

**PENGEMBANGAN MODEL PREDIKSI DAN
DECISION SUPPORT SYSTEM BERBASIS
GENERATIVE AI DAN *MACHINE LEARNING*
UNTUK MANAJEMEN CADANGAN BERAS
PEMERINTAH DI BULOG**

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang pesat dalam beberapa dekade terakhir telah menyebabkan transformasi besar di berbagai sektor (Anantrasirichai & Bull, 2022). Salah satu pendorong utama transformasi ini adalah kemajuan teknologi kecerdasan buatan (*Artificial Intelligence/AI*) (C. Zhang & Lu, 2021), khususnya di bidang *Machine Learning* (ML) (Alpaydin, 2020). Teknologi ini telah merevolusi berbagai sektor, termasuk kesehatan (Arumugam, Naved, Shinde, Leiva-Chauca, Huaman-Osorio, & Gonzales-Yanac, 2023; Kaur, Singh, & Verma, 2023), energi (C. Zhao, Wan, & Song, 2021), dan manufaktur (Kotsiopoulos, Sarigiannidis, Ioannidis, & Tzovaras, 2021) dengan meningkatkan efisiensi operasional melalui analisis *big data*, pengenalan pola, dan pengambilan keputusan yang lebih baik (Mehmood, Bashir, Rabie, Broderick, & Davies, 2023; Rai, Tiwari, Ivanov, & Dolgui, 2021).

Sektor pertanian dan pangan juga tidak luput dari dampak positif *Machine Learning* (ML). Teknologi ini telah terbukti efektif dalam mengoptimalkan berbagai operasi, termasuk perencanaan, manajemen persediaan, dan distribusi produk, yang berkontribusi pada peningkatan efisiensi operasional di sektor pertanian dan pangan (Kuradusenge, Hitimana, Hanyurwimfura, Rukundo, Mtonga, Mukasine, Uwitonze, Ngabonziza, & Uwamahoro, 2023; Lutoslawski, Hernes, Radomska, Hajdas, Walaszczyk, & Kozina, 2021).

Pangan merupakan kebutuhan dasar manusia yang harus dipenuhi. Hak untuk mendapatkan pangan merupakan salah satu hak asasi manusia, seperti yang dijelaskan dalam Pasal 27 Undang-Undang Dasar 1945 dan Deklarasi Roma 1996. Sebagai kebutuhan dasar dan salah satu hak asasi manusia, pangan memainkan peran penting dalam stabilitas sosial, ekonomi, dan politik suatu negara (BULOG, 2024; Clapp, Moseley, Burlingame, & Termine, 2022).

Negara memiliki kewajiban untuk memenuhi kebutuhan pangan penduduknya sebagaimana pemenuhan hak-hak lainnya. Mengabaikan hak ini

dapat dianggap sebagai pelanggaran hak asasi manusia. Ketika hak atas pangan terus menerus diabaikan, maka hal itu sama saja dengan kejahatan genosida (Maskun, Naswar, Ahmad, Ilmar, Napang, & Nugraha, 2021). Hal ini menyoroti perlunya mengamankan ketahanan pangan nasional.

Menurut UU No. 18 Tahun 2012, ketahanan pangan adalah kondisi terpenuhinya pangan yang cukup dan berkualitas bagi seluruh masyarakat. Gangguan pada ketahanan pangan, seperti meroketnya harga pangan pada saat krisis ekonomi 1997/1998, dapat memicu gejolak sosial yang membahayakan stabilitas ekonomi dan keamanan nasional (BULOG, 2024). Kejadian tersebut menunjukkan pentingnya menjaga ketahanan pangan untuk menjaga stabilitas ekonomi, politik, dan keamanan nasional (BULOG, 2024; Clapp et al., 2022).

Di Indonesia, pangan sering dikaitkan dengan komoditas beras karena beras merupakan makanan pokok dan memegang peran penting dalam ketahanan pangan. Hal ini tercermin dari tingkat konsumsi beras yang berkisar antara 85,88% hingga 91,73% per tahun (Alfa & Subagyo, 2018) dan fakta bahwa masyarakat Indonesia mengalokasikan lebih dari 50% dari total pengeluarannya untuk membeli beras (Hafizah, Hakim, Harianto, & Nurmawati, 2020). Hal ini menunjukkan pentingnya menjaga ketersediaan beras untuk ketahanan pangan (Octania, 2021).

Menurut Permendag Nomor 4 Tahun 2012 dan Instruksi Presiden Nomor 5 Tahun 2015, Cadangan Beras Pemerintah (CBP) yang dikelola oleh Perum BULOG merupakan instrumen kunci dalam menjaga stabilitas harga beras, ketersediaan pangan, dan ketahanan pangan nasional. CBP berfungsi sebagai *buffer stock* untuk mengintervensi pasar ketika terjadi lonjakan harga dan memenuhi kebutuhan pangan masyarakat yang terdampak bencana alam.

Terlepas dari pentingnya CBP bagi ketahanan pangan, manajemen CBP di Indonesia menghadapi beberapa tantangan yang kompleks, yang meliputi pertumbuhan populasi, penurunan produksi, dan perubahan iklim. Pertumbuhan populasi yang pesat menyebabkan peningkatan permintaan beras yang signifikan. Hal ini tidak sejalan dengan penurunan produksi yang terus terjadi akibat dari perubahan iklim, bencana alam, dan alih fungsi lahan pertanian yang mengancam

ketersediaan persediaan beras (Beltran-Peña, Rosa, & D'Odorico, 2020; Lee, Lin, & Petway, 2018). Hal ini seringkali menyebabkan fluktuasi harga yang tidak terkendali, sehingga membebani masyarakat, terutama pada masyarakat dengan ekonomi rendah, dan dapat mengganggu stabilitas ketahanan pangan nasional dan menyebabkan kerawanan pangan (Octania, 2021).

Perubahan iklim menyebabkan kondisi cuaca ekstrem dan perubahan pola hujan, yang secara langsung berdampak pada pola tanam dan produksi pangan, serta meningkatkan risiko gagal panen dan ketidakstabilan persediaan (Amir, Saqib, Khan, Khan, Bokhari, Zaman-ul-Haq, & Majid, 2020; Farooq, Uzair, Raza, Habib, Xu, Yousuf, Yang, & Ramzan Khan, 2022; Hasegawa, Sakurai, Fujimori, Takahashi, Hijioka, & Masui, 2021; Mukherjee, 2021). Faktor-faktor ini secara akumulatif menciptakan ketidakpastian yang tinggi dalam manajemen CBP sehingga menyulitkan pengambilan keputusan yang akurat dan tepat waktu.

Sesuai Pasal 5 Peraturan Badan Pangan Nasional Nomor 12 Tahun 2022, manajemen Cadangan Beras Pemerintah (CBP) melibatkan beberapa tahapan penting, meliputi pengadaan, penyimpanan, dan distribusi. Tahapan pengadaan dimulai dengan mengidentifikasi kebutuhan persediaan berdasarkan prediksi permintaan (kebutuhan) dan produksi, diikuti dengan perencanaan pengadaan yang melibatkan tugas-tugas penting seperti penentuan jumlah cadangan persediaan.

Metode konvensional yang digunakan dalam memprediksi produksi dan permintaan beras saat ini, seperti metode logika fuzzy dan regresi linier berganda seringkali kurang akurat karena memiliki keterbatasan dalam menangani data yang kompleks, tidak pasti, dan tidak mampu menangkap berbagai faktor eksternal seperti pertumbuhan populasi, penurunan produksi, dan perubahan iklim (Sari, Ananda, & Rani, 2020). Keterbatasan ini menyulitkan dalam penentuan prediksi yang akurat, penentuan jumlah cadangan yang optimal, dan pengambilan keputusan pengadaan. Kesalahan dalam memprediksi produksi dan permintaan dapat menyebabkan kelebihan atau kekurangan persediaan, yang berdampak pada naiknya harga beras dan dapat mengganggu stabilitas ekonomi serta sosial

(Octania, 2021). Ketiadaan sistem pendukung keputusan yang komprehensif juga menghambat pengambilan keputusan strategis terkait pengadaan CBP.

Menghadapi tantangan tersebut, diperlukan pendekatan yang lebih inovatif dan berbasis data. Teknologi *Machine Learning* (ML) menawarkan potensi besar dalam meningkatkan akurasi prediksi, optimasi jumlah cadangan, dan pengambilan keputusan yang tepat. Pemanfaatan teknologi ini dapat membantu mengatasi tantangan kerawanan pangan secara efektif, memastikan ketersediaan pangan, dan mendukung stabilitas harga di pasar nasional (Mehmood et al., 2023; Rai et al., 2021).

Penelitian terdahulu telah mengusulkan berbagai model prediksi produksi dan permintaan. Noorunnahar, Chowdhury, and Mila (2023) membandingkan dua model yaitu ARIMA dan XGBoost untuk memprediksi produksi beras di Bangladesh. Hasil penelitian menunjukkan bahwa model XGBoost mengungguli model ARIMA dengan nilai MAE dan RMSE sebesar 2.779.742 dan 3.195.985, namun model ini tidak melibatkan variabel lain seperti suhu dan curah hujan, serta nilai *error* masih tinggi.

Satpathi, Setiya, Das, Nain, Jha, Singh, and Singh (2023) membandingkan lima metode, yaitu SMLR, ANN, LASSO, ELNET, Ridge Regression, dan *ensemble models* untuk memprediksi hasil panen padi di Chhattisgarh. Hasil penelitian menunjukkan bahwa model prediksi dengan ANN mengungguli model lainnya dengan nilai R^2 sebesar 1 dan 0,99 serta nilai RMSE sebesar 0,003 dan 0,004 yang menunjukkan kemampuan model dalam menangani berbagai variabel prediksi. Namun terdapat beberapa model yang menunjukkan kinerja baik selama pelatihan tetapi tidak selama proses pengujian, yang disebabkan oleh *overfitting*.

Nassibi, Fasihuddin, and Hsairi (2023) membandingkan dua metode yaitu Long Short-Term Memory (LSTM) dan Support Vector Machine (SVM) untuk memprediksi permintaan produk makanan. Hasil penelitian ini menunjukkan bahwa model LSTM mengungguli model SVM dengan nilai MAPE dan RMSE antara 3,78-15,3 dan 2,24-4,64 dan berkontribusi pada pengurangan limbah. Namun, kinerja model ini bergantung pada kualitas dan kelengkapan data.

Penelitian terdahulu telah menunjukkan bahwa *machine learning* (ML) dapat meramalkan produksi dan permintaan dengan lebih akurat daripada metode konvensional seperti ARIMA, namun model-model ini masih memiliki keterbatasan, seperti risiko *overfitting* dan tidak mempertimbangkan variabel eksternal, seperti perubahan iklim. Di sisi lain, kinerja model-model ini juga sangat bergantung pada kualitas dan kuantitas data, yang sering kali menghasilkan prediksi yang kurang akurat.

Solusi untuk mengatasi keterbatasan ini adalah dengan mengintegrasikan *Generative AI* (Gen AI) ke dalam model prediksi berbasis ML. Gen AI dapat menghasilkan data sintetis yang berkualitas tinggi untuk melengkapi dan memperbaiki data yang ada, sehingga meningkatkan kualitas *dataset* yang digunakan untuk prediksi (Yu & Guo, 2023). Teknologi ini juga dapat berguna untuk menciptakan skenario pelatihan yang beragam, yang membantu model belajar dari data yang lebih beragam dan mengurangi *overfitting* (Manjunath & Palayyan, 2023).

Penelitian ini bertujuan untuk mengembangkan model prediksi yang lebih akurat dengan memanfaatkan teknologi gen AI dan ML. Model ini akan memanfaatkan GAN, sebuah tipe *generative AI* yang dapat menghasilkan data yang realistis, sehingga dapat mengatasi ketergantungan pada data berkualitas tinggi dalam ML (Harshvardhan, Gourisaria, Pandey, & Rautaray, 2020). Algoritma ML yang digunakan adalah XGBoost yang telah menunjukkan kemampuan untuk menangani distribusi data yang kompleks, efisien secara komputasi, dan dapat mengatasi *overfitting* melalui pengenalan istilah regularisasi dalam *loss function* (C. Qin, Zhang, Bao, Zhang, Liu, & Liu, 2021; Jiayi Wang & Zhou, 2024; J. Wu, Li, & Ma, 2021). XGBoost ini digunakan untuk menganalisis data historis produksi dan konsumsi beras, data cuaca dari BMKG (Badan Meteorologi, Klimatologi, dan Geofisika), data luas tanam dan produktivitas dari Kementerian Pertanian, dan data lainnya. Model ini juga akan mempertimbangkan variabel lain yang relevan, seperti harga beras (Iqbal, 2019; Sossou & Igue, 2019) dan pendapatan per kapita (Bashir & Yuliana, 2019; Yusuf, Yusuf, Adesope, & Adebayo, 2020).

Penelitian ini juga akan menghitung cadangan beras pemerintah yang optimal dengan menggunakan *Deep Reinforcement learning* (DRL) dengan mempertimbangkan produksi nasional, keadaan darurat, stabilisasi harga, perjanjian internasional, dan kecukupan gizi sesuai dengan Pasal 2 dari Peraturan Badan Pangan Nasional Nomor 12 Tahun 2022. Penggunaan DRL memungkinkan penentuan kebijakan yang adaptif dan dinamis untuk mengoptimalkan cadangan beras dalam menghadapi ketidakpastian (Abu Zwaida, Pham, & Beauregard, 2021; Perez, Hubbs, Li, & Grossmann, 2021).

Penelitian ini juga akan membangun sistem pendukung keputusan (DSS) yang interaktif dan berbasis data untuk mendukung pengambilan keputusan strategis. DSS ini akan mengintegrasikan model prediksi, perhitungan cadangan optimal, dan informasi relevan lainnya untuk memberikan rekomendasi yang komprehensif kepada Perum Bulog terkait pengadaan CBP. DSS ini akan dilengkapi dengan antarmuka pengguna yang intuitif dan visualisasi data yang interaktif, seperti grafik, tabel, dan peta, sehingga memudahkan pengguna dalam memahami informasi dan mengambil keputusan. Melalui pengintegrasian teknologi-teknologi ini, diharapkan model dan sistem yang diusulkan dapat meningkatkan efisiensi operasional Perum BULOG dan menguatkan ketahanan pangan.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana mengembangkan model *generative AI* yang mampu menghasilkan data sintetis untuk meningkatkan akurasi model prediksi permintaan dan produksi beras?
2. Bagaimana mengembangkan model *machine learning* yang efektif untuk memprediksi permintaan dan produksi beras?
3. Bagaimana merancang model optimasi yang efektif untuk menentukan jumlah cadangan beras pemerintah yang optimal?
4. Bagaimana membangun *prototype decision support system* untuk cadangan beras pemerintah?

1.3 Tujuan Penelitian

Berdasarkan pokok permasalahan yang telah dirumuskan dalam rumusan masalah, maka tujuan penelitian ini adalah sebagai berikut:

1. Mengembangkan model *generative AI* untuk menghasilkan data sintetis yang realistis yang dapat digunakan sebagai data pelatihan untuk model prediksi permintaan dan produksi beras.
2. Mengembangkan model *machine learning* yang efektif dalam memprediksi permintaan dan produksi beras.
3. Merancang model optimasi yang efektif untuk menentukan jumlah cadangan beras pemerintah yang optimal.
4. Membangun *prototype decision support system* untuk cadangan beras pemerintah.

Penelitian terdahulu telah menunjukkan berbagai pendekatan dan metode yang berbeda untuk memprediksi produksi dan permintaan, optimasi persediaan, serta sistem Pendukung Keputusan (DSS) di berbagai sektor. Masing-masing model tersebut memiliki keterbatasan yang perlu diatasi, seperti bergantung pada kualitas data dan *overfitting*. Di sisi lain, integrasi teknologi Gen AI dan pembelajaran mesin dalam sektor pertanian dan pangan masih tergolong baru dan belum banyak dieksplorasi. Penelitian ini bertujuan untuk mengembangkan model yang lebih adaptif, yang mampu mengatasi fluktuasi dalam produksi dan permintaan, serta memberikan rekomendasi untuk jumlah cadangan beras pemerintah yang optimal. Penelitian ini juga akan membangun DSS yang akan memberikan rekomendasi strategi pengadaan beras yang efektif untuk menjamin ketahanan pangan. Penelitian ini akan menggunakan teknologi Gen AI dan ML dengan mempertimbangkan variabel-variabel yang mempengaruhi manajemen persediaan cadangan beras pemerintah. Gambar 2.5 memperlihatkan *fishbone diagram* yang menggambarkan usulan penelitian ini.

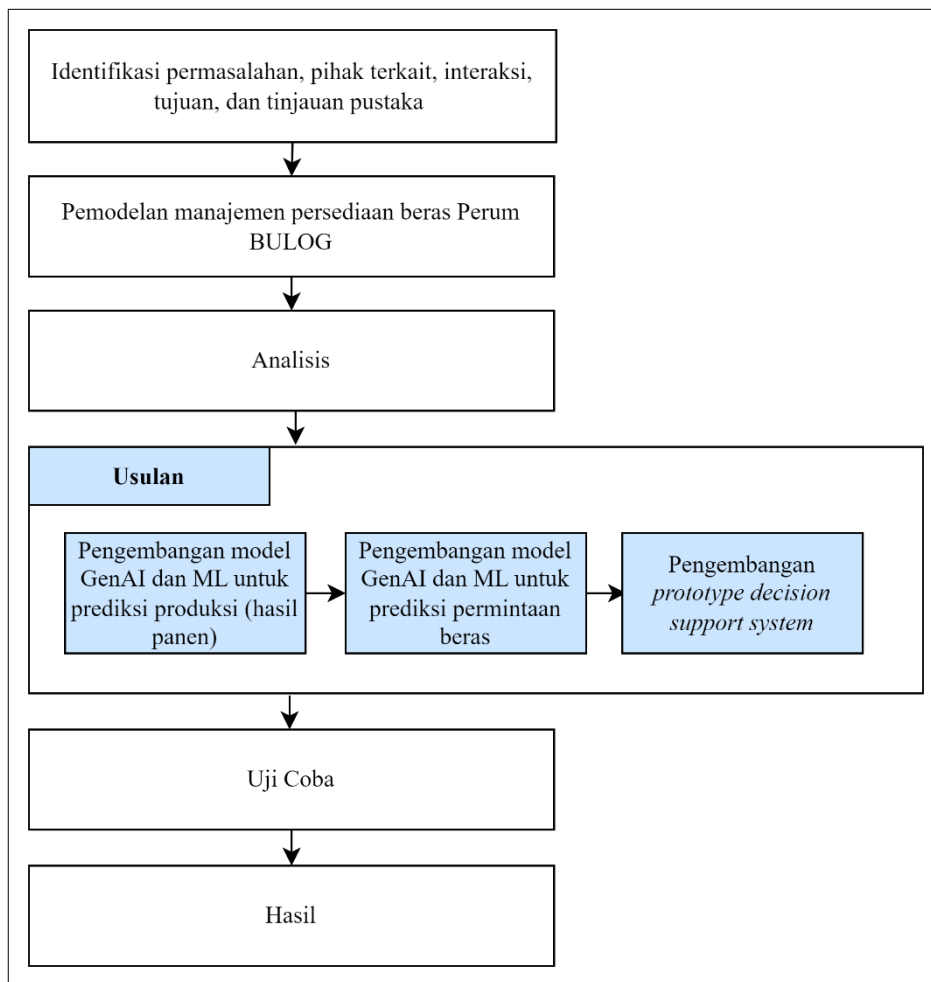
BAB 3

METODE PENELITIAN

3.1 Tahapan Penelitian

Tahapan penelitian merupakan gambaran dari langkah-langkah atau proses yang akan dilakukan dalam suatu penelitian. Penelitian ini terdiri dari lima tahapan. Tahap pertama adalah mengidentifikasi permasalahan, pihak terkait, interaksi, tujuan, dan tinjauan pustaka. Tahapan kedua adalah membuat model manajemen persediaan beras Perum BULOG berdasarkan hasil wawancara awal dengan pihak terkait. Tahap ketiga adalah melakukan analisis terhadap model manajemen persediaan beras Perum BULOG untuk mengidentifikasi area yang perlu ditingkatkan. Hasil analisis ini akan digunakan untuk merumuskan solusi terhadap permasalahan yang ada.

Tahap keempat adalah pengembangan solusi berbasis teknologi yang terdiri dari pengembangan berbagai model dan *prototype* sistem yang akan diuji. Usulan yang pertama adalah model *generative* AI untuk menghasilkan data sintetis yang realistis yang dapat digunakan sebagai data pelatihan untuk model prediksi. Model ini kemudian diintegrasikan ke dalam model ML prediksi produksi (hasil panen). Usulan yang kedua adalah model prediksi permintaan beras yang merupakan model yang mirip dengan model prediksi produksi (hasil panen) beras dengan beberapa penyesuaian agar sesuai dengan karakteristik data untuk prediksi permintaan beras. Usulan yang ketiga adalah pengembangan *prototype decision support system* yang mengintegrasikan model prediksi dan optimasi untuk mendukung kebijakan terkait pengadaan cadangan beras. Tahap kelima adalah uji coba terhadap *prototype decision support system*. Gambar 3.1 adalah tahapan pada penelitian ini.



Gambar 3. 1 Tahapan Penelitian

3.2 Pemodelan Manajemen Persediaan Beras Perum BULOG

Manajemen persediaan cadangan beras nasional telah menjadi perhatian penting dalam beberapa tahun terakhir karena meningkatnya permintaan pangan global, perubahan iklim, dan ketidakstabilan ekonomi. Cadangan ini merupakan stok strategis yang diawasi oleh pemerintah untuk menstabilkan persediaan dan harga beras, memberikan bantuan saat terjadi kekurangan pangan, dan mendukung tujuan ketahanan pangan nasional. Manajemen persediaan cadangan beras yang efektif sangat penting untuk memitigasi risiko yang terkait dengan gangguan persediaan dan fluktuasi harga beras (Yulianis & Rachman, 2021), yang pada akhirnya akan menjamin ketahanan pangan dan stabilitas ekonomi (Octania, 2021; Usdianto & Setiyowati, 2023).

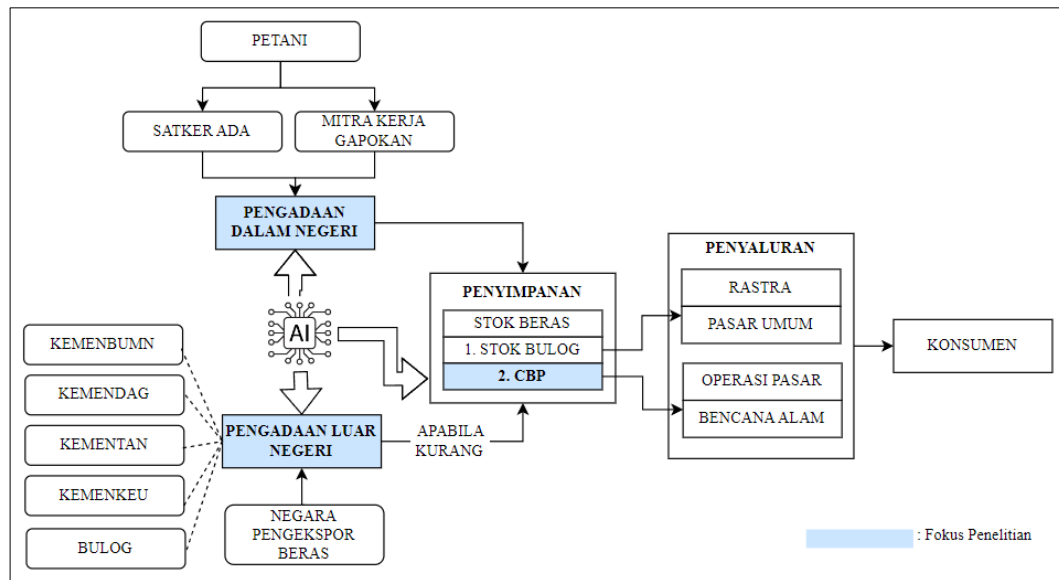
Para pemangku kepentingan yang terlibat dalam manajemen cadangan beras ini termasuk Kementerian Pertanian, Kementerian Perdagangan, Kementerian Keuangan, dan Kementerian Badan Usaha Milik Negara sebagai regulator, serta Perum BULOG yang bertanggung jawab untuk mengelola persediaan beras pemerintah dan stabilisasi harga di tingkat produsen dan konsumen (Octania, 2021). Perum BULOG bertanggung jawab untuk menjaga stabilitas harga dengan membeli gabah dan beras dari petani dengan harga yang ditentukan pemerintah ketika harga beli gabah turun, sehingga melindungi petani dari kerugian, dan menjual beras dengan harga yang lebih rendah daripada harga pasar ketika terjadi kenaikan harga beras, untuk memastikan keterjangkauan harga beras bagi masyarakat (Octania, 2021). Lembaga ini bertanggung jawab atas manajemen salah satu komponen cadangan beras nasional, yaitu Cadangan Beras Pemerintah (CBP), termasuk pada pengadaan (dalam negeri dan impor), penyimpanan, dan penyaluran beras untuk kebutuhan stabilisasi harga, bantuan pangan, dan keadaan darurat (Fang, Chen, Zhang, Pei, Gao, & Wang, 2020; Octania, 2021).

Beberapa penelitian telah menekankan peran penting Perum BULOG dalam manajemen persediaan cadangan beras di Indonesia. Melalui manajemen CBP, Perum BULOG memainkan peran penting dalam menjaga ketahanan pangan nasional, terutama saat terjadi fluktuasi harga atau gangguan persediaan. Keberadaan CBP yang dikelola Perum BULOG tidak hanya menstabilkan harga beras di pasar, tetapi juga menjamin ketersediaan beras bagi masyarakat, sehingga berkontribusi terhadap stabilitas ekonomi nasional (Octania, 2021; Putro, Purwaningsih, Sensuse, & Suryono, 2022; Silalahi et al., 2019). Mengingat peran penting ini, penerapan teknologi AI dapat membantu Perum BULOG dalam mengoptimalkan berbagai aspek dalam manajemen cadangan beras pemerintah, seperti prediksi permintaan dan produksi (hasil panen), optimasi cadangan beras, dan pengambilan keputusan yang lebih baik.

AI dapat digunakan untuk menganalisis data historis dan *real-time* guna menghasilkan prediksi yang akurat mengenai permintaan dan produksi (hasil panen) beras (Mehmood et al., 2023; Rai et al., 2021), sehingga memungkinkan

Perum BULOG untuk mengoptimalkan cadangan beras, menghindari kelebihan atau kekurangan cadangan beras, dan mengambil keputusan yang lebih baik dalam manajemen cadangan beras pemerintah (H. Qin, 2023).

Berdasarkan kajian model manajemen persediaan beras Perum BULOG, maka penelitian ini berfokus pada pemanfaatan teknologi AI untuk efektivitas manajemen cadangan beras pemerintah, terutama pada proses pengadaan. Gambar 3.1 menggambarkan model manajemen persediaan beras Perum BULOG.



Gambar 3. 2 Model Manajemen Persediaan Perum BULOG

3.3. Analisis

Analisis ini bertujuan untuk mengungkap kelemahan dan proses yang kompleks dalam manajemen persediaan cadangan beras pemerintah di Perum BULOG. Penelitian ini menggunakan metode analisis SWOT untuk mengidentifikasi titik-titik lemah yang krusial dalam manajemen persediaan cadangan beras pemerintah dan mengembangkan strategi untuk meningkatkan efisiensi dan efektivitas pengelolaan persediaan cadangan beras di Indonesia

Analisis SWOT bertujuan untuk mengetahui kekuatan, kelemahan, peluang, dan ancaman bisnis. Analisis lingkungan internal difokuskan untuk mengetahui kekuatan dan kelemahan, sedangkan analisis lingkungan eksternal difokuskan untuk mengetahui peluang dan ancaman (Putra, Pujangkoro, & Situmorang, 2022).

1. Analisis SWOT

1. *Strengths* (Kekuatan):

S1: Dukungan Pemerintah

Perum BULOG didukung oleh berbagai kebijakan pemerintah yang bertujuan menjaga stabilitas harga dan ketersediaan beras, seperti yang diatur dalam UU No. 18 Tahun 2012 tentang Pangan dan Perpres No. 48 Tahun 2016 tentang penugasan kepada Perum BULOG. Hal ini memberikan akses terhadap dukungan kebijakan dan finansial yang kuat, termasuk alokasi anggaran khusus untuk pengadaan CBP (Anggraini, Faqih, Sangadji, Kadarisman, & Revany, 2021; Octania, 2021; Utomo, 2020).

S2: Infrastruktur Logistik yang Memadai

Perum BULOG memiliki infrastruktur logistik yang cukup baik, termasuk gudang penyimpanan yang tersebar di berbagai daerah, yang berperan penting dalam menjaga kecukupan persediaan cadangan beras (Octania, 2021; Utomo, 2020)

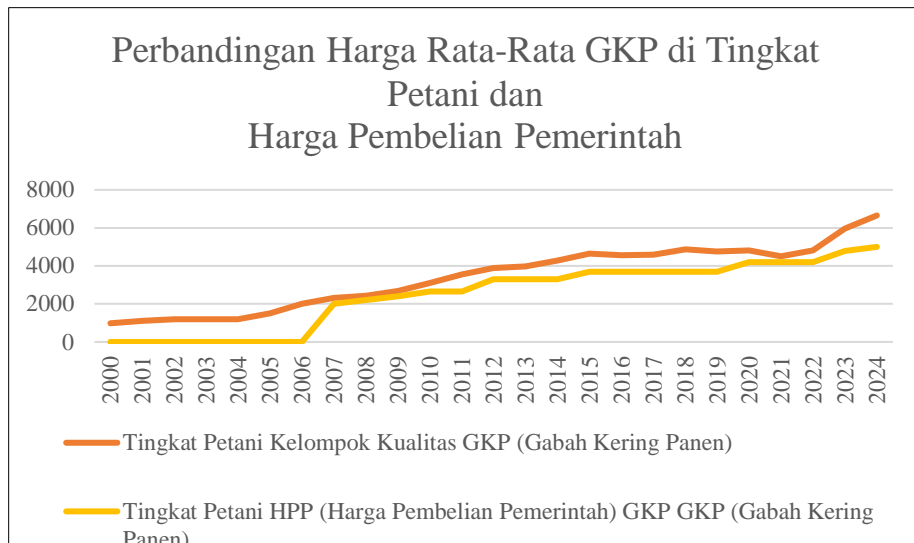
S3: Pengalaman dan Keahlian

Perum BULOG memiliki pengalaman puluhan tahun dalam manajemen persediaan beras, mulai dari pengadaan hingga distribusi, yang menjadi keunggulan dalam menjaga stabilitas harga dan persediaan beras (Anggraini et al., 2021; Utomo, 2020).

2. *Weaknesses* (Kelemahan):

W1: Ketidakmampuan untuk Bersaing dengan Sektor Swasta

Perum BULOG sering menghadapi tantangan dalam bersaing dengan sektor swasta, yang mampu menawarkan harga lebih tinggi kepada petani (Octania, 2021). Gambar 3.3 menggambarkan perbedaan yang signifikan antara harga gabah di tingkat petani dengan harga beli yang ditetapkan pemerintah. Kondisi ini menyebabkan petani lebih memilih menjual hasil panennya ke sektor swasta.



Gambar 3. 3 Perbandingan Harga Gabah (BPS, 2023)

W2: Ketergantungan pada Impor

Meskipun Perum BULOG memprioritaskan pengadaan dalam negeri untuk memenuhi CBP, tetapi masih terdapat ketergantungan pada impor beras, terutama selama periode penurunan produksi dalam negeri, yang membuat persediaan CBP rentan terhadap fluktuasi harga dan kebijakan perdagangan internasional (Octania, 2021; Utomo, 2020). Tabel 3.1 dan Gambar 3.4 menunjukkan banyaknya jumlah impor beras yang dilakukan Perum BULOG setiap tahunnya.

Tabel 3. 1 Jumlah Impor Beras (BPS, 2024)

Negara Asal	2017	2018	2019	2020	2021	2022	2023
	Berat Bersih: Ton						
India	32209.7	337999	7973.3	10594.4	215386.46	178533.57	69715.7
Thailand	108944.8	795600.1	53278	88593.1	69360.037	80182.506	1381921.2
Vietnam	16599.9	767180.9	33133.1	88716.4	65692.874	81828.039	1147705.3
Pakistan	87500	310990	182564.9	110516.5	52479.011	84407	309309.7
Myanmar	57475	41820	166700.6	57841.4	3790	3830	141204
Jepang	72.1	0.2	90	0.3	230.291	56.087	61.5
Tiongkok	2419	227.7	24.3	23.8	42.601	6	7
Lainnya	54.3	6.5ssss	744.6	0.3	760.146	364.065	12933.3
Total	305274.8	2253824.4	444508.8	356286.2	407741.42	429207.27	3062857.6



**Gambar 3. 4 Jumlah Impor Beras
(BPS, 2024)**

W3: Manajemen Persediaan yang Tidak Efisien

Terdapat ketidakmampuan dalam memprediksi permintaan dan persediaan secara akurat dan *real-time*, yang menyebabkan kelebihan dan kekurangan persediaan (Anggraini et al., 2021; Utomo, 2020).

W4: Keterbatasan Teknologi dan Transparansi

Sistem yang ada saat ini tidak memiliki kapasitas untuk mencatat transaksi secara transparan, yang mengakibatkan meningkatnya risiko kebocoran dan penipuan di sepanjang rantai pasok (Anggraini et al., 2021; Utomo, 2020).

W5: Koordinasi Antar Lembaga

Koordinasi antara Perum BULOG dan lembaga pemerintah lainnya, seperti Kementerian Pertanian dan Kementerian Perdagangan, terkadang menghadapi tantangan karena adanya perbedaan data, sehingga berdampak pada kelancaran pengambilan keputusan kebijakan impor (Siahaan, 2023).

3. *Opportunities* (Peluang):

O1: Memanfaatkan Teknologi Canggih

Integrasi teknologi canggih seperti *blockchain* dan AI berpotensi meningkatkan presisi, transparansi, dan efisiensi dalam mengelola persediaan CBP (Anggraini et al., 2021; Putro et al., 2022; Utomo, 2020).

O2: Pengembangan Sistem yang Terintegrasi

Terdapat peluang untuk mengembangkan sistem yang lebih terintegrasi dan canggih yang dapat memfasilitasi pengelolaan CBP yang lebih baik (Anggraini et al., 2021; Putro et al., 2022; Utomo, 2020).

4. *Threats* (Ancaman):

T1: Ketidakstabilan Harga dan Persediaan Global

Ketidakstabilan harga dan persediaan beras di pasar global dapat mempengaruhi kemampuan Perum BULOG untuk mengimpor beras dalam jumlah yang cukup dan dengan harga yang stabil (Octania, 2021).

T2: Dampak Perubahan Iklim

Perubahan iklim yang ekstrim dapat mengganggu produksi beras dalam negeri, sehingga meningkatkan risiko ketidakcukupan CBP dan fluktuasi harga beras di tingkat konsumen (Octania, 2021; Saud, Wang, Fahad, Alharby, Bamagoos, Mjrashi, Alabdallah, AlZahrani, AbdElgawad, & Adnan, 2022).

2. Strategi

Setelah dilakukan analisis SWOT, maka dapat dirumuskan strategi yang dapat digunakan Perum BULOG untuk meningkatkan kekuatan, mengatasi kelemahan, memanfaatkan peluang, dan meminimalkan potensi ancaman. Berikut adalah rincian strategi-strategi tersebut.

a. Strategi S-O (*Leveraging Strengths to Optimize Opportunities*)

SO1: Pemanfaatan Teknologi *Blockchain* untuk Transparansi dan Efisiensi

Memanfaatkan dukungan pemerintah untuk mengadopsi teknologi *blockchain* dalam manajemen pangan yang dapat meningkatkan transparansi, efisiensi, dan keamanan dalam transaksi dan pencatatan. Teknologi ini membantu dalam pencatatan, mengotomatisasi kontrak, memverifikasi transaksi, dan melacak keaslian produk dari produsen ke konsumen (S1, O1).

SO2: Pengembangan Sistem Terintegrasi

Meningkatkan infrastruktur yang sudah ada dengan mengembangkan sistem manajemen data yang terintegrasi dan canggih untuk memperkuat pengelolaan CBP secara lebih efektif (S2, O2).

SO3: Pengembangan Model Prediksi Permintaan dan Produksi (Hasil Panen)

Memanfaatkan AI untuk mengembangkan model prediksi permintaan dan produksi (hasil panen) beras yang akurat, sehingga memungkinkan pengambilan keputusan pengadaan yang lebih tepat untuk menghindari kelebihan atau kekurangan persediaan (S3, O1, O2).

SO4: Optimalisasi Program Peningkatan Produksi Pangan

Memanfaatkan tenaga ahli dan infrastruktur yang ada untuk mendukung program pemerintah dalam meningkatkan produksi pangan lokal, mengurangi ketergantungan impor, serta memperkuat stabilitas harga dan persediaan beras (S1, S3, O3).

b. Strategi S-T (*Using Strengths to Counter Threats*)

ST1: Pengembangan Alat Pendukung Keputusan untuk Pengadaan Impor
Mengembangkan *Decision Support Tool* (DST) dengan menggunakan input dari model prediksi untuk membantu menentukan kebijakan impor beras. DST akan membantu mengidentifikasi jumlah impor yang optimal berdasarkan analisis kebutuhan dan persediaan beras dalam negeri, sehingga mengurangi risiko ketidakstabilan persediaan dan harga beras (S2, T1).

ST2: Optimalisasi Manajemen Krisis dengan Prediksi Produksi (Hasil Panen)

Menerapkan model prediksi hasil panen untuk mempersiapkan dan merespons secara efektif dampak perubahan iklim pada produksi beras. Model prediksi ini dimanfaatkan untuk meningkatkan ketahanan pangan dan kesiapan dalam menghadapi fluktuasi hasil panen yang tidak terduga (S3, T2).

c. Strategi W-O (*Minimizing Weaknesses by Seizing Opportunities*)

WO1: Mengadopsi Teknologi *Blockchain* untuk Meningkatkan Kepercayaan dan Efisiensi

Mengadopsi teknologi *blockchain* untuk mengatasi keterbatasan teknologi saat ini, seperti sistem yang kurang transparan. *Blockchain* akan meningkatkan kepercayaan dan kredibilitas dalam operasi Perum BULOG, memfasilitasi transaksi yang lebih aman dan audit yang dapat diverifikasi (W1, W4, O1).

WO2: Peningkatan Koordinasi Antar Lembaga Melalui Sistem Terintegrasi

Membangun sistem terintegrasi yang melibatkan semua lembaga terkait untuk mengatasi masalah kurangnya koordinasi dan perbedaan data, serta memfasilitasi pengambilan keputusan yang lebih cepat dan akurat (W5, O2).

d. Strategi W-T (*Minimizing Weaknesses and Avoiding Threats*)

WT1: Mengoptimalkan Kebijakan Impor dengan Model Prediksi

Menggunakan model prediksi untuk mengurangi ketergantungan pada impor dengan mengidentifikasi jumlah produksi (hasil panen) dalam negeri yang dapat memenuhi permintaan, sekaligus mengetahui perlu atau tidaknya dilakukan impor (W2, W3, T1).

WT2: Peningkatan Manajemen Persediaan Melalui Analisis Tingkat Lanjut

Meningkatkan sistem manajemen persediaan dengan model prediksi agar lebih responsif terhadap perubahan permintaan dan kondisi darurat, serta mengurangi risiko kekurangan persediaan dan mengatasi fluktuasi harga (W3, T2).

Setelah melakukan analisis SWOT dan merumuskan strategi, maka hasilnya akan dirangkum dalam bentuk matriks SWOT yang terlampir pada Gambar 3.5.

<div> Faktor Internal </div> <div> Faktor Eksternal </div>	Kekuatan (S) <ol style="list-style-type: none"> 1. Dukungan Pemerintah (S1) 2. Infrastruktur Logistik yang Memadai (S2) 3. Pengalaman dan Keahlian (S3) 	Kelemahan (W) <ol style="list-style-type: none"> 1. Ketidakmampuan untuk Bersaing dengan Sektor Swasta (W1) 2. Ketergantungan pada Impor (W2) 3. Manajemen Persediaan yang Tidak Efisien (W3) 4. Keterbatasan Teknologi dan Transparansi (W4) 5. Kurangnya Koordinasi Antar Lembaga (W5)
	Strategi S-O <ol style="list-style-type: none"> 1. Pemanfaatan Teknologi <i>Blockchain</i> untuk Transparansi dan Efisiensi (S1, O1) 2. Pengembangan Sistem Terintegrasi (S2, O2) 3. Pengembangan Model Prediksi Permintaan dan Produksi (Hasil Panen) (S3, O1, O2) 	Strategi W-O <ol style="list-style-type: none"> 1. Mengadopsi Teknologi <i>Blockchain</i> untuk Meningkatkan Kepercayaan dan Efisiensi (W1, W4, O1) 2. Peningkatan Koordinasi Antar Lembaga Melalui Sistem Terintegrasi (W5, O2)
Ancaman (T) <ol style="list-style-type: none"> 1. Ketidakpastian Harga dan Persediaan Global (T1) 2. Dampak Perubahan Iklim (T2) 	Strategi S-T <ol style="list-style-type: none"> 1. Pengembangan Alat Pendukung Keputusan untuk Pengadaan Impor (S2, T1) 2. Optimalisasi Manajemen Krisis dengan Prediksi Produksi (Hasil Panen) (S3, T2) 	Strategi W-T <ol style="list-style-type: none"> 1. Mengoptimalkan Kebijakan Impor dengan Model Prediksi (W2, W3, T1) 2. Peningkatan Manajemen Persediaan Melalui Analisis Tingkat Lanjut (W3, T2)

Gambar 3. 5 Matriks SWOT

Berdasarkan analisis SWOT yang dilakukan terhadap Perum BULOG, telah diidentifikasi beberapa strategi penting yang menjadi fokus penelitian ini. Strategi tersebut meliputi pengembangan model prediksi permintaan dan produksi, optimasi manajemen krisis dengan prediksi produksi, pembuatan alat pendukung keputusan untuk pengadaan, optimasi kebijakan impor dengan model prediksi, dan peningkatan manajemen persediaan melalui analisis lanjutan. Strategi-strategi ini bertujuan untuk meningkatkan kinerja Perum BULOG dalam manajemen Cadangan Beras Pemerintah (CBP) dan dapat memberikan rekomendasi berbasis data untuk mendukung kebijakan pengadaan cadangan beras.

3.3.1 Regulasi Terkait Manajemen Cadangan Beras Pemerintah

Pemerintah telah menetapkan regulasi dalam manajemen cadangan beras untuk menjamin ketahanan pangan dan menangani potensi keadaan darurat.

Regulasi-regulasi ini bertujuan untuk menjamin ketersediaan beras selama keadaan darurat dan memanfaatkan cadangan beras yang dikelola pemerintah. Berikut adalah regulasi-regulasi utama yang berkaitan dengan cadangan beras pemerintah.

1. Peraturan Presiden Nomor 125 Tahun 2022 tentang Penyelenggaraan Cadangan Pangan Pemerintah

Pasal 2 Ayat (1): Pemerintah menetapkan jenis dan jumlah Cadangan Pangan Pemerintah (CPP) untuk menjamin ketersediaan pangan di seluruh wilayah Indonesia.

Pasal 4 Ayat (1): Penugasan kepada badan usaha milik negara untuk mengelola CPP, termasuk Perum BULOG dan/atau BUMN Pangan.

Pasal 8 Ayat (1): Pendanaan untuk penyelenggaraan CPP bersumber dari APBN dan/atau sumber pendanaan lain yang sah dan tidak mengikat.

2. Peraturan Badan Pangan Nasional Nomor 12 Tahun 2022 tentang Penyelenggaraan Cadangan Beras Pemerintah

Pasal 2: Penetapan jumlah CBP dilakukan dengan mempertimbangkan:

- a. Produksi beras dan/atau gabah secara nasional
- b. Penanggulangan keadaan darurat dan kerawanan pangan
- c. Pengendalian dan stabilisasi harga dan pasokan beras dan/atau gabah pada tingkat produsen dan konsumen
- d. Pelaksanaan perjanjian internasional dan bantuan pangan kerja sama internasional
- e. Angka kecukupan gizi yang dianjurkan

Pasal 3 Ayat (1): Penetapan jumlah CBP sebagaimana dimaksud dalam Pasal 2 ditetapkan oleh Kepala Badan.

Pasal 3 Ayat (4): Penetapan jumlah CBP sebagaimana dimaksud pada ayat (1) dilakukan paling sedikit (satu) kali dalam 1 (satu) tahaun.

Pasal 4 Ayat (3): Target pengadaan CBP terdiri atas volume pengadaa dalam negeri dan/atau pengadaan luar negeri.

Pasal 5 Ayat (1): Penyelenggaraan CBP dilakukan melalui pengadaan, pengelolaan, dan penyaluran.

Pasal 5 Ayat (2): Penyelenggaraan CBP sebagaimana dimaksud pada ayat (1) dilaksanakan oleh Badan Pangan Nasional melalui penugasan kepada Perum BULOG.

Pasal 10 Ayat (1): Dalam hal pengadaan CBP melalui produksi dalam negeri tidak mencukupi untuk pemenuhan cadangan, menjaga stabilitas harga dalam negeri, dan/atau memenuhi kebutuhan Pemerintah, dapat dilakukan pengadaan CBP dari luar negeri dengan tetap menjaga kepentingan produsen dan konsumen dalam negeri.

3. Peraturan Menteri Perdagangan Nomor 127 Tahun 2018 tentang Pengelolaan Cadangan Beras Pemerintah

Pasal 2 Ayat (1): Pengelolaan cadangan beras pemerintah untuk menjamin ketersediaan pasokan dan stabilisasi harga.

Pasal 5 Ayat (1): Penugasan kepada Perum BULOG untuk melakukan pengadaan beras dari produksi dalam negeri berdasarkan Harga Pembelian Pemerintah (HPP).

4. Peraturan Menteri Koordinator Bidang Perekonomian Republik Indonesia Nomor 5 Tahun 2018 tentang Koordinasi Pengelolaan Cadangan Beras Pemerintah untuk Stabilisasi Harga

Pasal 1 Ayat (3): Rapat koordinasi adalah rapat yang dipimpin oleh Menteri Koordinator Bidang Perekonomian.

Pasal 2 Ayat (2): Rapat koordinasi sebagaimana dimaksud pada ayat (1) paling sedikit melibatkan Menteri Pertanian, Menteri Perdagangan, Menteri Keuangan, Menteri Badan Usaha Milik Negara, dan Direktur Utama Perum BULOG.

5. Keputusan Kepala Badan Pangan Nasional Nomor 379.1/TS.03.03/K/11/2023:

Menetapkan jumlah persediaan minimal berbagai komoditas pangan yang harus dimiliki pemerintah hingga akhir tahun 2024. Untuk beras, jumlah minimal yang harus dimiliki adalah 2,4 juta ton dengan persediaan akhir tahun minimal 1,2 juta ton.

3.3.2 Parameter yang Dipertimbangkan dalam Menentukan Jumlah CBP

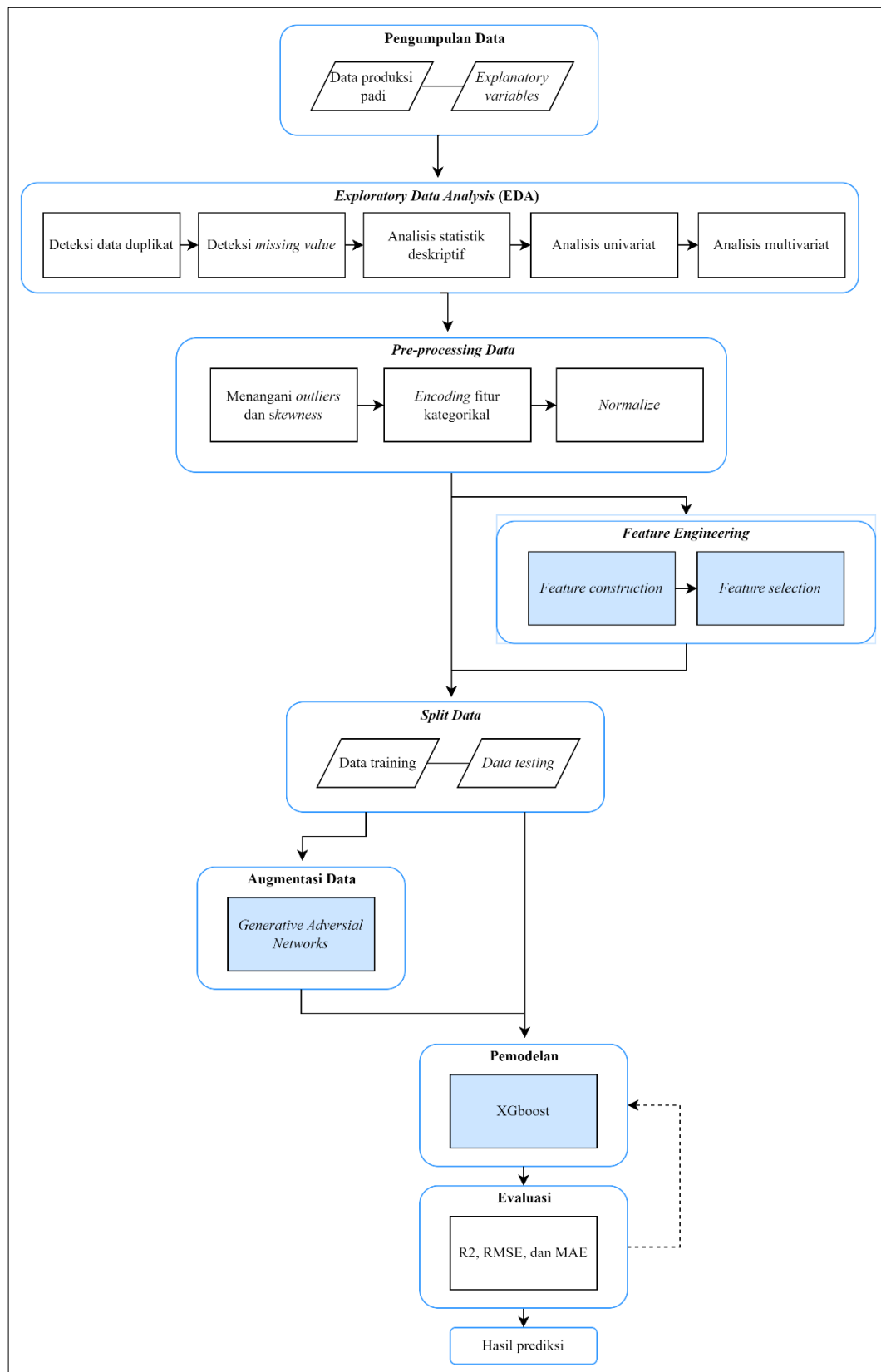
Berdasarkan Pasal 2 dari Peraturan Badan Pangan Nasional Nomor 12 Tahun 2022, terdapat beberapa parameter kunci yang dipertimbangkan dalam penetapan jumlah CBP. Berikut adalah parameter-parameter tersebut:

1. Produksi beras dan/atau gabah secara nasional, merupakan total hasil produksi beras atau gabah di seluruh negeri.
2. Penanggulangan keadaan darurat dan kerawanan pangan, merupakan potensi dan realitas situasi darurat atau krisis pangan yang memerlukan penggunaan cadangan.
3. Pengendalian dan stabilisasi harga dan pasokan beras dan/atau gabah, merupakan kemampuan cadangan untuk mengontrol fluktuasi harga dan pasokan beras di pasar.
4. Pelaksanaan perjanjian internasional dan bantuan pangan kerja sama internasional, merupakan kewajiban dan komitmen internasional yang mempengaruhi jumlah beras yang perlu disimpan.
5. Angka kecukupan gizi yang dianjurkan, merupakan kebutuhan gizi yang harus dipenuhi melalui konsumsi beras oleh populasi.

3.4. Pengembangan Model Prediksi Produksi (Hasil Panen) Beras

Gambar 3.6 menggambarkan delapan tahapan yang dijalani dalam pengembangan model prediksi untuk produksi (hasil panen) beras pada penelitian ini. Setiap tahap dirancang untuk memastikan keakuratan dan efektivitas model dalam memprediksi hasil panen.

Tahap pertama adalah pengumpulan data, yang meliputi data historis produksi padi dan data *explanatory variable*. Tahap kedua adalah explorasi data (EDA), yang meliputi penanganan data duplikat, penanganan nilai yang hilang (*missing value*), analisis statistik deskriptif, analisis univariat, dan analisis multivariat.



Gambar 3. 6 Tahapan penelitian prediksi permintaan yang diusulkan

Tahap ketiga adalah *preprocessing data* (pra-pemrosesan), yang meliputi penanganan *outlier* dan *skewness*, pengkodean (*encoding*) fitur kategorikal, serta normalisasi (*normalize*). Tahap keempat adalah *feature engineering* (rekayasa fitur), yang melibatkan konstruksi fitur dan seleksi fitur. Tahap kelima adalah membagi data (*split data*) menjadi data pelatihan (*training*) dan data pengujian (*testing*).

Tahap keenam adalah augmentasi data. Tahap ketujuh adalah pemodelan menggunakan algoritma XGBoost, yang dilatih menggunakan *dataset* pelatihan. Tahap kedelapan adalah evaluasi hasil, yang meliputi pengukuran performa model yang dapat dilihat dari nilai R^2 , *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE). Jika performa model perlu ditingkatkan, *hyperparameter* dapat disesuaikan kembali di tahap ketujuh. Setelah performa terbaik tercapai, model dapat diintegrasikan ke dalam model *decision support system*.

3.4.1 Pengumpulan Data

Pengembangan model prediksi produksi (hasil panen) ini akan menggunakan data yang dikumpulkan dari *website* Badan Pusat Statistik (BPS) dan BMKG. *Dataset* ini berisi variabel provinsi, tahun, produksi, luas panen, curah hujan, kelembapan, dan suhu rata-rata untuk periode 1993-2020. Penjelasan mengenai masing-masing variabel dapat dilihat di Tabel 3.2 dan sampel dari *dataset* ini dapat dilihat di Lampiran 1.

Tabel 3. 2 Deskripsi Variabel

Variabel	Deskripsi	Tipe Data
Provinsi	Informasi mengenai nama provinsi yang menghasilkan padi/beras	<i>Categorical</i>
Tahun	Tahun produksi (panen)	<i>Date</i>
Produksi	Jumlah hasil produksi (hasil panen) dalam satu tahun	<i>Numerical</i>
Luas Panen	Ukuran luas lahan yang dipanen di provinsi tersebut	<i>Numerical</i>
Curah Hujan	Curah hujan yang diterima provinsi tersebut dalam satu tahun	<i>Numerical</i>
Kelembapan	Tingkat kelembapan udara rata-rata di provinsi tersebut dalam satu tahun	<i>Numerical</i>
Suhu Rata-rata	Suhu udara rata-rata di provinsi tersebut dalam satu tahun	<i>Numerical</i>

Dataset yang digunakan dalam penelitian ini mencakup berbagai informasi yang berkaitan dengan hasil panen dari berbagai provinsi. Setiap baris mencakup beberapa variabel yang penting untuk analisis prediksi produksi (hasil panen).

Variabel 'Provinsi' menunjukkan nama provinsi di Indonesia di mana produksi padi/beras diukur, sementara variabel 'Tahun' menunjukkan tahun pengumpulan data produksi padi/beras. Variabel 'Produksi' menunjukkan jumlah total produksi (panen) beras dalam ton yang diproduksi di provinsi tersebut pada tahun tertentu. Variabel 'Luas Panen' menunjukkan luas lahan pertanian di provinsi tersebut yang ditanami padi dan dipanen untuk menghasilkan beras dalam satuan hektar.

Variabel 'Curah Hujan' menunjukkan rata-rata curah hujan dalam milimeter yang diterima oleh provinsi tersebut pada tahun tersebut, sedangkan variabel 'Kelembaban' menampilkan tingkat kelembaban udara rata-rata di provinsi tersebut dalam persentase. Terakhir, variabel 'Suhu Rata-rata' menunjukkan suhu udara rata-rata di provinsi tersebut dalam derajat *Celcius*, yang diukur selama satu tahun.

3.4.2 *Exploratory Data Analysis (EDA)*

Exploratory Data Analysis (EDA) adalah tahapan penting dalam pemodelan prediksi. Hal ini melibatkan analisis data historis secara menyeluruh, menggunakan statistik deskriptif, visualisasi data univariat dan multivariat seperti plot untuk mengidentifikasi pola, tren, dan korelasi. EDA memainkan peran penting dalam mengelola kualitas data. EDA juga memandu seleksi dan rekayasa fitur, yang sangat penting untuk mengembangkan model prediksi yang akurat (Santhoshkumar & Vanila, 2024).

EDA memungkinkan prediksi produksi (hasil panen) yang tepat dengan membangun fondasi yang kuat untuk pemodelan tingkat lanjut. Penelitian ini menggunakan Python untuk EDA.

a. Deteksi data duplikat

Sangat penting untuk mendeteksi dan menghapus data duplikat untuk menjaga keakuratan analisis dan pengembangan model prediktif. Data duplikat dapat menyebabkan hasil yang bias dan tidak dapat diandalkan (Koumarelas, Jiang, & Naumann, 2020). Pada penelitian ini, tidak ditemukan adanya data duplikat dalam *dataset* yang digunakan.

b. Deteksi nilai yang hilang (*missing value*)

Proses deteksi dan penanganan data hilang (*missing value*) dilakukan untuk memastikan keberlanjutan dalam data deret waktu (*time series*). Hal ini penting untuk prediksi produksi (hasil panen) karena memungkinkan model untuk menangkap pola dan *trend* dengan akurat, atau dengan kata lain tanpa melakukan pengisian nilai nol ini, analisis yang dilakukan kemungkinan besar akan menghasilkan hasil yang tidak akurat atau bias (Vallés-Pérez et al., 2022). Pada penelitian ini tidak dilakukan proses *replace missing value* karena *dataset* yang digunakan tidak memuat *missing value* seperti yang terlihat pada Tabel 3.3.

Tabel 3. 3 Missing value

Variabel	Missing value (%)
Provinsi	0
Tahun	0
Produksi	0
Luas Panen	0
Curah Hujan	0
Kelembapan	0
Suhu Rata-rata	0

c. Analisis statistik deskriptif

Analisis statistik deskriptif dilakukan untuk mendapatkan wawasan awal mengenai properti statistik dari *dataset*. Hal ini termasuk penghitungan rata-rata, median, standar deviasi, dan lainnya yang memberikan gambaran umum tentang bentuk dan distribusi data. Tabel 3.4 adalah ringkasan analisis statistik deskriptif *dataset* hasil panen.

Tabel 3. 4 Ringkasan Statistik

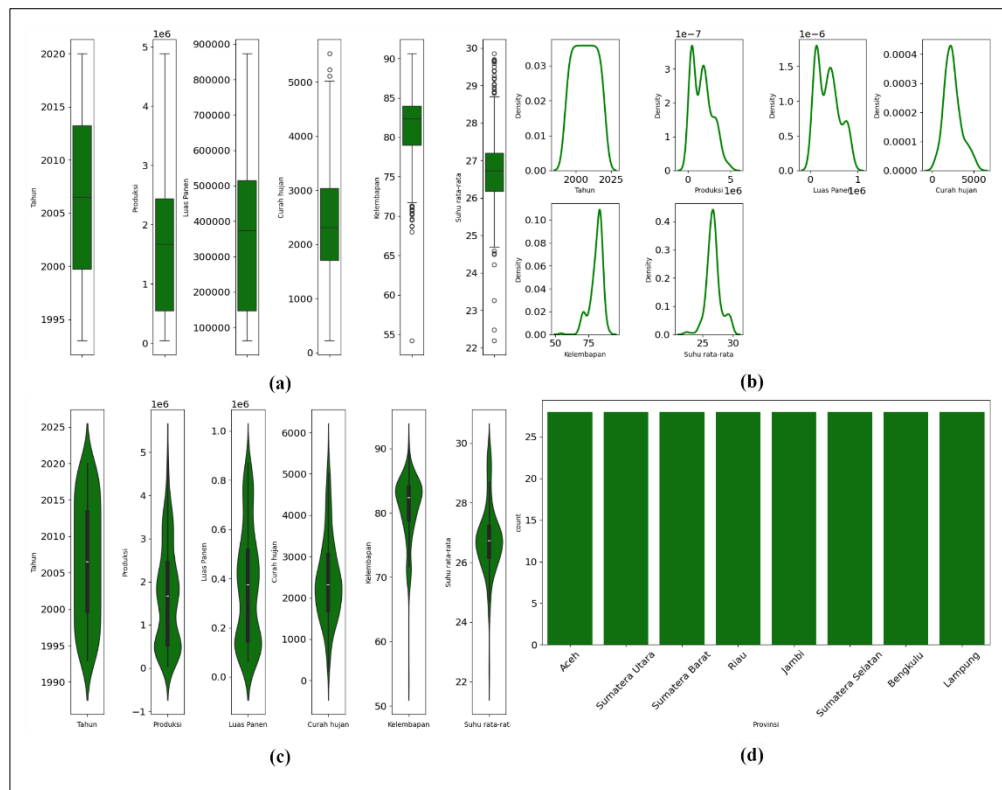
	Tahun	Produksi	Luas Panen	Curah hujan	Kelembapan	Suhu rata-rata
<i>Count</i>	224	224	224	224	224	224
<i>Mean</i>	2006,5	1,679701e+06	374349,96692	2452.490759	80,948705	26,801964
<i>Std</i>	8,095838	1,161387e+06	232751,161987	1031,972625	4,87868	1,197041
<i>Min</i>	1993	4,293800e+04	63142,04	222,5	54,2	22,19
25%	1999,75	5,488570e+05	146919,5	1703,525	78,975	26,1775
50%	2006,5	1,667773e+06	373551,5	2315,7	82,375	26,73
75%	2013,25	2,436851e+06	514570,25	3039,7	84	27,2
<i>Max</i>	2020	4,881089e+06	872737	5522	90,6	29,85

Berdasarkan Tabel 3.4 diketahui bahwa distribusi data cenderung normal karena perbedaan antara nilai *mean* dan *median* tidak terlalu besar. Produksi dan luas panen menunjukkan variasi yang besar dengan adanya *outlier*, yang

ditunjukkan oleh perbedaan antara *mean* dan *median* yang cukup besar. Curah hujan menunjukkan rentang nilai yang luas, namun perbedaan antara *mean* dan *median* relatif kecil. Kelembapan menunjukkan variasi yang cukup stabil dengan perbedaan antara *mean* dan *median* yang tidak besar. Suhu rata-rata menunjukkan data yang konsisten dengan perbedaan minimal antara *mean* dan *median*.

d. Analisis univariat

Analisis univariat dilakukan untuk setiap variabel dalam *dataset*, menggunakan teknik *box plots*, *density plot*, *violin plots*, dan *count plot*. Hal ini membantu dalam memvisualisasikan distribusi, mendeteksi *outliers*, dan memahami variasi dalam data (Shabdin, Ya'acob, & Sjarif, 2020).



Gambar 3.7 (a) Box Plots, (b) Density Plot, (c) Violin Plots, dan (d) Count Plot

Box Plot memberikan visualisasi distribusi berbasis waktu dari berbagai variabel seperti produksi, area panen, curah hujan, kelembapan, dan suhu rata-rata. Produksi dan luas panen menunjukkan tren yang meningkat dari waktu ke waktu, sementara curah hujan, kelembapan, dan suhu rata-rata

relatif stabil, dengan *outlier* yang terlihat pada variabel curah hujan, kelembaban, dan suhu rata-rata.

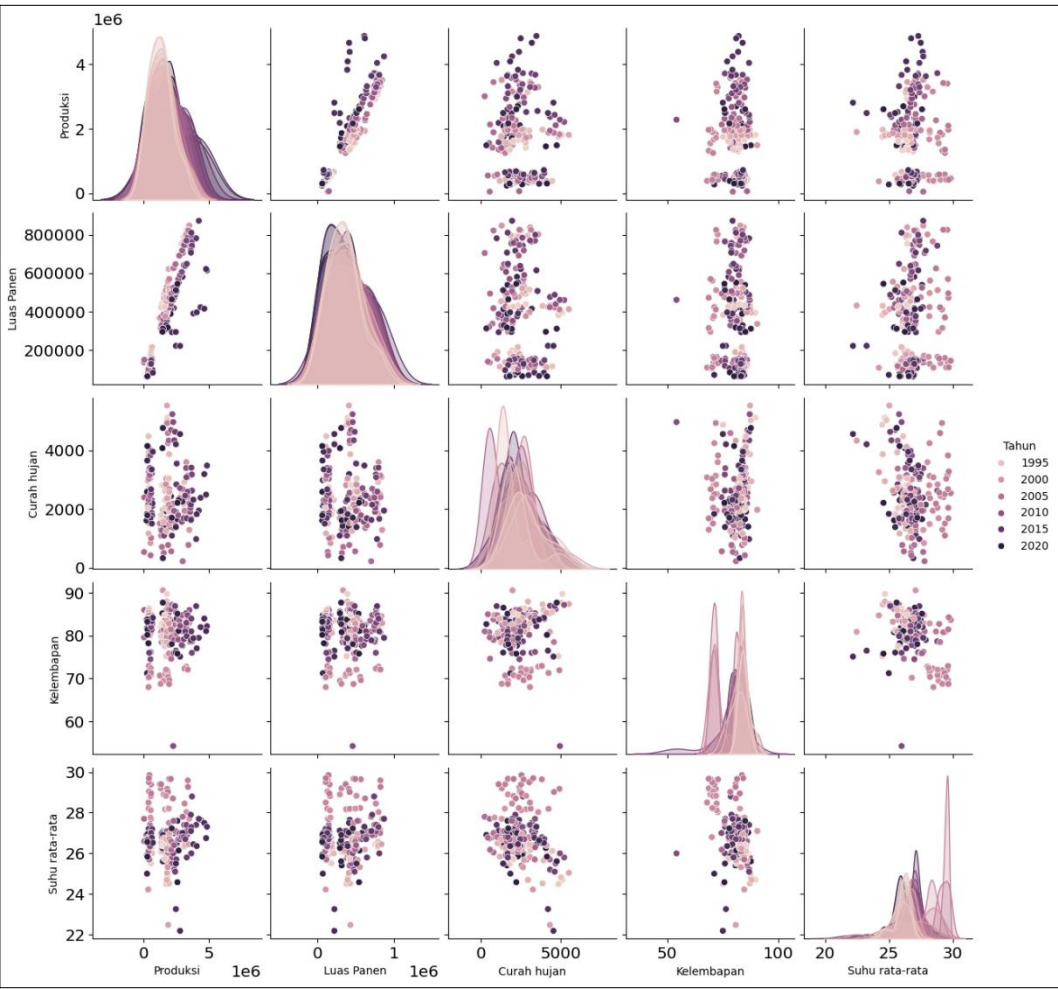
Density Plot menggambarkan distribusi probabilitas dari setiap variabel. Distribusi produksi dan luas panen menunjukkan pola miring ke kanan (*right-skewed*), yang mengindikasikan adanya konsentrasi nilai di batas bawah kisaran dan beberapa nilai yang sangat tinggi. Distribusi curah hujan dan suhu terlihat lebih simetris, sementara distribusi kelembaban menunjukkan bukti bimodalitas dengan dua puncak yang berbeda.

Hal serupa juga terjadi pada *violin plot* yang memberikan wawasan tentang distribusi probabilitas data. Variabel produksi dan luas panen menunjukkan distribusi yang condong ke kanan. Sebaliknya, variabel curah hujan dan suhu menunjukkan distribusi yang lebih simetris, sedangkan variabel kelembaban menunjukkan pola bimodal. Terakhir, *count plot* menunjukkan jumlah pengamatan, terlihat bahwa semua provinsi memiliki jumlah observasi yang sama.

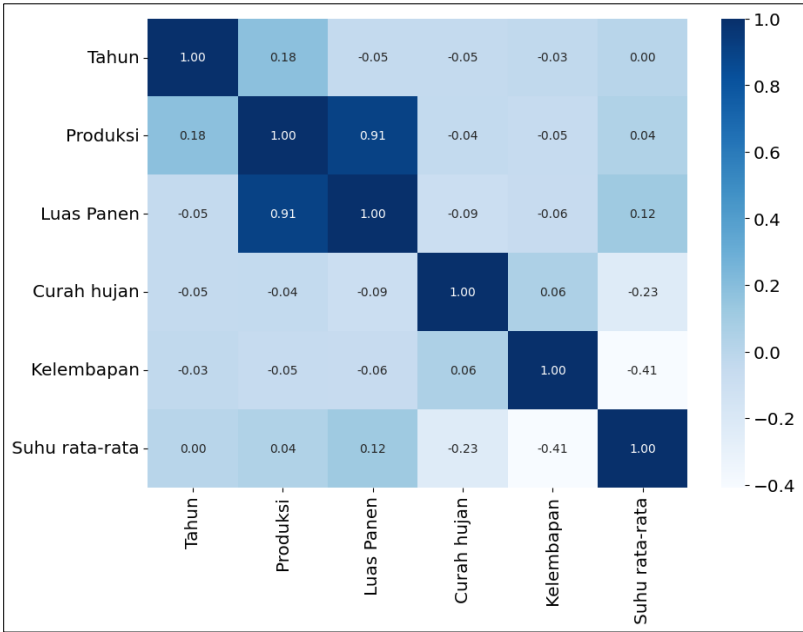
e. Analisis multivariat

Analisis multivariat melibatkan visualisasi data multivariat menggunakan *pair plot* dan *heatmap* korelasi untuk mengeksplorasi hubungan antar variabel. Visualisasi ini membantu mengidentifikasi pola dan hubungan yang mungkin tidak terlihat ketika menganalisis variabel secara terpisah (Ratilal & Reddy, 2023). Gambar 3.8 menunjukkan *pair plot* dan Gambar 3.9 menunjukkan *heatmap* korelasi.

Pair plot menggambarkan distribusi dan hubungan antara variabel, dengan fokus pada korelasi yang kuat antara produksi dan luas panen, yang mengindikasikan bahwa luas panen secara signifikan mempengaruhi tingkat produksi. Hubungan ini diperkuat oleh *heatmap* korelasi, yang menunjukkan koefisien korelasi sebesar 0,91, yang menandakan adanya korelasi yang kuat antara kedua variabel tersebut.



Gambar 3. 8 Pair Plot



Gambar 3. 9 Heatmap korelasi

3.4.3 *Preprocessing Data*

Tahap *preprocessing* dilakukan untuk mempersiapkan *dataset* untuk keperluan analisis *machine learning*. Berdasarkan hasil EDA yang telah dilakukan, maka tahapan *preprocessing* yang dilakukan meliputi penanganan *outlier* dan *skewness*, pengkodean (*encoding*) fitur kategorikal, serta normalisasi (*normalize*). Tahap ini sangat penting untuk menghasilkan model prediksi berbasis *machine learning* yang akurat dan efektif.

a. Penanganan *outliers* dan *skewness*

Berdasarkan analisis *box plots* dan *violin plots*, terlihat bahwa *dataset* yang digunakan mengandung banyak *outlier*, sehingga perlu menerapkan metode untuk mengatasinya. Di sisi lain, analisis *density plot* menunjukkan bahwa terdapat data yang *skew* atau miring pada *dataset* yang dapat mempengaruhi kinerja model. Untuk menangani keduanya, *transformasi log* akan digunakan dalam penelitian ini (Choi et al., 2022).

b. Pengkodean (*encoding*) fitur kategorikal

Untuk memproses variabel input kategorikal dengan cara yang dapat dipahami dan digunakan untuk analisis, maka digunakan representasi *label encoding*. Representasi ini mengubah variabel kategorikal menjadi representasi numerik berdasarkan penomoran kategori secara berurutan, di mana setiap kategori diberi nilai numerik yang unik (Dahouda & Joe, 2021). Pada penelitian ini *label encoding* diterapkan pada variabel 'Provinsi'.

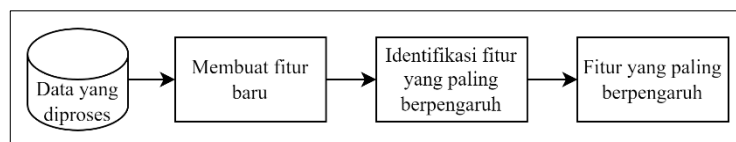
c. Normalisasi (*normalize*)

Pada penelitian ini, proses normalisasi data dilakukan dengan menggunakan metode *Interquartile Range* (IQR) atau *robust normalization*, yang melibatkan penghitungan rentang interkuartil data dan menggunakannya untuk menormalkan data. IQR adalah rentang antara kuartil pertama (Q1) dan kuartil ketiga (Q3), yang mewakili 50% bagian tengah data. Dengan menormalkan data, metode IQR dapat membantu meningkatkan akurasi prediksi (Vaitheeshwari & SathieshKumar, 2019).

3.4.4 *Features Engineering*

Feature engineering dilakukan untuk meningkatkan interpretabilitas model dengan membuat fitur-fitur yang lebih mudah dimengerti. Fitur-fitur yang terstruktur dengan baik dapat membantu dalam menjelaskan alasan dibalik prediksi yang dibuat oleh model. Proses *feature engineering* terdiri dari lima teknik utama, yaitu *feature improvement*, *feature construction*, *feature selection*, *feature extraction*, dan *feature learning* (Ozdemir, 2022).

Pada penelitian ini dilakukan dua perlakuan, dimana model dibangun dengan dan tanpa proses *feature engineering*. Hal ini dilakukan untuk membuktikan hasil penelitian Swaminathan and Venkitasubramony (2023) yang menyatakan bahwa penggunaan *feature engineering* sangat diperlukan untuk pemodelan prediksi untuk menghasilkan prediksi yang lebih akurat dan dapat diandalkan. Selain itu, penggunaan *features engineering*, keterlibatan *explanatory variabel* dan data yang lebih banyak akan menjadi salah satu fokus utama pada penelitian ini. Penelitian ini melibatkan *feature construction* dan *feature selection* seperti yang terlihat pada Gambar 3.10.



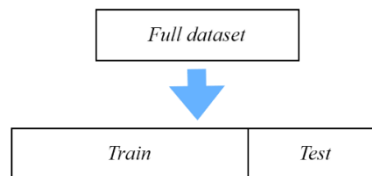
Gambar 3. 10 Flowchart feature engineering

Feature construction melibatkan proses pembuatan fitur baru dari data yang sudah ada. Hal ini melibatkan penggabungan data untuk membuat fitur yang lebih informatif dan relevan dengan masalah yang sedang dihadapi (Ozdemir, 2022). Pada penelitian ini dibuat fitur ‘Efisiensi Produksi’ dari ‘Produksi’ dan ‘Luas Panen’.

Seleksi fitur (*feature selection*) berperan penting dalam meningkatkan akurasi dalam model prediksi. Teknik ini melibatkan identifikasi fitur yang paling berpengaruh terhadap hasil panen. Pemilihan fitur yang tepat dapat memfasilitasi prediksi hasil panen padi/beras masa depan. Proses seleksi ini esensial untuk menyesuaikan model prediksi dengan dinamika hasil panen yang berubah-ubah (Bedi, 2023).

3.4.5 Pembagian Data (*Split Data*)

Pada proses pembuatan model *machine Learning*, data dapat dibagi menjadi dua set, yaitu data pelatihan (*training*) dan data pengujian (*testing*). Mesin dilatih dengan data pelatihan dan diuji dengan data pengujian untuk mengevaluasi keandalan dan performanya dalam mencapai tujuan penelitian. Untuk menentukan rasio pembagian data yang optimal, penelitian ini menggunakan tiga rasio yang umum digunakan yaitu 70:30, 80:20, dan 90:10 untuk data pelatihan dan pengujian (Grigorev, 2021; Nguyen, Ly, Ho, Al-Ansari, Le, Tran, Prakash, & Pham, 2021).



Gambar 3. 11 Membagi data menjadi pelatihan dan pengujian (Grigorev, 2021)

3.4.6 Augmentasi Data

Augmentasi data (*Data Augmentation/ DA*) adalah teknik yang digunakan dalam *machine learning* dan *deep learning* untuk meningkatkan ukuran *dataset* pelatihan dengan membuat versi modifikasi dari data yang ada. Proses ini sangat penting untuk meningkatkan performa model, terutama dalam skenario di mana jumlah data pelatihan yang tersedia terbatas (Bayer et al., 2022; Moreno-Barea et al., 2020; Onishi & Meguro, 2023). *Generative AI* (GenAI) akan dimanfaatkan untuk menambah data *training* dalam penelitian ini, khususnya dengan menggunakan *Generative Adversarial Networks* (GAN).

3.4.7 Pemodelan dengan *GridSearchCV*

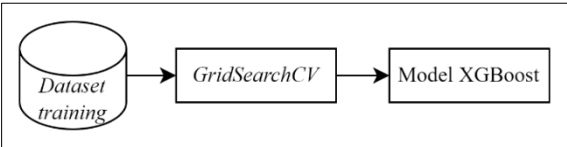
XGBoost merupakan model *machine learning* canggih yang membangun ansambel pohon keputusan secara berurutan, mengoptimalkan *loss function* dan menangani ketidakseimbangan data, *missing value*, serta *overfitting*. Prediksi produksi beras menggunakan algoritma XGBoost membutuhkan penyetingan

beberapa *hyperparameter* sebagai penunjang pemodelan. Proses untuk memilih *hyperparameter* terbaik dalam model *machine learning* untuk meningkatkan kinerja dikenal sebagai penyetelan *hyperparameter* (*tuning*) (Soleymani & Mohammadzadeh, 2023). Model prediksi pada penelitian ini menggunakan parameter *n_estimator*, *max_depth*, dan *learning_rate* yang diadopsi dari penelitian Ørebæk and Geitle (2021). *Range* parameter yang dipertimbangkan pada penelitian ini dapat dilihat di Tabel 3.5.

**Tabel 3. 5 Parameter Yang Dipilih Dan Nilainya
(Ørebæk & Geitle, 2021)**

<i>Hyperparameter</i>	<i>Value Range</i>
<i>Learning_rate</i>	0,1-1
<i>n_estimator</i>	100-250
<i>max_depth</i>	1- 14

Uji coba satu persatu kombinasi yang telah diatur dan ditambah dengan *GridSearchCV* akan mampu menentukan kombinasi mana yang menghasilkan performa terbaiknya. Pada konteks *machine learning*, *10-fold cross-validation* seringkali lebih disukai karena memberikan estimasi yang lebih dapat diandalkan untuk kinerja model, terutama ketika ukuran *dataset* terbatas (Marcot & Hanea, 2021). Untuk itu, pada penelitian ini nilai *CV* disetting 10 yang mana setiap kombinasi model dan parameter divalidasi sebanyak 10 kali dan data akan dibagi menjadi 10 bagian sama besar secara acak, yang mana 9 bagian akan digunakan untuk *training* dan 1 bagian untuk *validasi*. Dari hasil uji coba secara acak, didapatkan kombinasi terbaik yaitu *n_estimator* 115, *learning_rate* 0,28, serta *max_depth* 1. Kombinasi ini yang dimasukan ke dalam algoritma XGBoost seperti yang terlihat di Gambar 3.12.



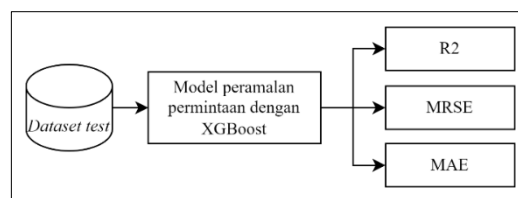
Gambar 3. 12 Flowchart pemodelan

Jika ingin meningkatkan performa model, maka dapat dilakukan penyetelan ulang atau tuning model. Proses ini melibatkan penyesuaian *hyperparameter* dan konfigurasi model untuk mencapai hasil yang lebih baik. Dengan penyetelan ulang, dapat mengeksplorasi kombinasi *hyperparameter* yang berbeda atau menggunakan teknik optimasi lanjut, seperti *GridserchCV*

(Liyanage, Basnayake, Gamage, Prabhashi, Kasthuriarachchi, & Abeywardhana, 2023). Meskipun memerlukan upaya yang cukup besar, penyetelan ulang dapat menghasilkan peningkatan signifikan dalam performa prediksi. Untuk itu, dalam konteks penelitian ini, akan terus dilakukan juga evaluasi dan penelitian pada tahap pemodelan akan terus dilakukan untuk memastikan bahwa model sesuai dengan tujuan penelitian.

3.4.8 Evaluasi Model

Setelah model dilatih, model tersebut menjalani tahap evaluasi di mana tingkat kesalahan diuji untuk memastikan model berfungsi dengan baik. Penelitian ini menggunakan *R-squared* (R^2), *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE) untuk menguji tingkat kesalahan. R^2 mengukur proporsi varians dalam variabel dan berkisar antara 0 – 1, dimana semakin mendekati 1 maka semakin layak suatu model untuk digunakan. RMSE mengukur besarnya tingkat kesalahan prediksi, dimana semakin rendah nilainya (mendekati nol), maka hasil prediksi akan semakin akurat. Di sisi lain, MAE menghitung rata-rata kesalahan kuadrat antara nilai aktual dan nilai yang diramalkan, dengan nilai yang rendah atau mendekati nol menunjukkan kecocokan yang lebih baik antara hasil permintaan dan data aktual (Chicco et al., 2021). Gambar 3.13 adalah skema evaluasi model pada penelitian ini.



Gambar 3. 13 Flowchart evaluasi model

Pada tahap evaluasi, model juga diuji dengan dua rancangan yang berbeda, yaitu model tanpa *features engineering*, serta pada model dengan *features engineering*. Tabel 3.6 adalah hasil evaluasi model pada penelitian ini.

Tabel 3. 6 Hasil Evaluasi Model

Dataset	Kondisi tanpa <i>feature engineering</i>			Kondisi dengan <i>feature engineering</i>		
	R^2	RMSE	MAE	R^2	RMSE	MAE
70:30	0,773	0,441	0,176	0,933	0,209	0,115
80:20	0,795	0,418	0,155	0,956	0,163	0,099
90:10	0,976	0,111	0,079	0,976	0,109	0,076

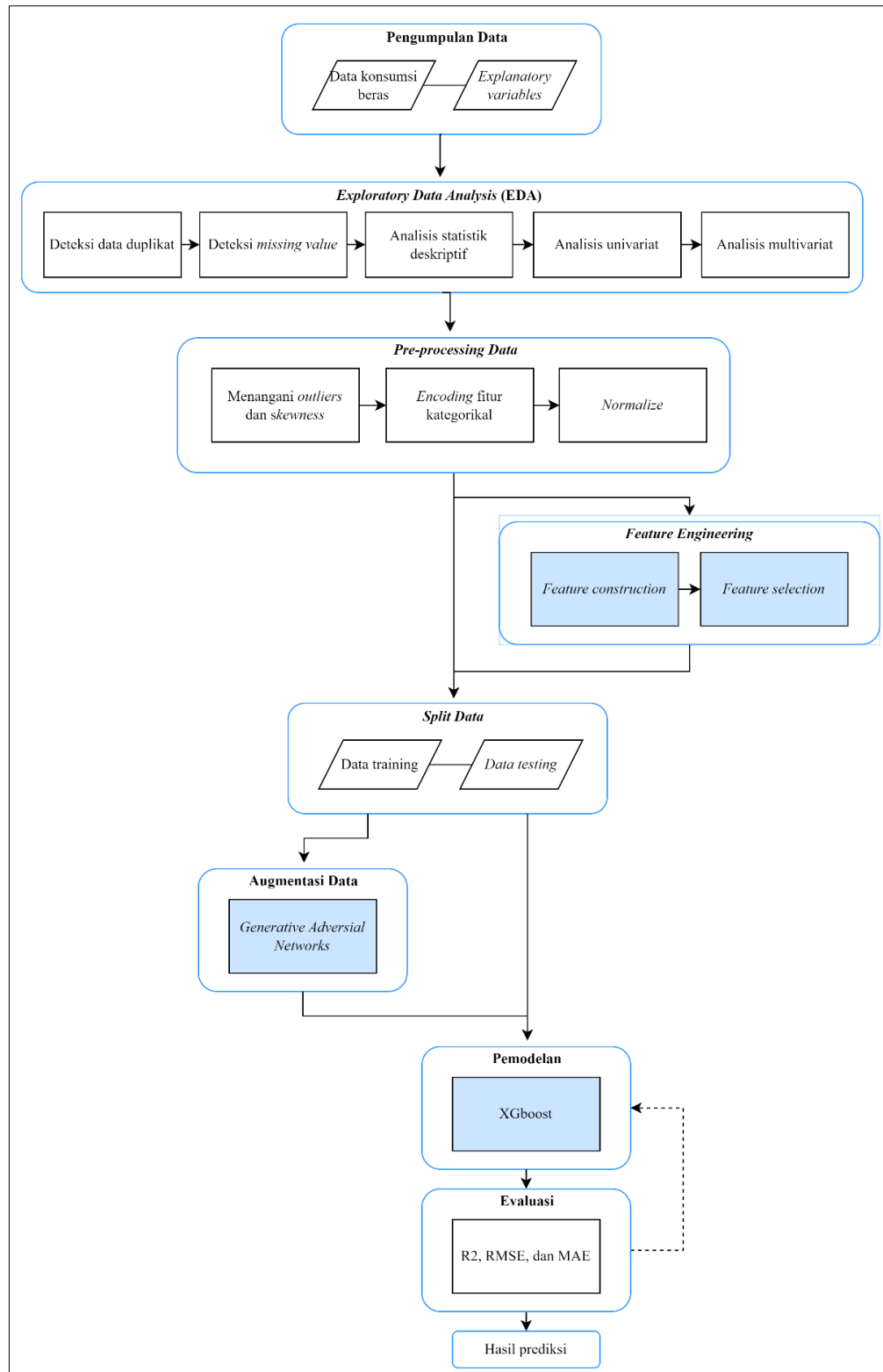
Hasil evaluasi menunjukkan pengaruh positif dari penggunaan *Feature Engineering* (FE) terhadap kinerja model *machine learning* dalam skema pembagian *dataset* yang berbeda (70:30, 80:20, dan 90:10). Khususnya, penggunaan *feature engineering* membantu meningkatkan nilai R^2 , yang menandakan peningkatan kemampuan model dalam menjelaskan variabilitas data yang diamati. Selanjutnya, nilai RMSE dan MAE menurun dengan implementasi FE, mengindikasikan bahwa kesalahan pada model prediksi menjadi lebih kecil dan prediksi menjadi lebih akurat secara umum. Analisis lebih lanjut pada perbedaan skema pembagian data menunjukkan bahwa model dengan proporsi data pelatihan yang lebih besar (90:10) menunjukkan hasil yang paling stabil dan akurat. Hal ini menyoroti pentingnya FE dalam meningkatkan efektivitas model dan menunjukkan keuntungan dari alokasi yang lebih besar pada data pelatihan dalam pengembangan model prediksi berbasis *machine learning*.

Nilai RMSE terendah yang diperoleh adalah 0,109, yang menunjukkan potensi model untuk menghasilkan prediksi produksi yang sangat akurat jika terus dikembangkan. Nosratabadi et al. (2021) mendapatkan nilai RMSE 33.575.595,74 untuk model prediksinya, hal ini menunjukkan bahwa model prediksi yang diusulkan memiliki potensi untuk menyaingi model yang ada dalam hal keakuratan hasil prediksi. Meskipun hal ini tidak mutlak karena model ini dibangun dengan algoritma dan data yang berbeda. Namun, potensi model ini menghasilkan prediksi yang sangat akurat masih terlihat. Berdasarkan hasil ini, penelitian ini akan fokus pada penambahan data pelatihan menggunakan *Generative Adversarial Network* (GAN) dan pengembangan lebih lanjut pada teknik *feature engineering* untuk meningkatkan kinerja prediksi.

3.5 Pengembangan Model Prediksi Permintaan

Model prediksi yang digunakan untuk prediksi permintaan adalah model yang sama dengan yang digunakan untuk memprediksi produksi atau hasil panen. Namun, beberapa penyesuaian dilakukan untuk memastikan bahwa model sesuai dengan karakteristik data permintaan beras, terutama pada tahapan *features engineering*, augmentasi data, dan pemodelan dengan *GridSearchCV*. Gambar

3.14 adalah tahapan pengembangan model prediksi permintaan pada penelitian ini.



Gambar 3. 14 Tahapan penelitian prediksi permintaan yang diusulkan

Pengembangan model prediksi permintaan beras ini terdiri dari delapan tahap. Tahap pertama adalah pengumpulan data, yang meliputi data historis konsumsi beras (Bashir & Yuliana, 2019) dan data *explanatory variable*, seperti harga beras (Iqbal, 2019; Sossou & Igue, 2019), pendapatan per kapita (Bashir & Yuliana, 2019; Yusuf et al., 2020), populasi (Bashir & Yuliana, 2019), rasio kemiskinan (Qian, Ito, & Zhao, 2020), dan peristiwa khusus seperti bulan ramadhan dan hari raya (Hasanah, 2020). Tahap kedua adalah eksplorasi data (EDA), yang meliputi penanganan data duplikat, penanganan nilai yang hilang (*missing value*), analisis statistik deskriptif, analisis univariat, dan analisis multivariat.

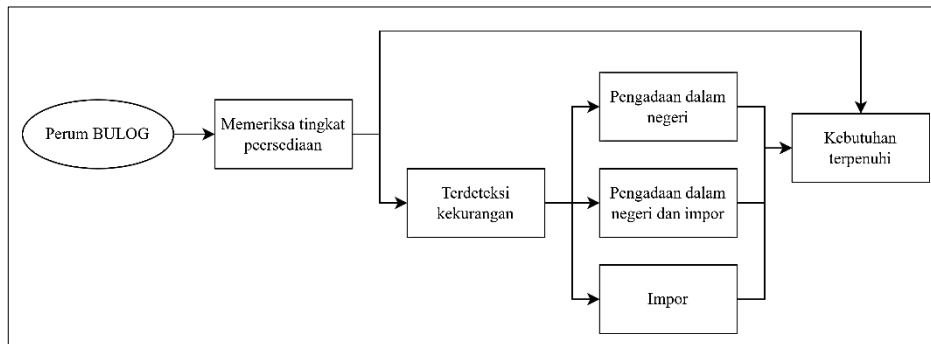
Tahap ketiga adalah *preprocessing data* (pra-pemrosesan), yang meliputi penanganan *outlier* dan *skewness*, pengkodean (*encoding*) fitur kategorikal, dan normalisasi (*normalize*). Tahap keempat adalah *feature engineering* (rekayasa fitur), yang melibatkan konstruksi fitur dan seleksi fitur. Tahap kelima adalah membagi data (*split data*) menjadi data pelatihan (*training*) dan data pengujian (*testing*).

Tahap keenam adalah augmentasi data. Tahap ketujuh adalah pemodelan menggunakan algoritma XGBoost, yang dilatih menggunakan *dataset* pelatihan. Tahap kedelapan adalah evaluasi hasil, yang meliputi pengukuran performa model yang dapat dilihat dari nilai R^2 , *Root Mean Squared Error* (RMSE), dan *Mean Absolute Error* (MAE). Jika performa model perlu ditingkatkan, *hyperparameter* dapat disesuaikan kembali di tahap ketujuh. Setelah performa terbaik tercapai, model dapat diintegrasikan ke dalam model *decision support system*.

3.6 *Prototype Decision Support System (DSS) untuk Pengadaan Cadangan Beras Pemerintah*

Decision Support System (DSS) pada penelitian ini akan digunakan untuk mendukung proses pengambilan keputusan internal terkait kebijakan pengadaan cadangan beras pemerintah. DSS ini akan mengintegrasikan data hasil prediksi produksi (hasil panen), permintaan, dan persediaan aktual untuk menentukan variabel keputusan pengadaan yang optimal. Alat ini memungkinkan Perum

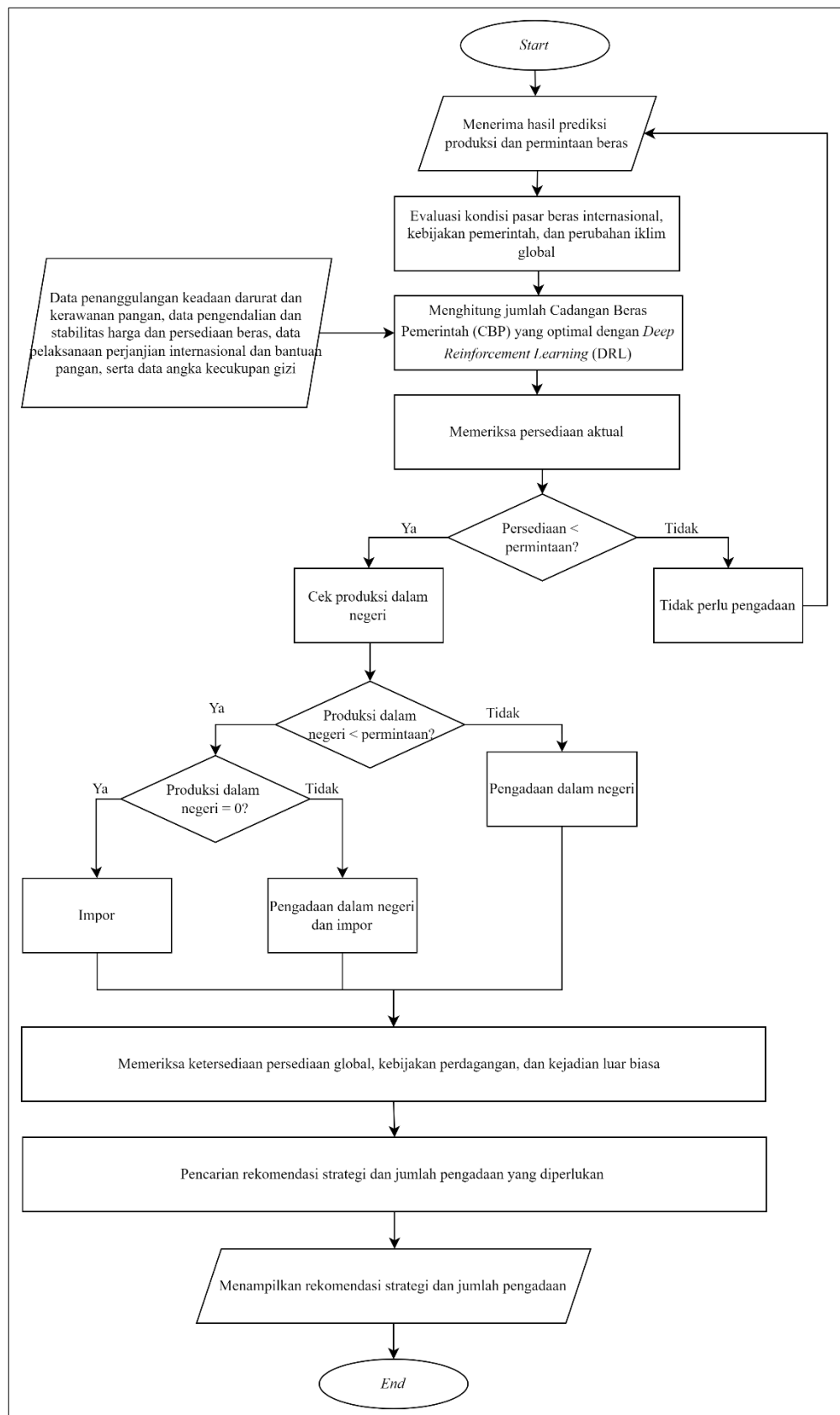
BULOG untuk merencanakan dan mengelola Cadangan Beras Pemerintah (CBP) secara efisien, sesuai dengan kebutuhan masyarakat. Gambar 3.15 adalah ilustrasi dari proses pemenuhan CBP.



Gambar 3. 15 Ilustrasi Pemenuhan CBP

Gambar 3.15 menggambarkan alur proses pemenuhan CBP yang dilakukan oleh Perum BULOG dan lembaga terkait. Proses ini dimulai dengan Perum BULOG memeriksa tingkat persediaan CBP. Jika ditemukan adanya kekurangan persediaan, maka langkah selanjutnya adalah menyusun strategi untuk mengatasi kekurangan tersebut. Terdapat beberapa opsi untuk mengatasi kekurangan tersebut, meliputi pengadaan dalam negeri, kombinasi antara pengadaan dalam negeri dan impor, atau impor penuh jika pengadaan dalam negeri tidak mencukupi.

Gambar 3.16 menguraikan proses pengambilan keputusan untuk mengelola persediaan CBP. Proses ini dimulai dengan penerimaan data tentang prediksi produksi (hasil panen) dalam negeri dan kebutuhan beras mendatang. Selanjutnya, evaluasi kondisi pasar beras internasional, kebijakan pemerintah, dan perubahan iklim global. Lalu dilakukan perhitungan terhadap jumlah Cadangan Beras Pemerintah (CBP) yang optimal, yang menentukan jumlah beras yang harus disimpan untuk kebutuhan darurat. Untuk mencapai optimasi ini, penelitian ini akan menggunakan metode *Deep Reinforcement Learning* (DRL). Berbeda dengan metode heuristik, DRL lebih kuat dengan hasil konvergensi yang stabil dan lebih cocok untuk masalah pengambilan keputusan (Z. Zhang, Zhang, & Qiu, 2019). Tahap berikutnya adalah memeriksa jumlah persediaan aktual.



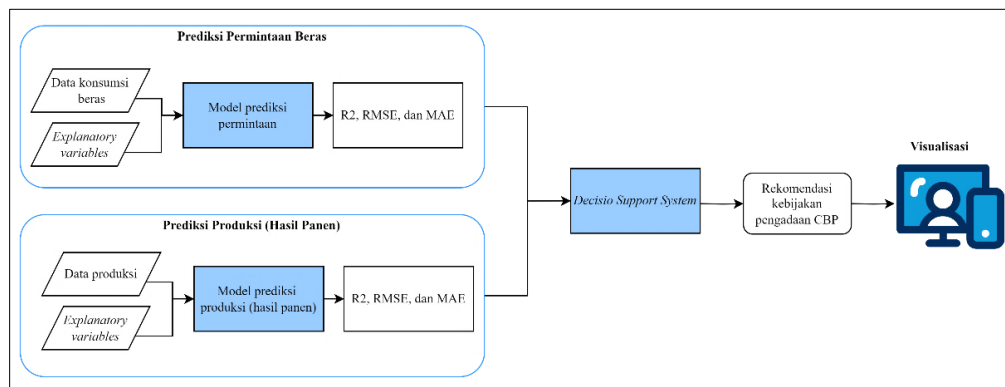
Gambar 3. 16 Diagram Alir Skenario Dasar Pengambilan Keputusan

Berdasarkan data persediaan aktual, dilakukan evaluasi apakah jumlah tersebut sudah mencukupi untuk memenuhi prediksi permintaan atau kebutuhan beras mendatang. Jika persediaan cukup, maka tidak diperlukan tindakan pengadaan tambahan. Sebaliknya, jika persediaan dinilai belum mencukupi kebutuhan, maka akan dilakukan penyerapan atau pengadaan dari produksi (hasil panen) dalam negeri. Pada kondisi di mana produksi (hasil panen) dalam negeri yang ada pada petani atau mitra kerja Perum BULOG sangat minim, maka akan dilakukan pengadaan dari sumber dalam negeri dan impor untuk mengisi kekurangan tersebut. Jika kondisi di mana produksi (hasil panen) dalam negeri kosong.

Penentuan jumlah pengadaan didasarkan pada penghitungan jumlah CBP yang diperlukan untuk memenuhi kekurangan antara kebutuhan dan persediaan yang tersedia dengan turut mempertimbangkan ketersediaan persediaan global, kebijakan perdagangan, dan kejadian luar biasa, seperti bencana alam atau gangguan politik yang bisa mempengaruhi pengadaan.

Output dari DSS ini adalah sebuah rekomendasi untuk strategi dan jumlah pengadaan beras. Rekomendasi ini dapat digunakan oleh Perum BULOG untuk membuat kebijakan pengadaan beras, khususnya dalam menentukan kebutuhan akan impor beras. Keseluruhan proses ini penting untuk memastikan ketersediaan CBP dan menghindari potensi krisis pangan.

DSS yang diusulkan adalah sebuah *platform* yang mengintegrasikan dua model prediksi dan proses optimasi untuk cadangan persediaan beras pemerintah. Sistem ini membantu Perum BULOG dan lembaga terkait dalam mengambil kebijakan strategis terkait pengadaan cadangan beras pemerintah. Hal ini untuk memastikan ketersediaan Cadangan Beras Pemerintah (CBP), sehingga mendukung stabilitas harga dan ketahanan pangan nasional. Gambar 3.17 adalah *prototype* DSS yang diusulkan.



Gambar 3. 17 *Prototype DSS yang Diusulkan*

3.7 Uji Coba

Uji coba sistem bertujuan untuk memvalidasi apakah sistem yang dikembangkan sesuai dengan tujuan awal dan layak untuk digunakan. Pada tahap ini, *prototype* sistem akan menjalani evaluasi secara menyeluruh untuk memastikan bahwa fungsionalitasnya telah memenuhi standar yang diharapkan dan untuk mengidentifikasi potensi *error* (Sutiah & Supriyono, 2021). Tahap uji coba pada penelitian ini akan menggunakan metode *black box testing*, yang dapat menemukan *error* secara cepat dan efisien di berbagai kategori. Hal ini termasuk fungsi yang salah atau hilang, kesalahan *interface*, masalah dengan struktur data atau akses *database* eksternal, kesalahan kinerja, serta kesalahan yang terkait dengan operasi dan *shutdown* sistem (Corso, Moss, Koren, Lee, & Kochenderfer, 2021).