

BAB II

LANDASAN TEORI

2.1 Pengadaan Barang dan Jasa (PBJ) Pemerintah

Pengadaan, menurut Sihaya (2013), adalah proses untuk memperoleh barang dan jasa yang diperlukan dengan pendekatan yang logis dan sistematis, mengikuti norma, etika, dan prosedur pengadaan yang telah ditetapkan. Yukins & Schooner (2007) menyatakan bahwa pengadaan adalah kegiatan yang bertujuan untuk mendapatkan barang atau jasa dengan cara yang transparan, efektif, dan efisien, yang sesuai dengan kebutuhan dan keinginan pengguna barang atau jasa tersebut.

Pengadaan Barang dan Jasa (PBJ) Pemerintah adalah proses yang dilakukan oleh instansi pemerintah untuk memperoleh barang atau jasa yang diperlukan dalam menjalankan tugas dan fungsinya (Arrowsmith & Kunzlik, 2009). Proses ini meliputi berbagai tahapan mulai dari perencanaan kebutuhan, persiapan anggaran, pemilihan penyedia, pelaksanaan kontrak, hingga penyelesaian dan pelaporan. PBJ memiliki peran penting karena tidak hanya bertujuan memenuhi kebutuhan pemerintah, tetapi juga mendukung pembangunan nasional, mempercepat penyediaan infrastruktur, dan memaksimalkan penggunaan anggaran secara transparan dan bertanggung jawab.

2.1.1 Tahapan Pengadaan Barang dan Jasa (PBJ) Pemerintah

Tahapan PBJ sesuai Peraturan Presiden Nomor 12 Tahun 2021 tentang Perubahan atas Peraturan Presiden Nomor 16 Tahun 2018 tentang Pengadaan Barang/Jasa Pemerintah dapat dirinci sebagai berikut:

a. Perencanaan Kebutuhan

Tahapan perencanaan kebutuhan berfungsi untuk mengidentifikasi kebutuhan riil instansi terkait agar pengadaan yang dilakukan benar-benar sesuai dengan sasaran. Dalam tahapan ini, instansi harus melakukan analisis kebutuhan dengan jelas, termasuk jenis barang atau jasa yang dibutuhkan, jumlahnya, serta perkiraan biaya. Menurut Setiyarto & Rapiyadi (2022), identifikasi kebutuhan yang tepat menjadi langkah penting dalam menghindari kesalahan pengadaan yang bisa menghambat kinerja organisasi pemerintah. Perencanaan yang akurat akan menjadi dasar bagi seluruh tahapan pengadaan selanjutnya.

b. Persiapan Pengadaan

Tahapan berikutnya adalah persiapan pengadaan. Pada tahap ini, panitia pengadaan menyusun berbagai dokumen yang diperlukan, termasuk Rencana Umum Pengadaan (RUP) yang mencakup detail kebutuhan, jadwal, dan anggaran yang telah dialokasikan. RUP harus diumumkan secara terbuka melalui Sistem Informasi Rencana Umum Pengadaan (SIRUP) untuk menjaga transparansi dan akuntabilitas. Tahap persiapan juga mencakup analisis pasar yang bertujuan untuk memahami ketersediaan dan harga barang atau jasa yang

dibutuhkan. Hal ini penting agar proses pengadaan dapat berlangsung dengan efisien serta melibatkan penyedia yang kompeten.

c. Pemilihan Penyedia

Tahap pemilihan penyedia menjadi salah satu proses yang krusial dalam PBJ. Di tahap ini, satuan kerja melakukan seleksi terhadap penyedia yang mengajukan penawaran sesuai kebutuhan yang telah direncanakan. Peraturan Presiden Nomor 12 Tahun 2021 tentang Perubahan atas Peraturan Presiden Nomor 16 Tahun 2018 tentang Pengadaan Barang/Jasa Pemerintah mengatur bahwa proses pemilihan ini harus dilaksanakan melalui sistem elektronik atau e-procurement, yang dirancang untuk meminimalkan potensi penyimpangan dan menjamin persaingan yang sehat di antara penyedia. Febrina (2017) dalam penelitiannya menekankan pentingnya keterbukaan dalam proses pemilihan, agar transparansi terjaga dan penyedia yang dipilih benar-benar mampu memberikan layanan yang optimal.

d. Pelaksanaan Kontrak

Setelah penyedia terpilih, tahap pelaksanaan kontrak dimulai dengan penandatanganan kontrak yang berisi kesepakatan terkait waktu, biaya, dan kuantitas serta kualitas barang atau jasa yang akan disediakan. Kontrak juga mencakup ketentuan pembayaran serta mekanisme penyelesaian jika terjadi perselisihan antara pihak penyedia dan pemerintah. Kontrak harus mencerminkan prinsip *good governance*, seperti transparansi dan akuntabilitas, serta mencegah potensi konflik kepentingan yang dapat merugikan negara.

e. Pengawasan dan Penyelesaian

Tahapan terakhir dalam PBJ adalah pengawasan dan penyelesaian pengadaan untuk memastikan bahwa barang atau jasa yang diterima sesuai dengan spesifikasi yang telah disepakati. Usman (2017) dalam penelitiannya mengemukakan bahwa pengawasan yang ketat, termasuk melalui *probity audit*, sangat penting untuk menjamin bahwa pengadaan dilakukan sesuai prosedur dan menghindari penyimpangan yang merugikan negara. Tahapan penyelesaian pengadaan ini juga mencakup proses evaluasi akhir dan pembuatan laporan untuk menutup kegiatan pengadaan secara administratif.

2.1.2 Kecurangan pada Pengadaan Barang dan Jasa (PBJ) Pemerintah

Kecurangan dalam proses pengadaan barang dan jasa pemerintah sering terjadi pada tahap perencanaan dan pemilihan penyedia. Tahap perencanaan memungkinkan adanya manipulasi data kebutuhan barang atau jasa, sedangkan tahap pemilihan tender sering melibatkan kolusi antara penyedia dan panitia pengadaan untuk memenangkan kontrak. Undang-Undang Nomor 31 Tahun 1999 tentang Pemberantasan Tindak Pidana Korupsi. menetapkan bahwa tindakan manipulatif dalam proses pengadaan dapat dikategorikan sebagai tindak pidana korupsi, termasuk di dalamnya praktik suap dan kolusi. Penelitian Rahman dkk. (2020) di Kejaksaan Negeri Jeneponto menunjukkan bahwa terdapat banyak kasus korupsi di bidang PBJ yang sering terjadi akibat kurangnya pengawasan dan pemanfaatan sistem audit yang lemah.

2.2 Teori Kecurangan

Berbagai teori telah dikembangkan untuk memahami faktor-faktor yang mendorong terjadinya kecurangan yang dapat dikaitkan dengan isu pengadaan

barang dan jasa pemerintah, antara lain:

2.2.1 Teori Fraud Triangle

Cressey (1973) dalam bukunya *Other People's Money* mengidentifikasi tiga elemen utama yang mendorong individu melakukan kecurangan:

1. Tekanan (Pressure): Dorongan atau tekanan yang dirasakan individu, seperti masalah keuangan atau target yang tidak realistis.
2. Kesempatan (Opportunity): Situasi yang memungkinkan individu melakukan kecurangan tanpa terdeteksi, sering kali akibat kelemahan dalam sistem pengendalian internal.
3. Rasionalisasi (Rationalization): Proses di mana pelaku membenarkan tindakan curangnya sebagai sesuatu yang dapat diterima.

Dalam konteks pengadaan barang dan jasa pemerintah, kelemahan dalam sistem dan prosedur pengadaan dapat menciptakan peluang bagi individu untuk melakukan kecurangan.

2.2.2 Teori Fraud Diamond

Wolfe & Hermanson (2004) mengembangkan Fraud Diamond dengan menambahkan elemen keempat yaitu “kemampuan” (capability). Kemampuan individu diartikan sebagai kemampuan untuk mengenali peluang dan memanfaatkan situasi untuk melakukan kecurangan. Penambahan elemen ini menekankan pentingnya peran kemampuan individu dalam melakukan kecurangan, terutama dalam lingkungan pengadaan yang kompleks.

2.2.3 Teori Fraud Pentagon

Howarth (2011) memperluas konsep ini menjadi Fraud Pentagon dengan menambahkan elemen kelima yaitu “arogansi” (arrogance) yang diartikan sebagai sikap superioritas yang membuat individu merasa kebal terhadap aturan dan kontrol. Menurut Lestari & Fadjrih Asyik (2023), arogansi dapat mendorong individu untuk mengabaikan prosedur pengadaan dan melakukan manipulasi demi keuntungan pribadi.

Selain ketiga teori di atas, menurut Zulaikha & Hadiprajitno (2016), fraud dalam pengadaan barang dan jasa pemerintah dipengaruhi secara signifikan oleh kelemahan dalam sistem pengadaan, kurangnya kualitas pelaksana pengadaan, serta adanya niat fraud dari pelaksana pengadaan.

2.3 Data Mining

Data mining adalah proses mengidentifikasi pola, hubungan, atau anomali dalam jumlah data yang besar menggunakan teknik komputasi canggih untuk membantu pengambilan keputusan (Dhawan, 2014). Teknik utama yang digunakan dalam *data mining* mencakup algoritma seperti klasifikasi, klastering, estimasi, prediksi, dan asosiasi. Prasetyo (2012) mengungkapkan bahwa *data mining* dapat dimanfaatkan dalam fungsi klasifikasi, analisis kelompok/klastering, deteksi anomali, dan analisis asosiasi. Setiap fungsi ini memiliki peran masing-masing dalam membantu proses pengambilan keputusan berbasis data:

a. Klasifikasi

Fungsi klasifikasi dalam *data mining* bertujuan untuk mengelompokkan data ke dalam kategori yang telah ditentukan sebelumnya. Teknik ini memanfaatkan algoritma seperti Decision Tree, Naive Bayes, Support Vector Machine, dan Neural Network untuk memprediksi kategori suatu data berdasarkan pola yang ditemukan dari data latih.

b. Analisis Kelompok (Klastering)

Analisis kelompok atau Klastering digunakan untuk mengelompokkan data ke dalam kelompok-kelompok yang memiliki karakteristik serupa tanpa perlu adanya kategori yang ditentukan sebelumnya. Algoritma seperti K-Means dan DBSCAN sering digunakan dalam clustering untuk mengidentifikasi kelompok atau pola tersembunyi dalam data.

c. Deteksi Anomali

Fungsi deteksi anomali atau *outlier detection* berfokus pada identifikasi data yang tidak sesuai atau berbeda dari pola umum dalam dataset. Deteksi anomali sangat penting untuk mendeteksi aktivitas yang mencurigakan atau abnormal.

d. Analisis Asosiasi

Analisis asosiasi bertujuan untuk menemukan aturan atau hubungan antara item dalam dataset, sering kali diterapkan dalam *market basket analysis* untuk mengidentifikasi pola pembelian produk yang sering dibeli bersama. Algoritma seperti Apriori dan FP-Growth digunakan dalam analisis ini untuk mengidentifikasi item-item yang memiliki asosiasi kuat dan relevan.

2.3.1 Anomali pada Data Mining

Dalam *data science*, anomali adalah titik atau pola data yang secara signifikan menyimpang dari perilaku yang diharapkan atau normal dalam suatu kumpulan data. Anomali, yang sering disebut juga outlier, dapat mengindikasikan peristiwa penting atau abnormal, seperti kesalahan sistem, penipuan, atau perubahan kritis dalam tren (Foorthuis, 2021). Dalam konteks data mining, algoritma deteksi anomali adalah pilihan yang tepat untuk mendeteksi kecurangan, yang secara inheren melibatkan aktivitas abnormal (Bolton & Hand, 2002).

Deteksi anomali dalam *data mining* dapat dilakukan mengidentifikasi titik data yang menyimpang dari pola normal dengan metode berbasis *threshold* dan kluster. Chandola dkk. (2009) mengungkapkan bahwa thresholding digunakan untuk membedakan anomali dengan menetapkan batas nilai; data yang melebihi batas ini dianggap sebagai anomali. Dalam deteksi anomali, threshold yang efektif seringkali ditentukan dengan distribusi statistik menggunakan ukuran seperti Z-Score atau rentang interkuartil.

Aggarwal (2015) menjelaskan peran klustering mendeteksi anomali dengan menganalisis kepadatan dan jarak dalam kelompok titik data yang serupa. Misalnya, Density-Based Spatial Clustering of Applications with Noise (DBSCAN) efektif karena mengelompokkan area padat sambil memberi label titik-titik yang terisolasi dan jarang sebagai outlier. Demikian pula, K-means clustering yang menilai jarak antara titik dan pusat kluster terdekatnya, lalu menandai titik-titik dengan jarak yang sangat tinggi sebagai anomali. Teknik pengelompokan tersebut dapat membedakan antara kluster padat data umum dan wilayah jarang

yang kemungkinan mengandung anomali.

2.3.2 Evaluasi Performa

Evaluasi model dalam konteks machine learning dan data mining memiliki tujuan untuk mengukur efektivitas berbagai metode dalam mengidentifikasi data yang menyimpang. Sathyanarayanan (2024) mengungkapkan, untuk mengevaluasi efektivitas dan akurasi sebuah model dapat menggunakan *confusion matrix* yang memiliki beberapa metrik evaluasi seperti *Accuracy*, *Precision*, *Recall*, *Specificity*, *F1-Score*, serta kurva *ROC-AUC* akan memvisualisasikannya.

a. *Confusion Matrix*

Confusion Matrix adalah tabel yang digunakan untuk menggambarkan bagaimana performa kinerja suatu sistem klasifikasi ketika nilai aktualnya diketahui. *Confusion matrix* mempresentasikan prediksi dan kondisi aktual dari algoritma *machine learning* terutama model klasifikasi (Kristiawan & Widjaja, 2021). Matriks ini memungkinkan kita untuk melihat secara jelas berapa banyak klasifikasi yang benar dan salah dilakukan oleh model pada setiap kelas (positif dan negatif). Elemen-elemen ini kemudian digunakan untuk menghitung berbagai metrik lain seperti Precision, Recall, dan F1-score.

Tabel II. 1 *Confusion Matrix*

		<i>Actual Values</i>	
		<i>Positive</i>	<i>Negative</i>
<i>Predicted Values</i>	<i>Positive</i>	<i>True Positive (TP)</i>	<i>False Positive (Type I Error)</i>
	<i>Negative</i>	<i>False Negative (Type II Error)</i>	<i>True Negative (TN)</i>

Sumber: (Swaminathan & Tantri, 2024)

Di dalam *confusion matrix* terdapat 4 kondisi, yaitu:

- *True Positive (TP)* prediksi *positif*, kondisi aktualnya *positif*
- *True Negative (TN)* prediksi *negatif*, kondisi aktualnya *negatif*
- *False Positive (FP)* prediksi *positif*, kondisi aktualnya *negatif (Type I Error)*
- *False Negative (FN)* prediksi *negatif*, kondisi aktualnya *positif (Type II Error)*

b. Akurasi (*Accuracy*)

Akurasi adalah rasio prediksi benar yang terdiri dari positif dan negatif dengan keseluruhan data.

$$Akurasi = \frac{TP + TN}{TP + FP + TN + FN}$$

Akurasi sangat berguna ketika kelas dalam dataset seimbang, namun memiliki keterbatasan dalam menangani dataset yang tidak seimbang. Dalam kasus seperti itu, model mungkin menunjukkan akurasi tinggi hanya karena lebih sering memprediksi kelas mayoritas. Namun, menurut Sathyanarayanan (2024), kemampuannya dalam mengidentifikasi kelas minoritas bisa jadi rendah. Akurasi juga bisa memberikan gambaran yang keliru jika distribusi kelas dalam dataset tidak merata.

c. Precision

Precision adalah salah satu metrik utama yang digunakan untuk mengevaluasi performa model dalam konteks klasifikasi biner. Precision mengukur proporsi prediksi positif yang benar. Precision sangat berguna ketika

biaya yang keluar ketika terjadi false alarm besar. Secara matematis, Precision dapat dihitung dengan:

$$Precision = \frac{TP}{TP + FP}$$

Sathyanarayanan (2024) mengungkapkan presisi tidak menilai kinerja model dalam menangani kasus negatif serta *false negative* sehingga dapat gagal mendeteksi banyak kasus yang benar-benar positif.

d. Recall

Recall mengukur proporsi aktual positif yang berhasil terdeteksi oleh model (Sathyanarayanan, 2024). Metrik ini sangat penting dalam deteksi anomali karena model yang lebih sensitif terhadap anomali dapat menangkap lebih banyak kejadian anomali, meskipun dengan kemungkinan menghasilkan lebih banyak False Positives. Recall dihitung dengan rumus:

$$Recall = \frac{TP}{TP + FN}$$

e. F1-score

Menurut Sokolova & Lapalme (2009) Precision dan Recall sering kali saling bertolak belakang, meningkatkan satu metrik dapat menurunkan yang lainnya. Untuk mengatasi ketidakseimbangan antara Precision dan Recall, F1-Score digunakan sebagai metrik gabungan yang merupakan *harmonic mean* dari Precision dan Recall:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

F1-Score memberikan gambaran yang lebih seimbang ketika Precision dan Recall memiliki bobot yang sama, dan sangat berguna ketika data tidak seimbang atau ketika False Positives dan False Negatives memiliki konsekuensi yang sama.

f. ROC (*Receiver Operating Characteristic*)

Di dalam *confusion matrix* hanya menyajikan informasi dalam bentuk angka. Untuk memvisualisasikan kinerja algoritma klasifikasi secara grafis, dapat menggunakan *Receiver Operating Characteristic* (ROC). ROC merupakan kurva yang menampilkan ilustrasi performansi dari *binary classifier system* dalam menghasilkan sebuah prediksi.

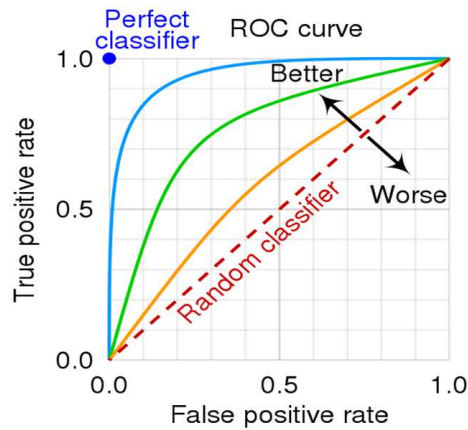
Kurva ROC awalnya dikembangkan untuk membedakan antara *signal* (*True Positive Result*) dan *noise* (*False Positive Result*) dalam menganalisis sinyal pada radar selama Perang Dunia II. Selanjutnya, ROC ini diterapkan dalam berbagai macam bidang termasuk kedokteran dan *machine learning*. (Yang & Berdine, 2017)

Kurva ROC dibuat berdasarkan perhitungan menggunakan *confussion matrix*. Di dalamnya menampilkan dua parameter utama yaitu *True Positive Rate* (TPR) yang digambarkan pada sumbu *y* dan *False Positive Rate* (FPR) *x*. Kurva ROC berperan dalam menentukan ambang batas (*threshold*) terbaik bagi model klasifikasi untuk memaksimalkan *True Positives* dan meminimalkan *False Positives*.

Secara ideal, kurva ROC menunjukkan True Positive Rate (TPR) 100% dengan False Positives (FP) nol. Kurva ini lebih sesuai untuk digunakan pada

dataset yang seimbang.

Gambar II. 1 Kurva Receiver Operating Characteristic



Sumber: Swaminathan & Tantri (2024)

g. Area Under the ROC Curve (AUC)

Area Under the ROC Curve (AUC) mengukur kemampuan model dalam membedakan antara kelas positif dan negatif. ROC adalah grafik yang memplot True Positive Rate (TPR) terhadap False Positive Rate (FPR) pada berbagai nilai threshold Fawcett (2006).. AUC memiliki nilai antara 0 dan 1. AUC bernilai 1 untuk model yang sepenuhnya benar, yang berarti model tersebut adalah *classifier* sempurna. Sebaliknya, model yang sepenuhnya salah memiliki AUC sebesar 0. Semakin tinggi nilai AUC, semakin baik model dalam membedakan antara dua kelas. Kriteria nilai untuk AUC adalah sebagai berikut:

Tabel II. 2 Intrepetasi Skor AUC ROC

Nilai AUC	Kualitas Model
0,9 – 1,0	Sangat Baik
0,8 – 0,9	Baik
0,7 – 0,8	Cukup
0,6 – 0,7	Kurang
0,5 – 0,6	Gagal
0,5	Acak / Random

Sumber: Yang & Berdine (2017)

h. Matthews Correlation Coefficient (MCC)

Matthews Correlation Coefficient (MCC) adalah salah satu metrik yang digunakan untuk mengevaluasi kinerja model klasifikasi, khususnya pada masalah yang melibatkan kelas yang tidak seimbang (imbalanced class). MCC memberikan gambaran yang lebih menyeluruh tentang kualitas model klasifikasi dengan mempertimbangkan semua kemungkinan hasil klasifikasi, termasuk true positives (TP), true negatives (TN), false positives (FP), dan false negatives (FN). MCC dihitung dengan rumus:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Nilai MCC berkisar antara -1 hingga +1, di mana nilai +1 menunjukkan prediksi yang sempurna, 0 menunjukkan tidak ada korelasi, dan -1 menunjukkan prediksi yang sepenuhnya salah (prediksi yang berlawanan dengan kenyataan). Menurut Chicco dkk. (2021) salah satu kelebihan MCC dibandingkan dengan metrik lainnya seperti akurasi atau F1-score adalah kemampuannya untuk memberikan gambaran yang lebih objektif pada data

yang tidak seimbang, di mana metrik seperti akurasi bisa memberikan hasil yang menyesatkan .

Pentingnya MCC dalam konteks evaluasi model klasifikasi semakin diakui di berbagai bidang, mulai dari bioinformatika hingga analisis data besar, karena dapat menghindari bias yang sering muncul pada model dengan distribusi kelas yang tidak seimbang (Baldi dkk., 2000). Oleh karena itu, MCC dianggap sebagai salah satu metrik yang sangat berguna dalam evaluasi kinerja model yang lebih komprehensif.

2.4 Penelitian Terdahulu

Dalam topik pendekatan *data mining* dalam audit, Werner dkk. (2021) mengkaji penggunaan *process mining* untuk meningkatkan keandalan audit. Penelitian mereka menemukan bahwa *process mining* memungkinkan auditor untuk menelusuri dan memvalidasi jejak audit, mengurangi intervensi manual, dan meningkatkan akurasi hasil audit. Penggunaan otomatisasi ini berulang kali muncul dalam inovasi audit modern karena memungkinkan tingkat ketepatan, konsistensi, dan ketepatan waktu yang lebih baik.

Bi (2023) menggunakan algoritma genetika berbasis *data mining* untuk meningkatkan kualitas analisis data keuangan, menunjukkan bahwa pendekatan ini sangat efektif dalam menganalisis kumpulan data yang kompleks. Demikian pula, Zhou (2022) yang mengeksplorasi penggunaan algoritma neural network dalam *smart audit system* yang membantu dalam mengevaluasi variabel-variabel keuangan. Penelitian mereka menyimpulkan bahwa penerapan *data mining* dalam proses audit mendukung pengambilan keputusan yang lebih baik dan deteksi

kesalahan yang lebih dini, menawarkan pendekatan proaktif terhadap audit keuangan.

Niessen dkk. (2020) mengusulkan penggunaan data mining terhadap Open Contracting Data Standard (OCDS) untuk mendeteksi anomali pada transaksi pengadaan. Dengan menggunakan algoritma isolation forest untuk mendeteksi anomali, penulis mengemukakan bahwa isolation forest dapat menjadi alat yang efektif untuk mengidentifikasi transaksi yang berpotensi mengindikasikan adanya fraud atau penyimpangan dalam proses pengadaan.

Studi lainnya, oleh Potin dkk. (2023) mengungkapkan potensi adopsi pendekatan *machine learning* berbasis graf dalam mendeteksi anomali pada jaringan hubungan antar entitas yang terlibat dalam pengadaan publik. Pendekatan ini memanfaatkan teknik pattern mining untuk menemukan pola-pola yang tidak biasa dalam graf transaksi, yang dapat mengindikasikan adanya tindakan manipulasi. Pemetaan entitas-entitas yang berinteraksi dalam pengadaan ini (seperti penyedia barang/jasa, pejabat pengadaan, dan lain-lain), memungkinkan analisis graf sehingga memberikan wawasan yang lebih mendalam terkait hubungan yang mencurigakan antar entitas.

Penelitian yang dilakukan oleh Torres-Berru & López Batista (2021) mengungkapkan bahwa, berbagai algoritma seperti clustering (K-Means), Self-Organizing Map (SOM), Support Vector Machine (SVM), dan Principal Component Analysis (PCA). SOM dapat digunakan untuk menentukan tingkat pengaruh dari setiap parameter penilaian pada berbagai fase pengadaan. K-Means digunakan untuk membuat kelompok melalui clustering, serta semi-supervised

learning dengan SVM dan PCA digunakan untuk membangun model guna mendeteksi anomali dalam proses pengadaan. Hasil penelitian menunjukkan bahwa model berbasis data mining dapat membantu mendeteksi potensi anomali yang mungkin terlewatkan oleh metode konvensional (akurasi 95%), dengan demikian meminimalkan potensi korupsi dan penyalahgunaan anggaran.

Sangil (2020) juga membahas tema serupa dengan meneliti *process mining* berbasis heuristik dalam pengadaan publik, mengidentifikasi penyimpangan dari prosedur yang seharusnya. Pendekatan ini meningkatkan kepatuhan pada standar yang akan memperkuat unsur akuntabilitas dalam pengadaan publik. Perng & Chang (2004) juga menerapkan *data mining* dalam pengadaan konstruksi pemerintah untuk meningkatkan kepatuhan terhadap standar internasional, yang memperkuat transparansi dan akuntabilitas.

Kirkos dkk. (2007) menggunakan pendekatan *data mining* klasifikasi dengan menggunakan decision trees, neural network, dan Bayesian network untuk mendeteksi *fraud* melalui analisis rasio keuangan. Temuan mereka menunjukkan bahwa model *machine learning* ini memperkuat deteksi *fraud*, terutama melalui kemampuan prediktif yang tinggi. Byrnes (2019) dan Wan dkk. (2021) juga mendukung otomatisasi analisis yang dilakukan menggunakan algoritma klustering. Hasil penelitian menunjukkan bahwa metode ini meningkatkan akurasi deteksi *fraud* dengan mengungkap pola data tersembunyi di berbagai kumpulan *big data* dan memperkuat efektivitas audit.

Penelitian Ghedini Ralha & Sarmiento Silva (2012) yang mengembangkan

sistem multi-agent *data mining* yang menggabungkan clustering dan aturan asosiasi untuk mendeteksi kartel dalam data pengadaan, mencapai akurasi 90% dalam mendeteksi kartel. Di Indonesia, Barkhah (2023b) menggunakan pendekatan berbasis data untuk mendeteksi perilaku kolusif baik vertikal maupun horizontal dalam pengadaan pemerintah, mengidentifikasi klaster perusahaan dengan pola penawaran yang mencurigakan. Hal ini sejalan dengan penelitian Nugroho et al. (2019), yang menerapkan web mining dan analisis statistik pada data e-procurement di Indonesia, mengidentifikasi pola-pola yang menunjukkan indikasi *fraud* dalam data tender.

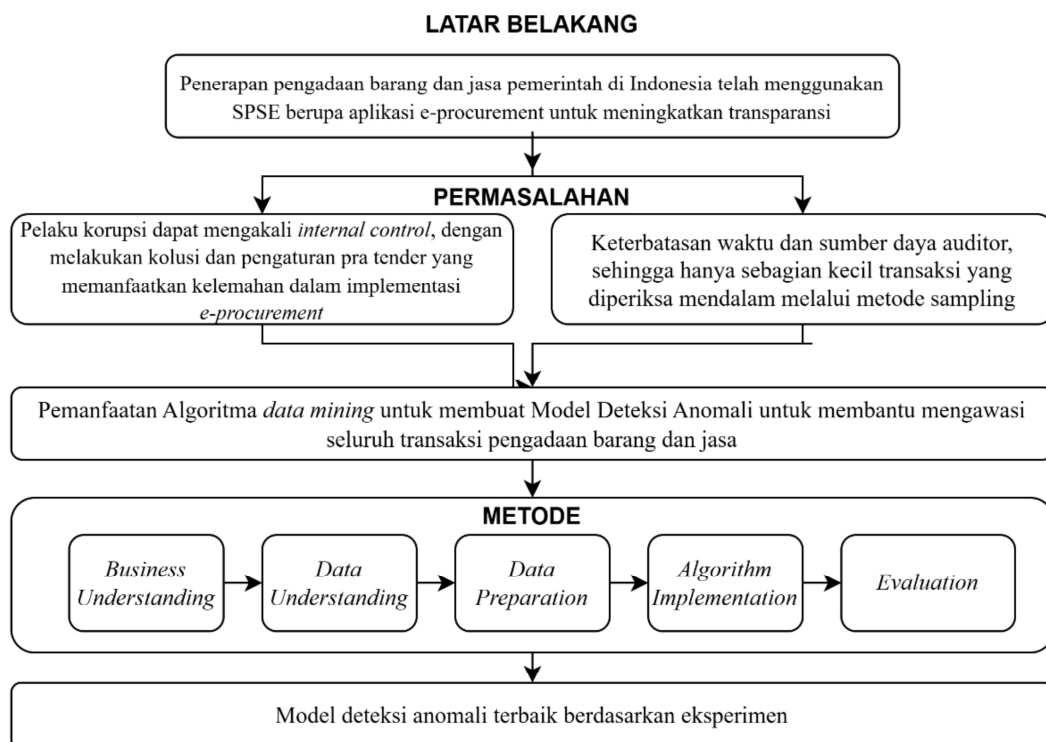
Dari hasil studi literatur di atas, dapat diketahui bahwa pembahasan mengenai penggunaan *data mining* telah banyak dibahas oleh peneliti dari berbagai negara termasuk Indonesia, baik yang dikaitkan dengan kualitas audit maupun sebagai alat bantu dalam mendeteksi anomali. Penelitian-penelitian tersebut saling mendukung satu sama lain bahwa penggunaan *data mining* dapat meningkatkan efektivitas dan efisiensi dalam proses audit, namun belum ada penelitian yang menggunakan pendekatan berbagai algoritma supervised learning dan unsupervised learning secara langsung dalam membangun model deteksi sehingga dapat memberikan gambaran yang lebih komprehensif terhadap potensi algoritma data mining dalam membantu mendeteksi anomali. Data rinci studi literatur tertuang pada lampiran 1

2.5 Kerangka Pemikiran

Setelah memahami referensi yang memiliki keterkaitan dengan topik skripsi, untuk memudahkan dalam penelitian ke depan maka penulis merancang sebuah kerangka pemikiran. Sugiyono (2013) mengungkapkan bahwa kerangka pemikiran

merupakan alur berpikir penulis yang bertujuan untuk membentuk alur penelitian yang jelas dan dapat diterima oleh akal. Kerangka pemikiran ini memberikan gambaran komprehensif mengenai latar belakang, masalah, metode pengumpulan dan pengolahan data, serta hasil penelitian.

Gambar II. 2 Kerangka Pemikiran



Sumber: Diolah Penulis (2025)