melakukan perbandingan tiga metode dalam memprediksi jumlah penduduk. Dari ketiga metode didapatkan hasil nilai standar deviasi terkecil dihasilkan oleh metode *Least Square* yakni 2532,27 daripada kedua metode (metode aritmatik dan metode geometrik) yang lain [4].

Indah Rahmawati dan Rony Wijanarko melakukan penelitian dengan menggunakan metode *Least Square* pada data penjualan obat di Apotek Demak Farma Jaya. Bertujuan untuk menghindari tragedi kehabisan persediaan obat dan kesalahan pelaporan [5].

Medyantiwi Rahmawita dan Ilham Fazri pada penelitiannya tentang meramalkan penjualan obat-obatan di Rumah. Hasil penelitiannya mempunyai *error* (tingkat kesalahan) yang kecil [6].

Nindian Puspa Dewi dan Indah Listiowarni dengan menggunakan data dari CV. Rizky Mulya mencapai nilai akurasi 90% pada tahun 2014 sampai 2016 [7].

Johanna Sindya Widjaya dkk, dalam memprediksi jumlah pasien Covid-19 menghasilkan nilai MAPE yang cukup tinggi 59,2%. Nilai tersebut merupakan penjumlahan dari persentase prediksi pasien positif, sembuh, dan meninggal lalu di rata-rata [8].

Ranti Wilda Nur Octavia dan Umi Chotijah menggunakan metode *Least Square* untuk memprediksi penjualan. Pengujian performance menggunakan tiga metode yakni MAD sebesar 0,38, MSE sebesar 1,707 sedangkan MAPE sebesar 0,03602% [9].

Daniel Swanjaya dan Danar Putra Pamungkas dengan menggunakan pengujian performance melalui dua metode yakni korelasi dan MAPE. Penelitiannya menghasilkan nilai korelasi yang rendah yakni 0 sampai 0,2 sedangkan MAPE 41,8% dan 56,3% dengan menggunakan data sebanyak 15 data [10].

Dzurrotun Nasyika dkk, dalam memprediksi jumlah permintaan produk darah dengan menggunakan *Least Square* lalu diuji menggunakan metode MAPE menghasilkan 14,40% untuk prediksi permintaan golongan darah A sedangkan 180,66% untuk prediksi permintaan golongan darah AB [11].

Sitti Raodah Garuda dan Yuliantoro Baliadi di peramalan toleransi komoditas pangan. Data yang digunakan merupakan data produksi dalam periode 1970-2015 dengan empat (4) jenis tanaman yakni Padi, Jagung, Kedelai dan Ubi Jalar. Hasil dari peramalan keempat metode tersebut dibandingkan dan dicari mana yang terbaik ketika pengujian performanya (*MSD/MSE*, *MAD/MAE* dan *MAPE*). Model *tren moving average ordo 2* sesuai untuk memprediksi produksi padi dan kedelai. Sedangkan Model *tren Quadratic* lebih sesuai untuk memprediksi produksi jagung dan ubi jalar [12].

Ahmad Rifa'i melakukan penelitian tersebut untuk meramalkan jumlah pendaftar dari tahun ajaran 2023/2024 sampai dengan tahun ajaran 2026/2027. Data yang digunakan yakni tahun akademik dan jumlah pendaftar pada tahun akademik tersebut. Metode yang digunakan adalah metode *Quadratic/parabolic*. Hasil peramalannya menyatakan bahwa Jumlah calon mahasiswa baru yang akan mendaftar pada periode tahun ajaran 2023/2024 sebanyak 651 Mahasiswa, 2024/2025 sebanyak 818 Mahasiswa, 2025/2026 sebanyak 1001 Mahasiswa, 2026/2027 sebanyak 1200 Mahasiswa [13].

Dayu Pratama dan Maulana Ardiansyah dalam memprediksi penjualan minuman di kedai. Dalam penelitiannya membandingkan dua metode yakni *Least Square* dan *Quadratic/Prabolic*. *Least Square* menghasilkan nilai MAPE sebesar 3,16% , sedangkan 3,05% [14].

Indah Simamora dalam penelitiannya yang menggunakan metode tren non linear menghasilkan penurunan jumlah penduduk dalam hasil prediksinya. Sebanyak 318923 jiwa yang diramalkan untuk tahun 2018 [15].

Maka dari itu, penelitian ini menggunakan dua metode yakni metode Least Square dan metode Quadratic untuk dibandingkan, metode mana yang cocok dalam prediksi jumlah pendaftar mahasiswa baru. Dimana nilai prediksi yang diperoleh metode tersebut dapat membantu dan menjadi bahan pertimbangan dalam penentuan kuota/pagu jumlah penerimaan mahasiswa baru di tahun akademik mendatang. Dalam menentukan metode mana yang terbaik, maka pengujian performance pada penelitian ini menggunakan pengujian performance dengan tiga metode yakni MAE, MAPE dan MSE.

## 2. METODE PENELITIAN

### 2.1. Jenis, Sifat dan Pendekatan Penelitian

Pada penelitian ini menggunakan jenis penelitian eksperimen yakni melakukan pengujian performa dari hasil prediksi yang menggunakan Metode *Least Square* dan *Quadratic* pada data jumlah pendaftar mahasiswa baru.

Penelitian ini dilakukan secara mandiri menggunakan sifat deskriptif dimana data yang diperoleh akan dibandingkan dengan data yang di uji. Penelitian ini menggunakan pendekatan kuantitatif.

### 2.2. Teknik Pengumpulan Data

Peneliti memperoleh data pendaftar mahasiswa baru dari akun jalur PMB Nasional yakni jalur SPAN-PTKIN dan UM-PTKIN. Serta jalur PMB lokal yang diadakan oleh kampus sendiri yakni Mandiri Reguler. Data jumlah pendaftar tersebut diambil dari Tahun Akademik 2014/2015 sampai 2022/2023.

### 2.3. Teknik Analisis Data

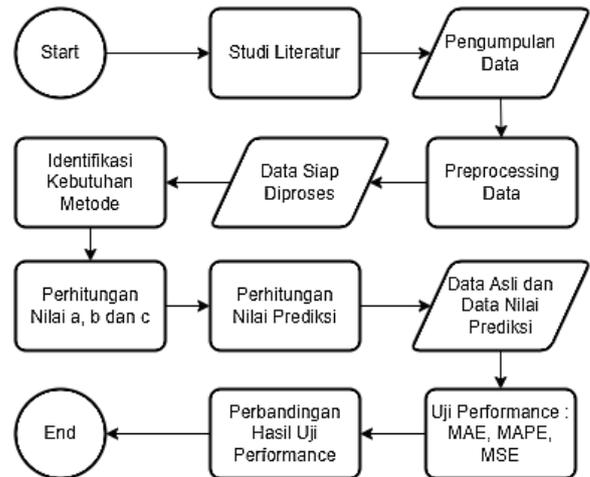
Analisis Data pada penelitian ini menggunakan pengujian *performance* pada hasil prediksi metode tersebut..

### 2.4. Lokasi Penelitian

Data yang akan digunakan merupakan data pendaftar mahasiswa baru yang mendaftar di IAIN Madura.

#### 2.1. Alur Penelitian

Berikut merupakan Alur Penelitian yang dilakukan, tertera pada **Gambar 1** :



Gambar 1. Alur penelitian

Penjelasan tahapan pada alur penelitian yang terdapat pada **Gambar 1** :

#### a. Studi literatur

Mencari, mempelajari dan mengumpulkan referensi/landasan teori yang berkaitan dengan metode yang digunakan dalam penelitian dari berbagai sumber jurnal maupun buku untuk melengkapi konsep dan teori dalam penulisan.

#### b. Pengumpulan data

Mengumpulkan data dari sumber yang resmi untuk digunakan dalam penelitian.

#### c. Preprocessing data

Melakukan pengecekan data terlebih dahulu, seperti data kosong dan data double harus dihilangkan sebelum data tersebut diolah.

#### d. Data siap diproses

Data yang sudah melalui tahap preprocessing berarti data ini sudah siap untuk diproses dan diolah oleh sistem.

#### e. Identifikasi kebutuhan metode

Bagian ini, sistem melakukan perhitungan untuk mendapatkan nilai dari  $\sum Y$ ,  $\sum XY$ ,  $\sum X^2$ ,  $\sum X^4$ , dan  $\sum X^2 Y$ .

#### f. Perhitungan nilai a, b dan c

Hasil dari tahapan identifikasi kebutuhan metode tersebut disubstitusikan ke persamaan (1), (2), (3), (4) dan (5) untuk mendapatkan nilai konstanta (a), variabilitas (b), variabilitas per X (b) dan variabilitas per X kuadrat (c).

#### g. Perhitungan nilai prediksi

Hasil perolehan dari perhitungan a, b dan c selanjutnya disubstitusikan ke persamaan (6) dan (7) untuk mendapatkan nilai prediksi.

#### h. Data asli dan data nilai prediksi

Bagian ini, sudah diperoleh nilai data asli dan data nilai prediksi di perunit waktunya (nilai prediksi di setiap tahun akademik berdasarkan urutan dari tahun akademik di data asli).

#### i. Uji performance : MAE, MAPE, MSE

Melakukan perhitungan selisih antara nilai data asli dan data nilai prediksi. Selanjutnya disubstitusikan ke persamaan (8), (9) dan (10) untuk mendapatkan nilai MAE, MAPE dan MSE.

j. Perbandingan hasil uji performance  
 Dari tahap uji performance didapat nilai MAE, MAPE dan MSE dari kedua metode. Nilai tersebut dibandingkan mana yang nilainya lebih kecil.

**2.6. Landasan Teori**

a. *Machine Learning*

*Machine Learning* / Pembelajaran Mesin merupakan studi ilmiah tentang algoritma yang membangun model matematika ataupun statistik komputasi berdasarkan data sampel / data *training* untuk membuat suatu prediksi ataupun keputusan tanpa diprogram secara eksplisit dalam melakukan tugasnya [16]. Dalam *Machine Learning* terdapat 3 jenis yakni :

1) *Supervised Learning*

Teknik yang mengidentifikasi hubungan / mengenali pola antara input data dan label output. Contoh dari teknik ini salah satunya ada prediksi / peramalan.

2) *Unsupervised Learning*

Teknik yang mengidentifikasi hubungan / mengenali pola yang sebelumnya belum dapat ditemukan pada data yang tidak berlabel / tidak memiliki output yang suda ada sebelumnya dengan pengawasan manusia yang minimal.

3) *Semi-Unsupervised Learning*

Teknik yang menggabungkan sejumlah kecil data berlabel dan data tidak berlabel selama dilakukan training data untuk menemukan pola diantara keduanya.

b. *Forecasting* (Peramalan / Prediksi)

*Forecasting* / peramalan / prediksi merupakan proses untuk menduga suatu kejadian mendatang berdasarkan historis dan pengalaman dari data yang digunakan dengan menggunakan perhitungan matematis untuk memperkecil resiko kesalahan [17].

Tujuan dari *Forecasting* / Peramalan / Prediksi untuk mengurangi / meminimalkan ketidakpastian sehingga dapat diperoleh estimasi yang mendekati nilai aktual [18].

c. *Analisis Time Series*

*Analisis Time Series* / Analisis Deret Berkala / Analisis Deret Waktu ialah prosedur analitis yang dapat digunakan untuk mengidentifikasi perubahan atau perkembangan nilai suatu variabel perubahan dari waktu ke waktu.

Terdapat empat (4) komponen yang terdapat pada Time Series Data, antara lain :

- 1) Trend : tendensi atau kecenderungan pergerakan data yakni bisa cenderung naik, turun atau tetap.
- 2) Seasonality / Variasi Musiman : pola yang ditemukan pada data, dimana data tersebut mengalami kenaikan/penurunan secara berulang dan sifatnya periodik.
- 3) Cyclical / Variasi Siklis : di dalam data terdapat pola yang berulang namun tidak bersifat periodik (reguler).

4) Random Variations / Variasi Random : gerakan data yang tidak memiliki pola atau tidak beraturan (irregular).

d. *Least Square*

*Least Square* / kuadrat terkecil adalah metode penentuan *trend* data yang mencakup analisis deret waktu dengan dua kasus yakni data genap atau ganjil [14].

Persamaan (1) dan (2) merupakan rumus dalam pencarian *konstanta* a dan *variabelitas* b dalam Metode *Least Square* [19]:

$$a = \frac{\sum Y}{n} \tag{1}$$

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

Dimana :

- Y : Nilai asli
- n : Banyak data
- X : Unit waktu

e. *Quadratic*

*Quadratic* / Kuadratik / Parabolik adalah trend dimana nilai variabel dependen meningkat atau menurun secara linear atau berbentuk parabola saat data diubah menjadi scatter plot [20].

Persamaan (3), (4) dan (5) merupakan rumus dalam pencarian konstanta a, variabelitas b dan c variabelitas per X kuadrat, dalam Metode *Quadratic*:

$$a = \frac{(\sum Y \cdot \sum X^4) - (\sum X^2 Y \cdot \sum X^2)}{(n \cdot \sum X^4) - (\sum X^2)^2} \tag{3}$$

$$b = \frac{\sum XY}{\sum X^2} \tag{4}$$

$$c = \frac{(n \cdot \sum X^2 Y) - (\sum X^2 \cdot \sum Y)}{(n \cdot \sum X^4) - (\sum X^2)^2} \tag{5}$$

Dimana :

- Y : Nilai asli
- a : *Konstanta*
- X : Unit waktu
- b : *Variabelitas* per X
- c : *Variabelitas* per X kuadrat

f. Persamaan Garis Linear dan Non Linear/Parabolic

Persamaan umum garis linear [3] dalam menghasilkan nilai prediksi, pada persamaan (6) :

$$Y' = a + bX \tag{6}$$

Dimana :

- Y' : Nilai peramalan
- a : *Konstanta*
- X : Unit waktu
- b : *Variabelitas* per X

Persamaan umum garis non linear/parabolic [13] dalam menghasilkan nilai prediksi, pada persamaan (7) :

$$Y' = a + bX + cX^2 \tag{7}$$

Dimana :

- Y' : Nilai peramalan
- X : Unit waktu
- a : Konstanta
- b : Variabelitas per X
- c : Variabelitas per X kuadrat

g. *Mean Absolute Error (MAE)*

*Mean Absolute Error (MAE) / Mean Absolute Deviation (MAD)* merupakan perhitungan dari perubahan mutlak antara nilai asli dan peramalan serta rata-rata dari hasil keduanya [21].

Dengan menggunakan data riil dan data hasil peramalan, maka substitusikan ke persamaan (8):

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y - Y'| \tag{8}$$

Dimana :

- Y : Nilai asli
- Y^' : Nilai peramalan/prediksi
- n : Banyaknya data

h. *Mean Absolute Percentage Error (MAPE)*

*Mean Absolute Percentage (MAPE)* ialah nilai persentase dari perbedaan antara nilai aktual dan peramalan [22]. Perhitungan rerata kesalahan dengan *MAPE* biasanya lebih banyak dipakai dikarenakan lebih mudah dipahami karena nilainya dalam bentuk persentase [8].

Pada persamaan (9) untuk mendapat nilai *MAPE*:

$$MAPE = \left(\frac{100\%}{n}\right) \cdot \sum_{i=1}^n \frac{|Y - Y'|}{Y} \tag{9}$$

Dimana :

- Y : Nilai asli
- Y^' : Nilai peramalan/prediksi
- n : Banyaknya data

Tabel 1 merupakan interpretasi nilai *MAPE* tentang kategori akurasi peramalan beserta rentang nilai *MAPE* dalam bentuk persentasenya :

Tabel 1. Kategori akurasi MAPE

Nilai MAPE	Akurasi Peramalan
MAPE ≤ 10%	Tinggi
10% < MAPE ≤ 20%	Baik
20% < MAPE ≤ 50%	Cukup
MAPE > 50%	Rendah

i. *Mean Square Error (MSE)*

*Mean Squared Error (MSE)* adalah perhitungan dari kuadrat perubahan antara nilai asli dan peramalan serta rata-rata dari hasil keduanya [9].

Pada persamaan (10) untuk mendapat nilai *MSE*:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y - Y')^2 \tag{10}$$

Dimana :

- Y : Nilai asli
- Y^' : Nilai peramalan/prediksi
- n : Banyaknya data

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Pengumpulan Data

Data yang digunakan merupakan data jumlah pendaftar mahasiswa baru di tiga jalur masuk dari Tahun Akademik 2014/2015 sampai 2022/2023. **Tabel 2** merupakan data asli atau belum dilakukan proses preprocessing (penghapusan data kosong dan double), pada tabel tersebut telah dirinci tiga jalur masuk (SPAN-PTKIN, UM-PTKIN dan Mandiri Reguler):

Tabel 2. Data sebelum preprocessing

Tahun Akademik	SPAN-PTKIN	UM-PTKIN	Mandiri Reguler
2014/2015	563	827	468
2015/2016	1596	1668	1041
2016/2017	1866	1649	1039
2017/2018	1887	2125	1384
2018/2019	1746	1665	1962
2019/2020	3049	2086	1879
2020/2021	2386	1776	995
2021/2022	2085	1198	773
2022/2023	1743	1141	479

#### 3.2. Preprocessing Data

Pada tahap ini, harus dilakukan pengecekan apakah terdapat data kosong atau ada pendaftar yang melakukan pendaftaran lebih dari satu dalam satu jalur yang sama. Ini bertujuan agar tidak ada data double pada data yang akan digunakan nantinya. **Tabel 3** merupakan rincian jumlah data kosong/double yang ditemukan:

Tabel 3. Jumlah data kosong/double

Tahun Akademik	SPAN-PTKIN	UM-PTKIN	Mandiri Reguler
2014/2015	0	0	0
2015/2016	0	0	0
2016/2017	0	0	0
2017/2018	0	0	0
2018/2019	0	0	728
2019/2020	0	0	591
2020/2021	0	0	394
2021/2022	0	0	399
2022/2023	0	0	215

**Tabel 3** pada kolom SPAN-PTKIN dan UM-PTKIN tertulis nol, dikarenakan data dari kedua jalur tersebut merupakan jalur yang digelar oleh tingkat nasional dan yang melakukan pendaftaran merupakan dari pihak sekolah sehingga tidak ada double data. Sedangkan jalur Mandiri Reguler hanya dari Tahun Akademik 2018/2019 sampai 2022/2023 yang teridentifikasi data-data double.

**Tabel 4** merupakan data pendaftar yang sudah melalui penghapusan data double, lalu ketiga jalur tersebut dijumlahkan pertahun akademik:

Tabel 4. Data pendaftar mahasiswa baru

Tahun Akademik	SPAN-PTKIN	UM-PTKIN	Mandiri Reguler	Total
2014/2015	563	827	468	1858
2015/2016	1596	1668	1041	4305
2016/2017	1866	1649	1039	4554
2017/2018	1887	2125	1384	5396
2018/2019	1746	1665	1234	4645
2019/2020	3049	2086	1288	6423
2020/2021	2386	1776	601	4763
2021/2022	2085	1198	374	3657
2022/2023	1743	1141	264	3148

**3.3. Data Siap Diproses**

**Tabel 5** merupakan tabel data jumlah pendaftar pertahun akademik yang sudah melalui tahapan preprocessing data, sehingga data tersebut siap untuk diproses ke tahap identifikasi kebutuhan metode:

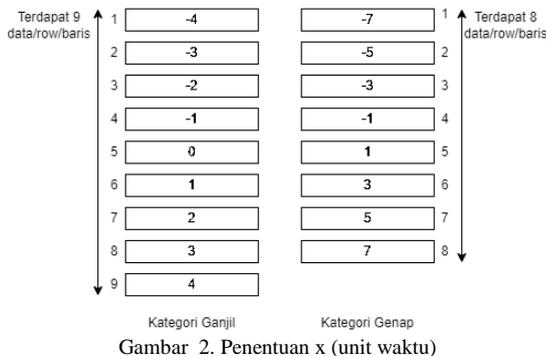
Tabel 5. Data total pendaftar mahasiswa baru

Tahun Akademik	Jumlah Pendaftar
2014/2015	1858
2015/2016	4305
2016/2017	4554
2017/2018	5396
2018/2019	4645
2019/2020	6423
2020/2021	4763
2021/2022	3657
2022/2023	3148

**3.4. Identifikasi Kebutuhan Metode**

Pada bagian ini merupakan mengidentifikasi sekaligus melakukan perhitungan berdasarkan keperluan untuk pencarian konstanta a, variabelitas b dan variabelitas kuadrat c.

Pada metode ini dalam penentuan nilai X (Unit Waktu) harus mengecek terlebih dahulu banyaknya data yang digunakan merupakan kategori ganjil atau genap. Ilustrasi penentuan X (Unit Waktu) digambarkan pada **Gambar 2** berlaku pada *Least Square* maupun *Quadratic*:



Gambar 2. Penentuan x (unit waktu)

Penjelasan mengenai **Gambar 2** dalam penentuan X (Unit Waktu) [2] :

- a. Kategori Ganjil / Data Ganjil  
Misal terdapat data sebanyak 9 data / baris / row, dibagi menjadi 2 bagian sehingga di dapat 4,5.

Maka data dari baris 1 – 4 merupakan baris yang nilai X (Unit Waktu) kurang dari 0. Baris ke 5 merupakan nilai X (Unit Waktu) yang bernilai 0 karena berada ditengah. Selanjutnya, baris ke 6 – 9 mendapat nilai X (Unit Waktu) lebih dari 0. Jika diurutkan dari data ke 1 – ke 9 maka nilai X (Unit Waktu) nya yakni -4,-3,-2,-1,0,1,2,3,4.

- b. Kategori Genap / Data Genap  
Misal terdapat data sebanyak 8 data / baris / row, dibagi menjadi 2 bagian sehingga di dapat 4. Maka nilai X (Unit Waktu) yang nilainya 0 berada di antara data baris ke 4 dan 5. Sehingga, baris ke 4 – 1 akan memiliki nilai X (Unit Waktu) kurang dari 0. Sedangkan baris ke 5 – 8 akan memiliki nilai X (Unit Waktu) lebih dari 0. Jika diurutkan dari data ke 1 – ke 8 maka nilai X (Unit Waktu) nya yakni -7,-5,-3,-1,1,3,5,7.

Berdasarkan penjelasan mengenai **Gambar 2**, **Tabel 6** merupakan hasil perhitungan untuk keperluan metode *Least Square*. Y merupakan data asli dari jumlah pendaftar (total pendaftar dari ketiga jalur), X merupakan unit waktu berdasarkan ketentuan yang dijelaskan **Gambar 2**. XY merupakan hasil perkalian antara nilai jumlah pendaftar dan unit waktu, selanjutnya  $X^2$  merupakan hasil kuadrat dari unit waktu:

Tabel 6. Hasil perhitungan least square

Tahun Akademik	Y	X	XY	$X^2$
2014/2015	1858	-4	-7432	16
2015/2016	4305	-3	-12915	9
2016/2017	4554	-2	-9108	4
2017/2018	5396	-1	-5396	1
2018/2019	4645	0	0	0
2019/2020	6423	1	6423	1
2020/2021	4763	2	9526	4
2021/2022	3657	3	10110	9
2022/2023	3148	4	11984	16
Total	38749		4661	60

**Tabel 7** merupakan hasil perhitungan untuk keperluan metode *Quadratic*. Y merupakan data asli dari total pendaftar (total pendaftar dari ketiga jalur), X merupakan unit waktu berdasarkan ketentuan yang dijelaskan **Gambar 2**. XY merupakan hasil perkalian antara nilai jumlah pendaftar dan unit waktu, selanjutnya  $X^2$  merupakan hasil kuadrat dari unit waktu.  $X^2Y$  merupakan hasil dari perkalian kolom  $X^2$  dan Y.  $X^4$  merupakan hasil dari kolom unit waktu yang dikuadratkan empat:

Tabel 7. Hasil perhitungan quadratic

Tahun Akademik	Y	X	XY	$X^2$	$X^2Y$	$X^4$
2014/2015	1858	-4	-7432	16	29728	256
2015/2016	4305	-3	-12915	9	38745	81
2016/2017	4554	-2	-9108	4	18216	16
2017/2018	5396	-1	-5396	1	5396	1
2018/2019	4645	0	0	0	0	0
2019/2020	6423	1	6423	1	6423	1
2020/2021	4763	2	9526	4	19052	16
2021/2022	3657	3	10110	9	32913	81
2022/2023	3148	4	11984	16	50368	256
Total	38749		4661	60	200841	708

**3.5. Perhitungan Nilai a, b, dan c**

a. Perhitungan pada *Least Square*

Selanjutnya, melakukan substitusi ke persamaan umum untuk mendapatkan nilai prediksi pada metode *Least Square*.

Berdasarkan hasil dari Identifikasi Kebutuhan Metode yang diperoleh melalui **Tabel 6**, berikut proses substitusi ke **persamaan (1)** dan **(2)**:

$$\begin{aligned}
 a &= \frac{\sum Y}{n} \\
 &= \frac{38749}{9} = 4305,444 \\
 b &= \frac{\sum XY}{\sum X^2} \\
 &= \frac{4661}{60} = 77,683
 \end{aligned}$$

Setelah didapatkan nilai *konstanta* a dan *variabelitas* b, selanjutnya melakukan perhitungan nilai prediksi jumlah pendaftar mahasiswa baru pada tiap Tahun Akademiknya.

Berikut contoh perhitungan nilai prediksi dari *Least Square* untuk Tahun Akademik 2014/2015 yang dimana X atau unit waktu untuk Tahun Akademik tersebut adalah -4 berdasarkan hasil dari **Tabel 6**. Selanjutnya disubstitusikan ke **persamaan (6)** seperti di bawah ini:

$$\begin{aligned}
 Y' &= a + bX \\
 &= 4305,444 + (77,683 * (-4)) \\
 &= 4305,444 + (-310,733) \\
 &= 3994,711
 \end{aligned}$$

3994,711 merupakan hasil nilai prediksi jumlah pendaftar untuk Tahun Akademik 2014/2015. Perhitungan nilai prediksi untuk Tahun-tahun Akademik selanjutnya menggunakan cara yang sama.

**Tabel 8** merupakan data jumlah pendaftar mahasiswa baru berdasarkan data asli (Y) dan hasil prediksinya (Y') dengan menggunakan metode *Least Square*. Setelah didapat nilai prediksi, perlu dilakukan perhitungan selisih antar keduanya (Y - Y') karena hasil selisih ini akan dijumlah dan digunakan untuk pengujian performance nantinya:

Tabel 8. Nilai aktual dan nilai prediksi (least square)

Tahun Akademik	Y	Y'	Y - Y'
2014/2015	1858	3994,711	2136,711
2015/2016	4305	4072,394	232,605
2016/2017	4554	4150,078	403,922
2017/2018	5396	4227,761	1168,239
2018/2019	4645	4305,444	339,555
2019/2020	6423	4383,128	2039,872
2020/2021	4763	4460,811	302,188
2021/2022	3657	4538,494	881,494
2022/2023	3148	4616,178	1468,178
Total			8972,767

b. Perhitungan pada *Quadratic*

Selanjutnya, melakukan substitusi ke persamaan umum untuk mendapatkan nilai prediksi pada metode *Quadratic* dengan menggunakan data pada **Tabel 7**.

Data penjumlahannya di substitusikan ke persamaan (3), (4) dan (5):

$$\begin{aligned}
 a &= \frac{(\sum Y \cdot \sum X^4) - (\sum X^2 Y \cdot \sum X^2)}{(n \cdot \sum X^4) - (\sum X^2)^2} \\
 &= \frac{27434292 - 12050460}{6372 - 3600} \\
 &= \frac{38749}{9} = 5549,723 \\
 b &= \frac{\sum XY}{\sum X^2} \\
 &= \frac{4661}{60} = 77,683 \\
 c &= \frac{(n \cdot \sum X^2 Y) - (\sum X^2 \cdot \sum Y)}{(n \cdot \sum X^4) - (\sum X^2)^2} \\
 &= \frac{1807569 - 2324940}{6372 - 3600} \\
 &= \frac{-517371}{2772} = -186,642
 \end{aligned}$$

Setelah didapatkan nilai *konstanta* a, *variabelitas* b dan *variabelitas* kuadrat c, selanjutnya melakukan perhitungan nilai prediksi jumlah pendaftar mahasiswa baru pada tiap Tahun Akademiknya.

Berikut contoh perhitungan nilai prediksi dari *Quadratic* untuk Tahun Akademik 2014/2015 yang dimana X atau unit waktu untuk Tahun Akademik tersebut adalah -4 sedangkan X<sup>2</sup>nya 16 atau (-4)<sup>2</sup> berdasarkan hasil dari **Tabel 7**. Selanjutnya disubstitusikan ke **persamaan (7)** seperti di bawah ini:

$$\begin{aligned}
 Y' &= a + bX + cX^2 \\
 &= 5549,723 + (77,683 * (-4)) \\
 &\quad + (-186,642 * (-4)^2) \\
 &= 5549,723 + (-310,733) \\
 &\quad + (-2986,27) \\
 &= 2252,721
 \end{aligned}$$

Perhitungan nilai prediksi untuk Tahun-tahun Akademik selanjutnya menggunakan cara yang sama.

**Tabel 9** merupakan data jumlah pendaftar mahasiswa baru berdasarkan data asli (Y) dan hasil prediksinya (Y') dengan menggunakan metode *Quadratic*. Setelah didapat nilai prediksi, perlu dilakukan perhitungan selisih antar keduanya (Y - Y') karena hasil selisih ini akan dijumlah dan digunakan untuk pengujian performance nantinya:

Tabel 9. Nilai aktual dan nilai prediksi (quadratic)

Tahun Akademik	Y	Y'	Y - Y'
2014/2015	1858	2252,721	394,7212
2015/2016	4305	3636,897	668,103
2016/2017	4554	4647,789	93,78918
2017/2018	5396	5285,398	110,6022
2018/2019	4645	5549,723	904,7229
2019/2020	6423	5440,765	982,2355
2020/2021	4763	4958,523	195,5225
2021/2022	3657	4102,997	445,997
2022/2023	3148	2874,188	273,8121
Total			4069,506

### 3.6. Uji Performance

Pada bagian ini, data yang didapat pada **Tabel 8** digunakan untuk perhitungan uji performance menggunakan metode *MAE*, *MAPE* dan *MSE* dengan mensubstitusikan data tabel tersebut ke persamaan (8), (9) dan (10). **Gambar 3** merupakan hasil uji performance dari data **Tabel 8** (*Least Square*):

Mean squared error / MSE : 1494205.36  
 Mean absolute error / MAE : 996.97  
 Mean Absolute Percentage Error / MAPE : 0.30  
 Gambar 3. Uji performance (least square)

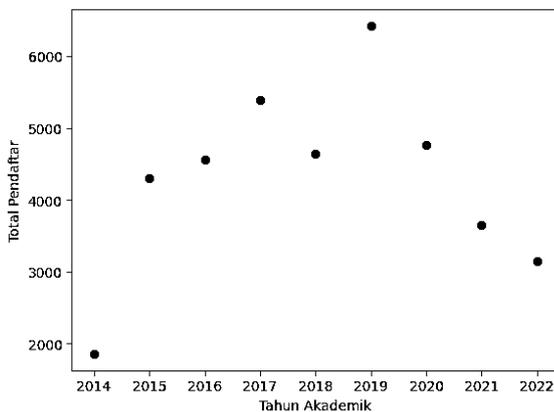
Sedangkan untuk *Quadratic* menggunakan **Tabel 9** dalam melakukan pengujian performancenya. **Gambar 4** merupakan hasil uji performance dari data **Tabel 9** (*Quadratic*):

Mean squared error / MSE : 302069.04  
 Mean absolute error / MAE : 452.17  
 Mean Absolute Percentage Error / MAPE : 0.11  
 Gambar 4. Uji performance (quadratic)

Hasil perhitungan uji performance MAPE (*Mean Absolute Percentage Error*) diubah dalam bentuk persen. Jadi hasil MAPE pada *Least Square* sebesar 30% sedangkan pada *Quadratic* sebesar 11%.

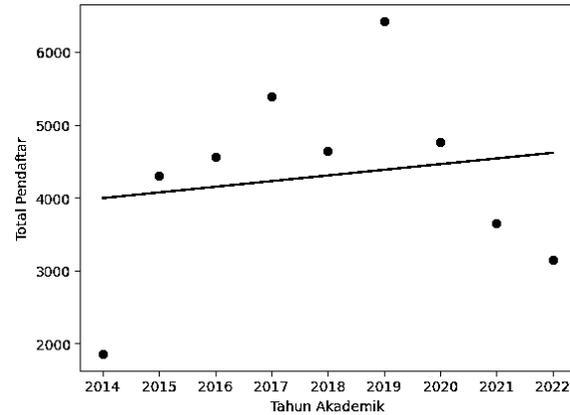
### 3.7. Visualisasi Data

**Gambar 5** merupakan visualisasi berdasarkan data asli/aktualnya dimana gambar lingkaran/titik tersebut menandakan Tahun Akademik dengan jumlah pendaftar tertentu berdasarkan data pada **Tabel 5** yang merupakan data yang sudah siap diproses :



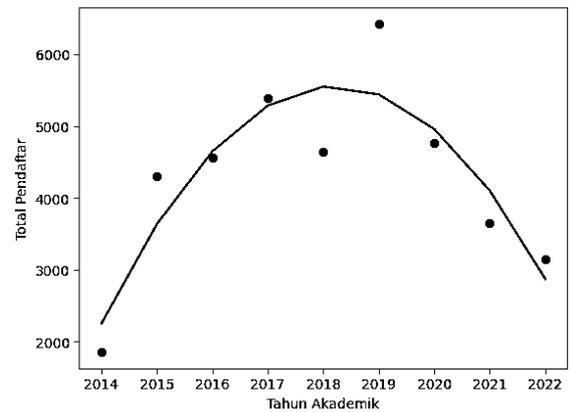
Gambar 5. Visualisasi data asli/aktual

Pada **Gambar 6** menggunakan data pada **Tabel 8** yang merupakan perpaduan dua data yakni data asli/aktual (kolom *Y*) digambarkan menggunakan lingkaran/titik dan data hasil prediksi (kolom *Y'*) digambarkan menggunakan garis yang merupakan hasil dari perhitungan *Least Square*:



Gambar 6. Visualisasi data aktual dan data prediksi per tahun (*least square*)

Pada **Gambar 7** menggunakan data pada **Tabel 9** yang merupakan perpaduan dua data yakni data asli/aktual (kolom *Y*) digambarkan menggunakan lingkaran/titik dan data hasil prediksi (kolom *Y'*) digambarkan menggunakan garis yang merupakan hasil dari perhitungan *Quadratic*:



Gambar 7. Visualisasi data aktual dan data prediksi per tahun (*quadratic*).

## 4. DISKUSI

Berdasarkan apa yang sudah dipaparkan pada hasil dan pembahasan di atas. Keberhasilan penggunaan metode *Least Square* yang telah dilakukan di penelitian sebelumnya seperti yang dilakukan oleh Sagung Putri Chandra Astiti yakni *Least Square* memperoleh nilai standart deviasi terkecil [4]. Selain itu, Nindian Puspa Dewi yang akurasi dari prediksinya mencapai 90% [7]. Namun ada juga yang hasilnya cukup buruk bahkan nilai MAPE yang dihasilkan melebihi dari nilai MAPE yang diperoleh oleh penelitian ini yang hanya bernilai 30%. Diantaranya yang dilakukan oleh Johanna Sindy Widjaya [8] dan penelitian milik Dzurrotun Nasyika [11] yang membahas tentang prediksi jumlah pasien Covid dan jumlah permintaan produk darah dengan nilai MAPEnya berada di atas 30% dari nilai MAPE di penelitian ini. Jika dilihat dari kasus-kasus yang berhasil dan kurang berhasil mayoritas merupakan tentang data penjualan untuk yang

kategori berhasil sedangkan data yang bersifat naik turun/non linear menyebabkan hasil MAPENya besar atau tergolong buruk. Begitu pula yang dilakukan di penelitian ini yang menggunakan data jumlah pendaftar, namun nilai MAPE yang didapat pada penelitian ini dari hasil metode *Least Square* masih di bawah nilai dari penelitian-penelitian sebelumnya yang masuk dalam kategori kurang berhasil dalam penelitiannya.

Sedangkan pada penelitian Dayu Pratama [14] yang membandingkan metode Least Square dan Quadratic memperoleh hasil MAPE yang beda tipis. Kedua metode tersebut masuk dalam kategori akurasi peralamalan "Tinggi" dengan menggunakan data sebanyak 10 data penjualan. Namun pada penelitian ini hanya mencapai ke kategori "Baik". Mungkin saja penelitian yang telah dilakukan di penelitian ini dapat juga masuk ke kategori akurasi peralamalan "Tinggi" jika data yang digunakan lebih banyak daripada yang digunakan sekarang. Selain itu, Sitti Raodah [12] yang meneliti tentang produksi pangan menyatakan metode Quadratic baik digunakan dalam meramalkan produksi ubi jalar.

## 5. KESIMPULAN

Berdasarkan pemaparan, metode yang lebih cocok pada penelitian ini yakni metode *Quadratic* dengan nilai dari hasil uji performance (MAE, MAPE, MSE) lebih kecil daripada nilai hasil uji performance yang didapat dari metode *Least Square*.

Metode *Quadratic* lebih cocok dikarenakan dari bentuk data yang digunakan pada penelitian, dapat dilihat pada **Gambar 5** (Visualisasi Data Asli/Aktual) yang memiliki pola seperti parabola (mengalami kenaikan lalu mengalami penurunan di beberapa tahun terakhir).

Jika ingin tetap data dalam penelitian ini untuk digunakan pada penelitian selanjutnya, saran yang dapat diberikan yakni mencobanya untuk melakukan kombinasi dengan metode lain. Namun tentunya harus memperhatikan bentuk data yang digunakan dan metode yang digunakan termasuk dalam penggunaan data linear atau non linear.

## DAFTAR PUSTAKA

- [1] I. S. Machfiroh, W. A. Alam Sur, and R. T. Pangestu, "Trend Semi Average and Least Square in Forecasting Yamaha Motorcycle Sales," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 16, no. 1, pp. 343–354, 2022, doi: 10.30598/barekengvol16iss1pp341-352.
- [2] A. P. Silalahi and H. G. Simanullang, "Prediksi Jumlah Pasien Covid-19 Di Indonesia Menggunakan Least Square Method Berbasis Android," *J. Inform. Manaj. dan Komput.*, vol. 14, no. 1, pp. 86–93, 2022, doi: 10.36723/juri.v14i1.328.
- [3] A. Ridwan, A. Faisol, and F. Santi Wahyuni, "Penerapan Metode Least Square Untuk Prediksi Penjualan Berbasis Web Pada Doni Sport Malang," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 4, no. 1, pp. 129–136, 2020, doi: 10.36040/jati.v4i1.2745.
- [4] S. Putri and C. Astiti, "Penerapan Metode Least Square Dalam Perhitungan Proyeksi Jumlah Penduduk," *SEPREN (Journal Math. Educ. Applied)*, vol. 04, no. 02, pp. 147–154, 2023, doi: 10.36655/sepren.v4i1.
- [5] I. Rahmawati and R. Wijanarko, "Implementasi Prediksi Penjualan Obat Menggunakan Metode Least Square Pada Apotek Demak Farma Jaya," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 1, no. 1, pp. 19–23, 2019, doi: 10.36499/jinrpl.v1i1.2762.
- [6] M. Rahmawita and I. Fazri, "Aplikasi Peramalan Penjualan Obat Menggunakan Metode Least Squaredi Rumah Sakit Bhayangkara," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 4, no. 2, p. 201, 2018, doi: 10.24014/rmsi.v4i2.5685.
- [7] I. Dewi, Nindian Puspa; Listiowarni, "Peramalan Harga Bahan Proyek Menggunakan Metode Least Square (Studi Kasus : CV Rizky Mulya)," *J-TFA (Jurnal Teknol. Inform.)*, vol. 2, no. 1, pp. 1–33, 2019, [Online]. Available: <https://jurnal.umm.ac.id/index.php/J-TIFA/article/view/240/150>
- [8] J. S. Widjaya, D. Agushinta R, and S. R. Puspita Sari, "Sistem Prediksi Jumlah Pasien Covid-19 Menggunakan Metode Trend Least Square Berbasis Web," *Sistemasi*, vol. 10, no. 1, pp. 39–51, 2021, doi: 10.32520/stmsi.v10i1.1036.
- [9] R. W. N. Octavia and U. Chotijah, "Implementasi Metode Least Square Untuk Prediksi Penjualan Kue Donat dan Bomboloni," *Jutisi (Jurnal Ilm. Tek. Inform. dan Sist. Informasi)*, vol. 11, no. 1, pp. 251–262, 2022, doi: 10.35889/jutisi.v11i1.802.
- [10] D. Swanjaya and D. Putra Pamungkas, "Analisa Hasil Prediksi Metode Least Square menggunakan Korelasi dan MAPE pada Toko PS," *Gener. J.*, vol. 5, no. 1, pp. 11–18, 2021, doi: 10.29407/gj.v5i1.15440.
- [11] D. Nasyika, S. Slamim, and P. Pandunata, "Sistem Prediksi Jumlah Permintaan Produk Darah Menggunakan Metode Least Square Regression Line (Studi Kasus : Utd Pmi Kabupaten Jombang)," *INFORMAL Informatics J.*, vol. 3, no. 2, pp. 48–55, 2018, doi: 10.19184/isj.v3i2.9989.
- [12] S. R. Garuda and Y. Baliadi, "Aplikasi Metode Tren Waktu Satu Ragam Dalam Peramalan Toleransi Komoditas Pangan Terhadap Perubahan Iklim Di Papua," *J. Inform. Pertanian.*, vol. 27, no. 1, pp. 35–46,

- 2018.
- [13] A. Rifa'i, "The Statistical Parabolic Projection Method Untuk Forecasting," *J. Manaj. dan Bisni*, vol. 8, no. 2, pp. 354–365, 2019, doi: 10.34006/jmbi.v8i2.137.
- [14] D. Pratama and M. Ardiansyah, "Analisis Perbandingan Metode Least Square dan Parabolik untuk Perhitungan Forecasting Penjualan Minuman Pada Kedai Rumah Celoteh," *INFORMATIKA*, vol. 3, no. 1, pp. 78–91, 2022.
- [15] I. Simamora, "Metode Trend Non Linear Untuk Forecasting Komposisi Penduduk Kabupaten Tapanuli Tengah Menurut Jenis Kelamin Tahun 2006-2019," *J. Curere*, vol. 02, no. 02, pp. 175–183, 2018, doi: 10.36764/jc.v2i2.
- [16] R. R. Pratama, "Analisis Model Machine Learning Terhadap Pengenalan Aktifitas Manusia," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 19, no. 2, pp. 302–311, 2020, doi: 10.30812/matrik.v19i2.688.
- [17] W. Kurniadi, "Pendukung Keputusan Dalam Peramalan Penjualan Ayam Broiler Dengan Metode Trend Moment Dan Simple Moving Average Pada CV. Merdeka Adi Perkasa," *J. Media Inform. Budidarma*, vol. 2, no. 3, pp. 76–90, 2018, doi: 10.30865/mib.v2i3.652.
- [18] M. Hatta and A. Fauziah Fitri, "Sistem Prediksi Persediaan Stok Darah Dengan Metode Least Square Pada Unit Transfusi Darah Studi Kasus PMI Kota Cirebon," *J. Ilm. Ilmu Komput.*, vol. 6, no. 1, pp. 41–45, 2020, doi: 10.35329/jiik.v6i1.130.
- [19] R. Maulana Fauzi and D. Iskandar Mulyana, "Implementasi Data Mining Menggunakan Metode Least Square untuk Memprediksi Penjualan Lampu LED pada PT. Sumber Dinamika Solusitama," *J. Sos. Teknol.*, vol. 1, no. 8, pp. 907–919, 2021, doi: 10.36418/jurnalsostech.v1i8.182.
- [20] I. Zikri, S. Safrida, E. Susanti, and R. A. Putri, "Analysis of trend and determinant factors of imported soybean in the period of 2003-2022," *Adv. Food Sci. Sustain. Agric. Agroindustrial Eng.*, vol. 3, no. 1, pp. 17–24, 2020, doi: 10.21776/ub.afssae.2020.003.01.3.
- [21] A. Navlani, A. Fandango, and I. Idris, *Python Data Analysis: Perform Data Collection, Data Processing, Wrangling, Visualization, and Model Building Using Python*, Third edit. Birmingham, Mumbai: Pact Publishing Ltd, 2021.
- [22] W. M. Baihaqi, M. Dianingrum, and K. A. N. Ramadhan, "Regresi Linier Sederhana Untuk Memprediksi Kunjungan Pasien Di Rumah Sakit Berdasarkan Jenis Layanan Dan Umur Pasien," *J. Simetris*, vol. 10, no. 2, pp. 671–680, 2019, [Online]. Available: <https://www.jurnal.umk.ac.id/index.php/simet/article/view/3484>.