

ARIMA (AUTOREGRESSIVE INTEGRATED MOVING AVERAGE) SEBAGAI ALGORITMA PREDIKSI FAKTOR PEMBATALAN HEREGISTRASI MAHASISWA BARU DI UNIVERSITAS X

Verra Sri Yulia Rahmawati^{a,1}, Iwan Rizal Setiawan^{b,2}, Asriyanik^{c,3}

^{a,b,c} Universitas Muhammadiyah Sukabumi, Jl. R. Syamsudin, S.H. No. 50, Sukabumi, Jawa Barat, Indonesia

¹ verrasyr084@ummi.ac.id*, ² metalizer_5150@ummi.ac.id, ³ asriyanik263@ummi.ac.id

* corresponding author

ARTICLE INFO

Keywords

ARIMA
SEMMA
Prediction
Heregistration
Time Series Analysis

ABSTRACT

The digitalization era has transformed education, impacting new student admissions in Indonesia, which has various universities: State, Official, Religious, and Private. These universities share common procedures for admitting new students: file selection, test selection, and re-registration. However, many new students cancel their re-registration due to financial constraints, distance, or choosing another campus. Some students neglect the re-registration process until the deadline passes, affecting the accreditation of study programs and the reputation of the campus. To address this issue, a prediction model for re-registration cancellation rates can evaluate campus performance in attracting new students. The ARIMA algorithm (AutoRegressive Integrated Moving Average) is proposed as a suitable prediction model for time series data. This model can help universities identify and address factors leading to re-registration cancellations, thereby improving their performance and reputation. Using the SEMMA (Sample, Explore, Modify, Model, Assess) data mining methodology, the research produced an evaluation matrix with RMSE (Root Mean Square Error) values for various features: "non_registered" (145.77), "parents_income" (0.84), "parents_job" (4.07), and "entrance" (0.16). Additionally, the correlation matrix revealed two variables with a high influence on the target: "entrance" (0.85) and "parents_income" (0.68).

1. Pendahuluan

Seleksi mahasiswa baru di era zaman sekarang secara menyeluruh sudah menggunakan *platform* seperti aplikasi dan *web*. Alur pendaftaran mahasiswa baru telah diatur oleh peraturan yang dikeluarkan oleh Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi No 48 Tahun 2022 Pasal 4 ialah seperti Seleksi Nasional Berdasarkan Prestasi (SNBP), Seleksi Nasional Berdasarkan Tes (SNBT), dan Seleksi Mandiri. Secara garis besar, mahasiswa baru perlu melalui tahapan seleksi, seperti seleksi administrasi, seleksi tes, dan registrasi ulang. Registrasi ulang mahasiswa baru adalah langkah penting yang menentukan keberlanjutan pendidikan, tetapi fenomena pembatalan registrasi ulang telah menjadi masalah yang semakin meningkat di lingkungan pendidikan tinggi Indonesia. Fenomena ini tidak hanya berdampak pada institusi pendidikan tinggi tertentu, tetapi juga berdampak pada sistem pendidikan tinggi nasional secara keseluruhan. namun isu yang diangkat mencerminkan tantangan yang dihadapi oleh banyak perguruan tinggi swasta di Indonesia.

Penelitian ini berfokus pada salah satu universitas swasta di Jawa Barat, namun isu yang diangkat mencerminkan tantangan yang dihadapi oleh banyak perguruan tinggi swasta di Indonesia. Menurut [1] terdapat empat isu utama yang dihadapi oleh perguruan tinggi swasta yaitu sukarnya dukungan dana, relatif tinggi pada biaya kuliah, kualitas yang bervariasi, dan perluasan akses. Kemudian, ditemukannya fenomena seperti mahasiswa baru yang dengan sengaja mengabaikan proses pendaftarannya hingga periode selesai [2]. Hal tersebut dapat mempengaruhi akreditasi dari program studi serta berkurangnya nilai dari suatu universitas [3]. Berdasarkan permasalahan di atas, hal tersebut memiliki korelasi dengan permasalahan yang terjadi di Universitas X, terlebih pada permasalahan biaya kuliah yang relatif tinggi, dan mahasiswa baru cenderung memilih Perguruan Tinggi Negeri.

Seperti halnya pada penelitian terkait penggunaan ARIMA yang telah dilakukan oleh Dienda dan Christine dengan judul “*Perbandingan Algoritma ARIMA, Prophet, dan LSTM dalam Prediksi Penjualan Tiket Wisata Taman Hiburan (Studi Kasus: Saloka Theme Park)*” menunjukkan bahwa algoritma ARIMA (1,0,1) memiliki kinerja yang lebih unggul jika dibandingkan dengan LSTM dan Prophet dengan skor RMSE sebesar 762.009 [4]. Selain itu, penelitian lain telah dilakukan oleh [5] dengan judul “*Perbandingan Prediksi Harga Saham Dengan Model ARIMA Dan Artificial Neural Network*” menyatakan bahwa skor RMSE ARIMA lebih kecil daripada ANN, yakni dengan skor 1.3738 sedangkan ANN memiliki skor 4.6814. Penelitian lain telah dilakukan oleh [6], yang memiliki judul “*Sales Forecasting pada Dealer Motor X Dengan LSTM, ARIMA dan Holt-Winters Exponential Smoothing*” menunjukkan bahwa ARIMA menghasilkan skor RMSE paling kecil untuk enam jenis motor jika disandingkan Holt-Winters Exponential Smoothing dan LSTM, serta pada tiga jenis motor ARIMA menghasilkan skor MAPE yang rendah, LSTM pada penelitian ini kurang cocok digunakan karena jumlah *data set* yang hanya 84 data. Sehingga dapat disimpulkan penggunaan ARIMA dalam penelitian kali ini adalah pilihan yang terbaik.

Meninjau pada pemaparan dari masalah yang sudah disampaikan sebelumnya, penelitian ini bertujuan untuk membuat model menggunakan ARIMA agar dapat memprediksikan faktor pembatalan registrasi ulang mahasiswa baru selama lima tahun ke depan serta menentukan variabel mana yang paling mempengaruhi pembatalan registrasi ulang mahasiswa baru di Universitas X, sehingga dapat membantu pihak Penerimaan Mahasiswa Baru (PMB) dalam menentukan strategi pemasaran yang efektif untuk menggait kepercayaan mahasiswa baru. Secara umum, temuan dalam penelitian ini meliputi mengidentifikasi faktor-faktor yang dapat mempengaruhi keputusan mahasiswa baru dalam melanjutkan atau membatalkan registrasi ulang. Evaluasi model penelitian ini menggunakan *Root Mean Square Error* (RMSE). Adapun kebaruan yang dapat ditawarkan yaitu dari objek penelitiannya. ARIMA sering kali digunakan untuk memprediksikan harga atau penjualan, dalam kasus penelitian ini ARIMA digunakan untuk memprediksikan faktor pembatalan registrasi ulang mahasiswa baru.

2. Metodologi Penelitian

Agar mendukung keberlangsungan penelitian, diperlukannya sebuah metode untuk menciptakan penelitian yang terstruktur serta terarah. Penelitian ini menggunakan metodologi *data mining* SEMMA (*Sample, Explore, Modify, Model, Assess*). Pemilihan SEMMA dikarenakan metode ini lebih ringkas daripada metode lainnya selain itu SEMMA juga memiliki sifat yang lebih fleksibel jika dibandingkan dengan CRISP-DM dan KDD. SEMMA mengurangi kompleksitasnya dalam mengolah model sehingga tidak ditemukannya fitur seperti manajemen proyek seperti di CRISP-DM. *Tools* yang digunakan pada penelitian ini dengan menggunakan bahasa pemrograman Python dan juga Google Colab. Proses-proses dalam SEMMA sebagai berikut:

2.1 *Sample*

Tahap ini dimulai dengan mengambil sampel data yang mewakili dari keseluruhan data yang ada. Sampel data ditampilkan dengan jumlah data sebanyak 5 yang merepresentasikan data teratas. Data historis yang mencakup atribut-atribut yang mempengaruhi faktor pembatalan mahasiswa baru diambil dari *ICT Center* Universitas X. Sebanyak 8137 data digunakan untuk penelitian ini.

2.2 *Explore*

Data perlu diperiksa untuk menilai kualitasnya, termasuk kelengkapan dan konsistensinya. Pada tahap ini, dilakukan eksplorasi data untuk memahami pola, karakteristik, dan struktur data. Tujuan eksplorasi ini adalah untuk mengidentifikasi distribusi data, mendeteksi *outlier*.

2.3 *Modify*

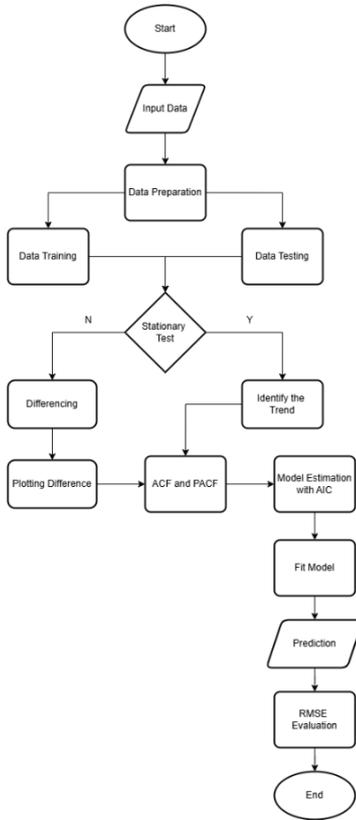
Selama tahap ini, dilakukan pembersihan data dengan menghilangkan data-data yang hilang, terduplikat, atau tidak konsisten. Transformasi data seperti konversi jenis data kategorial menjadi numerik. Selain itu, seleksi fitur dilakukan untuk memilih atribut-atribut relevan dalam memprediksi faktor pembatalan mahasiswa baru.

2.4 *Modelling*

Model penelitian ini menggunakan metode ARIMA. Metode ARIMA dipilih sebagai algoritma prediksi karena kemampuannya dalam memprediksi data *timeseries*. Seperti halnya menurut [7] perbandingan ARIMA dengan LSTM yakni ARIMA memiliki keunggulan dalam pengimplementasian serta interpretasi yang mudah, kemudian ARIMA lebih efektif menangani data yang bersifat linear, juga ARIMA memiliki proses yang lebih cepat dalam pelatihannya. Model dibuat berdasarkan pada dua jenis data yaitu data latih untuk melatih model dan uji untuk mengevaluasi kinerja model. Alur model dapat dilihat pada Gambar 1.

2.5 *Assess*

Setelah model dibuat, langkah selanjutnya adalah mengevaluasinya menggunakan Root Mean Square Error (RMSE) sebagai matriks evaluasi. Evaluasi dilakukan untuk mengukur kinerja model dengan memperhatikan tingkat kesalahan yang dapat menjadi acuan untuk tingkat akurasi model.



Gambar 1. Flowchart Model

3. Hasil dan Pembahasan

3.1 Sample

Penelitian ini menggunakan data yang bersumber dari basis data *ICT Center* Universitas X. Data yang didapatkan merupakan data mahasiswa baru dari tahun 2015 hingga 2024 dengan jumlah atribut sebanyak 8 atribut. Pemilihan atribut yang dijadikan target ialah pada atribut “*is_registered*”. Data mahasiswa baru yang dimaksud tertera pada Gambar 2. yang mana merupakan sampel dari 5 data teratas, data sebagai berikut:

	department	faculty	is_registered	parents_income	parents_job	registration_date	accreditation	entrance
0	TEKNIK INFORMATIKA	FAKULTAS SAINS DAN TEKNOLOGI	1	3.000.000 - 5.000.000	Wiraswasta	4/3/2022	Baik Sekali	Non Tes
1	KEPERAWATAN	FAKULTAS KESEHATAN	1	3.000.000 - 5.000.000	Pegawai Negeri	6/3/2021	Baik	Tes
2	PENDIDIKAN BAHASA DAN SASTRA INDONESIA	FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN	1	3.000.000 - 5.000.000	Wiraswasta	8/5/2019	B	Non Tes
3	ADMINISTRASI BISNIS	FAKULTAS ILMU SOSIAL	1	< 3.000.000	Lain-lain	8/25/2022	Baik Sekali	Non Tes
4	AKUNTANSI	FAKULTAS EKONOMI	1	< 3.000.000	Lain-lain	4/12/2021	B	Non Tes

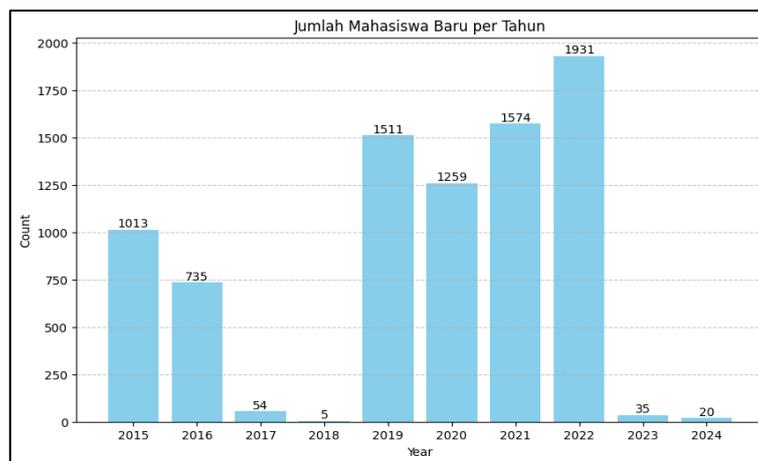
Gambar 2. Sampel Data Mahasiswa Baru

3.2 Explore

Eksplorasi data pada penelitian ini dengan melakukan identifikasi tipe data seperti pada Gambar 3., yang menghasilkan 7 tipe data *object* dan 1 tipe data *integer*. Lalu, melakukan penyajian data berdasarkan distribusi kelas dari setiap atribut. Serta, mengidentifikasi adanya *outlier* pada data dengan melihat Gambar 4., yang mana terdapat data yang *outlier* yaitu pada tahun 2017, 2018, 2023, dan 2024. Namun, penelitian ini hanya mengambil data 5 tahun ke belakang yakni 2019 hingga 2023. Tahun 2023 tetap digunakan karena atribut “*registration_date*” bukanlah target dari penelitian ini.

department	object
faculty	object
is_registered	int64
parents_income	object
parents_job	object
registration_date	object
accreditation	object
entrance	object
dtype:	object

Gambar 3. Tipe Data



Gambar 4. Diagram Jumlah Mahasiswa Baru Per Tahun

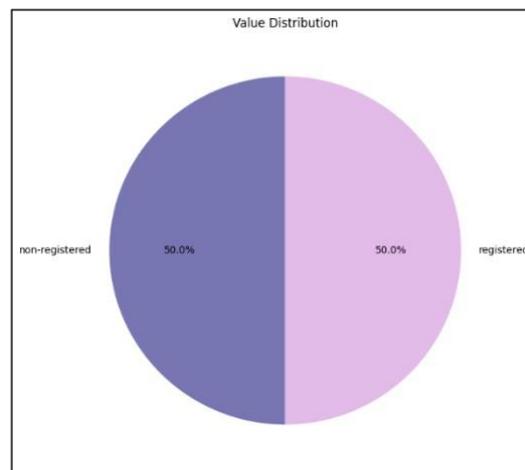
3.3 Modify

Proses *modify* dilakukan agar menghindari data bermasalah sebelum dilakukannya *modelling*. *Modify* yang dilakukan seperti pembersihan data, transformasi data, seleksi fitur dan menyeimbangkan data. Pada proses pembersihan data, dibagi menjadi tiga tahapan yaitu mencari nilai kosong dengan hasil sebanyak 5104 data penanganan nilai kosong dengan mengisi nilai modus pada 2 atribut, 1 atribut menyesuaikan dengan atribut lain, dan 2 atribut dilakukan penghapusan. Selanjutnya, mencari nilai duplikat, ditemukannya 245 baris data yang terduplikat. Kemudian, tahap transformasi data, mengubah data kategorial ke numerik menggunakan fungsi *OrdinalEncode* agar hasil transformasi tidak terbatas hanya 1 dan 0. Tahap selanjutnya yaitu seleksi fitur menggunakan ANOVA yang menghasilkan nilai seperti pada Gambar 5., berdasarkan hasil seleksi fitur menggunakan ANOVA, dipilihnya fitur “*parents_job*”, “*entrance*”, “*parents_income*” dan “*registration_date*”. Terakhir, data dilakukan penyeimbangan

dengan metode *undersampling*, hasil *undersampling* menghasilkan rasio 50:50 dengan jumlah data menjadi 2200 pada target yang dapat dilihat pada Gambar 6.

Feature Scores (ANOVA statistics):			
	Feature	Score	P-value
3	parents_job	156.026680	2.184533e-35
6	entrance	126.498929	4.515751e-29
2	parents_income	13.116708	2.949727e-04
4	registration_date	7.256903	7.081591e-03
1	faculty	2.351899	1.251806e-01
5	accreditation	1.704942	1.916905e-01
0	department	0.410982	5.214950e-01

Gambar 5. Hasil Seleksi Fitur



Gambar 6. Hasil Undersampling

3.4 Modelling

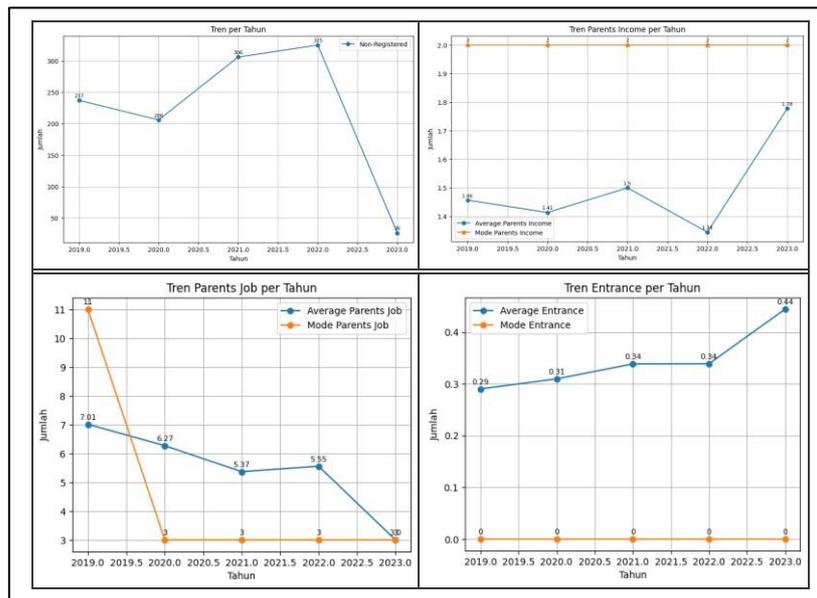
Proses *modelling* mengikuti langkah-langkah pada Gambar 1., dimulai dengan pembagian data dengan rasio 70:30 di mana 70% adalah data *training* yang berasal dari data 2019 – 2021, dan 30% data *testing* yang berasal dari data 2022 – 2023. Kemudian, data *timeseries* perlu dilakukannya uji stasioneritas agar hasil akurasi dapat maksimal, uji KPSS (*Kwiatkowski Phillips Schmidt Shin*) pada penelitian ini untuk empat fitur yang disajikan pada Tabel 1.

Tabel 1. Hasil Uji Stasioner

Fitur	KPSS Statistik	P-Value	Nilai Kritis			
			1%	2.5%	5%	10%
<i>non_registered</i>	0.20957495429616088	0.1				
<i>parents_income</i>	0.40000000000000005	0.07715517241379288	0.73	0.57	0.46	0.34
<i>parents_job</i>	0.42918088320896247	0.06457720551337826				
<i>entrance</i>	0.41906737764921337	0.06893647515120113				

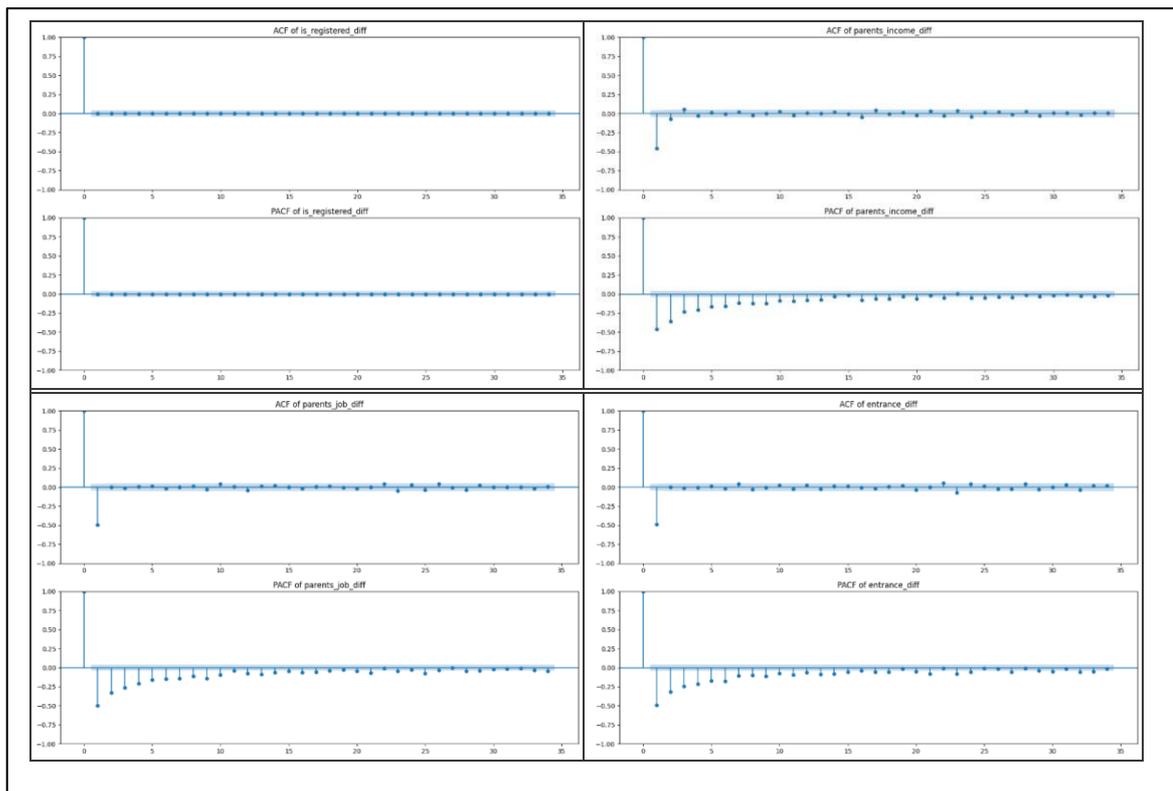
Kemudian mengidentifikasi tren per tahun dari seluruh fitur, parameter tren dibagi menjadi dua yaitu berdasarkan rata-rata dan modus, parameter ini hanya berlaku pada tiga fitur yaitu "*parents_income*",

“parents_job” dan “entrance”, tren dapat dilihat pada beberapa Gambar 7., yang dapat disimpulkan jika parameter data menggunakan rata-rata grafik cenderung variatif, sedangkan jika parameter data menggunakan modus bentuk grafik cenderung konstan.



Gambar 7. Tren Per Tahun Seluruh Fitur

Selanjutnya, untuk mencari nilai parameter p (*AutoRegressive*) dan q (*Moving Average*) menggunakan plot ACF (*Autocorrelation*) dan PACF (*Partial Autocorrelation*). Menurut [8] penggunaan plot ACF untuk mengukur keterhubungan dari nilai deret waktu antara waktu sekarang dan periode sebelumnya, sedangkan plot PACF berfungsi untuk mengukur korelasi antara variabel serta deret waktu yang kemudian dipertimbangkannya pengaruh dalam variabel. Plot ACF dan PACF dapat dilihat pada Gambar 8.



Gambar 8. Hasil Plot ACF dan PACF

Berdasarkan Gambar 8., seluruh fitur pada $lag\ 0 = 1$, hal ini karena pengukuran hubungan terhadap deret waktu tersebut mengukur dirinya sendiri [9]. Lalu, pada fitur “*is_registered*” menunjukkan nilai $lag\ 1$ dan selanjutnya = 0, menunjukkan data sudah stasioner. Selanjutnya, pada fitur “*parents_income*”, “*parents_job*”, dan “*entrance*” menunjukkan adanya korelasi negatif signifikan, namun kian menurun pada $lag\ selanjutnya$.

Tahap selanjutnya yaitu melakukan estimasi model ARIMA dengan tolak ukur skor AIC (*Akaike Information Criterion*). Apabila skor AIC semakin minim, maka model tersebutlah yang akan dipilih, hasil estimasi model ARIMA dapat dilihat pada Gambar 9.

ARIMA(0, 0, 0) AIC=64.85589079411294	ARIMA(0, 0, 0) AIC=-0.8705886583421734	ARIMA(0, 0, 0) AIC=21.194828290218993	ARIMA(0, 0, 0) AIC=-11.12452814513248
ARIMA(0, 0, 1) AIC=64.59024059907436	ARIMA(0, 0, 1) AIC=-1.0952088650965601	ARIMA(0, 0, 1) AIC=23.0385265713771	ARIMA(0, 0, 1) AIC=-9.317615483812288
ARIMA(0, 0, 2) AIC=65.14776825768924	ARIMA(0, 0, 2) AIC=-0.15358444586495956	ARIMA(0, 0, 2) AIC=24.04321180431403	ARIMA(0, 0, 2) AIC=-8.024082699928164
ARIMA(0, 0, 3) AIC=66.13767185362413	ARIMA(0, 0, 3) AIC=-1.378241913369516	ARIMA(0, 0, 3) AIC=25.83386672538738	ARIMA(0, 0, 3) AIC=-6.388285176749503
ARIMA(0, 0, 4) AIC=67.39692313312875	ARIMA(0, 0, 4) AIC=3.6183514442349605	ARIMA(0, 0, 4) AIC=27.13400954414297	ARIMA(0, 0, 4) AIC=-5.052064841355545
ARIMA(0, 1, 0) AIC=53.88684876945882	ARIMA(0, 1, 0) AIC=1.7638132158762234	ARIMA(0, 1, 0) AIC=16.08508559727681	ARIMA(0, 1, 0) AIC=-9.766355949530168
ARIMA(0, 1, 1) AIC=55.16879142152544	ARIMA(0, 1, 1) AIC=2.6053911385368006	ARIMA(0, 1, 1) AIC=18.085085786425566	ARIMA(0, 1, 1) AIC=-7.778468299085557
ARIMA(0, 1, 2) AIC=56.99855020720294	ARIMA(0, 1, 2) AIC=4.02232805965054	ARIMA(0, 1, 2) AIC=20.90744838317383	ARIMA(0, 1, 2) AIC=-6.466294914478534
ARIMA(0, 1, 3) AIC=57.753898647427	ARIMA(0, 1, 3) AIC=5.800182516421699	ARIMA(0, 1, 3) AIC=20.874112527697953	ARIMA(0, 1, 3) AIC=-4.521218177463382
ARIMA(0, 1, 4) AIC=60.66292368247956	ARIMA(0, 1, 4) AIC=7.6764121555820815	ARIMA(0, 1, 4) AIC=22.59551799663948	ARIMA(0, 1, 4) AIC=-2.537553013836653
ARIMA(1, 0, 0) AIC=66.1733038361383	ARIMA(1, 0, 0) AIC=-1.0741905135418381	ARIMA(1, 0, 0) AIC=22.987680881601466	ARIMA(1, 0, 0) AIC=-9.3529424137466
ARIMA(1, 0, 1) AIC=66.12355467989235	ARIMA(1, 0, 1) AIC=-0.2536342252592334	ARIMA(1, 0, 1) AIC=24.9872702978249	ARIMA(1, 0, 1) AIC=-7.36099724195650055
ARIMA(1, 0, 2) AIC=68.52269651178466	ARIMA(1, 0, 2) AIC=1.3696964274251329	ARIMA(1, 0, 2) AIC=25.958893785632398	ARIMA(1, 0, 2) AIC=-6.068553319966437
ARIMA(1, 0, 3) AIC=67.8570207338802	ARIMA(1, 0, 3) AIC=3.4100585689516567	ARIMA(1, 0, 3) AIC=27.512973821892384	ARIMA(1, 0, 3) AIC=-4.636255274728352
ARIMA(1, 1, 0) AIC=69.2811921209244	ARIMA(1, 1, 0) AIC=5.201706628275583	ARIMA(1, 1, 0) AIC=29.13117896817218	ARIMA(1, 1, 0) AIC=-3.052450822447701
ARIMA(1, 1, 1) AIC=67.55140063799804	ARIMA(1, 1, 1) AIC=4.8552509139537685	ARIMA(1, 1, 1) AIC=28.085015558560727	ARIMA(1, 1, 1) AIC=-7.79574986874834
ARIMA(1, 1, 2) AIC=58.9029490151288	ARIMA(1, 1, 2) AIC=3.7153633023783352	ARIMA(1, 1, 2) AIC=29.532145041278024	ARIMA(1, 1, 2) AIC=-6.209617680882355
ARIMA(1, 1, 3) AIC=60.71013087209171	ARIMA(1, 1, 3) AIC=5.66828077470717	ARIMA(1, 1, 3) AIC=28.72443921246744	ARIMA(1, 1, 3) AIC=-4.550758253858389
ARIMA(1, 1, 4) AIC=62.65018599205937	ARIMA(1, 1, 4) AIC=7.668280916228006	ARIMA(1, 1, 4) AIC=22.875560371698977	ARIMA(1, 1, 4) AIC=-2.5123864113056586
ARIMA(2, 0, 0) AIC=66.197732682757	ARIMA(2, 0, 0) AIC=9.668281004859445	ARIMA(2, 0, 0) AIC=24.368395469934088	ARIMA(2, 0, 0) AIC=-0.608553319966437
ARIMA(2, 0, 1) AIC=68.10946647963785	ARIMA(2, 0, 1) AIC=0.49628022333256486	ARIMA(2, 0, 1) AIC=24.985914334974556	ARIMA(2, 0, 1) AIC=-7.38454319641773
ARIMA(2, 0, 2) AIC=69.85525131893152	ARIMA(2, 0, 2) AIC=2.6009224327893614	ARIMA(2, 0, 2) AIC=26.98149933632378	ARIMA(2, 0, 2) AIC=-5.95743849377619
ARIMA(2, 0, 3) AIC=70.66424581839328	ARIMA(2, 0, 3) AIC=3.3078931706912413	ARIMA(2, 0, 3) AIC=27.113407577143732	ARIMA(2, 0, 3) AIC=-5.23744495676808585
ARIMA(2, 0, 4) AIC=71.8646482255472	ARIMA(2, 0, 4) AIC=5.58135408111131	ARIMA(2, 0, 4) AIC=28.29806289264632	ARIMA(2, 0, 4) AIC=-3.980296867608366
ARIMA(2, 1, 0) AIC=57.6259627814646	ARIMA(2, 1, 0) AIC=6.949085248148278	ARIMA(2, 1, 0) AIC=31.112515811537525	ARIMA(2, 1, 0) AIC=-1.052871921670716
ARIMA(2, 1, 1) AIC=58.68064718329071	ARIMA(2, 1, 1) AIC=1437.14431667254	ARIMA(2, 1, 1) AIC=52358.85170755006	ARIMA(2, 1, 1) AIC=-0.34015440485547
ARIMA(2, 1, 2) AIC=60.6175043026679	ARIMA(2, 1, 2) AIC=5.668281423930933	ARIMA(2, 1, 2) AIC=20.61170513887848	ARIMA(2, 1, 2) AIC=-2.55487376157635
ARIMA(2, 1, 3) AIC=62.58179957545707	ARIMA(2, 1, 3) AIC=7.668280761542834	ARIMA(2, 1, 3) AIC=22.380421927675098	ARIMA(2, 1, 3) AIC=-0.554828057454313
ARIMA(2, 1, 4) AIC=64.64410647794847	ARIMA(2, 1, 4) AIC=9.66828092410446	ARIMA(2, 1, 4) AIC=24.364736158392457	ARIMA(2, 1, 4) AIC=-0.554828057454313
ARIMA(3, 0, 0) AIC=69.6016455380923	ARIMA(3, 0, 0) AIC=11.668280862164266	ARIMA(3, 0, 0) AIC=26.406975524008903	ARIMA(3, 0, 0) AIC=-4.451745673758705
ARIMA(3, 0, 1) AIC=70.3638966749488	ARIMA(3, 0, 1) AIC=3.3078931706912413	ARIMA(3, 0, 1) AIC=26.47452063727814	ARIMA(3, 0, 1) AIC=-5.962454223824629
ARIMA(3, 0, 2) AIC=70.44118806283332	ARIMA(3, 0, 2) AIC=4.453782816980707	ARIMA(3, 0, 2) AIC=27.17829945726607	ARIMA(3, 0, 2) AIC=-5.174391967490472
ARIMA(3, 0, 3) AIC=71.52489257700174	ARIMA(3, 0, 3) AIC=5.251185935056662	ARIMA(3, 0, 3) AIC=28.38662239968631	ARIMA(3, 0, 3) AIC=-3.7496638116045027
ARIMA(3, 0, 4) AIC=72.7554828280543	ARIMA(3, 0, 4) AIC=7.108182803887939	ARIMA(3, 0, 4) AIC=29.89291419020918	ARIMA(3, 0, 4) AIC=-2.3009050712215344
ARIMA(3, 1, 0) AIC=53.40995343913988	ARIMA(3, 1, 0) AIC=8.79257743722201	ARIMA(3, 1, 0) AIC=32.2109687989599	ARIMA(3, 1, 0) AIC=-0.04283521347319663
ARIMA(3, 1, 1) AIC=59.24766565977623	ARIMA(3, 1, 1) AIC=5.668280846516426	ARIMA(3, 1, 1) AIC=30.364726543612313	ARIMA(3, 1, 1) AIC=-4.54826124931321
ARIMA(3, 1, 2) AIC=58.46399701815117	ARIMA(3, 1, 2) AIC=7.668280847911273	ARIMA(3, 1, 2) AIC=32.36472796218205	ARIMA(3, 1, 2) AIC=-2.554828057454313
ARIMA(3, 1, 3) AIC=62.92822174626105	ARIMA(3, 1, 3) AIC=9.66828092410446	ARIMA(3, 1, 3) AIC=34.364726543612313	ARIMA(3, 1, 3) AIC=-0.554828057454313
ARIMA(3, 1, 4) AIC=66.58799394540093	ARIMA(3, 1, 4) AIC=13.668280866344411	ARIMA(3, 1, 4) AIC=36.4727165161647	ARIMA(3, 1, 4) AIC=-10.569817243437981
ARIMA(4, 0, 0) AIC=54.39291296689596	ARIMA(4, 0, 0) AIC=-0.854721244843569	ARIMA(4, 0, 0) AIC=20.894106671735827	ARIMA(4, 0, 0) AIC=-9.8418383065025
ARIMA(4, 0, 1) AIC=58.2934982647294	ARIMA(4, 0, 1) AIC=-3.1462835195847493	ARIMA(4, 0, 1) AIC=22.064962644349933	ARIMA(4, 0, 1) AIC=-6.416127295790709
ARIMA(4, 0, 2) AIC=72.573444599795	ARIMA(4, 0, 2) AIC=3.7799391074012245	ARIMA(4, 0, 2) AIC=24.86075152280882	ARIMA(4, 0, 2) AIC=-3.708236383494073
ARIMA(4, 0, 3) AIC=72.94078382406175	ARIMA(4, 0, 3) AIC=5.709178540744225	ARIMA(4, 0, 3) AIC=27.115724768785256	ARIMA(4, 0, 3) AIC=-2.0079699352954243
ARIMA(4, 0, 4) AIC=76.64455031012253	ARIMA(4, 0, 4) AIC=7.168367057056308	ARIMA(4, 0, 4) AIC=28.182717553062567	ARIMA(4, 0, 4) AIC=-0.554828057454313
ARIMA(4, 1, 0) AIC=56.2378228098184	ARIMA(4, 1, 0) AIC=9.668280866344411	ARIMA(4, 1, 0) AIC=32.364726543612313	ARIMA(4, 1, 0) AIC=-2.554828057454313
ARIMA(4, 1, 1) AIC=62.44828212857446	ARIMA(4, 1, 1) AIC=11.668280866344411	ARIMA(4, 1, 1) AIC=34.364726543612313	ARIMA(4, 1, 1) AIC=-0.554828057454313
ARIMA(4, 1, 2) AIC=62.557933441836596	ARIMA(4, 1, 2) AIC=11.6682809565739	ARIMA(4, 1, 2) AIC=26.364730295472086	ARIMA(4, 1, 2) AIC=-4.451745673758705
ARIMA(4, 1, 3) AIC=66.62319246282877	ARIMA(4, 1, 3) AIC=13.668280818078772	ARIMA(4, 1, 3) AIC=28.36472922650244	ARIMA(4, 1, 3) AIC=-3.445172754981868
ARIMA(4, 1, 4) AIC=68.65024110071818	ARIMA(4, 1, 4) AIC=15.668280821022812	ARIMA(4, 1, 4) AIC=30.36472906249978	ARIMA(4, 1, 4) AIC=-5.445172121425581

Gambar 9. Estimasi Model ARIMA Seluruh Fitur

Hasil estimasi model pada Gambar 9., didapat satu persatu model ARIMA terbaik dari setiap fitur. Rincian model ARIMA terbaik dapat dilihat pada Tabel 2. Seluruh *order* akan diuji dan dilihat hasil prediksi mana yang paling efektif.

Tabel 2. Hasil Pemilihan Model ARIMA

Fitur	Order	AIC
<i>non_registered</i>	(3,1,0)	53.40995343913988
<i>parents_income</i>	(0,0,1)	-1.0952088650965601
<i>parents_job</i>	(0,1,0)	16.08508559727681
<i>entrance</i>	(0,0,0)	-11.12452814513248

Selanjutnya, yaitu tahap inisiasi model ARIMA, satu per satu *order* pada Tabel 2., diuji. Di bawah ini merupakan persamaan dan kode dari ARIMA, yang dimulai dari *differencing*, estimasi komponen AR, estimasi komponen MA, dan kombinasi antar AR dan MA:

a. *Differencing*

$$y'_t = y_t - y_{t-1} \tag{1}$$

b. *Estimasi Komponen AR (AutoRegressive)*

$$y'_t = \phi_1 y'_{t-1} + \epsilon_t \tag{2}$$

c. *Estimasi Komponen MA (Moving Average)*

$$y'_t = \epsilon_t + \theta_1 y'_{t-1} \tag{3}$$

d. *Kombinasi AR dan MA untuk ARIMA*

$$y_t - y_{t-1} = \phi_1 (y_{t-1} - y_{t-2}) + \epsilon_t + \theta_1 \epsilon_{t-1} \tag{4}$$

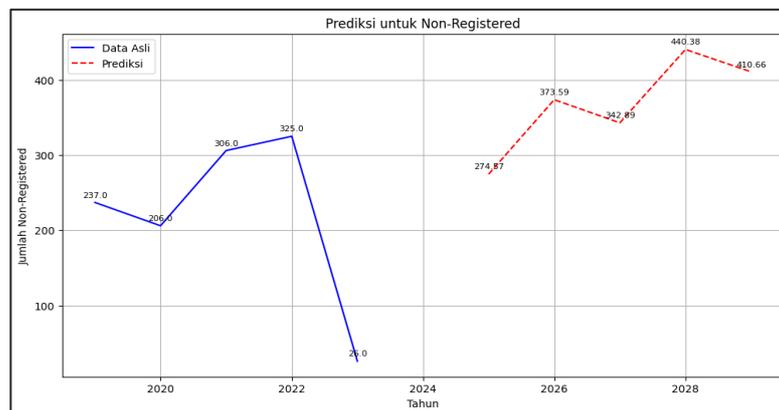
Keterangan:

- y_t : Nilai seri waktu pada waktu t
- y_{t-1} : Nilai seri waktu pada waktu $t - 1$ (satu periode sebelumnya)
- y_{t-2} : Nilai seri waktu pada waktu $t - 2$ (dua periode sebelumnya)
- ϕ_1 : Koefisien komponen *AutoRegressive* (AR)
- ϵ_t : Galat pada waktu t
- ϵ_{t-1} : Galat pada waktu t (satu periode sebelumnya)
- θ_1 : Koefisien komponen *Moving Average* (MA)

Setelah pemodelan dilakukan, didapati hasil dari masing-masing *order* yang disajikan pada Gambar 11., secara berturut-turut mulai dari kiri atas hingga kanan bawah, *order* yang dipilih ialah (3,1,0) karena hasil prediksinya yang bervariasi serta tidak menghasilkan nilai negatif. Berbeda dengan *order* lain yang menghasilkan prediksi dengan nilai yang konstan dan juga negatif. Hasil visualisasi prediksi dapat dilihat pada Gambar 12.

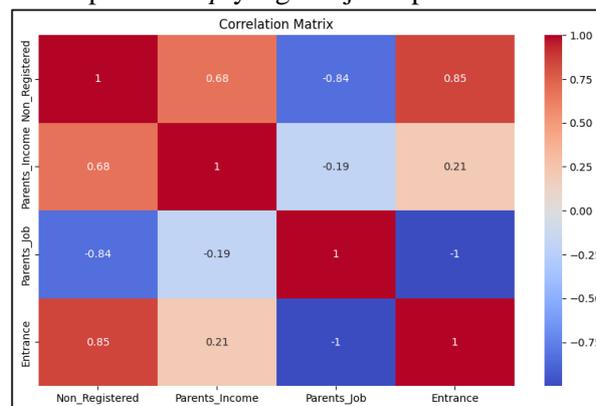
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Predicted_Non_Registration</th> <th>Predicted_Avg_Parents_Income</th> </tr> </thead> <tbody> <tr><td>3 2025</td><td>274.573348</td><td>1.438916</td></tr> <tr><td>4 2026</td><td>373.590577</td><td>1.486226</td></tr> <tr><td>5 2027</td><td>342.887855</td><td>1.441943</td></tr> <tr><td>6 2028</td><td>440.377004</td><td>1.477397</td></tr> <tr><td>7 2029</td><td>410.655135</td><td>1.449812</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Avg_Parents_Job</th> <th>Predicted_Avg_Entrance</th> </tr> </thead> <tbody> <tr><td>3</td><td>0.365129</td></tr> <tr><td>4</td><td>0.389633</td></tr> <tr><td>5</td><td>0.412216</td></tr> <tr><td>6</td><td>0.433829</td></tr> <tr><td>7</td><td>0.452211</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Mode_Parents_Income</th> <th>Predicted_Mode_Parents_Job</th> </tr> </thead> <tbody> <tr><td>3</td><td>3.003636</td></tr> <tr><td>4</td><td>3.007294</td></tr> <tr><td>5</td><td>3.007292</td></tr> <tr><td>6</td><td>3.007289</td></tr> <tr><td>7</td><td>3.007287</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Mode_Entrance</th> </tr> </thead> <tbody> <tr><td>3</td><td>0.0</td></tr> <tr><td>4</td><td>0.0</td></tr> <tr><td>5</td><td>0.0</td></tr> <tr><td>6</td><td>0.0</td></tr> <tr><td>7</td><td>0.0</td></tr> </tbody> </table>	Year	Predicted_Non_Registration	Predicted_Avg_Parents_Income	3 2025	274.573348	1.438916	4 2026	373.590577	1.486226	5 2027	342.887855	1.441943	6 2028	440.377004	1.477397	7 2029	410.655135	1.449812	Predicted_Avg_Parents_Job	Predicted_Avg_Entrance	3	0.365129	4	0.389633	5	0.412216	6	0.433829	7	0.452211	Predicted_Mode_Parents_Income	Predicted_Mode_Parents_Job	3	3.003636	4	3.007294	5	3.007292	6	3.007289	7	3.007287	Predicted_Mode_Entrance	3	0.0	4	0.0	5	0.0	6	0.0	7	0.0	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Predicted_Non_Registration</th> <th>Predicted_Avg_Parents_Income</th> </tr> </thead> <tbody> <tr><td>3 2025</td><td>221.498427</td><td>1.435275</td></tr> <tr><td>4 2026</td><td>245.299352</td><td>1.452539</td></tr> <tr><td>5 2027</td><td>245.299352</td><td>1.452539</td></tr> <tr><td>6 2028</td><td>245.299352</td><td>1.452539</td></tr> <tr><td>7 2029</td><td>245.299352</td><td>1.452539</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Avg_Parents_Job</th> <th>Predicted_Avg_Entrance</th> </tr> </thead> <tbody> <tr><td>3</td><td>0.299934</td></tr> <tr><td>4</td><td>0.312494</td></tr> <tr><td>5</td><td>0.312494</td></tr> <tr><td>6</td><td>0.312494</td></tr> <tr><td>7</td><td>0.312494</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Mode_Parents_Income</th> <th>Predicted_Mode_Parents_Job</th> </tr> </thead> <tbody> <tr><td>3</td><td>7.000335</td></tr> <tr><td>4</td><td>5.400136</td></tr> <tr><td>5</td><td>5.400136</td></tr> <tr><td>6</td><td>5.400136</td></tr> <tr><td>7</td><td>5.400136</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Mode_Entrance</th> </tr> </thead> <tbody> <tr><td>3</td><td>-0.000005</td></tr> <tr><td>4</td><td>-0.000005</td></tr> <tr><td>5</td><td>-0.000005</td></tr> <tr><td>6</td><td>-0.000005</td></tr> <tr><td>7</td><td>-0.000005</td></tr> </tbody> </table>	Year	Predicted_Non_Registration	Predicted_Avg_Parents_Income	3 2025	221.498427	1.435275	4 2026	245.299352	1.452539	5 2027	245.299352	1.452539	6 2028	245.299352	1.452539	7 2029	245.299352	1.452539	Predicted_Avg_Parents_Job	Predicted_Avg_Entrance	3	0.299934	4	0.312494	5	0.312494	6	0.312494	7	0.312494	Predicted_Mode_Parents_Income	Predicted_Mode_Parents_Job	3	7.000335	4	5.400136	5	5.400136	6	5.400136	7	5.400136	Predicted_Mode_Entrance	3	-0.000005	4	-0.000005	5	-0.000005	6	-0.000005	7	-0.000005
Year	Predicted_Non_Registration	Predicted_Avg_Parents_Income																																																																																																									
3 2025	274.573348	1.438916																																																																																																									
4 2026	373.590577	1.486226																																																																																																									
5 2027	342.887855	1.441943																																																																																																									
6 2028	440.377004	1.477397																																																																																																									
7 2029	410.655135	1.449812																																																																																																									
Predicted_Avg_Parents_Job	Predicted_Avg_Entrance																																																																																																										
3	0.365129																																																																																																										
4	0.389633																																																																																																										
5	0.412216																																																																																																										
6	0.433829																																																																																																										
7	0.452211																																																																																																										
Predicted_Mode_Parents_Income	Predicted_Mode_Parents_Job																																																																																																										
3	3.003636																																																																																																										
4	3.007294																																																																																																										
5	3.007292																																																																																																										
6	3.007289																																																																																																										
7	3.007287																																																																																																										
Predicted_Mode_Entrance																																																																																																											
3	0.0																																																																																																										
4	0.0																																																																																																										
5	0.0																																																																																																										
6	0.0																																																																																																										
7	0.0																																																																																																										
Year	Predicted_Non_Registration	Predicted_Avg_Parents_Income																																																																																																									
3 2025	221.498427	1.435275																																																																																																									
4 2026	245.299352	1.452539																																																																																																									
5 2027	245.299352	1.452539																																																																																																									
6 2028	245.299352	1.452539																																																																																																									
7 2029	245.299352	1.452539																																																																																																									
Predicted_Avg_Parents_Job	Predicted_Avg_Entrance																																																																																																										
3	0.299934																																																																																																										
4	0.312494																																																																																																										
5	0.312494																																																																																																										
6	0.312494																																																																																																										
7	0.312494																																																																																																										
Predicted_Mode_Parents_Income	Predicted_Mode_Parents_Job																																																																																																										
3	7.000335																																																																																																										
4	5.400136																																																																																																										
5	5.400136																																																																																																										
6	5.400136																																																																																																										
7	5.400136																																																																																																										
Predicted_Mode_Entrance																																																																																																											
3	-0.000005																																																																																																										
4	-0.000005																																																																																																										
5	-0.000005																																																																																																										
6	-0.000005																																																																																																										
7	-0.000005																																																																																																										
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Predicted_Non_Registration</th> <th>Predicted_Avg_Parents_Income</th> </tr> </thead> <tbody> <tr><td>3 2025</td><td>306.0</td><td>1.5</td></tr> <tr><td>4 2026</td><td>306.0</td><td>1.5</td></tr> <tr><td>5 2027</td><td>306.0</td><td>1.5</td></tr> <tr><td>6 2028</td><td>306.0</td><td>1.5</td></tr> <tr><td>7 2029</td><td>306.0</td><td>1.5</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Avg_Parents_Job</th> <th>Predicted_Avg_Entrance</th> </tr> </thead> <tbody> <tr><td>3</td><td>0.338542</td></tr> <tr><td>4</td><td>0.338542</td></tr> <tr><td>5</td><td>0.338542</td></tr> <tr><td>6</td><td>0.338542</td></tr> <tr><td>7</td><td>0.338542</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Mode_Parents_Income</th> <th>Predicted_Mode_Parents_Job</th> </tr> </thead> <tbody> <tr><td>3</td><td>3.0</td></tr> <tr><td>4</td><td>3.0</td></tr> <tr><td>5</td><td>3.0</td></tr> <tr><td>6</td><td>3.0</td></tr> <tr><td>7</td><td>3.0</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Mode_Entrance</th> </tr> </thead> <tbody> <tr><td>3</td><td>0.0</td></tr> <tr><td>4</td><td>0.0</td></tr> <tr><td>5</td><td>0.0</td></tr> <tr><td>6</td><td>0.0</td></tr> <tr><td>7</td><td>0.0</td></tr> </tbody> </table>	Year	Predicted_Non_Registration	Predicted_Avg_Parents_Income	3 2025	306.0	1.5	4 2026	306.0	1.5	5 2027	306.0	1.5	6 2028	306.0	1.5	7 2029	306.0	1.5	Predicted_Avg_Parents_Job	Predicted_Avg_Entrance	3	0.338542	4	0.338542	5	0.338542	6	0.338542	7	0.338542	Predicted_Mode_Parents_Income	Predicted_Mode_Parents_Job	3	3.0	4	3.0	5	3.0	6	3.0	7	3.0	Predicted_Mode_Entrance	3	0.0	4	0.0	5	0.0	6	0.0	7	0.0	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Predicted_Non_Registration</th> <th>Predicted_Avg_Parents_Income</th> </tr> </thead> <tbody> <tr><td>3 2025</td><td>249.666662</td><td>1.456853</td></tr> <tr><td>4 2026</td><td>249.666662</td><td>1.456853</td></tr> <tr><td>5 2027</td><td>249.666662</td><td>1.456853</td></tr> <tr><td>6 2028</td><td>249.666662</td><td>1.456853</td></tr> <tr><td>7 2029</td><td>249.666662</td><td>1.456853</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Avg_Parents_Job</th> <th>Predicted_Avg_Entrance</th> </tr> </thead> <tbody> <tr><td>3</td><td>0.312805</td></tr> <tr><td>4</td><td>0.312805</td></tr> <tr><td>5</td><td>0.312805</td></tr> <tr><td>6</td><td>0.312805</td></tr> <tr><td>7</td><td>0.312805</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Mode_Parents_Income</th> <th>Predicted_Mode_Parents_Job</th> </tr> </thead> <tbody> <tr><td>3</td><td>5.666662</td></tr> <tr><td>4</td><td>5.666662</td></tr> <tr><td>5</td><td>5.666662</td></tr> <tr><td>6</td><td>5.666662</td></tr> <tr><td>7</td><td>5.666662</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Predicted_Mode_Entrance</th> </tr> </thead> <tbody> <tr><td>3</td><td>-0.000005</td></tr> <tr><td>4</td><td>-0.000005</td></tr> <tr><td>5</td><td>-0.000005</td></tr> <tr><td>6</td><td>-0.000005</td></tr> <tr><td>7</td><td>-0.000005</td></tr> </tbody> </table>	Year	Predicted_Non_Registration	Predicted_Avg_Parents_Income	3 2025	249.666662	1.456853	4 2026	249.666662	1.456853	5 2027	249.666662	1.456853	6 2028	249.666662	1.456853	7 2029	249.666662	1.456853	Predicted_Avg_Parents_Job	Predicted_Avg_Entrance	3	0.312805	4	0.312805	5	0.312805	6	0.312805	7	0.312805	Predicted_Mode_Parents_Income	Predicted_Mode_Parents_Job	3	5.666662	4	5.666662	5	5.666662	6	5.666662	7	5.666662	Predicted_Mode_Entrance	3	-0.000005	4	-0.000005	5	-0.000005	6	-0.000005	7	-0.000005
Year	Predicted_Non_Registration	Predicted_Avg_Parents_Income																																																																																																									
3 2025	306.0	1.5																																																																																																									
4 2026	306.0	1.5																																																																																																									
5 2027	306.0	1.5																																																																																																									
6 2028	306.0	1.5																																																																																																									
7 2029	306.0	1.5																																																																																																									
Predicted_Avg_Parents_Job	Predicted_Avg_Entrance																																																																																																										
3	0.338542																																																																																																										
4	0.338542																																																																																																										
5	0.338542																																																																																																										
6	0.338542																																																																																																										
7	0.338542																																																																																																										
Predicted_Mode_Parents_Income	Predicted_Mode_Parents_Job																																																																																																										
3	3.0																																																																																																										
4	3.0																																																																																																										
5	3.0																																																																																																										
6	3.0																																																																																																										
7	3.0																																																																																																										
Predicted_Mode_Entrance																																																																																																											
3	0.0																																																																																																										
4	0.0																																																																																																										
5	0.0																																																																																																										
6	0.0																																																																																																										
7	0.0																																																																																																										
Year	Predicted_Non_Registration	Predicted_Avg_Parents_Income																																																																																																									
3 2025	249.666662	1.456853																																																																																																									
4 2026	249.666662	1.456853																																																																																																									
5 2027	249.666662	1.456853																																																																																																									
6 2028	249.666662	1.456853																																																																																																									
7 2029	249.666662	1.456853																																																																																																									
Predicted_Avg_Parents_Job	Predicted_Avg_Entrance																																																																																																										
3	0.312805																																																																																																										
4	0.312805																																																																																																										
5	0.312805																																																																																																										
6	0.312805																																																																																																										
7	0.312805																																																																																																										
Predicted_Mode_Parents_Income	Predicted_Mode_Parents_Job																																																																																																										
3	5.666662																																																																																																										
4	5.666662																																																																																																										
5	5.666662																																																																																																										
6	5.666662																																																																																																										
7	5.666662																																																																																																										
Predicted_Mode_Entrance																																																																																																											
3	-0.000005																																																																																																										
4	-0.000005																																																																																																										
5	-0.000005																																																																																																										
6	-0.000005																																																																																																										
7	-0.000005																																																																																																										

Gambar 11. Hasil Prediksi ARIMA Per Order



Gambar 12. Grafik Prediksi

Berdasarkan Gambar 12, dapat dilihat bahwa hasil prediksi mahasiswa baru yang tidak registrasi menggunakan ARIMA ditunjukkan pada grafik garis merah putus-putus. Hasil prediksi cenderung naik dan turun setiap tahunnya. Kemudian, penelitian ini mengukur faktor pembatalan registrasi ulang mahasiswa baru, yang mana untuk membuktikan pengaruh fitur terhadap target digunakannya metode *correlation matrix* untuk mengukur korelasi antar dua variabel yakni fitur dan target (*non_registered*). Visualisasi *correlation matrix* berupa *heatmap* yang disajikan pada Gambar 13.



Gambar 13. Korelasi Fitur dan Target

Berdasarkan visualisasi di Gambar 13., menunjukkan hasil matriks korelasi positif target “*non_registered*” dengan fitur “*entrance*” atau jalur masuk sebesar 0.85, dan fitur “*parents_income*” atau penghasilan orang tua sebesar 0.68. Sehingga, dapat disimpulkan dari hasil matriks korelasi terhadap faktor pengaruh pembatalan registrasi ulang mahasiswa baru di Universitas X adalah jalur masuk dan penghasilan orang tua. Jalur masuk mempengaruhi keputusan mahasiswa dalam melanjutkan studinya. Kemudian penghasilan orang tua menjadi faktor yang mempengaruhi kesanggupan dalam biaya perkuliahan selama masa studi.

Temuan dari penelitian ini menunjukkan keunggulan model ARIMA yaitu secara efektif dapat memprediksi faktor pembatalan registrasi ulang mahasiswa baru di Universitas X. Dengan penggunaan model ini, universitas dapat mengidentifikasi tren dan pola yang dapat memberikan kontribusi terhadap keputusan mahasiswa baru untuk membatalkan registrasi ulang. Hasil dari model tersebut dapat digunakan

universitas untuk meningkatkan komunikasi dengan mahasiswa baru, menawarkan lebih banyak beasiswa, serta perancangan strategi pemasaran yang efektif.

Studi serupa yang telah dilakukan oleh [10] memiliki perbedaan yang signifikan yakni pemilihan *order* terbaik ARIMA yang tidak dilakukannya perbandingan dari beberapa *order* terbaik, sedangkan penelitian ini melakukan perbandingan dari beberapa *order* terbaik ARIMA, sehingga terdapat validasi model yang lebih efektif.

3.5 Assess

Setelah dilakukan pemodelan, perlu adanya matriks evaluasi untuk mengukur tingkat kesalahan serta menjadi tolak ukur akurasi dari model tersebut. Penelitian ini menggunakan matriks evaluasi *Root Mean Square* (RMSE) yang mana uraian persamaan dan hasilnya pada Gambar 14., sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{6}$$

Keterangan:

- n : Jumlah *data set*
- y_i : Nilai sebenarnya
- \hat{y}_i : Nilai prediksi

RMSE Non-Registration: 145.7757813297085
RMSE Average Parents Income: 0.8420190389486322
RMSE Average Parents Job: 4.070971190815872
RMSE Average Entrance: 0.1680399435299618
RMSE Mode Parents Income: 1.1547005383792515
RMSE Mode Parents Job: 7.852812657387451
RMSE Mode Entrance: 0.0

Gambar 14. Hasil RMSE Model

Berdasarkan Gambar 14., hasil RMSE dari target atau “*non_registered*” menghasilkan nilai yang relatif tinggi sebesar 145.77, menunjukkan tingkat kesalahan yang cukup tinggi dan akurasi yang cukup rendah. Selanjutnya pada fitur “*parents_income*” dengan parameter rata-rata dan modus menghasilkan nilai sebesar 0.84 dan 1.15, menunjukkan tingkat kesalahan yang relatif rendah. Lalu, pada fitur “*parents_job*” secara berturut-turut menghasilkan nilai sebesar 4.07 dan 7.85, menunjukkan tingkat kesalahan cukup tinggi di urutan kedua. Diakhiri oleh fitur “*entrance*” dengan nilai berturut-turut 0.16 dan 0 yang menunjukkan tingkat kesalahan dalam prediksi terbilang sangat rendah bahkan sempurna.

4. Kesimpulan

Penelitian ini menemukan bahwa model ARIMA efektif untuk memprediksi faktor pembatalan registrasi ulang siswa baru, khususnya untuk data set waktu jangka pendek hingga menengah. Model ARIMA dengan order (3,1,0), yang dipilih berdasarkan nilai AIC dan hasil diagnosis yang menunjukkan bahwa model tidak memiliki *white noise*, dengan *p-value* > 0,05, adalah estimasi terbaik yang dihasilkan. Variabel “*entrance*” dan “*parents_income*” memiliki pengaruh terbesar terhadap pembatalan registrasi ulang, masing-masing dengan nilai 0,85 dan 0,68. Mengubah prosedur pendaftaran, menggunakan pendekatan pemasaran yang sesuai dengan tren, dan memberikan beasiswa kepada orang berpenghasilan menengah adalah beberapa contoh solusi yang dapat diterapkan. Selain itu, penelitian ini menunjukkan nilai RMSE yang beragam untuk empat fitur utama: fitur Non-Register memiliki nilai RMSE 145.775, dan nilai RMSE lainnya untuk rata-rata, modus penghasilan, pekerjaan orang tua, dan jalur masuk berkisar antara 0,168 dan 7,852.

Saran yang dapat diterapkan untuk mengatasi masalah pembatalan registrasi adalah dengan merancang serta menguji strategi pemasaran yang efektif seperti bukti dukungan finansial, layanan

kesejahteraan mahasiswa. Dilanjut dengan adanya transparansi informasi biaya kuliah, dan membangun kemitraan dengan pihak-pihak yang dapat menjangkau calon mahasiswa baru.

Daftar Pustaka

- [1] S. Wulandari and I. V. Bangun, "Isu Utama Perguruan Tinggi Swasta," *Kesejaht. Rakyat Budg. Issue Br.*, vol. 01, no. 16, pp. 1–2, 2021, [Online]. Available: <https://binus.ac.id/character-building/2020/05/kesejahteraan-rakyat/>
- [2] P. Reji, R. U. Fajri, and N. S. Puspita, "Klasifikasi Daftar Ulang Calon Mahasiswa Baru Dengan Menggunakan Metode Klasifikasi Naive Bayes," *INDEX (Informatics Digit. Expert.*, vol. 4, no. 2, pp. 70–74, 2022, [Online]. Available: <https://e-journal.unper.ac.id/index.php/informatics>
- [3] N. H. Harani and C. Prianto, "Penerapan Algoritma Adaboost Guna Menentukan Pola Masuknya Calon Mahasiswa," *J. Transform.*, vol. 18, no. 1, p. 123, 2020, doi: 10.26623/transformatika.v18i1.1606.
- [4] D. R. H. Roosaputri and C. Dewi, "Perbandingan Algoritma ARIMA, Prophet, dan LSTM dalam Prediksi Penjualan Tiket Wisata Taman Hiburan (Studi Kasus: Saloka Theme Park)," *J. Penerapan Sist. Informatika (Komputer Manajemen)*, vol. 4, no. 3, pp. 507–517, 2023.
- [5] B. yafitra Pandji, Indwiarti, and A. A. Rohmawati, "Perbandingan Prediksi Harga Saham Dengan Model Arima Dan Artificial Neural Network," *Ind. Comput.*, vol. 4, no. 2, pp. 189–198, 2019, doi: 10.21108/indojc.2019.4.2.344.
- [6] J. Soeryawinata, H. N. Palit, and L. W. Santoso, "Sales Forecasting pada Dealer Motor X Dengan LSTM, ARIMA dan Holt-Winters Exponential Smoothing," *J. Infra*, vol. 10, no. 2, pp. 1–4, 2022.
- [7] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," *Proc. - 17th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018*, no. December, pp. 1394–1401, 2018, doi: 10.1109/ICMLA.2018.00227.
- [8] E. I. Sihombing, C. D. Suhendra, and L. F. Marini, "Analisis Data Time Series Untuk Prediksi Harga Komoditas Pangan Menggunakan Autoregressive Integrated Moving Average," vol. 4, no. 6, pp. 2711–2720, 2024, doi: 10.30865/klik.v4i6.1863.
- [9] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed., vol. 19, no. 4. Monash: Otexts, 2018. doi: 10.2307/1054108.
- [10] A. P., "Higher Education Institution (HEI) Enrollment Forecasting Using Data Mining Technique," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 2, pp. 2060–2064, 2020, doi: 10.30534/ijatcse/2020/179922020.