



## PENERAPAN DEEP LEARNING UNTUK ANALISIS BIG DATA KEPENDUDUKAN: MENUJU EFISIENSI DAN AKURASI PELAYANAN PEMERINTAH

Muh Rivandy Setiawan

Universitas Muhammadiyah Bandung, Bandung, Indonesia  
*Email:* muhrivandysetiawan@gmail.com

Received: (22-09-2025), Accepted: (29-12-2025), Published Online: (31-12-2025)

### ABSTRAK

Perkembangan big data kependudukan menuntut adanya pendekatan analitik yang lebih canggih untuk menjawab tantangan konsistensi, akurasi, dan ketepatan kebijakan publik. Penelitian ini bertujuan untuk menerapkan metode deep learning, khususnya *Autoencoder* dan *Artificial Neural Network (ANN)*, dalam menganalisis data kependudukan Kota Bandung periode 2017–2024. *Autoencoder* digunakan untuk mendeteksi anomali dan inkonsistensi antara data Badan Pusat Statistik (BPS) dan Dinas Kependudukan dan Pencatatan Sipil (Disdukcapil), sedangkan ANN difokuskan pada prediksi kepemilikan e-KTP berbasis variabel demografi dan spasial. Hasil penelitian menunjukkan bahwa tren pertumbuhan penduduk Kota Bandung relatif stabil dengan laju yang lambat, menandakan pentingnya perencanaan berbasis data. Tingkat kepemilikan e-KTP menunjukkan disparitas antar kecamatan, dengan beberapa wilayah sudah mencapai lebih dari 96% sementara lainnya masih di bawah 90%. *Autoencoder* berhasil mendeteksi anomali dengan reconstruction loss sekitar 0.06, yang mengindikasikan peningkatan kualitas data setelah proses rekonstruksi. Sementara itu, ANN menghasilkan prediksi yang mendekati data aktual dengan rata-rata *error* yang kecil, membuktikan efektivitasnya dalam mengidentifikasi kecamatan yang berisiko rendah maupun tinggi dalam capaian kepemilikan e-KTP. Secara teoritis, penelitian ini memperkaya literatur tentang penerapan *deep learning* dalam administrasi publik berbasis data. Secara praktis, temuan ini memberikan rekomendasi aplikatif bagi pemerintah daerah, khususnya dalam memperkuat *e-government*, meningkatkan akurasi pelayanan publik, serta merancang intervensi kebijakan yang lebih tepat sasaran. Penerapan *deep learning* terbukti mampu mendukung tata kelola kependudukan yang lebih efisien, akurat, dan responsif.

**Kata kunci:** Deep Learning, Autoencoder, Artificial Neural Network, Big Data Kependudukan, e-KTP, Kota Bandung

### ABSTRACT

The advancement of population big data necessitates a more sophisticated analytical approach to address the challenges of consistency, accuracy, and the precision of public policy. This research aims to apply deep learning methods, specifically *Autoencoder* and *Artificial Neural Network (ANN)*, to analyze population data for Bandung City from 2017–2024. The *Autoencoder* was used to detect anomalies and inconsistencies between data from the Central Bureau of Statistics (BPS) and the

*Department of Population and Civil Registration (Disdukcapil), while the ANN focused on predicting e-ID card ownership based on demographic and spatial variables. The research findings show that Bandung City's population growth trend is relatively stable with a slowing rate, indicating the importance of data-driven planning. e-ID card ownership rates show disparities between sub-districts, with some areas having reached over 96% while others are still below 90%. The Autoencoder successfully detected anomalies with a reconstruction loss of approximately 0.06, which indicates an improvement in data quality after the reconstruction process. Meanwhile, the ANN produced predictions that closely approximated the actual data with a small average error, proving its effectiveness in identifying sub-districts at low and high risk for e-ID card ownership attainment. Theoretically, this research enriches the literature on the application of deep learning in data-driven public administration. Practically, the findings provide applicable recommendations for local government, particularly in strengthening e-government, improving the accuracy of public services, and designing more precisely targeted policy interventions. The application of deep learning has been proven capable of supporting population governance that is more efficient, accurate, and responsive.*

**Keywords:** Deep Learning, Autoencoder, Artificial Neural Network, Population Big Data, e-KTP, Bandung City

## PENDAHULUAN

Transformasi digital telah menjadi pendorong utama perubahan dalam administrasi publik, memicu kebutuhan data yang lebih akurat, terkini, dan konsisten. Di Indonesia, kebijakan kependudukan melalui perekaman e-KTP dan administrasi kependudukan digital telah dianggap sebagai langkah penting menuju pemerintahan yang lebih efektif dan transparan. Namun kenyataannya, masih ada celah signifikan antara target kepemilikan identitas digital dan realitas di lapangan. Kekurangan dalam sinkronisasi data antar lembaga, ketidakteraturan dalam memperbarui data, serta munculnya anomali-anomali administratif menjadi hambatan yang serius dalam pemanfaatan data tersebut untuk kebijakan publik. Penelitian oleh Didin, Haedar, & Aliah (2023) menyebut bahwa meskipun teknologi AI telah mulai diadopsi, kualitas data yang rendah — seperti data duplikat, input yang tidak konsisten, dan kecepatan pembaruan — menghambat efektivitasnya.

Kota Bandung, sebagai pusat pendidikan, budaya, dan ekonomi kreatif di Jawa Barat, memiliki dinamika demografi yang cukup kompleks. Mobilitas penduduk dari luar daerah, mahasiswa yang datang dan pergi, serta pekerja kontrak menambah tingkat fluktuasi pada data kependudukan. Misalnya, dalam periode yang tidak lama setelah pandemi COVID-19, banyak perpindahan penduduk yang tidak tercatat atau terlambat pencatatannya. Beberapa laporan investigatif menunjukkan bahwa kepemilikan e-KTP di beberapa daerah belum merata, dan tidak semua penduduk wajib KTP telah terekam dengan benar di Disdukcapil, yang menyebabkan gap antara jumlah wajib KTP dan kepemilikan e-KTP. Sebuah studi di Manado menunjukkan bahwa meskipun perekaman telah dilakukan, pelayanan di Kecamatan Tumiting belum optimal karena prosedur dan akses layanan yang masih terbatas (Alhasni, Onibala, & Rantung, 2025). ini menunjukkan bahwa meskipun regulasi sudah jelas, praktek di lapangan bisa jauh dari ideal.

Data dari BPS menunjukkan bahwa persentase kepemilikan e-KTP dari 2019 sampai 2023 berkisar sekitar 75-76% untuk beberapa kota/kabupaten (FK Putri et al.,

2024) — jauh dibawah target 100%. Hal ini mempertegas bahwa meskipun legalitas dan kebijakan telah mendukung, implementasi, distribusi, dan kualitas data masih sering mengalami hambatan. Situasi ini memperkuat risiko bahwa tanpa deteksi anomali dan pembersihan data, kebijakan publik bisa salah sasaran, misalnya bantuan sosial yang dibagikan tidak merata atau fasilitas yang direncanakan berdasarkan data yang usang.

Salah satu cara untuk mengatasi masalah ini adalah dengan penerapan metode deep learning, khususnya Autoencoder untuk mendeteksi anomali dan *ANN (Artificial Neural Network)* untuk prediksi gap antara wajib KTP dan kepemilikan e-KTP. Autoencoder dapat membantu mengidentifikasi entri data yang tidak wajar, misalnya data duplikat, jumlah populasi yang melonjak di satu kecamatan tanpa migrasi resmi, atau input yang tidak konsisten antar sumber. ANN dapat digunakan untuk memprediksi kepemilikan e-KTP berdasarkan variabel demografi seperti umur, pendidikan, jenis kelamin, pekerjaan, dan lokasi geografis (kecamatan/kelurahan), sehingga pemerintah bisa menentukan kecamatan mana yang butuh intervensi tambahan.

Penelitian di berbagai kota menunjukkan bahwa inovasi pelayanan publik memegang peran penting dalam mengurangi kesenjangan kepemilikan dokumen kependudukan. Studi mengenai "Inovasi Layanan Teman Dilan" di Mojokerto, misalnya, telah memfasilitasi perekaman e-KTP secara langsung di rumah bagi masyarakat rentan seperti lansia dan penyandang disabilitas (Anindya et al., 2024). Hasil studi tersebut membuktikan bahwa kemudahan akses fisik bagi masyarakat sangat membantu meningkatkan cakupan kepemilikan dokumen identitas. Fenomena ini menegaskan bahwa solusi terhadap permasalahan kependudukan tidak hanya bersifat teknis seperti pembersihan data dan prediksi, melainkan juga harus menyentuh aspek administratif serta operasional.

Berdasarkan tinjauan literatur dan kondisi di lapangan, dapat diidentifikasi gap penelitian yang penting: sedikit studi yang menggunakan teknik *deep learning* (*Autoencoder* dan *ANN*) untuk menggabungkan deteksi anomali dan prediksi gap kepemilikan e-KTP di level kota besar di Indonesia; juga sedikit yang membahas secara spesifik dampak migrasi internal dan keberhasilan perekaman di kecamatan/kelurahan sebagai unit analisis. Penelitian ini hadir untuk mengisi gap tersebut dengan menggunakan data kependudukan Kota Bandung 2017–2024 sebagai kasus, fokus pada inkonsistensi antar lembaga, anomali data, fluktuasi migrasi, dan prediksi gap e-KTP.

Penelitian ini bertujuan untuk mengidentifikasi tingkat inkonsistensi data kependudukan Kota Bandung dari tahun ke tahun, mendeteksi anomali di dalam dataset, serta memahami pengaruh data terhadap migrasi yang memengaruhi fluktuasi populasi dan kepemilikan e-KTP. Selain itu, penelitian ini juga dilakukan untuk memprediksi celah (*gap*) kepemilikan e-KTP di tingkat kecamatan atau kelurahan agar intervensi pemerintah menjadi lebih tepat sasaran. Melalui pendekatan tersebut, hasil penelitian ini diharapkan tidak hanya memberikan kontribusi secara akademis, tetapi juga menghasilkan rekomendasi kebijakan operasional yang aplikatif bagi Disdukcapil Kota Bandung serta para pemangku kebijakan terkait.

## METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan penerapan metode deep learning, yaitu kombinasi *Autoencoder* dan *Artificial Neural Network (ANN)*. Fokus penelitian adalah mendeteksi inkonsistensi dan anomali data kependudukan serta memprediksi celah (*gap*) kepemilikan e-KTP di Kota Bandung selama periode 2017–2024. *Autoencoder* digunakan karena efektif dalam mendeteksi pola data yang menyimpang pada dataset besar tanpa label, sedangkan ANN digunakan untuk memprediksi kepemilikan e-KTP berdasarkan variabel demografis dan spasial. Dengan pendekatan ini, penelitian bersifat eksploratif sekaligus prediktif untuk mendukung tata kelola kependudukan berbasis data (Priatna et al., 2024).

Tahap awal penelitian adalah pengumpulan data sekunder dari Badan Pusat Statistik (BPS) Kota Bandung dan Dinas Kependudukan dan Pencatatan Sipil (Dukcapil) Kota Bandung. Data yang dikumpulkan mencakup jumlah penduduk total, jumlah wajib KTP, jumlah pemilik e-KTP, serta karakteristik demografi penduduk per kecamatan atau kelurahan. Unit analisis utama adalah wilayah administratif (kecamatan dan kelurahan) per tahun. Penggunaan dua sumber data bertujuan untuk memungkinkan analisis perbedaan dan inkonsistensi antar lembaga, yang masih sering ditemukan dalam pengelolaan data kependudukan di tingkat daerah (Didin, Haedar, & Aliah, 2024).

Prosedur *preprocessing* disusun secara sistematis karena tahap ini sangat menentukan kualitas hasil deep learning. Preprocessing meliputi pembersihan data ganda, penanganan missing values, transformasi data kategori ke numerik, serta normalisasi skala data. Proses ini dilakukan menggunakan bahasa pemrograman *Python* dengan pustaka pandas, *numpy*, dan *scikit-learn*, sedangkan pengembangan model *deep learning* menggunakan *TensorFlow* dan Keras. Model dilatih pada perangkat berbasis CPU dan dapat memanfaatkan GPU untuk mempercepat proses pelatihan. *Autoencoder* dilatih dengan *batch size* 32 dan *learning rate* 0.001 selama maksimal 200 *epoch*, sedangkan ANN dilatih dengan *batch size* 16 dan *learning rate* 0.001 hingga maksimal 400 *epoch*. Untuk menjaga stabilitas model dan mencegah *overfitting*, digunakan mekanisme *early stopping* dan *learning rate reduction*. Konfigurasi ini dinilai memadai untuk dataset kependudukan berskala menengah yang digunakan dalam penelitian ini.

Tahap kedua adalah deteksi anomali data menggunakan *Autoencoder*. Model dilatih untuk mempelajari pola umum data kependudukan dan merekonstruksi kembali data input. Selisih antara data asli dan hasil rekonstruksi dihitung sebagai *reconstruction error*. Data dengan nilai *error* yang melebihi ambang batas tertentu dikategorikan sebagai anomali, seperti jumlah wajib KTP yang melampaui total penduduk atau perbedaan signifikan antara data BPS dan Dukcapil. Indikator utama pada tahap ini meliputi jumlah anomali yang terdeteksi, distribusi *reconstruction error*, dan proporsi data yang tidak konsisten.

Arsitektur *Autoencoder* yang digunakan berupa jaringan *feedforward* dengan beberapa lapisan *dense*, terdiri atas lapisan *encoder* yang mereduksi dimensi data, *bottleneck* layer sebagai representasi *latent*, serta *decoder* yang merekonstruksi data ke bentuk semula. Fungsi aktivasi ReLU digunakan pada lapisan tersembunyi, sedangkan

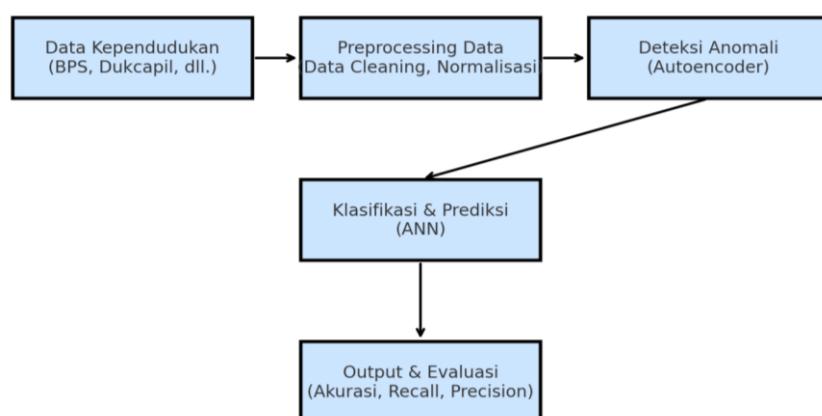
lapisan keluaran menggunakan aktivasi linear. Proses pelatihan dioptimalkan menggunakan *Adam optimizer* dan *loss function Mean Squared Error (MSE)*. Untuk meningkatkan keandalan hasil, anomali yang terdeteksi diverifikasi kembali menggunakan aturan logika kependudukan dan data administratif pendukung bila tersedia (Aslam, 2024; Neloy, 2024).

Tahap ketiga adalah prediksi *gap* kepemilikan e-KTP menggunakan *Artificial Neural Network (ANN) tipe Multi-Layer Perceptron (MLP)*. Variabel input meliputi indikator demografi (umur, jenis kelamin, pendidikan, pekerjaan) dan variabel spasial (kecamatan/kelurahan), sedangkan variabel *output* berupa prediksi jumlah atau rasio kepemilikan e-KTP. ANN dirancang dengan beberapa lapisan tersembunyi dan fungsi aktivasi ReLU untuk menangkap hubungan nonlinier antar variabel demografi yang kompleks (Magazzino, 2024; Grossman, 2023).

Evaluasi kinerja model dilakukan menggunakan beberapa metrik kuantitatif. Untuk Autoencoder, evaluasi difokuskan pada distribusi *reconstruction error* dan jumlah anomali yang teridentifikasi. Untuk ANN, metrik evaluasi meliputi *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)*, *Mean Absolute Error (MAE)*, dan *koefisien determinasi ( $R^2$ )*. Validasi dilakukan dengan metode *train-test split (70:30)* dan *k-fold cross-validation* untuk memastikan stabilitas dan kemampuan generalisasi model (Alfahrizi et al., 2023).

Aspek etika dan tata kelola data menjadi bagian penting dalam metode penelitian ini. Seluruh data diproses dalam bentuk agregat dan telah dianonimkan sebelum analisis. Informasi identitas pribadi seperti nama dan NIK tidak digunakan dalam proses pemodelan. Penelitian ini mematuhi Undang-Undang Nomor 27 Tahun 2022 tentang Perlindungan Data Pribadi serta pedoman etika penelitian sosial. Hasil model tidak dimaksudkan untuk menggantikan proses verifikasi administratif, melainkan sebagai alat bantu analitik dalam penentuan prioritas kebijakan pelayanan e-KTP.

Guna memperjelas alur penelitian, kerangka metode penelitian divisualisasikan seperti pada Gambar 1 berikut:



Sumber: Priatna et al., 1014

**Gambar 1. Kerangka metode penelitian deep learning untuk data kependudukan**

Berikut ini ringkasan indikator penelitian disajikan dalam Tabel 1 untuk memperjelas apa saja yang diukur adalah sebagai berikut:

**Tabel 1. Indikator Pengukuran Penelitian**

No.	Tahapan Analisis	Indikator yang Diukur
1	<i>Preprocessing</i>	Jumlah missing values, data ganda, hasil normalisasi data
2	<i>AutoEncoder</i>	Reconstruction error, jumlah anomali, proporsi data yang tidak konsisten
3	ANN (Prediksi e-KTP)	Akurasi prediksi, MSE, RMSE, error relatif terhadap data aktual

Sumber: Priatna et al., 2024

Tahap akhir penelitian adalah analisis hasil dan interpretasi kebijakan. Data hasil prediksi dibandingkan dengan kondisi aktual untuk melihat seberapa jauh teknologi deep learning mampu memberikan solusi. Dari sisi praktis, indikator yang relevan bagi pemerintah daerah adalah besarnya gap kepemilikan e-KTP dan distribusi anomali per kecamatan. Penelitian ini juga mengukur aspek teknis (kinerja model), aspek administratif (tingkat konsistensi data), serta aspek sosial (proporsi warga yang belum memiliki e-KTP). Etika penelitian tetap dijaga dengan menghapus informasi identitas pribadi, menjaga anonimitas, serta mematuhi aturan perlindungan data penduduk di Indonesia.

## HASIL DAN PEMBAHASAN

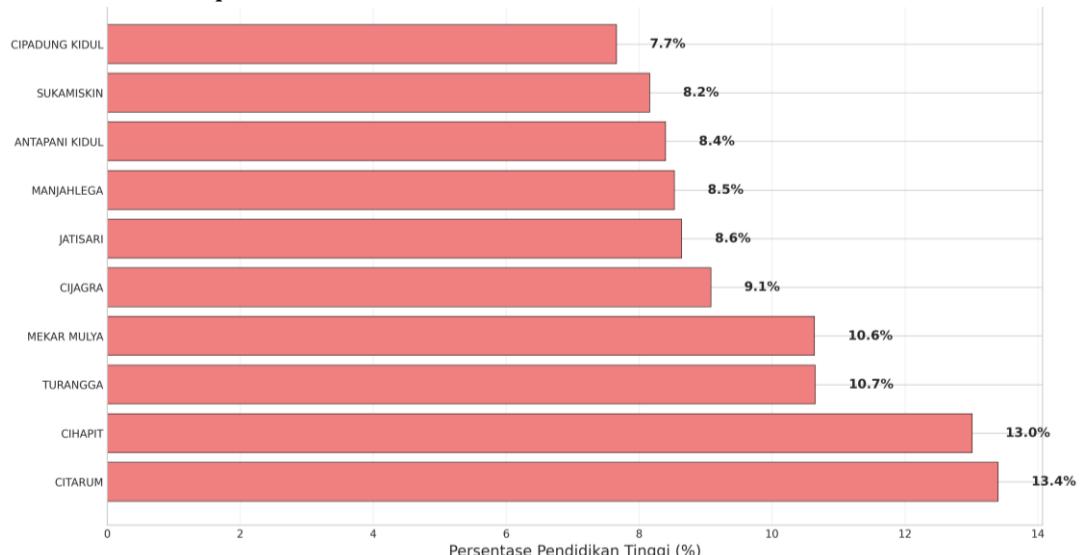
Pembahasan berikut ini menggambarkan tren demografi, tingkat pendidikan, serta ketersediaan identitas digital di Kota Bandung. Dari data yang tersedia, terlihat adanya pola pertumbuhan populasi yang stabil namun tetap menunjukkan peningkatan signifikan, disertai dengan disparitas dalam akses layanan publik seperti e-KTP dan tingkat pendidikan antar wilayah. Analisis ini memberikan gambaran komprehensif tentang kondisi sosial-ekonomi dan kebijakan pemerintah daerah dalam memastikan inklusi dan kualitas hidup masyarakat.

**Tabel 2. Tren Pertumbuhan Penduduk Kota Bandung (2017-2024)**

Tahun	Jumlah Penduduk	Growth Absolute	Growth_Percentage
2017	2.412.458	-	-
2018	2.452.179	39.721.0	1.65
2019	2.480.464	28.285.0	1.15
2020	2.500.967	20.503.0	0.83
2021	2.526.476	25.509.0	1.02
2022	2.545.005	18.529.0	0.73
2023	2.569.107	24.102.0	0.95
2024	2.591.763	22.656.0	0.88

Sumber: Disdukcapil, 2025

Tabel 2 menyajikan tren pertumbuhan jumlah penduduk Kota Bandung dari Disdukcapil periode 2017–2024 yang menunjukkan pertumbuhan positif dengan fluktuasi pada laju pertumbuhannya. Jumlah penduduk meningkat secara konsisten dari 2.412.458 jiwa pada 2017 menjadi 2.591.763 jiwa pada 2024, dengan rata-rata pertumbuhan absolut sebesar 23.000 jiwa per tahun. Laju pertumbuhan persentase mencapai titik tertinggi pada 2018 (1,65%) dan mengalami penurunan bertahap hingga mencapai 0,88% pada 2024. Penurunan laju pertumbuhan ini secara empiris dipengaruhi oleh dinamika migrasi, tingkat fertilitas, serta implementasi kebijakan urbanisasi di Kota Bandung. Proyeksi jumlah penduduk untuk tahun 2025 diperkirakan mencapai 2.613.899 jiwa, yang mengonfirmasi bahwa tren pertumbuhan tetap berlanjut meskipun dengan intensitas yang lebih rendah. Data ini menjadi instrumen krusial dalam perumusan kebijakan perencanaan infrastruktur, layanan kesehatan, dan sektor pendidikan masa depan.



**Gambar 2. Top 10 Kelurahan dengan Persentase Pendidikan Tertinggi (Diploma IV/Strata I ke atas)**

Berdasarkan Gambar 2, grafik batang horizontal tersebut menunjukkan 10 Kelurahan dengan persentase pendidikan tertinggi (kategori Diploma IV/Strata I ke atas) yang disusun secara meningkat. Kelurahan Citarum menempati posisi puncak dengan persentase tertinggi sebesar 13,4%, diikuti oleh Cihapit (13,0%) dan Turangga (10,7%) sebagai tiga besar wilayah dengan tingkat pendidikan tinggi paling dominan. Sementara itu, kelurahan lainnya seperti Mekar Mulya, Cijagra, Jatisari, Manjahlega, Antapani Kidul, dan Sukamiskin berada di kisaran 8% hingga 10%, dengan Kelurahan Cipadung Kidul berada di posisi kesepuluh pada angka 7,7%. Secara keseluruhan, data ini menggambarkan sebaran konsentrasi penduduk berpendidikan tinggi di berbagai wilayah kelurahan tersebut.

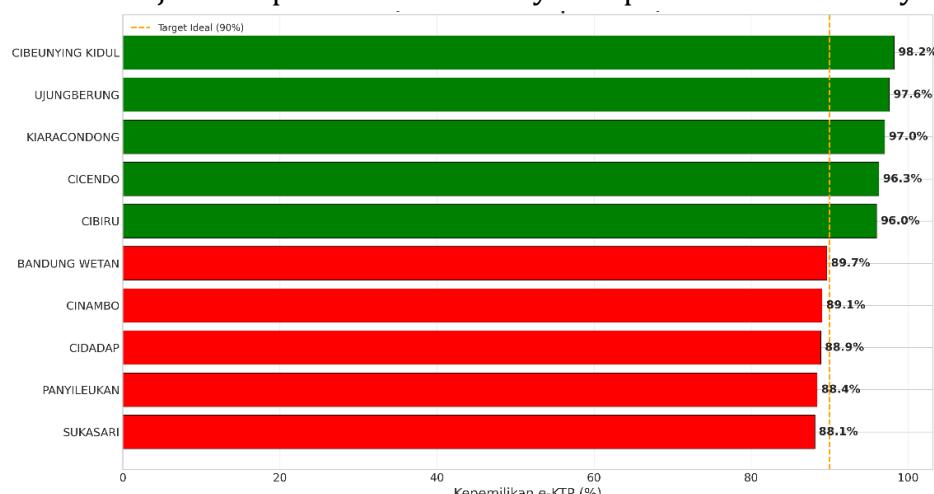
**Tabel 3. Rata-Rata Penduduk, Kepemilikan e-KTP, dan Indeks Ketimpangan Status Kawin per Kecamatan**

No.	Kecamatan	Rata-Rata Penduduk	Rata-Rata Kepemilikan E KTP	Indeks Ketimpangan Status Kawin
1.	Babakan Ciparay	141210	93.66	4.652
2.	Batununggal	121032	94.62	4.423
3.	Bojongloa Kaler	123973	95.22	4.542
4.	Bandung Kulon	135031	95.49	4.642
5.	Kiaracondong	130090	97.01	4.559

Sumber: Disdukcapil, 2025

Tabel 3 menyajikan rata-rata jumlah penduduk, tingkat *coverage* e-KTP, serta indeks ketimpangan status kawin per kecamatan di Kota Bandung. Rata-rata penduduk dihitung sebagai nilai rerata tahunan jumlah penduduk total per kecamatan selama periode pengamatan. *Coverage* e-KTP didefinisikan sebagai rasio jumlah penduduk wajib KTP (usia  $\geq 17$  tahun atau sudah/pernah kawin) yang telah memiliki e-KTP terhadap total penduduk wajib KTP, dinyatakan dalam persentase. Dengan definisi ini, nilai *coverage* tidak dihitung terhadap total penduduk, melainkan secara spesifik terhadap populasi dewasa yang secara administratif diwajibkan memiliki e-KTP, sehingga lebih relevan untuk evaluasi kinerja pelayanan administrasi kependudukan.

Indeks ketimpangan status kawin merupakan ukuran komposit yang menggambarkan tingkat ketidakseimbangan distribusi status perkawinan (belum kawin, kawin, cerai hidup, cerai mati) antar kelurahan dalam satu kecamatan. Indeks ini dihitung dari data proporsi status kawin yang telah dinormalisasi dan ditransformasikan dalam skala indeks, sehingga nilainya bersifat relatif (bukan persentase langsung) dan digunakan untuk keperluan perbandingan antar wilayah. Nilai indeks yang lebih tinggi menunjukkan distribusi status kawin yang semakin tidak merata antar kelurahan. Dengan demikian, meskipun tingkat *coverage* e-KTP relatif tinggi di seluruh kecamatan, nilai indeks ketimpangan yang masih cukup besar pada beberapa wilayah mengindikasikan adanya variasi kondisi sosial-demografis yang perlu diperhatikan dalam perencanaan kebijakan kependudukan dan layanan publik berbasis wilayah.



**Gambar 3. Top 5 Kecamatan dengan Kepemilikan e-KTP Terendah dan Tertinggi (Basis: Rata-rata per Tahun)**

Gambar 3 memperlihatkan perbandingan kepemilikan e-KTP antara kecamatan dengan nilai tertinggi dan terendah, dengan garis kuning sebagai target ideal (90%). Kecamatan Cibeunying Kidul, Ujungberung, Kiaracondong, Cicendo, dan Cibiru memiliki coverage di atas 96%, menunjukkan pencapaian yang sangat baik dalam program e-KTP. Sebaliknya, kecamatan seperti Sukasari (88.1%), Panyileukan (88.4%), dan Cidadap (88.9%) masih berada di bawah target, dengan warna merah yang menandakan potensi isu aksesibilitas. Ketidakseimbangan ini bisa disebabkan oleh faktor geografis, infrastruktur teknologi, atau kurangnya kesadaran masyarakat. Kecamatan dengan coverage rendah memerlukan intervensi khusus, seperti kampanye edukasi dan fasilitasi layanan di titik-titik strategis, agar seluruh warga dapat memanfaatkan identitas digital secara optimal.

**Tabel 4. Pertumbuhan Penduduk dan Indeks Pendidikan Tinggi di Beberapa Kelurahan**

No.	Kelurahan	Pertumbuhan Penduduk (%)	Indeks Pendidikan Tinggi (Persen)
1.	Pasirjati	106.71	113.95
2.	Pasir Biru	101.47	113.88
3.	Cisantran Kidul	84.65	236.3
4.	Merdeka	34.06	269.14
5.	Cisaranten Kulon	29.99	251.93
6.	Rancabolang	27.74	309.14
7.	Pasirwangi	25.64	103.62
8.	Mekarmulya	21.34	35.99
9.	Pasanggrahan	18.88	132.62
10.	Pasir Impun	18.59	116.73

Sumber: Disdukcapil, 2025

Tabel 4 menampilkan data pertumbuhan penduduk dan indeks pendidikan tinggi di sepuluh kelurahan terpilih. Kelurahan Pasirjati menunjukkan pertumbuhan penduduk tertinggi (106.71%), yang kemungkinan besar didorong oleh faktor urbanisasi dan pembangunan perumahan. Namun, meskipun pertumbuhan populasi tinggi, indeks pendidikan tinggi di Pasirjati juga tinggi (113.95), menunjukkan bahwa pertumbuhan populasi tidak hanya bersifat numerik, tetapi juga berkualitas. Di sisi lain, kelurahan seperti Mekarmulya dan Pasir Impun memiliki pertumbuhan yang lebih rendah tetapi indeks pendidikan tinggi yang sangat tinggi, menunjukkan bahwa populasi di wilayah ini cenderung terdidik dan mungkin merupakan komunitas profesional. Pola ini mengindikasikan bahwa pengembangan wilayah di Bandung tidak hanya ditandai oleh pertumbuhan jumlah, tetapi juga oleh kualitas kapasitas manusia.



**Gambar 4. Tren Pertumbuhan Penduduk Kota Bandung**

Gambar 4 menggambarkan tren pertumbuhan penduduk Kota Bandung dari tahun 2017 hingga prediksi tahun 2025. Data aktual menunjukkan tren linier yang stabil dengan pertumbuhan yang terus meningkat, dari 2.412.458 jiwa pada 2017 menjadi 2.591.763 jiwa pada 2024. Prediksi untuk tahun 2025 mencapai 2.613.899 jiwa, yang menunjukkan bahwa pertumbuhan akan terus berlanjut meskipun dengan laju yang lebih lambat. Kurva yang mulus menunjukkan konsistensi dalam dinamika populasi, yang dipengaruhi oleh faktor-faktor seperti migrasi internal, natalitas, dan urbanisasi. Prediksi ini penting untuk perencanaan kota, termasuk pengembangan infrastruktur transportasi, pemukiman, dan layanan publik yang harus mampu menampung pertumbuhan populasi yang terus berlangsung.

**Tabel 5. Hasil Pengukuran Model AutoEncoder**

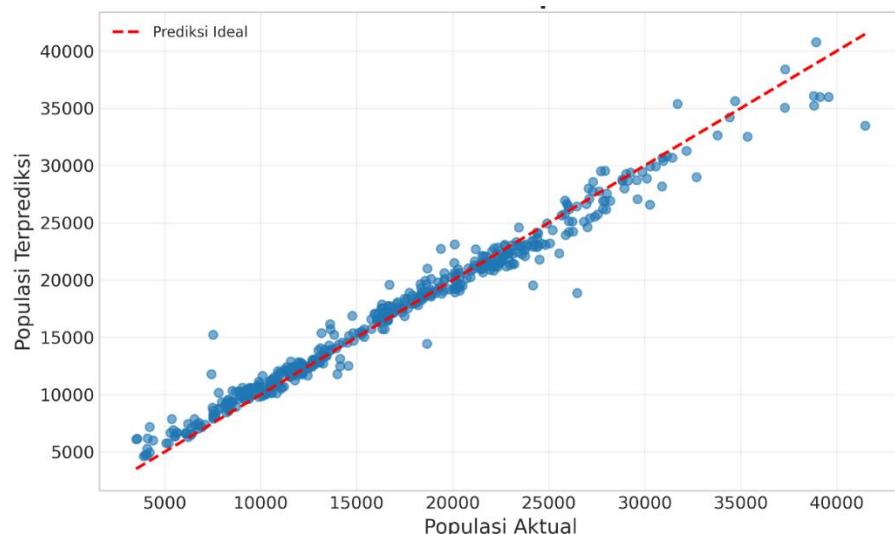
No.	Pengukuran	Hasil
1.	Train Loss	0.020825
2.	Test Loss	0.023820

Tabel 5 menunjukkan bahwa *AutoEncoder* berfungsi dengan sangat baik dalam tugas rekonstruksi data. 'Loss' di sini mengacu pada *Mean Squared Error (MSE)* antara input asli dan *output* yang direkonstruksi oleh *AutoEncoder*. Nilai yang sangat rendah (mendekati 0) mengindikasikan bahwa *AutoEncoder* mampu menangkap dan merepresentasikan fitur-fitur penting dari data input dengan presisi tinggi. Selain itu, perbedaan yang sangat kecil antara *Loss Latih* dan *Loss Uji* menunjukkan bahwa model tidak mengalami *overfitting*. Ini berarti *AutoEncoder* tidak hanya menghafal data pelatihan, tetapi juga menggeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya (data uji). Dengan demikian, fitur-fitur *terenkode* yang dihasilkan oleh *AutoEncoder* adalah representasi yang solid dan efisien dari data asli, sangat cocok untuk digunakan sebagai input bagi model ANN selanjutnya.

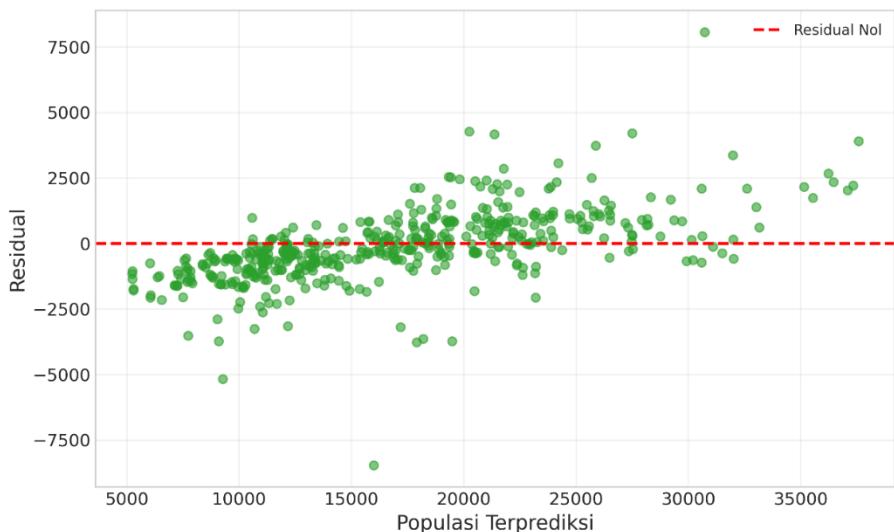
**Tabel 6. Hasil Pengukuran Model ANN**

No.	Pengukuran	Hasil
1.	R <sup>2</sup>	0.9682
2.	MAE	968
3.	MSE	1,828
4.	RMSE	1,348

Tabel 6 model menunjukkan kinerja yang sangat kuat dalam memprediksi populasi. Dengan nilai R<sup>2</sup> sebesar 0.9682, model ini mampu menjelaskan sekitar 96.82% dari variabilitas dalam data populasi, yang menandakan tingkat kecocokan yang luar biasa tinggi antara prediksi dan nilai aktual. Selain itu, nilai MAE (*Mean Absolute Error*) sebesar 968 menunjukkan bahwa rata-rata kesalahan prediksi model adalah 968 individu, sementara RMSE (*Root Mean Squared Error*) sebesar 1,348 mengindikasikan bahwa deviasi standar residu adalah 1,348 individu. Nilai-nilai *error* yang relatif rendah ini, bersama dengan R<sup>2</sup> yang sangat tinggi, menegaskan bahwa model regresi yang dikembangkan sangat akurat dan dapat diandalkan untuk memprediksi jumlah penduduk di kelurahan.

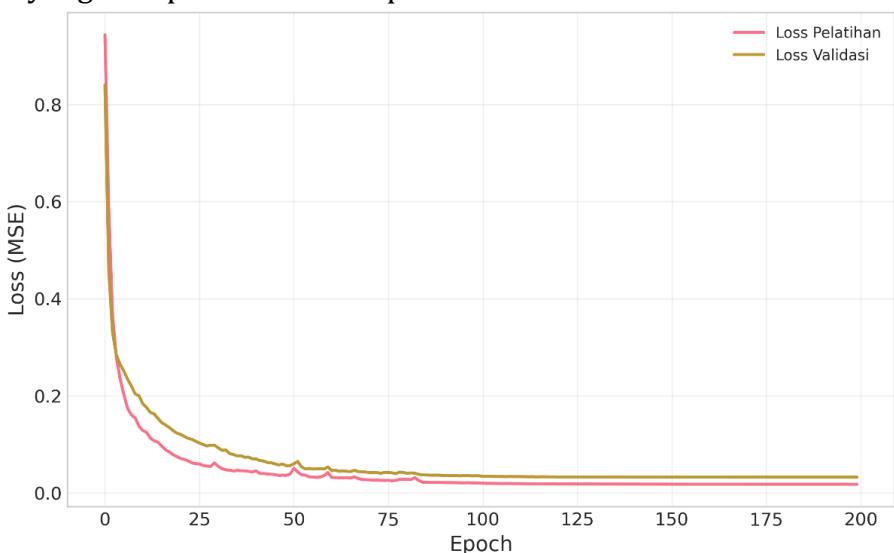
**Gambar 5. Prediksi ANN vs Sebenarnya**

Gambar 5 menampilkan scatter plot hubungan antara populasi aktual (sumbu x) dan prediksi model ANN (sumbu y), dengan garis diagonal merah putus-putus ("Prediksi Ideal",  $y = x$ ) sebagai acuan sempurna. Sebagian besar titik data biru transparan terdistribusi dekat garis tersebut, menunjukkan prediksi yang sangat akurat—terutama pada rentang populasi menengah (10.000–30.000 penduduk)—meskipun ada beberapa outlier di ujung atas yang mengindikasikan sedikit over-prediction pada kelurahan berpopulasi tinggi ( $>30.000$ ). Secara keseluruhan, model ANN berhasil memprediksi jumlah penduduk dengan baik, didukung rata-rata kesalahan hanya 968 orang, yang relatif kecil terhadap skala data.



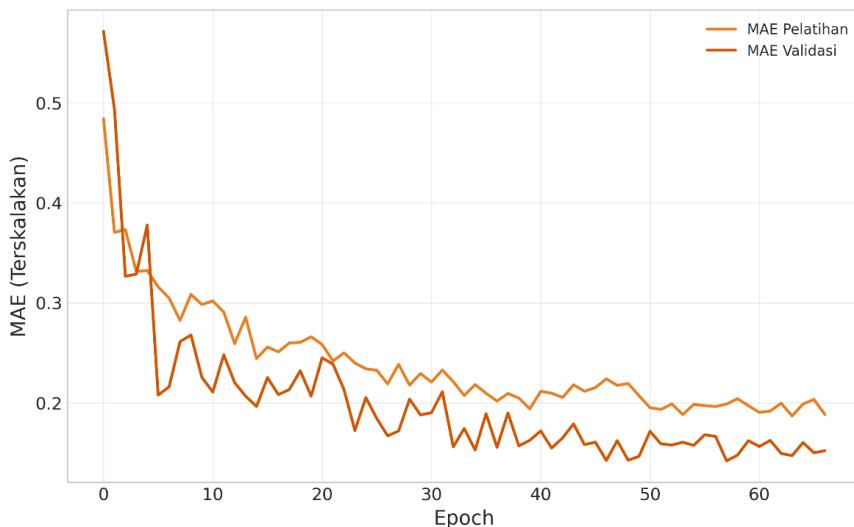
**Gambar 6. ANN Residual**

Gambar 6 menampilkan selisih antara nilai aktual dan prediksi terhadap nilai prediksi. Distribusi titik-titik berada secara acak di sekitar garis nol, tanpa pola sistematis seperti tren naik-turun atau bentuk parabola, yang menunjukkan bahwa model tidak memiliki bias arah tertentu. Meskipun ada beberapa residual ekstrem (hingga  $\pm 15.000$ ), mereka bersifat sporadis dan tidak membentuk pola yang mengancam validitas model. Hal ini mengindikasikan bahwa *error* model bersifat acak dan model telah belajar representasi yang cukup baik dari data pelatihan.



**Gambar 7. Loss Rekonstruksi AutoEncoder**

Grafik ini menunjukkan bahwa *loss* pelatihan dan *loss* validasi menurun secara konsisten dan signifikan seiring bertambahnya *epoch*, yang menunjukkan bahwa model *AutoEncoder* berhasil mempelajari representasi data dengan baik. Kedua kurva loss ini juga tetap berdekatan dan stabil pada nilai yang sangat rendah, sekitar 0.02 hingga 0.03, mengindikasikan bahwa *AutoEncoder* tidak mengalami *overfitting* dan mampu menggeneralisasi dengan efektif pada data yang belum pernah dilihat. Hal ini menegaskan bahwa proses kompresi fitur oleh *AutoEncoder* berjalan sangat efisien dan menghasilkan representasi data yang berkualitas tinggi.



**Gambar 8. Progres Pelatihan ANN**

Gambar 8 menunjukkan bahwa baik MAE pelatihan maupun MAE validasi menurun secara signifikan dan relatif stabil seiring bertambahnya epoch. Kurva validasi MAE menunjukkan penurunan yang sehat dan tetap berdekatan dengan kurva pelatihan MAE, mengindikasikan bahwa model ANN tidak mengalami overfitting yang parah dan memiliki kemampuan generalisasi yang baik. Penurunan *Learning Rate* dan *EarlyStopping* telah membantu model mencapai konvergensi yang optimal, menghasilkan MAE yang rendah dan menunjukkan bahwa model telah belajar untuk membuat prediksi yang akurat terhadap jumlah populasi.

Hasil penelitian menunjukkan bahwa pertumbuhan penduduk Kota Bandung selama periode 2017–2024 relatif stabil, dengan rata-rata pertambahan sekitar 23.000 jiwa per tahun. Temuan ini sejalan dengan teori *urban demographic transition* yang menyatakan bahwa kota besar cenderung memasuki fase konsolidasi populasi akibat urbanisasi selektif, migrasi sirkuler, dan penurunan laju fertilitas (Sarker, 2021; Selmy, 2024). Dalam konteks ini, stabilitas pertumbuhan justru menuntut sistem data kependudukan yang lebih presisi untuk menangkap mobilitas internal dan dinamika administratif yang tidak selalu tercermin dalam angka agregat.

Analisis coverage e-KTP memperlihatkan adanya disparitas antar kecamatan, meskipun secara umum tingkat kepemilikan sudah tinggi. Temuan ini mendukung teori *uneven access to public services*, yang menyatakan bahwa pemerataan kebijakan di tingkat pusat tidak selalu menghasilkan pemerataan hasil di tingkat lokal karena perbedaan akses, kapasitas layanan, dan karakteristik sosial masyarakat (Anindya et al., 2024). Dengan demikian, hasil ini menegaskan bahwa persoalan kepemilikan e-KTP tidak semata administratif, tetapi juga struktural dan spasial.

Dari sisi metodologis, *Autoencoder* terbukti efektif dalam mendeteksi anomali dan inkonsistensi data lintas sumber (BPS dan Dukcapil). Temuan ini memperkuat literatur terbaru yang menyebutkan bahwa *Autoencoder* mampu meningkatkan kualitas data administratif dengan mengidentifikasi pola penyimpangan yang sulit terdeteksi melalui metode statistik konvensional (Neloy, 2024; Aslam, 2024). Dalam konteks tata kelola

pemerintahan, kemampuan ini krusial karena kualitas data merupakan fondasi utama kebijakan berbasis bukti.

Sementara itu, ANN menunjukkan kinerja prediksi yang baik dalam memodelkan hubungan non-linear antara variabel demografi dan kepemilikan e-KTP. Hal ini sejalan dengan teori non-linear demographic modeling yang menekankan bahwa fenomena sosial tidak selalu mengikuti pola linier sederhana (Magazzino, 2024; Grossman, 2023). Integrasi *Autoencoder* dan ANN dalam satu kerangka analitik menjadi *novelty* utama penelitian ini, karena tidak hanya meningkatkan kualitas data (melalui deteksi anomali), tetapi juga langsung memanfaatkannya untuk prediksi kebijakan yang aplikatif di level kecamatan.

## KESIMPULAN

Penelitian ini menjawab permasalahan utama terkait kualitas data kependudukan dan ketimpangan kepemilikan e-KTP di Kota Bandung melalui pendekatan deep learning. Hasil penelitian menunjukkan bahwa meskipun pertumbuhan penduduk relatif stabil, masih terdapat disparitas kepemilikan e-KTP antar kecamatan yang berpotensi menghambat pemerataan layanan publik.

Penerapan *Autoencoder* terbukti efektif dalam mendeteksi anomali dan inkonsistensi data antar sumber, sehingga mampu meningkatkan reliabilitas data kependudukan. Sementara itu, ANN mampu memprediksi kepemilikan e-KTP dengan tingkat akurasi yang tinggi, sehingga dapat digunakan sebagai alat bantu perencanaan berbasis proyeksi. Dengan demikian, kombinasi *Autoencoder* dan ANN tidak hanya menjawab persoalan teknis kualitas data, tetapi juga menyediakan dasar analitik yang kuat untuk mendukung pengambilan keputusan kebijakan kependudukan di tingkat daerah.

Untuk pengembangan penelitian selanjutnya, disarankan agar model diperluas dengan memasukkan data mikro atau data temporal yang lebih rinci (misalnya data bulanan), sehingga dinamika perubahan kependudukan dapat dianalisis secara lebih sensitif. Selain itu, eksplorasi metode hybrid lain seperti kombinasi deep learning dengan pendekatan spasial (GIS) berpotensi meningkatkan ketepatan analisis wilayah.

Secara aplikatif, pemerintah daerah—khususnya Disdukcapil—dapat memanfaatkan hasil penelitian ini sebagai sistem pendukung keputusan (*decision support system*) untuk memprioritaskan intervensi layanan e-KTP pada kecamatan dengan coverage rendah atau tingkat anomali data tinggi. Integrasi model ini ke dalam sistem *e-government* diharapkan dapat meningkatkan efisiensi, akurasi, dan responsivitas pelayanan administrasi kependudukan.

## Glosarium:

### *Artificial Neural Network (ANN)*

Model komputasi yang meniru cara kerja jaringan saraf manusia untuk mengenali pola dan membuat prediksi berdasarkan data. Dalam penelitian ini, ANN digunakan untuk memprediksi tingkat kepemilikan e-KTP berdasarkan karakteristik demografi dan

wilayah.

#### *Autoencoder*

Jenis model deep learning yang digunakan untuk mempelajari pola data dan mendeteksi data yang tidak wajar (anomali). Autoencoder bekerja dengan cara membandingkan data asli dengan hasil rekonstruksinya.

#### Anomali Data

Data yang nilainya tidak wajar atau menyimpang dari pola umum, misalnya jumlah wajib KTP yang lebih besar dari total penduduk atau perbedaan signifikan antar sumber data.

#### *Batch Size*

Jumlah data yang diproses oleh model dalam satu kali proses pembelajaran. Pengaturan ini memengaruhi kecepatan dan stabilitas pelatihan model.

#### *Deep Learning*

Pendekatan pembelajaran mesin berbasis jaringan saraf berlapis yang mampu mengenali pola kompleks dalam data berukuran besar. Deep learning banyak digunakan dalam analisis data kependudukan dan pelayanan publik.

#### *Early Stopping*

Teknik penghentian otomatis pelatihan model ketika kinerja model tidak lagi meningkat, untuk mencegah hasil yang terlalu menyesuaikan data pelatihan (overfitting).

#### *Encoding (Reduksi Dimensi)*

Proses merangkum banyak variabel data menjadi representasi yang lebih ringkas tanpa kehilangan informasi penting, sehingga analisis menjadi lebih efisien.

#### *Gap Kepemilikan e-KTP*

Selisih antara jumlah penduduk yang wajib memiliki e-KTP dan jumlah penduduk yang telah memiliki e-KTP pada suatu wilayah tertentu.

#### *Learning Rate*

Parameter yang mengatur seberapa cepat model memperbarui pembelajarannya. Nilai yang terlalu besar dapat membuat model tidak stabil, sedangkan nilai yang terlalu kecil memperlambat proses pelatihan.

#### *Mean Absolute Error (MAE)*

Ukuran rata-rata selisih absolut antara hasil prediksi model dan data aktual. Metrik ini menunjukkan seberapa besar kesalahan prediksi secara umum.

#### *Mean Squared Error (MSE)*

Ukuran kesalahan prediksi yang menghitung rata-rata kuadrat selisih antara nilai prediksi dan nilai aktual. Digunakan untuk menilai ketepatan model regresi.

#### *Normalisasi Data*

Proses menyamakan skala nilai data agar berada dalam rentang tertentu sehingga dapat diproses secara optimal oleh model analitik.

#### *Overfitting*

Kondisi ketika model terlalu menyesuaikan diri dengan data pelatihan sehingga kinerjanya menurun saat digunakan pada data baru.

#### *Preprocessing Data*

Tahapan awal pengolahan data yang mencakup pembersihan data ganda, pengisian data kosong, dan penyesuaian format data agar siap dianalisis. Reconstruction Error Selisih antara data asli dan data hasil rekonstruksi Autoencoder. Nilai error yang tinggi menandakan adanya potensi anomali.

#### *Root Mean Squared Error (RMSE)*

Akar dari MSE yang memberikan gambaran kesalahan prediksi dalam satuan yang sama dengan data asli, sehingga lebih mudah diinterpretasikan.

#### *Train-Test Split*

Metode pembagian data menjadi data pelatihan dan data pengujian untuk mengukur kemampuan model dalam memprediksi data baru.

#### *Validasi Silang (Cross-Validation)*

Teknik evaluasi model dengan cara menguji model pada beberapa pembagian data yang berbeda untuk memastikan hasil yang stabil dan tidak bergantung pada satu pembagian data saja.

## **DAFTAR PUSTAKA**

- FK Putri, et al. (2024). Persentase Kepemilikan E-KTP Tahun 2019-2023 di Beberapa Kota/Kabupaten. *Jurnal Penelitian & Pengembangan Masyarakat Regional*.
- NB Yuvanti. (2024). Efektivitas Kepemilikan Identitas Kependudukan Digital (IKD) melalui perekaman e-KTP di Kabupaten Jember. *Jurnal Ilmu Administrasi Publik*.
- Anindya, N. T., dkk. (2024). Kajian Kualitas Layanan KTP-Elektronik dalam Inovasi Layanan Teman Dilan di Kabupaten Mojokerto. *Publika: Jurnal Ilmu Kependudukan & Administrasi Publik*.
- Data “Evaluasi Pelayanan E-KTP di Manado” (Alhasni, Onibala, & Rantung, 2025). *Administro: Jurnal Kebijakan dan Ilmu Administrasi Negara*.
- Putri, Y., dkk. (2024). Pelayanan Publik pada Disdukcapil di Kabupaten Kuantan Singgingi: Analisis Wajib vs Kepemilikan E-KTP. *Jurnal Inovasi Sektor Publik*, 4(2).
- Tamimatul Hasanah, I. R., Siska Rahayu Indri Tazkiya, Putri Indriani, & lainnya. (2024). Penggunaan KTP Elektronik Dilihat Dari Perspektif Public Value di Dinas Kependudukan Kota Tangerang. *Innovative: Journal of Social Science Research*, 4(4), 6863-6879.
- Sagita, S. S. S. (2025). Inovasi Pelayanan Publik Kependudukan: Tantangan Perekaman E-KTP di Kabupaten Sumbawa. *SIMBOL: Jurnal Administrasi & Pelayanan Publik*.
- Aslam, M. M. (2024). An improved autoencoder-based approach for anomaly detection in industrial control systems. *International Journal of Critical Infrastructure Protection*. Advance online publication.
- Grossman, I. (2023). Forecasting small area populations with long short-term memory networks. *Population Studies*, 77(2), 345–361.
- Lee, S., Jin, H., Hussain, S., Nengroo, T., Doh, Y., & Heo, C. (2021). Smart metering system capable of anomaly detection by bi-directional LSTM Autoencoder. arXiv preprint.
- Magazzino, C. (2024). An artificial neural network experiment on the prediction of

- unemployment forecasting. *Journal of Economic Dynamics*, 12(1), 45–62.
- Neloy, A. A. (2024). A comprehensive study of auto-encoders for anomaly detection: architectures and reconstruction ability. *Journal of Big Data Research*, 5, 101234.
- Selmy, H. A. (2024). Big data analytics, deep learning techniques and applications: A review. *Information Systems*, 115, 102318.
- Sarker, I. H. (2021). Data science and analytics: An overview from data-driven perspectives. *Annals of Data Science*, 8, 1–24.
- State of the Republic of Indonesia. (2022). Law No. 27 of 2022 concerning Personal Data Protection. Jakarta: Government Gazette. (PDP Law). Retrieved from <https://jdih.setkab.go.id>
- Tang, M., & colleagues. (2022). Machine learning and deep learning for big data analytics: A review of methods and applications. *International Journal of Data Science*, 3(2), 123–145.
- Reuters. (2024, June 28). Indonesia orders audit of government data centres after cyberattack. Reuters. <https://www.reuters.com/technology/cybersecurity/bulk-indonesia-data-hit-by-cyberattack-not-backed-up-officials-say-2024-06-28/>
- Yuvianti, N. B. (2024). Evaluating effectiveness of e-KTP recording and identity services in Jember Regency. *Jurnal Ilmu Administrasi Publik*, 10(1), 33–44.