

**KLASIFIKASI ANGGARAN BERDASARKAN JENIS BELANJA DAN
SUMBER DANA TERHADAP PELAKSANAAN ANGGARAN
MENGGUNAKAN METODE RANDOM FOREST
DAN SUPPORT VECTOR MECHINE**

TESIS

Oleh:
NOVARDY
NIM. 230605210009



**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

**KLASIFIKASI ANGGARAN BERDASARKAN JENIS BELANJA DAN
SUMBER DANA TERHADAP PELAKSANAAN ANGGARAN
MENGGUNAKAN METODE RANDOM FOREST
DAN SUPPORT VECTOR MECHINE**

HALAMAN JUDUL

TESIS

Diajukan kepada:

**Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam Memperoleh gelar
Magister Komputer (M.Kom)**

**OLEH:
NOVARDY
NIM. 230605210009**

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

**KLASIFIKASI ANGGARAN BERDASARKAN JENIS BELANJA DAN
SUMBER DANA TERHADAP PELAKSANAAN ANGGARAN
MENGGUNAKAN METODE RANDOM FOREST
DAN SUPPORT VECTOR MECHINE**

HALAMAN PERSETUJUAN

TESIS

**NOVARDY
NIM. 230605210009**

Telah diperiksa dan disetujui untuk diuji:
Tanggal:

Pembimbing I,



Dr. Ririen Kusumawati, S.Si., M.Kom
NIP. 19720309 200501 2 002

Pembimbing II,



Dr. M. Amin Hariyadi, M.T
NIP. 19670118 200501 1 001

Mengetahui,

Ketua Program Studi Magister Informatika

Fakultas Sains dan Teknologi

Universitas Islam Negeri Maulana Malik Ibrahim Malang



Prof. Dr. Ir. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

**KLASIFIKASI ANGgarAN BERDASARKAN JENIS BELANJA DAN
SUMBER DANA TERHADAP PELAKSANAAN ANGgarAN
MENGGUNAKAN METODE RANDOM FOREST
DAN SUPPORT VECTOR MECHINE**
HALAMAN PENGESAHAN

TESIS

**Oleh:
NOVARDY
NIM. 230605210909**

Telah Dipertahankan di Depan Dewan Penguji Thesis
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Magister Komputer (M.Kom)
Tanggal: 10 Desember 2025

Susunan Dewan Penguji

- | | |
|----------------------|--|
| Penguji I | : Prof. Dr. Sri Harini, M. Si
NIP. 19731014 200112 2 002 |
| Penguji II | : Dr. Mochamad Imamudin, Lc., MA
NIP. 19740602 200901 1 010 |
| Pembimbing I | : Dr. Ririen Kusumawati, S.Si., M.Kom
NIP. 19720309 200501 2 002 |
| Pembimbing II | : Dr. M. Amin Hariyadi, M.T
NIP. 19670118 200501 1 001 |

Tanda Tangan

()

()
)

Mengetahui dan Mengesahkan
Ketua Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Prof. Dr. Ir. Muhammad Faisal, M.T
NIP. 19740510 200501 1 007

**KLASIFIKASI ANGGARAN BERDASARKAN JENIS BELANJA DAN
SUMBER DANA TERHADAP PELAKSANAAN ANGGARAN
MENGGUNAKAN METODE RANDOM FOREST
DAN SUPPORT VECTOR MECHINE**

TESIS

Diajukan kepada:

**Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Magister Komputer (M.Kom)**

**Oleh:
NOVARDY
NIM. 230605210009**

**PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Novardy
NIM : 230605210009
Program Studi : Magister Informatika
Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Thesis yang saya tulis ini benar-banar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan Thesis ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 10 Desember 2025

...g membuat pernyataan,



Novardy
NIM. 230605210009

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh

Syukur *Alhamdulillah* penulis hantarkan kehadiran Allah SWT yang telah melimpahkan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan studi di Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang sekaligus menyelesaikan Thesis ini dengan baik.

Selanjutnya penulis haturkan ucapan terima kasih seiring do'a dan harapan jazakumullah ahsanal jaza' kepada semua pihak yang telah membantu terselesaikannya Thesis ini. Ucapan terima kasih ini penulis sampaikan kepada:

1. Bapak Dr. Muhammad Faisal, M.T, selaku Ketua Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Ibu Dr. Ririen Kusumawati, S.Si., M.Kom dan Bapak Dr. M. Amin Hariyadi, M.T selaku dosen pembimbing Thesis, yang telah banyak memberikan pengarahan dan pengalaman yang berharga.
3. Prof. Dr. Sri Harini, M. Si dan Dr. Mochamad Imamudin, Lc., MA selaku penguji Thesis yang telah banyak memberikan masukan dan saran dalam penulisan Thesis ini
4. Segenap sivitas akademika Program Studi Magister Informatika, terutama seluruh Bapak/ Ibu dosen, terima kasih atas segenap ilmu dan bimbingannya.
5. Ayahanda dan Ibunda tercinta yang senantiasa memberikan doa dan restunya kepada penulis dalam menuntut ilmu.
6. Kakak dan adik penulis yang selalu memberikan semangat kepada penulis untuk menyelesaikan Thesis ini.
7. Semua pihak yang ikut membantu dalam menyelesaikan Thesis ini baik berupa materiil maupun moril.

Penulis menyadari bahwa dalam penyusunan Thesis ini masih terdapat kekurangan dan penulis berharap semoga Thesis ini bias memberikan manfaat kepada para pembaca khususnya bagi penulis secara pribadi. *Aamian Yaa Rabbal 'Alamin.*

Wa 'alaikumsalam warahmatullahi wabarakatuh

Malang, 25 November 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	I
HALAMAN PERSETUJUAN.....	II
HALAMAN PENGESAHAN.....	III
HALAMAN PERNYATAAN	V
KATA PENGANTAR	VI
DAFTAR ISI.....	VII
DAFTAR GAMBAR	X
DAFTAR TABEL.....	XI
ABSTRAK	XII
ABSTRACT	XIII
الملخص	XIV
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Pernyataan masalah.....	7
1.3 Tujuan Penelitian	8
1.4 Manfaat Penelitian	8
1.5 Batasan Masalah.....	9
BAB II TINJAUAN PUSTAKA DAN KERANGKA TEORI.....	11
2.1 Konsep Klasifikasi Anggaran	11
2.2 Akuntabilitas Fiskal	11
2.3 Pengelolaan APBN K/L	12
2.4 Klasifikasi Berdasarkan Jenis Belanja	13

2.5	Klasifikasi Berdasarkan Sumber Dana.....	13
2.6	Implikasi dalam Pengelolaan APBN K/L	14
2.7	Pemodelan Klasifikasi.....	15
2.8	Random Forest Classifier (RFC).....	15
2.9	Support Vector Machine (SVM).....	16
2.10	Evaluation Model.....	18
2.11	Kerangka Teori.....	19
BAB III METODOLOGI PENELITIAN		34
3.1	Jenis Penelitian.....	34
3.2	Lokasi dan Waktu Penelitian	34
3.3	Jenis dan Sumber Data	34
3.4	Teknik Pengumpulan Data.....	35
3.5	Fokus Penelitian	35
3.6	Variabel Penelitian dan Proses Pelabelan Data.....	35
3.7	Analisis Data dengan Machine Learning	35
3.8	Desain Sistem.....	36
3.9	Kerangka Konsep Penelitian	37
BAB IV ANALISIS DATA		39
4.1	Data Penelitian	39
4.1.1	Profil data.....	40
4.1.2	Statistik Deskriptif	40
4.1.3	Distribusi Data	41
4.1.4	Distribusi Persentase Realisasi Anggaran.....	42
4.2	Preprocessing Data.....	43
BAB V RANDOM FOREST CLASSIFIER (RFC)		47
5.1	Metode Random Forest.....	47

5.2	Desain Model Random Forest.....	47
5.3	Pengujian 1 dengan perbandingan data 90:10.....	48
5.4	Pengujian 2 dengan perbandingan data 85:15.....	49
5.5	Pengujian 3 dengan perbandingan data 80:20.....	50
5.6	Pengujian 4 dengan perbandingan data 70:30.....	50
5.7	Pengujian 5 dengan perbandingan data 60:40.....	51
5.8	Kesimpulan	51
BAB VI SUPPORT VECTOR MACHINE (SVM).....		53
6.1	Metode Support Vector Machine (SVM).....	53
6.2	Desain Model Support Vector Machine.....	54
6.3	Pengujian 1 SVM dengan perbandingan data 90:10	55
6.4	Pengujian 2 SVM dengan perbandingan data 85:15	56
6.5	Pengujian 3 SVM dengan perbandingan data 80:20	56
6.6	Pengujian 4 SVM dengan perbandingan data 70:30	57
6.7	Pengujian 5 SVM dengan perbandingan data 60:40	58
6.8	Kesimpulan	58
BAB VII PEMBAHASAN		60
7.1	Analisis Kinerja Model Klasifikasi.....	60
7.2	Hubungan Penelitian dengan Pandangan Islam	61
BAB VIII KESIMPULAN DAN SARAN.....		65
8.1	Kesimpulan	65
DAFTAR PUSTAKA		67
LAMPIRAN		70

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Random Forest	16
Gambar 2. 2 Optimal Hyperplane dengan Maximum Margin.	17
Gambar 2. 3 Kerangka Teori.....	29
Gambar 3. 1 Desain Sistem.....	36
Gambar 3. 2 Kerangka Konsep	37
Gambar 4. 1 Distribusi Data	41
Gambar 4. 2 Distribusi Persentase Realisasi Anggaran	42
Gambar 4. 3 Hapus Duplikasi	43
Gambar 4. 4 Hapus Nilai Kososng.....	43
Gambar 4. 5 Rasio Realisasi	43
Gambar 4. 6 Rasio Data	44
Gambar 4. 7 Membuat kolom kategori pelaksanaan.....	44
Gambar 4. 8 Hasil Pengkategorian Tingkat Serapan Anggaran	45
Gambar 4. 9 Heatmap Heatmap Korelasi Antar Variabel Numerik	45
Gambar 5. 1 Desain Model Random Forest.....	47

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	18
Tabel 2. 2 Perbandingan Akurasi Penelitian Terkait	28
Tabel 2. 3 Tabel Alur Sistem Klasifikasi	31
Tabel 4. 1 Dataset.....	39
Tabel 4. 2 Deskriptif data.....	40
Tabel 5. 1 Pengujian 1 Random Forest	48
Tabel 5. 2 Pengujian 2 Random Forest	49
Tabel 5. 3 Pengujian 3 Random Forest	50
Tabel 5. 4 Pengujian 4 Random Forest	50
Tabel 5. 5 Pengujian 5 Random Forest	51
Tabel 5. 6 Rata-Rata Accuracy, Precision, Recall, f1-score	52
Tabel 6. 1 Pengujian 1 SVM	55
Tabel 6. 2 Pengujian 2 SVM	56
Tabel 6. 3 Pengujian 3 SVM	56
Tabel 6. 4 Pengujian 4 SVM	57
Tabel 6. 5 Pengujian 5 SVM	58
Tabel 6. 6 Rata-rata Akurasi, Presisi, recall, dan f1-score	59
Tabel 7. 1 Perbandingan Accuracy, Precision, Recall Random Forest dan SVM	60

ABSTRAK

Novardy. 2025. Klasifikasi Anggaran Berdasarkan Jenis Belanja Dan Sumber Dana Terhadap Pelaksanaan Anggaran Menggunakan Metode Random Forest Classifier (RFC) Dan Support Vector Machine (SVM). Thesis. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Ririen Kusumawati, S.Si., M. Kom (II) Dr. M. Amin Hariyadi, M.T.

Kata kunci: Anggaran, Klasifikasi, Random Forest, Support Vector Machine, Pelaksanaan Anggaran

Klasifikasi anggaran memainkan peran penting terhadap perencanaan, pengelolaan serta penganggaran hingga implementasi dan akuntabilitas. Kita membuat anggaran dengan mempertimbangkan berbagai jenis pengeluaran dan sumber dana. Setiap jenis pengeluaran, seperti gaji pegawai, barang, modal, hibah, bantuan sosial, subsidi, bunga, dan pendapatan non-pajak (PNBP) atau badan pelayanan publik (BLU), memiliki seperangkat aturan dan cara pelacakan uang tersendiri. Studi ini bertujuan untuk menunjukkan bagaimana klasifikasi anggaran, berdasarkan jenis pengeluaran dan sumber dana, diterapkan dalam implementasi Anggaran Pendapatan. Penelitian ini bertujuan untuk menilai kinerja klasifikasi dengan dua model yaitu Random Forest Classifier (RFC) dan Support Vector Machine (SVM) berdasarkan data historis serta mengevaluasi kinerja masing-masing model. Uji coba menghasilkan model Random Forest secara konsisten lebih baik daripada model SVM untuk setiap proporsi data, dengan perbandingan 90:10 hingga 60:40. Model Random Forest mencapai performa terbaiknya pada pembagian data 80:20, dengan skor akurasi 94 persen, skor presisi 94 persen, skor recall 94 persen, dan skor F1 87 persen sedangkan Skor akurasi rata-rata hasil pengujian SVM adalah 80 persen.

ABSTRACT

Novardy, 2025: Budget Classification Based on Expenditure Type and Funding Source on Budget Execution Using the Random Forest Classifier (RFC) and Support Vector Machine (SVM) Methods. Master Thesis of Magister Informatika Study Program, Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Supervisors: (I) Dr. Ririen Kusumawati, S.Si., M.Kom (II) Dr. M. Amin Hariyadi, M.T.

Keywords: Budget, Classification, Random Forest, Support Vector Machine, Budget Execution

Budget classification plays a crucial role in planning, management, and budgeting, from implementation to accountability. We create budgets by considering various types of expenditures and funding sources. Each type of expenditure, such as employee salaries, goods, capital, grants, social assistance, subsidies, interest, and non-tax revenue (PNBP) or public service agencies (BLU), has its own set of rules and methods for tracking money. This study aims to demonstrate how budget classification, based on expenditure types and funding sources, is applied in the implementation of the Revenue Budget. This study aims to assess the classification performance of two models, namely the Random Forest Classifier (RFC) and Support Vector Machine (SVM), based on historical data and evaluate the performance of each model. Tests show that the Random Forest model consistently outperforms the SVM model for each data proportion, with a ratio of 90:10 to 60:40. The Random Forest model achieved its best performance at the 80:20 data split, with an accuracy score of 94 percent, a precision score of 94 percent, a recall score of 94 percent, and an F1 score of 87 percent. The average accuracy score of the SVM test results was 80 percent.

الملخص

نوفاردي، 2025: تصنيف الميزانية بناءً على نوع الإنفاق ومصدر التمويل في تنفيذ الميزانية باستخدام مصنف الغابة العشوائية (RFC) وطريقة آلة المتجهات الداعمة (SVM) رسالة ماجستير في برنامج دراسات المعلوماتية، كلية العلوم والتكنولوجيا، جامعة الإسلامية الحكومية مولانا مالك إبراهيم في مالانج. المشرفون: (1). ريرين كوسوماواتي، بكالوريوس علوم، ماجستير علوم حاسوب (2). محمد أمين هاريادي، ماجستير هندسة

الكلمات المفتاحية: الميزانية، التصنيف، الغابة العشوائية، آلة المتجهات الداعمة، تنفيذ الميزانية

يلعب تصنيف الميزانية دوراً محورياً في إدارة المالية العامة للدولة، بدءاً من التخطيط وإعداد الميزانية وصولاً إلى التنفيذ والمساءلة. يتم إعداد الميزانية مع مراعاة مختلف أنواع الإنفاق ومصادر التمويل. وكل نوع من أنواع الإنفاق، كرواتب الموظفين، والسلع، والمشاريع الرأسمالية، والمنح، والمساعدات الاجتماعية، والإعانات، والفوائد، والإيرادات غير الضريبية، أو إيرادات هيئات الخدمات العامة، مجموعة قواعده الخاصة وأدوات تتبع الأموال. تهدف هذه الدراسة إلى توضيح كيفية تطبيق تصنيف الميزانية، بناءً على أنواع الإنفاق ومصادر التمويل، في تنفيذ ميزانية الدولة. ويهدف هذا البحث إلى تقييم أداء نماذجين للتصنيف، وهما مصنف الغابات العشوائية (RFC) وآلة المتجهات الداعمة (SVM)، باستخدام تقنيات تستند إلى بيانات تاريخية لتقييم أداء كل نموذج. وتبين النتائج أن نموذج الغابات العشوائية يتفوق باستمرار على نموذج آلة المتجهات الداعمة في جميع نسب البيانات، من 90:10 إلى 40:60. حقق نموذج الغابة العشوائية أفضل أداء له عند تقسيم البيانات بنسبة 80:20، حيث بلغت دقة النموذج 94%， ودقة التصنيف 94%， ودقة الاستدعاء 94%， وقيمة F1 87%. أما متوسط أداء نموذج آلة المتجهات الداعمة (SVM) فكان 80%.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dalam mengelola distribusi, pembagian, dan stabilisasi ekonomi nasional pemerintah memerlukan alat utama yaitu APBN. Di dalamnya tercermin kebijakan anggaran pemerintah dan hal tersebut dituangkan dalam wujud rencana pendapatan, belanja, dan pembiayaan negara. APBN tidak hanya menjadi dokumen keuangan, namun juga merupakan instrumen kebijakan untuk mendukung pencapaian tujuan pembangunan nasional. Agar APBN dapat dilaksanakan secara efektif dan akuntabel, diperlukan suatu sistem klasifikasi anggaran yang jelas, konsisten, dan sesuai standar internasional.

Dalam konteks pelaksanaan anggaran di Kementerian/Lembaga (K/L), klasifikasi anggaran menjadi sangat penting karena berfungsi sebagai dasar pengelolaan keuangan negara dari perencanaan, penganggaran, pelaksanaan, dan pertanggungjawaban. Dua dimensi klasifikasi yang paling menentukan dalam tahap pelaksanaan adalah klasifikasi berdasarkan jenis belanja dan klasifikasi berdasarkan sumber dana. Kedua dimensi ini bukan hanya sekadar administrasi teknis, melainkan berimplikasi langsung terhadap tata kelola anggaran, kelancaran penyerapan, pencatatan aset negara, serta penyajian laporan keuangan pemerintah pusat.

Klasifikasi berdasarkan jenis belanja merupakan pengelompokan anggaran menurut sifat atau tujuan pengeluaran. Bagan Akun Standar (BAS) yang dikelola oleh Kementerian Keuangan mengatur bahwa jenis belanja K/L terbagi dalam beberapa kategori utama, diantaranya belanja untuk pegawai, barang, modal, hibah, bantuan sosial, subsidi, belanja bunga, serta belanja lainnya. Pengklasifikasian ini penting karena setiap jenis belanja memiliki karakteristik dan perlakuan akuntansi yang berbeda. Misalnya, belanja barang bersifat habis pakai atau untuk kegiatan operasional, sementara belanja modal menghasilkan aset tetap yang menambah kekayaan negara. Salah klasifikasi antara belanja untuk barang dan pembelanjaan modal akan berpengaruh signifikan pada laporan keuangan, khususnya neraca pemerintah, karena aset bisa tidak tercatat atau justru terjadi salah saji.

Di sisi lain, klasifikasi berdasarkan sumber dana menunjukkan asal pembiayaan dari suatu belanja. Dalam APBN K/L, terdapat beberapa sumber dana utama yang digunakan, yaitu Rupiah Murni (RM), Penerimaan (PNBP), pendapatan (BLU), Surat Berharga (SBSN), Pinjaman Luar Negeri atau Hibah Luar Negeri (PHLN), Pinjaman Dalam Negeri atau Hibah Dalam Negeri (PDN), serta hibah secara langsung. Masing-masing sumber dana memiliki aturan penggunaan dan pelaporan yang berbeda. Sebagai contoh, belanja yang menggunakan dana SBSN harus benar-benar dialokasikan untuk kegiatan atau proyek yang bersifat produktif dan menghasilkan aset, sementara belanja dengan sumber dana PHLN harus sesuai dengan ketentuan perjanjian dengan pemberi pinjaman atau hibah, termasuk dalam hal jenis belanja yang diperbolehkan (*eligible expenditure*), mekanisme penarikan dana, dan kewajiban pelaporan.

Relevansi klasifikasi jenis belanja dan sumber dana ini sangat terasa pada tahap pelaksanaan anggaran di K/L. Pada tahap penarikan dana misalnya, dokumen untuk mengajukan pencairan dana (SPM) yang diajukan pada Kantor KPPN harus sesuai dengan kode akun belanja dan kode sumber dana sebagaimana tercantum dalam DIPA. Kesalahan pengkodean akan menyebabkan penolakan SPM sehingga menghambat realisasi anggaran. Demikian pula, dalam proses pencatatan Barang Milik Negara (BMN), belanja yang seharusnya dicatat sebagai belanja modal tidak boleh diklasifikasikan sebagai belanja barang, karena akan menimbulkan salah saji aset dan berpotensi menjadi temuan audit.

Tidak jarang pula terjadi masalah ketika suatu proyek dibiayai oleh sumber dana tertentu, tetapi pada pelaksanaannya justru dibebankan ke sumber dana lain. Hal ini berpotensi menimbulkan *ineligible expenditure* pada proyek yang didanai oleh pinjaman atau hibah luar negeri, sehingga klaim penggantian dana tidak dapat diproses oleh pemberi pinjaman/hibah. Akibatnya, pemerintah menanggung kerugian ganda, yaitu dana proyek tidak dapat direimburse dan pelaporan keuangan menjadi tidak valid. Kasus serupa juga dapat terjadi pada penggunaan dana PNBP atau BLU, di mana belanja yang seharusnya bersumber dari PNBP justru dibebankan pada Rupiah Murni, sehingga mengacaukan keseimbangan penggunaan dana K/L.

Beberapa contoh nyata dapat menggambarkan pentingnya ketepatan klasifikasi ini. Suatu Kementerian/Lembaga (KL/L), melaksanakan pembangunan infrastruktur melalui belanja modal dengan sumber dana SBSN. Apabila belanja tersebut salah diklasifikasikan sebagai belanja barang, maka tidak hanya aset infrastruktur yang tidak tercatat dalam neraca BMN, tetapi juga laporan kinerja proyek SBSN menjadi tidak valid. Demikian pula, Kementerian/Lembaga (KL/L) lain dalam pengadaan vaksin menggunakan dana PHLN yang harus dicatat secara spesifik sesuai perjanjian dengan donor. Salah klasifikasi belanja akan menyebabkan belanja tersebut tidak diakui sebagai *ineligible expenditure*, sehingga klaim penggantian dana ditolak.

Kondisi-kondisi tersebut menunjukkan bahwa klasifikasi anggaran menurut jenis belanja dan sumber dana tidak hanya sekadar aspek teknis administrasi, tetapi juga berkaitan erat dengan aspek akuntabilitas, transparansi, dan efektivitas penggunaan APBN. Kesalahan dalam pengklasifikasian dapat berdampak sistemik, mulai dari keterlambatan realisasi anggaran, tidak tercatatnya aset negara, laporan keuangan yang salah saji, hingga berkurangnya kepercayaan publik maupun donor internasional terhadap kredibilitas pengelolaan keuangan negara.

Dengan demikian, kajian mengenai klasifikasi anggaran berdasarkan jenis belanja dan sumber dana dalam pelaksanaan anggaran pada K/L menjadi sangat penting untuk dilakukan. Kajian ini akan membantu mengidentifikasi sejauh mana klasifikasi tersebut memengaruhi kelancaran pelaksanaan anggaran, kualitas pertanggungjawaban keuangan, dan kepatuhan pada peraturan perundang-undangan. Disisi lain, kajian ini juga dapat memberikan rekomendasi perbaikan tata kelola anggaran, baik dari sisi sistem aplikasi, regulasi, maupun peningkatan kapasitas sumber daya manusia di lingkungan K/L.

Secara prinsipil, pengelolaan keuangan negara merupakan sebuah amanah (trust) yang dititipkan oleh rakyat kepada para pengelolanya. Dalam Surat An-Nisa' ayat 58 Allah SWT berfirman:

إِنَّ اللَّهَ يَأْمُرُكُمْ أَنْ تُؤْدُوا الْأَمْانَاتِ إِلَى أَهْلِهَا وَإِذَا حَكَمْتُمْ بَيْنَ النَّاسِ أَنْ تَحْكُمُوا بِالْعُدْلِ إِنَّ اللَّهَ نِعِمَّا يَعِظُكُمْ بِهِ إِنَّ اللَّهَ كَانَ سَمِيعًا بَصِيرًا

"Sesungguhnya Allah menyuruh kamu menyampaikan amanat kepada yang berhak menerimanya, dan (menyuruh kamu) apabila menetapkan hukum di antara manusia supaya kamu menetapkan dengan adil. Sesungguhnya Allah memberi pengajaran yang sebaik-baiknya kepadamu. Sesungguhnya Allah adalah Maha Mendengar lagi Maha Melihat."

Ayat ini memiliki dua dimensi utama:

1. Dimensi Amanah, yaitu kewajiban untuk menunaikan setiap titipan sesuai dengan haknya (Nikmah, *et al.*, 2022). Dalam pengelolaan APBN Kementerian/Lembaga (K/L), amanah dapat dipahami sebagai kewajiban pejabat pengelola anggaran untuk memastikan bahwa alokasi dan penggunaan dana negara tepat sasaran sesuai peruntukannya, baik berdasarkan jenis belanja (pegawai, barang, modal, bansos) maupun sumber dana (PNBP, rupiah murni, hibah, BLU, pinjaman, SBSN).
2. Dimensi Keadilan, yaitu kewajiban untuk menegakkan keadilan dalam setiap keputusan. Hal ini mengandung makna bahwa pelaksanaan anggaran tidak boleh dipengaruhi kepentingan pribadi, kelompok, atau politik, tetapi harus berpijak pada asas manfaat, pemerataan, serta kepentingan masyarakat luas.

Dalam literatur tafsir klasik seperti *Tafsir Ibn Katsir*, makna ayat ini, menyatakan bahwa amanat dimaknai umum dan diwajibkan untuk semua orang, baik yang bertakwa ataupun durhaka. Sebagian besar mufassirin menyatakan bahwa ayat diturunkan berkenaan tentang Usman ibnu Talhah ibnu Abu Talhah. Abdullah ibnu Abdul Uzza ibnu Usman ibnu Abdud Dar ibnu Qusai ibnu Kilab Al-Qurasyi Al-Abdari adalah nama Abu Talhah. Menurut kisah terkenal, ayat ini diturunkan tentang peristiwa tersebut. Karena itu, Muhammad ibnul Hanafiyah dan Ibnu Abbas menyatakan bahwa amanat ini ditujukan kepada baik orang yang berbuat baik maupun orang yang berbuat tidak baik. Dengan kata lain, semua orang harus bersifat umum. Ini merupakan perintah dari Allah SWT berkenaan meminta agar hukum ditetapkan secara adil di antara manusia. Karena itu, Muhammad ibnu Ka'b, Zaid ibnu Aslam, dan Syahr ibnu Hausyab menjelaskan bahwa ayat ini hanya diturunkan untuk para hakim yang memutuskan perkara di antara manusia. Amanah pada ayat ini tidak hanya mencakup titipan individu, tetapi juga tanggung jawab publik, termasuk jabatan, wewenang, dan harta. Para mufassir menekankan bahwa seorang pemimpin yang diberi kewenangan atas urusan umat wajib

menyalurkan hak secara benar, adil, dan tidak zalim. Relevansi ayat tersebut dalam penelitian terkait klasifikasi anggaran adalah bahwa setiap bentuk pengelolaan anggaran negara sejatinya merupakan amanah publik yang harus dipertanggungjawabkan dengan prinsip akuntabilitas, transparansi, dan keadilan. Ayat ini dapat menjadi landasan normatif dan etik yang memperkuat urgensi penelitian mengenai kesesuaian klasifikasi anggaran dengan implementasinya pada K/L (Al-Imam Abul Fida Isma'il Kasir Ad-Dimasyqi 2000).

Ayat ini juga menegaskan kewajiban untuk mengelola amanat dengan benar, disampaikan kepada pemiliknya (dalam hal ini rakyat), dan ditegakkan berdasarkan landasan keadilan. Akuntabilitas, transparasi dan efektivitas dalam mengelola APBN merupakan aplikasi nyata dalam menjalankan amanat tersebut. Keterbukaan informasi (transparansi) adalah cara menyampaikan amanat kepada pemiliknya, sedangkan pertanggungjawaban yang benar (akuntabilitas) dan penggunaan yang tepat sasaran (efektivitas) adalah wujud dari berbuat adil dalam pembangunan.

Dalam penyelenggaraan pemerintahan dan pengelolaan urusan publik, dalam Surat An-Nisa ayat 58 menegaskan dua landasan utama yaitu amanah serta keadilan. Amanah harus disampaikan kepada yang berhak, dan keputusan ditetapkan dengan landasan keadilan. Amanah dalam konteks ini tidak hanya terbatas pada titipan materi, tetapi juga mencakup tanggung jawab jabatan, kewenangan, serta pengelolaan harta publik seperti anggaran negara. Penegasan ayat ini diperkuat oleh sejumlah hadis shahih. Nabi SAW bersabda:

عَنْ أَبِي ذِرٍّ رَضِيَ اللَّهُ عَنْهُ قَالَ: قُلْتُ: يَا رَسُولَ اللَّهِ، أَلَا تَسْتَعْمِلُنِي؟ فَضَرَبَ بِيَدِهِ عَلَى مَنْكِبِي، ثُمَّ قَالَ: يَا أَبَا ذِرٍّ، إِنَّكَ ضَعِيفٌ، وَإِنَّهَا أَمَانَةٌ، وَإِنَّهَا يَوْمَ الْقِيَامَةِ حُرْبٌ وَنَدَامَةٌ، إِلَّا مَنْ أَخْذَهَا بِحَقِّهَا، وَأَدَى الَّذِي عَلَيْهِ فِيهَا

“Dari Abu Dzar ia berkata: Aku berkata, "Wahai Rasulullah, tidakkah engkau memberiku jabatan (untuk memimpin)?" Maka beliau menepuk pundakku dengan tangannya, kemudian bersabda; "Wahai Abu Dzar, sesungguhnya engkau lemah, sedangkan jabatan itu adalah amanah. Dan pada hari kiamat, jabatan itu akan menjadi kehinaan dan penyesalan, kecuali bagi orang yang mengambilnya dengan haknya dan menunaikan kewajiban di dalamnya". (HR. Muslim, no. 1825).

Hadis ini memberikan peringatan bahwa setiap jabatan dan wewenang (Lu'ul Isnaini *et al.*, 2024), termasuk dalam pengelolaan anggaran negara, bukanlah sekadar posisi kehormatan, melainkan sebuah amanah yang berat dan akan dimintai pertanggungjawaban. Selain itu, Rasulullah SAW juga bersabda:

فَإِذَا ضُيِّعَتِ الْأَمَانَةُ فَأَنْتَظِرْ السَّاعَةَ قَالَ كَيْفَ إِضَاعَتُهَا قَالَ إِذَا وُسِّدَ الْأَمْرُ إِلَى غَيْرِ
أَهْلِهِ فَأَنْتَظِرْ السَّاعَةَ

“Apabila amanah sudah disia-siakan maka tunggulah kiamat. Seseorang bertanya: “Bagaimana disia-siakannya amanat itu?”. Nabi SAW menjawab: “Jika urusan diserahkan bukan kepada ahlinya, maka tunggulah kiamat”.

(HR Bukhari No. 59 dari Abu Hurairah Kitab al-Ilmi Bab *Man Suila 'Ilman wa huwa Muataghillun fii Hadiitsihi*)(KH Zakky Mubarak 2021).

Hadis ini selaras dengan perintah Allah dalam surat An-Nisa ayat 58, yang menegaskan pentingnya menempatkan amanah kepada orang yang berhak dan memiliki kompetensi. Dalam konteks pengelolaan APBN K/L, hal ini berarti penempatan pejabat pengelola anggaran harus berdasarkan kapabilitas, integritas, serta kemampuan mempertanggungjawabkan tugasnya. Lebih jauh, Rasulullah ﷺ menekankan keutamaan keadilan dengan sabdanya:

عَنْ عَبْدِ اللَّهِ بْنِ عَمْرُو بْنِ الْعَاصِ رَضِيَ اللَّهُ تَعَالَى عَنْهُمَا قَالَ: قَالَ رَسُولُ اللَّهِ صَلَّى
اللَّهُ عَلَيْهِ وَسَلَّمَ : إِنَّ الْمُقْسِطِينَ عِنْدَ اللَّهِ عَلَى مَنَابِرِ مِنْ نُورٍ، عَنْ يَمِينِ الرَّحْمَنِ
الَّذِينَ يَعْدِلُونَ فِي حُكْمِهِمْ وَأَهْلِهِمْ وَمَا وَلَوْا

“Abdullah bin Amr r.a. berkata, sesungguhnya Rasulullah SAW bersabda: “Penegak keadilan itu di sisi Allah berada di atas “mimbar cahaya”. Mereka itu adalah para penegak keadilan dalam menetapkan hukum terhadap keluarga dan rakyat yang berada di bawah pimpinannya.” (HR. Muslim, no. 1827).

Hadis ini memberikan penghargaan tinggi kepada pemimpin atau pejabat yang berlaku adil, selaras dengan perintah ayat untuk menetapkan hukum dengan adil. Dalam praktik pengelolaan anggaran, keadilan diwujudkan melalui distribusi

anggaran yang transparan, tidak diskriminatif, dan berorientasi pada kemaslahatan rakyat.

Dengan demikian, surat An-Nisa ayat 58 beserta hadis-hadis pendukungnya menegaskan bahwa pengelolaan anggaran negara, khususnya dalam Kementerian/Lembaga, adalah kewajiban yang harus dilaksanakan dengan penuh akuntabilitas, profesionalisme, dan keadilan. Prinsip ini menjadi landasan normatif bagi penelitian mengenai klasifikasi anggaran berdasarkan jenis belanja dan sumber dana, sekaligus mengingatkan bahwa setiap keputusan dalam pelaksanaan anggaran akan dipertanggungjawabkan, baik di dunia maupun di akherat.

Penelitian tentang klasifikasi anggaran menurut jenis belanja dan sumber dana dalam konteks pelaksanaan anggaran K/L diharapkan dapat memberikan kontribusi terhadap perbaikan manajemen keuangan negara. Guna meningkatkan akurasi dan objektivitas dalam proses klasifikasi tersebut, penelitian ini akan memanfaatkan teknik pembelajaran mesin, khususnya *Random Forest Classifier* dan *Support Vector Machine* (SVM).

Kedua teknik tersebut dipilih karena kemampuan mereka dalam menganalisis dataset yang kompleks dan multidimensi seperti data anggaran. Random Forest mampu memberikan insight mengenai faktor-faktor penentu klasifikasi (mendukung aspek akuntabel dan transparan), sementara SVM dikenal robust dalam menghasilkan model klasifikasi dengan akurasi tinggi (mendukung aspek efektif).

Penelitian ini tidak hanya bermanfaat secara akademis sebagai referensi ilmiah dalam membandingkan kinerja kedua algoritma tersebut, tetapi juga memiliki nilai praktis yang signifikan. Model klasifikasi yang dihasilkan diharapkan dapat menjadi alat bantu yang mendukung terwujudnya pengelolaan APBN yang lebih transparan, akuntabel, dan efektif.

1.2 Pernyataan masalah

- a. Bagaimana penerapan klasifikasi anggaran berdasarkan jenis belanja dan sumber dana dalam pelaksanaan anggaran APBN pada K/L?

- b. Bagaimana tingkat kinerja model Random Forest Classifier dan Support Vector Machine (SVM) dalam mengelompokkan jenis belanja dan sumber dana anggaran K/L?
- c. Dari dua model klasifikasi Random Forest Classifier dan Support Vector Machine (SVM), manakah kinerja terbaik untuk klasifikasi data anggaran K/L sesuai akurasi, presisi, recall, dan F1-score?
- d. Rekomendasi apa saja yang dapat disusun untuk meningkatkan transparansi, akuntabilitas, efisiensi, dan kesesuaian tujuan pembangunan dalam pelaksanaan anggaran K/L?

1.3 Tujuan Penelitian

- a. Mengidentifikasi dan mengevaluasi penerapan klasifikasi anggaran berdasarkan jenis belanja dan sumber dana dalam pelaksanaan APBN pada K/L melalui analisis data historis dan business rules yang berlaku.
- b. Mengukur dan menganalisis tingkat kinerja model Random Forest Classifier dan Support Vector Machine (SVM) dalam mengelompokkan jenis belanja dan sumber dana anggaran K/L menggunakan metrik evaluasi standar.
- c. Membandingkan performa kedua model klasifikasi guna menentukan model terbaik berdasarkan analisis komparatif akurasi, presisi, recall, dan F1-score pada data anggaran K/L.
- d. Menyusun rekomendasi perbaikan agar pelaksanaan anggaran lebih transparan, akuntabel, efisien, dan sesuai tujuan pembangunan.

1.4 Manfaat Penelitian

- a. Manfaat Akademik
 1. Menambah literatur bidang manajemen keuangan publik, akuntansi pemerintahan, dan kebijakan fiskal.
 2. Menjadi referensi untuk penelitian lanjutan tentang transparansi, akuntabilitas, dan efektivitas pengelolaan anggaran.
- b. Manfaat Bagi Pemerintah
 1. Menjadi dasar evaluasi dalam merancang dan melaksanakan anggaran sesuai klasifikasi belanja dan sumber dana.

2. Membantu optimalisasi perencanaan, pelaksanaan, hingga revisi anggaran agar tepat sasaran dan sesuai peraturan.
- c. Manfaat Bagi Lembaga Pengawasan dan Audit
 1. Memberikan perspektif praktis tentang potensi permasalahan dalam klasifikasi anggaran yang dapat memicu temuan audit.
 2. Mendukung fungsi pengawasan dengan data dan analisis yang lebih tajam terkait belanja dan sumber dana.
- d. Manfaat Bagi Masyarakat
 1. Meningkatkan pemahaman publik tentang asal-usul dan peruntukan dana negara/daerah.
 2. Memperkuat partisipasi masyarakat dalam mengawasi pelaksanaan anggaran.

1.5 Batasan Masalah

Pembatasan masalah dalam penelitian ini agar mampu memecahkan permasalahan utama, yaitu sebagai berikut:

- a. Ruang Lingkup Kelembagaan
Penelitian hanya difokuskan pada Kementerian/Lembaga (K/L) pemerintah pusat sebagai unit pengguna anggaran APBN.
- b. Objek Klasifikasi Anggaran
Penelitian difokuskan pada klasifikasi anggaran berdasarkan jenis belanja (Barang, Pegawai, Modal, Bansos, dan Belanja Lainnya) serta sumber dana (Rupiah Murni, PNBP, Hibah, Pinjaman, BLU, dan sumber dana lain sesuai regulasi APBN).
- c. Aspek yang Dianalisis
 1. Ketepatan penggunaan klasifikasi jenis belanja serta sumber dana dalam dokumen Daftar Isian Pelaksanaan Anggaran (DIPA).
 2. Dampak kesalahan klasifikasi terhadap akuntabilitas fiskal dan laporan keuangan Kementerian/Lembaga (K/L).
 3. Upaya atau strategi perbaikan klasifikasi anggaran oleh Kementerian/Lembaga (K/L).

d. Periode Penelitian

Penelitian dibatasi pada pelaksanaan anggaran dalam kurun waktu tiga tahun terakhir (Tahun Anggaran 2022-2024), dengan pertimbangan ketersediaan data dan relevansi terhadap regulasi terbaru.

e. Data yang digunakan adalah dokumen APBN, DIPA, laporan keuangan K/L, dan peraturan perundangan yang berlaku.

f. Fokus Evaluasi

Penelitian tidak membahas secara detail aspek politik anggaran, kebijakan makro fiskal, maupun teknis detail pelaksanaan belanja (procurement, kontrak, dsb.), tetapi dibatasi pada klasifikasi anggaran dan dampaknya pada akuntabilitas.

BAB II

TINJAUAN PUSTAKA DAN KERANGKA TEORI

2.1 Konsep Klasifikasi Anggaran

Klasifikasi anggaran merupakan sistem pengelompokan pendapatan dan belanja negara sesuai dengan karakteristik, tujuan, maupun sumber dananya. Menurut *Government Finance Statistics (GFS) Manual* yang diterbitkan oleh IMF, klasifikasi anggaran bertujuan agar laporan keuangan pemerintah dapat disajikan secara konsisten, komparatif, dan dapat dipahami oleh berbagai pemangku kepentingan, baik domestik maupun internasional (PMK 102 /PMK.02/2018 2018)

Pemerintah menetapkan landasan klasifikasi anggaran melalui Kementerian Keuangan yang tercantum dalam Undang-Undang Nomor 17 Tahun 2003 tentang Keuangan Negara, dan Undang-Undang Nomor 1 Tahun 2004 tentang Perbendaharaan Negara, serta diturunkan lebih lanjut dalam Bagan Akun Standar (BAS). BAS menjadi acuan utama dalam pengkodean belanja, pendapatan, transfer, dan pemberian pada APBN(UU No. 17 Tahun 2003).

Khusus pada dimensi belanja, Bagan Akun Standar (BAS) mengklasifikasikan belanja pemerintah ke dalam beberapa kategori belanja, diantaranya: pegawai, barang, modal, bunga, subsidi, hibah, bantuan sosial, serta belanja lainnya. Sementara dari sisi sumber dana, pengelompokan dilakukan berdasarkan asal dana yang membiayai suatu kegiatan, seperti Rupiah Murni, PNBP/BLU, SBSN, PHLN/PDN, maupun hibah.

Dengan adanya klasifikasi yang jelas, setiap transaksi anggaran dapat ditelusuri dengan akurat, sehingga meningkatkan transparansi, memudahkan proses audit, dan memastikan bahwa dana digunakan sesuai peruntukannya.

2.2 Akuntabilitas Fiskal

Akuntabilitas fiskal adalah salah satu landasan utama dalam mengelola keuangan negara. Mardiasmo (2009) menyatakan bahwa akuntabilitas publik yaitu Semua tindakan yang bertanggung jawab kepada pemegang amanah harus disajikan, dilaporkan, dan diumumkan (Putri e Subardjo 2017).

Dalam konteks pengelolaan APBN, akuntabilitas fiskal dicapai dengan membuat laporan keuangan pemerintah yang akurat, jelas, dan mengikuti aturan akuntansi pemerintah. Laporan keuangan tersebut kemudian diaudit oleh BPK untuk memastikan kewajaran penyajian.

Ketepatan klasifikasi jenis belanja dan sumber dana berperan penting dalam menjamin akuntabilitas fiskal. Misalnya, salah dalam mencantumkan klasifikasi belanja barang yang seharusnya belanja modal dapat mengakibatkan aset negara tidak tercatat dalam neraca, sehingga laporan keuangan menjadi salah saji. Selain itu, kesalahan dalam menyebutkan jenis sumber pendanaan yang ada dapat menyebabkan pengeluaran yang tidak tepat pada proyek-proyek yang dibiayai oleh pinjaman atau hibah luar negeri, yang dapat membuat investor kurang yakin akan integritas pemerintah (PMK 62 Tahun 2023).

Dengan demikian, akuntabilitas fiskal bukan hanya menyangkut pertanggungjawaban hasil pembangunan, tetapi juga mencakup akurasi pencatatan setiap transaksi sesuai klasifikasi anggaran yang berlaku.

2.3 Pengelolaan APBN K/L

Pengelolaan APBN pada Kementerian/Lembaga (K/L) adalah bagian menyeluruh dari manajemen keuangan negara dalam mewujudkan efektivitas, efisiensi, transparansi, dan akuntabilitas dalam pelaksanaan program pemerintah. K/L sebagai pengguna anggaran memiliki tanggung jawab penuh terhadap perencanaan, penganggaran, pelaksanaan, penatausahaan, pelaporan, hingga pertanggungjawaban anggaran. Proses ini tidak dapat dilepaskan dari sistem klasifikasi anggaran, yang menentukan bagaimana setiap transaksi dicatat, dikelompokkan, dan dilaporkan (PMK 62 Tahun 2023)

2.4 Klasifikasi Berdasarkan Jenis Belanja

Jenis belanja dalam APBN K/L merupakan kategori ekonomi yang menggambarkan sifat penggunaan anggaran (PMK 102 /PMK.02/2018 2018). Sesuai dengan regulasi penganggaran, jenis belanja utama yang digunakan dalam dokumen anggaran adalah:

- a. **Belanja Pegawai**, merupakan kompensasi gaji beserta tunjangan untuk Aparatur Sipil Negara (ASN).
- b. **Belanja Barang**, yang merupakan belanja operasional dan non-operasional penunjang kegiatan K/L.
- c. **Belanja Modal**, Belanja Modal: Ini adalah biaya untuk membeli aset tetap dan barang yang akan berguna lebih dari satu tahun anggaran.
- d. **Belanja Bantuan Sosial (Bansos)**, Belanja Bantuan Sosial (Bansos), yang merupakan sumbangan finansial atau non-finansial kepada masyarakat untuk membantu atau melindungi mereka.
- e. **Belanja Lainnya**, Pengeluaran lain yang tidak termasuk dalam kategori di atas tetapi tunduk pada peraturan..

Klasifikasi jenis belanja ini berfungsi sebagai instrumen untuk mengukur efisiensi dan efektivitas program, serta sebagai basis penyusunan laporan keuangan pemerintah. Kesalahan klasifikasi dapat menimbulkan distorsi informasi, baik dalam laporan realisasi anggaran maupun laporan keuangan, sehingga berimplikasi pada kualitas pengambilan keputusan.

2.5 Klasifikasi Berdasarkan Sumber Dana

Selain menurut jenis belanja, pengelolaan APBN K/L juga memerlukan klasifikasi berdasarkan sumber dana, yang menggambarkan asal pembiayaan suatu kegiatan (PMK 160/PMK.02/2012) Beberapa sumber dana utama adalah:

- a. **Rupiah Murni (RM)**, yakni dana yang berasal dari penerimaan dalam negeri (pajak dan non-pajak) tanpa ikatan penggunaan khusus.
- b. **Penerimaan Negara Bukan Pajak (PNBP)**, yang bersumber dari layanan K/L kepada masyarakat, seperti biaya pendidikan, pelayanan kesehatan, atau jasa lainnya.

- c. **Pinjaman dan Hibah Luar Negeri (PHLN)**, yang dialokasikan untuk mendukung pembiayaan program tertentu sesuai perjanjian dengan mitra internasional.
- d. **Badan Layanan Umum (BLU)**, yang memiliki fleksibilitas dalam mengelola pendapatan dan belanja berdasarkan pola keuangan tersendiri.
- e. **Sumber Dana Lain**, sesuai ketentuan APBN (misalnya SBSN/Surat Berharga Sukuk Negara untuk pembiayaan proyek).

Pengelompokan menurut sumber dana menjadi penting untuk memastikan **kesesuaian penggunaan anggaran dengan asal pembiayaan**, menjaga kepatuhan pada peraturan, serta memudahkan pemantauan kinerja fiskal.

2.6 Implikasi dalam Pengelolaan APBN K/L

Klasifikasi anggaran berdasarkan jenis belanja dan sumber dana memiliki implikasi strategis dalam pengelolaan APBN K/L:

- a. **Perencanaan dan Penganggaran**, yaitu memberikan kerangka kerja yang jelas bagi perancangan RKA-K/L dan penetapan DIPA.
- b. **Pelaksanaan Anggaran**, yaitu memastikan penggunaan dana sesuai klasifikasi, sehingga mencegah terjadinya deviasi penggunaan belanja.
- c. **Akuntabilitas Fiskal**, yaitu mendukung dalam memberikan informasi keuangan yang akurat dan dapat dipercaya dan sesuai dengan Standar Akuntansi Pemerintahan (SAP).
- d. **Evaluasi Kinerja**, yaitu memudahkan penilaian efisiensi dan efektivitas program berdasarkan alokasi belanja dan sumber pendanaannya.
- e. **Audit dan Pengawasan**, yaitu menjadi basis bagi Badan Pemeriksa Keuangan (BPK) maupun Aparat Pengawas Intern Pemerintah (APIP) dalam menilai kepatuhan dan kualitas tata kelola anggaran.

Dengan demikian, pengelolaan APBN K/L tidak sekadar soal besaran anggaran, tetapi juga bagaimana **klasifikasi anggaran** dapat mendukung akuntabilitas, transparansi, serta efektivitas pelaksanaan program pemerintah.

2.7 Pemodelan Klasifikasi

Klasifikasi adalah proses menciptakan model yang dapat memisahkan dan mengidentifikasi berbagai jenis data atau konsep sehingga model tersebut dapat menebak jenis yang tidak diketahui dari suatu item yang dapat diamati (Sihombing e Yuliati 2021). Proses klasifikasi terkait erat dengan penambangan data, berfungsi sebagai metode untuk mengenali pola dan menggali informasi dari sebuah data mentah menggunakan algoritma pembelajaran mesin, sehingga memfasilitasi pengambilan keputusan yang tepat. Analisis klasifikasi data melibatkan pembuatan model yang menjelaskan penggolongan data, yang digunakan dalam memprediksi kelas atau kategori data yang belum diproses. (Louis M e Dian B 2022).

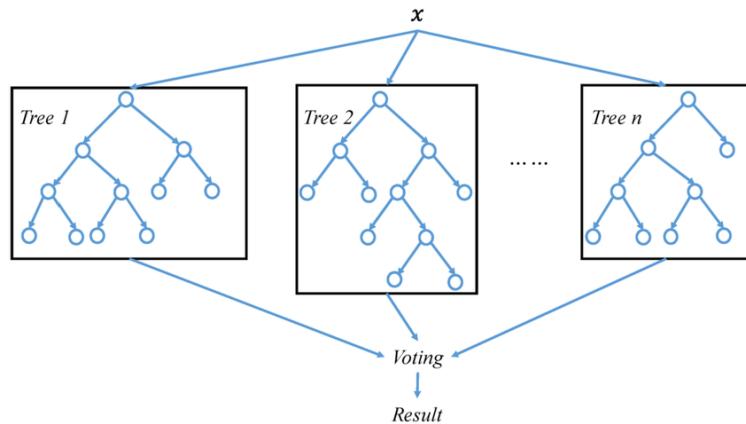
2.8 Random Forest Classifier (RFC)

Random Forest digunakan yang secara konsisten dalam menghasilkan output yang menguntungkan. Random Forest terdiri dari beberapa pohon, masing-masing dikembangkan menggunakan metode keacakan dan proses lainnya. Cara kerjanya adalah dengan membentuk beberapa pohon keputusan selama fase pelatihan dan menghasilkan label yang mewakili kelas yang sering muncul atau perkiraan rata-rata setiap pohon (Khasanah et al. 2021). Pendekatan ini merupakan varian dari algoritma Pohon Keputusan, termasuk ensemble pohon keputusan yang digunakan sebagai pengklasifikasi dasar terpadu. Metode ini digunakan untuk klasifikasi dan analisis regresi. (Louis M e Dian B 2022).

Tahapan pada algoritma Random Forest (Khasanah *et al.* 2021) diantaranya:

- a. Menentukan jumlah pohon (k) yang akan dibuat. Jumlah k ditentukan dari total karakteristik m , dengan k kurang dari m .
- b. Probabilitas sampel dari kumpulan data N pengamatan (dengan penggantian) untuk setiap pohon.
- c. Subset prediktor acak dipilih untuk setiap pohon. Di mana m kurang dari p , dengan p adalah jumlah variabel prediktor.
- d. Tahap kedua dan ketiga dilakukan hingga k pohon dihasilkan.
- e. Dalam klasifikasi, hasil prediksi Random Forest diperoleh dari suara mayoritas hasil klasifikasi dari ensemble pohon.

Random Forest membangun setiap pohon dari sampel acak data pelatihan menggunakan pengambilan sampel bootstrap, sama seperti metode bagging normal lainnya. Pendekatan ini juga membahas bagaimana fitur-fitur dapat bersifat acak. Saat membangun pohon, pendekatan bagging standar akan mengambil pemisahan terbaik dari semua fitur saat membagi sebuah node.



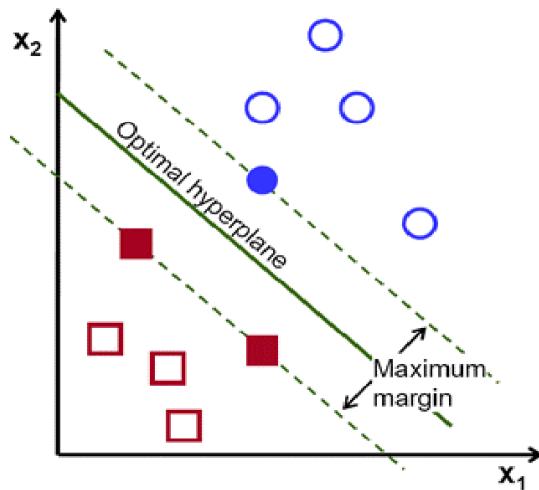
Gambar 2. 1 Arsitektur Random Forest

Dalam Random Forest, pemisahan optimal dipilih dari sekelompok fitur yang acak. Hal-hal acak ini membuat pohon dasar lebih berbeda, menurunkan varians, dan membuat model secara keseluruhan lebih baik (Wang *et al.* 2019).

2.9 Support Vector Machine (SVM)

Support Vector Machines (SVM) adalah metode pembelajaran mesin baru yang digunakan untuk analisis prediktif dalam skenario klasifikasi dan regresi. SVM mengidentifikasi pola melalui metodologi statistic. Support Vector Machines (SVM) adalah cara baru untuk menerapkan pembelajaran mesin untuk analisis prediktif dalam situasi klasifikasi dan regresi. SVM menggunakan metode statistik untuk menemukan pola. SVM mengidentifikasi pola melalui metodologi statistik (Tantika e Kudus 2022). Support Vector Machine (SVM) akan mengidentifikasi vektor pendukung optimal yang memisahkan dua kelas berbeda dengan memaksimalkan margin (Tantika e Kudus 2022). Teknik Support Vector Machine (SVM) digunakan dalam proses klasifikasi untuk mengidentifikasi pola dalam data pelatihan berlabel dan kemudian menghasilkan model pembelajaran mesin (Husada e Paramita 2021). Support Vector Machine (SVM) adalah pendekatan prediktif yang dikembangkan oleh Vapnik pada tahun 1995, yang dapat diterapkan pada

tugas klasifikasi dan regresi. Pendekatan ini merupakan teknik pembelajaran mesin yang dikategorikan sebagai pembelajaran terawasi. SVM dapat diterapkan untuk klasifikasi data linier dan nonlinier. Konsep dasarnya adalah mengoptimalkan *hyperplane* (Sihombing e Yuliaty 2021).



Gambar 2. 2 Optimal Hyperplane dengan Maximum Margin.

Untuk menentukan hyperplane yang paling efisien untuk mengklasifikasikan data ke dalam dua kategori, margin *hyperplane* dihitung untuk mengidentifikasi titik maksimum (Husada e Paramita 2021). Dalam menentukan *hyperplane* pada SVM, dapat menggunakan persamaan (1).

$$(w \cdot x_i) + b = 0 \quad (1).$$

Di dalam data x_i , yang termasuk pada kelas -1 dapat dirumuskan seperti pada persamaan (2).

$$(w \cdot x_i + b) \leq 1, y_i = -1 \quad (2).$$

Sedangkan data x_i yang termasuk pada kelas +1 dapat dirumuskan seperti pada persamaan (3).

$$(w \cdot x_i + b) \geq 1, y_i = 1 \quad (3).$$

2.10 Evaluation Model

Evaluasi kinerja model kategorisasi didasarkan pada penilaian item yang akurat dan tidak akurat. Penelitian ini menggunakan *Confusion matrix* untuk menghitung kinerja klasifikasi, termasuk perhitungan hasil klasifikasi aktual yang diprediksi (Louis M and Dian B 2022). Penelitian ini dilakukan untuk menentukan metodologi pembagian dan kategorisasi dataset yang paling efektif, yang dinilai menggunakan indikator kinerja klasifikasi. Penelitian ini menggunakan matriks konfusi sebagai ukuran kinerja klasifikasi. Matriks konfusi adalah cara yang baik untuk melihat seberapa baik sistem klasifikasi dapat membedakan antara objek dari kelas yang berbeda (Sihombing e Yuliaty 2021).

Akurasi

Akurasi mengukur efektivitas model klasifikasi, yang mewakili proporsi instance yang diprediksi dengan benar relatif terhadap keseluruhan dataset. Kriteria untuk menentukan klasifikasi yaitu dengan menggunakan confusion matrix (Dewi, Purwono, e Kurniawan Dwi 2022) yang dapat dilihat pada Tabel 2.1.

Tabel 2. 1 Confusion Matrix

Kelas	Diklasifikasikan sebagai Positif	Diklasifikasikan sebagai Negatif
Positive	True Positive (TP)	False Positive (FP)
Negative	True Negative (TN)	False Negative (FN)

Pada Tabel 2.1 menjelaskan tentang kategorisasi *confusion matrix*. True positive (TP) menunjukkan bahwa model klasifikasi secara akurat mengidentifikasi jumlah true positive. True negative (TN) menunjukkan bahwa model klasifikasi secara akurat mengidentifikasi jumlah instance negatif. False positive (FP) menunjukkan bahwa model klasifikasi secara keliru mengkategorikan instance negatif. False negative (FN) menunjukkan bahwa model klasifikasi secara keliru mengkategorikan instance positif sebagai negatif. Akurasi mengukur kinerja model klasifikasi, yang mewakili proporsi instance yang diprediksi dengan benar relatif

terhadap keseluruhan dataset. Persamaan untuk menghitung ketelitian dapat dilihat pada Persamaan 1.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

Pengukuran kinerja mencakup presisi, recall, dan F1-score. Presisi adalah rasio yang menguntungkan atau tingkat keandalan, khususnya persentase perkiraan positif yang akurat terhadap jumlah keseluruhan prediksi positif. (Dewi, Purwono, e Kurniawan Dwi 2022). Persamaan untuk menghitung presisi dapat dilihat pada Persamaan 2.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall terkadang disebut sebagai tingkat positif sejati atau sensitivitas. Ini juga dikenal sebagai keandalan model dalam mengidentifikasi data yang diberi label positif secara akurat (Dewi, Purwono, e Kurniawan Dwi 2022). Persamaan untuk menghitung recall dapat dilihat pada persamaan 3.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F1-score F1 merangkum hasil perhitungan akurasi dan recall dengan membentuk rata-rata harmonik (Dewi, Purwono, e Kurniawan Dwi 2022). Persamaan untuk menghitung F1-score dapat dilihat pada persamaan 4.

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

2.11 Kerangka Teori

Untuk memperoleh pemahaman tentang variable-variabel penting dalam penelitian ini, maka diperlukan kerangka teori. Teori-teori ini digunakan sebagai acuan saat merancang metodologi, membangun hipotesis, dan menganalisis dan menginterpretasikan temuan penelitian. Penelitian berikut yang berkaitan dengan

penelitian ini digunakan untuk memperoleh pemahaman tentang metode yang digunakan di antaranya:

Nurul Khasanah, Dkk (2021)

Melakukan penelitian yang berkonsentrasi pada klasifikasi kanker kulit menggunakan metode pengajaran mesin, terutama algoritma Random Forest. Jenis kanker kulit (ganas, malignant, atau jinak/benign) digunakan sebagai variabel independen dan fitur ekstraksi warna (histogram), bentuk (moments hue), dan tekstur (Haralick). Teori yang mendasari penelitian ini termasuk teori klasifikasi dalam pembelajaran mesin, teori ekstraksi fitur dari gambar digital, dan evaluasi kinerja model dengan menggunakan metrik seperti akurasi, ketepatan, recall, dan kappa statistic. Proses ekstraksi fitur dari gambar kulit memungkinkan algoritma Random Forest untuk melakukan klasifikasi untuk menunjukkan hubungan antar variabel. Tujuan penelitian adalah menciptakan sistem klasifikasi yang efisien dan akurat untuk mendeteksi berbagai kanker kulit tanpa melakukan biopsi invasif. Hasil penelitian menunjukkan akurasi ekstraksi fitur hue moment sebesar 84,2% lebih tinggi dari hasil penelitian sebelumnya. Ini membangun kerangka teori bahwa penggabungan ekstraksi fitur multidomain (warna, bentuk, dan tekstur) dengan algoritma pembelajaran kelompok dapat meningkatkan hasil klasifikasi kanker kulit (Khasanah *et al.* 2021).

Pramesti Dewi, Dkk (2022)

Penelitian ini fokus pada penerapan teknologi pembelajaran mesin, khususnya algoritma Random Forest, untuk mengkategorikan jenis hipertensi berdasarkan karakteristik individu seperti jenis kelamin, usia, tinggi badan, berat badan, indeks massa tubuh (IMT), tekanan darah sistolik dan diastolik, serta denyut jantung. Kerangka teoritis penelitian ini mencakup pembelajaran terawasi, algoritma klasifikasi ensemble Random Forest, dan metodologi optimasi hiperparameter yang menggunakan GridSearch CV dengan validasi silang k-fold. Penelitian ini mengkaji hubungan prediktif antar variabel, dengan memanfaatkan karakteristik pribadi sebagai variabel independen untuk memprakirakan variabel dependen, khususnya klasifikasi hipertensi (normal, prahipertensi, hipertensi

stadium 1, dan hipertensi stadium 2). Tujuan penelitian ini adalah untuk meningkatkan akurasi klasifikasi dengan menyempurnakan hiperparameter dalam model Random Forest. Hasilnya menunjukkan peningkatan substansial dalam akurasi model, meningkat dari 72,3% menjadi 86,1% pasca-penerapan GridSearch CV. Parameter optimal yang diidentifikasi adalah kedalaman maksimum=80, fitur maksimum = 3, minimal tiga sampel, pembagian sampel minimum=8, dan n estimator sama dengan 1000. Kerangka teoritis yang dibangun menggarisbawahi penggabungan prapemrosesan data, pemodelan klasifikasi, dan optimasi hiperparameter untuk meningkatkan kinerja prediktif dalam domain kesehatan, khususnya dalam klasifikasi hipertensi (Dewi, Purwono, e Kurniawan Dwi, 2022).

Yulia Aryani (2021)

Topik penelitian ini berfokus pada klasifikasi sinyal radar dari ionosfer menjadi dua kategori, yaitu "**good**" (berguna) dan "**bad**" (tidak berguna), dengan menggunakan tiga algoritma *machine learning*. Variabel prediktor terdiri dari 34 atribut kontinu yang merepresentasikan karakteristik pantulan radar, sementara variabel target adalah atribut ke-35 yang bersifat kategorik. Teori yang mendasari penelitian ini meliputi **Support Vector Machine (SVM)** dengan pemilihan *kernel* yang optimal, **Naïve Bayes** sebagai metode probabilistik berdasarkan teorema Bayes, dan **Random Forest** sebagai ensemble classifier yang menggabungkan banyak pohon keputusan. Hubungan antar variabel bersifat prediktif, di mana atribut radar digunakan untuk memprediksi kategori pantulan. Tujuan penelitian adalah membandingkan kinerja ketiga metode tersebut dalam mengklasifikasikan data ionosfer dari *UCI Machine Learning Repository* guna menemukan model dengan akurasi, presisi, dan recall terbaik. Hasil penelitian menunjukkan bahwa **Random Forest** mencapai kinerja tertinggi dengan akurasi 90,57%, presisi 94,12%, dan sensitivitas 91,43%, mengungguli SVM dan Naïve Bayes, sehingga dapat dijadikan sebagai kerangka teori yang efektif untuk klasifikasi sinyal radar ionosfer (Aryani e Wijayanto 2021).

Jalis Dwi Muthohhar, dkk (2023)

Berdasarkan jurnal tersebut, penelitian ini fokus pada analisis klasifikasi tiga klasifikasi algoritma dalam mesin pembelajaran, yaitu Decision Tree, Naive Bayes, dan Random Forest Classifier, untuk memprediksi penyakit jantung menggunakan dataset Heart Attack Analysis & Prediction dari Kaggle. Variabel yang digunakan meliputi 12 atribut antara lain umur, jenis kelamin, jenis nyeri dada, tekanan darah, kadar kolesterol, gula darah puasa, hasil EKG, denyut jantung maksimal, nyeri dada saat berolahraga, depresi ST, kemiringan segmen ST, dan variabel sasaran yaitu adanya penyakit jantung. Teori yang mendasari penelitian ini meliputi teorema Bayes pada Naive Bayes, struktur pohon keputusan pada Decision Tree, dan metode pemungutan suara ansambel pada Random Forest. Hubungan antar variabel diperiksa melalui pemilihan fitur menggunakan RandomForestRegressor untuk data numerik dan metode KBest untuk data kategorikal, yang menunjukkan bahwa atribut RestingECG memiliki dampak minimal dan oleh karena itu dikecualikan. Tujuan dari penelitian ini adalah untuk menemukan model terbaik dengan parameter F1-Score setelah melakukan fine-tuning dengan Random Search dan Grid Search CV untuk menghindari overfitting. Hasil penelitian yang dilakukan menunjukkan bahwa Random Forest Classifier mencapai kinerja tertinggi dengan F1-Score sebesar 0,868 setelah penyetelan, melampaui Decision Tree (0,844) dan Naive Bayes (0,85) (Muthohhar e Prihanto 2023).

Widya Apriliah, dkk (2021)

Penelitian ini berfokus pada pengembangan model prediksi diabetes mellitus menggunakan pendekatan machine learning. Variabel yang digunakan mencakup 17 atribut klinis dan demografis, termasuk usia, jenis kelamin, dan gejala seperti poliuria, polidipsia, penurunan berat badan mendadak, kelemahan, antara lain, dengan variabel target berupa klasifikasi diabetes sebagai positif atau negatif. Teori yang mendasari penelitian ini adalah algoritma klasifikasi dalam penambangan data, khususnya Support Vector Machine (SVM), Naive Bayes, dan Random Forest, yang diuji untuk mengidentifikasi pola hubungan antara gejala dan faktor risiko yang terkait dengan kemungkinan terkena diabetes. Tujuan penelitian

ini adalah untuk mengevaluasi kinerja tiga algoritma untuk menentukan model dengan akurasi tertinggi dalam prediksi dini diabetes. Hasil penelitian menghasilkan Random Forest menunjukkan kinerja yang unggul, mencapai akurasi 97,88%, presisi 0,979, recall 0,979, dan nilai ROC 0,998, melampaui SVM dan Naive Bayes. Akibatnya, ini dapat digunakan sebagai kerangka teoritis yang efektif untuk pengembangan sistem prediksi diabetes berbasis pembelajaran mesin (Apriliah *et al.* 2021).

Louis Madaerdo Sotarjua, dkk (2022)

Studi ini didasarkan pada jurnal dan berfokus pada perbandingan kinerja tiga algoritma pembelajaran mesin K-Nearest Neighbor (KNN), Decision Tree, dan Random Forest dalam mengklasifikasikan promosi jabatan karyawan menggunakan data yang memiliki kelas yang tidak seimbang. Variabel independen yang digunakan adalah wilayah, pendidikan, jumlah pelatihan, usia, peringkat tahun sebelumnya, lama masa kerja, penghargaan yang diraih, dan skor pelatihan rata-rata. Variabel dependen adalah status promosi jabatan karyawan (dipromosikan atau tidak). Teori yang mendasari penelitian ini mencakup konsep penambangan data dan klasifikasi dalam pembelajaran mesin, serta teknik untuk mengatasi ketidakseimbangan kelas melalui Synthetic Minority Over-sampling Technique (SMOTE). Hubungan antar variabel dijelaskan melalui proses pemilihan fitur menggunakan uji statistik ANOVA untuk data numerik dan uji Chi-Squared untuk data kategorikal, yang mencoba menemukan faktor-faktor yang paling berpengaruh terhadap hasil klasifikasi. Tujuan penelitian ini adalah untuk mengevaluasi efektivitas tiga algoritma pada data tidak seimbang dan mengidentifikasi metode yang paling optimal berdasarkan kriteria akurasi, presisi, recall, dan skor F1. Hasil penelitian menunjukkan bahwa algoritma KNN mencapai kinerja ideal, dengan meraih skor F1 sebesar 85,88% pada data uji, tanpa menunjukkan tanda-tanda overfitting maupun underfitting. Dengan demikian, algoritma ini dapat digunakan sebagai model yang andal untuk klasifikasi promosi karyawan dalam konteks ketidakseimbangan kelas (Louis M e Dian B, 2022).

Dimas Aditiya, dkk (2023)

Berdasarkan jurnal tersebut, penelitian ini berfokus pada pengujian efektivitas penerapan klasifikasi pembelajaran mesin untuk menganalisis survei kepuasan pelanggan Maskapai Penerbangan X. Variabel yang digunakan meliputi atribut layanan seperti Layanan Wifi Dalam Pesawat, Kenyamanan Kursi, Makanan dan Minuman, Penanganan Bagasi, serta variabel demografis dan operasional seperti Jenis Pelanggan, Jenis Perjalanan, Kelas, dan Jarak Penerbangan. Penelitian ini didasarkan pada teori penambangan data dan pembelajaran terawasi, memanfaatkan algoritma klasifikasi seperti K-Nearest Neighbor (KNN), Decision Tree, dan Random Forest. Hubungan antar variabel dijelaskan melalui proses klasifikasi yang menghubungkan fitur layanan dengan variabel target, yaitu kepuasan (puas atau netral/tidak puas). Tujuan penelitian ini adalah untuk membandingkan kinerja ketiga model pembelajaran mesin dalam mengklasifikasikan tingkat kepuasan pelanggan dan memberikan rekomendasi untuk peningkatan layanan. Temuan penelitian menunjukkan bahwa Decision Tree dan Random Forest adalah model yang paling efektif, mencapai akurasi masing-masing 94,5% dan 96%, berdasarkan validasi silang. Akibatnya, keduanya dapat berfungsi sebagai kerangka kerja teoritis untuk implementasi sistem klasifikasi kepuasan pelanggan yang otomatis dan efisien (Aditiya e Latifa 2023).

Rina Silviany Tantika (2022)

Tujuan dari penelitian ini adalah mengklasifikasikan pasien penyakit tiroid ke dalam tiga kategori, yaitu euthyroidism (normal), hypothyroidism, dan hyperthyroidism, dengan menggunakan metode Support Vector Machine (SVM) multiclass. Variabel prediktor yang digunakan meliputi lima indikator medis: X1 (persentase hasil uji asam T3), X2 (total serum thyroxin), X3 (total serum triiodothyronin), X4 (hormon basal thyroid stimulating), dan X5 (perbedaan maksimal absolut nilai TSH setelah disuntik), sementara variabel target (Y) adalah kelas kondisi tiroid. Teori utama yang mendasari penelitian ini adalah SVM yang awalnya dirancang untuk klasifikasi biner, namun dikembangkan untuk masalah multiclass melalui dua pendekatan: One Against All (OAA) dan One Against One (OAO). Hubungan antar variabel dijembatani melalui proses normalisasi z-score

untuk menyeragamkan skala data, sehingga meningkatkan kinerja model. Tujuan penelitian adalah membandingkan akurasi kedua pendekatan SVM tersebut dalam mengklasifikasikan data tiroid. Hasil penelitian menunjukkan bahwa pendekatan OAO dengan kernel linear menghasilkan akurasi tertinggi sebesar 99,53%, jauh lebih unggul dibandingkan OAA yang hanya mencapai 87,69%, sehingga dapat disimpulkan bahwa OAO lebih efektif untuk klasifikasi multiclass pada data medis tiroid (Tantika e Kudus 2022).

Hendry Cipta Husada, dkk (2021)

Penelitian ini berfokus pada penerapan analisis sentimen terhadap opini publik mengenai maskapai penerbangan yang diunggah di platform Twitter. Variabel utama yang digunakan mencakup data teks berupa *tweet* sebagai variabel independen, serta kategori sentimen (positif, negatif, netral) sebagai variabel dependen. Teori yang mendasari penelitian ini meliputi *text mining* dengan tahapan *preprocessing* seperti *cleansing*, *case folding*, *tokenizing*, *filtering*, dan *stemming*, dilanjutkan dengan pembobotan kata menggunakan metode TF-IDF untuk mengubah data teks menjadi representasi numerik. Algoritma *Support Vector Machine* (SVM) diterapkan sebagai metode klasifikasi dengan berbagai fungsi kernel (*linear*, *polynomial*, *RBF*, *sigmoid*) dan parameter tuning untuk meningkatkan kinerja model. Hubungan antar variabel terlihat dari bagaimana data teks yang telah diproses dan diberi bobot digunakan oleh SVM untuk memetakan pola sentimen. Tujuan penelitian adalah menghasilkan model klasifikasi otomatis yang akurat dalam mengkategorikan opini pengguna Twitter, sehingga dapat menjadi bahan evaluasi bagi maskapai penerbangan. Hasil penelitian menunjukkan bahwa kernel RBF dengan parameter $C = 10$ dan $\gamma = 1$ memberikan kinerja terbaik dengan akurasi 84,37%, serta nilai *precision*, *recall*, dan *f-score* di atas 84%, mengungguli kernel lainnya dan algoritma pembanding seperti Naïve Bayes (Husada e Paramita 2021).

Hidayat, dkk (2023)

Penelitian ini, berdasarkan jurnal "Heart Disease Classification Using Random Forest Classifier", bertujuan untuk meningkatkan presisi deteksi penyakit

jantung dengan menggunakan algoritma Random Forest, yang dilengkapi dengan teknik pra-pemrosesan data. Variabel yang digunakan mencakup karakteristik klinis dari dataset penyakit jantung yang diperoleh dari Kaggle, termasuk usia, jenis kelamin, kadar kolesterol, dan detak jantung maksimum. Variabel-variabel ini diproses melalui tahapan penanganan nilai yang hilang dan normalisasi data menggunakan metode Min-Max. Teori utama yang mendasari penelitian ini adalah algoritma Random Forest, sebuah komponen pembelajaran ensemble, yang menghasilkan prediksi yang lebih stabil dan presisi dengan membangun beberapa pohon keputusan. Kami mengamati bagaimana variabel-variabel saling terkait dengan menjalankan sejumlah eksperimen yang mengubah parameter seperti `n_estimators`, `max_depth`, dan `min_samples_leaf`. Kami juga mengamati dua skenario pembagian data: 80:20 dan 70:30. Tujuan utamanya adalah untuk meningkatkan kinerja model agar dapat mencapai akurasi yang lebih baik daripada 90% pada penelitian sebelumnya. Hasil studi menunjukkan bahwa kombinasi pra-pemrosesan data, normalisasi, dan penyetelan parameter meningkatkan akurasi hingga 94% pada pembagian data 80:20. Ini berarti metode ini efektif untuk klasifikasi biner penyakit jantung dan dapat digunakan untuk membangun kerangka kerja teoretis bagi sistem pendukung diagnosis medis (Hidayat, Sunyoto, e Al Fatta 2023).

Istiqomatul Fajriyah Yuliati, dkk (2021)

Penelitian ini mengkaji pemanfaatan teknik pembelajaran mesin untuk mengkategorikan risiko Berat Badan Lahir Rendah (BBLR) di Indonesia, menggunakan data dari Survei Demografi dan Kesehatan Indonesia (SDKI) 2017. Variabel respons adalah status berat badan lahir bayi (normal atau rendah), sedangkan variabel prediktor meliputi faktor obstetrik seperti jarak dan urutan kelahiran, perawatan prenatal, dan faktor demografi dan perilaku termasuk usia ibu, pendidikan, status ekonomi, kebiasaan merokok, dan sumber air minum. Pekerjaan ini didasarkan pada metode klasifikasi dalam pembelajaran mesin, termasuk Pohon Klasifikasi dan Regresi (CART), Hutan Acak, Naïve Bayes, dan Mesin Vektor Pendukung (SVM), yang menggunakan teknik resampling untuk mengatasi ketidakseimbangan data. Korelasi antar variabel diperiksa untuk menentukan

faktor-faktor utama yang memengaruhi terjadinya berat badan lahir rendah (BBLR). Penelitian ini bertujuan untuk mengevaluasi kemanjuran berbagai model kategorisasi untuk mengidentifikasi model yang paling akurat untuk mendeteksi risiko berat badan lahir rendah, terutama pada kelas minoritas. Temuan ini menunjukkan bahwa model Hutan Acak yang memanfaatkan kedua strategi pengambilan sampel menghasilkan kinerja optimal, sebagaimana dibuktikan oleh nilai sensitivitas, spesifisitas, G-mean, dan AUC tertinggi, dengan tiga faktor paling signifikan adalah jarak dan urutan kelahiran, perawatan prenatal, dan usia ibu. Temuan ini dapat menjadi dasar untuk membangun kerangka kerja teoretis bagi intervensi kesehatan yang lebih terfokus yang bertujuan untuk mengurangi insiden bayi berat lahir rendah (Sihombing e Yuliati 2021).

Berdasarkan kesepuluh jurnal yang dirangkum, dapat ditarik beberapa benang merah dan poin penting mengenai penerapan algoritma machine learning, khususnya Random Forest, serta perbandingannya dengan algoritma lain. Dalam hampir semua penelitian yang membandingkan beberapa algoritma, Random Forest (RF) konsisten menunjukkan kinerja terbaik. Berikut adalah sintesis dan analisisnya:

1. Klasifikasi Medis: RF unggul dalam memprediksi diabetes (Akurasi 97.88%), penyakit jantung (F1-Score 0.868; Akurasi 94%), dan kanker kulit (Akurasi 84.2%).
2. Klasifikasi Sinyal: RF juga terbukti paling akurat (90.57%) dalam mengklasifikasikan sinyal radar ionosfer dibandingkan SVM dan Naïve Bayes.
3. Alasan Keunggulan: Kelebihan RF terletak pada metode ensemble-nya, yang menggabungkan banyak pohon keputusan untuk mengurangi overfitting dan meningkatkan stabilitas prediksi. Beberapa penelitian (seperti Dewi 2022 dan Hidayat 2023) juga menekankan bahwa kinerja RF dapat dioptimalkan lebih lanjut melalui tuning hyperparameter.

Tabel 2. 2 Perbandingan Akurasi Penelitian Terkait

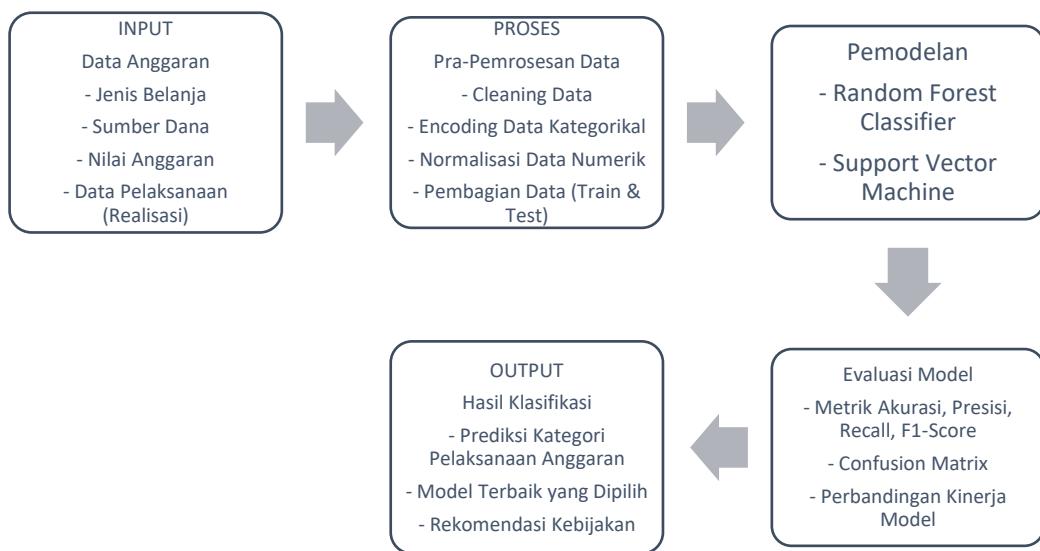
Peneliti	Thn	Model	Variabel	Metode	Akurasi	Keterangan
Nurul Khasanah, dkk	2021	Random Forest	Fitur ekstraksi warna, bentuk, tekstur	Ekstraksi fitur gambar	84.2%	Klasifikasi kanker kulit
Pramesti Dewi, dkk	2022	Random Forest	Jenis kelamin, usia, IMT, tekanan darah, denyut jantung	GridSearch CV, k-fold validation	86.1%	Klasifikasi hipertensi
Yulia Aryani	2021	RFC vs SVM vs Naïve Bayes	34 atribut radar	Perbandingan 3 algoritma	90.57%	Klasifikasi sinyal ionosfer
Jalis Dwi Muthohhar, dkk	2023	RFC vs DT vs Naïve Bayes	12 atribut klinis jantung	Feature selection, Random Search	F1: 86.8%	Prediksi penyakit jantung
Widya Apriliah, dkk	2021	RFC vs SVM vs Naïve Bayes	17 atribut klinis diabetes	Perbandingan 3 algoritma	97.88%	Prediksi diabetes early stage
Louis Madaerdo Sotarjua, dkk	2022	KNN vs DT vs RFC	8 atribut karyawan	SMOTE, ANOVA, Chi-Squared	F1: 85.88%	Klasifikasi promosi jabatan
Dimas Aditiya, dkk	2023	KNN vs DT vs RFC	8 atribut layanan penerbangan	Cross-validation	96%	Analisis kepuasan pelanggan
Rina Silviany Tantika	2022	SVM Multiclass	5 indikator medis tiroid	Z-score, OAA vs OAO	99.53%	Klasifikasi penyakit tiroid
Hendry Cipta Husada, dkk	2021	SVM vs Naïve Bayes	Data teks tweet	TF-IDF, text mining	84.37%	Analisis sentimen Twitter
Hidayat, dkk	2023	Random Forest	Karakteristik klinis jantung	Min-Max normalization	94%	Klasifikasi penyakit jantung

Random Forest adalah algoritma yang sangat kuat yang sering menjadi pilihan terbaik untuk tugas klasifikasi di berbagai bidang, terutama di bidang kesehatan, dan kumpulan jurnal ini memberikan kerangka teoritis dan praktis yang kuat. Keunggulan ini, bagaimanapun, tidak mutlak. Sebuah model sangat bergantung pada "tiga tindakan" untuk berhasil:

1. Pilihan algoritma yang sesuai dengan masalah dan karakteristik data.
2. Pra-pemrosesan data yang teliti untuk "membersihkan" dan menyiapkan data.

3. Tuning hyperparameter dan teknik pengelolaan data seperti SMOTE untuk kasus ketidakseimbangan kelas digunakan untuk mengoptimalkan model.

Kerangka teori ini disusun dalam tiga tahap utama (Input, Proses, Output) untuk memberikan pijakan yang jelas bagi penelitian. Konsep dasar dari penelitian yang akan dilakukan Adalah mengklasifikasikan atau memprediksi pelaksanaan anggaran (sebagai variabel target/output) berdasarkan jenis belanja dan sumber dana (sebagai variabel input/fitur). Klasifikasi ini menggunakan model Machine Learning untuk menemukan pola historis yang dapat membantu dalam pengambilan keputusan dan evaluasi keuangan.



Gambar 2. 3 Kerangka Teori

Dari Gambar 2.3 menggambarkan kerangka alur system yang terbagi dalam tiga tahap yaitu:

1. INPUT (Data Masukan)

- **Data Anggaran:** Data historis anggaran yang mencakup:
 - **Jenis Belanja:** Kategori pengeluaran (contoh: belanja pegawai, belanja barang, belanja modal)
 - **Sumber Dana:** Asal dana (contoh: APBD, APBN, Dana Desa)
 - **Nilai Anggaran:** Besaran anggaran yang dialokasikan
 - **Data Pelaksanaan:** Realisasi anggaran yang telah terjadi (sebagai variabel target)

2. PROSES

- **Pra-Pemrosesan Data:**
 - **Cleaning Data:** Menangani missing value, outlier, dll.
 - **Encoding Data Kategorikal:** Mengubah data kategorikal (seperti jenis belanja dan sumber dana) menjadi numerik.
 - **Normalisasi Data Numerik:** Menyamakan skala data numerik (seperti nilai anggaran) agar tidak bias.
 - **Pembagian Data:** Membagi data menjadi data latih (train) dan data uji (test) dengan rasio tertentu (contoh: 80:20).
- **Pemodelan:**
 - **Random Forest Classifier:** Membangun model dengan algoritma ensemble dari pohon keputusan.
 - **Support Vector Machine:** Membangun model dengan mencari hyperplane terbaik untuk pemisah kelas.
- **Evaluasi Model:**
 - Menggunakan metrik seperti akurasi, presisi, recall, dan F1-score.
 - Membuat confusion matrix untuk melihat kinerja model per kelas.
 - Membandingkan kinerja kedua model untuk menentukan model terbaik.

3. OUTPUT (Keluaran)

- **Hasil Klasifikasi:** Prediksi kategori pelaksanaan anggaran (contoh: Tepat Waktu, Terlambat, Optimal).
- **Model Terbaik:** Model yang dipilih berdasarkan evaluasi kinerja.
- **Rekomendasi Kebijakan:** Insight yang dapat digunakan untuk perbaikan proses anggaran.

Pada Tabel 2.2 di bawah ini dijelaskan lebih detail terkait tahapan-tahapannya dalam mengklasifikasikan atau memprediksi pelaksanaan anggaran (sebagai variabel target/output) berdasarkan jenis belanja dan sumber dana (sebagai variabel input/fitur). Klasifikasi ini menggunakan model Machine Learning untuk menemukan pola historis yang dapat membantu dalam pengambilan keputusan dan evaluasi keuangan.

Tabel 2. 3 Tabel Alur Sistem Klasifikasi

Tahap	Komponen	Deskripsi
INPUT	<p>1. Data Historis Anggaran</p> <ul style="list-style-type: none"> - Data Master: Kode Rekening/Jenis Belanja (e.g., Belanja Pegawai, Belanja Barang, Belanja Modal), Sumber Dana (e.g., APBD, Bansos, Pinjaman). - Data Transaksi: Realisasi nilai anggaran, tanggal transaksi, unit kerja. - Data Target/Kategori Pelaksanaan: Label historis yang mencatat hasil pelaksanaan (e.g., " Serapan Tinggi ", " Serapan Sedang ", " Serapan Rendah "). 	<p>Tahap ini merupakan pengumpulan bahan baku. Data yang lengkap dan berkualitas adalah kunci keberhasilan model. Variabel independen (Jenis Belanja, Sumber Dana) dan variabel dependen (Pelaksanaan Anggaran) diambil dari sumber-sumber seperti Sistem Keuangan Daerah (SIPKD) atau laporan realisasi anggaran.</p>
PROSES	<p>2. Pra-Pemrosesan Data</p> <ul style="list-style-type: none"> - Pembersihan Data (Data Cleaning): Menangani data yang hilang (missing values), outlier, atau inkonsistensi. - Transformasi Data: Mengubah data kategorikal (seperti nama Jenis Belanja) menjadi data numerik menggunakan teknik seperti Label Encoding atau One-Hot Encoding agar dapat diproses oleh algoritma. - Normalisasi/Standarisasi: Terutama penting untuk SVM, yaitu menyamakan skala fitur numerik (seperti nilai anggaran) untuk mencegah dominasi fitur dengan skala besar. - Pembagian Data: Membagi dataset menjadi data latih (training set, e.g., 80%) dan data uji (testing set, e.g., 20%). 	<p>Tahap ini adalah tahap preparasi untuk memastikan data siap dimasukkan ke dalam model. Hasil dari tahap ini adalah dataset yang bersih dan terstruktur.</p>
	<p>3. Ekstraksi Fitur & Pemodelan</p> <ul style="list-style-type: none"> - Pemilihan Fitur (Feature Selection): Mengidentifikasi fitur paling berpengaruh (misalnya, apakah kombinasi jenis belanja tertentu dengan sumber dana tertentu lebih signifikan?). Teknik seperti Importance Score dari Random Forest dapat digunakan. - Pembuatan Model Klasifikasi: <ul style="list-style-type: none"> a. Random Forest Classifier (RFC): Membangun banyak pohon keputusan (ensemble) dan hasil klasifikasi 	<p>Ini adalah inti dari penelitian. Dua algoritma berbeda diterapkan pada dataset yang sama. Teori ensemble learning (RFC) dan maximum margin classification (SVM) menjadi landasan pada tahap ini.</p>

Tahap	Komponen	Deskripsi
	<p>ditentukan oleh voting dari semua pohon. Metode ini robust terhadap overfitting.</p> <p>b. Support Vector Machine (SVM): Mencari hyperplane terbaik dalam ruang dimensi tinggi untuk memisahkan kelas-kelas pelaksanaan anggaran. Kernel (linear, RBF, dll) akan dieksplorasi untuk menemukan pemisah non-linear yang optimal.</p>	
	<p>4. Evaluasi dan Optimasi Model</p> <ul style="list-style-type: none"> - Pelatihan Model (Training): Model RFC dan SVM dipelajari menggunakan data latih untuk mengenali pola. - Pengujian Model (Testing): Model yang telah dilatih diuji dengan data uji yang belum pernah dilihat untuk mengukur kemampuannya menggeneralisasi. - Optimasi Hyperparameter: Menggunakan teknik seperti GridSearchCV untuk mencari kombinasi parameter terbaik (contoh: n_estimators untuk RFC, C dan gamma untuk SVM) guna memaksimalkan kinerja. - Perbandingan Kinerja: Membandingkan kedua model berdasarkan metrik akurasi, presisi, recall, F1-Score, dan AUC-ROC. 	<p>Tahap ini untuk memastikan model yang dibangun tidak hanya akurat pada data latih tetapi juga reliable pada data baru. Model terbaik akan dipilih untuk tahap output.</p>
OUTPUT	<p>5. Hasil Klasifikasi</p> <ul style="list-style-type: none"> - Label Klasifikasi Baru: Model yang telah terlatih dan teruji dapat menerima input data anggaran baru (Jenis Belanja dan Sumber Dana) untuk memprediksi kategori Pelaksanaan Anggaran-nya (misal: "Diprediksi Tepat Waktu"). - Laporan Klasifikasi: Berupa confusion matrix dan laporan metrik evaluasi yang menunjukkan seberapa baik kinerja model. - Model Final: Model machine learning (baik RFC atau SVM yang terbukti lebih unggul) yang siap diimplementasikan dalam sistem pendukung keputusan. - Rekomendasi: Insight tentang kombinasi jenis belanja dan sumber 	<p>Tahap akhir adalah menghasilkan keluaran yang memiliki nilai guna. Output utama adalah kemampuan prediktif dan insight yang dapat membantu meningkatkan efisiensi dan efektivitas pelaksaa</p>

Tahap	Komponen	Deskripsi
	dana seperti apa yang cenderung menghasilkan pelaksanaan anggaran yang optimal, yang dapat menjadi bahan pertimbangan bagi perencana anggaran.	

BAB III

METODOLOGI PENELITIAN

3.1 Jenis Penelitian

Penelitian ini mengambil judul “Klasifikasi Anggaran Berdasarkan Jenis Belanja Dan Sumber Dana Terhadap Pelaksanaan Anggaran Menggunakan Metode Random Forest Classifier (RFC) Dan Support Vector Machine (SVM).” Sesuai dengan rumusan masalah dan tujuan penelitian, maka jenis penelitian yang dipilih adalah penelitian kuantitatif dengan pendekatan machine learning untuk analisis klasifikasi. Pemilihan metode ini didasarkan pada tujuan penelitian, yaitu untuk mendeskripsikan, menganalisis, dan mengevaluasi ketepatan klasifikasi anggaran serta dampaknya terhadap akuntabilitas pengelolaan APBN di Kementerian / Lembaga.

3.2 Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan pada salah satu satuan kerja di lingkungan Kementerian Agama, yaitu pada **Universitas Islam Negeri Maulana Malik Ibrahim Malang**. Lokasi penelitian ini dipilih dengan pertimbangan:

- a. Memiliki volume anggaran yang signifikan.
- b. Mengelola lebih dari satu sumber dana (Rupiah Murni, PNBP, dan PHLN).
- c. Pernah menghadapi kendala atau temuan terkait klasifikasi anggaran dalam laporan keuangan.

Waktu penelitian direncanakan berlangsung selama ±3 bulan, mulai dari tahap pengumpulan data, analisis, hingga penyusunan laporan hasil penelitian.

3.3 Jenis dan Sumber Data

- a. Dokumen DIPA dan revisinya.
- b. Laporan Realisasi Anggaran (LRA) K/L.
- c. Laporan Keuangan Pemerintah Pusat (LKPP) dan Laporan Hasil Pemeriksaan BPK.
- d. Regulasi terkait pengelolaan anggaran (UU, PP, PMK, Perdirjen Perbendaharaan, BAS).

3.4 Teknik Pengumpulan Data

Studi Dokumentasi: mengumpulkan dokumen resmi (DIPA, laporan realisasi, BAS, laporan keuangan, hasil audit) dan revisinya dari beberapa tahun terakhir serta Laporan Realisasi Anggaran (LRA) atau data transaksi keuangan historis yang detail (hingga level bendahara pengeluaran/Surat Perintah Membayar - SPM) yang memuat informasi: Kode Akun, Uraian Belanja, Sumber Dana, Pagu, Realisasi, Kode Program/Kegiatan, dan Unit Organisasi.

3.5 Fokus Penelitian

- a. Ketepatan klasifikasi **jenis belanja** (pegawai, barang, modal, hibah, bansos, dll.) dalam pelaksanaan APBN K/L.
- b. Ketepatan klasifikasi **sumber dana** (Rupiah Murni, PNBP, BLU, SBSN, PHLN/PDN, hibah langsung) dalam mendanai kegiatan K/L.
- c. Dampak kesalahan klasifikasi terhadap pelaksanaan anggaran, pencatatan aset, serta kualitas laporan keuangan.
- d. Upaya dan strategi perbaikan klasifikasi anggaran di tingkat K/L maupun dalam sistem perpendaharaan.

3.6 Variabel Penelitian dan Proses Pelabelan Data

- a. Variabel Target (Label): Jenis Belanja (contoh: Belanja Pegawai, Barang, Modal) dan Sumber Dana (contoh: APBN, BLU).
- b. Variabel Fitur: Variabel lain dalam dataset (Pagu, Realisasi, Kode Program, Kode Kegiatan, Waktu, Unit Organisasi) yang akan digunakan model untuk memprediksi variabel target.
- c. Pelabelan Data: Proses memastikan data historis sudah memiliki label yang benar untuk dijadikan dasar pelatihan model.

3.7 Analisis Data dengan Machine Learning

Langkah-langkah analisis data kuantitatif adalah sebagai berikut:

- a. **Preprocessing Data:** Menangani missing value, outliers, dan normalisasi data.
- b. **Feature Engineering:** Memilih fitur yang paling relevan untuk model.

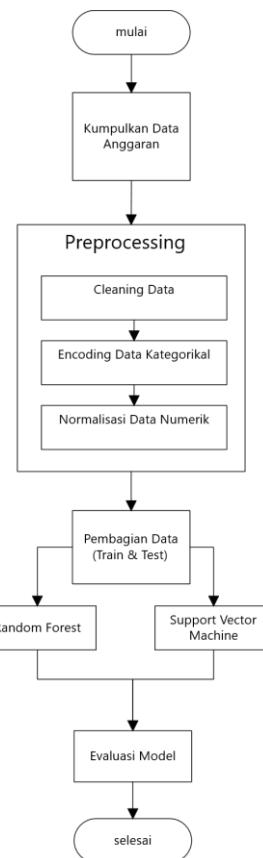
c. **Pelatihan dan Pengujian Model:**

- Data dibagi menjadi data latih (train set) dan data uji (test set).
- Model **Random Forest Classifier** dan **Support Vector Machine (SVM)** dilatih dengan data latih.
- Hyperparameter tuning dilakukan untuk mengoptimalkan kinerja model.

d. **Evaluasi Model:**

- Kinerja kedua model dievaluasi dan dibandingkan menggunakan metrik **Akurasi, Presisi, Recall, dan F1-Score**.
- Analisis **Feature Importance** dari model Random Forest dilakukan untuk mengidentifikasi variabel paling berpengaruh dalam klasifikasi.

3.8 Desain Sistem



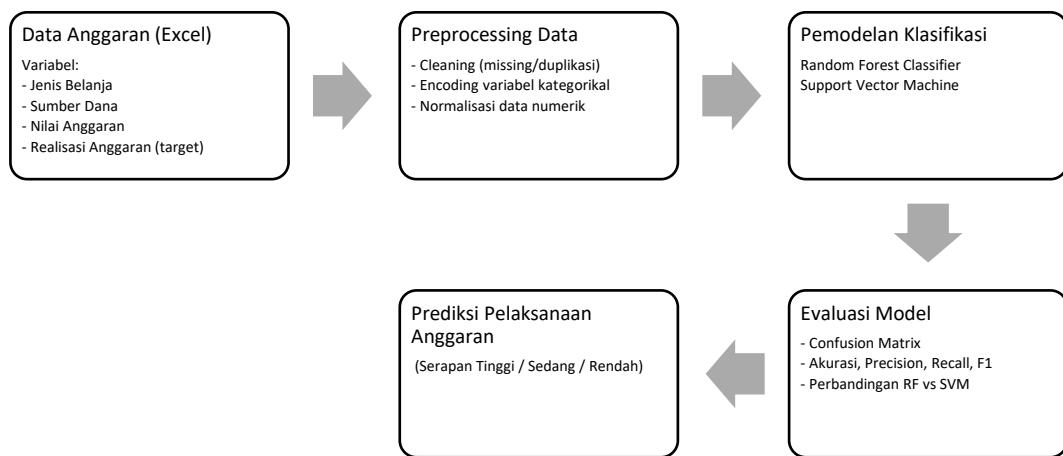
Gambar 3. 1 Desain Sistem

Penjelasan Singkat Alur:

- Kumpulkan Data Anggaran:** Mengumpulkan data transaksi keuangan K/L yang berisi informasi tentang jenis belanja, sumber dana, dll.
- Bersihkan & Siapkan Data:** Membersihkan data dari kesalahan dan mengaturnya agar bisa dipahami oleh model AI.
- Latih Model AI:** Melatih dua model (Random Forest dan SVM) untuk mempelajari pola dalam data.
- Evaluasi & Bandingkan:** Menguji kedua model untuk melihat mana yang lebih akurat dalam mengklasifikasikan anggaran.
- Analisis Hasil & Buat Rekomendasi:** Menganalisis model terbaik untuk memahami polanya dan membuat saran perbaikan untuk kebijakan.

3.9 Kerangka Konsep Penelitian

Gambaran alur berpikir dalam penelitian ini adalah menganalisis data anggaran berdasarkan jenis belanja dan sumber dana untuk memprediksi pelaksanaan anggaran.



Gambar 3.2 Kerangka Konsep

Alur Kerangka Konsep dapat dijelaskan sebagai berikut:

1. Data Anggaran

Dataset yang digunakan berisi variabel:

- kode_kegiatan
- uraian
- nilai_anggaran
- jenis_belanja
- sumber_dana
- realisasi_anggaran (target variabel Y)

2. Variabel Input (X)

- Jenis Belanja
- Sumber Dana
- Nilai Anggaran (jika diperlukan sebagai fitur tambahan)

3. Preprocessing Data

- Membersihkan data dari duplikasi/missing value
- Encoding data kategorikal (Jenis Belanja, Sumber Dana)
- Normalisasi atau standarisasi data numerik

4. Metode Klasifikasi

- Random Forest Classifier (RFC)
- Support Vector Machine (SVM)

5. Evaluasi Model

- Confusion Matrix
- Akurasi, Precision, Recall, dan F1-Score
- Perbandingan performa RFC vs SVM

6. Prediksi Pelaksanaan Anggaran (Y)

- Model terbaik digunakan untuk memprediksi realisasi_anggaran

BAB IV

ANALISIS DATA

Bab ini menyajikan pengolahan data sebelum di proses pada machine learning yang di buat yang terdiri dari preprocessing data, Feature Engineering, Pelatihan dan Pengujian Model, Evaluasi Model.

4.1 Data Penelitian

Data yang diperoleh dalam penelitian ini sebanyak 544 data, hal ini dikarenakan keterbatasan resource yang digunakan pada penelitian ini. Berikut beberapa data yang ditampilkan dalam laporan ini dari dokumen resmi (DIPA, laporan realisasi, BAS, laporan keuangan, hasil audit) dan revisinya dari beberapa tahun terakhir serta Laporan Realisasi Anggaran (LRA) atau data transaksi keuangan historis yang detail (hingga level bendahara pengeluaran/Surat Perintah Membayar - SPM) yang memuat informasi: Kode Akun, Uraian Belanja, Sumber Dana, Pagu, Realisasi, Kode Program/Kegiatan, dan Unit Organisasi dan telah diolah menjadi dataset seperti dapat dilihat pada Tabel 4.1.

Tabel 4. 1 Dataset

kode_kegiatan	uraian	nilai_anggaran	jenis_belanja	sumber_dana	realisasi_anggaran	%_realisasi
0	101	Operasional dan layanan perkantoran Pendidikan	496580000	Belanja Barang	RM BOPTN	75118000
1	102	Bimbingan Teknis Penulisan Karya Ilmiah Bagi M...	50880000	Belanja Barang	RM BOPTN	47394500
2	103	Pengabdian Kepada Masyarakat	165070000	Belanja Barang	RM BOPTN	114854000
3	104	Evaluasi Kurikulum Program Studi di FITK	16200000	Belanja Barang	RM BOPTN	0
4	105	Penjaminan Mutu FITK	40390000	Belanja Barang	RM BOPTN	0
...
539	640	Penyelenggaraan Operasional Fakultas Saintek	249875000	Belanja Barang	RM	166580329
540	641	PENYELENGGARAAN OPERASIONAL FAKULTAS KEDOKTERAN	66054000	Belanja Barang	RM	36518000
541	642	Pemeliharaan Gedung Pascasarjana	128144000	Belanja Barang	RM	0
542	643	Layanan Manajemen dan Tugas Teknis Bidang Pere...	98379000	Belanja Barang	RM	56115002
543	644	Penyelenggaran Operasional Mahad	195619000	Belanja Barang	RM	16993000

544 rows × 7 columns

```

*** Informasi Dataset ***
Jumlah Baris: 544
Jumlah Kolom: 7

Nama Kolom:
['kode_kegiatan', 'uraian', 'nilai_anggaran', 'jenis_belanja', 'sumber_dana', 'realisasi_anggaran', '%_realisasi']

```

Dari Tabel 4.1 ditampilkan data sebanyak sepuluh baris dengan nama dan 7 kolom diantaranya: 'kode_kegiatan', 'uraian', 'nilai_anggaran', 'jenis_belanja', 'sumber_dana', 'realisasi_anggaran', '%_realisasi'

4.1.1 Profil data

Dataset yang digunakan dalam penelitian ini terdiri dari **544 entri** data anggaran dengan variabel-variabel sebagai berikut:

Variabel Input:

- `kode_kegiatan`: Identifier unik untuk setiap kegiatan
- `uraian`: Deskripsi detail dari kegiatan
- `nilai_anggaran`: Pagu anggaran yang dialokasikan (dalam satuan mata uang)
- `jenis_belanja`: Klasifikasi jenis pengeluaran
- `sumber_dana`: Asal pendanaan kegiatan
- `realisasi_anggaran`: Jumlah anggaran yang benar-benar terealisasi

Variabel Target:

- `%_realisasi`: Persentase realisasi anggaran ($\text{Realisasi}/\text{Pagu} \times 100\%$)

4.1.2 Statistik Deskriptif

Pada Statistik Deskriptif ini mendeskripsikan karakteristik utama dari kumpulan data yang digunakan seperti pada Tabel 4.2.

Tabel 4. 2 Deskriptif data

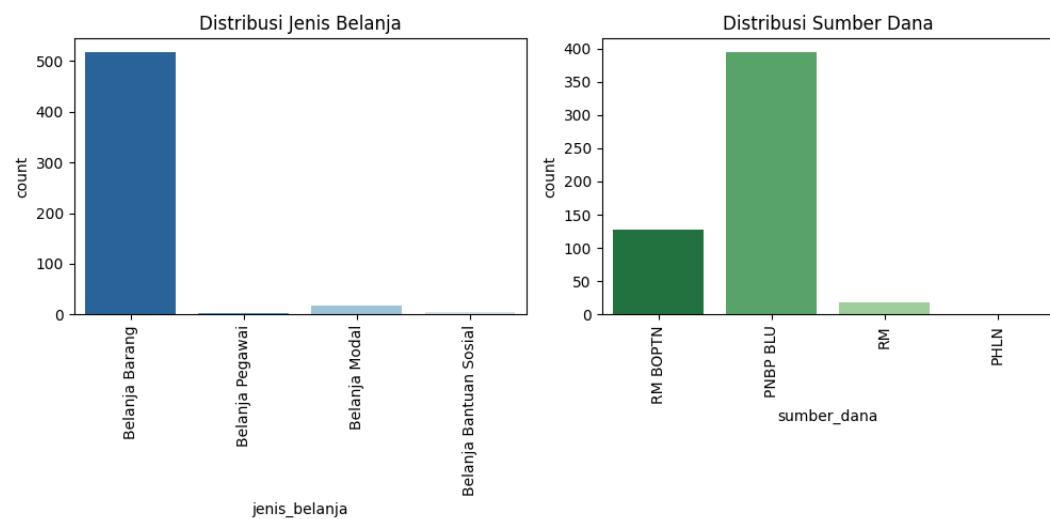
Statistik	<code>kode_kegiatan</code>	<code>nilai_anggaran</code>	<code>realisasi_anggaran</code>	<code>%_realisasi</code>
count	544	544	544	544
mean	372.5	1.479.330.000	711.728.700	0.2106
std (deviasi standar)	157.18	17.497.140.000	9.132.957.000	0.3026
min (minimum)	101	1.060.000	0	0.00
25% (kuartil 1)	236.75	25.457.500	0	0.00
50% (median)	372.5	80.000.000	0	0.00
75% (kuartil 3)	508.25	217.520.000	30.650.460	0.3749
max (maksimum)	644	380.376.100.000	199.633.500.000	1.00

Dari Tabel 4.2 menjelaskan bahwa sebagian besar kegiatan gagal menyerap anggaran (bahkan 50%-nya nol), sehingga rata-rata realisasi hanya 21%. Namun,

ada beberapa proyek beranggaran sangat besar yang berhasil direalisasi, yang membuat statistik rata-rata menjadi tinggi. Sebagian besar kegiatan justru bernilai kecil dan tidak terserap.

4.1.3 Distribusi Data

Distribusi merupakan penyebaran atau sebaran data dalam suatu variabel yang terbagi ke dalam kategori atau rentang nilai tertentu. Seperti dijelaskan pada Gambar 4.1



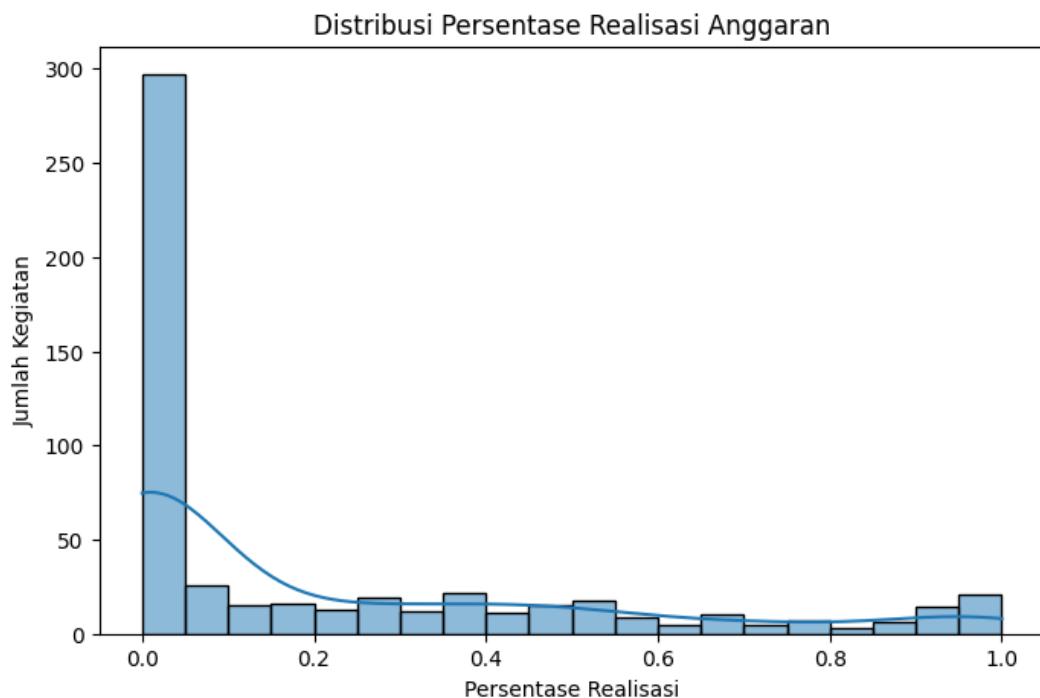
Gambar 4. 1 Distribusi Data

Pada Gambar 4.1 menunjukkan penyebaran dua variabel kategoris dalam kumpulan data anggaran, yaitu jenis belanja dan sumber dana. Diagram batang menunjukkan bahwa kategori Belanja Barang memiliki entri data terbanyak, dengan lebih dari 500 baris. Ini berarti sebagian besar dana dalam anggaran dialokasikan untuk membeli barang dan memenuhi kebutuhan unit kerja. Di sisi lain, kategori Belanja Modal hanya memiliki sekitar dua puluh item data. Ini berarti tidak banyak kegiatan yang berkaitan dengan pembelian aset tetap atau pembangunan. Di saat yang sama, informasi Belanja Pegawai dan Belanja Bantuan Sosial tampaknya hampir seluruhnya hilang. Sumber pendanaan terbesar, dengan hampir 380 kegiatan yang berasal dari Penerimaan Negara Bukan Pajak BLU. Dengan lebih dari 130 kegiatan, BOPTN (Bantuan Operasional Rupiah Murni Perguruan Tinggi Negeri) RM berada di urutan kedua. Kegiatan-kegiatan ini sangat penting untuk mendapatkan dana bagi

kebutuhan akademik dan operasional perguruan tinggi. Sementara itu, dana BOPTN (Bantuan Operasional Rupiah Murni) hanya digunakan untuk beberapa kegiatan operasional, dan dana PHLN (Pinjaman/Hibah Luar Negeri) sama sekali tidak ada dalam dataset. Pola distribusi ini menunjukkan bahwa sebagian besar dana berasal dari bantuan operasional pemerintah dan dana internal perguruan tinggi. Sebagian kecil berasal dari investasi dan kegiatan yang terjadi di luar negeri.

4.1.4 Distribusi Persentase Realisasi Anggaran

Distribusi Persentase Realisasi Anggaran menunjukkan penyebaran persentase realisasi anggaran untuk semua kegiatan dalam kumpulan data ini dilihat pada Gambar 4.2.



Gambar 4. 2 Distribusi Persentase Realisasi Anggaran

Pada Gambar 4.2 menunjukkan bahwa beberapa kegiatan tidak terlalu sering terjadi. Hal ini ditunjukkan oleh batang tertinggi dalam rentang 0,0–0,1, yang memiliki lebih dari 300 kegiatan. Ini berarti bahwa sebagian besar kegiatan belum dilakukan atau hanya menggunakan sebagian kecil dari keseluruhan anggaran.

Distribusi kemudian miring tajam ke kanan, yang berarti bahwa terdapat lebih sedikit kegiatan dengan tingkat realisasi yang lebih tinggi. Hanya ada

beberapa kegiatan yang memiliki persentase realisasi mendekati 1,0 (100%). Pola ini memiliki distribusi yang miring positif, yang berarti bahwa sebagian besar nilainya berada pada persentase realisasi yang rendah.

Secara keseluruhan, angka-angka ini menunjukkan bahwa jumlah penyerapan anggaran selama periode observasi masih cukup rendah karena disebabkan oleh berbagai faktor, seperti keterlambatan pelaksanaan kegiatan, proses administrasi yang belum selesai, atau permasalahan teknis dan regulasi terkait penggunaan anggaran. Anda dapat mengetahui jenis pengeluaran atau sumber pembiayaan mana yang paling bertanggung jawab atas rendahnya realisasi ini dengan mencermatinya lebih lanjut.

4.2 Preprocessing Data

Preprocessing merupakan proses penting dalam mempermudah pengklasifikasian diantaranya:

1. Menghapus data dari duplikasi/missing value.

Berikut adalah kode program menghapus duplikasi data dalam Gambar 4.3.

```
duplikat = data.duplicated().sum()
data = data.drop_duplicates()
print(f"\nJumlah data duplikat yang dihapus: {duplikat}")
```

Gambar 4. 3 Hapus Duplikasi

2. Menghapus nilai kosong (NaN)

Berikut adalah kode program menghapus nilai kosong diperlihatkan dalam Gambar 4.4.

```
missing_rows = data.isnull().sum().sum()
data = data.dropna()
print(f"\nJumlah nilai kosong yang dihapus: {missing_rows}")
```

Gambar 4. 4 Hapus Nilai Kososng

3. Hitung rasio realisasi

Berikut adalah kode program menghapus nilai kosong dalam Gambar 4.5.

```
data["rasio"]=(data["realisasi_anggaran"]/data["nilai_anggaran"])*100
display(data.head(10)).
```

Gambar 4. 5 Rasio Realisasi

Hasilnya dapat di lihat pada Gambar 4.6

kode_kegiatan	uraian	nilai_anggaran	jenis_belanja	sumber_dana	realisasi_anggaran	%_realisasi	ratio
0	101 Operasional dan layanan perkantoran Pendidikan	486588000	Belanja Barang	RM BOPTN	75118000	0.15	15.44
1	102 Bimbingan Teknis Penulisan Karya Ilmiah Bagi M...	50880000	Belanja Barang	RM BOPTN	47394500	0.93	93.15
2	103 Pengabdian Kepada Masyarakat	165070000	Belanja Barang	RM BOPTN	114854000	0.70	69.58
3	104 Evaluasi Kurikulum Program Studi di FITK	16200000	Belanja Barang	RM BOPTN	0	0.00	0.00
4	105 Penjaminan Mutu FITK	40390000	Belanja Barang	RM BOPTN	0	0.00	0.00
5	106 Pendampingan Publikasi Ilmiah FITK	82000000	Belanja Barang	RM BOPTN	990000	0.01	1.21
6	107 Peningkatan Mutu Riset dan Academic Writing	105446000	Belanja Barang	RM BOPTN	0	0.00	0.00
7	108 Program Pembinaan Prestasi Mahasiswa	169470000	Belanja Barang	RM BOPTN	50400000	0.30	29.74
8	109 Character Building Ormawa	35200000	Belanja Barang	RM BOPTN	0	0.00	0.00
9	110 Kerjasama Seleksi dan Penempatan Lulusan	30400000	Belanja Barang	RM BOPTN	22300000	0.73	73.36

Gambar 4. 6 Rasio Data

4. Membuat kolom kategori pelaksanaan (Target / Y).

Berikut adalah kode program menghapus nilai kosong diperlihatkan dalam Gambar 4.7.

```

if 'ratio' in data.columns:
    def kategori(ratio):
        if ratio >= 90:
            return "Serapan Tinggi"
        elif ratio >= 60:
            return "Serapan Sedang"
        else:
            return "Serapan Rendah"

    data["pelaksanaan"] = data["ratio"].apply(kategori)
    print("\nKolom 'pelaksanaan' berhasil ditambahkan berdasarkan nilai rasio.")
else:
    print ("\nKolom 'ratio' tidak ditemukan! Pastikan nama kolom sesuai dengan file Excel Anda.")

```

Gambar 4. 7 Membuat kolom kategori pelaksanaan

Berikut adalah penjelasan mengenai proses yang dilakukan:

1. Penambahan Kolom Kategori Pelaksanaan (Target / Y) ram melakukan pengecekan keberadaan kolom 'ratio' dalam dataset

- Menggunakan fungsi kategori () untuk mengklasifikasikan tingkat serapan anggaran berdasarkan nilai rasio:
- Serapan Tinggi: rasio $\geq 90\%$
- Serapan Sedang: rasio $\geq 60\%$ dan $< 90\%$
- Serapan Rendah: rasio $< 60\%$
- Hasil klasifikasi ditambahkan sebagai kolom baru 'pelaksanaan' menggunakan metode apply ()

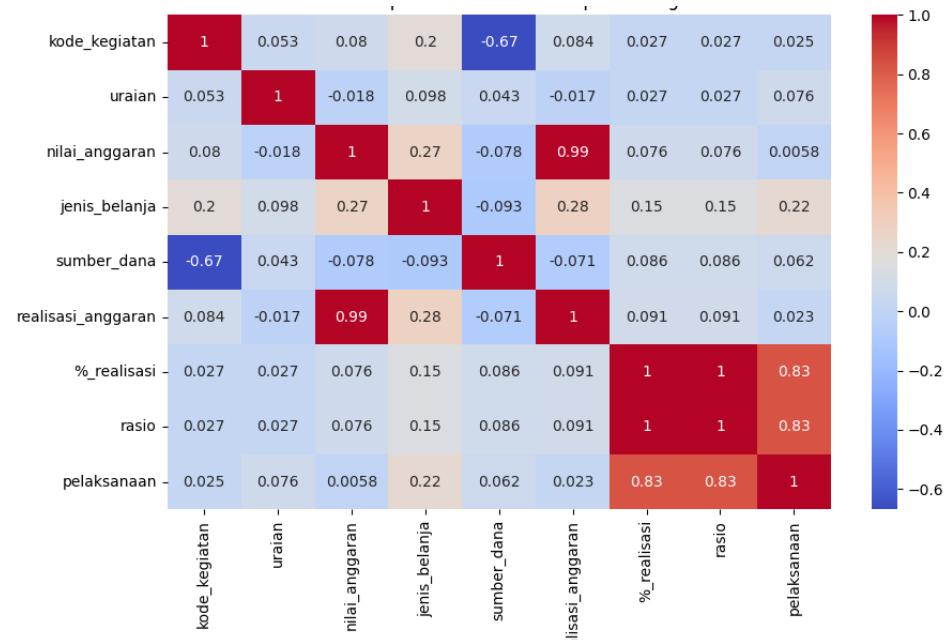
2. Hasil pada Data:

- Kegiatan 102 memiliki rasio 93.15% → Serapan Tinggi
- Kegiatan 103 dan 110 memiliki rasio 69.58% dan 73.38% → Serapan Sedang
- Kegiatan lainnya dengan rasio di bawah 60% → Serapan Rendah
- Hasil tampilan setelah dijalankan seperti dijelaskan pada Gambar 4.4 di bawah ini:

kode_kegiatan	uraian	nilai_anggaran	jenis_belanja	sumber_dana	realisasi_anggaran	%_realisasi	ratio	pelaksanaan
0	101 Operasional dan layanan perkantoran Pendidikan	486588000	Belanja Barang	RM BOPTN	75118000	0.15	15.44	Serapan Rendah
1	102 Bimbingan Teknis Penulisan Karya Ilmiah Bagi M...	50880000	Belanja Barang	RM BOPTN	47394500	0.93	93.15	Serapan Tinggi
2	103 Pengabdian Kepada Masyarakat	165070000	Belanja Barang	RM BOPTN	114854000	0.70	69.58	Serapan Sedang
3	104 Evaluasi Kurikulum Program Studi di FITK	16200000	Belanja Barang	RM BOPTN	0	0.00	0.00	Serapan Rendah
4	105 Penjaminan Mutu FITK	40390000	Belanja Barang	RM BOPTN	0	0.00	0.00	Serapan Rendah
5	106 Pendampingan Publikasi Ilmiah FITK	82000000	Belanja Barang	RM BOPTN	990000	0.01	1.21	Serapan Rendah
6	107 Peningkatan Mutu Riset dan Academic Writing	105446000	Belanja Barang	RM BOPTN	0	0.00	0.00	Serapan Rendah
7	108 Program Pembinaan Prestasi Mahasiswa	169470000	Belanja Barang	RM BOPTN	50400000	0.30	29.74	Serapan Rendah
8	109 Character Building Ormawa	35200000	Belanja Barang	RM BOPTN	0	0.00	0.00	Serapan Rendah
9	110 Kerjasama Seleksi dan Penempatan Lulusan	30400000	Belanja Barang	RM BOPTN	22300000	0.73	73.36	Serapan Sedang

Gambar 4. 8 Hasil Pengkategorian Tingkat Serapan Anggaran

Setelah mengkategorikan tingkat serapan anggaran, analisis dilanjutkan dengan mengeksplorasi hubungan linear antar variabel numerik melalui heatmap korelasi, yang mengungkapkan pola keterkaitan yang menarik antara alokasi anggaran, realisasi, dan efisiensi penyerapannya seperti dijelaskan pada Gambar 4. 9 Heatmap Korelasi Antar Variabel Numerik di bawah ini.



Gambar 4. 9 Heatmap Korelasi Antar Variabel Numerik

Berdasarkan Gambar 4. 9 yang ditampilkan analisis hubungan antar variabel numerik yaitu hubungan signifikan antara sejumlah karakteristik dan variabel implementasi tujuan dikonfirmasi oleh temuan analisis korelasi, yang ditampilkan sebagai peta panas. Indikator kinerja penyerapan anggaran merupakan prediktor penting dalam mengidentifikasi jenis implementasi aktual, seperti yang ditunjukkan oleh asosiasi positif yang sangat signifikan antara %_realisasi dan rasio serta implementasi (nilai korelasi 0,83). Namun, faktor kategorikal seperti sumber pendanaan dan jenis pengeluaran memberikan kontribusi prediktif yang sederhana; yang terakhir memiliki korelasi 0,22 dengan implementasi. Selain itu, peta panas menunjukkan multikolinearitas substansial antara %_realisasi dan rasio (0,83) dan antara realisasi anggaran dan nilai anggaran (0,99), yang harus diperhitungkan saat menafsirkan bobot fitur. Meskipun hubungan variabel keuangan (realisasi dan rasio) jauh lebih menonjol, secara keseluruhan, struktur korelasi ini mendukung kesesuaian penggunaan jenis pengeluaran dan sumber pendanaan sebagai karakteristik independen yang relevan dalam memprediksi pelaksanaan anggaran.

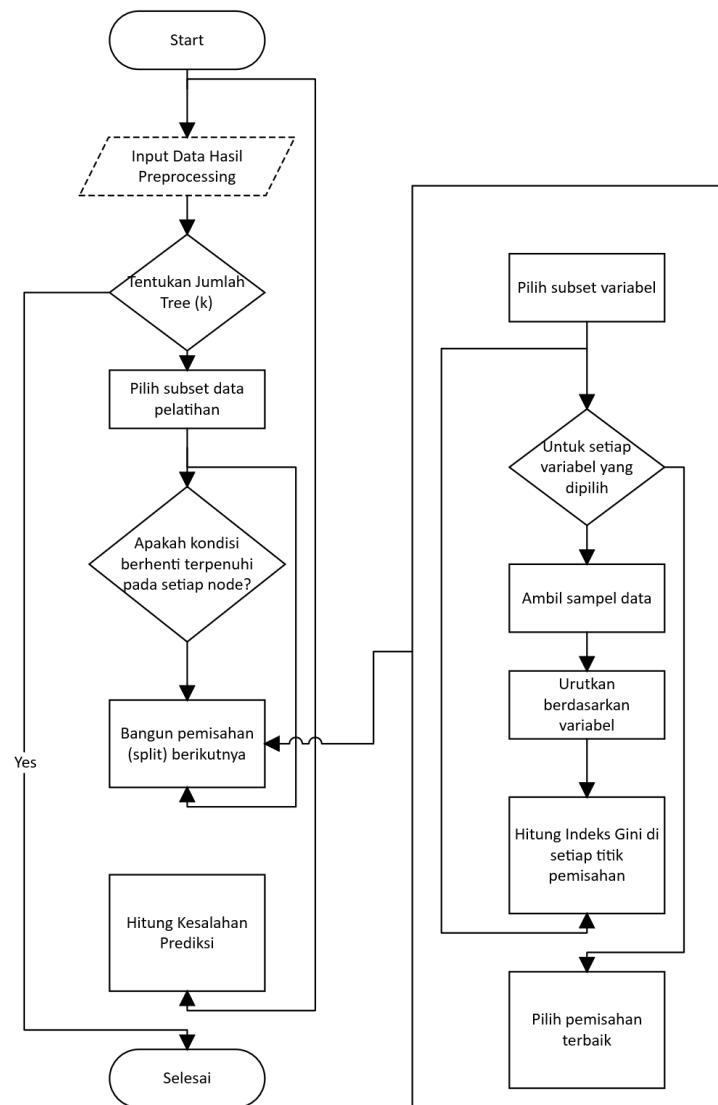
BAB V

RANDOM FOREST CLASSIFIER (RFC)

5.1 Metode Random Forest

Pendekatan ini merupakan varian dari algoritma Pohon Keputusan, termasuk ensemble pohon keputusan yang digunakan sebagai pengklasifikasi dasar terpadu. Metode ini digunakan untuk klasifikasi dan analisis regresi. (Louis M e Dian B, 2022).

5.2 Desain Model Random Forest



Gambar 5. 1 Desain Model Random Forest

Desain pada penelitian ini menjelaskan metodologi yang digunakan oleh peneliti untuk menarik kesimpulan yang bertujuan menjawab permasalahan yang diidentifikasi. Penelitian ini menggunakan model Random Forest (RF) dengan kerangka kerja pelatihan untuk dalam penganggaran berdasarkan kategori pengeluaran dan sumber pembiayaan dalam konteks pelaksanaan anggaran.

Setelah data melalui preprocessing maka Langkah selanjutnya Adalah melakukan proses klasifikasi proses tiga kali percobaan yaitu dengan system pembagian data training dan testing 90:10, 85:15, 80:20, 70:30, 60:40. Percobaan dilakukan untuk menilai kinerja model Random Forest dengan membandingkan data pelatihan dan data uji, dengan fokus pada Akurasi, Presisi, dan Recall. Eksperimen ini bertujuan untuk memastikan dampak volume data pelatihan dan data uji terhadap efektivitas klasifikasi Random Forest. Teknik evaluasi yang digunakan adalah Matriks Konfusi, dan model yang dihasilkan ditunjukkan pada Tabel 5.1.

5.3 Pengujian 1 dengan perbandingan data 90:10

Tabel 5. 1 Pengujian 1 Random Forest

	precision	recall	f1-score	support
0	0.88	1.00	0.94	423
1	0.00	0.00	0.00	33
2	0.33	0.09	0.15	32
accuracy			0.87	488
macro avg	0.41	0.36	0.36	488
weighted avg	0.79	0.87	0.82	488

Pengujian Random Forest dengan komposisi data 90:10, model ini berhasil mengidentifikasi kelas mayoritas (kelas 0) dengan sangat baik, tetapi masih belum berhasil mengidentifikasi kelas minoritas (kelas 1 dan 2). Pada kelas 0, akurasinya mencapai 88% dan recall-nya mencapai 100%. Ini berarti model tersebut mengidentifikasi semua data yang berasal dari kelas 0 dengan benar, dengan tingkat kesalahan perkiraan yang sangat rendah. Di sisi lain, pada kelas 1, model tersebut tidak dapat menemukan satu pun dari 33 data yang diketahui, yang ditunjukkan oleh angka akurasi dan recall sebesar 0%. Ini berarti semua informasi dari kelas 1 dimasukkan ke dalam kelas yang salah. Model tersebut juga tidak berhasil pada

kelas 2, dengan akurasi hanya 33% dan recall hanya 9%. Ini berarti sebagian besar data kelas 2 tidak dikenali dengan benar. Meskipun model ini secara keseluruhan akurat sebesar 87%, nilai F1-score rata-rata global yang rendah, yaitu hanya 36%, menunjukkan bahwa kinerjanya tidak merata di seluruh kelas dan sangat dipengaruhi oleh banyaknya data di kelas 0. Dapat dikatakan bahwa model Random Forest masih bias terhadap kelas mayoritas dan belum mampu mengenali kelas minoritas sebaik yang seharusnya dengan rasio 90:10.

5.4 Pengujian 2 dengan perbandingan data 85:15

Tabel 5. 2 Pengujian 2 Random Forest

	precision	recall	f1-score	support
0	0.88	1.00	0.94	400
1	0.50	0.03	0.06	31
2	0.25	0.07	0.11	30
accuracy			0.87	461
macro avg	0.54	0.37	0.37	461
weighted avg	0.82	0.87	0.82	461

Pengujian dengan menggunakan komposisi data 85:15 untuk menguji model Random Forest. Hasilnya menunjukkan bahwa model tersebut masih bekerja dengan sangat baik di kelas 0, yang merupakan kelas mayoritas, tetapi tidak begitu baik di kelas 1 dan 2, yang merupakan kelas minoritas. Model tersebut akurat 88% dari waktu di kelas 0 dan secara konsisten akurat 100% dari waktu. Ini berarti bahwa model tersebut dengan tepat memprediksi semua data yang sebenarnya berasal dari kelas 0. Namun, model tersebut masih tidak terlalu bagus di kelas 1. Model tersebut memiliki akurasi hanya 50% dan recall hanya 3%, yang berarti bahwa hanya 1 dari 31 set data kelas 1 yang diidentifikasi dengan benar. Yang lainnya salah dimasukkan ke kelas lain. Model tersebut juga tidak bekerja dengan baik di kelas 2, dengan akurasi 25% dan recall 7%. Ini berarti bahwa sebagian besar data di kelas ini tidak dapat diprediksi dengan benar berdasarkan kelompoknya. Model ini masih sangat akurat (87% dari waktu), tetapi nilai F1-score rata-rata global yang hanya 37% menunjukkan bahwa kinerjanya tidak merata antar kelas dan masih sangat dipengaruhi oleh besarnya jumlah data di kelas 0. Secara

keseluruhan, hasil ini menunjukkan bahwa model Random Forest masih bias terhadap kelas mayoritas pada pembagian data 85:15 dan belum mampu menggambarkan kelas minoritas dengan tepat.

5.5 Pengujian 3 dengan perbandingan data 80:20

Tabel 5. 3 Pengujian 3 Random Forest

Kelas	precision	recall	f1-score	support
0	0.94	1.00	0.97	95
1	0.75	0.43	0.55	7
2	1.00	0.57	0.73	7
accuracy			0.94	109
macro avg	0.90	0.67	0.75	109
weighted avg	0.93	0.94	0.93	109

Dari hasil Confusion Matrik di atas di dapat informasi bahwa Model ini mengenali kelas ini dengan sangat baik dengan tingkat kesalahan yang rendah. Berdasarkan hasil pengujian pertama dengan pembagian data 80% untuk pelatihan dan 20% untuk pengujian, model Random Forest menunjukkan peningkatan performa dengan akurasi 93,58%. Nilai precision dan recall tertinggi diperoleh pada kelas 0, model mampu mengenali seluruh data kelas ini dengan sempurna (recall = 1.00). Meskipun demikian, performa pada kelas 1 dan kelas 2 masih belum maksimal akibat ketidakseimbangan jumlah data. Secara keseluruhan, pengujian ini membuktikan bahwa peningkatan proporsi data latih berdampak positif terhadap kemampuan generalisasi model dalam melakukan klasifikasi pelaksanaan anggaran.

5.6 Pengujian 4 dengan perbandingan data 70:30

Tabel 5. 4 Pengujian 4 Random Forest

Kelas	precision	recall	f1-score	support
0	0.93	0.97	0.95	141
1	0.40	0.36	0.38	11
2	1.00	0.45	0.62	11
accuracy			0.90	163
macro avg	0.78	0.60	0.65	163
weighted avg	0.90	0.90	0.89	163

Berdasarkan hasil pengujian ke dua dengan pembagian data 70% untuk pelatihan dan 30% untuk pengujian, model Random Forest menghasilkan tingkat akurasi sebesar 89,57%. Nilai precision dan recall tertinggi diperoleh pada kelas 0, menunjukkan model sangat baik dalam mengenali kelas ini. Namun, performa pada kelas 1 dan kelas 2 masih rendah karena jumlah datanya relatif sedikit. Secara keseluruhan, model Random Forest dapat dikategorikan memiliki performa yang baik dalam mengklasifikasikan data pelaksanaan anggaran, meskipun masih perlu dilakukan penyeimbangan data untuk meningkatkan kinerja pada kelas minoritas.

5.7 Pengujian 5 dengan perbandingan data 60:40

Tabel 5. 5 Pengujian 5 Random Forest

Kelas	precision	recall	f1-score	support
0	0.94	0.97	0.96	188
1	0.43	0.40	0.41	15
2	0.88	0.50	0.64	14
accuracy			0.90	217
macro avg	0.75	0.62	0.67	217
weighted avg	0.90	0.90	0.90	217

Tabel 5.3 menunjukkan hasil pengujian ketiga menggunakan metode Random Forest dengan tingkat akurasi sebesar 90% dari total 217 data uji. Berdasarkan hasil evaluasi, model menunjukkan performa sangat baik pada kelas 0 dengan precision 0.94, recall 0.97, dan f1-score 0.96, menandakan kemampuan model yang tinggi dalam mengenali kelas mayoritas. Namun, performa menurun pada kelas 1 dan kelas 2 dengan f1-score masing-masing 0.41 dan 0.64, yang menunjukkan bahwa model masih kesulitan mengklasifikasikan kelas dengan jumlah data relatif sedikit. Nilai macro average sebesar 0.75 untuk precision, 0.62 untuk recall, dan 0.67 untuk f1-score menunjukkan bahwa secara rata-rata performa antar kelas belum seimbang, sedangkan weighted average sebesar 0.90 untuk ketiga metrik mengindikasikan bahwa model secara keseluruhan tetap memiliki kinerja yang baik meskipun terjadi ketidakseimbangan data antar kelas.

5.8 Kesimpulan

Dari 5 ujicoba yang telah dilakukan, diperoleh nilai rata-rata untuk Akurasi, Presisi, recall, dan f1-score yang tercantum pada Tabel 5.6.

Tabel 5. 6 Rata-Rata **Accuracy, Precision, Recall, f1-score**

Training	Testing	accuracy	precision	recall	f1-score
90	10	87%	79%	87%	0.82
85	15	87%	82%	87%	0.82
80	20	94%	93%	94%	0.93
70	30	90%	90%	90%	0.89
60	40	90%	90%	90%	0.90
Rata-Rata		90%	90%	87%	90%

Setelah lima pengujian klasifikasi menggunakan algoritma Random Forest dan pembagian data Pelatihan dan Pengujian yang berbeda (dari 90:10 hingga 60:40), model ini secara konsisten berkinerja sangat baik. Angka Akurasi rata-rata yang ditemukan adalah 90%, yang berarti model tersebut dapat menebak kelas dengan benar 90% dari waktu. Keberhasilan ini didukung oleh skor Presisi rata-rata 90% dan skor F1 rata-rata 90%, yang menunjukkan perpaduan yang baik antara pengurangan False Positives dan False Negatives. Model tersebut melewatkannya sekitar 13% dari true positives, seperti yang ditunjukkan oleh Recall rata-rata yang sedikit lebih rendah yaitu 87%. Namun, fakta bahwa ukuran rata-rata secara keseluruhan di atas 87% menunjukkan bahwa Random Forest adalah model yang baik dan andal untuk pekerjaan klasifikasi ini.

BAB VI

SUPPORT VECTOR MACHINE (SVM)

6.1 Metode Support Vector Machine (SVM)

Metode ini digunakan untuk tujuan prediksi dalam tugas klasifikasi dan regresi. Support Vector Machines (SVM) menyediakan metodologi statistik untuk identifikasi pola. Support Vector Machine (SVM) mengidentifikasi vektor pendukung optimal yang membatasi dua kelas berbeda dengan memaksimalkan margin (Tantika e Kudus 2022). Teknik Support Vector Machine (SVM) digunakan dalam proses klasifikasi untuk mengidentifikasi pola dari data pelatihan berlabel kemudian menghasilkan model pembelajaran mesin (Husada e Paramita, 2021).

Metode ini merupakan metode machine learning dan masuk kelas supervised learning digunakan untuk klasifikasi baik data linear atau non-linear yang Ide dasarnya adalah memaksimalkan batas *hyperplane* (Sihombing e Yuliati, 2021).

Untuk mendapatkan *hyperplane* yang paling ideal dalam mengklasifikasikan data ke dalam dua kategori, maka perhitungan margin *hyperplane* digunakan untuk mengidentifikasi titik terbesar (Husada e Paramita 2021). Dalam memperoleh *hyperplane* pada SVM, dapat menggunakan persamaan (1).

$$(w \cdot x_i) + b = 0 \quad (1).$$

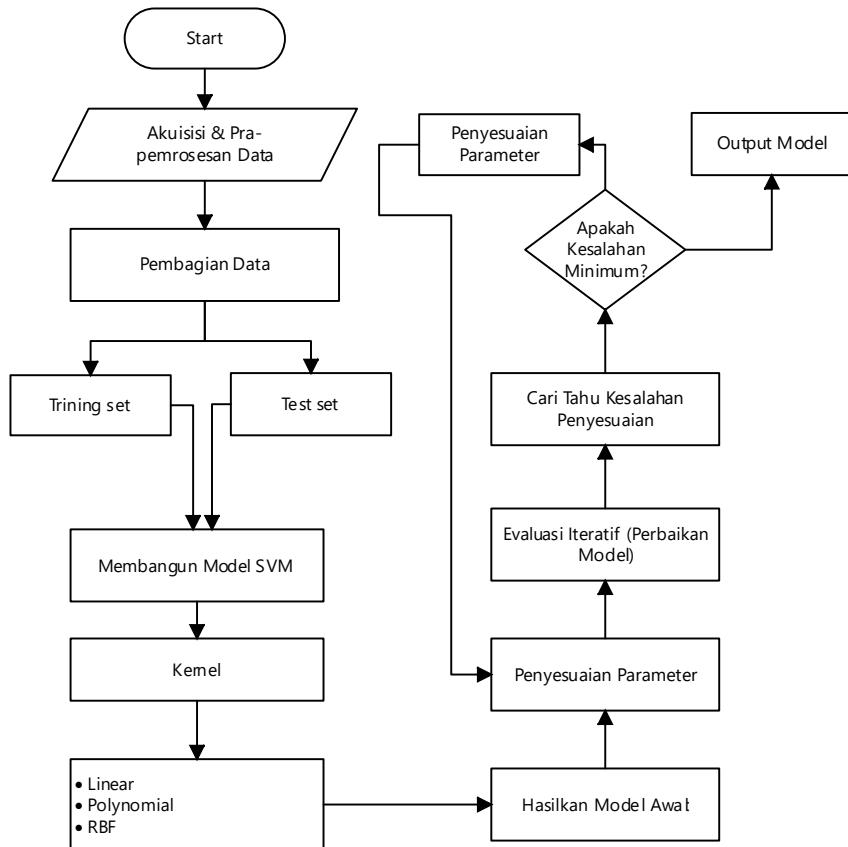
Di dalam data x_i , yang termasuk pada kelas -1 dapat dirumuskan seperti pada persamaan (2).

$$(w \cdot x_i + b) \leq 1, y_i = -1 \quad (2).$$

Sedangkan data x_i yang termasuk pada kelas +1 dapat dirumuskan seperti pada persamaan (3).

$$(w \cdot x_i + b) \geq 1, y_i = 1 \quad (3).$$

6.2 Desain Model Support Vector Machine



Gambar 6. 1 Flowchart Desain Model Support Vector Machine

Pada Gambar 6.1 merupakan gambaran dasar tentang bagaimana model SVM dibangun. Untuk membuat pengklasifikasi SVM, memerlukan fungsi kernel dan nilainya. Langkah pertama dalam membuat model SVM adalah mengumpulkan dataset dan melakukan beberapa persiapan. Ini termasuk menangani nilai yang hilang, menyimpan variabel kategori, dan menskalakan fitur numerik. Langkah-langkah ini diperlukan untuk memastikan bahwa struktur data baik untuk melatih model pembelajaran mesin. Berbagai fungsi kernel diantaranya fungsi polinomial, fungsi basis radial, dan pengklasifikasi linear. Data pelatihan dimasukkan ke dalam algoritma SVM, yang mencoba menemukan hyperplane terbaik dalam ruang fitur untuk memisahkan kelas sebanyak mungkin. Validasi silang digunakan untuk menemukan angka terbaik untuk hiperparameter. Akurasi pengklasifikasi SVM bergantung pada seberapa baik ia dapat menemukan hyperplane terbaik untuk membedakan antara kelas-kelas yang berbeda. Menemukan celah terbesar antara dua hyperplane inilah yang menentukan model klasifikasi mana yang paling akurat.

Untuk memulai proses klasifikasi, data dipisah untuk pelatihan dan pengujian dibuat. Setelah itu, penelitian mencakup pengujian untuk membuat model dengan mendapatkan selisih margin terbesar.

Seperti yang dilakukan pada Random Forest berlaku juga pada *Support Vector Machine* (SVM) yaitu setelah melalui preprocessing Langkah berikutnya melakukan proses klasifikasi proses tiga kali percobaan yaitu dengan sistem pembagian data training dan testing 80:20, 70:30, 60:40. Tahap ini dilakukan percobaan dengan mengatur perbandingan antara data pelatihan dan pengujian untuk mengevaluasi performa model *Support Vector Machine* (SVM), berdasarkan nilai Accuracy, Precision, dan Recall. Ujicoba ini dilakukan untuk melihat sejauh mana jumlah data pelatihan dan pengujian dapat memengaruhi klasifikasi yang dilakukan oleh Random Forest. Hasil ujicoba dilihat pada Tabel 6.1.

6.3 Pengujian 1 SVM dengan perbandingan data 90:10

Tabel 6. 1 Pengujian 1 SVM

	precision	recall	f1-score	support
0	0.88	0.98	0.93	423
1	0.00	0.00	0.00	33
2	0.39	0.22	0.28	32
accuracy			0.86	488
macro avg	0.42	0.40	0.40	488
weighted avg	0.79	0.86	0.82	488

Hasil pengujian klasifikasi Support Vector Machine (SVM) dengan perbandingan data 90:10 menunjukkan adanya ketidakseimbangan kinerja yang signifikan di antara tiga kelas, di mana model mencapai performa yang superior pada Kelas 0 (kelas mayoritas, support 423) dengan f1-score 0.93, namun menunjukkan kegagalan total (zero-performance) dalam mengklasifikasikan Kelas 1 (f1-score 0.00) dan kinerja yang marginal pada Kelas 2 (f1-score 0.28). Meskipun akurasi (accuracy) model mencapai 0.86, nilai ini bias dan didorong oleh keberhasilan pada Kelas 0, sementara Macro Average yang rendah (f1-score 0.40) mengindikasikan bahwa model memiliki masalah serius dalam generalisasi dan identifikasi pada kelas minoritas, yang mana isu class imbalance ini memerlukan intervensi metodologis lebih lanjut untuk mencapai kinerja yang merata dan robust.

6.4 Pengujian 2 SVM dengan perbandingan data 85:15

Tabel 6. 2 Pengujian 2 SVM

	precision	recall	f1-score	support
0	0.87	1.00	0.93	400
1	0.00	0.00	0.00	31
2	0.00	0.00	0.00	30
accuracy			0.87	461
macro avg	0.29	0.33	0.31	461
weighted avg	0.75	0.87	0.81	461

Hasil pengujian klasifikasi Support Vector Machine (SVM) dengan perbandingan data 85:15 menunjukkan dominasi performa yang ekstrem pada Kelas 0 (kelas mayoritas, support 400), di mana model mencapai kinerja yang superior dengan f1-score 0.93 (precision 0.87 dan recall 1.00). Namun, model menunjukkan kegagalan total (zero-performance) dalam mengklasifikasikan kedua kelas minoritas, yaitu Kelas 1 (support 31) dan Kelas 2 (support 30), karena semua metrik (precision, recall, dan f1-score) untuk kedua kelas tersebut bernilai 0.00. Meskipun akurasi (accuracy) model mencapai 0.87 (87%), nilai ini bias karena didorong oleh keberhasilan pada Kelas 0, sementara nilai Macro Average yang sangat rendah (f1-score 0.31) secara tegas mengindikasikan bahwa model tidak memiliki kemampuan generalisasi yang memadai dan sangat rentan terhadap isu class imbalance.

6.5 Pengujian 3 SVM dengan perbandingan data 80:20

Tabel 6. 3 Pengujian 3 SVM

	precision	recall	f1-score	support
0	0.87	1.00	0.93	95
1	0.00	0.00	0.00	7
2	0.00	0.00	0.00	7
accuracy			0.87	109
macro avg	0.29	0.33	0.31	109
weighted avg	0.76	0.87	0.81	109

Hasil pengujian model Support Vector Machine (SVM) dengan pembagian data latih dan data uji sebesar 80:20 menunjukkan bahwa model memiliki akurasi

keseluruhan sebesar 87%. Namun, performa klasifikasi tidak merata pada setiap kelas. Kelas 0 berhasil diprediksi dengan sangat baik dengan nilai precision 0,87, recall 1,00, dan f1-score 0,93, sedangkan untuk kelas 1 dan 2, model gagal melakukan prediksi yang benar dengan nilai precision, recall, dan f1-score sebesar 0,00. Nilai macro average (0,29 precision, 0,33 recall, dan 0,31 f1-score) menunjukkan ketidakseimbangan performa antar kelas, sedangkan weighted average (0,76 precision, 0,87 recall, 0,81 f1-score) masih cukup baik karena didominasi oleh kelas mayoritas. Hal ini mengindikasikan bahwa model SVM cenderung overfitting terhadap kelas dominan dan kurang mampu mengenali kelas minoritas, sehingga diperlukan penyeimbangan data atau tuning parameter agar hasil klasifikasi menjadi lebih seimbang.

6.6 Pengujian 4 SVM dengan perbandingan data 70:30

Tabel 6. 4 Pengujian 4 SVM

	precision	recall	f1-score	support
0	0.87	1.00	0.93	141
1	0.00	0.00	0.00	11
2	0.00	0.00	0.00	11
accuracy			0.87	163
macro avg	0.29	0.33	0.31	163
weighted avg	0.75	0.87	0.80	163

Berdasarkan Tabel 6.1 Confusion Matrix Pengujian 2 menggunakan algoritma SVM (Support Vector Machine) dengan pembagian data latih dan uji sebesar 70:30, diperoleh akurasi keseluruhan sebesar 87%. Hasil ini menunjukkan bahwa model cukup baik dalam mengklasifikasikan data secara umum, namun kinerjanya tidak merata antar kelas. Kelas 0 memiliki performa sangat baik dengan precision 0,87, recall 1,00, dan f1-score 0,93, menandakan bahwa hampir semua data kelas ini terkласifikasi dengan benar. Sebaliknya, kelas 1 dan 2 sama sekali tidak berhasil diprediksi oleh model dengan nilai precision, recall, dan f1-score sebesar 0,00, yang menunjukkan bahwa SVM gagal mengenali data dari kedua kelas tersebut. Nilai macro average (precision 0,29; recall 0,33; f1-score 0,31) mempertegas adanya

ketimpangan antar kelas, sementara weighted average (precision 0,75; recall 0,87; f1-score 0,80) masih relatif tinggi karena dominasi kelas 0. Secara keseluruhan, hasil ini mengindikasikan bahwa model SVM cenderung bias terhadap kelas mayoritas dan belum optimal dalam menangani kelas minoritas.

6.7 Pengujian 5 SVM dengan perbandingan data 60:40

Tabel 6. 5 Pengujian 5 SVM

Kelas	precision	recall	f1-score	support
0	0.87	1.00	0.93	188
1	0.00	0.00	0.00	15
2	0.00	0.00	0.00	14
accuracy			0.87	217
macro avg	0.29	0.33	0.31	217
weighted avg	0.75	0.87	0.80	217

Berdasarkan Tabel 6.1 Confusion Matrix Pengujian 2 menggunakan algoritma SVM (Support Vector Machine) dengan komposisi data latih dan uji yang sama yaitu 60:40, diperoleh hasil yang identik dengan pengujian pertama, yaitu akurasi 87%. Model masih menunjukkan kinerja yang sangat baik dalam mengenali kelas 0 dengan precision 0,87, recall 1,00, dan f1-score 0,93, namun kembali gagal mendeteksi kelas 1 dan kelas 2, yang masing-masing memiliki nilai precision, recall, dan f1-score sebesar 0,00. Nilai macro average (0,29; 0,33; 0,31) dan weighted average (0,75; 0,87; 0,80) juga tetap sama dengan pengujian sebelumnya. Kesamaan hasil ini menunjukkan bahwa model SVM menghasilkan pola klasifikasi yang konsisten.

6.8 Kesimpulan

Dari 5 ujicoba yang telah dilakukan, diperoleh nilai rata-rata untuk Akurasi, Presisi, recall, dan f1-score yang tercantum pada Tabel 6.6

Tabel 6. 6 Rata-rata Akurasi, Presisi, recall, dan f1-score

Training	Testing	accuracy	precision	recall	f1-score
90	10	86%	79%	86%	82%
85	15	87%	75%	87%	81%
80	20	87%	76%	87%	81%
70	30	87%	75%	87%	80%
60	40	87%	75%	87%	80%
Rata-Rata		87%	87%	75%	87%

Dalam Tabel 6.6 menerangkan hasil klasifikasi Support Vector Machine (SVM) diuji lima kali dengan rasio pembagian data pelatihan dan pengujian yang berbeda (dari 90:10 hingga 60:40). Model SVM selalu menunjukkan hasil yang stabil. Program ini mampu membuat prediksi akurat sebanyak 87% dari waktu, yang merupakan tingkat yang sangat tinggi. Skor Presisi rata-rata adalah 75%, skor Recall rata-rata adalah 87%, dan skor F1 rata-rata adalah 80%. Meskipun angka Presisi hanya 75%, ini berarti ada lebih banyak false positive dalam prediksi. Namun, total skor F1 sebesar 80% menunjukkan bahwa model SVM mencapai keseimbangan yang baik antara Presisi dan Recall untuk pekerjaan klasifikasi ini.

BAB VII

PEMBAHASAN

7.1 Analisis Kinerja Model Klasifikasi

Bagian ini akan mengkaji hasil klasifikasi dari dua model pendekatan Random Forest dan Support Vector Machine, menggunakan data uji dari bab sebelumnya. Untuk menggunakan metodologi Random Forest dan Support Vector Machine, data anggaran yang diperoleh harus melalui banyak tahapan persiapan, termasuk penghapusan data kosong, penghapusan nilai yang hilang, dan konversi tipe data ke format numerik. Prosedur ini harus dilakukan sebelum data digunakan dalam proses analitik untuk kedua pendekatan tersebut. Hasil pengujian memberikan nilai akurasi untuk setiap pendekatan, yang ditunjukkan pada Tabel 7.1.

Tabel 7. 1 Perbandingan Accuracy, Precision, Recall Random Forest dan SVM

Pengujian		Accuracy		precision		recall		f1-score	
Training	Testing	RFC	SVM	RFC	SVM	RFS	SVM	RFC	SVM
90	10	87%	86%	79%	79%	87%	86%	0.82	82%
85	15	87%	87%	82%	75%	87%	87%	0.82	81%
80	20	94%	87%	93%	76%	94%	87%	0.93	81%
70	30	90%	87%	90%	75%	90%	87%	0.89	80%
60	40	90%	87%	90%	75%	90%	87%	0.90	80%
Rata-Rata		90%	87%	90%	75%	87%	87%	90%	80%

Berdasarkan hasil pengujian pada Tabel 7.1, Dengan menggunakan dua model utama—Random Forest Classifier (RFC) dan Support Vector Machine (SVM) uji klasifikasi untuk menebak bagaimana anggaran akan dijalankan berdasarkan jenis pengeluaran dan sumber dana menunjukkan tingkat keberhasilan yang berbeda. Model Random Forest secara keseluruhan berkinerja sangat baik, dengan rata-rata Recall 85%, Akurasi 90%, dan Presisi 90%. Model ini juga mendapatkan skor F1 rata-rata 90%. Berdasarkan hasil ini, model Random Forest sangat baik dalam menemukan berbagai jenis aplikasi anggaran di dunia nyata. Selain itu, model SVM bekerja dengan baik dan konsisten, dengan Akurasi 87% dan Recall 87%, serta Presisi 75% dan skor F1 80%. Angka-angka ini menunjukkan bahwa kedua model dapat menebak dengan benar bagaimana anggaran akan

digunakan; Random Forest berkinerja lebih baik dalam Akurasi dan Presisi, sementara SVM berkinerja konsisten di berbagai rasio data uji..

Model Random Forest Classifier (RFC) secara teratur berkinerja lebih baik daripada model Support Vector Machine (SVM), seperti yang ditunjukkan oleh hasil pengujian. Skor F1 rata-rata, yang merupakan cara utama untuk membandingkan Presisi dan Recall, menunjukkan bahwa RFC 90% lebih baik daripada SVM. Alasan RFC lebih baik adalah karena ini adalah metode ensemble yang mencampur hasil dari beberapa pohon keputusan. Metode ini berhasil menurunkan variasi dan meningkatkan generalisasi, yang sangat penting untuk data eksekusi anggaran yang biasanya rumit dan tidak linier.

Di sisi lain, model SVM memiliki Presisi rata-rata yang jauh lebih rendah (75%), meskipun memiliki Akurasi rata-rata dan stabilitas Recall yang tinggi (87%). Ini menunjukkan bahwa model SVM memiliki tingkat False Positive yang lebih tinggi meskipun bagus dalam menemukan kasus positif (Recall tinggi). Dalam kehidupan nyata, ini berarti bahwa SVM cenderung mengatakan bahwa implementasi anggaran akan berjalan dengan baik meskipun mungkin tidak demikian. Ini mungkin karena pemilihan kernel bukanlah cara terbaik untuk membedakan antara kelas dalam ruang fitur yang kompleks.

7.2 Hubungan Penelitian dengan Pandangan Islam

Mengelola anggaran negara bukan hanya tugas administratif; melainkan juga mencakup aspek moral, etika, dan spiritual. Dari sudut pandang Islam, semua jenis pengelolaan aset publik merupakan amanah yang harus dikelola dengan penuh akuntabilitas, kejujuran, dan kehati-hatian. Hal ini berkaitan dengan kajian klasifikasi anggaran berdasarkan jenis belanja dan sumber pembiayaan, karena akurasi, presisi, dan efisiensi dalam klasifikasi anggaran sangat memengaruhi kualitas pelaksanaan APBN di Kementerian/Lembaga.

Bagian ini membahas bagaimana Sains dan Islam dapat bekerja sama dengan berfokus pada dua prinsip penting yang memandu penelitian ini: (1) nilai kejujuran

dan (2) nilai efisiensi (tidak membuang-buang waktu). Kedua konsep ini memberikan kerangka etika yang menggarisbawahi perlunya penelitian tentang kategorisasi anggaran dan penggunaan teknik pembelajaran mesin untuk meningkatkan akuntabilitas dalam pengelolaan keuangan negara.

a. Pentingnya Kejujuran dalam Penyelenggaraan Anggaran

Islam sangat menjunjung tinggi kejujuran, yang erat kaitannya dengan seberapa baik Anda mengelola uang. Dalam surat Al Baqarah Ayat 42 Allah SWT berfirman:

وَلَا تُلْبِسُوا الْحَقَّ بِالْبَاطِلِ وَلَا تُكْثِرُوا الْحَقَّ وَلَا تَعْلَمُونَ ﴿٤٢﴾

“Janganlah kamu campuradukkan kebenaran dengan kebatilan dan (jangan pula) kamu sembunyikan kebenaran, sedangkan kamu mengetahui(-nya) ”.

Perspektif Ibn Kathir menjelaskan larangan mencampuradukkan kebenaran dengan tipu daya, khususnya mengenai integrasi doktrin Yahudi dan Kristen dengan ajaran Islam. Bersamaan dengan itu, mereka mengakui bahwa agama Allah adalah Islam. Bagian ini juga menginstruksikan untuk tidak menyembunyikan informasi yang Anda miliki tentang kebenaran Rasulullah dan ajarannya, Sedangkan mereka mendapatkannya tertulis dalam kitab-kitab yang berada di tangannya (Munthe et al. 2024). Pada tafsir bisa diambil makna bahwa setiap prosedur yang berkaitan dengan aset publik harus bebas dari manipulasi, penyembunyian informasi, atau penyalahgunaan wewenang. Kejujuran mencakup seluruh tahapan pengelolaan anggaran, termasuk pengajuan, keputusan, pelaksanaan, pencatatan, dan pelaporan.

b. Pentingnya Efisiensi dan Larangan Pemborosan dalam Pengelolaan Anggaran

Efisiensi merupakan konsep yang sangat di junjung tinggi dalam Islam. Anggaran negara harus digunakan secara efisien, tepat sasaran, dan tidak berlebihan. Dalam surat Al-A'raf Ayat 31 Allah SWT berfirman:

يَٰٰيُّ أَدَمَ خُذُوا زِينَتَكُمْ عِنْدَ كُلِّ مَسْجِدٍ وَكُلُّوا وَاشْرَبُوا وَلَا تُسْرِفُوا إِنَّ اللَّهَ لَا يُحِبُّ الْمُسْرِفِينَ ﴿٢٦﴾

“Wahai anak cucu Adam, pakailah pakaianmu yang indah pada setiap (memasuki) masjid dan makan serta minumlah, tetapi janganlah berlebihan. Sesungguhnya Dia tidak menyukai orang-orang yang berlebihan”.

Tafsir Al Misbah Volume 5 menunjukkan bahwa ayat ini diturunkan ketika beberapa sahabat Nabi berusaha meniru kelompok al-Hummas, khususnya suku Quraisy dan keturunannya, yang menunjukkan semangat keagamaan yang tinggi, menahan diri dari melakukan Tawaf kecuali dengan mengenakan pakaian baru yang tidak ternoda dosa, dan menunjukkan kriteria ketat mengenai pemilihan dan jumlah makanan selama Haji. Para sahabat Nabi berkata, "Lebih wajar bagi kita untuk melakukan ini daripada al-Hummas." Ayat tersebut di atas menasihati dan memberikan petunjuk tentang tata cara pelaksanaannya (Dr. M. Quraish Shihab 2002). Ayat di atas bisa diambil pelajaran tentang larangan berlebihan dalam berbagai hal, tidak hanya dalam hal makanan dan minuman, akan tetapi juga ajaran inti tentang efisiensi dan larangan pemborosan dalam pengelolaan keuangan atau penganggaran. Larangan ini menyoroti gagasan keseimbangan, yang menyatakan bahwa setiap orang harus hidup sesuai kemampuan mereka (bukan "di luar jangkauan") dan menahan diri dari pengeluaran lebih dari penghasilan mereka, karena hal ini dapat mengakibatkan kerugian. Untuk meraih kesuksesan di dunia ini dan akhirat, ajaran ini merupakan pedoman etika yang menyeluruh yang menekankan nilai pengendalian diri, pengelolaan sumber daya yang bijaksana, dan menjauhi segala sesuatu yang dilarang (haram).

Dalam surat Al-Isra' Ayat 26 Allah SWT berfirman

وَاتَّدَا الْفُرْبَى حَقَّهُ وَالْمِسْكِينُونَ وَابْنَ السَّيِّئِينَ وَلَا تُنْبَدِرْ تَبَذِّيرًا ﴿٢٦﴾

“Berikanlah kepada kerabat dekat haknya, (juga kepada) orang miskin, dan orang yang dalam perjalanan. Janganlah kamu menghambur-hamburkan (hartamu) secara boros”.

Dari sudut pandang Islam, materialisme yang berlebihan disebut israf, yang dilarang dalam Al-Quran. Surah Al-Isra', ayat 26-27, menggarisbawahi larangan menghambur-hamburkan kekayaan dan menyebut orang-orang yang boros sebagai saudara setan. Ayat ini tidak hanya melarang pemborosan materi tetapi juga mengutuk pola pikir ketidakpuasan yang terus-menerus. Israf menandakan kurangnya rasa syukur dan berpotensi mengganggu tatanan sosial. Oleh karena itu,

penting untuk memahami dan menerapkan prinsip-prinsip kesederhanaan dalam kehidupan sehari-hari (Hamzah Muchtar 2025). Ayat ini bisa diambil Pelajaran tentang gagasan dasar ekonomi dan menyatakan bahwa semua jenis uang, termasuk uang yang diberikan masyarakat kepada pemerintah, tidak boleh disia-siakan (israf). Pemborosan terjadi ketika uang tidak digunakan di tempat yang tepat, pada waktu yang tepat, atau dalam jumlah yang tepat. Langkah pertama dalam menghentikan pemborosan adalah menempatkan anggaran dalam kategori yang tepat berdasarkan bagaimana uang itu dibelanjakan dan dari mana asalnya.

BAB VIII

KESIMPULAN DAN SARAN

8.1 Kesimpulan

Kesimpulan studi ini didasarkan pada penelitian yang telah dilakukan dan temuan-temuannya. Pada tahap kesimpulan dilakukan untuk memenuhi tujuan penelitian yang telah rancang sebelumnya, yaitu untuk melihat seberapa baik klasifikasi anggaran dan model pembelajaran mesin dapat memprediksi bagaimana anggaran akan digunakan.

Penerapan klasifikasi anggaran berdasarkan jenis belanja dan sumber dana pada pelaksanaan APBN di Kementerian/Lembaga dapat diidentifikasi dengan baik melalui analisis data historis dan aturan yang berlaku. Hasil analisis menunjukkan bahwa jenis belanja dan sumber dana memiliki hubungan yang signifikan terhadap tingkat pelaksanaan anggaran. Belanja rutin seperti operasional dan honorarium cenderung lebih stabil, sedangkan belanja modal atau pemeliharaan bersifat lebih bervariasi tergantung pada proses pengadaan dan kondisi di lapangan.

1. Pengukuran dan Analisis Kinerja Model Machine Learning

Random Forest Classifier (RFC) menunjukkan tingkat kinerja yang sangat tinggi dan unggul, dengan rata-rata Akurasi sebesar 90%, Presisi 90%, Recall 87%, dan F1-score mencapai 90%. Support Vector Machine (SVM) menunjukkan kinerja yang stabil, namun lebih rendah dibandingkan RFC, dengan rata-rata Akurasi 87%, Recall 87%, Presisi 75%, dan F1-score 80%.

2. Model Klasifikasi Terbaik

Random Forest Classifier (RFC) terbukti menjadi model terbaik untuk mengurutkan data anggaran K/L berdasarkan perbandingan ukuran peringkat. Hal ini didukung oleh fakta bahwa RFC memiliki skor F1 rata-rata tertinggi (90%) dan Presisi tertinggi (90%), yang menunjukkan bahwa ia dapat menemukan

keseimbangan terbaik antara mengurangi False Positives dan False Negatives serta memberikan hasil prediksi yang lebih akurat untuk pelaksanaan anggaran.

3. Rekomendasi Kebijakan

Ketika anggaran K/L diimplementasikan, saran-saran berikut dapat dibuat untuk menjadikannya lebih terbuka, akuntabel, efisien, dan selaras dengan tujuan pertumbuhan:

1. Membangun Sistem Peringatan Dini: Gunakan model Random Forest Classifier sebagai mesin utama dalam sistem peringatan dini yang dapat dengan cepat mengetahui kemungkinan kegagalan kinerja anggaran.
2. Fokus pada Pengawasan Berbasis Prediksi: Beri tahu para pembuat undang-undang untuk memfokuskan upaya audit dan pengawasan mereka pada kelompok jenis pengeluaran dan sumber pendanaan yang menurut model kemungkinan berisiko tinggi (berdasarkan analisis Pentingnya Fitur RFC).

DAFTAR PUSTAKA

- Aditiya, Dimas, e Ulinnuha Latifa. 2023. "Uji Efektivitas Penerapan Machine Learning Classification Untuk Survey Kepuasan Pelanggan Maskapai Penerbangan X". *Barometer* 8(1). doi:10.35261/barometer.v8i1.6566.
- Al-Imam Abul Fida Isma'il Kasir Ad-Dimasyqi. 2000. *Tafsir Ibnu Kasir juz 5*. SInar Baru Algensindo.
- Apriliah, Widya, Ilham Kurniawan, Muhamad Baydhowi, e Tri Haryati. 2021. "Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest". *SISTEMASI* 10(1). doi:10.32520/stmsi.v10i1.1129.
- Aryani, Yulia, e Arie Wahyu Wijayanto. 2021. "Klasifikasi Pengembalian Radar dari Ionosfer Menggunakan SVM, NaÃ¯ve Bayes dan Random Forest". *Komputika : Jurnal Sistem Komputer* 10(2). doi:10.34010/komputika.v10i2.4347.
- Dewi, Pramesti, Purwono Purwono, e Safar Kurniawan Dwi. 2022. "Pemanfaatan Teknologi Machine Learning pada Klasifikasi Jenis Hipertensi Berdasarkan Fitur Pribadi". *Smart Comp: Jurnalnya Orang Pintar Komputer* 11(3). doi:10.30591/smartcomp.v11i3.3721.
- Dr. M. Quraish Shihab. 2002. "Tafsir Al-Mishbah Jilid 05".
- Hamzah Muchtar, Evan. 2025. "Tafsir Maudhu'i q.s. Al-isra' ayat 26–27: Solusi Qur'ani Fenomena Fomo dan Gaya Hidup Konsumtif Gen-z". *JIQTa: Jurnal Ilmu Al-Qur'an dan Tafsir* 4(1): 1–14. doi:10.36769/jiqta.v4i1.937.
- Hidayat, Hidayat, Andi Sunyoto, e Hanif Al Fatta. 2023. "Klasifikasi Penyakit Jantung Menggunakan Random Forest Clasifier". *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)* 7(1). doi:10.47970/siskom-kb.v7i1.464.
- Husada, Hendry Cipta, e Adi Suryaputra Paramita. 2021. "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)". *Teknika* 10(1). doi:10.34148/teknika.v10i1.311.
- KH Zakky Mubarak. 2021. "Naungan Allah dan Perlindungan-Nya".
- Khasanah, Nurul, Rachman Komarudin, Nurul Afni, Yana Iqbal Maulana, e Agus Salim. 2021. "Skin Cancer Classification Using Random Forest Algorithm". *SISFOTENIKA* 11(2). doi:10.30700/jst.v11i2.1122.
- Louis M, e Dian B. 2022. "Perbandingan Algoritma Knn, Decision Tree,Dan Random Forest Pada Data Imbalanced Class Untuk Klasifikasi Promosi". *Jurnal INSTEK (Informatika Sains dan Teknologi)* 7(2). doi:10.24252/instek.v7i2.31385.

Lu'ul, Isnaini, Atim Muthoharoh, MAg Choirul Huda Maulidin, Allaen Firmansyah Defi Tri Wulandari, Anis Farika Citra Puspaningrum Kurniasari, e Ahmad Fakhri Hidayat Felix Gastiadirrijal Faruq Eka Ammilatul Putri Wahyuningtias. 2024. *Konsep Kepemimpinan Modern Perspektif Hadis Nabi*. www.academiapublication.com.

Munthe, Cinta Rohaini, Fadwa Nabilah, Liza Aulia, Br Manurung, Mala Purnawati, Nursumayyah Damanik, e Info Artikel. 2024. "Sulawesi Tenggara Educational Journal Etika Akademis dalam Perspektif Surat Al-Baqarah Ayat 42 dan Kaitannya dengan Asbabun Nuzul (Larangan Mencampur adukkan Kebenaran dan Kebatilan)". 4(3). [http://jurnal-unsultra.ac.id/index.php/seduj](http://jurnal-unsultra.ac.id/index.php/sedujhttp://jurnal-unsultra.ac.id/index.php/seduj)

Muthohhar, Jalis Dwi, e Agus Prihanto. 2023. "Analisis Perbandingan Algoritma Klasifikasi untuk Penyakit Jantung". *Journal of Informatics and Computer Science (JINACS)*. doi:10.26740/jinacs.v4n03.p298-304.

Nikmah, Ckamilatun, Firdausi Amalia Khoir, e Hesty Ova Noviandani Fakultas Syariah dan Hukum UIN Sunan Ampel. 2022. *Konsep Wadiah Menurut Fikih Dan (Khes)*.

PMK 62 Tahun 2023. 2023. *Peraturan Menteri Keuangan Nomor 62 Tahun 2023 tentang Perencanaan Anggaran, Pelaksanaan Anggaran, serta Akuntansi dan Pelaporan Keuangan*.

PMK 102 /PMK.02/2018. 2018. *Peraturan Menteri Keuangan RI Nomor 102 Tahun 2018 tentang klasifikasi anggaran*. www.jdih.kemenkeu.go.id (14 de setembro de 2025).

PMK 160/PMK.02/2012. *Lampiran i Peraturan menteri keuangan republik indonesia Nomor 160/pmk.02/2012 tentang petunjuk penyusunan dan pengesahan daftar isian pelaksanaan anggaran*. www.djpp.depumham.go.id.

Putri, Ridha Rahmadita, e Anang Subardjo. 2017. "Pengaruh Akuntabilitas, Transparansi, Dan Pengawasan Terhadap Kinerja Anggaran Pada Sektor Publik".

Sihombing, Pardomuan Robinson, e Istiqomatul Fajriyah Yuliati. 2021. "Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia". *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer* 20(2). doi:10.30812/matrik.v20i2.1174.

Tantika, Rina Silviany, e Abdul Kudus. 2022. "Penggunaan Metode Support Vector Machine Klasifikasi Multiclass pada Data Pasien Penyakit Tiroid". *Bandung Conference Series: Statistics* 2(2). doi:10.29313/bcss.v2i2.3590.

UU No. 17 Tahun 2003. 2023. *Undang-undang (UU) Nomor 17 Tahun 2003 tentang Keuangan Negara*. <https://peraturan.bpk.go.id/Details/43017/uu-no-17-tahun-2003> (14 de setembro de 2025).

Wang, Y., Z. Pan, J. Zheng, L. Qian, e M. Li. 2019. “A hybrid ensemble method for pulsar candidate classification”. *Astrophysics and Space Science* 364(8). doi:10.1007/s10509-019-3602-4.

LAMPIRAN

Beberapa Data Hasil Encoding

index	kode_kegiatan	urai	nilai_anggaran	jenis_belanja	sumber_dana	realisasi_anggaran	%_realisasi	ratio	pelaksanaan
0	101	191	48658800	0	1	3	75118000	0.15437	15.437
1	102	33	50880000	16507000	1	3	47394500	0.93149	93.149
2	103	269	0	16200000	1	3	114854000	0.69579	69.578
3	104	59	16200000	40390000	1	3	0	0	0
4	105	326	40390000	82000000	1	3	0	0.01207	1.2073
5	106	253	10544600	16947000	1	3	990000	3	17
6	107	313	0	35200000	1	3	0	0	0
7	108	403	0	30400000	1	3	50400000	0.29739	29.739
8	109	42	35200000	9660000	1	3	0	0	0
9	110	125	57720300	57720300	1	3	22300000	0.73355	73.355
10	111	182	9660000	96639540	1	3	8070000	3	26
11	112	144	0	25350000	1	3	0	0.83540	83.540
12	113	458	10970000	16660000	1	3	0	0.16742	16.742
13	114	30	0	21020000	1	3	76800000	7	73
14	115	506	35820000	30456500	1	3	0	0	0
15	116	223	21020000	38556000	1	3	0	0	0
16	117	299	15790000	23810200	1	3	0	0	0
17	118	74	97380000	30456500	1	3	25934562	0.26632	26.632
18	119	93	23810200	38556000	1	3	0	3	33
19	120	84	0	40500000	1	3	0	0	0
20	121	148	0	88800000	1	3	69126940	0.22696	22.696
21	122	518	40500000	30456500	1	3	22107500	0.54586	54.586
22	123	504	88800000	38556000	1	3	23041700	0.25947	25.947
23	124	136	17000000	29816000	1	3	0	9	86
24	125	511	14520000	30456500	1	3	0	0	0
25	126	449	57600000	38556000	1	3	51730000	0.89809	89.809
26	127	436	29816000	30456500	1	3	29060000	0.97464	97.464
27	128	453	15800000	38556000	1	3	0	4	45
28	129	475	5250000	4200000	1	3	4080000	0.77714	77.714
29	130	56	4200000	13280000	1	3	2100000	0.5	50
30	131	76	26494000	26494000	1	3	0	0.34964	34.964
31	132	99	0	92635320	1	3	0	6	64
32	133	491	18960000	92635320	1	3	0	0	0

33	134	24	80000000	1	3	40000000	0.5	50	0
34	135	256	10000000 25850000	1	3	0	0	0	0
35	136	117	0	1	3	0	0	0	0
36	137	81	22200000	1	3	0	0	0	0
37	138	90	90000000	1	3	90000000	1	100	2
38	139	89	54400000	1	3	6800000	0.125 0.27083	12.5 27.083	0
39	140	284	43200000 16650000	1	3	11700000	3 0.47567	33 47.567	0
40	141	490	0	1	3	79200000	6 0.37469	57 37.469	0
41	142	459	49800000	1	3	18660000	9	88	0
42	143	2	20400000	1	3	0	0	0	0
43	144	87	6800000	1	3	3400000	0.5	50	0
44	145	172	28400000	1	3	0	0	0	0
45	146	88	17229000	1	3	0	0	0	0
46	147	12	20600000	1	3	0	0	0	0
47	148	13	75550000	1	3	0	0	0	0
48	149	77	20150000	1	3	0	0	0	0
49	150	129	16000000	1	3	0	0	0	0
50	151	376	73760000	1	3	0	0	0	0
51	152	444	40000000	1	3	0	0	0	0
52	153	5	21200000	1	3	0	0 0.92429	0 92.429	0
53	154	418	27000000	1	3	24956000	6 0.09459	63 9.4594	2
54	155	79	37000000	1	3	3500000	5 0.36287	59 36.287	0
55	156	190	90000000	1	3	32658880	6 0.18776	64 18.776	0
56	157	515	43520000	1	3	8171500	4	42	0
57	158	264	54500000	1	3	0	0 0.92903	0 92.903	0
58	159	452	27000000	1	3	25084000	7	7	2
59	160	365	42000000	1	3	0	0 0.51057	0 51.057	0
60	161	357	99900000	1	3	51006587	6	64	0
61	162	328	5000000	1	3	0	0 0.19722	0 19.722	0
62	163	274	36000000	1	3	7100000	2	22	0
63	164	32	15400000 12946600	1	3	5700000	0.37013 0.55628	99 55.628	0
64	165	57	0	1	3	72020000	5	5	0
65	166	455	10950000	1	3	0	0	0	0
66	167	476	3650000	1	3	0	0	0	0
67	168	225	10000000	1	3	0	0	0	0
68	169	250	22500000	1	3	0	0	0	0
69	170	363	22600000	1	3	0	0	0	0
70	171	364	22600000	1	3	0	0 0.60437	0 60.437	0
71	172	406	17820000	1	3	10770000	7	71	1
72	173	474	4600000 23834800	1	3	0	0 0.28607	0 28.607	0
73	174	185	0	1	3	68184500	1	12	0
74	175	448	8680000	1	3	0	0	0	0

75	176	489	3210000	1	3	0	0	0	0
76	177	217	18600000	1	3	4500000	5	55	0
77	178	505	16360000	1	3	0	0	0	0
78	179	408	58931000	1	3	20250000	2	22	0
79	180	478	5873000	1	3	0	0	0	0
80	181	440	11808000	1	3	0	0	0	0
81	182	7	49640000	1	3	27040000	2	2	0
82	183	417	17010000	1	3	0	0	0	0
83	184	437	61700000	1	3	15000000	2	18	0
84	185	301	15870000	1	3	10660000	8	76	1
85	186	101	5650000	1	3	0	0	0	0
86	187	410	23796000	1	3	6950000	6	59	0
87	188	497	13275000	1	3	0	0	0	0
88	189	307	30000000	1	3	28447000	3	33	2
89	190	509	18885000	1	3	0	0	0	0
90	191	392	17200000	1	3	4800000	0.27907	98	0
91	192	210	4140000	1	3	0	0	0	0
92	193	111	13700000	1	3	6450000	3	29	0
93	194	207	10225000	1	3	4875000	3	26	0
94	195	440	5550000	1	3	5250000	6	59	2
95	196	433	14320000	1	3	13600000	1	07	2
96	197	488	6715000	1	3	6225000	9	9	2
97	198	283	41280000	1	3	0	0	0	0
98	199	295	51390000	1	3	0	0	0	0
99	200	296	17475000	1	3	0	0	0	0
100	201	265	36000000	1	3	0	0	0	0
101	202	379	18900000	1	3	0	0	0	0
102	203	124	41680000	1	3	13284000	4	4	0
103	204	216	21316000	0	1	0	0	0	0
104	205	106	54800000	1	3	0	0	0	0
105	206	122	2800000	1	3	2500000	7	71	1
106	207	208	4250000	1	3	0	0	0	0
107	208	442	6195000	1	3	0	0	0	0
108	209	289	19830000	0	1	0	0	0	0
109	210	281	13900000	1	3	0	0	0	0
110	211	416	57880000	1	3	0	0	0	0
111	212	462	10088000	0	1	17400000	2	22	0
112	213	372	26500000	1	3	0	0	0	0
113	214	292	85000000	1	3	0	0	0	0
114	215	334	39700000	0	1	251530750	9	87	1

115	216	29	78488900	0	1	3	367217610	0.46785	9	46.785	93	0
-----	-----	----	----------	---	---	---	-----------	---------	---	--------	----	---

Beberapa Data Hasil Prediksi

kode_kegiatan	uraian	nilai_anggaran	jenis_belanja	sumber_dana	realisasi_anggaran	%_realisasi	ratio	pelaksanaan	Prediksi_RF	Prediksi_SVM
		4865880								
101	191	00	1	3	75118000	0.1543	15.4	0	0	0
		5088000				0.9314	93.1			
102	33	0	1	3	47394500	96	4957	2	0	0
		1650700				0.6957	69.5			
103	269	00	1	3	1.15E+08	9	7897	1	1	0
		1620000								
104	59	0	1	3	0	0	0	0	0	0
		4039000								
105	326	0	1	3	0	0	0	0	0	0
		8200000				0.0120	1.20			
106	253	0	1	3	990000	73	7317	0	0	0
		1054460								
107	313	00	1	3	0	0	0	0	0	0
		1694700				0.2973	29.7			
108	403	00	1	3	50400000	98	3978	0	0	0
		3520000								
109	42	0	1	3	0	0	0	0	0	0
		3040000				0.7335	73.3			
110	125	0	1	3	22300000	53	5526	1	1	0
						0.8354	83.5			
111	182	9660000	1	3	8070000	04	4037	1	1	0
		5772030				0.1674	16.7			
112	144	00	1	3	96639540	27	4273	0	0	0
		2535000								
113	458	0	1	3	0	0	0	0	0	0
		1097000				0.7000	70.0			
114	30	00	1	3	76800000	91	0912	1	1	0
		3582000								
115	506	0	1	3	0	0	0	0	0	0
		2102000								
116	223	0	1	3	0	0	0	0	0	0
		1579000								
117	299	0	1	3	0	0	0	0	0	0
		9738000				0.2663	26.6			
118	74	0	1	3	25934562	23	3233	0	0	0
		3855600								
119	93	0	1	3	0	0	0	0	0	0
		2381020								
120	84	00	1	3	0	0	0	0	0	0
		3045650				0.2269	22.6			
121	148	00	1	3	69126940	69	9694	0	0	0
		4050000				0.5458	54.5			
122	518	0	1	3	22107500	64	8642	0	0	0
		8880000				0.2594	25.9			
123	504	0	1	3	23041700	79	4786	0	0	0
		1700000								
124	136	0	1	3	0	0	0	0	0	0
		1452000								
125	511	0	1	3	0	0	0	0	0	0
		5760000				0.8980	89.8			
126	449	0	1	3	51730000	9	0903	1	0	0
		2981600				0.9746	97.4			
127	436	0	1	3	29060000	44	6445	2	2	0
		1580000								
128	453	0	1	3	0	0	0	0	0	0
						0.7771	77.7			
129	475	5250000	1	3	4080000	43	1429	1	1	0
130	56	4200000	1	3	2100000	0.5	50	0	0	0
		1328000								
131	76	0	1	3	0	0	0	0	0	0

132	99	2649400 00 1896000	1	3	92635320	0.3496 46	34.9 6464	0	0	0
133	491	0 8000000	1	3	0	0	0	0	0	0
134	24	0 1000000	1	3	40000000	0.5	50	0	0	0
135	256	0 2585000	1	3	0	0	0	0	0	0
136	117	00 2220000	1	3	0	0	0	0	0	0
137	81	0 9000000	1	3	0	0	0	0	0	0
138	90	0 5440000	1	3	90000000	1	100	2	0	0
139	89	0 4320000	1	3	6800000	0.125 0.2708	12.5 27.0	0	0	0
140	284	0 1665000	1	3	11700000	33	8333	0	0	0
141	490	00 4980000	1	3	79200000	0.4756 0.3746	47.5 37.4	0	0	0
142	459	0 2040000	1	3	18660000	99	6988	0	0	0
143	2	0	1	3	0	0	0	0	0	0
144	87	6800000 2840000	1	3	3400000	0.5	50	0	1	0
145	172	0 1722900	1	3	0	0	0	0	0	0
146	88	0 2060000	1	3	0	0	0	0	0	0
147	12	0 7555000	1	3	0	0	0	0	0	0
148	13	0 2015000	1	3	0	0	0	0	0	0
149	77	0 1600000	1	3	0	0	0	0	0	0
150	129	0 7376000	1	3	0	0	0	0	0	0
151	376	0 4000000	1	3	0	0	0	0	0	0
152	444	0 2120000	1	3	0	0	0	0	0	0
153	5	0 2700000	1	3	0	0	0	0	0	0
154	418	0 3700000	1	3	24956000	0.9242 96	92.4 2963	2	2	0
155	79	0 9000000	1	3	3500000	0.0945 95	9.45 9459	0	0	0
156	190	0 4352000	1	3	32658880	0.3628 76	36.2 8764	0	0	0
157	515	0 5450000	1	3	8171500	0.1877 64	18.7 7642	0	0	0
158	264	0 2700000	1	3	0	0	0	0	0	0
159	452	0 4200000	1	3	25084000	0.9290 37	92.9 037	2	2	0
160	365	0 9990000	1	3	0	0	0	0	0	0
161	357	0	1	3	51006587	0.5105 76	51.0 5764	0	0	0
162	328	5000000 3600000	1	3	0	0	0	0	0	0
163	274	0 1540000	1	3	7100000	0.1972 22	19.7 2222	0	0	0
164	32	0 1294660	1	3	5700000	0.3701 3	37.0 1299	0	0	0
165	57	00 1095000	1	3	72020000	0.5562 85	55.6 285	0	0	0
166	455	0	1	3	0	0	0	0	0	0
167	476	3650000 1000000	1	3	0	0	0	0	0	0
168	225	0	1	3	0	0	0	0	0	0

169	250	2250000 0	1	3	0	0	0	0	0	0
170	363	2260000 0	1	3	0	0	0	0	0	0
171	364	2260000 0	1	3	0	0	0	0	0	0
172	406	1782000 0	1	3	10770000	77	3771	1	1	0
173	474	4600000 2383480	1	3	0	0	0	0	0	0
174	185	00	1	3	68184500	71	0712	0	0	0
175	448	8680000	1	3	0	0	0	0	0	0
176	489	3210000 1860000	1	3	0	0	0	0	0	0
177	217	0	1	3	4500000	35	9355	0	0	0
178	505	1636000 0	1	3	0	0	0	0	0	0
179	408	5893100 0	1	3	20250000	22	6222	0	0	0
180	478	5873000 1180800	1	3	0	0	0	0	0	0
181	440	0	1	3	0	0	0	0	0	0
182	7	4964000 0	1	3	27040000	22	722	0	0	0
183	417	1701000 0	1	3	0	0	0	0	0	0
184	437	6170000 0	1	3	15000000	12	1118	0	0	0
185	301	1587000 0	1	3	10660000	08	7076	1	1	0
186	101	5650000 2379600	1	3	0	0	0	0	0	0
187	410	0	1	3	6950000	66	0659	0	0	0
188	497	1327500 0	1	3	0	0	0	0	0	0
189	307	3000000 0	1	3	28447000	33	2333	2	2	0
190	509	1888500 0	1	3	0	0	0	0	0	0
191	392	1720000 0	1	3	4800000	7	0698	0	0	0
192	210	4140000 1370000	1	3	0	0	0	0	0	0
193	111	0	1	3	6450000	03	8029	0	0	0
194	207	1022500 0	1	3	4875000	73	7726	0	0	0
195	440	0	1	3	6225000	29	029	2	2	0
196	433	5550000 1432000	1	3	5250000	46	9459	2	2	0
197	488	0	1	3	0	0	0	0	0	0
198	283	0	1	3	0	0	0	0	0	0
199	295	5139000 0	1	3	0	0	0	0	0	0
200	296	1747500 0	1	3	0	0	0	0	0	0
201	265	3600000 0	1	3	0	0	0	0	0	0
202	379	1890000 0	1	3	0	0	0	0	0	0
203	124	4168000 0	1	3	0	0	0	0	0	0
204	216	2131600 00	1	3	0	0	0	0	0	0
205	106	5480000 0	1	3	0	0	0	0	0	0

206	122	2800000	1	3	2500000	0.8928	89.2	8571	1	1	0
207	208	4250000	1	3	0	0	0	0	0	0	0
208	442	6195000	1	3	0	0	0	0	0	0	0
		1983000									
209	289	00	1	3	0	0	0	0	0	0	0
		1390000									
210	281	0	1	3	0	0	0	0	0	0	0
		5788000									
211	416	0	1	3	0	0	0	0	0	0	0
		1008800				0.1724	17.2				
212	462	00	1	3	17400000	82	4822		0	0	0
		2650000									
213	372	0	1	3	0	0	0	0	0	0	0
		8500000									
214	292	0	1	3	0	0	0	0	0	0	0
		3970000				0.6335	63.3				
215	334	00	1	3	2.52E+08	79	5787		1	1	0
		7848890				0.4678	46.7				
216	29	00	1	3	3.67E+08	59	8593		0	0	0
		7860000				0.2187	21.8				
217	389	0	1	3	17190000	02	7023		0	0	0
		4480000				0.5543	55.4				
218	1	00	1	3	2.48E+08	88	388		0	1	0
		4107400									
219	345	0	1	3	0	0	0	0	0	0	0
		4978600				0.2835	28.3				
220	517	0	1	3	14116800	5	5496		0	0	0
		1105900									
221	361	0	1	3	0	0	0	0	0	0	0
		2034900				0.8278	82.7				
222	68	0	1	3	16846000	54	8539		1	1	0
		1131700									
223	431	0	1	3	0	0	0	0	0	0	0
		1272400									
224	176	0	1	3	0	0	0	0	0	0	0
		2748000									
225	19	3161000	1	3	0	0	0	0	0	0	0
		3161000				0.9889	98.8				
226	261	000	1	3	3.13E+09	28	9276		2	2	0
		1272000				0.2948	29.4				
227	419	0	1	3	3750000	11	8113		0	0	0
		2432000				0.4874	48.7				
230	293	00	1	1	1.19E+08	34	4342		0	0	0
		4000000									
231	20	00	1	1	0	0	0	0	0	0	0
		1100000									
232	21	000	1	1	0	0	0	0	0	0	0
		3224500				0.9944	99.4				
233	294	00	1	1	3.21E+08	18	4177		2	1	0
		1.09264				0.3575	35.7				
234	164	E+11	1	1	3.91E+10	97	5966		0	0	0
		2082150				0.4111	41.1				
235	409	00	1	1	85600000	14	1135		0	0	0
		5677040				0.3884	38.8				
236	43	000	1	1	2.21E+09	07	4066		0	0	0
		1850800									
237	178	0	1	1	0	0	0	0	0	0	0
		2233750				0.3907	39.0				
238	189	00	1	1	87280000	33	7331		0	0	0
		2573167				0.2571	25.7				
239	432	000	1	1	6.62E+08	15	1151		0	0	0
		3673379									
240	471	000	1	1	0	0	0	0	0	0	0
		3050674				0.7375	73.7				
241	273	000	1	1	2.25E+09	42	5419		1	1	0
		1414393				0.4286	42.8				
242	186	000	1	1	6.06E+08	13	6127		0	0	0
		3271250				0.7415	74.1				
243	25	000	1	1	2.43E+09	21	5208		1	1	0