

Jakarta Smart City: Pengembangan Model Smart Mobility Prediction Menggunakan GHMM-ARIMA

Vina Ayumi

Fakultas Ilmu Komputer, Universitas Mercu Buana, Jakarta, Indonesia

vina.ayumi@mercubuana.ac.id

Article Info

Article history:

Received, xxx xx xxxx

Revised, xxx xx xxxx

Accepted, xxx xx xxx

Kata Kunci:

GHMM

ARIMA

RMSE

Jakarta Smart City

ABSTRAK

Pemerintah Jakarta telah menggunakan infrastruktur digital, seperti *platform online* dan aplikasi perangkat lunak, untuk mengimplementasikan elemen-elemen ini. Namun, masih ada ruang untuk perbaikan dalam memaksimalkan manfaat bagi kota dan penduduknya. Salah satu area yang perlu dioptimalkan adalah pengembangan model *smart mobility prediction* dan meningkatkan kinerja dari model yang ada sebelumnya. Pada penelitian ini, algoritma *gaussian hidden markov model* (GHMM) dan *autoregressive integrated moving average* (ARIMA) digunakan untuk pemantauan mobilitas dapat diprediksi untuk menguraikan kemacetan di Jakarta. Parameter yang digunakan pada pendeteksian *stay points* adalah *time threshold* 20 menit, dan *distance threshold* 200 meter. Hasil evaluasi menunjukkan pengujian model ARIMA mendapatkan nilai *root mean square error* (RMSE) sebesar 162.766, menunjukkan *error* yang cukup tinggi. Hasil evaluasi untuk prediksi menggunakan GHMM memprediksi mobilitas untuk mendukung program Jakarta Smart City pada data pengujian dihitung menggunakan model akurasi dan model RMSE. Kinerja dari GHMM mendapatkan akurasi sebesar 76,90% dan RMSE sebesar 1.641. Nilai evaluasi dari GHMM dapat dikatakan cukup baik untuk memodelkan data mobilitas.

ABSTRACT

The Jakarta government has used digital infrastructure, such as online platforms and software applications, to implement these elements. However, there is still room for improvement in maximizing benefits for the city and its residents. One area that needs to be optimized is the development of smart mobility prediction models and improving the performance of existing models. In this study, Gaussian hidden markov model (GMM) and autoregressive integrated moving average (ARIMA) algorithms were used for predictable mobility monitoring to decipher congestion in Jakarta. The parameters used in detecting stay points are a time threshold of 20 minutes, and a distance threshold of 200 meters. The evaluation results showed that the ARIMA model test obtained a root mean square error (RMSE) value of 162,766, showing a fairly high error. The evaluation results for prediction using GHMM predicting mobility to support the Jakarta Smart City program on the test data were calculated using the accuracy model and RMSE model. The performance of GHMM gets an accuracy of 76.90% and RMSE of 1,641. The evaluation value of GHMM can be said to be good enough to model mobility data.

Keywords:

GHMM

ARIMA

RMSE

Jakarta Smart City

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



Penulis Korespondensi:

Vina Ayumi,

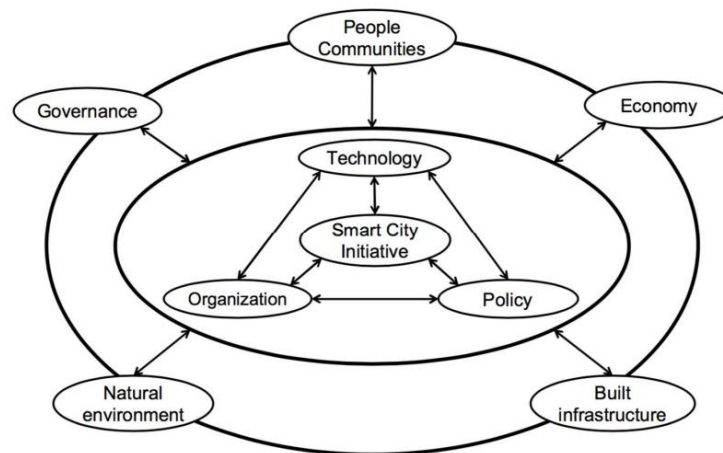
Fakultas Ilmu Komputer

Universitas Mercu Buana, Jakarta, Indonesia

Email: vina.ayumi@mercubuana.ac.id

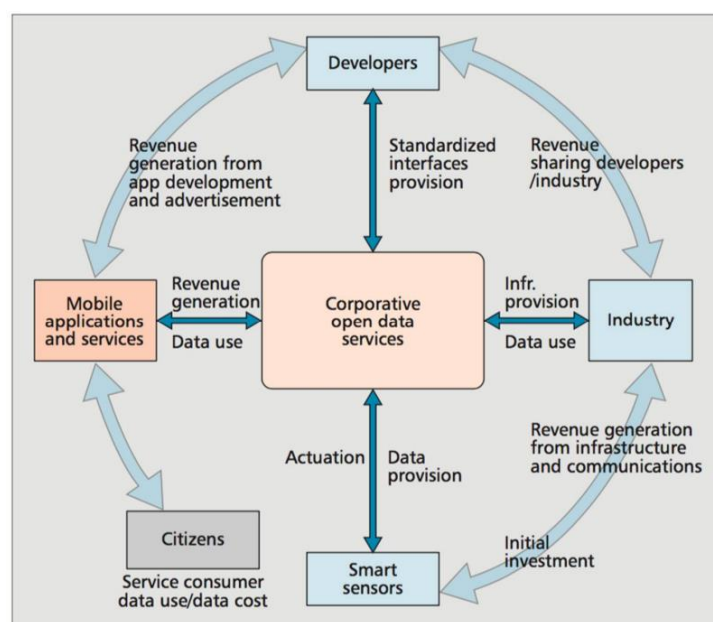
1. PENDAHULUAN

Jakarta telah menerapkan konsep *smart city*, dengan fokus pada enam elemen: *Smart Economy*, *Smart Governance*, *Smart People*, *Smart Mobility*, *Smart Environment*, dan *Smart Living*. Pemerintah Jakarta telah menggunakan infrastruktur digital, seperti platform online dan aplikasi perangkat lunak, untuk mengimplementasikan elemen-elemen ini. Namun, masih ada ruang untuk perbaikan dalam memaksimalkan manfaat bagi kota dan penduduknya. Salah satu area yang perlu dioptimalkan adalah pengembangan model *smart mobility prediction* dan meningkatkan kinerja dari model yang ada sebelumnya [1]–[5]. Setiap elemen *smart city* akan saling mendukung sehingga perlu adanya peningkatan kinerja di setiap elemen, seperti yang terlihat pada **Gambar 1**.



Gambar 2 Elemen smart city framework [6]

Komponen *smart mobility* adalah aspek penting dari Jakarta Smart City. Komponen *smart mobility* ini bertujuan untuk meningkatkan infrastruktur transportasi dan memfasilitasi akses publik ke transportasi umum, yang pada akhirnya mengarah pada manajemen lalu lintas yang lebih baik dan peningkatan minat masyarakat untuk menggunakan transportasi umum. Implementasi mobilitas cerdas di Jakarta mencakup penggunaan teknologi, informasi, dan komunikasi untuk mengelola dan mengatasi masalah transportasi. Program *Jakarta Smart City* juga berfokus pada peningkatan kualitas hidup warga kota dengan meningkatkan mobilitas dan memperkenalkan sistem pemantauan lalu lintas cerdas. Implementasi *smart mobility* di Jakarta perlu didukung inovasi yang terbaru agar pemantauan mobilitas dapat diprediksi untuk menguraikan kemacetan di Jakarta [7]–[13]. Pengembangan pada *smart mobility* akan melalui beberapa proses seperti yang terlihat pada **Gambar 2**.



Gambar 3 Proses pengembangan layanan teknologi smart city [14]

Prediksi mobilitas dalam *Jakarta Smart City* dilakukan untuk mengetahui pola masyarakat Jakarta yang berpindah di daerah perkotaan, contohnya seperti berpergian baik dengan berjalan atau menggunakan kendaraan pribadi atau transportasi publik ke suatu tempat atau dengan cara lainnya Riset untuk memahami pola mobilitas manusia sangat penting untuk mendukung *Jakarta Smart City*. Data *human mobility* bersifat time-series dan sequential. Beberapa riset mengenai prediksi mobilitas manusia telah berhasil dilakukan diantaranya menggunakan algoritma machine learning berbasis time-series dan sequential. Salah satu algoritma time-series yang populer adalah ARIMA, dan algoritma *sequential* adalah *Gaussian Hidden Markov Model* (GMM) [15], [16], [25], [17]–[24].

ARIMA adalah algoritma *time-series predictor* yang merupakan singkatan dari *autoregressive integrated moving-average*. ARIMA dapat digunakan untuk mengolah dataset dalam mencari pola prediksi. Markov order- k $O(k)$ dapat dimanfaatkan untuk mengetahui pola prediksi lokasi selanjutnya yang dikunjungi berdasarkan urutan riwayat lokasi yang sudah pernah dikunjungi. Meskipun telah banyak digunakan dalam beberapa kasus, namun algoritma *gaussian hidden markov model* (GHMM) atau algoritma ARIMA belum banyak diimplementasikan dalam mengolah data mobilitas manusia. Berdasarkan latar belakang diatas, penelitian ini mencoba untuk menggunakan algoritma ARIMA dan GHMM untuk memprediksi mobilitas untuk mendukung program *Jakarta Smart City* dalam menguraikan kemacetan di Jakarta.

2. METODE PENELITIAN

Pada penelitian ini, algoritma *gaussian hidden markov model* (GMM) dan *autoregressive integrated moving average* (ARIMA) digunakan untuk mengetahui pola masyarakat Jakarta yang berpindah di daerah perkotaan, contohnya seperti berpergian baik dengan berjalan atau menggunakan kendaraan pribadi atau transportasi publik ke suatu tempat. Tahap pertama adalah menggabungkan seluruh data trajectory individual user. Terdapat 182 user dengan total trajectory 17,621 masing-masing trajectory dengan total distance 1,292,951kilometers total durasi 50,176 jam. Struktur *data trajectory* yang digunakan dapat dilihat pada **Gambar 1**.

	lat	lon	date	time1	time2
0	39.975476	116.330303	2007-04-12	09:31:32	11:13:43
1	39.975608	116.330298	2007-04-12	10:21:16	14:42:42
2	39.975596	116.329675	2007-04-12	13:46:21	14:25:09
3	39.975313	116.330825	2007-04-13	01:04:56	05:19:33
4	39.977055	116.330136	2007-04-13	02:06:44	05:20:57

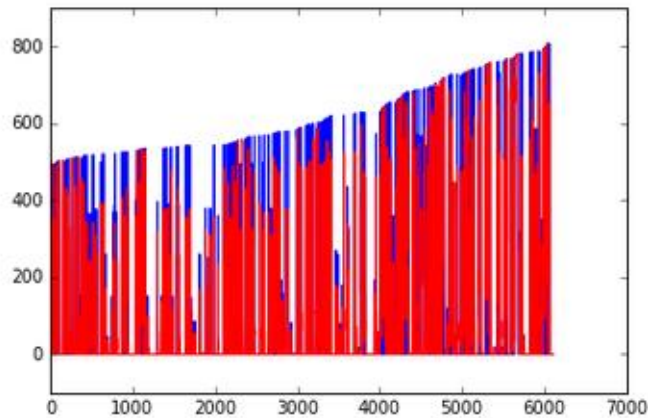
Gambar 4 Struktur *Data Trajectory*

Tahap kedua adalah pembersihan data dengan mendeteksi *stay points* dan menghilangkan data yang bukan *stay points*. Parameter yang digunakan pada pendeteksian *stay points* adalah *time threshold* 20 menit, dan *distance threshold* 200 meter. Tahap ketiga adalah ekstraksi region of interest dengan melakukan clustering pada *datastay points*.

Selanjutnya, *clustering* dilakukan dengan metode DBScan untuk mengetahui jumlah cluster dari seluruh *stay points*. Suatu point dianggap masuk kedalam satu cluster apabila jaraknya kurang dari 2 km dari centroid cluster. Tahap selanjutnya adalah melakukan pembangunan model pada data menggunakan metode *gaussian hidden markov model* (GMM) dan *autoregressive integrated moving average* (ARIMA). Pembangunan model dilakukan dengan melakukan pembelajaran pada data pelatihan dengan input waktu dan output label lokasi.

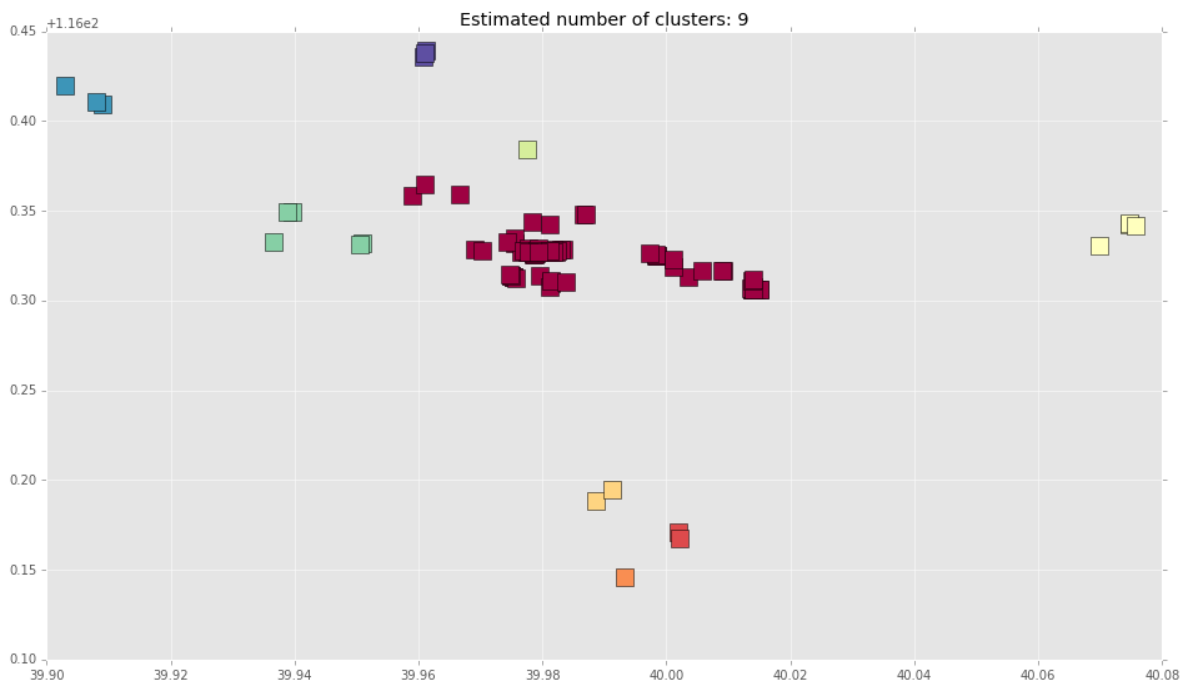
3. HASIL DAN PEMBAHASAN

Pada eksperimen pertama dilakukan pembangunan model ARIMA memprediksi mobilitas untuk mendukung program *Jakarta Smart City* dalam menguraikan kemacetan di Jakarta. Pembangunan model dilakukan dengan melakukan pembelajaran pada data pelatihan dengan input waktu dan output label lokasi. Setelah model diperoleh selanjutnya dilakukan prediksi pada data uji. Hasil evaluasi menunjukkan pengujian model ARIMA mendapatkan nilai *root mean square error* (RMSE) sebesar 162.766, menunjukkan *error* yang cukup tinggi. Hal ini mengindikasikan bahwa pemodelan yang dilakukan dengan seluruh data user kurang signifikan dari pada jika dilakukan pemodelan per individu. Grafik nilai RMSE pada model ARIMA dapat dilihat pada **Gambar 1**.



Gambar 5 Grafik RMSE dari model ARIMA

Pada eksperimen kedua, klusterisasi digunakan sebagai label kelas pada pelatihan model. Pada tahap pelatihan model selanjutnya mengimplementasikan algoritma *gaussian hidden markov model* (GHMM). Implementasi ini menggunakan *library hmmlearn* yang ada pada *sklearn*. Pada model GHMM, parameter yang digunakan adalah *n_components* dengan nilai 5, *covariance_type* dengan nilai 'spherical', dan *decoder_algorithm* dengan nilai 'viterbi'. Hasil klusterisasi menggunakan *gaussian hidden markov model* (GHMM) dapat dilihat pada **Gambar 1**.



Gambar 6 Klusterisasi menggunakan GHMM

Hasil evaluasi untuk prediksi menggunakan GHMM memprediksi mobilitas untuk mendukung program Jakarta Smart City pada data pengujian dihitung menggunakan model akurasi dan model RMSE. Kinerja dari GHMM mendapatkan akurasi sebesar 76,90% dan RMSE sebesar 1.641. Nilai evaluasi dari GHMM dapat dikatakan cukup baik untuk memodelkan data mobilitas.

4. KESIMPULAN

Penelitian ini menganalisis data mobilitas dari 182 orang dengan total *trajectory* 17,621 masing-masing *trajectory* dengan total *distance* 1,292,951kilometers total durasi 50,176 jam. Pembangunan model dilakukan dengan melakukan pembelajaran pada data pelatihan dengan input waktu dan output label lokasi. Setelah model diperoleh selanjutnya dilakukan prediksi pada data uji. Hasil evaluasi menunjukkan pengujian model ARIMA mendapatkan nilai *root mean square error* (RMSE) sebesar 162.766, menunjukkan *error* yang cukup tinggi. Hasil evaluasi untuk prediksi menggunakan GHMM memprediksi mobilitas untuk mendukung program Jakarta Smart City pada data pengujian dihitung menggunakan model akurasi dan model RMSE. Kinerja dari GHMM mendapatkan akurasi sebesar 76,90% dan RMSE sebesar 1.641. Nilai evaluasi dari GHMM dapat dikatakan cukup baik untuk memodelkan data mobilitas.

UCAPAN TERIMA KASIH

Terima kasih kepada Pusat Penelitian dan Fakultas Ilmu Komputer, Universitas Mercu Buana yang telah mendukung pelaksanaan penelitian ini.

REFERENSI

- [1] P. Cáceres, C. E. Cuesta, B. Vela, J. M. Cavero, and A. Sierra, "Smart data at play: improving accessibility in the urban transport system," *Behav. Inf. Technol.*, vol. 39, no. 6, pp. 681–694, 2020.
- [2] A. De Paola, A. Giammanco, G. Lo Re, and M. Morana, "Human mobility simulator for smart applications," pp. 203–210, 2019.
- [3] M. Sadikin, R. Y. Harwikarya, and L. Mitaliska, "Combining ITAF and ISO 27004 to Perform IS Audit in Higher Education Institution," in *International Conference on Recent Innovations in Computer Science and Information Technology (ICRICSIT-2015)*, New York, USA, 2015, pp. 5–6.
- [4] D. Fitrihanah, R. A. M. Zen, and N. H. Praptono, "Prediction Framework for Potential Tuna Fish Zone in Indonesia," in *Proceedings of the 2019 Asia Pacific Information Technology Conference*, 2019, pp. 51–55.
- [5] I. Nurhaida *et al.*, "Implementation of Deep Learning Predictor (LSTM) Algorithm for Human Mobility Prediction," *Int. J. Interact. Mob. Technol.*, vol. 14, no. 18, p. 132, Nov. 2020.
- [6] H. Chourabi *et al.*, "Understanding smart cities: An integrative framework," in *2012 45th Hawaii international conference on system sciences*, 2012, pp. 2289–2297.
- [7] B. N. Avianto, "Implementation of 'Jakarta Smart City' In One-Stop-Service: Daily Need In The City South Jakarta?," *J. Res. Soc. Sci. Econ. Manag.*, vol. 2, no. 09, 2023.
- [8] E. W. Handoyo, "Implementasi Kebijakan Jakarta Smart City dalam Mewujudkan Smart Governance pada Pemerintahan Provinsi DKI Jakarta," *Unnes Polit. Sci. J.*, vol. 6, no. 2, pp. 42–47, 2023.
- [9] R. Oktorini and L. S. Barus, "Integration of Public Transportation in Smart Transportation System (Smart Transportation System) in Jakarta," *Konfrontasi J. Kult. Ekon. dan Perubahan Sos.*, vol. 9, no. 2, pp. 341–347, 2022.
- [10] D. I. Sensuse, P. Prima, E. Cahyaningsih, and H. Noprisson, "Knowledge management practices in e-Government," in *2017 3rd International Conference on Science in Information Technology (ICSITech)*, 2017.
- [11] H. Noprisson, "Analysis and Design e-Government Website for Special Allocation Fund," *Int. J. Comput. Sci. Eng.*, vol. 8, no. 02, 2019.
- [12] M. Mishbah, D. I. Sensuse, and H. Noprisson, "Information system implementation in smart cities based on types, region, sub-area," *2017 Int. Conf. Inf. Technol. Syst. Innov. ICITSI 2017 - Proc.*, vol. 2018-Janua, pp. 155–161, 2017.
- [13] V. Ayumi, "Smart Village and Its Development in Research," *Int. J. Adv. Stud. Comput. Sci. Eng.*, vol. 9, no. 10, pp. 1–5, 2020.
- [14] I. Vilajosana, J. Llosa, B. Martinez, M. Domingo-Prieto, A. Angles, and X. Vilajosana, "Bootstrapping smart cities through a self-sustainable model based on big data flows," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 128–134, 2013.
- [15] K. Zhao, S. Tarkoma, S. Liu, and H. Vo, "Urban human mobility data mining: An overview," *2016 IEEE Int. Conf. Big Data (Big Data)*, pp. 1911–1920, 2016.
- [16] Y. Zhang *et al.*, "A human mobility dataset collected via LBSLab," *Data Br.*, vol. 46, p. 108898, 2023.
- [17] H. F. Chan, A. Skali, and B. Torgler, "A global dataset of human mobility," *Res. Pap. Econ.*, 2020.
- [18] X. Guo, R. Zhang, X. Liu, and J. Huai, "Human mobility semantics analysis: a probabilistic and scalable approach," *Geoinformatica*, vol. 22, no. 3, pp. 507–539, 2018.
- [19] Y. Gao, J. Cheng, H. Meng, and Y. Liu, "Measuring spatio-temporal autocorrelation in time series data

- of collective human mobility,” *Geo-spatial Inf. Sci.*, vol. 22, no. 3, pp. 166–173, 2019.
- [20] Y. K. Qarout, Y. P. Raykov, and M. A. Little, “Probabilistic Modelling for Unsupervised Analysis of Human Behaviour in Smart Cities,” *Sensors*, vol. 20, no. 3, p. 784, 2020.
- [21] R. Singh and Y. Chen, “Learning Gaussian Hidden Markov Models From Aggregate Data,” *IEEE Control Syst. Lett.*, vol. 7, pp. 478–483, 2023.
- [22] N. Bátfai, R. Besenczi, P. Jeszenszky, M. Szabó, and M. Ispány, “Markov modeling of traffic flow in Smart Cities,” *Appl. Med. Informatics*, vol. 53, pp. 21–44, 2021.
- [23] “Applications, Modern Trends, and Challenges of Multiscale Modeling in Smart Cities,” *Adv. data Min. database Manag. B. Ser.*, pp. 1–23, 2023.
- [24] F. Zhang, S. Han, H. Gao, and T. Wang, “A Gaussian mixture based hidden Markov model for motion recognition with 3D vision device,” *Comput. Electr. Eng.*, vol. 83, p. 106603, 2020.
- [25] Y. Jung and J. Park, “Scalable Inference for Hybrid Bayesian Hidden Markov Model Using Gaussian Process Emission,” *J. Comput. Graph. Stat.*, vol. 31, no. 3, pp. 666–683, 2022.