Masuk: 25 April 2026 | Revisi: 14 mei 2026 | Diterima: 22 Juni 2026
e-SSIN: 3031-3082; DOI: https://doi.org/10.26618/1wvarw57

JUMPTECH

Model *Autoencoder* untuk Deteksi Anomali pada *Log Email* Mahasiswa Universitas Muhammadiyah Makassar

Alvina Damayanti[1] | Fahrim Irhamna Rachman[*2] | Darniati[2]

1 Mahasiswa Program Studi Arsitektur, Fakultas Teknik, Universitas Muhammadiyah Makassar, Indonesia.
Email: 105841111521@student.unismuh.ac.id

2 Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Makassar, Indonesia.
Email: fachrim141020@unismuh.ac.id; darniati@unismuh.ac.id

Korespondensi:
*Fahrim Irhamna Rachman
fachrim141020@unismuh.ac.id

ABSTRAK
Perkembangan teknologi informasi meningkatkan penggunaan email institusi sebagai sarana komunikasi akademik, namun juga menimbulkan risiko keamanan seperti *spam*, *phishing*, dan akses tidak sah. Penelitian ini bertujuan untuk mengimplementasikan model *Autoencoder* dalam mendeteksi anomali pada log email mahasiswa Universitas Muhammadiyah Makassar. Metode yang digunakan adalah pendekatan kuantitatif berbasis *unsupervised learning* dengan memanfaatkan data log email yang telah melalui tahap *preprocessing* dan *feature engineering*. Model *Autoencoder* dirancang menggunakan arsitektur *encoder-decoder* dengan lima *hidden layer* untuk mempelajari pola aktivitas email normal. Proses deteksi anomali dilakukan menggunakan nilai *reconstruction error* dengan *threshold* pada persentil ke-99. Hasil penelitian menunjukkan bahwa model mampu mendeteksi aktivitas anomali dengan baik, di mana sekitar 1% data teridentifikasi sebagai anomali dari keseluruhan *dataset*. Temuan ini menunjukkan bahwa metode *Autoencoder* efektif digunakan untuk mendeteksi aktivitas mencurigakan pada sistem email institusi dan berpotensi mendukung peningkatan keamanan sistem informasi di lingkungan perguruan tinggi.

Kata Kunci:
Autoencoder, deteksi anomali, log email, *unsupervised learning*, keamanan sistem informasi.

ABSTRACT
The development of information technology has increased the use of institutional email as a medium for academic communication, but it has also introduced security risks such as spam, phishing, and unauthorized access. This study aims to implement an Autoencoder model for anomaly detection in student email logs at Universitas Muhammadiyah Makassar. The research employed a quantitative approach based on unsupervised learning using email log data that had undergone preprocessing and feature engineering. The Autoencoder model was designed using an encoder-decoder architecture with five hidden layers to learn normal email activity patterns. Anomaly detection was performed using reconstruction error values with a threshold set at the 99th percentile. The results showed that the model was able to detect anomalous activities effectively, where approximately 1% of the data were identified as anomalies from the entire dataset. These findings indicate that the Autoencoder method is effective for detecting suspicious activities in institutional email systems and has the potential to enhance information system security in higher education environments.

Keyworsds:
Autoencoder, anomaly detection, email logs, unsupervised learning, information system security.

1 | PENDAHULUAN

Perkembangan teknologi informasi telah membawa transformasi besar dalam sistem komunikasi akademik, termasuk penggunaan email institusi dalam aktivitas mahasiswa (Ortega-Fernandez, 2024). Universitas Muhammadiyah Makassar sebagai salah satu institusi pendidikan tinggi telah menyediakan layanan email resmi kepada mahasiswanya sebagai sarana komunikasi, pengiriman tugas, informasi akademik, dan administrasi kampus (Rinaldi, 2021). Namun, dalam penggunaan email tersebut, terdapat potensi penyalahgunaan atau aktivitas mencurigakan yang dapat merugikan institusi, seperti

pengiriman *spam*, percobaan *phishing*, hingga akses tidak sah (Bu & Cho, 2021).

Di era digital saat ini, sistem keamanan informasi telah menjadi prioritas utama bagi organisasi di seluruh dunia. Perkembangan teknologi *cloud computing* dan peningkatan jumlah perangkat yang terhubung ke internet menciptakan tantangan besar dalam mempertahankan keamanan sistem informasi (Zhao et al., 2022). Menurut survei industri terkini, hampir 85% perusahaan global sedang mengembangkan teknologi deteksi anomali untuk mengamankan infrastruktur digital mereka (Kumari et al., 2024). Ancaman keamanan siber yang semakin canggih memerlukan pendekatan proaktif dalam mengidentifikasi aktivitas mencurigakan sebelum berkembang menjadi serangan yang merugikan (Ferrag et al., 2020). Deteksi anomali telah menjadi komponen kritis dalam strategi keamanan siber modern. Sistem log, khususnya log *email*, mengandung informasi berharga yang dapat digunakan untuk mengidentifikasi pola aktivitas normal dan abnormal dalam jaringan organisasi (Chung et al., 2023). Penelitian menunjukkan bahwa analisis log secara manual tidak lagi efektif mengingat volume data yang terus meningkat dan kompleksitas serangan yang berkembang pesat (Yang & Shami, 2024). Oleh karena itu, implementasi sistem deteksi anomali otomatis menjadi kebutuhan mendesak untuk memastikan keamanan dan keandalan sistem informasi.

Autoencoder, sebagai salah satu arsitektur *deep learning*, telah terbukti efektif dalam tugas deteksi anomali tanpa supervisi. Kemampuan *Autoencoder* dalam merekonstruksi data normal dengan tingkat *error* rendah membuatnya ideal untuk mengidentifikasi anomali berdasarkan *reconstruction error* yang tinggi (Campos et al., 2021). Penelitian terbaru menunjukkan bahwa pendekatan berbasis *Autoencoder* dapat mengatasi keterbatasan metode tradisional dalam menangani *dataset* yang tidak seimbang dan pola anomali yang kompleks (Zhang et al., 2025). Variational *Autoencoder* (VAE) dan *Convolutional Autoencoder* telah menunjukkan hasil yang menjanjikan dalam berbagai domain deteksi anomali (Masuda et al., 2021).

Implementasi sistem deteksi anomali di Indonesia juga telah diterapkan dalam berbagai sektor, termasuk pemerintahan dan institusi pendidikan. Penelitian terbaru menunjukkan penggunaan algoritma *Isolation Forest* untuk deteksi anomali pada log sistem *web server* memberikan hasil yang optimal dalam mengidentifikasi aktivitas mencurigakan (Al Kahfi et al., 2025). Demikian pula, implementasi sistem monitoring keamanan dengan notifikasi *WhatsApp* telah terbukti efektif dalam mendeteksi anomali jaringan di lingkungan perusahaan (Khotimah & Maharjan, 2025).

Sistem email merupakan salah satu vektor serangan utama dalam lingkungan akademik, namun penelitian khusus tentang deteksi anomali pada log email masih terbatas. Aktivitas email yang tidak normal dapat mengindikasikan berbagai ancaman keamanan, mulai dari *phishing*, *malware distribution*, hingga data *exfiltration* (Yu et al., 2025). Oleh karena itu, diperlukan penelitian yang mengeksplorasi penerapan model *Autoencoder* untuk mengidentifikasi pola anomali dalam log email secara otomatis dan akurat (Studiawan & Sohel, 2021). Selain itu, pendekatan *unsupervised learning* berbasis *Autoencoder* dinilai mampu mendeteksi pola aktivitas yang tidak biasa tanpa memerlukan pelabelan data secara manual. Metode ini bekerja dengan mempelajari karakteristik aktivitas normal, kemudian mengidentifikasi penyimpangan berdasarkan nilai *reconstruction error* yang dihasilkan model. Sejumlah penelitian terdahulu menunjukkan bahwa pendekatan *Autoencoder* memiliki kemampuan yang baik dalam mendeteksi anomali pada data jaringan, sistem log, dan aktivitas keamanan siber dengan tingkat akurasi yang tinggi serta mampu menangani data dalam jumlah besar dan kompleks (Torabi et al., 2023). Selain itu, penelitian lain juga menyebutkan bahwa model berbasis *Autoencoder* efektif digunakan dalam aktivitas mencurigakan pada sistem jaringan modern karena mampu mengenali pola anomali yang sebelumnya tidak diketahui (*unknown attack*) secara lebih adaptif (Al-Zubidi & Farhan, 2025).

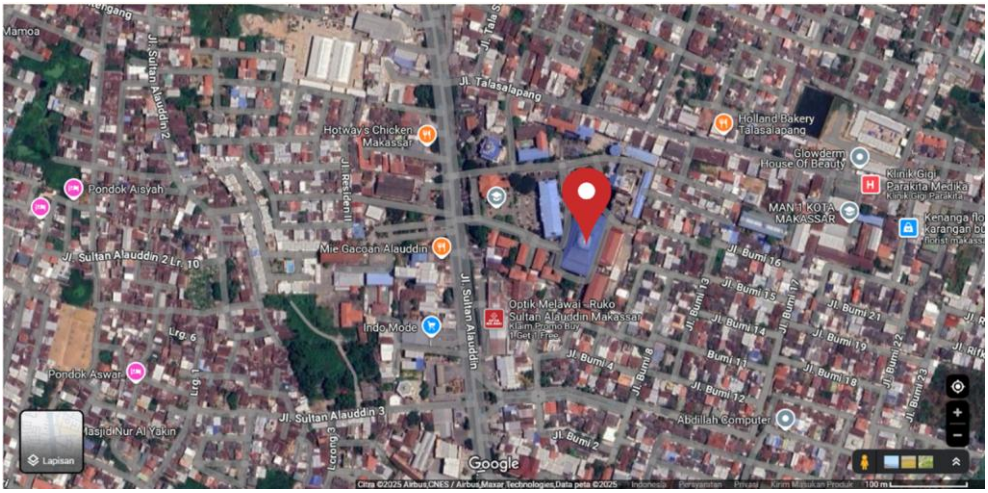
Penelitian ini diharapkan dapat memberikan kontribusi praktis bagi peningkatan keamanan sistem informasi di lingkungan Universitas Muhammadiyah Makassar khususnya, dan institusi pendidikan di Indonesia pada umumnya. Selain itu, hasil penelitian ini juga dapat menjadi referensi untuk pengembangan sistem keamanan email yang lebih canggih dan adaptif terhadap ancaman siber yang terus berkembang.

2 | METODE

Metodologi penelitian ini disusun untuk mendukung penerapan model *Autoencoder* dalam mendeteksi anomali pada log email mahasiswa Universitas Muhammadiyah Makassar. Penelitian menggunakan pendekatan kuantitatif dengan tahapan pengumpulan data, *preprocessing*, perancangan model, pelatihan, hingga evaluasi hasil deteksi anomali. Metode yang digunakan berfokus pada analisis pola aktivitas email mahasiswa berdasarkan data log yang diperoleh dari sistem email institusi. Proses penelitian dilakukan secara bertahap untuk memastikan model mampu mengenali pola aktivitas normal dan mendeteksi aktivitas yang menyimpang secara otomatis. Selain itu, penelitian memanfaatkan teknik *unsupervised learning* sehingga model dapat mempelajari karakteristik data tanpa memerlukan pelabelan data secara manual.

2.1 | Lokasi Penelitian

Penelitian ini dilakukan di Universitas Muhammadiyah Makassar dengan objek kajian berupa sistem email institusi yang digunakan oleh mahasiswa sebagai sarana komunikasi akademik. Lokasi penelitian dipilih karena tingginya intensitas penggunaan email resmi kampus yang berpotensi menimbulkan aktivitas anomali, seperti *spam*, *phishing*, dan penyalahgunaan akses. Data yang digunakan berupa log aktivitas email mahasiswa yang telah dianonimkan untuk menjaga privasi pengguna dan mencerminkan pola komunikasi yang dinamis. Data tersebut melalui tahap pra-pemrosesan dan rekayasa fitur untuk merepresentasikan aspek waktu, konten, dan perilaku pengguna, kemudian digunakan dalam penerapan model *Autoencoder* berbasis *unsupervised learning* guna mendeteksi aktivitas anomali secara otomatis dan adaptif.



GAMBAR 1 Loaksi Penelitian

2.2 | Teknik Pengumpulan Data

Metode pengumpulan data dalam penelitian ini dilakukan dengan menggunakan data sekunder berupa log aktivitas email mahasiswa Universitas Muhammadiyah Makassar. Data log diperoleh dari sistem email institusi yang mencatat aktivitas pengiriman dan penerimaan email mahasiswa dalam periode tertentu. Untuk menjaga keamanan dan privasi pengguna, data yang digunakan telah melalui proses anonimisasi dengan menghilangkan informasi identitas pribadi. Selanjutnya, data log email diproses melalui tahap pra-pemrosesan, yang meliputi pembersihan data, penghapusan duplikasi, serta penyesuaian format data agar siap digunakan dalam proses analisis. Selain itu, dilakukan rekayasa fitur untuk menghasilkan atribut-atribut yang mempresentasikan aspek waktu, konten, dan perilaku pengguna, sehingga data yang diperoleh relevan dan representatif untuk keperluan deteksi anomali.

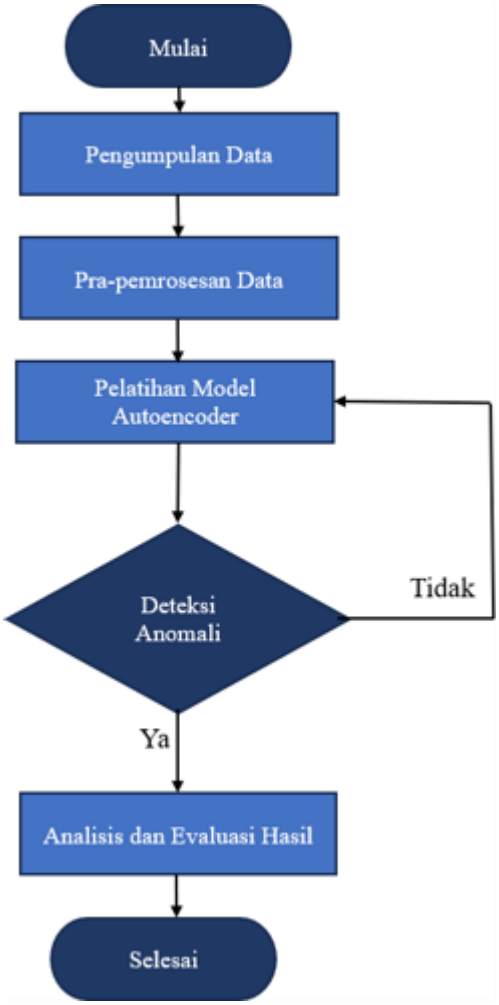
Analisis sistem dalam penelitian ini dilakukan dengan merancang sistem deteksi anomali berbasis *unsupervised learning* menggunakan model *Autoencoder*. Sistem dirancang untuk mempelajari pola aktivitas email normal melalui proses pelatihan menggunakan data log yang diasumsikan sebagai aktivitas normal. Model *Autoencoder* bekerja dengan cara merekonstruksi data *input* dan menghitung nilai *reconstruction error* sebagai indikator penyimpangan aktivitas. Aktivitas email dengan nilai *reconstruction error* yang melebihi ambang batas tertentu diklasifikasikan sebagai anomali. Analisis sistem mencakup tahapan perancangan arsitektur model, penentuan parameter pelatihan, serta evaluasi hasil deteksi anomali berdasarkan pola yang dihasilkan. Dengan pendekatan ini, sistem mampu mendeteksi aktivitas mencurigakan secara otomatis tanpa memerlukan data berlabel dan adaptif terhadap perubahan karakteristik data log email.

2.3 | Perancangan Metode *Autoencoder*

Perancangan sistem dimulai dari pengumpulan data log email yang mencakup informasi waktu akses, pengirim, penerima, dan metadata lainnya. Selanjutnya dilakukan tahap *preprocessing* untuk membersihkan data, menangani data yang tidak valid, serta melakukan *feature engineering* agar data dapat diproses oleh model *machine learning*.

Data yang telah diproses kemudian dibagi menjadi data pelatihan dan data pengujian. Model *autoencoder* dilatih menggunakan data normal untuk mempelajari pola aktivitas *email* yang umum terjadi. Setelah proses pelatihan selesai, model digunakan untuk mendeteksi anomali pada data baru berdasarkan nilai *reconstruction error*. Data dengan nilai *error* melebihi

threshold yang ditentukan diklasifikasikan sebagai anomali. Hasil evaluasi kemudian digunakan sebagai dasar dalam penarikan kesimpulan dan rekomendasi pengembangan sistem deteksi anomali.



GAMBAR 4 Pespektif Pusat Kebudayaan

Dua Berdasarkan **GAMBAR 2**, proses perancangan sistem deteksi anomali dimulai dari tahap pengumpulan data log email yang berisi informasi aktivitas pengguna, kemudian dilanjutkan dengan tahap *preprocessing* dan *feature engineering* untuk membersihkan serta menyiapkan data agar dapat diproses oleh model *machine learning*. Setelah itu, data dibagi menjadi data pelatihan dan data pengujian, di mana model *Autoencoder* dilatih menggunakan data normal untuk mempelajari pola aktivitas email yang umum terjadi. Model yang telah dilatih kemudian digunakan untuk mendeteksi aktivitas anomali berdasarkan nilai *reconstruction error*, sehingga data dengan nilai kesalahan rekonstruksi melebihi *threshold* dikategorikan sebagai anomal.

2.4 | Teknik Pengujian Sistem

Pengujian sistem dilakukan untuk mengevaluasi kemampuan model *Autoencoder* dalam mendeteksi anomali pada log email mahasiswa. Data pengujian yang digunakan merupakan data yang telah melalui tahap *preprocessing* dan tidak digunakan saat pelatihan model, sehingga proses evaluasi dapat dilakukan secara objektif terhadap kemampuan model dalam mengenali pola baru. Pada tahap ini, model melakukan proses rekonstruksi terhadap data pengujian, kemudian dihitung nilai *reconstruction error* dari setiap data log untuk mengetahui tingkat perbedaan antara data asli dan hasil rekonstruksi model.

Proses pengujian dilakukan dengan membandingkan nilai *reconstruction error* terhadap nilai *threshold* yang telah ditentukan sebelumnya. Data yang memiliki nilai kesalahan rekonstruksi melebihi *threshold* diklasifikasikan sebagai anomali karena dianggap memiliki pola aktivitas yang berbeda dari aktivitas normal yang telah dipelajari model. Sebaliknya, data dengan nilai *error* di bawah *threshold* dikategorikan sebagai aktivitas normal. Dengan adanya proses pengujian tersebut, dapat diketahui tingkat performa model *Autoencoder* dalam mendeteksi aktivitas mencurigakan pada sistem email mahasiswa.

2.5 | Teknik Analisis Data

Teknik analisis data pada penelitian ini dilakukan secara kuantitatif untuk menilai efektivitas model *Autoencoder* dalam mendeteksi anomali pada log email mahasiswa. Tahap analisis dimulai dari proses pengumpulan data log email yang diperoleh dari sistem email kampus, kemudian dilakukan tahap *preprocessing* untuk membersihkan data dari nilai kosong (*missing value*), data duplikat, serta data yang tidak valid agar kualitas *dataset* menjadi lebih baik. Setelah itu dilakukan proses *feature engineering* untuk menghasilkan fitur-fitur yang mampu mempresentasikan pola aktivitas pengguna email, seperti pola waktu penggunaan, frekuensi aktivitas, ukuran pesan, dan aktivitas pengguna berdasarkan periode tertentu. Data yang telah diproses kemudian dinormalisasi agar setiap fitur memiliki rentang nilai yang seimbang sebelum digunakan pada proses pelatihan model.

Selanjutnya, *dataset* dibagi menjadi data pelatihan, data validasi, dan data pengujian untuk memastikan model dapat dievaluasi secara objektif. Data pelatihan digunakan untuk melatih model dalam mempelajari pola aktivitas email normal, sedangkan data validasi digunakan untuk memantau performa model selama proses pelatihan dan mencegah terjadinya *overfitting*. Data pengujian digunakan untuk mengevaluasi kemampuan model dalam mendeteksi anomali pada data yang belum pernah dipelajari sebelumnya. Pada tahap ini juga ditentukan beberapa parameter model, seperti *learning rate*, *batch size*, jumlah *epoch*, dan ukuran *latent space* untuk mengoptimalkan proses pembelajaran model *Autoencoder*. Parameter tersebut dipilih agar model mampu menghasilkan proses rekonstruksi data yang stabil dan akurat.

Model *Autoencoder* kemudian diterapkan untuk mempelajari pola aktivitas normal melalui proses *encoding* dan *decoding*. Setelah proses pelatihan selesai, model menghasilkan nilai *reconstruction error* yang digunakan sebagai dasar dalam mendeteksi anomali. Data yang memiliki nilai *reconstruction error* tinggi dianggap memiliki pola aktivitas yang berbeda dari data normal sehingga dikategorikan sebagai anomali. Sebaliknya, data dengan nilai kesalahan rekonstruksi rendah dikategorikan sebagai aktivitas normal karena model mampu merekonstruksi data tersebut dengan baik.

3 | Hasil Dan Pembahasan

Berdasarkan metodologi penelitian yang telah diterapkan, tahap selanjutnya adalah penyajian hasil penelitian dan pembahasan terhadap implementasi model deteksi anomali pada aktivitas email mahasiswa. Pada bagian ini dijelaskan proses pengumpulan data, implementasi model *Autoencoder*, serta analisis hasil deteksi berdasarkan nilai *reconstruction error* yang dihasilkan model. Selain itu, dilakukan pembahasan terhadap karakteristik data yang teridentifikasi sebagai aktivitas normal maupun anomali guna mengevaluasi kemampuan model dalam mendeteksi aktivitas *email* yang menyimpang pada sistem *email* institusi.

3.1 | Pengumpulan Data

Penelitian ini dilakukan di Universitas Muhammadiyah Makassar melalui Lembaga Sistem Informasi Teknologi (LSIT) yang mengelola layanan email resmi mahasiswa. Data yang digunakan berupa data log aktivitas email mahasiswa yang diperoleh dari server email kampus dan disimpan dalam format *.csv*. *Dataset* utama yang digunakan adalah *audit-gmail-2025.csv*, yang berisi rekaman aktivitas email mahasiswa selama periode tertentu pada tahun 2025 dan dihasilkan oleh sistem audit keamanan kampus. Data log tersebut mencakup berbagai metadata aktivitas email, seperti waktu aktivitas, alamat IP pengguna, jenis aktivitas email, ukuran pesan, identitas pengguna, serta informasi lain yang berkaitan dengan aktivitas akses email mahasiswa. Pengumpulan data dilakukan untuk memperoleh pola aktivitas email yang dapat digunakan dalam analisis perilaku pengguna serta mendeteksi aktivitas yang tidak normal pada sistem email kampus.

Sebelum digunakan dalam proses pelatihan model, data log email terlebih dahulu melalui tahap *preprocessing* yang meliputi pembersihan data, penanganan nilai kosong (*missing value*), penghapusan data duplikat, serta transformasi data agar sesuai dengan format yang dibutuhkan oleh model *machine learning*. Selanjutnya dilakukan proses *feature engineering* untuk menghasilkan atribut tambahan, seperti frekuensi penggunaan email, panjang pesan, ukuran pesan, serta pola aktivitas berdasarkan waktu penggunaan. Data yang telah diproses kemudian dinormalisasi agar setiap fitur memiliki rentang nilai yang seimbang sebelum digunakan sebagai masukan dalam pelatihan model *Autoencoder* dengan pendekatan *unsupervised learning*. Pendekatan ini memungkinkan model mempelajari pola aktivitas email normal dan mendeteksi aktivitas yang menyimpang berdasarkan nilai *reconstruction error* yang dihasilkan.

3.2 | Analisis Kebutuhan Ruang

Implementasi dilakukan dengan menerapkan model deteksi anomali berbasis *Autoencoder* pada data log aktivitas email mahasiswa Universitas Muhammadiyah Makassar tahun 2025 yang telah melalui tahap pra-pemrosesan dan rekayasa fitur.

Model dilatih menggunakan data aktivitas email normal untuk mempelajari pola dasar perilaku pengguna, kemudian nilai *reconstruction error* digunakan sebagai indikator dalam mengidentifikasi aktivitas anomali. Sebagai masukan model, dipilih sembilan fitur utama yang mempresentasikan perilaku email mahasiswa dapat ditunjukkan pada **TABEL 1**. Model *Autoencoder* yang digunakan memiliki arsitektur *multilayer perceptron* (MLP) dengan lima *hidden layer*, dengan rincian arsitektur ditunjukkan pada **TABEL 2**.

TABEL 1. Model *Autoencoder*

No	Nama Fitur	Kategori	Deskripsi
1	*Hour*	Waktu	Jam saat aktivitas terjadi
2	*Dayofweek*	Waktu	Hari aktivitas (0=Senin, 6=Minggu)
3	*is_weekend*	Waktu	Penanda apakah akhir pekan (1/0)
4	*text_len*	Konten	Panjang teks *email* (subjek atau isi)
5	*size_bytes*	Konten	Ukuran *email* dalam *byte*
6	*ip_last_octet*	Jaringan	*Oktet* terakhir alamat IP
7	*user_count*	Perilaku	Total aktivitas per pengguna
8	*event_count*	Perilaku	Total aktivitas per *event*
9	*user_hour_count*	Perilaku	Total aktivitas per pengguna dalam satu jam tertentu

Berdasarkan **TABEL 1**, fitur yang digunakan pada model deteksi anomali terdiri atas empat kategori utama, yaitu waktu, konten, jaringan, dan perilaku pengguna. Fitur waktu seperti *Hour*, *Dayofweek*, dan *is_weekend* digunakan untuk mengidentifikasi pola aktivitas email berdasarkan waktu tertentu, karena aktivitas anomali umumnya terjadi di luar jam normal penggunaan. Selanjutnya, fitur konten seperti *text_len* dan *size_bytes* digunakan untuk menganalisis karakteristik isi email, di mana email dengan ukuran atau panjang teks yang tidak biasa dapat menjadi indikasi aktivitas mencurigakan. Pada kategori jaringan, fitur *ip_last_octet* dimanfaatkan untuk melihat pola akses berdasarkan alamat IP pengguna. Sementara itu, fitur perilaku seperti *user_count*, *event_count*, dan *user_hour_count* digunakan untuk mengukur intensitas aktivitas pengguna, sehingga model dapat mengenali lonjakan aktivitas yang berbeda dari pola normal. Kombinasi seluruh fitur tersebut diharapkan mampu meningkatkan kemampuan model *Autoencoder* dalam membedakan aktivitas normal dan aktivitas anomali pada log email mahasiswa.

TABEL 2. *Model Autoencoder*

Layer	**Jumlah *Neuron***	***Activation Function***
Input	9	-
Hidden Layer 1	64	*ReLU*
Hidden Layer 2	32	*ReLU*
Bottleneck	8	*Linear*
Hidden Layer 3	32	*ReLU*
Hidden Layer 4	64	*ReLU*
Output	9	*Linear*

Berdasarkan **TABEL 2**, model *Autoencoder* dibangun menggunakan arsitektur *Multilayer Perceptron* (*MLP*) yang terdiri atas lapisan *encoder* dan *decoder*. Lapisan *input* memiliki 9 *neuron* yang sesuai dengan jumlah fitur masukan pada data log email. Pada tahap *encoder*, jumlah *neuron* dikurangi secara bertahap dari 64 *neuron* pada *Hidden Layer* 1 menjadi 32 *neuron* pada *Hidden Layer* 2, hingga mencapai lapisan *Bottleneck* dengan 8 *neuron*. Proses reduksi dimensi ini bertujuan untuk menghasilkan representasi fitur penting dari data aktivitas email. Selanjutnya, pada tahap *decoder*, jumlah *neuron* ditingkatkan kembali dari 32 *neuron* pada *Hidden Layer* 3 menjadi 64 *neuron* pada *Hidden Layer* 4 untuk merekonstruksi data ke bentuk semula pada lapisan *output*. Fungsi aktivasi *ReLU* digunakan pada *hidden layer* karena mampu meningkatkan efisiensi pembelajaran pola *nonlinier,* sedangkan fungsi *linear* pada *bottleneck* dan *output layer* digunakan agar proses rekonstruksi data numerik dapat dilakukan dengan lebih stabil. Struktur arsitektur tersebut memungkinkan model mempelajari pola aktivitas normal dan mendeteksi penyimpangan berdasarkan nilai *reconstruction error* yang dihasilkan.

3.3 | Analisis dan Pembahasan

Setelah proses pelatihan model *Autoencoder* menggunakan data pelatihan, dilakukan tahap deteksi anomali berdasarkan nilai *reconstruction error*, yaitu selisih antara data *input* dan hasil rekonstruksi model. Semakin besar nilai *reconstruction error,* semakin besar kemungkinan suatu aktivitas email diklasifikasikan sebagai anomali. Dalam penelitian ini, ambang batas deteksi ditetapkan pada persentil ke-99 dari distribusi *reconstruction error* untuk memisahkan aktivitas normal dan anomali.

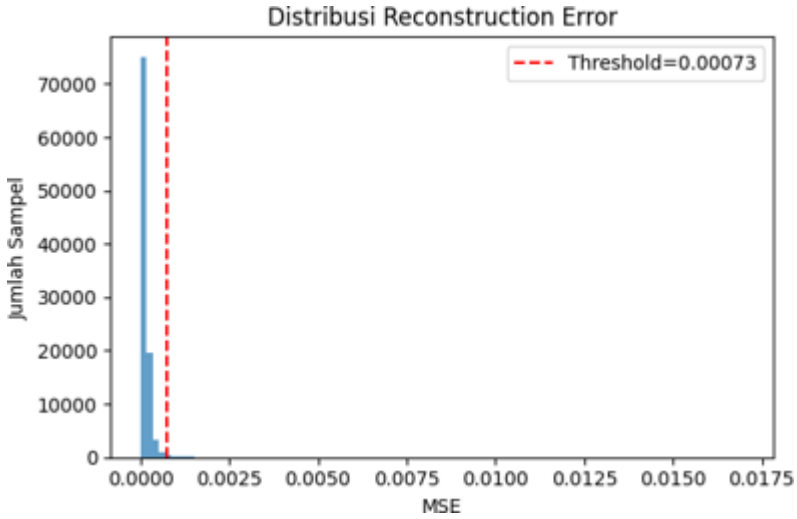
TABEL 3. Hasil Klasifikasi Deteksi Anomali

Kategori	Jumlah Data	*Persentase*
Normal	99	99.00%
Anomali	1	1.00%
Total	100	100%

Berdasarkan **TABEL 3**, hasil klasifikasi menunjukkan bahwa sebagian besar aktivitas email berhasil dikategorikan sebagai aktivitas normal, yaitu sebanyak 99 data atau sebesar 99,00% dari total data pengujian. Sementara itu, terdapat 1 data atau sebesar 1,00% yang terdeteksi sebagai anomali berdasarkan nilai *reconstruction error* yang melebihi ambang batas yang telah ditentukan. Hasil tersebut menunjukkan bahwa model *Autoencoder* mampu mempelajari pola aktivitas email normal dengan baik sehingga dapat mengidentifikasi data yang memiliki penyimpangan pola secara efektif. *Persentase* data anomali yang relatif kecil juga mengindikasikan bahwa sebagian besar aktivitas email mahasiswa berada dalam kondisi normal, sedangkan data yang terdeteksi sebagai anomali berpotensi mempresentasikan aktivitas mencurigakan atau tidak biasa pada sistem email.

3.4 | Analisis Hasil Klasifikasi Deteksi Anomali

Untuk mengevaluasi performa model *Autoencoder* dalam mendeteksi anomali, dilakukan analisis terhadap distribusi nilai *reconstruction error* menggunakan *Mean Squared Error* (MSE). Distribusi nilai MSE digunakan untuk melihat perbedaan karakteristik antara data normal dan data anomali berdasarkan tingkat kesalahan rekonstruksi yang dihasilkan model. Pada penelitian ini digunakan nilai *threshold* sebesar 0,00073 sebagai batas pemisah antara aktivitas normal dan aktivitas anomali. Berdasarkan hasil analisis, sebagian besar data memiliki nilai MSE yang rendah dan berada di bawah *threshold*, yang menunjukkan bahwa model mampu merekonstruksi pola aktivitas normal dengan baik. Sementara itu, data dengan nilai MSE yang melebihi *threshold* dikategorikan sebagai anomali karena memiliki pola aktivitas yang berbeda dari data normal. Adapun distribusi *reconstruction error* dapat ditunjukkan pada **GAMBAR 3** berikut.



GAMBAR 3. Distribusi *Reconstruction Error*

Berdasarkan **GAMBAR 3**, distribusi nilai *reconstruction error* menunjukkan bahwa sebagian besar data terkonsentrasi pada nilai MSE yang rendah dan berada di bawah nilai *threshold* sebesar 0,00073. Hal ini menandakan bahwa model *Autoencoder* mampu melakukan proses rekonstruksi data normal dengan tingkat kesalahan yang kecil, sehingga pola aktivitas email normal dapat dipelajari dengan baik oleh model. Sebaliknya, terdapat sejumlah kecil data yang memiliki nilai MSE di atas *threshold*, yang menunjukkan adanya perbedaan pola aktivitas dibandingkan dengan data normal. Data tersebut dikategorikan sebagai anomali karena model mengalami kesulitan dalam merekonstruksi pola aktivitasnya secara akurat.

Distribusi nilai *reconstruction error* yang cenderung terkumpul di bawah *threshold* juga menunjukkan bahwa model memiliki kemampuan yang baik dalam membedakan aktivitas normal dan aktivitas anomali. Semakin jauh nilai MSE suatu data dari batas *threshold*, maka semakin besar tingkat penyimpangan aktivitas tersebut terhadap pola normal yang telah dipelajari model. Dengan demikian, hasil visualisasi distribusi *reconstruction error* membuktikan bahwa metode *Autoencoder* efektif digunakan dalam mendeteksi aktivitas email yang tidak normal pada data log email mahasiswa.

3.5 | Perbandingan Hasil Sebelum dan Sesudah Optimasi

Pada tahap ini dilakukan analisis lanjutan terhadap hasil deteksi anomali yang dihasilkan oleh model *Autoencoder* berdasarkan nilai *reconstruction error*. Dari keseluruhan data yang dianalisis, sebanyak 1% data terklasifikasi sebagai anomali, sedangkan 99% lainnya tergolong aktivitas normal. Untuk mengkaji karakteristik perbedaan antara kedua kategori tersebut, dilakukan pengambilan masing-masing sepuluh sampel dengan nilai *reconstruction error* tertinggi sebagai representasi data anomali dan sepuluh sampel dengan nilai *error* terendah sebagai representasi data normal. Analisis ini bertujuan untuk mengidentifikasi pola penyimpangan pada aktivitas email serta membandingkannya dengan perilaku normal, sehingga memberikan pemahaman lebih lanjut terkait potensi aktivitas mencurigakan dalam log email.

TABEL 4. Hasil Klasifikasi Deteksi Anomali

No	Tanggal & Waktu	*Event*	Email Tujuan	Terdeteksi
1	2025-08-22T08:00:09+08:00	*View*	andi_agung@student.unismuh.ac.id	1
2	2025-08-22T08:00:09+08:00	*Open*	andi_agung@student.unismuh.ac.id	1
3	2025-09-20T07:24:13+08:00	*Receive*	105841112320@student.unismuh.ac.id	1
4	2025-10-04T07:42:01+08:00	*Receive*	reskiawalia.s@student.unismuh.ac.id	1
5	2025-10-04T07:44:56+08:00	*Receive*	Tritiara@student.unismuh.ac.id	1
6	2025-09-01T03:18:10+08:00	*Open*	105841109221@student.unismuh.ac.id	1
7	2025-09-05T07:58:37+08:00	*Receive*	jumriati@student.unismuh.ac.id	1
8	2025-09-22T07:46:21+08:00	*Move to Trash*	105841107522@student.unismuh.ac.id	1
9	2025-09-22T07:46:20+08:00	*Move to Trash*	105841104922@student.unismuh.ac.id	1
10	2025-08-25T06:40:40+08:00	*Delete*	105841105421@student.unismuh.ac.id	1

Berdasarkan **TABEL 4**, data yang terdeteksi sebagai anomali menunjukkan karakteristik aktivitas email yang berbeda dibandingkan pola normal. Aktivitas anomali didominasi oleh *event* seperti *View*, *Open*, *Move to Trash*, dan *Delete*, yang muncul pada waktu tertentu dengan pola akses yang tidak umum. Selain itu, beberapa aktivitas *Receive* juga terdeteksi sebagai anomali karena memiliki nilai *reconstruction error* yang tinggi, sehingga dianggap menyimpang dari pola mayoritas data pelatihan. Hal ini menunjukkan bahwa model *Autoencoder* mampu mengenali aktivitas email yang memiliki karakteristik berbeda dari perilaku normal pengguna. Aktivitas seperti penghapusan email secara berurutan (*Move to Trash* dan *Delete*) maupun pembukaan email pada waktu tertentu dapat mengindikasikan adanya perilaku mencurigakan atau aktivitas yang jarang terjadi pada sistem email mahasiswa.

TABEL 5. Hasil Klasifikasi Deteksi Anomali

No	Tanggal & Waktu	*Event*	To (Email)	Terdeteksi
1	2025-10-15 16:11:34+00:00	*Receive*	105841101721@student.unismuh.ac.id	Normal
2	2025-10-15 15:56:43+00:00	*Receive*	105841102922@student.unismuh.ac.id	Normal
3	2025-09-17 15:47:14+00:00	*Receive*	105841100121@student.unismuh.ac.id	Normal
4	2025-10-02 17:40:47+00:00	*Receive*	105841110722@student.unismuh.ac.id	Normal
5	2025-10-02 17:36:41+00:00	*Receive*	105841106221@student.unismuh.ac.id	Normal
6	2025-10-02 17:37:24+00:00	*Receive*	105841116522@student.unismuh.ac.id	Normal
7	2025-10-02 17:16:10+00:00	*Receive*	105841103622@student.unismuh.ac.id	Normal
8	2025-10-02 17:19:16+00:00	*Receive*	105841104321@student.unismuh.ac.id	Normal
9	2025-10-15 15:56:25+00:00	*Receive*	105841116522@student.unismuh.ac.id	Normal
10	2025-09-02 15:55:26+00:00	*Receive*	anwardhyndha@student.unismuh.ac.id	Normal

4 | KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, model *autoencoder* berhasil diimplementasikan untuk mendeteksi anomali pada data log email mahasiswa Universitas Muhammadiyah Makassar. Model dirancang menggunakan arsitektur *encoder-decoder* dengan lima lapisan tersembunyi yang mampu mempelajari pola aktivitas email normal berdasarkan data pelatihan. Dalam prosesnya, beberapa fitur penting seperti waktu akses, ukuran email, panjang teks, dan frekuensi aktivitas pengguna digunakan sebagai parameter utama untuk mendukung proses identifikasi pola perilaku pengguna. Hasil pengujian menunjukkan bahwa model memiliki kemampuan yang baik dalam membedakan aktivitas normal dan aktivitas yang menyimpang.

Dengan menggunakan nilai *threshold* berdasarkan *reconstruction error* pada persentil ke-99, model berhasil mengidentifikasi sekitar 1% data sebagai anomali dari keseluruhan *dataset*. Temuan ini menunjukkan bahwa sebagian besar aktivitas email mahasiswa berada dalam kondisi normal, sementara sebagian kecil aktivitas terdeteksi memiliki pola yang tidak biasa dan berpotensi menjadi indikasi ancaman keamanan atau penyalahgunaan sistem email.

Daftar Pustaka

Al-Zubidi, A. F., & Farhan, A. K. (2025). *Multi-Class Anomaly Detection in Network Intrusion Detection Using Variational Autoencoder*. *15*(6), 1265–1278.

Al Kahfi, M., Buaton, R., & Prahmana, G. (2025). *Journal of Artificial Intelligence and Engineering Applications Implementation of Isolation Forest-Based Machine Learning in Batch Anomaly Detection on Zeek Log Data (Case Study: Langkat Regency Communication and Information Agency)*. *5*(1), 2808–4519.

Bu, S.-J., & Cho, S.-B. (2021). Deep Character-Level Anomaly Detection Based on a Convolutional Autoencoder for Zero-Day Phishing URL Detection. *Electronics*, *10*(12). https://doi.org/10.3390/electronics10121492

Campos, D., Kieu, T., Guo, C., Huang, F., Zheng, K., Yang, B., & Jensen, C. S. (2021). Unsupervised Time Series Outlier Detection with Diversity-Driven Convolutional Ensembles - Extended Version. *CoRR*, *abs/2111.11108*.

Chung, M.-H., Wang, L., Li, S., Yang, Y., Giang, C., Jerath, K., Raman, A., Lie, D., & Chignell, M. (2023). *Implementing Active Learning in Cybersecurity: Detecting Anomalies in Redacted Emails*.

Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, *50*, 102419. https://doi.org/https://doi.org/10.1016/j.jisa.2019.102419

Khotimah, K., & Maharjan, R. (2025). *Improving Detection Accuracy of Brute-Force Attacks on MariaDB Using Standard Isolation Forest : A Comparative Analysis with Rotated Variant*. *25*(1), 145–160. https://doi.org/10.30812/matrik.v25i1.5817

Kumari, S., Prabha, C., Karim, A., Hassan, M. M., & Azam, S. (2024). A Comprehensive Investigation of Anomaly Detection Methods in Deep Learning and Machine Learning: 2019–2023. *IET Information Security*, *2024*(1). https://doi.org/10.1049/2024/8821891

Masuda, M., Hachiuma, R., Fujii, R., Saito, H., & Sekikawa, Y. (2021). Toward Unsupervised 3D Point Cloud Anomaly Detection Using Variational Autoencoder. *Proceedings - International Conference on Image Processing, ICIP*, *2021-September*, 3118–3122. https://doi.org/10.1109/ICIP42928.2021.9506795

Ortega-Fernandez, I. (2024). *Network intrusion detection system for DDoS attacks in ICS using deep autoencoders*. *3*, 5059–5075. https://doi.org/10.1007/s11276-022-03214-3

Rinaldi, L. (2021). *Never missing the whole picture: Intellectual development from a neuroconstructivist perspective* (pp. 280–295). https://doi.org/10.4324/9780429445590-19-19

Studiawan, H., & Sohel, F. (2021). Anomaly detection in a forensic timeline with deep autoencoders. *Journal of Information Security and Applications*, *63*, 103002. https://doi.org/https://doi.org/10.1016/j.jisa.2021.103002

Torabi, H., Mirtaheri, S. L., & Greco, S. (2023). Practical autoencoderbased anomaly detection by using vector reconstruction error. *Cybersecurity*, 1–13. https://doi.org/10.1186/s42400-022-00134-9

Yang, L., & Shami, A. (2024). Towards Autonomous Cybersecurity: An Intelligent AutoML Framework for Autonomous Intrusion Detection. In *AutonomousCyber 2024 - Proceedings of the Workshop on Autonomous Cybersecurity, Co-Located with: CCS 2024* (Vol. 1, Issue 1). Association for Computing Machinery. https://doi.org/10.1145/3689933.3690833

Yu, H., Song, S., Sun, L., Su, W., Yang, X., & Liu, C. (2025). All-directional Disparity Estimation for Real-world QPD Images. *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 21836–21846. https://doi.org/10.1109/CVPR52734.2025.02034

Zhang, B., Kieu, T., Qiu, X., Guo, C., Hu, J., Zhou, A., Jensen, C. S., & Yang, B. (2025). An Encode-then-Decompose Approach to Unsupervised Time Series Anomaly Detection on Contaminated Training Data – Extended Version. *Icde*, 1–15.

Zhao, H., Benomar, Z., Pfandzelter, T., & Georgantas, N. (2022). Supporting Multi-Cloud in Serverless Computing. *Proceedings - 2022 IEEE/ACM 15th International Conference on Utility and Cloud Computing, UCC 2022*, 285–290. https://doi.org/10.1109/UCC56403.2022.00051

y of Guangdong Lion Dance in Chinese Intangible Cultural Heritage Through Anthropological Perspectives. *Ijsasr*, *4*(4), 391-400. https://doi.org/10.60027/ijsasr.2024.4511