INTEGRASI MACHINE LEARNING DAN MODEL ARIMA UNTUK PREDIKSI INFLASI DAN PDRB DI SULAWESI SELATAN



JACQLIEN PATRICIA MAWERU RUNTU D071201038



INTEGRASI MACHINE LEARNING DAN MODEL ARIMA UNTUK PREDIKSI INFLASI DAN PDRB DI SULAWESI SELATAN



JACQLIEN PATRICIA MAWERU RUNTU D071201038



INTEGRASI MACHINE LEARNING DAN MODEL ARIMA UNTUK PREDIKSI INFLASI DAN PDRB DI SULAWESI SELATAN



JACQLIEN PATRICIA MAWERU RUNTU D071201038

INTEGRASI MACHINE LEARNING DAN MODEL ARIMA UNTUK PREDIKSI INFLASI DAN PDRB DI SULAWESI SELATAN

PERNYATAAN PENGAJUAN

JACQLIEN PATRICIA MAWERU RUNTU D071201038

Skripsi

Sebagai salah satu syarat untuk mencapai gelar sarjana

Program Studi Teknik Industri

Pada

SKRIPSI

INTEGRASI MACHINE LEARNING DAN MODEL ARIMA UNTUK PREDIKSI INFLASI DAN PDRB DI SULAWESI SELATAN

HALAMAN PENGESAHAN

JACQLIEN PATRICIA MAWERU RUNTU D071201038

Skripsi,

Telah dipertahankan di depan Panitia Ujian Sarjana pada 22 November 2024 dan dinyatakan telah memenuhi syarat kelulusan Pada

Program Studi Teknik Industri Fakultas Teknik Universitas Hasanuddin Makassar

Mengesahkan: Pembimbing tugas akhir,



Dr. Ir. Saiful, S.T., M.T., IPU, ASEAN. Eng. NIP. 19810606 200604 1 004 Mengetahui: Ketua Program Studi,



Ir. Kifayah Amar, ST.,M.Sc.,Ph.D,IPU NIP. 19740621 200604 2 001

PERNYATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi berjudul "Integrasi Machine Learning dan Model ARIMA untuk Prediksi Inflasi dan PDRB di Sulawesi Selatan" adalah benar karya saya dengan arahan dari pembimbing Dr. Ir. Saiful, S.T., M.T., IPU, ASEAN. Eng. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi manapun. Sumber informasi berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan daLam daftar pustaka skripsi ini. Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut berdasarkan aturan yang berlaku.

Dengan ini saya melimpahkan hak cipta (hak ekonomis) karya tulis saya berupa skripsi ini kepada Universitas Hasanuddin.

Gowa, 22 November 2024

ME THE THE TABLE OF THE TABLE O

Jacqlien Patricia Maweru Runtu NIM D071201038

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan kehadirat Tuhan yang Maha Esa atas segala berkat dan anugerahNya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Integrasi *Machine Learning* dan Model ARIMA untuk Prediksi Inflasi dan PDRB di Sulawesi Selatan". Skripsi ini disusun dalam rangka untuk memperoleh gelar sarjana pada program studi Teknik Industri Fakultas Teknik Universitas Hasanuddin. Penulis menyadari bahwa selama penyusunan tugas akhir ini, terdapat banyak dukungan dan motivasi dari berbagai pihak. Penulis sangat merasa berterima kasih kepada:

- Tuhan Yang Maha Esa atas anugerah dan kasih yang diberikan-Nya kepada penulis serta kesempatan untuk menempuh studi S1 dengan segala berkat kemurahan-Nya. Terima kasih Tuhan untuk penyertaanMu sehingga penulis senantiasa diberikan kelancaran dan dimampukan dalam menyelesaikan segala tanggung jawab.
- 2. Orang tua saya (Papi, Mami, Daddy dan Mami Irene), Adik-adik tercinta (Christiano, Jeslyn dan Nadine), Oma Lelan, serta keluarga yang senantiasa mendoakan, membimbing, serta mendukung baik secara moral dan materil kepada penulis.
- 3. Ibu Ir. Kifayah Amar, ST, M.Sc, Ph.D,IPU selaku Ketua Departemen Teknik Industri Fakultas Teknik Universitas Hasanuddin Universitas Hasanuddin.
- 4. Bapak Dr. Ir. Saiful, ST., MT., IPU, ASEAN. Eng selaku dosen pembimbing yang sangat membantu penulis dalam menyelesaikan tugas akhir ini. Terima kasih untuk setiap bimbingan, arahan, dan masukan selama penyusunan skripsi ini.
- 5. Bapak Ir. Armin Darmawan, ST., MT., Ph.D., IPM dan Ibu Ir. A. Besse Riyani Indah, ST., MT., IPM selaku dosen penguji dalam penelitian ini. Saya berterima kasih atas kesediaan bapak dan ibu yang telah meluangkan waktu dan memberikan masukan yang sangat berharga dalam penelitian ini.
- 6. Dosen dan staff Departemen Teknik Industri Fakultas Teknik Universitas Hasanuddin.
- 7. Mas Fadhil Muhammad selaku pihak Bank Indonesia yang sangat begitu bermurah hati membantu saya selama proses penelitian berlangsung.
- 8. Terima kasih untuk Gentry Brief Senaen, terima kasih telah meluangkan waktunya buat saya selama saya mengerjakan skripsi. Dukungan, doa dan cinta yang kamu berikan telah menjadi kekuatan dalam menyelesaikan skripsi ini. Terima kasih sudah menjadi pengigat untuk selalu percaya diri dan berusaha.
- 9. Terima kasih untuk sahabat-sahabat terbaikku; *Better than your ex* (Icha, Angie, Ellen dan Uni), *Graduation* 2024 (Jenny dan Tirza). Terima kasih atas tawa, semangat, dan kehadiran kalian yang selalu membuat hari-hari lebih ringan.
- 10. Teman-teman seperjuangan Re20source yang saling membantu dan mendukung selama perkuliahan saya selama 4 tahun.
- 11. Semua pihak yang tidak dituliskan satu per satu yang telah membantu dan mendukung penulis selama ini.

ABSTRAK

JACQLIEN PATRICIA MAWERU RUNTU. Integrasi Machine Learning dan Model ARIMA untuk Prediksi Inflasi dan PDRB di Sulawesi Selatan (dibimbing oleh Saiful).

Latar Belakang. Pentingnya perkembangan teknologi informasi dan Machine Learning (ML) dalam analisis ekonomi, terutama dalam meningkatkan efisiensi dan akurasi prediksi variabel makroekonomi seperti inflasi dan Produk Domestik Regional Bruto (PDRB). ML mampu mengolah data besar dan kompleks tanpa memerlukan asumsi yang ketat, sehingga dapat mengidentifikasi pola dan hubungan yang tidak terlihat dengan metode tradisional. Implementasi ML di Bank Indonesia diharapkan mampu menghasilkan prediksi yang lebih konsisten dan akurat, mendukung pengambilan keputusan yang lebih tepat dalam kebijakan ekonomi, serta memperkuat stabilitas dan pertumbuhan ekonomi di Sulawesi Selatan. Tujuan. Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model prediksi inflasi dan PDRB yang lebih efektif menggunakan machine learning, serta membandingkannya akurasi model ARIMA, dengan tujuan akhir untuk memperdalam pemahaman mengenai dinamika ekonomi di Sulawesi Selatan melalui analisis grafis dari data yang dihasilkan. Metode. Penelitian ini menggunakan metode Autoregressive Integrated Moving Average (ARIMA) dan implementasi ML untuk pemodelan otomatis. Hasil. penerapan model ARIMA menunjukkan efektivitas dalam memprediksi inflasi dan PDRB di Sulawesi Selatan. Dalam prediksi inflasi, model ARIMA manual (1,1,1) menunjukkan performa yang lebih baik dibandingkan model Auto ARIMA, dengan nilai RMSE sebesar 0.5795 dibandingkan 0.5813 pada Auto ARIMA, serta MAE 0.3980 berbanding 0.4032. Sebaliknya, untuk prediksi PDRB, model Auto ARIMA (1,1,0)(0,0,1) menghasilkan prediksi yang lebih akurat dengan RMSE 0.0199 dibandingkan model manual ARIMA yang memiliki RMSE 0.0217. Analisis tren menunjukkan pola musiman yang konsisten, dengan inflasi mengalami peningkatan signifikan hingga 2023 dan stabilisasi pada 2024.

Kata kunci: Machine Learning, ARIMA, inflasi, Produk Domestik Regional Bruto (PDRB).

ABSTRACT

JACQLIEN PATRICIA MAWERU RUNTU. *Integration of Machine Learning and ARIMA Model for Inflation and GRDP Prediction in South Sulawesi* (Supervised by Saiful).

Background. The importance of the development of information technology and machine learning (ML) in economic analysis, especially in improving the efficiency and accuracy of predicting macroeconomic variables such as inflation and Gross Regional Domestic Product (GRDP). ML is able to process large and complex data without requiring strict assumptions, so it can identify patterns and relationships that are not visible with traditional methods. The implementation of ML at Bank Indonesia is expected to produce more consistent and accurate predictions, support more informed decision making in economic policy, and strengthen economic stability and growth in South Sulawesi. Objectives. This study aims to develop and evaluate a more effective inflation and GRDP prediction model using machine learning, and compare it to the accuracy of the ARIMA model, with the ultimate goal of deepening understanding of economic dynamics in South Sulawesi through graphical analysis of the resulting data. Methods. This research uses the Autoregressive Integrated Moving Average (ARIMA) method and ML implementation for automatic modeling. Results, the application of the ARIMA model shows effectiveness in predicting inflation and GRDP in South Sulawesi. In predicting inflation, the manual ARIMA (1.1.1) model shows better performance than the Auto ARIMA model, with an RMSE value of 0.5795 compared to 0.5813 in Auto ARIMA, and MAE 0.3980 versus 0.4032. In contrast, for the prediction of GRDP, the Auto ARIMA (1,1,0)(0,0,1) model produces more accurate predictions with an RMSE of 0.0199 compared to the manual ARIMA model which has an RMSE of 0.0217. Trend analysis shows a consistent seasonal pattern, with inflation experiencing a significant increase until 2023 and stabilizing in 2024.

Keywords: Machine Learning, ARIMA, inflation, Gross Regional Domestic Product (GRDP).

DAFTAR ISI

DEDAN	VATA ANI DENIGA III ANI	Halaman
	ATAAN PENGAJUAN	
	MAN PENGESAHAN	
	ATAAN KEASLIAN SKRIPSI DAN PELIMPAHAN HAK CIPTA	
	AN TERIMA KASIH	
	RAK	
	RACT	
DAFTA	R ISI	ix
	R TABEL	
	R GAMBAR	
	R LAMPIRAN	
BAB I	PENDAHULUAN	1
1.1	Latar Belakang	
1.2	Teori	4
1.3	Rumusan Masalah	16
1.4	Tujuan Penelitian	16
1.5	Batasan Masalah	16
1.6	Manfaat Penelitian	16
BAB II	METODOLOGI PENELITIAN	18
2.1	Objek Penelitian	18
2.2	Teknik Pengumpulan dan Sumber Data	18
2.3	Teknik Pengolahan Data dan Analisis	19
2.4	Flowchart Penelitian	20
2.5	Kerangka Berpikir	23
BAB III	HASIL DAN PEMBAHASAN	24
3.1	Hasil Penelitian	24
3.2	Pembahasan	60
BAB IV	KESIMPULAN	68
4.1	Kesimpulan	68
4.2	Saran	69
4.3	Keterbatasan Penelitian	69
DAFTA	R PUSTAKA	70
LAMPII	RAN	72

DAFTAR TABEL

Tabel 1 Jenis Transformasi	9
Tabel 2. Bentuk Fungsi ACF dan PACF	10
Tabel 3. Penelitian Terdahulu	
Tabel 4. Time Series Inflasi	24
Tabel 5. Time Series PDRB	24

DAFTAR GAMBAR

Gambar 1. Jenis pola peramalan5	
Gambar 2. Flowchart Penelitian	
Gambar 3. Kerangka pikir penelitian23	
Gambar 4. Datetime Inflasi	
Gambar 5. Datetime PDRB27	
Gambar 6. Grafik Time Series Inflasi Sulawesi Selatan28	
Gambar 7. Dekomposisi Musiman dari Time Series Inflasi	
Gambar 8. Grafik Time Series PDRB Sulawesi Selatan	
Gambar 9. Dekomposisi Musiman dari Time Series PDRB30	
Gambar 10. Uji ADF Time Series Inflasi31	
Gambar 11. Uji ADF Time Series PDRB	
Gambar 12. Output Jumlah Differencing Setiap Tes Time Series Inflasi33	
Gambar 13. Plot ACF dan PACF Time Series Inflasi34	
Gambar 14. Hasil Differencing Time Series Inflasi	ŀ
Gambar 15. Plot ACF Hasil Differencing Time Series Inflasi	,
Gambar 16. Plot PACF Hasil Differencing Time Series Inflasi36	j
Gambar 17. Output Jumlah Differencing Setiap Tes Time Series PDRB36	j
Gambar 18. Plot PACF dan ACF Time Series PDRB	
Gambar 19. Hasil Differencing Time Series PDRB37	•
Gambar 20. Plot ACF dan PACF Hasil Differencing Time Series PDRB38	,
Gambar 21. Parameter Auto ARIMA Inflasi	
Gambar 22. Parameter Auto ARIMA PDRB40	
Gambar 23. Output Fitting Model Auto ARIMA Inflasi41	
Gambar 24. Output Fitting Model Manual ARIMA Inflasi42	
Gambar 25. Output Fitting Model Auto ARIMA PDRB43	
Gambar 26. Output Fitting Model Manual ARIMA PDRB45	
Gambar 27. Output Forecasting Model Auto ARIMA Inflasi	
Gambar 28. Output Forecasting Model Manual ARIMA Inflasi	
Gambar 29. Output Forecasting Model Auto ARIMA PDRB47	•
Gambar 30. Output Forecasting Model Manual ARIMA PDRB48	
Gambar 31. Prediksi Model Auto ARIMA Inflasi49	
Gambar 32. Prediksi Model Manual ARIMA Inflasi50	
Gambar 33. Prediksi Model Auto ARIMA PDRB51	
Gambar 34. Prediksi Model Manual ARIMA PDRB52	
Gambar 35. Diagnostic plot dari Model Auto Arima Inflasi	,
Gambar 36. Diagnostic plot dari Model Manual Arima Inflasi54	
Gambar 37. Hasil Evaluasi Model Inflasi55	
Gambar 38. Diagnostic plot dari Model Auto Arima PDRB56	
Gambar 39. Diagnostic plot dari Model Manual Arima PDRB	
Gambar 40. Hasil Evaluasi Model PDRB58	
Gambar 41. Hasil Model Terbaik Inflasi	
Gambar 42. Hasil Model Terbaik PDRB59	
Gambar 43. Metrik Evaluasi Inflasi60	
Gambar 44. Metrik Evaluasi PDRB61	
Gambar 45. SARIMAX Results Model Auto ARIMA Inflasi	
Gambar 46. SARIMAX Results Model Manual ARIMA Inflasi	
Gambar 47. SARIMAX Results Model Auto ARIMA PDRB	
Gambar 48. SARIMAX Results Model Manual ARIMA PDRB63	
Gambar 49. Plot Data Time Series Inflasi Model Auto ARIMA	
Gambar 50. Plot Data Time Series Inflasi Model Manual ARIMA66	j

Gambar 51. Plot Data Time Series	s PDRB Model Auto ARIMA6	36
Gambar 52. Plot Data Time Series	s PDRB Model Manual ARIMA6	36

DAFTAR LAMPIRAN

Lampiran 1. Coding Program Inflasi	72
Lampiran 2. Coding Program PDRB	80
Lampiran 3. Diskusi bersama ahli ekonom Bank Indonesia KPw Sulawesi Sela	atan87

BAB I PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komputasi telah mengalami akselerasi signifikan dalam beberapa dekade terakhir, perubahan mendalam di berbagai sektor termasuk ekonomi. Teknologi informasi yang semakin maju memungkinkan pengumpulan, penyimpanan, dan analisis data dalam jumlah besar. Salah satu cabang teknologi yang berkembang pesat dalam konteks ini adalah Machine Learning (ML) adalah bagian dari Artificial Intelligence (AI) yang fokus pada pengembangan algoritma yang memungkinkan komputer untuk belajar dari data dan membuat prediksi atau keputusan tanpa harus diprogram secara eksplisit. Fokus utamanya adalah membangun sebuah aplikasi komputer yang dapat mempelajari data, lalu membuat sebuah model yang siap digunakan untuk memecahkan kasus tertentu. ML mengikuti metode analisis data yang bertanggung jawab untuk mengotomatisasi pembuatan model secara analitis. Menggunakan algoritma yang secara iteratif untuk mendapatkan pengetahuan dari data dan dalam proses ini memungkinkan komputer menemukan pengetahuan yang tampaknya tersembunyi tanpa bantuan dari eksternal. Untuk mengeyaluasi kineria aplikasi ML maka diperlukan sebuah metode pengukuran kinerja yang kualitatif. Aplikasi atau model ML mendapatkan experience berdasarkan dataset yang disediakan pada proses pelatihan. Sebuah dataset adalah kumpulan contoh-contoh yang harus dipelajari oleh komputer (Id, 2021). Dalam konteks analisis ekonomi, ML menawarkan keunggulan signifikan dibandingkan metode tradisional karena kemampuannya mengolah dalam jumlah besar dan kompleks, serta mengidentifikasi pola yang tidak mudah terlihat oleh analisis konvensional.

Relevansi ML dalam analisis ekonomi sangat tinggi, telah menunjukkan potensinya yang besar dalam meningkatkan efisiensi dan akurasi analisis data. Dengan kemampuan untuk mengolah dan menganalsis data yang kompleks dan beragam, ML dalam ekonomi memungkinkan analisis yang lebih mendalam dan keputusan yang lebih tepat berbasis data, yang sangat bermanfaat dalam perencanaan dan pengambilan kebijakan ekonomi. Salah satu keunggulan utama ML dibandingkan metode statistik tradisional dalam analisis data ekonomi adalah kemampuannya untuk menangani volume data yang sangat besar dan kompleks tanpa memerlukan asumsi linearitas atau distribusi tertentu. Algoritma ML seperti pohon keputusan, jaringan saraf tiruan, dan model ensemble dapat mengungkapkan pola dan hubungan dalam data yang mungkin tidak terlihat dengan metode statistik konvensional (Mahesh, 2020). Kemaiuan teknologi komputasi dan ketersediaan data yang semakin melimpah membuat penerapan ML menjadi lebih efektif. Selain itu, model ML memiliki kemampuan adaptif yang memungkinkan model untuk terus diperbarui dan ditingkankan seiring dengan bertambahnya data baru, sehingga prediksi yang dihasilkan lebih akurat dan relevan dengan kondisi terkini. Menurut Mullainathan dan Spiess (2017), ML menawarkan pendekatan yang lebih flesibel dan robust dalam menghadapi data yang heterogen dan dinamis, membuatnya sangat cocok untuk aplikasi ekonomi yang kompleks dan beragam. Fleksibilitas ML terletak pada kemampuannya memproses dan menganalisis berbagai jenis data tanpa

memerlukan asumsi awal yang ketat, seperti yang sering diperlukan dalam metode statistik tradisional. Data ekonomi sering kali bersifat heterogen, berasal dari berbagai sumber dengan format dan struktur yang berbeda, serta mencakup variabel-variabel yang beragam. Algoritma ML dapat mengintegrasikan dan mengolah data tersebut secara efisien, mengenali pola yang kompleks dan non-linear yang mungkin terlewatkan oleh pendekatan tradisional. Selain itu, ML sangat *robust*, artinya algoritma ini dapat beradaptasi dengan cepat terhadap perubahan dalam data, memungkinkan model untuk terus diperbarui dan disesuaikan dengan data terbaru. Hal ini sangat penting dalam ekonomi yang dinamis, dimana kondisi pasar dan faktor-faktor ekonomi dapat berubah dengan cepat. Fleksibilitas dan ketahan ini menjadikan ML sangat efektif dalam aplikasi ekonomi yang membutuhkan analisis canggih dan prediksi yang akurat.

Implementasi ML dalam bidang ekonomi, seperti memprediksi variabel-variabel makroekonomi dengan setepat mungkin bagi pihak berwenang dalam pengambilan keputusan agar kebijakan ekonomi menjadi lebih efektif dan efisien. Hal ini akan memberikan kontribusi positif terhadap kepercayaan perekonomian dengan mengurangi ketidakpastian pada indikator-indikator utama perekonomian. Salah satu indikator penting dari upaya pembangunan dan perkembangan ekonomi adalah pertumbuhan ekonomi, yang dapat meningkatkan kesejahteraan masyarakat suatu negara sesuai dengan tingkat pendapatan per kapita. Oleh karena itu, pertumbuhan ekonomi dapat digunakan sebagai tolak ukur keberhasilan suatu negara dalam hal pembangunan ekonomi. Pada akhirnya, pertumbuhan ekonomi juga dapat digunakan untuk menjelaskan hal-hal yang terkait dengan ekonomi (Hodijah & Angelina, 2021). Pertumbuhan ekonomi merujuk pada peningkatan dalam barang dan jasa yang diproduksi oleh suatu kawasan selama periode waktu tertentu. Ini diukur melalui pertumbuhan Produk Domestik Regional Bruto (PDRB), yang merupakan nilai semua barang dan jasa yang dihasilkan dalam suatu ekonomi dalam periode waktu tertentu (Ernita, 2021). Dalam konteks PDRB, inflasi juga berpengaruh. PDRB atas dasar harga berlaku dapat dipengaruhi oleh perubahan harga akibat inflasi, sehingga untuk mendapatkan gambaran yang lebih akurat mengenai pertumbuhan ekonomi riil, digunakan PDRB atas dasar konstan yang menghilangkan pengaruh inflasi. Dengan demikian, memahami dan mengelola inflasi menjadi krusial dalam upaya mempertahankan pertumbuhan ekonomi yang berkelanjutan dan stabil.

Inflasi dan Produk Domestik Regional Bruto (PDRB) adalah dua indikator ekonomi yang saling terkait dan sangat penting dalam analisis ekonomi makro. Inflasi mencerminkan tingkat perubahan harga barang dan jasa dalam suatu ekonomi, yang dapat mempengaruhi daya beli masyarakat dan stabilitas ekonomi secara keseluruhan (Anzelia, dkk., 2021). Sementara itu, PDRB merupakan ukuran total nilai barang dan jasa yang dihasilkan dalam suatu wilayah dalam periode tertentu, yang mencerminkan kesehatan ekonomi dan pertumbuhan. Hubungan antara inflasi dan PDRB sering kali kompleks; inflasi yang tinggi dapat mengurangi konsumsi dan investasi, yang pada gilirannya dapat memperlambat pertumbuhan PDRB. Sebaliknya, pertumbuhan PDRB yang kuat dapat menyebabkan peningkatan permintaan agregat, yang dapat mendorong inflasi. Oleh karena itu, pemahaman yang mendalam tentang dinamika antara inflasi dan PDRB sangat penting untuk merumuskan kebijakan ekonomi yang efektif dan responsif.

Hubungan antara inflasi dan PDRB merupakan aspek krusial dalam pemahaman dinamika ekonomi suatu negara atau wilayah. Ketika inflasi meningkat, daya beli masyarakat cenderung menurun, yang dapat mengakibatkan penurunan permintaan barang dan jasa, serta menghambat pertumbuhan investasi. Hal ini sering kali mengarah pada penurunan PDRB, di mana sektor-sektor ekonomi mengalami perlambatan akibat berkurangnya konsumsi. Sebaliknya, peningkatan PDRB yang menunjukkan pertumbuhan ekonomi yang solid biasanya disertai dengan peningkatan permintaan agregat. Namun, jika permintaan ini tidak diimbangi dengan pasokan yang memadai, hal tersebut dapat memicu inflasi yang lebih tinggi (Silitonga, 2021). Oleh karena itu, para pembuat kebijakan perlu memahami interaksi ini secara menyeluruh, agar dapat merancang strategi yang tidak hanya mendukung pertumbuhan ekonomi yang berkelanjutan, tetapi juga menjaga stabilitas harga dan kesejahteraan masyarakat.

Saat ini, prediksi indikator makroekonomi masih dilakukan secara manual melalui pembaruan worksheet dan penyusunan analisis, dengan hasil prediksi yang terkadang kurang akurat dan memiliki deviasi akibat kurangnya penerapan metode ilmiah. Pendekatan integrasi machine learning dan ARIMA (Autoregressive Integrated Moving Average) memungkinkan prediksi yang lebih efisien dan akurat dengan memanfaatkan data historis dan pola variabel ekonomi secara statistik, sehingga mampu menghemat waktu kerja sekaligus meningkatkan ketepatan prediksi. Metode ini membantu mengidentifikasi pola dan tren musiman pada inflasi dan PDRB, memungkinkan intervensi pengendalian yang lebih tepat sasaran dan sesuai waktu, serta mendukung pengambilan keputusan pre-emptive. Keunggulan machine learning dalam pendekatan ini dibandingkan metode manual terletak pada kemampuannya mengolah data dalam volume besar secara otomatis, mengenali pola-pola kompleks, dan beradaptasi dengan perubahan data yang dinamis. *Machine learning* memungkinkan analisis variabel-variabel makroekonomi yang beragam, yang sering kali sulit dipetakan secara manual, dengan memanfaatkan algoritma yang mendeteksi korelasi dan anomali di luar pola linier. Hasil prediksi yang lebih cepat dan akurat mendukung pengambilan keputusan yang proaktif, memungkinkan pengendalian inflasi dan penguatan PDRB dengan strategi yang lebih efektif (Chauhan & Singh, 2018). Selain itu, otomatisasi melalui machine learning mengurangi kesalahan manusia dalam proses analisis manual, sehingga meminimalkan deviasi prediksi dan meningkatkan keandalan hasil, yang pada akhirnya membantu satuan kerja mencapai kinerja yang lebih optimal dalam pengendalian ekonomi.

Dalam konteks *machine learning*, ARIMA sering dianggap sebagai pendekatan *unsupervised learning* karena model ini tidak memerlukan label target untuk setiap data yang digunakan dalam pelatihan, melainkan hanya membutuhkan data historis dari variabel yang ingin diprediksi. Hal ini berbeda dengan *supervised learning*, yang biasanya membutuhkan data yang berlabel untuk membangun hubungan antara variabel *input* dan *output*. ARIMA memanfaatkan sifat data *time series* untuk mengenali pola internal, seperti tren dan musiman, yang ada dalam data tersebut tanpa harus mengetahui nilai target dari variabel independen lainnya (Hyndman & Athanasopoulos, 2018). Dengan demikian, ARIMA sangat berguna untuk kasus prediksi ekonomi, di mana pola waktu sangat mendominasi, namun tidak selalu tersedia data lengkap untuk setiap

faktor eksternal yang mungkin memengaruhi prediksi. Penggunaan ARIMA dalam model integrasi yang mengombinasikan *machine learning* menawarkan fleksibilitas tambahan dalam menghasilkan prediksi yang lebih akurat dan relevan. Misalnya, *machine learning* dapat digunakan untuk mendeteksi pola non-linear atau interaksi variabel yang kompleks, sedangkan ARIMA akan berfokus pada pola linear dan musiman yang kuat dalam data *time series* (Ospina, dkk., 2023). Kombinasi ini sangat cocok untuk prediksi ekonomi, karena ekonomi dipengaruhi oleh banyak faktor internal maupun eksternal yang berubah dari waktu ke waktu. Model integrasi ini memungkinkan peningkatan performa prediksi, terutama dalam situasi yang menuntut respons cepat terhadap dinamika pasar dan ekonomi global yang fluktuatif. Penerapan ini juga membantu mengurangi kebutuhan intervensi manual dan potensi bias subjektif dalam proses prediksi.

Integrasi ARIMA dan *machine learning* memiliki keunggulan signifikan dibandingkan metode deret waktu lainnya seperti metode eksponensial smoothing atau model VAR (Vector Autoregressive). Eksponensial smoothing, misalnya, sering kali efektif untuk data dengan tren yang relatif stabil, tetapi kurang mampu menangkap fluktuasi musiman dan pola autokorelasi jangka panjang yang umum dalam data ekonomi. Sebaliknya, ARIMA unggul dalam mengidentifikasi pola-pola tersebut, sedangkan machine learning membantu melengkapi kelemahan ARIMA dalam menangkap pola non-linier yang sering muncul pada data ekonomi yang lebih dinamis. Kombinasi ini menciptakan model yang lebih responsif terhadap perubahan mendadak dalam data, sehingga dapat memberikan prediksi yang lebih akurat untuk variabel seperti inflasi dan PDRB. Model VAR, yang banyak digunakan dalam analisis ekonomi multivariat, mampu menangani variabel yang saling mempengaruhi secara simultan, tetapi metode ini memiliki keterbatasan ketika menghadapi data ekonomi yang nonstasioner atau ketika harus memprediksi jangka panjang dalam situasi yang kompleks (Hyndman & Athanasopoulos, 2018). Dengan bantuan machine learning, integrasi ini mampu menyaring noise dalam data, memberikan estimasi yang lebih stabil dan presisi yang lebih tinggi dibandingkan model yang mengandalkan VAR saja. Model integrasi ini membantu pengambil kebijakan dan analis ekonomi untuk memperoleh proyeksi yang lebih tepat, sehingga memungkinkan perencanaan ekonomi yang lebih baik dan pengambilan keputusan yang lebih tepat untuk stabilitas ekonomi jangka panjang di Sulawesi Selatan.

1.2 Teori

Pada penelitian ini beberapa teori yang akan membahas konsep dasar forecasting, machine learning, analisis deret waktu (time series analysis), identifikasi model, uji signifikansi parameter, pengujian diagnostik, kriteria pemilihan model terbaik, Autoregressive Integrated Moving Average (ARIMA).

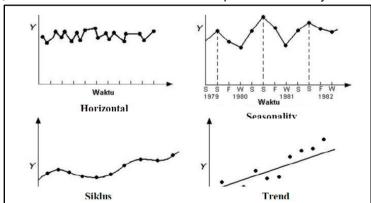
1.2.1 Forecasting

Forecasting (peramalan) merupakan suatu perkiraan tentang keadaan di masa yang akan datang, forecasting (peramalan) dapat dilakukan dengan berbagai metode yang disesuaikan dengan pola data yang ada. Hal ini dapat dilakukan dengan melibatkan pengambilan data historis dan memprediksikan ke masa mendatang dengan suatu

bentuk model matematis (kuantitatif), atau bisa juga merupakan prodiksi intuisi yang bersifat subjektif (kualitatif). Peramalan kuantitatif menggunakan metode statistik atau model matematis yang beragam dengan data masa lalu dan variabel seba akibat untuk meramalkan permintaan. Peramalan subjektif atau kualitatif menggabungkan faktor, seperti intuisi, emosi, pengalaman pribadi dan sistem nilai pengambil keputusan untuk meramal. *Forecasting* (peramalan) memiliki manfaat yang signifikan, termasuk membantu pemerintah dan perusahaan dalam perencanaan dan pengambilan keputusan yang lebih baik (Lusiana & Yuliarty, 2020).

Pada umumnya, ada 3 jenis *forecasting* (peramalan), yaitu: peramalan ekonomi (berhubung dengan indikator-indikator ekonomi secara makro, misalnya); peramalan teknologi (berkaitan dengan laju perkembangan teknologi); dan peramalan permintaan (berkaitan dengan prediksi permintaan untuk produk atau layanan suatu perusahaan). Selain itu, terdapat beberapa komponen pola data *forecasting* (peramalan) antara lain (Seto,2016):

- a. *Trend* (T), terjadi bila ada kenaikan atau penurunan dari data secara gradual dari gerakan datanya dalam kurun waktu panjang.
- b. Seasonality (S), pola musiman terjadi bila pola datanya berulang sesudah suatu periode tertentu: hari, mingguan, bulanan, triwulan dan tahun.
- c. Cycles (C), siklus adalah suatu pola data yang terjadi setiap beberapa tahun, biasanya dipengaruhi oleh fluktuasi ekonomi jangka panajng berkaitan dengan siklus bisnis.
- d. *Horizontal* (H)/Stasioner, terjadi bila nilai data berfluktuasi disekitar nilai rata-rata yang tetap, stabil atau disebut stasioner terhadap nilai rata-ratanya.



Gambar 1. Jenis pola peramalan Sumber: Seto (2016)

Dalam praktiknya, terdapat berbagai metode *forecasting* yang dapat digunakan, tergantung pada karakteristik data dan tujuan analisis. Berikut adalah penjelasan lengkap mengenai jenis-jenis *forecasting*.

1. Simple Average

Simple Average merupakan salah satu metode forecasting yang paling dasar. Metode ini menghitung rata-rata dari data historis dan menggunakan nilai ini sebagai prediksi untuk periode yang akan datang. Cocok untuk data tanpa tren atau musiman.

- Kelebihan: Mudah diterapkan dan dipahami, tidak memerlukan data yang kompleks.
- Kekurangan: Metode ini tidak mempertimbangkan tren atau musiman, sehingga kurang akurat jika data memiliki pola (Makridakis, 2020).

2. Moving Average

Moving Average menghitung rata-rata dari sejumlah nilai dalam periode tertentu. Jumlah periode yang digunakan dapat disesuaikan untuk menghaluskan fluktuasi data. Berguna untuk mengidentifikasi tren jangka pendek.

- Kelebihan: Dapat menghilangkan fluktuasi jangka pendek, sehingga menghasilkan prediksi yang lebih stabil.
- Kekurangan: Memerlukan data historis yang cukup untuk menghasilkan hasil yang baik dan tidak menangkap perubahan mendadak (Hyndman & Athanasopoulos, 2018).

3. Exponential Smoothing

Exponential Smoothing memberikan bobot lebih pada data terbaru. Metode ini menggunakan parameter smoothing untuk menentukan sejauh mana peramalan dipengaruhi oleh data historis.

- Kelebihan: Sangat efektif untuk data yang stabil dan memungkinkan penyesuaian cepat terhadap perubahan tren.
- Kekurangan: Memerlukan pemilihan parameter yang tepat untuk mencapai hasil optimal (Chatfield, 2019).

4. Model Regresi

Model regresi adalah teknik statistik yang digunakan untuk memodelkan hubungan antara variabel independen dan variabel dependen. Dalam konteks forecasting, model regresi dapat digunakan untuk memperkirakan nilai masa depan berdasarkan hubungan historis.

- Kelebihan: Dapat mempertimbangkan beberapa variabel pada saat yang sama dan hubungan antara variabel.
- Kekurangan: Hasil prediksi sangat tergantung pada kekuatan hubungan antara variabel yang dipertimbangkan (Wooldridge, 2016).

5. Model ARIMA (Autoregressive Integrated Moving Average)

ARIMA terdiri dari tiga komponen: *Autoregressive* (AR), *Integrated* (I), dan *Moving Average* (MA). Secara umum, model ARIMA dirumuskan dengan notasi p (orde/derajat *autoregressive*), d (orde/ derajat *differencing*), q (orde/derajat *moving average*).

- Kelebihan: Dapat menangkap berbagai pola dalam data time series, termasuk tren dan musiman setelah melalui proses differencing.
- Kekurangan: Proses identifikasi dan parameterisasi model bisa kompleks dan memerlukan pemahaman statistik yang mendalam (Putra & Kurniawati, 2021).

1.2.2 Machine Learning (ML)

Machine Learning (ML) merupakan bidang studi yang fokus kepada desain dan analisis algoritma sehingga memungkinkan komputer untuk dapat belajar. Menurut

Samuel (dalam Id, 2021), berisi sebuah algoritma yang bersifat *generic* (umum) dimana algoritma tersebut dapat menghasilkan sesuatu yang menarik atau bermanfaat dari sejumlah data tanpa harus menulis kode yang spesifik. Pada intinya, algoritma yang generik tersebut ketika diberikan sejumlah data maka ia dapat membangun sebuah aturan atau model atau inferensi dari data tersebut. Sebagai contoh sebuah algoritma untuk mengenali tulisan tangan dapat digunakan untuk mendeteksi email yang berisi spam dan bukan spam tanpa mengganti kode. Algoritma yang sama ketika diberikan data pelatihan yang berbeda menghasilkan logika klasifikasi yang berbeda.

ML dapat dikelompokkan berdasarkan cara mesin belajar dan memproses data sehingga dapat melakukan tugasnya. Pembagian ML berdasarkan cara belajarnya dibagi menjadi tiga kelompok yaitu:

a. Supervised Learning

Supervised learning adalah model dilatih menggunakan data terlabel, artinya setiap input dilengkapi dengan output yang benar. Jika dianalogikan pada proses pembelajaran, komputer atau mesin akan mempelajari data training yang berisi label. Tujuan dari model ini adalah untuk memprediksi output berdasarkan input baru. Supervised learning sering digunakan untuk tugas-tugas seperti klasifikasi dan regresi.

b. Unsupervised Learning

Unsupervised learning adalah data yang tidal berlabel. Model mencoba menemukan struktur atau pola yang mendasari data. Secara matematis, Unsupervised learning terjadi ketika memiliki sejumlah data input dan tanpa variabel output yang berhubungan. Unsupervised learning biasanya digunakan untuk clustering dan reduksi dimensi.

c. Reinforcement Learning

Reinforcement learning adalah model yang belajar dengan berinteraksi dengan lingkungan dan menerima umpan balik dalam bentuk reward atau punishment. Melalui sebuah algoritma, mesin akan mempelajari bagaimana membuat keputusan yang spesifik berdasarkan lingkungan yang berubah-ubah. Reinforcement learning banyak digunakan dalam game dan robotika.

1.2.3 Analisis Deret Waktu (Time Series Analysis)

Menurut Ekanada (2014), analisis deret waktu (time teries analysis) adalah pendekatan time series yang digunakan untuk meramalkan masa depan berdasarkan data historis dan kesalahan masa lalu. Time series adalah suatu runtutan data yang berdasarkan pada waktu. Time series ini kemudian disusun berdasarkan urutan waktu atau data yang dikumpulkan dari waktu ke waktu sebagai data berkala. Data waktu yang digunakan dapa berupa data mingguan, bulanan tahunan, atau triwulan. Data ini berhubungan dengan data statistik yang dicatat dan diamati dalam batas-batas waktu tertentu. Time series sangat penting dalam statistika karena berguna untuk melakukan peramalan pada hal-hal yang ingin diramalkan dalam pengambilan keputusan. Salah satu cara untuk meramalkan adalah dengan menggunakan analisis pola hubungan antara variabel yang diperkirakan dan variabel waktu. Beberapa contoh metode time series adalah metode smoothing, metode ARIMA, metode prediksi trend dengan regresi. Metode time series menjadi pilihan yang lebih efisien untuk menghasilkan forecast yang

akurat. Ada beberapa hal yang perlu diperhatikan dalam mengolah data time series, yaitu (Wardhno, dkk., 2019):

- 1. Didasarkan pada asumsi stasioneritas
- 2. Bila asumsi stasioneritas tidak terpenuhi, menyebabkan timbulnya masalah *Auto*korelasi.
- 3. Regresi dengan nilai R^2 tinggi lebih dari 0,9 menunjukkan hubungan yang tidak signifikan atau *spurious regression*.
- 4. Adanya fenomena *random walk*. Misal: harga saham besok sama dengan harga saham saat ini ditambah error yang random.
- 5. Regresi dengan data time series seringkali digunakan untuk *forecasting*.
- 6. Pengujian untuk stasioneritas dilakukan sebelum uji kausalitas.

Stasioneritas data adalah masalah utama dalam analisis *time series*. Ini berkaitan dengan konsistensi pergerakan data rangkaian waktu. Data yang stasioner terjadi ketika nilai varian dan rata-ratanya tetap sepanjang waktu dan diikuti oleh nilai varian antar dua periode yang hanya bergantung pada jarak. Mereka akan bergerak secara konstan dan konvergen di sekitar nilai rata-rata dengan deviasi kecil tanpa pergerakan tren yang positif atau negatif. Regresi palsu, juga dikenal sebagai *spurious regression*, dihasilkan dari data yang tidak stasioner. Uji kointegrasi adalah metode awal untuk menghindari *spurious regression*. Ini melibatkan hubungan jangka panjang antara variabel yang tidak stasioner. Ini menghasilkan kombinasi linier, yang menghasilkan kondisi yang stasioner atau kondisi keseimbangan dalam jangka panjang. *Error Correction Model* (ECM) adalah model dinamis yang digunakan untuk mengkoreksi persamaan regresi variabel yang tidak stasioner sehingga dapat kembali ke kondisi keseimbangan dalam jangka panjang, asalkan ada hubungan kointegrasi antarvariabel (Wardhno, dkk., 2019).

1.2.4 Identifikasi Model

Identifikasi model digunakan untuk mengetahui apakah data tersebut stasioner atau tidak. Kestasioneran data dapat dibuktikan dengan menganalisis plot Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF). Model linier time series yang stasioner adalah model-model yang dapat digunakan untuk data-data yang stasioner. Data stasioner yaitu data yang mempunyai mean nilainya tidak berubah setiap waktu, sedangkan data tidak stasioner terdapat pola data trend atau pola musiman. Setelah data sudah stasioner, langkah berikutnya adalah menetapkan model ARIMA (p,d,q). Jika tidak terjadi proses differencing maka d diberi nilai 0, jika menjadi stasioner setelah differencing pertama maka nilai d=1, jika data stasioner setelah differencing kedua, maka d=2 dan seterusnya. Berikut terdapat beberapa langkah dalam proses pembentukan model ARIMA yaitu sebagai berikut:

a. Stasioneritas

Stasioneritas data dibagi menjadi dua, yakni stasioner dalam *mean* dan stasioner dalam *varians*. Data dikatakan stasioner jika tidak mengalami perubahan yang signifikan. Jika data yang digunakan tidak memenuhi stasioner dalam *mean* maka dilakukan *differencing* atau pembedaan antara data pengamatan pada waktu dengan data pengamatan pada waktu sebelumnya. Pada umumnya data time series tidak stasioner, maka perlu adanya pengamatan plot data *time series* dengan cara mengamati dari nilai-nilai *Auto*korelasi pada plot ACF. Jika plot cenderung

memperlihatkan *trend* searah diagonal atau nilai-nilai *Auto*korelasinya signifikan berbeda dari nol atau menjauh dari nol untuk beberapa periode waktu maka data tersebut belum stasioner. Jika plot cenderung konstan tidak terjadi penurunan atau pertumbuhan maka data dapat dikatakan stasioner atau nilai-nilai korelasinya dari data stasioner akan turun sampai nol sesudah lag kedua atau ketiga (Waititu & Kiboro, 2015). Apabila data yang digunakan belum memenuhi kestasioneran dalam varians maka dilakukan transformasi.

$$T(Z_t) = \frac{Z_t^{\lambda} - 1}{\lambda}$$
, untuk $\lambda \neq 0$(1)

Nilai λ disebut juga dengan parameter transformasi. Berikut ini tabel transformasi yang harus dilakukan apabila data yang dianalisis memiliki nilai λ tertentu.

Tabel 1 Jenis Transformasi		
Nilai λ	Jenis transformasi	
1	Z_t (tidak ada transformasi)	
0,5	$\sqrt{Z_t}$	
0	lnZ_t	
-0,5	$\frac{1}{\sqrt{Z_t}}$	
-1	$\frac{1}{Z_t}$	

Sumber: (Saputri,2019)

b. Autocorrelation Function (ACF)

Identifikasi model selanjutnya dalam metode *time series* adalah dengan menggunakan *Autocorrelation Function* (ACF). ACF merupakan suatu korelasi fungsi yang mendefinisikan ikatan dari dua variabel dalam *time series*. Secara sederhana, fungsi *Auto*korelasi merupakan suatu hubungan data dengan data lain pada suatu penelitian pada suatu deret berkala (Tantika, 2018). Data dikatakan tidak stasioner dalam *mean* apabila diagram ACF cenderung turun lambat atau turun secara linier. Secara umum fungsi ACF dirumuskan sebagai berikut (Wei, 2006):

$$P_{k} = \frac{Cov(Z_{t}Z_{t+k})}{\sqrt{Var(Z_{t})}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_{k}}{\gamma_{O}}.$$
 (2)

Secara grafis ditunjukkan nilai p_k dengan lag k. Nilai ACF berkisar antara $-1 \le p_k \le 1$. Selanjutnya $p_k = p_{-k}$ yang menggambarkan nilai hubungan korelasi antara nilai tersebut, tanda positif menunjukkan bahwa hubungannya adalah berkorelasi positif dan tanda negatif menunjukkan bahwa hubungannya adalah berkorelasi negatif. Hubungan yang kuat didapatkan apabila nilai korelasinya 1 dan semakin melemah ke nilai 0.

c. Partial Auto Correlative Funtion (PACF)

Partial Auto Correlative Funtion (PACF) adalah korelasi antar deret pengamatan pada lag-lag yang mengukur keeratan antar pengamatan suatu deret waktu. PACF juga dapat diartikan sebagai himpunan Autokorelasi parsial untuk berbagai lag k. Fungsi dari PACF sendiri digunakan sebagai pengukur setiap lag 1,2,3. Menurut dari

data time series iika sampel ACF turun sangat lambat dan data sampel PACF terputus setelah lag maka hal tersebut menunjukkan bahwa data tidak stasioner dan perlu dilakukan differencing (Fahmi, 2017). Secara umum, fungsi PACF dirumuskan seperti pada persamaan berikut (Wei, 2006):

$$\widehat{\phi}_{k+1,k+1} = \frac{\widehat{\rho}_{k+1} - \sum\limits_{j=1}^{k} \widehat{\phi}_{k,j} \ \widehat{\rho}_{k+1-j}}{1 - \sum\limits_{j=1}^{k} \widehat{\phi}_{k,j} \ \widehat{\rho}_{j}}$$
(3) Berikut adalah karakteristik untuk menentukan orde berdasarkan plot ACF dan PACF

Berikut adalah karakteristik untuk menentukan orde berdasarkan plot ACF dan PACF (Wei, 2006).

Tabel 2. Bentuk Fungsi ACF dan PACF	Ξ
-------------------------------------	---

Model	ACF	PACF
AR (p)	Naik/ turun secara eksponensial	Terpotong pada lag q
MA(q)	Terpotong pada lag q	Naik/ turun secara
		eksponensial
ARMA(p,q)	Naik/ turun secara	Naik/ turun secara
	eksponensial	eksponensial

1.2.5 Uji Signifikansi Parameter

Setelah melakukan proses identifikasi dan memperoleh model sementara, selanjutnya yaitu menaksir parameter model yang digunakan untuk mengetahui apakah parameter yang digunakan signifikan terhadap model atau tidak. Setiap penaksiran parameter pada model Box-Jenkins dapat menggunakan Thitung. Pengujian hipotesis terhadap parameter adalah sebagai berikut (Cryer & Chan, 2008):

 H_0 : $\alpha = 0$ atau $\theta = 0$ (Parameter tidak signifikan dalam model)

 H_1 : $\propto \neq 0$ atau $\theta \neq 0$ (Parameter signifikan dalam model)

Statistik uji:

Statistic uji.
$$t = \frac{\hat{\alpha}}{sE\hat{\alpha}} \ atau \ t = \frac{\hat{\theta}}{sE\hat{\theta}} \tag{5}$$

 $\widehat{\alpha}$: nilai taksiran dari parameter \propto $\widehat{\theta}$: nilai taksiran dari parameter θ

 $SE\hat{\alpha}$: standar kesalahan dari nilai taksiran $\hat{\alpha}$ $SE\widehat{\theta}$: standar kesalahan dari nilai taksiran $\widehat{\theta}$

Jika taraf signifikan yang digunakan adalah sebesar ∝ maka H_o ditolak, jika $|t_{hitung}| > t_{\alpha/2;n-p}$ atau jika nilai $P_{value} < \infty$.

1.2.6 Pengujian Diagnostik

Pengujian diagnostik dilakukan dengan menguji spesifikasi model apakah model persamaan tersebut telah memadai untuk dijadikan model peramalan. Dilakukan uji Ljung-Box untuk mendeteksi adanya korelasi antar-residual, asumsi bahwa residual telah memenuhi whine noise yaitu deret variabel acak yang independen (tidak

berkorelasi), identik dan terdistribusi normal dengan rata-rata mendekati 0 (μ =0) dan standar deviasi tertentu (σ). Dilakukan uji korelasi untuk mendeteksi independensi residul dan uji kenormalan residual model. Pengujian untuk melihat residual telah *white noise* dengan menggunakan hipotesis sebagai berikut:

Hipotesis

 H_0 : $\propto = 0$ atau $\theta = 0$ (Parameter tidak signifikan dalam model)

 H_1 : $\propto \neq 0$ atau $\theta \neq 0$ (Parameter signifikan dalam model)

Statistik uji menggunakan Ljung-box:

$$Q_t = T(T+2) \sum_{k=1}^{k} \frac{\widehat{\rho}^{k^2}}{T-k}$$
 (6)

T: banyaknya data

K: banyaknya lag yang diuji

 $\widehat{\rho k}^2$: Sdugaan *Auto*korelasi

Kriteria keputusan yaitu H_0 ditolak jika $Q_t > X_{(\infty,df)}^2$ dengan derajat kebebasan K dikurangi banyaknya parameter pada model atau $P_{value} < \beta$ artinya \in_t adalah barisan yang tidak memiliki korelasi (Gooijer&Hyndman,2006).

1.2.7 Kriteria Pemilihan Model terbaik

Analisis deret waktu memberikan peramalan nilai masa depan berdasarkan data masa lalu. Tingkat keberhasilan dan keakuratan peramalan dapat diukur dengan menghitung kesalahan peramalan, di mana kriteria pemilihan model didasarkan pada residual dan kesalahan tersebut. Dalam peramalan, terdapat ketidakpastian yang diakomodasi dengan memasukkan unsur kesalahan dalam perumusan peramalan deret waktu. Penyimpangan dalam peramalan bukan hanya disebabkan oleh unsur kesalahan, tetapi juga oleh ketidakmampuan model peramalan dalam menangkap unsur lain dalam deret waktu, yang turut mempengaruhi besarnya penyimpangan tersebut. Untuk memilih model terbaik, residual yang baik harus memenuhi asumsi white noise dan berdistribusi normal. Oleh karena itu, diperlukan kriteria tertentu untuk menentukan model yang akan digunakan dengan menghitung akurasi dari data out sample (Wei,2006). Pengukuran akurasi peramalan untuk data out sample dapat diukur menggunakan beberapa indikator kesalahan peramalan yakni dengan menggunakan kriteria Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), dan Mean Absolute Deviation (MAD).

1. Root Mean Square Error (RMSE)

RMSE mengukur rata-rata akar kuadrat dari selisih antara nilai yang diprediksi dengan aktual. Ini memberikan penilaian terhadap seberapa jauh prediksi dari nilai sebenarnya, dengan penekanan yang lebih besar pada kesalahan yang lebih besar. Rumus RMSE adalah sebagai berikut (Tantika, 2018):

÷

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}...$$
(7)

 y_i : nilai aktual. \hat{y}_i : nilai prediksi.

n: jumlah pengamatan.

2. Mean Absolute Percentage Error (MAPE)

MAPE mengukur rata-rata presentase kesalahan absolut antara nilai yang diprediksi dengan nilai aktual. Ini dinyatakan dalam bentuk presentase dan memberikan gambaran tentang seberapa besar kesalahan relatif terhadap nilai aktual. Rumus MAPE adalah sebagai berikut:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{y_i - \widehat{y_i}}{y_i} \right|$$
 (8)

 y_i : nilai aktual.

 \widehat{y}_i : nilai prediksi.

n: jumlah pengamatan.

3. Mean Absolute Deviation (MAD)

MAD mengukur rata-rata absolut dari selisih antara nilai yang diprediksi dengan nilai aktual. Berbeda dengan RMSE, MAD memberikan penilaian yang lebih seimbang terhadap kesalahan besar dan kecil karena menggunakan nilai absolut.

$$MAD = \frac{1}{n} \sum_{i=1}^{n} |y_i - \widehat{y}_i| \tag{9}$$

 y_i : nilai aktual.

 \widehat{y}_{l} : nilai prediksi.

 $\it n$: jumlah pengamatan.

(Munawaroh, 2010).

1.2.8 Penelitian Terdahulu

Tabel 3. Penelitian Terdahulu

No	Nama Peneliti	Metode Penelitian	Hasil Penelitian
	(Tahun) dan Judul		
1	Adolfo Rodriguez	Random forests,	Prediksi dari model LSTM dan
	Vargas (2020) yang	extreme gradient	KNN univariat adalah yang
	berjudul	boosting, Long	terbaik, diikuti oleh r <i>andom</i>
	"Forecasting Costa	Short-Term Memory	forests dan extreme gradient
	Rican Inflation with	(LTSM Networks,	boosting. Kombinasi prediksi dari
	Machine Learning	dan dua varian <i>K-</i>	ketiga metode ini meningkatkan
	Methods."	Nearest Neighbors	kinerja dibandingkan dengan
		(KNN).	prediksi individual pada semua
			horizon waktu yang diuji. Hasil
			evaluasi menunjukkan bahwa
			kombinasi prediksi ini tidak bias,
			kesalahan prediksi tidak
			berkorelasi pada horizon h=1,
			dan mampu memprediksi arah
			perubahan inflasi dengan akurasi
			lebih dari 50% pada semua
			horizon.
2	Adam Richardson,	Model Autoregresif	-
	Thomas V. F.	(AR), gradient	
	Mulder, dan Tugrul	boosting, ridge	regression, dan neural networks
	Vehbi (2021) yang	regression, lasso	dapat mengurangi kesalahan

	berjudul "Nowcasting GDP using machine- learning algorithms: A real-time assessment."	regression, elastic net regression, support vector machine regression (SVM), dan neural network.	nowcasting rata-rata 20-30% dibandingkan dengan benchmark AR. Selain itu, algoritma pembelajaran mesin juga mengungguli model faktor dinamis. Model boosted trees secara khusus terbukti lebih unggul daripada prakiraan resmi yang dihasilkan oleh Bank Sentral Selandia Baru. Secara keseluruhan, penelitian ini menunjukkan bahwa penggunaan algoritma machine learning dalam nowcasting PDB memberikan hasil yang lebih akurat.
3	Luigi Longo, Massimo Riccaboni dan Armando Rungi (2022) yang berjudul "A Neural Network Ensemble Approach for GDP Forecasting."	Reccurrent Neural Networks (RNNs) dan Dynamic Factor Model (DFM) dengan Generalized Autoregressive Score (GAS).	Metode ensemble yang menggabungkan DFM dengan GAS dan RNN meningkatkan akurasi prediksi pertumbuhan PDB AS secara signifikan. Selama krisis keuangan 2008-09, metode ini mengurangi kesalahan prediksi sebesar 20% dan selama krisis COVID-19, metode ini mampu menangkap peningkatan PDB pada kuartal ketiga 2020 yang sebelumnya tidak terdeteksi oleh model lain. Penggunaan integrated gradients menunjukkan bahwa indikator stres keuangan dan klaim pengangguran pasca lockdown adalah variabel kunci dalam memahami evolusi kedua krisis ini.
4	Niko Hauzenberger, Florian Huber dan Karin Klieber (2023) yang berjudul "Real- time Inflation Forecasting using non linear dimension reduction techniques."	Autoencoder, Model AR (Autoregressive), Bayesian Model Averaging (BMA), Dynamic Model Averaging (DMA).	Autoencoder secara signifikan meningkatkan akurasi prediksi inflasi dibandingkan model tradisional. Autoencoder dengan jumlah faktor optimal menunjukkan peningkatan kinerja yang mencolok dalam metrik LPL (log predictive likelihood) dan RMSE (Root Mean Squared

		Error). DMA menghasilkan akurasi prediksi lebih tinggi untuk satu bulan dan satu kuartal ke depan dibandingkan model AR sederhana. Penggunaan priors yang berbeda juga mempengaruhi kinerja, dengan hasil menunjukkan peningkatan peforma prediksi dengan pilihan prior yang tepat.
5 Gustavo Silva Araujo dan Wagne Piazza Gaglianone (2023) yang berjudu "Machine Learning Methods Fo Inflation Forecasting in Brazil: new contenders versus classical models."	Learning (ML) berbasis regularisasi (elastic net, lasso, adaptive lasso, ridge). Metode berbasis	Metode ML mengungguli model ekonometrik tradisional dalam hal mean-squared error. Terdapat bukti adanya non-linearitas dalam dinamika inflasi yang relevan untuk peramaln inflasi. Set metode peramalan terbaik seringkali termasuk kombinasi peramalan, metode berbasis pohon (random forest dan xgboost), breakeven inflation, dan ekspektasi berbasis survei, dapat menghasilkan metode kombinasi quantile baru dan pendekatan hybrid ML yang berhasil memperkirakan inflasi dengan ketidakpastian yang terukur melalui chart kipas. Dengan alat identifikasi variabel kunci yang paling berpengaruh dalam memprediksi inflasi.

Penelitian Adolfo Rodriguez Vargas (2020) yang berjudul "Forecasting Costa Rican Inflation with Machine Learning Methods" menggunakan berbagai metode machine learning seperti random forest, extreme gradient boosting, Long Short-Term Memory (LSTM) Networks, dan dua varian K-Nearest Neighbors (KNN). Penelitian ini menemukan bahwa model LSTM dan KNN univariat memberikan hasil prediksi terbaik, diikuti oleh random forest dan extreme gradient boosting. Kombinasi prediksi dari ketiga metode ini meningkatkan kinerja prediksi dibandingkan prediksi individual pada semua

horizon waktu yang diuji. Sebaliknya, penelitian yang akan dilakukan menggunakan metode ARIMA yang berfokus pada otomatisasi proyek prediksi inflasi dan PDRB di Sulawesi Selatan.

Penelitian Adam Richardson, Thomas V. F. Mulder, dan Tugrul Vehbi (2021) yang berjudul "Nowcasting GDP using machine-learning algorithms: A real-time assessment" mengeksplorasi penggunaan berbagai algoritma machine learning seperti gradient boosting, ridge regresseion, lasso regression, elastic net regression, support vector machine regression (SVM), dan neural network untuk nowcasting PDB. Hasil penelitian mereka menunjukkan bahwa model seperti boosted trees, SVM dan neural network dapat mengurangi kesalahan nowcasting rata-rata 20-30% dibandingkan dengan benchmark AR. Di sisi lain, penelitian yang akan dilakukan tidak menggunakan berbagai algoritma machine learning tetapi berfokus pada penerapan model ARIMA untuk otomatisasi prediksi.

Penelitian Luigi Longo, Massimo Riccaboni dan Armando Rungi (2022) yang berjudul " *A Neural Network Ensemble Approach for GDP Forecasting*" menggunakan metode ensemble yang menggabungkan *Dynamic Factor Model* (DFM) dengan *Generalized Autoregressive Score* (GAS) dan *Recurrent Neural Networks* (RNN) untuk meningkatkan akurasi prediksi pertumbuhan PDB AS. Penelitian ini menunjukkan bahwa metode *ensemble* ini secara signifikan meningkatkan akurasi prediksi selama krisis keuangan dan pandemi COVID-19. Sebaliknya, penelitian yang akan dilakukan untuk inflasi dan PDRB di Sulawesi Selatan menggunakan satu model ekonometrik, yaitu ARIMA yang tidak menggabungkan beberapa model seperti dalam pendekatan *ensemble*. Penelitian ini berfokus pada efektivitas dan efisiensi model ARIMA dalam konteks otomatisasi.

Penelitian Niko Hauzenberger, Florian Huber dan Karin Klieber (2023) yang berjudul "Real-time Inflation Forecasting using non linear dimension reduction techniques" meencorngeksplorasi penggunaan teknik reduksi dimensi non-linear untuk memprediksi inflasi secara real-time. Metode yang digunakan termasuk Autoencoder, model AR (Autoregressive), Bayesian Model Averaging (BMA), dan Dynamic Model Averaging (DMA) untuk memprediksi inflasi secara real-time. Hasil penelitian menunjukan bahwa Autoencoder secara signifikan meningkatkan akurasi prediksi inflasi dibandingkan model tradisional, dan DMA menghasilkan akurasi prediksi yang lebih tinggi dibandingkan model AR sederhana. Penelitian yang akan dilakukan untuk inflasi dan PDRB Sulawesi Selatan berbeda dengan penelitian ini karena menggunakan metode ARIMA tanpa menggunakan teknik reduksi dimensi non-linear atau model averaging yang kompleks. Fokus penelitian ARIMA adalah pada penerapan model ekonometrik untuk otomatisasi prediksi dalam konteks lokal.

Penelitian Gustavo Silva Araujo dan Wagner Piazza Gaglianone (2023) yang berjudul "Machine Learning Methods For Inflation Forecasting in Brazil: new contenders versus classical models" mengeksplorasi peramalan inflasi di Brazil dengan menggunakan berbagai metode machine leraning dan model ekonometrik tradisional seperti ARMA, VAR, dan model faktor. Hasil penelitian menunjukkan bahwa metode machine learning mengungguli model ekonometrik tradisional dalam hal mean-squared

error. Pada penelitian yang akan dilakukan berfokus pada efektivitas dan efisiensi model ARIMA dalam konteks otomatisasi di wilayah spesifik dengan fokus pada Sulawesi Selatan. Penelitian tersebut tidak hanya memperluas penerapan metode *machine learning* tetapi juga menawarkan solusi praktis untuk tantangan ekonomi di tingkat regional. Kombinasi metode yang akan digunakan bertujuan mencapai hasil yang lebih akurat dan efektif, yang langsung dapat diterapkan dalam perencanaan dan pengambilan keputusan ekonomi di Sulawesi Selatan.

1.3 Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini antara lain:

- a. Seberapa akurat model ARIMA dalam memprediksi inflasi dan PDRB di Sulawesi Selatan berdasarkan metrik evaluasi?
- b. Model prediksi apa yang paling efektif untuk memprediksi inflasi dan PDRB di Sulawesi Selatan?
- c. Bagaimana tren dan pola data inflasi serta PDRB yang dihasilkan dari model prediksi *machine learning*?

1.4 Tujuan Penelitian

Adapun tujuan dalam penelitian ini antara lain:

- a. Mengevaluasi akurasi model ARIMA dalam menghasilkan prediksi yang tepat berdasarkan metrik evaluasi.
- b. Menentukan model prediksi inflasi dan PDRB yang efektif dengan machine learning.
- c. Membuat dan menganalisis grafik yang menggambarkan tren dan pola data inflasi serta PDRB yang dihasilkan dari model prediksi untuk memberikan pemahaman yang lebih baik tentang dinamika ekonomi di Sulawesi Selatan.

1.5 Batasan Masalah

Agar penelitian mengarah pada tujuan yang diharapkan maka untuk permasalahan diatas maka ditetapkan pembatasan sebagai berikut:

- a. Objek penelitian terfokus pada prediksi inflasi dan PDRB di Sulawesi Selatan.
- b. Data yang digunakan dalam penelitian ini mencakup data historis inflasi dan PDRB dari data internal indikator makroekonomi Bank Indonesia KPw Sulawesi Selatan.
- c. Metode yang akan digunakan adalah *Autoregressive Integrated Moving Average* (ARIMA)
- d. Alat bantu yang digunakan dalam penelitian ini adalah bahasa pemrograman python.

1.6 Manfaat Penelitian

Adapun manfaat yang diharapkan dapat diperoleh dari penelitian ini adalah sebagai berikut:

a. Bagi Penulis

Menambah wawasan dan kemampuan dalam mengaplikasikan ilmu-ilmu teknik industri dalam memecahkan permasalahan nyata di dunia industri, khususnya dalam memprediksi inflasi dan PDRB Sulawesi Selatan menggunakan *machine learning* dengan *tools* Pyhton.

b. Bagi Bank Indonesia

Hasil Penelitian diharapkan dapat memberikan kontribusi signifikan bagi peningkatan proses analisis ekonomi di Bank Indonesia, mendukung pembuatan kebijakan yang lebih informatif dan dapat mendukung kebijakan ekonomi yang efektif dan responsif.

c. Bagi Akademik

Diharapkan dengan adanya penelitian ini dapat menjadi masukan dalam melakukan penelitian mengenai integrasi *machine learning* dan model ARIMA serta dapat memberikan kontribusi dalam menambah wawasan keilmuan kepada civitas akademik dalam bidang Teknik Industri khususnya mengenai otomasi prediksi inflasi dan PDRB Sulawesi Selatan menggunakan *machine learning* dengan *tools* Pyhton.

BAB II METODOLOGI PENELITIAN

2.1 Objek Penelitian

Penelitian ini berfokus pada implementasi teknik *machine learning* untuk otomatisasi prediksi inflasi dan PDRB di Sulawesi Selatan. Objek dari penelitian ini adalah data historis inflasi dan PDRB yang diperoleh dari data internal indikator makroekonomi oleh Bank Indonesia KPw Sulawesi Selatan. Penelitian ini dilaksanakan di Bank Indonesia pada bulan Juni 2024. Penelitian ini menggunakan pendekatan kuantitatif dengan mengumpul data inflasi tahun 2019-2024 dan data PDRB tahun 2010-2024.

2.2 Teknik Pengumpulan dan Sumber Data

Teknik pengumpulan data dalam penelitian ini dimulai dengan memperoleh data historis inflasi dan PDRB dari indikator makroekonomi internal yang disediakan oleh Kantor Perwakilan Bank Indonesia (KPw) Sulawesi Selatan. Data inflasi yang dikumpulkan mencakup periode dari tahun 2019 hingga 2024, sementara data PDRB mencakup periode dari tahun 2010 hingga 2024. Penelitian ini dilaksanakan di Bank Indonesia selama periode februari hingga juni 2024. Proses pengumpulan data ini melibatkan akses langsung ke data internal Bank Indonesia, memastikan keakuratan dan relevansi data yang digunakan dalam penelitian ini.

Setelah data terkumpul, data tersebut diolah dan dipersiapkan untuk analisis lebih lanjut menggunakan skrip python. Data yang tersimpan dalam format Excel dibaca ke dalam *DataFrame* menggunakan pustaka 'pandas'. Langkah selanjutnya adalah memproses data tanggal agar sesuai dengan format *datetime*, termasuk menetapkan akhir kuartal sebagai titik referensi waktu. Data kemudian diatur ulang dengan menghapus kolom tanggal asli dan menetapkan kolom tanggal baru sebagai indeks. Dengan persiapan data ini, penelitian memastikan bahwa data siap digunakan untuk analisis dan pemodelan, memungkinkan penerapan teknik *machine learning* untuk prediksi inflasi dan PDRB secara otomatis. Pada penelitian ini terdapat beberapa tahap yang dilakukan dalam pengumpulan data yang meliputi:

a. Wawancara (interview)

Pengumpulan data ini dilakukan dengan cara mengadakan tanya jawab secara langsung dengan pihak-pihak yang berhubungan dengan permasalahan yang diangkat dalam penelitian guna mendapatkan data dan keterangan yang berlandaskan kepada tujuan penelitian.

b. Review Literatur (Literature Review)

Mengkaji penelitian sebelumnya yang relevan untuk memahami konteks, metode dan temuan yang berkaitan dengan inflasi, PDRB dan penggunaan *machine learning* dalam prediksi ekonomi.

c. Perumusan Masalah

Pada tahap ini menuliskan terkait dengan pertanyaan-pertanyaan yang akan ditemukan jawabannya pada penelitian yang akan dilakukan.

d. Pengumpulan Data

Pengumpulan data dari data historis inflasi dan PDRB yang diperoleh dari data internal indikator makroekonomi oleh Bank Indonesia KPw Sulawesi Selatan.

2.3 Teknik Pengolahan Data dan Analisis

Prosedur pengolahan data dilakukan berdasarkan pada tahapan-tahapan yang telah diurutkan, sebagai berikut.

1. Tahapan Pendahuluan

Raw data dari database file Excel dimuat dan dilakukan transfomasi format data dalam skrip Python dengan menentukan path ke file Excel yang berisi data dan kemudian membacanya kedalam DataFrame untuk dianalisis dan pemodelan lebih lanjut.

2. Visualisasi Data

Raw data dari database divisualisasikan menggunakan plot time series. Plot ini menampilkan data asli, rolling mean (rata-rata bergerak), dan rolling standard deviation (standar deviasi bergerak) untuk melihat pola dan volatilitas dalam data. Selain itu, data dipecah menjadi komponen trend, seasonal, dan residual menggunakan metode dekomposisi aditif dengan bantuan pustaka 'statsmodels'. Visualisasi komponen-komponen ini memberikan pemahaman yang lebih mendalam tentang struktur data deret waktu, seperti tren jangka panjang dan pola musiman.

3. Uji Stasioneritas

Stasioneritas adalah prasyarat penting untuk penerapan model ARIMA. Uji stasioneritas dilakukan menggunakan *Augmented Dickey-Fuller Test* yang disediakan oleh pustaka '*statsmodels*'. Uji ini membantu menentukan apakah data memiliki *mean, varians*, dan *Auto*korelasi yang konstan sepanjang waktu. Data yang stasioner lebih mudah dimodelkan dan diprediksi dengan akurat.

4. Penentuan Parameter untuk Model ARIMA

Parameter ARIMA terdiri dari tiga komponen: p (Autoregressive), d (differencing), dan q (moving average). Pada tahap ini, parameter awal ditentukan secara manual berdasarkan analisis grafik ACF (Autocorrelation Function) dan PACF (Partial Autocorrelation Function). Selain itu, metode Auto ARIMA digunakan untuk secara otomatis mengidentifikasi parameter terbaik berdasarkan kriteria seperti Akaike Information Criterion (AIC). Auto ARIMA memudahkan proses pemilihan parameter dengan mengeksplorasi berbagai kombinasi parameter dan memilih yang memberikan model terbaik.

5. Membuat dan Melatih Model ARIMA

Model ARIMA dibuat dan dilatih menggunakan dua pendekatan: manual dan otomatis. Dalam pendekatan manual, model dibuat dengan parameter yang dipilih secara manual, sedangkan dalam pendekatan otomatis, parameter dipilih oleh algoritma *Auto* ARIMA. Model dilatih menggunakan data historis tingkat inflasi dan PDRB. Proses pelatihan melibatkan penyesuaian parameter model untuk meminimalkan kesalahan prediksi dan memastikan bahwa model dapat menangkap pola dalam data dengan baik.

6. Evaluasi Model

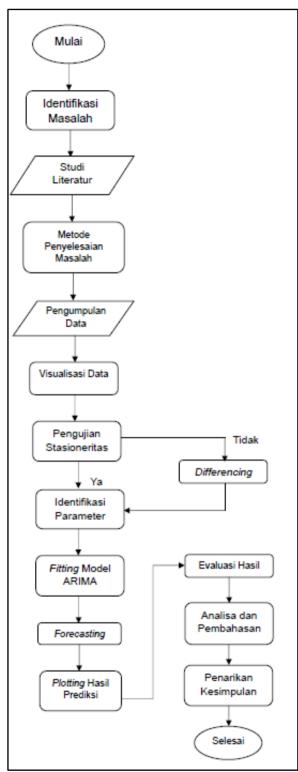
Model yang telah dilatih kemudian digunakan untuk melakukan prediksi nilai tingkat inflasi dan PDRB pada periode data yang tersedia. Hasil prediksi dibandingkan dengan nilai aktual untuk mengevaluasi kinerja model. Evaluasi dilakukan menggunakan metrik seperti *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan koefisien korelasi antara nilai prediksi dan nilai aktual. *Plot* diagnostik seperti *residual plot*, Q-Q *plot*, dan *Autocorrelation plot residual* digunakan untuk memeriksa kesesuaian model dan mengidentifikasi kemungkinan masalah dalam model.

7. Visualisasi Hasil

Hasil prediksi model ARIMA kemudian divisualisasikan bersama dengan data aktual untuk menilai kinerja model secara visual. Ini membantu dalam memahami seberapa baik model dalam menangkap pola dalam data *time series*.

2.4 Flowchart Penelitian

Flowchart penelitian merupakan salah satu cara bagi penulis dalam melakukan penelitian secara sistematis. Berikut pada gambar 3 adalah flowchart atau alur penelitian yang dilakukan oleh peneliti, antara lain:



Gambar 2. Flowchart Penelitian

a. Identifikasi masalah

Mengidentifikasi fenomena dari masalah yang disajikan pada latar belakang penelitian.

b. Studi literatur

Peneliti melakukan pencarian informasi mengenai teori-teori yang terkait dengan penelitian, yakni konsep dasar *machine learning, forecasting*, analisis deret waktu (*time series analysis*), identifikasi model, uji signifikansi parameter, pengujian diagnostik, kriteria pemilihan model terbaik, *Autoregressive Integrated Moving Average* (ARIMA).

c. Metode penyelesaian masalah

Pada tahap ini dilakukan penyusunan metode penyelesaian masalah berdasarkan hasil studi literatur yang telah dilakukan sebelumnya.

d. Pengumpulan data

Data ini adalah data data historis inflasi dan PDRB yang diperoleh dari data internal indikator makroekonomi oleh Bank Indonesia KPw Sulawesi Selatan.

e. Visualisasi data

Memvisualisasikan data untuk memahami pola, tren, dan musiman (seasonality) yang mungkin ada.

f. Pengujian stasioneritas

Uji stasioneritas seperti *Augmented Dickey-Fuller* (ADF) *test* biasanya digunakan untuk menguji apakah data sudah stasioner. Jika data tidak stasioner, langkah berikutnya adalah melakukan *differencing* untuk membuat data menjadi stasioner.

g. Differencing

Jika data tidak stasioner, maka dilakukan *differencing*, yaitu mengurangkan nilai-nilai data sebelumnya dari nilai-nilai data saat ini untuk menghilangkan tren atau pola musiman.

h. Identifikasi parameter

Nilai-nilai parameter ini bisa diidentifikasi menggunakan ACF (*Autocorrelation Function*) dan PACF (*Partial Autocorrelation Function*).

i. Fitting model ARIMA

Setelah diidentifikasi, dilakukan *fitting* data untuk menghitung parameter model yang optimal.

j. Forecasting

Jika model ARIMA telah memenuhi semua kriteria, maka model tersebut digunakan untuk memprediksi nilai masa depan dari deret waktu.

k. Plotting Hasil Prediksi

Visualisasi seberapa baik prediksi sesuai dengan data asli, dimana grafik akan menunjukkan perbandingan antara data aktual dan data prediksi.

I. Evaluasi Hasil

Hasil prediksi dibandingkan dengan data sebenarnya untuk mengevaluasi keakuratan model. Metrik seperti *Mean Absolute Error* (MAE) atau *Root Mean*

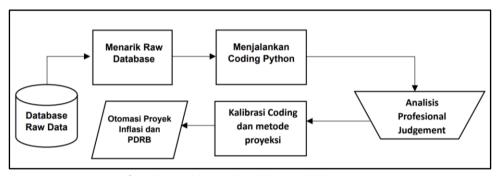
Squared Error (RMSE) bisa digunakan. Berdasarkan hasil evaluasi, model dengan peforma terbaik dipilih.

m. Analisis data dan Pembahasan Tahapan ini akan menjawab pertanyaan dari rumusan masalah pada penelitian ini.

n. Penarikan Kesimpulan Berdasarkan hasil analisis yang telah dilakukan, peneliti akan menarik kesimpulan dan memberikan saran.

2.5 Kerangka Berpikir

Kerangka berpikir adalah struktur konseptual yang digunakan untuk menjelaskan hubungan antara berbagai elemen dalam penelitian. Berikut merupakan kerangka pikir pada gambar 3.



Gambar 3. Kerangka pikir penelitian

Kerangka berpikir penelitian pada gambar 3 tersebut menunjukkan alur integrasi antara *Machine Learning* (ML) dan ARIMA dalam proses prediksi inflasi dan Produk Domestik Regional Bruto (PDRB). Proses dimulai dengan penarikan data mentah dari database, diikuti dengan menjalankan kode Python untuk melakukan analisis dan prediksi. Hasil prediksi kemudian dikalibrasi melalui penyesuaian kode dan metode prediksi agar sesuai dengan kondisi aktual. Selanjutnya, hasil prediksi ini dievaluasi menggunakan *profesional judgement* untuk memastikan ketepatan dan relevansi prediksi, sebelum akhirnya digunakan sebagai output akhir dalam proses pengambilan keputusan atau analisis lanjutan.