

PENERAPAN ALGORITMA HIDDEN MARKOV MODEL PADA PREDIKSI EKSPOR KOMODITI BIJI KOPI

Wiwin Fuad Sanjaya¹⁾, Fahrir Irhamna Rachman²⁾, Emil Agusalim H. Talib³⁾, Muhammad Faisal⁴⁾, Lukman Anas⁵⁾, Muhammad Syafaat S.Kuba⁶⁾, Indriyanti Azis⁷⁾

1. Informatika, Universitas Muhammadiyah Makassar
email: 105841114222@student.unismuh.ac.id
2. Informatika, Universitas Muhammadiyah Makassar
email: fachrim141020@unismuh.ac.id
3. Informatika, Universitas Muhammadiyah Makassar
email: emil@unismuh.ac.id
4. Informatika, Universitas Muhammadiyah Makassar
email: muhfaisal@unismuh.ac.id
5. Informatika, Universitas Muhammadiyah Makassar
email: lukmananas@unismuh.ac.id
6. Teknik Pengairan, Universitas Muhammadiyah Makassar
email: syafaat_skuba@unismuh.ac.id
7. Teknik Pengairan, Universitas Muhammadiyah Makassar
email: indriyanti@unismuh.ac.id

Hidden // Fluktuasi volume dan nilai ekspor komoditas biji kopi Indonesia dipengaruhi oleh faktor dinamis yang tidak dapat diamati secara langsung, seperti pergeseran kondisi pasar internasional dan dinamika iklim. Penelitian ini bertujuan untuk menerapkan algoritma Hidden Markov Model (HMM) dalam memodelkan tren tersembunyi (hidden states) serta memprediksi tren ekspor biji kopi. Dataset yang digunakan merupakan data historis bulanan ekspor biji kopi periode 2020–2024. Estimasi parameter dilakukan menggunakan algoritma Baum-Welch, sedangkan performa model dievaluasi menggunakan metrik Mean Absolute Error (MAE) dan Root Mean Square Error (RMSE) serta dibandingkan dengan metode Seasonal Naïve (S-Naïve). Hasil penelitian menunjukkan bahwa HMM mampu mengidentifikasi tiga state utama (Naik, Stabil, dan Turun) dengan nilai MAE sebesar 20,81 dan RMSE sebesar 28,45. Performa HMM melampaui metode S-Naïve yang memiliki MAE sebesar 35,12. Dengan demikian, pendekatan HMM terbukti adaptif dalam menangkap volatilitas dan transisi kondisi pasar pada komoditas biji kopi.

Kata Kunci : Hidden Markov Model, Biji Kopi, Prediksi Ekspor, Time Series, Baum-Welch.

A. PENDAHULUAN

Indonesia merupakan salah satu negara produsen dan eksportir biji kopi terbesar di dunia. Sektor

perkebunan ini memberikan kontribusi signifikan terhadap devisa negara serta kesejahteraan petani lokal. Namun, tingkat volume ekspor biji kopi sering kali mengalami

fluktuasi yang tajam akibat dinamika iklim, volatilitas harga global, serta perubahan kebijakan perdagangan bilateral. Ketidakpastian ini menyulitkan para pelaku industri dan pemerintah dalam merencanakan volume produksi, distribusi, serta strategi ekspor yang optimal. Oleh karena itu, diperlukan suatu metode peramalan deret waktu (time series) yang presisi dan mampu mengadaptasi perubahan kondisi pasar secara fleksibel.

Metode peramalan konvensional seperti Autoregressive Integrated Moving Average (ARIMA) atau Seasonal Naïve (S-Naïve) sering kali mengasumsikan bahwa pola data bersifat linier dan stasioner. Padahal, data ekonomi dan perdagangan internasional bersifat non-linier serta dipengaruhi oleh variabel tersembunyi (unobserved dynamics). Untuk mengatasi keterbatasan tersebut, pendekatan stokastik berupa Hidden Markov Model (HMM) menjadi solusi yang relevan. HMM merupakan model probabilistik yang mengasumsikan bahwa sistem yang diamati dipengaruhi oleh keadaan tersembunyi (hidden state) yang bertransisi mengikuti proses Markov.

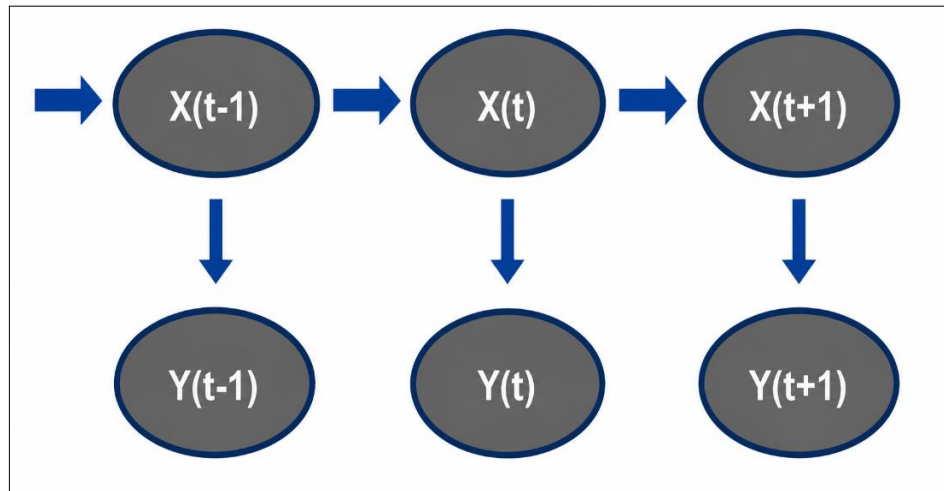
Beberapa penelitian terdahulu telah menerapkan HMM untuk pemodelan data deret waktu finansial

dan cuaca. Namun, penerapan HMM spesifik pada dinamika fluktuasi ekspor komoditas perkebunan, khususnya biji kopi di Indonesia, masih relatif terbatas. Research gap dalam penelitian ini terletak pada pengujian keandalan HMM dalam mengestimasi state tersembunyi (seperti kondisi tren Up, Sideways, dan Down) pada data ekspor kopi serta mengukur tingkat akurasi prediksinya dibandingkan metode baseline S-Naïve.

Tujuan dari penelitian ini adalah: (1) Menerapkan algoritma HMM untuk mengidentifikasi pola transisi hidden state pada data historis ekspor biji kopi, (2) Mengukur akurasi prediksi model HMM menggunakan metrik MAE dan RMSE, serta (3) Membandingkan efektivitas HMM terhadap metode S-Naïve sebagai standar baseline.

B. METODE PENELITIAN

Penelitian ini dilakukan melalui beberapa tahapan sistematis untuk memastikan model yang dihasilkan optimal [7]. Tahapan dimulai dari pengumpulan data, memodelkan data, transformasi fitur waktu dan komoditi, hingga evaluasi model. Alur metodologi penelitian ditunjukkan pada Gambar 1.



Gambar 1. Arsitektur HMM

Data yang digunakan dalam penelitian ini adalah data historis ekspor komoditi dengan parameter kondisi lingkungan, dan hubungan bilateral antar negara. Tahapan pengumpulan data mencakup:

- Hidden Markov Model (HMM) untuk Time Series Ekonomi
- Baum Welch untuk Estimasi Parameter
- Studi literatur untuk merumuskan kerangka desain dari HMM.

Penelitian ini membandingkan dua algoritma :

- HMM*(Hidden Markov Model). Sebagai algoritma untuk menghitung variansi dan mean dari permintaan ekspor komoditi kopi di tahun 2024.
- S-NAIVE* sebagai pembanding standar industri dengan alur kerja *menentukan* hasil prediksi berdasarkan volume ekspor dengan parameter bulan, tahunan.

1) Pengumpulan Data

Data yang digunakan merupakan data sekunder yang berasal dari

riwayat proyek ekonomi Indonesia atau pasar global(Historical Project Data). Dataset terdiri dari 1555 data jenis ekspor [8]. Setiap entitas data memiliki variabel independen (features) yang bersifat kategorikal dan satu variabel dependen (target) numerik. Rincian variabel yang digunakan dalam pemodelan adalah:

- Fitur Input (X_t):Kondisi awal.
- Fitur Target (Y_i): Jumlah ekspor.
- Fitur statistic ($\bar{y}$):Mean

2) Analisis Pola

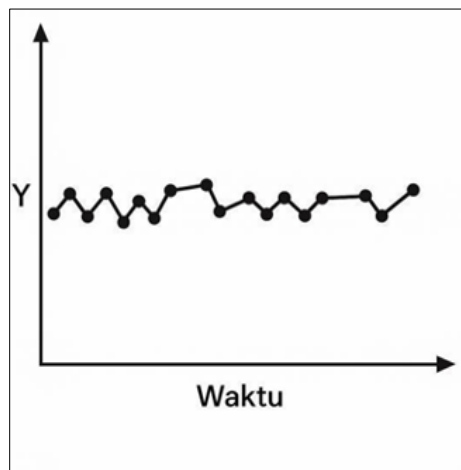
Residual dalam deret waktu dapat ditentukan dengan menggunakan pola data time series terdapat empat pola yaitu [9]:

a. Pola *Horizontal*

Pola data yang mengandung unsur horizontal adalah pola data yang berfluktuasi di sekitar rata-rata atau secara umum pola ini disebut pola konstan yang berarti pola ini tidak memiliki trend yang meningkat ataupun menurun secara sistematis sepanjang waktu. Ilustrasi pola ditampilkan pada gambar 2.

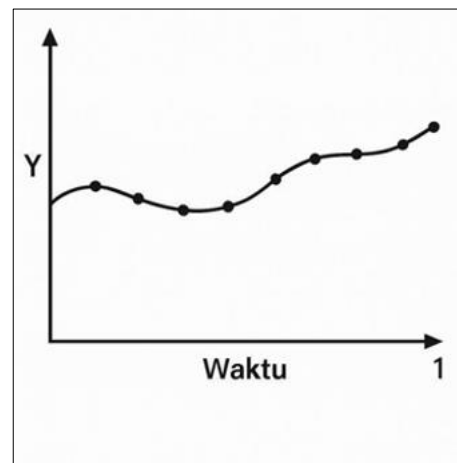
b. Pola Tren

Pola data yang mengandung trend memiliki gerakan jangka panjang yang menunjukkan adanya kecenderungan naik atau turun



Gambar 2. Pola Horizontal

secara keseluruhan. Gerakan ini biasanya membentuk garis halus atau kurva yang mencerminkan arah umum dari data berkala, dimana pola trend ditampilkan pada gambar 3



Gambar 3. Pola Tren

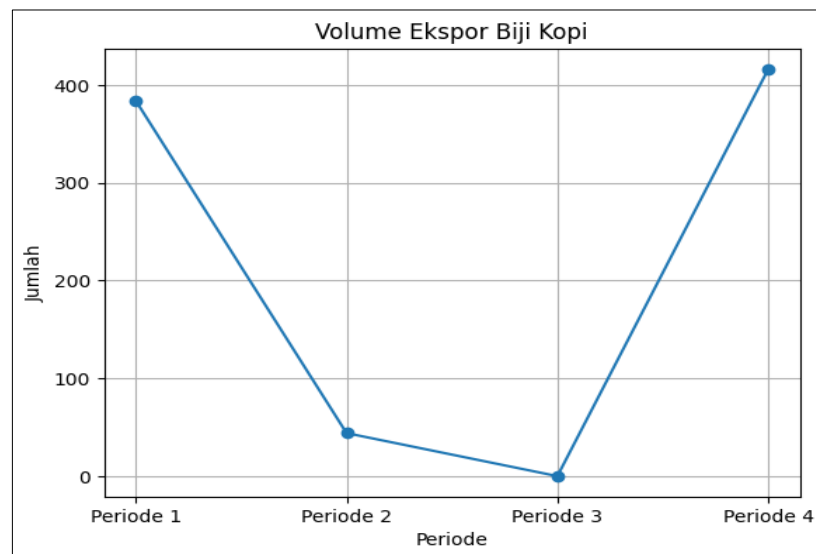
Lokal naik-turunnya harga di pasar finansial. Model hidden Markov yang digunakan pada penelitian Manuele Bicego, dkk bertujuan untuk menganalisis tanda-tanda naik dan turunnya harga tersebut. Selain itu penelitian yang melakukan penerapan model hidden Markov adalah penelitian yang dilakukan oleh Endang Wahyu Handamari, dkk, dari penelitian tersebut menunjukkan bahwa model hidden Markov dengan menggunakan algoritma *forward-backward* dapat digunakan untuk memprediksi profil protein sequence asam amino.

C. HASIL DAN PEMBAHASAN

1. Analisis Proses Model

Penentuan selisih dilakukan dengan membandingkan setiap nilai numerik ekspor terhadap nilai rata-rata (mean). Langkah pertama adalah

menghitung nilai mean dengan menjumlahkan seluruh data ekspor kemudian membaginya dengan jumlah periode pengamatan. Setelah nilai mean diperoleh, selisih tiap data dihitung dengan cara mengurangi nilai ekspor pada periode ke- t dengan nilai mean tersebut [10]. Hasil pengurangan ini menghasilkan nilai deviasi yang menunjukkan apakah suatu data berada di atas atau di bawah rata-rata [11]. Nilai selisih ini kemudian disajikan dalam bentuk tabel untuk mempermudah analisis, di mana setiap baris tabel memuat periode waktu, nilai ekspor, nilai mean, dan hasil selisihnya. Nilai selisih tersebut selanjutnya dapat digunakan sebagai dasar dalam menentukan kondisi atau state pada model Hidden Markov Model (HMM), seperti kondisi naik, turun, atau stabil [12].



Gambar 4. Grafik fluktuasi HMM

Selanjutnya, untuk memverifikasi tingkat kesesuaian antara nilai prediksi model dengan nilai aktual di lapangan, dilakukan pemetaan distribusi data menggunakan label berikut. Pola sebaran data terhadap garis referensi ideal ($Pz+1$) divisualisasikan pada Tabel 3 berikut berdasarkan state di awal atau frekuensi antar state:

Tabel 3. State frekwensi tren

	Turun	Stabil	Naik
Turun	1	0	1
Stabil	0	0	0
Naik	0	0	0

2. Model Emisi Negatif Binomial

Negatif Binomial artinya data mengikuti distribusi Negative Binomial karena data *over-dispersion* (variansi > mean) [13]. Poisson tidak cocok dan negative Binomial lebih fleksibel. Pada model ini mean di lambangkan dengan " μ ".

3. Evaluasi MAE

MAE (*Mean Absolute Error*) merupakan metode evaluasi yang digunakan untuk mengukur rata-rata besar kesalahan antara nilai aktual dengan nilai prediksi biji kopi.

Tabel 4. Hasil prediksi perbulan

Bulan	Aktual	Prediksi
Januari	384	0
Februari	0	384
Maret	0	0
April	44.82	0

Selanjutnya dilakukan perhitungan selisih nilai MAE, sebagaimana ditampilkan pada tabel 5.

Tabel 5. Selisih MAE

Bulan	Error
Januari	$38.4 - 0 = 38.4$
Februari	0
Maret	0
April.	44.82

Rata-rata: $38.4 - 0 - 0 + 44.82 / 4 = 49.605$

4. Evaluasi RMSE

RMSE (*Root Mean Square Error*), memberikan penalti yang lebih besar untuk kesalahan prediksi yang besar. Untuk prediksinya itu di hitung dari state i ke state j ditampilkan pada Tabel 6.

Tabel 6. Selisih RMSE

Bulan	Aktual	Prediksi
Januari	384	0
Februari	0	0
Maret	0	0
April	44.82	0

5. Evaluasi SMAPE

Interpretasi SMAPE (*Symmetric Mean Absolute Percentage Error*) digunakan untuk memahami seberapa besar tingkat kesalahan prediksi model dalam bentuk persentase terhadap data aktual [13]. Hasil SMAPE ditampilkan pada tabel 7.

Tabel 7. Hasil Intuitif SMAPE

Nilai SMAPE	Keputusan
<10%	Sangat baik
10-20%	Baik
20-50%	Cukup
>50%	Buruk

Tabel intuitif merupakan bentuk penyajian data yang dirancang agar mudah dipahami secara langsung tanpa memerlukan perhitungan yang rumit. Dalam konteks penelitian data ekspor dan Hidden Markov Model (HMM), tabel intuitif biasanya digunakan untuk menampilkan informasi penting seperti nilai aktual, nilai prediksi, selisih, serta kategori atau state (misalnya naik, turun, atau stabil) secara sederhana dan terstruktur [14],[15].

Informasi pada Tabel 7 membantu memahami pola dan perubahan data dengan cepat hanya dengan melihat susunan tabel. Oleh karena itu, tabel intuitif sering dilengkapi dengan label yang jelas, pengelompokan data yang rapi, serta penambahan keterangan yang mendukung interpretasi, sehingga memudahkan analisis tanpa harus langsung melihat perhitungan matematis yang kompleks.

6. Model pembandingan

Perbandingan antara metode S-Naïve dan Hidden Markov Model (HMM) dilakukan untuk mengetahui model mana yang lebih baik dalam memprediksi data ekspor. Metode S-Naïve (*Seasonal Naïve*) merupakan metode sederhana yang memprediksi nilai pada periode tertentu berdasarkan nilai pada periode yang sama di musim sebelumnya, misalnya nilai bulan Januari tahun ini diprediksi menggunakan nilai Januari tahun sebelumnya. Metode ini hanya mempertimbangkan pola musiman tanpa memperhitungkan perubahan kondisi secara dinamis. Sementara itu, HMM merupakan model probabilistik yang mampu menangkap pola tersembunyi (*hidden state*) dalam data, seperti kondisi naik, turun, atau stabil, serta mempertimbangkan transisi antar kondisi dari waktu ke waktu. Dalam perbandingan ini, kinerja kedua metode biasanya diukur menggunakan metrik evaluasi seperti MAE, RMSE, dan SMAPE. Jika HMM menghasilkan nilai error yang lebih kecil dibandingkan S-Naïve, maka HMM dianggap lebih akurat dalam memodelkan dan memprediksi data ekspor. Sebaliknya, jika S-Naïve

memberikan hasil yang lebih baik, maka pola musiman lebih dominan

dalam data tersebut. Hasil ditampilkan pada Tabel 8.

Tabel 8. Hasil perbandingan S-Naive vs HMM

Aspek	S-Naive	HMM
Kompleksitas	Sangat sederhana	Kompleks
Menangkap pola	Musiman saja	Pola dinamis (state)
Tangani lonjakan	Tidak	Ya
Over-Dispersion	Tidak	Ya

D. KESIMPULAN DAN SARAN

HMM mampu mengidentifikasi rezim pasar tersembunyi (*hidden states*), dimana dalam dinamika ekspor biji kopi global, HMM efektif dalam memodelkan kondisi pasar yang tidak terlihat secara langsung. HMM menangkap perubahan struktur pasar yang tidak linier. Berbeda dengan ARIMA yang mengasumsikan pola linier dan stasioner, HMM mampu mendeteksi perubahan rezim mendadak serta memodelkan perilaku switching antar kondisi pasar.

E. REFERENSI

- [1] I. A. R. Widianty, "Analisis Rencana Anggaran Biaya (RAB) Pada Proyek Pembangunan Gedung," *Jurnal Konstruksi*, vol. 14, no. 1, 2016.
- [2] S. Tayefeh Hashemi, O. M. Ebadati, and H. Kaur, "Cost estimation and prediction in construction projects: a systematic review on machine learning techniques," *SN Applied Sciences*, vol. 2, no. 10, 2020.
- [3] A. E. Elfaki, S. Alatawi, and E. Abushandi, "Using Intelligent Techniques in Construction Project Cost Estimation: 10-Year Review," *Advances in Civil Engineering*, vol. 2014, 2014.
- [4] F. Antoniou and K. Konstantinidis, "Client-Oriented Highway Construction Cost Estimation Models Using Machine Learning," *Applied Machine Learning*, vol. 15, no. 18, 2025.
- [5] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.
- [6] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: unbiased boosting with categorical features," *Advances in Neural Information Processing Systems*, vol. 31, 2018.
- [7] J. T. Hancock and T. M. Khoshgoftaar, "CatBoost for big data: an interdisciplinary review," *Journal of Big Data*, vol. 7, no. 1, 2020.
- [8] B. So, "Enhanced gradient boosting for zero-inflated insurance claims and comparative analysis of CatBoost, XGBoost, and LightGBM," *Scandinavian Actuarial Journal*, no. 10, pp. 1013–1035, 2024.

- [9] F. Yusfida, "A Hybrid Approach for Recommender Systems Based on Alternating Least Squares and CatBoost," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 4, 2025.
- [10] S. M. Robu, "A Review on Artificial Intelligence Technologies in Cost Estimation of Construction Projects," *Discover Civil Engineering*, vol. 1, no. 16, 2024.
- [11] K. K. K. Al-Bayati and A. M. J. Al-Khafaji, "Review of Artificial Intelligence Techniques for Cost Estimation in Construction Projects," *International Journal of Construction Management*, 2023.
- [12] Y. Wang, "Price Prediction of Second-Hand Housing in Chengdu Based on CatBoost Algorithm," in *2022 4th International Conference on Artificial Intelligence and Advanced Manufacture*, 2022.
- [13] J. Hancock and T. M. Khoshgoftaar, "Performance of CatBoost and XGBoost in Medicare Fraud Detection," in *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2020.
- [14] A. R. Putri, N. A. Santoso, and B. A. Santoso, "Implementasi Algoritma Regresi Linier dan ARIMA untuk Prediksi Harga Emas," *Jurnal Sistem Komputer dan Kecerdasan Buatan*, vol. 6, no. 2, 2023.
- [15] V. K. Gupta and A. Kumar, "Comparative Analysis of Machine Learning Models for Sales Prediction," *International Journal of Computer Applications*, vol. 175, no. 2, 2020.