

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Pendapatan negara adalah pemasukan yang diperoleh pemerintah dari berbagai sumber untuk membiayai pengeluaran negara, seperti pembangunan infrastruktur, layanan publik, pendidikan, kesehatan, serta pertahanan dan keamanan. Sumber utama pendapatan negara terdiri dari penerimaan pajak dan Penerimaan Negara Bukan Pajak (PNBP), di mana pajak memiliki peran dominan dalam menjaga stabilitas ekonomi dan fiskal serta sebagai alat pengatur kebijakan ekonomi. Pendapatan negara berperan penting dalam memastikan kelangsungan pembangunan dan kesejahteraan masyarakat, karena digunakan untuk meningkatkan kualitas hidup warga negara, menciptakan lapangan kerja, dan menjaga ketahanan ekonomi. Selain itu, pendapatan negara juga mendukung keberlanjutan program sosial dan kebijakan fiskal yang bertujuan untuk mengurangi kesenjangan ekonomi dan meningkatkan daya saing nasional di tingkat global [1].

Namun, permasalahan muncul ketika realisasi pendapatan negara sering kali tidak sesuai dengan estimasi yang telah direncanakan. Faktor-faktor seperti perlambatan ekonomi global, fluktuasi harga komoditas, serta perubahan kebijakan fiskal berkontribusi terhadap ketidakakuratan dalam proyeksi pendapatan negara. Menurut data dari Kementerian Keuangan, realisasi pendapatan negara pada tahun 2020 hanya mencapai 96,10% dari target yang telah ditetapkan, dengan penurunan tajam pada penerimaan pajak sebesar 19,67% dan Penerimaan Negara Bukan Pajak (PNBP) sebesar 16,42%. Ketidakakuratan dalam proyeksi pendapatan negara ini dapat menghambat perencanaan anggaran yang efektif, menyebabkan defisit fiskal, serta mengancam keberlanjutan program pembangunan dan kesejahteraan masyarakat [2]. Oleh karena itu, diperlukan metode peramalan yang lebih akurat dan adaptif untuk memproyeksikan realisasi pendapatan negara yang sesuai dengan kondisi ekonomi yang dinamis.

Perkembangan teknologi machine learning telah membawa transformasi yang signifikan dalam metode peramalan ekonomi modern. Penelitian terbaru

menunjukkan bahwa algoritma machine learning telah digunakan untuk menganalisis data pajak historis, mengidentifikasi pola, dan membuat prediksi yang akurat tentang tren pajak masa depan dengan kemampuan yang unggul dalam menangani dataset yang besar dan kompleks, yang mungkin sulit ditangani oleh metode tradisional [3]. Swenson (2024) dalam studinya tentang peramalan pajak korporasi menggunakan deep learning mendemonstrasikan bahwa neural networks, decision tree, dan model random forest memberikan peramalan yang robust meskipun mencakup periode pandemi COVID-19, menunjukkan ketahanan model machine learning terhadap gangguan ekonomi [3]. Penelitian ini juga menegaskan bahwa dari perspektif kebijakan, peramalan yang lebih akurat terhadap pembayaran pajak oleh perusahaan akan menghasilkan pendapatan pajak yang lebih dapat diprediksi bagi pemerintah [3].

Metode peramalan tradisional yang saat ini digunakan oleh institusi pemerintah, seperti model Autoregressive Integrated Moving Average (ARIMA) dan Vector Autoregression (VAR), memiliki keterbatasan fundamental dalam menangkap kompleksitas hubungan non-linear antar variabel ekonomi. Zhang et al. (2024) dalam tinjauan komprehensif mereka tentang model pembelajaran mendalam untuk peramalan deret waktu keuangan menjelaskan bahwa metode statistik tradisional mungkin tidak dapat menangkap pola non-linear atau hubungan kompleks yang ada dalam deret waktu keuangan dunia nyata [4]. Model-model tradisional mengandalkan struktur matematika yang telah ditentukan sebelumnya dan asumsi statistik untuk memodelkan dan menganalisis data deret waktu namun sering mengalami keterbatasan dalam menghadapi kondisi ekonomi yang dinamis [4].

Kemajuan terbaru dalam teknologi deep learning telah terbukti menghasilkan peningkatan akurasi yang signifikan dalam peramalan ekonomi. Zhang et al. (2024) menegaskan bahwa jaringan Long Short-Term Memory (LSTM), varian terkenal dari Recurrent Neural Networks (RNN), telah menunjukkan kesuksesan yang luar biasa dalam menangkap dependensi temporal antara titik data dalam deret waktu keuangan dan telah menggantikan metode statistik tradisional dan machine learning sebagai pilihan utama untuk tugas peramalan harga [4]. Lebih lanjut, mereka menjelaskan bahwa teknik deep learning memberikan pengaruh transformatif

dalam bidang ini karena kemampuannya untuk secara otomatis belajar dan beradaptasi dengan pola yang rumit sambil menangkap dependensi linear dan non-linear dalam data [4].

Longo et al. (2022) dalam penelitiannya tentang ensemble neural networks untuk peramalan GDP menunjukkan bahwa pendekatan ensemble learning yang menggabungkan Recurrent Neural Network (RNN) dan model Faktor Dinamis yang memperhitungkan variasi waktu dalam rata-rata dengan Generalized Autoregressive Score (DFM-GAS) dapat secara signifikan meningkatkan akurasi peramalan indikator makroekonomi [5]. Penelitian mereka mendemonstrasikan bahwa model ensemble mengungguli berbagai model alternatif dengan meramalkan tingkat pertumbuhan GDP AS kuartalan satu kuartal ke depan pada periode 2005Q2-2020Q1 dan menunjukkan bahwa komponen machine learning sangat berguna di masa krisis, ketika terjadi perubahan struktural [5].

Ketersediaan dataset ekonomi komprehensif dari International Monetary Fund (IMF) dengan metodologi standar dan cakupan historis yang diperluas membuka peluang untuk mengembangkan model peramalan yang lebih canggih. Dataset IMF World Economic Outlook Database menyediakan data yang konsisten dan dapat dibandingkan secara internasional untuk berbagai indikator ekonomi dengan jaminan kualitas yang tinggi [6]. Dataset ini mencakup empat indikator ekonomi fundamental: pendapatan pemerintah, pengeluaran pemerintah, tingkat pertumbuhan GDP riil, dan utang publik bruto sebagai persentase GDP, yang memungkinkan analisis hubungan ekonomi yang komprehensif. Dengan kedalaman historis yang mencakup berbagai siklus ekonomi termasuk berbagai periode krisis, dataset ini ideal untuk melatih model neural networks lanjutan yang dapat belajar dari pola ekonomi kompleks dan adaptif terhadap kondisi ekonomi yang berubah.

Model Deep Autoregressive (DeepAR) telah muncul sebagai metodologi terdepan untuk peramalan deret waktu probabilistik. Salinas et al. (2020) dalam makalah fundamental mereka menunjukkan bahwa DeepAR menghasilkan peramalan yang lebih akurat daripada metode canggih lainnya, sambil memerlukan pekerjaan manual yang minimal dan dapat menghasilkan peningkatan akurasi sekitar 15% dibandingkan dengan metode canggih lainnya [7]. Keunggulan utama

DeepAR terletak pada kemampuannya untuk menghasilkan peramalan probabilistik yang sangat penting untuk memiliki inventori yang tepat di waktu yang tepat dan di tempat yang tepat dalam konteks perencanaan dan manajemen risiko [7]. Model ini tidak hanya menghasilkan peramalan titik, tetapi juga menyediakan distribusi probabilitas yang memungkinkan optimalisasi proses bisnis dengan kuantifikasi ketidakpastian yang berharga untuk pengambilan keputusan dalam kondisi ketidakpastian [7].

Metode DeepAR digunakan dalam penelitian ini untuk memprediksi pendapatan negara Indonesia, karena kemampuannya yang telah terbukti dalam menangkap pola temporal jangka panjang, tren, dan pola musiman pada data multivariat yang kompleks, serta menghasilkan prediksi probabilistik yang mencakup distribusi kemungkinan dengan kuantifikasi ketidakpastian. Penelitian ini bertujuan untuk menganalisis performa model DeepAR dibandingkan dengan metode baseline tradisional menggunakan dataset komprehensif IMF periode 1980-2024, yang diharapkan dapat meningkatkan akurasi prediksi pendapatan negara secara signifikan dan mendukung perencanaan anggaran yang lebih stabil dan terukur, mengurangi risiko kesalahan dalam pengalokasian anggaran, serta memastikan pemanfaatan pendapatan negara yang optimal untuk program pembangunan melalui implementasi sistem peramalan yang siap produksi dan ramah pengguna.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam penelitian ini meliputi:

1. Bagaimana penerapan model DeepAR dalam memprediksi realisasi pendapatan negara Indonesia menggunakan dataset IMF dengan 4 variabel ekonomi periode 1980-2024?
2. Seberapa signifikan peningkatan performa model DeepAR yang dioptimasi menggunakan Bayesian optimization dibandingkan dengan model baseline DeepAR (hyperparameter default) dalam memprediksi realisasi pendapatan negara berdasarkan metrik evaluasi MAPE, RMSE, dan CRPS?
3. Bagaimana implementasi model DeepAR dalam bentuk aplikasi web Flask yang dapat digunakan untuk peramalan pendapatan negara?

### **1.3. Tujuan Penelitian**

Adapun tujuan dari penelitian ini dibedakan menjadi 2 tujuan yaitu, tujuan umum dan tujuan khusus:

#### **1. Tujuan Umum**

Untuk menganalisis dan mengevaluasi performa metode DeepAR dalam memprediksi pendapatan negara Indonesia menggunakan dataset multivariat IMF periode 1980-2024, membandingkannya dengan dua model baseline, serta mengimplementasikannya dalam bentuk aplikasi web untuk mendukung perencanaan fiskal yang lebih akurat dan berbasis data.

#### **2. Tujuan Khusus**

- a) Menerapkan metode DeepAR sebagai model peramalan probabilistik untuk memprediksi realisasi pendapatan negara Indonesia menggunakan dataset komprehensif IMF yang mencakup pendapatan pemerintah, pengeluaran pemerintah, tingkat pertumbuhan GDP riil, dan utang publik bruto sebagai persentase GDP periode 1980-2024.
- b) Mengevaluasi performa model menggunakan metrik MAPE dan RMSE untuk perbandingan umum dengan model baseline, serta metrik CRPS untuk mengukur keunggulan peramalan probabilistik DeepAR.
- c) Mengembangkan dan mengimplementasikan aplikasi web berbasis Flask yang mengintegrasikan model DeepAR optimal untuk memberikan layanan peramalan pendapatan negara.

### **1.4. Manfaat Penelitian**

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

#### **1. Bagi Akademisi**

Penelitian ini memberikan kontribusi ilmiah dalam pengembangan metode peramalan deret waktu ekonomi menggunakan pendekatan deep learning, khususnya model DeepAR dengan dataset multivariat IMF periode 1980-2024. Kajian ini dapat menjadi referensi bagi penelitian selanjutnya yang berfokus pada prediksi pendapatan negara menggunakan indikator ekonomi makro, serta memperkaya literatur terkait model probabilistik dan analisis ketidakpastian dalam forecasting ekonomi emerging market. Penelitian ini juga berkontribusi

pada pengembangan metodologi evaluasi komprehensif yang membandingkan model deep learning dengan baseline tradisional menggunakan metrik MAPE, RMSE, dan CRPS.

## 2. Bagi Praktisi

Penelitian ini dapat menjadi acuan bagi praktisi di bidang data science, ekonomi, dan perencanaan keuangan publik yang tertarik menerapkan model probabilistik untuk peramalan data ekonomi. Studi ini menyajikan implementasi DeepAR secara lengkap mulai dari preprocessing dataset IMF, optimisasi hyperparameter menggunakan Bayesian optimization, hingga deployment dalam bentuk aplikasi web Flask siap produksi, sehingga memberikan panduan praktis untuk pengembangan sistem peramalan ekonomi yang dapat diimplementasikan secara real-time.

### 1.5. Batasan Masalah

Agar penelitian ini lebih terfokus dan tidak meluas dari pembahasan yang dimaksudkan, maka penelitian ini membatasi beberapa aspek sebagai berikut:

1. Dataset yang digunakan terbatas pada data ekonomi Indonesia dari International Monetary Fund (IMF) dengan empat indikator fundamental: pendapatan pemerintah, pengeluaran pemerintah, tingkat pertumbuhan GDP riil, dan utang publik bruto sebagai persentase GDP, dengan rentang waktu dari tahun 1980 hingga 2024.
2. Evaluasi performa model dilakukan menggunakan tiga metrik utama: Mean Absolute Percentage Error (MAPE) dan Root Mean Squared Error (RMSE) untuk perbandingan umum dengan model baseline, serta Continuous Ranked Probability Score (CRPS) untuk mengukur keunggulan peramalan probabilistik DeepAR.
3. Model perbandingan dibatasi pada DeepAR baseline (hyperparameter default), dan DeepAR optimal hasil Bayesian optimization, tanpa melibatkan model deep learning lainnya.
4. Penelitian ini difokuskan pada peramalan pendapatan pemerintah sebagai persentase GDP sebagai variabel target, dengan tiga variabel ekonomi lainnya sebagai variabel penyerta, tanpa melibatkan variabel eksternal tambahan seperti indikator pasar saham, komoditas, atau faktor geopolitik.