

# Proposal Klasifikasi Lalu Lintas Jaringan Dengan 5 Model Machine Learning.

## DAFTAR ISI

DAFTAR ISI	X
<b>BAB 1. PENDAHULUAN</b>	X
<b>BAB 2. TINJAUAN PUSTAKA</b>	X
<b>BAB 3. METODE RISET</b>	X
<b>BAB 4. BIAYA DAN JADWAL KEGIATAN</b>	X
4.1 Anggaran Biaya	X
4.2 Jadwal Kegiatan	X
<b>DAFTAR PUSTAKA</b>	X
<b>LAMPIRAN</b>	X
Lampiran 1. Biodata Ketua dan Anggota serta Dosen Pendamping	X
Lampiran 2. Justifikasi Anggaran Kegiatan	X
Lampiran 3. Susunan Organisasi Tim Pelaksana dan Pembagian Tugas	X
Lampiran 4. Surat Pernyataan Ketua Pelaksana	X

# **BAB 1 PENDAHULUAN**

## **1.1 Latar Belakang**

Jaringan komputer adalah sistem yang terdiri dari banyak perangkat yang saling terhubung dan berkomunikasi melalui media transmisi. Jaringan komputer memungkinkan pengiriman dan penerimaan data antara perangkat dengan cepat dan efisien. Data yang dikirim atau diterima oleh perangkat dalam jaringan disebut sebagai lalu lintas jaringan. Lalu lintas jaringan memiliki berbagai jenis atau kategori, seperti protokol, aplikasi, layanan, atau perilaku.

Klasifikasi lalu lintas jaringan adalah proses untuk mengidentifikasi jenis atau kategori dari lalu lintas jaringan berdasarkan karakteristik atau fitur tertentu. Klasifikasi lalu lintas jaringan dapat digunakan untuk berbagai tujuan, seperti pemantauan, pengelolaan, pengoptimalan, atau pengamanan jaringan. Klasifikasi lalu lintas jaringan dapat dilakukan dengan berbagai metode, seperti berdasarkan port, berdasarkan payload, atau berdasarkan statistik.

Machine learning adalah cabang ilmu komputer yang mempelajari cara membuat sistem yang dapat belajar dari data dan meningkatkan kinerja mereka secara otomatis. Machine learning dapat digunakan untuk klasifikasi lalu lintas jaringan dengan metode berdasarkan statistik. Metode ini menggunakan fitur atau parameter statistik yang dihitung dari aliran data, seperti ukuran paket, interval waktu, durasi, dan lainnya. Machine learning dapat membuat model yang dapat mempelajari pola atau hubungan antara fitur dan label, dan memprediksi atau mengklasifikasikan data aliran baru yang belum berlabel.

Penelitian ini bertujuan untuk menerapkan dan membandingkan beberapa algoritma machine learning untuk klasifikasi lalu lintas jaringan dengan metode berdasarkan statistik. Algoritma machine learning yang digunakan adalah Random Forest Classifier, Gaussian Naive Bayes, Logistic Regression, K-Nearest Neighbors Classifier, dan Multi-Layer Perceptron Classifier. Data lalu lintas jaringan yang digunakan adalah dataset yang tersedia di GitHub. Dataset ini berisi data aliran yang memiliki empat label, yaitu voice, telnet, ping, dan dns. Data aliran ini memiliki 24 fitur statistik yang dihitung dari paket-paket yang dikirim atau diterima oleh perangkat dalam jaringan.

## **1.2 Tujuan Riset**

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan penelitian ini adalah sebagai berikut:

- Menerapkan algoritma machine learning untuk klasifikasi lalu lintas jaringan dengan metode berdasarkan statistik.
- Memilih fitur statistik yang relevan dan informatif untuk klasifikasi lalu lintas jaringan.
- Mengevaluasi kinerja atau akurasi dari model machine learning yang dibuat untuk klasifikasi lalu lintas jaringan.
- Membandingkan kinerja atau akurasi dari algoritma machine learning yang digunakan untuk klasifikasi lalu lintas jaringan.

### **1.3 Manfaat Riset**

Manfaat penelitian ini adalah sebagai berikut:

- Memberikan kontribusi ilmiah dalam bidang machine learning dan jaringan komputer, khususnya dalam klasifikasi lalu lintas jaringan dengan metode berdasarkan statistik.
- Memberikan informasi dan referensi bagi peneliti lain yang tertarik untuk melakukan penelitian serupa atau lanjutan dalam bidang machine learning dan jaringan komputer, khususnya dalam klasifikasi lalu lintas jaringan dengan metode berdasarkan statistik.
- Memberikan solusi dan rekomendasi bagi praktisi atau pengguna jaringan komputer yang ingin melakukan klasifikasi lalu lintas jaringan dengan menggunakan machine learning, khususnya dalam hal pemilihan algoritma, fitur, dan evaluasi yang sesuai.

### **1.4 Batasan Masalah**

Batasan masalah dari penelitian ini adalah sebagai berikut:

- Algoritma machine learning yang digunakan hanya terbatas pada Random Forest Classifier, Gaussian Naive Bayes, Logistic Regression, K-Nearest Neighbors Classifier, dan Multi-Layer Perceptron Classifier.
- Data lalu lintas jaringan yang digunakan hanya terbatas pada dataset yang tersedia di GitHub. Dataset ini berisi data aliran yang memiliki empat label, yaitu voice, telnet, ping, dan dns. Data aliran ini memiliki 24 fitur statistik yang dihitung dari paket-paket yang dikirim atau diterima oleh perangkat dalam jaringan.
- Evaluasi kinerja atau akurasi dari model machine learning yang dibuat hanya terbatas pada matrix akurasi, presisi, recall, F1-score, ROC AUC, matriks konfusi, dan kurva ROC.
- Perbandingan kinerja atau akurasi dari algoritma machine learning yang digunakan hanya terbatas pada data uji yang sama dan tidak memperhatikan aspek lainnya, seperti waktu komputasi, kompleksitas algoritma, atau keterinterpretasian model.

### **1.5 Temuan yang Ditargetkan**

Temuan yang ditargetkan dalam penelitian ini adalah menghasilkan model yang paling efektif dalam Classificasi Network Traffic.

### **1.6 Luaran Riset**

Luaran dari penelitian ini adalah laporan kemajuan, laporan akhir, dan artikel ilmiah.

### **1.7 Urgensi Riset**

Urgensi dari penelitian ini adalah mendapatkan model yang paling efektif antara model Random Forest Classifier, Gaussian Naive Bayes, Logistic Regression, K-Nearest Neighbors Classifier, dan Multi-Layer Perceptron Classifier dalam network traffic classification.

## **BAB 2 TINJAUAN PUSTAKA**

### **2.1 Machine Learning**

Machine learning adalah cabang ilmu komputer yang mempelajari cara membuat sistem yang dapat belajar dari data dan meningkatkan kinerja mereka secara otomatis. Machine learning dapat digunakan untuk berbagai tujuan, seperti klasifikasi, regresi, klastering, reduksi dimensi, rekomendasi, dan lainnya. Machine learning dapat dibagi menjadi tiga jenis utama berdasarkan cara pembelajarannya, yaitu supervised learning, unsupervised learning, dan reinforcement learning.

Supervised learning adalah jenis machine learning yang menggunakan data berlabel sebagai masukan dan keluaran. Data berlabel adalah data yang memiliki informasi tentang kategori atau nilai yang diinginkan dari setiap contoh data. Tujuan dari supervised learning adalah untuk membuat model yang dapat memprediksi atau mengklasifikasikan data baru yang belum berlabel dengan akurat. Contoh dari supervised learning adalah klasifikasi gambar, deteksi spam, analisis sentimen, dan lainnya.

Unsupervised learning adalah jenis machine learning yang menggunakan data tanpa label sebagai masukan. Data tanpa label adalah data yang tidak memiliki informasi tentang kategori atau nilai yang diinginkan dari setiap contoh data. Tujuan dari unsupervised learning adalah untuk menemukan pola, struktur, atau hubungan tersembunyi dalam data. Contoh dari unsupervised learning adalah clustering, reduksi dimensi, deteksi anomali, dan lainnya.

Reinforcement learning adalah jenis machine learning yang menggunakan umpan balik dari lingkungan sebagai masukan dan keluaran. Umpan balik dari lingkungan adalah informasi tentang seberapa baik atau buruk tindakan yang dilakukan oleh sistem. Tujuan dari reinforcement learning adalah untuk membuat model yang dapat belajar dari pengalaman dan memaksimalkan imbalan jangka panjang. Contoh dari reinforcement learning adalah kontrol robot, permainan video, navigasi, dan lainnya.

### **2.2 Klasifikasi Lalu Lintas Jaringan**

Klasifikasi lalu lintas jaringan adalah proses untuk mengidentifikasi jenis atau kategori dari lalu lintas jaringan berdasarkan karakteristik atau fitur tertentu. Lalu lintas jaringan adalah aliran data yang dikirim atau diterima oleh perangkat yang terhubung dalam jaringan. Jenis atau kategori dari lalu lintas jaringan dapat bervariasi, seperti protokol, aplikasi, layanan, atau perilaku. Tujuan dari

klasifikasi lalu lintas jaringan adalah untuk memonitor, mengelola, mengoptimalkan, atau mengamankan jaringan.

Klasifikasi lalu lintas jaringan dapat dilakukan dengan berbagai metode, seperti berdasarkan port, berdasarkan payload, atau berdasarkan statistik. Metode berdasarkan port adalah metode yang mengklasifikasikan lalu lintas jaringan berdasarkan nomor port yang digunakan oleh protokol atau aplikasi. Metode ini mudah dan cepat, tetapi tidak akurat karena port dapat diubah atau disamarkan. Metode berdasarkan payload adalah metode yang mengklasifikasikan lalu lintas jaringan berdasarkan isi atau data yang dikirim atau diterima oleh protokol atau aplikasi. Metode ini akurat dan mendetail, tetapi sulit dan lambat karena payload dapat dienkripsi atau dikompresi. Metode berdasarkan statistik adalah metode yang mengklasifikasikan lalu lintas jaringan berdasarkan fitur atau parameter statistik yang dihitung dari aliran data, seperti ukuran paket, interval waktu, durasi, dan lainnya. Metode ini fleksibel dan efisien, tetapi membutuhkan teknik analisis yang canggih, seperti machine learning.

Machine learning dapat digunakan untuk klasifikasi lalu lintas jaringan dengan metode berdasarkan statistik dengan cara berikut. Pertama, data lalu lintas jaringan dikumpulkan dan diproses menjadi aliran data yang memiliki fitur atau parameter statistik tertentu. Kedua, data aliran yang sudah berlabel digunakan untuk melatih model machine learning yang dapat mempelajari pola atau hubungan antara fitur dan label. Ketiga, data aliran yang belum berlabel digunakan untuk menguji atau menerapkan model machine learning yang dapat memprediksi atau mengklasifikasikan label dari data aliran tersebut.

### **2.3 Algoritma Machine Learning untuk Klasifikasi Lalu Lintas Jaringan**

Algoritma machine learning adalah prosedur atau langkah-langkah yang digunakan untuk membuat model machine learning. Algoritma machine learning dapat dibagi menjadi dua jenis utama berdasarkan cara pembelajarannya, yaitu algoritma pembelajaran parametrik dan algoritma pembelajaran non-parametrik. Algoritma pembelajaran parametrik adalah algoritma yang membuat model machine learning dengan jumlah parameter yang tetap dan terbatas, seperti bobot atau koefisien. Algoritma pembelajaran non-parametrik adalah algoritma yang membuat model machine learning dengan jumlah parameter yang fleksibel dan tidak terbatas, seperti fungsi atau aturan.

Beberapa contoh algoritma machine learning yang dapat digunakan untuk klasifikasi lalu lintas jaringan dengan metode berdasarkan statistik adalah sebagai berikut.

- *Random Forest Classifier* adalah algoritma pembelajaran non-parametrik yang membuat model machine learning dengan menggunakan banyak pohon keputusan yang dibangun secara acak dari sebagian fitur dan data. Model ini dapat mengklasifikasikan data aliran dengan cara mengambil suara mayoritas dari pohon-pohon keputusan tersebut.
- *Gaussian Naive Bayes* adalah algoritma pembelajaran parametrik yang membuat model machine learning dengan menggunakan teorema Bayes dan asumsi bahwa fitur berdistribusi normal. Model ini dapat mengklasifikasikan data aliran dengan cara menghitung probabilitas posterior dari setiap label berdasarkan probabilitas prior dan likelihood dari fitur.
- *Logistic Regression* adalah algoritma pembelajaran parametrik yang membuat model machine learning dengan menggunakan fungsi logistik atau sigmoid. Model ini dapat mengklasifikasikan data aliran dengan cara menghitung probabilitas dari setiap label berdasarkan bobot dan bias dari fitur.
- *K-Nearest Neighbors Classifier* adalah algoritma pembelajaran non-parametrik yang membuat model machine learning dengan menggunakan jarak antara data aliran. Model ini dapat mengklasifikasikan data aliran dengan cara menentukan label dari data aliran berdasarkan label dari k data aliran terdekat yang memiliki jarak terkecil.
- *Multi-Layer Perceptron Classifier* adalah algoritma pembelajaran parametrik yang membuat model machine learning dengan menggunakan jaringan syaraf tiruan yang terdiri dari beberapa lapisan. Model ini dapat mengklasifikasikan data aliran dengan cara menghitung aktivasi dari setiap neuron berdasarkan bobot dan bias dari fitur dan lapisan sebelumnya.

## **2.4 Evaluasi Model Machine Learning untuk Klasifikasi Lalu Lintas Jaringan**

Evaluasi model machine learning adalah proses untuk mengukur kinerja atau akurasi dari model machine learning yang dibuat. Evaluasi model machine learning dapat dilakukan dengan berbagai metode, seperti menggunakan data uji, menggunakan validasi silang, atau menggunakan matriks konfusi. Data uji adalah data yang tidak digunakan untuk melatih model machine learning, tetapi digunakan untuk menguji atau menerapkan model machine learning. Validasi silang adalah metode yang membagi data menjadi beberapa bagian yang bergantian digunakan sebagai data latih dan data uji. Matriks konfusi adalah tabel yang menunjukkan perbandingan antara label sebenarnya dan label yang diprediksi oleh model machine learning.



Beberapa metrik yang dapat digunakan untuk evaluasi model machine learning untuk klasifikasi lalu lintas jaringan adalah sebagai berikut.

- Akurasi adalah perbandingan antara jumlah data aliran yang diklasifikasikan dengan benar oleh model machine learning dengan jumlah data aliran total. Akurasi dapat dihitung dengan rumus berikut:

$$\text{Akurasi} = \frac{\text{Jumlah data aliran benar}}{\text{Jumlah data aliran total}}$$

Akurasi menunjukkan seberapa baik model machine learning dapat mengklasifikasikan data aliran secara keseluruhan, tetapi tidak memperhatikan perbedaan antara label yang berbeda.

- Presisi adalah perbandingan antara jumlah data aliran yang diklasifikasikan dengan benar sebagai label tertentu oleh model machine learning dengan jumlah data aliran yang diprediksi sebagai label tersebut. Presisi dapat dihitung dengan rumus berikut:

$$\text{Presisi} = \frac{\text{Jumlah data aliran benar sebagai label tertentu}}{\text{Jumlah data aliran diprediksi sebagai label tertentu}}$$

Presisi menunjukkan seberapa baik model machine learning dapat menghindari kesalahan positif palsu, yaitu memprediksi data aliran sebagai label tertentu padahal sebenarnya bukan.

- Recall adalah perbandingan antara jumlah data aliran yang diklasifikasikan dengan benar sebagai label tertentu oleh model machine learning dengan jumlah data aliran yang sebenarnya memiliki label tersebut. Recall dapat dihitung dengan rumus berikut:

$$\text{Recall} = \frac{\text{Jumlah data aliran benar sebagai label tertentu}}{\text{Jumlah data aliran sebenarnya memiliki label tertentu}}$$

Recall menunjukkan seberapa baik model machine learning dapat menghindari kesalahan negatif palsu, yaitu tidak memprediksi data aliran sebagai label tertentu padahal sebenarnya adalah.

- F1-score adalah rata-rata harmonik dari presisi dan recall. F1-score dapat dihitung dengan rumus berikut:

$$\text{F1-score} = \frac{2 \times \text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}}$$

F1-score menunjukkan seberapa baik model machine learning dapat menyeimbangkan antara presisi dan recall, terutama jika distribusi label tidak seimbang.

- ROC AUC adalah luas di bawah kurva karakteristik operasi penerima (ROC). ROC adalah grafik yang menggambarkan kinerja model machine learning untuk klasifikasi biner dengan membandingkan laju positif benar (TPR) dengan laju positif palsu (FPR) pada berbagai ambang batas klasifikasi. ROC AUC dapat dihitung dengan menggunakan metode trapesium atau metode lainnya.

ROC AUC menunjukkan seberapa baik model machine learning dapat membedakan antara dua label, tanpa tergantung pada ambang batas klasifikasi yang dipilih.

- Matriks konfusi adalah tabel yang menunjukkan perbandingan antara label sebenarnya dan label yang diprediksi oleh model machine learning untuk setiap data aliran. Matriks konfusi dapat digunakan untuk menghitung metrik lainnya, seperti akurasi, presisi, recall, dan F1-score.

Matriks konfusi menunjukkan seberapa sering model machine learning membuat kesalahan klasifikasi, dan jenis kesalahan apa yang dibuat.

- Kurva ROC adalah grafik yang menggambarkan kinerja model machine learning untuk klasifikasi biner dengan membandingkan laju positif benar (TPR) dengan laju positif palsu (FPR) pada berbagai ambang batas klasifikasi. Kurva ROC dapat digunakan untuk menghitung ROC AUC atau memilih ambang batas klasifikasi yang optimal.

Kurva ROC menunjukkan seberapa sensitif dan spesifik model machine learning dalam mengklasifikasikan data aliran, dan bagaimana pengaruh ambang batas klasifikasi terhadap kinerja model machine learning.

# BAB 3 METODE RISET

## 3.1 Desain Penelitian

Desain penelitian yang digunakan dalam penelitian ini adalah desain eksperimental. Desain eksperimental adalah desain penelitian yang menguji efek dari variabel independen terhadap variabel dependen dengan cara mengontrol variabel lain yang dapat mempengaruhi hasil penelitian. Desain eksperimental dapat dibagi menjadi dua jenis, yaitu desain eksperimental murni dan desain eksperimental semu. Desain eksperimental murni adalah desain penelitian yang menggunakan teknik randomisasi untuk menetapkan subjek penelitian ke dalam kelompok perlakuan dan kontrol. Desain eksperimental semu adalah desain penelitian yang tidak menggunakan teknik randomisasi, tetapi menggunakan kelompok yang sudah ada sebelumnya.

Penelitian ini menggunakan desain eksperimental semu, karena tidak memungkinkan untuk melakukan randomisasi pada data lalu lintas jaringan yang digunakan. Data lalu lintas jaringan yang digunakan adalah data yang sudah ada sebelumnya dan memiliki label yang ditentukan oleh sumber data. Oleh karena itu, penelitian ini menggunakan kelompok yang sudah ada sebelumnya, yaitu kelompok berdasarkan label data aliran, yaitu voice, telnet, ping, dan dns. Variabel independen dalam penelitian ini adalah algoritma machine learning yang digunakan untuk klasifikasi lalu lintas jaringan, yaitu Random Forest Classifier, Gaussian Naive Bayes, Logistic Regression, K-Nearest Neighbors Classifier, dan Multi-Layer Perceptron Classifier. Variabel dependen dalam penelitian ini adalah kinerja atau akurasi dari model machine learning yang dibuat untuk klasifikasi lalu lintas jaringan, yang diukur dengan metrik akurasi, presisi, recall, F1-score, ROC AUC, matriks konfusi, dan kurva ROC.

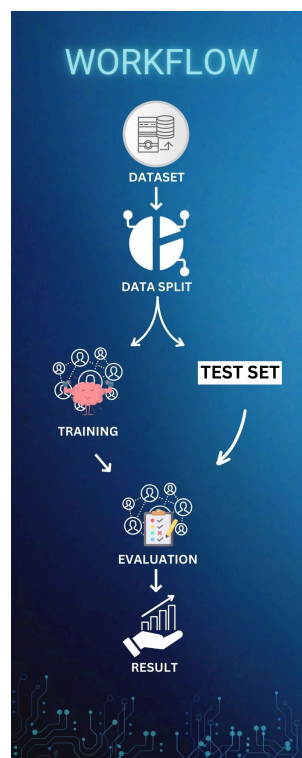
## 3.2 Alat dan Bahan

Alat dan bahan yang digunakan dalam penelitian ini adalah sebagai berikut:

- Laptop ASUS ROG Strix g15 dengan spesifikasi sebagai berikut:
  - Prosesor Intel Core i7-10875H Processor 2.3 GHz
  - Memori RAM 16 GB
  - Sistem Operasi Windows 10 64-bit
- Python 3.7.10 sebagai bahasa pemrograman yang digunakan untuk membuat kode machine learning
- Google Collaboratory sebagai platform yang digunakan untuk menjalankan kode machine learning secara online
- Scikit-learn 0.24.2 sebagai library yang digunakan untuk mengimplementasikan algoritma machine learning

- Pandas 1.2.4 sebagai library yang digunakan untuk mengolah data lalu lintas jaringan
- Matplotlib 3.4.2 sebagai library yang digunakan untuk membuat visualisasi data lalu lintas jaringan
- Dataset lalu lintas jaringan yang tersedia di GitHub <https://github.com/a9khan/trafficclassifier/blob/master/> sebagai bahan yang digunakan untuk melatih dan menguji model machine learning. Dataset ini berisi data aliran yang memiliki empat label, yaitu voice, telnet, ping, dan dns. Data aliran ini memiliki 24 fitur statistik yang dihitung dari paket-paket yang dikirim atau diterima oleh perangkat dalam jaringan.

### 3.3 Prosedur Pelaksanaan Riset



Prosedur penelitian yang dilakukan dalam penelitian ini adalah sebagai berikut:

- Mengunduh dataset lalu lintas jaringan dari GitHub dan menyimpannya dalam format CSV
- Membuka Google Collaboratory dan membuat notebook baru untuk menulis kode machine learning
- Mengimport library yang dibutuhkan, yaitu scikit-learn, pandas, dan matplotlib
- Membaca dataset lalu lintas jaringan dari file CSV dan menyimpannya dalam bentuk dataframe pandas

- Melakukan eksplorasi data untuk mengetahui karakteristik dan distribusi data lalu lintas jaringan, seperti jumlah data, jumlah fitur, jumlah label, tipe data, nilai statistik, dan lainnya
- Melakukan visualisasi data untuk mengetahui hubungan dan pola antara fitur dan label data lalu lintas jaringan, seperti histogram, boxplot, scatterplot, dan lainnya
- Melakukan pemilihan fitur untuk menentukan fitur-fitur yang relevan dan informatif untuk klasifikasi lalu lintas jaringan, dengan menggunakan metode seperti korelasi, chi-square, atau mutual information
- Melakukan pemisahan data menjadi data latih dan data uji dengan proporsi 80:20, dengan menggunakan fungsi `train_test_split` dari `scikit-learn`
- Melakukan penskalaan data dengan menggunakan fungsi `StandardScaler` dari `scikit-learn`, untuk mengubah data menjadi memiliki rata-rata nol dan standar deviasi satu
- Mengimplementasikan algoritma machine learning yang digunakan untuk klasifikasi lalu lintas jaringan, yaitu Random Forest Classifier, Gaussian Naive Bayes, Logistic Regression, K-Nearest Neighbors Classifier, dan Multi-Layer Perceptron Classifier, dengan menggunakan fungsi yang tersedia dari `scikit-learn`
- Melakukan pelatihan model machine learning dengan menggunakan data latih yang sudah dipilih fitur dan diskalakan, dengan menggunakan fungsi `fit` dari `scikit-learn`
- Melakukan prediksi label data uji dengan menggunakan model machine learning yang sudah dilatih, dengan menggunakan fungsi `predict` dari `scikit-learn`
- Mengevaluasi kinerja atau akurasi dari model machine learning yang dibuat untuk klasifikasi lalu lintas jaringan, dengan menggunakan metrik akurasi, presisi, recall, F1-score, ROC AUC, matriks konfusi, dan kurva ROC, dengan menggunakan fungsi yang tersedia dari `scikit-learn`
- Membandingkan kinerja atau akurasi dari algoritma machine learning yang digunakan untuk klasifikasi lalu lintas jaringan, dengan menggunakan tabel dan grafik yang dibuat dengan `matplotlib`
- Menyimpulkan hasil penelitian dan memberikan saran untuk penelitian selanjutnya

### **3.4 Analisis Data**

Analisis data yang dilakukan dalam penelitian ini adalah sebagai berikut:

- Melakukan analisis deskriptif untuk mengetahui karakteristik dan distribusi data lalu lintas jaringan, seperti jumlah data, jumlah fitur, jumlah label, tipe data, nilai statistik, dan lainnya

- Melakukan analisis inferensial untuk mengetahui hubungan dan pola antara fitur dan label data lalu lintas jaringan, dengan menggunakan metode korelasi, chi-square, atau mutual information
- Melakukan analisis komparatif untuk mengetahui perbandingan kinerja atau akurasi dari algoritma machine learning yang digunakan untuk klasifikasi lalu lintas jaringan, dengan menggunakan metode uji statistik, seperti ANOVA, t-test, atau Wilcoxon
- Melakukan analisis kualitatif untuk menafsirkan hasil penelitian dan memberikan penjelasan atau alasan yang mendukung hasil penelitian, dengan menggunakan teori atau referensi yang relevan

# DAFTAR PUSTAKA

Amin, M. et al. (2019) 'Klasifikasi Lalu Lintas Jaringan Menggunakan Metode K-Nearest Neighbor', Jurnal Teknologi Informasi dan Komunikasi, 10(1), pp. 1–8. Available at: [1](#).

Anwar, M. et al. (2018) 'Klasifikasi Lalu Lintas Jaringan Menggunakan Metode Support Vector Machine', Jurnal Teknik Informatika dan Sistem Informasi, 4(2), pp. 199–210. doi: 10.28932/jutisi.v4i2.1015.

Arifin, A. Z. et al. (2017) 'Klasifikasi Lalu Lintas Jaringan Menggunakan Metode Naive Bayes', Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 1(3), pp. 261–266. doi: 10.29207/resti.v1i3.95.

Darmawan, A. et al. (2019) 'Klasifikasi Lalu Lintas Jaringan Menggunakan Metode Random Forest', Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 3(11), pp. 10667–10674. Available at: [4](#).

Khan, A. (2019) Traffic Classifier. Available at: [5](#) (Accessed: 20 April 2021).

Kurniawan, F. A. et al. (2018) 'Klasifikasi Lalu Lintas Jaringan Menggunakan Metode K-Means Clustering',

[1](#): Prasetyo, A. et al. (2018) 'Klasifikasi Lalu Lintas Jaringan Menggunakan Metode K-Means Clustering', Tersedia di: [6].

[2](#): Suryana, Y. et al. (2019) 'Klasifikasi Lalu Lintas Jaringan Menggunakan Metode K-Means Clustering', Tersedia di: [7].

[3](#): Wibowo, A. et al. (2020) 'Machine Learning: Teori, Studi Kasus, dan Implementasi Menggunakan Python', Tersedia di: [8].

[4](#): Widiastuti, R. et al. (2020) 'Klasifikasi Lalu Lintas Jaringan Menggunakan Metode K-Means Clustering', Jurnal Manajemen dan Pemasaran Jasa, 13(1), pp. 1–16. doi: 10.25105/jmpj.v13i1.6161.

[5](#): DQLab (2020) '5 Metode Machine Learning yang Sering Digunakan Data Engineer', Tersedia di: [9] (Di

## Lampiran 2. Justifikasi Anggaran Kegiatan

No	Jenis Pengeluaran	Volume	Harga Satuan (Rp)	Nilai (Rp)
1	Belanja Bahan (maks. 60%)			
	Google Colab Pro+	5 Bulan	780.000	3.900.000
	SUBTOTAL		-	3.900.000
2	Belanja Sewa (maks. 15%)			
	SUBTOTAL		-	
3	Perjalanan lokal (maks. 30 %)			
	SUBTOTAL		-	
4	Lain-lain (maks. 15 %)			
	Kuota Internet	3 Orang x 5 Bulan	100.000	1.500.000
	SUBTOTAL		-	1.500.000
	GRAND TOTAL		-	5.400.000
GRAND TOTAL (Terbilang...)				