



Deteksi Anomali Traffic Pada Jaringan Komputer Menggunakan Naive Bayes, Decision Tree Dan Isolation Forest

Milka Justine Mendrofa^{1*}, Kristian Juri Damai Lase², Haeni Budiati³

Universitas Kristen Immanuel, Indonesia

Email: milka.justine.m@mail.ukrim.ac.id^{1*}, kristian@ukrim.ac.id², heni@ukrimuniversity.ac.id³

*Correspondence

ABSTRAK

Anomali jaringan komputer adalah pola atau aktivitas yang tidak biasa yang menyimpang dari perilaku normal jaringan. Deteksi anomali sangat penting untuk mengidentifikasi ancaman keamanan, gangguan operasional, atau kegagalan sistem yang mungkin terjadi. Tujuan penelitian ini adalah untuk mengembangkan dan mengevaluasi tiga metode utama untuk mendeteksi anomali dalam trafik jaringan komputer: Naive Bayes, Decision Tree, dan Isolation Forest. Berbasis pada Teorema Bayes, pendekatan naive Bayes memprediksi kemungkinan suatu kejadian berdasarkan data historis. Model berbasis pohon keputusan, decision tree membagi data secara iteratif berdasarkan karakteristik tertentu untuk mengklasifikasikan atau memprediksi hasil. Algoritma Isolasi Hutan adalah algoritma berbasis kelompok yang dimaksudkan untuk mendeteksi anomali dengan cepat dengan mengisolasi data anomali. Fokus utama penelitian ini adalah membandingkan kinerja ketiga metode tersebut dalam mendeteksi anomali trafik jaringan, termasuk kemampuan masing-masing metode untuk menemukan pola tidak normal secara akurat dan efisien. Tujuan dari penelitian ini adalah untuk memberikan pemahaman tentang metode yang paling efektif dalam hal deteksi anomali jaringan, sehingga dapat membantu membangun sistem keamanan jaringan yang lebih andal dan responsif terhadap ancaman.

Kata Kunci: anomali, metode naive bayes, decision tree, isolation forest, jaringan komputer.

ABSTRACT

Computer network anomalies refer to unusual patterns or activities that deviate from normal network behavior. Anomaly detection is crucial for identifying security threats, operational disruptions, or potential system failures. This study aims to develop and evaluate three primary methods for detecting anomalies in computer network traffic: Naive Bayes, Decision Tree, and Isolation Forest. Based on Bayes' Theorem, the Naive Bayes approach predicts the likelihood of an event occurring based on historical data. The Decision Tree model iteratively splits data based on specific characteristics to classify or predict outcomes. The Isolation Forest algorithm is a group-based approach designed to detect anomalies quickly by isolating anomalous data points. The main focus of this study is to compare the performance of these three methods in detecting network traffic anomalies, including their ability to accurately and efficiently identify abnormal patterns. The goal of this research is to provide insights into the most effective method for network anomaly detection, ultimately contributing to the development of a more reliable and responsive network security system.

Keywords: anomalies, naive bayes method, decision tree, isolation forest, computer networks.

PENDAHULUAN

Dalam era digital ini, Jaringan komputer sangat penting untuk mempertahankan kemandirian data bagi organisasi dan individu, terutama dengan berkembangnya internet dan teknologi (Prasetyo et al.,

2018). Internet dan teknologi terhubung untuk menjamin ketersediaan layanan di jaringan komputer, individu harus mengatasi ancaman dalam lalu lintas jaringan (Suwaryo et al., 2014). Lalu lintas jaringan adalah istilah yang mengacu pada data yang ada di jaringan saat jaringan beroperasi, dengan data yang ditukarkan sebagai cara untuk mendapatkan akses ke sumber daya layanan permintaan dan respons dari penyedia layanan (Pertiwi et al., 2023).

Salah satu cara untuk menemukan perilaku yang mencurigakan yang dapat menunjukkan serangan siber atau pelanggaran data adalah dengan mendeteksi anomali dalam lalu lintas jaringan (Ness et al., 2025);(Triana, 2024);(Beno et al., 2022). Oleh karena itu, sangat penting untuk mengembangkan teknik yang efisien untuk mendeteksi anomali ini. Dalam mendeteksi anomali ada tiga algoritma yang sering di gunakan ialah Naive Bayes, Decision Tree dan Isolation Forest.

Teorema Bayes, metode klasifikasi Naive Bayes menggunakan metode probabilitas dan statistik. Metode ini diciptakan oleh ilmuwan Inggris Thomas Bayes dan bertujuan untuk memprediksi peluang di masa depan berdasarkan pengalaman masa lalu. Metode ini dikenal sebagai teorema Bayes. Salah satu ciri utama klasifikator dari Naive Bayes ini adalah keyakinan yang sangat kuat (naif) bahwa setiap kondisi atau kejadian independen. Naive Bayes menjelaskan setiap kelas keputusan, menghitung probabilitas dengan asumsi bahwa kelas keputusan adalah benar, mengingat vektor informasi objek, menurut Olson Delen (2008). Algoritma ini menganggap atribut objek independen. Jumlah frekuensi dari "master" tabel keputusan adalah jumlah kemungkinan yang terlibat dalam membuat perkiraan akhir (Oleh et al., 2022);(Bahtiar et al., 2020);(Anwar, 2022).

Salah satu pendekatan yang paling mudah ditafsirkan oleh manusia adalah pohon keputusan (Decision Tree). Model prediksi yang menggunakan struktur pohon atau berhirarki disebut pohon keputusan (Decision Tree). Mengubah data menjadi pohon keputusan (Decision Tree) dan aturan keputusan adalah inti dari konsep pohon keputusan (Decision Tree). Manfaat utamanya adalah memecah proses pengambilan keputusan yang kompleks menjadi lebih sederhana, yang memungkinkan pengambil keputusan untuk lebih mudah menyelesaikan masalah yang ada. Kemampuan untuk menangani data yang besar dan kompleks dengan waktu komputasi yang relatif cepat adalah keuntungan utama dari metode ini. Untuk menyelesaikan masalah ini, penelitian ini mengembangkan dan mengevaluasi model model pohon keputusan yang efisien yang dapat mendeteksi anomali lalu lintas jaringan dan meningkatkan keamanan dan kinerja jaringan komputer (Manalu et al., 2024);(Pertiwi et al., 2023).

Algoritma Isolation Forest "mengisolasi" pengamatan dengan memilih fitur secara acak dan kemudian memilih nilai maksimum dan minimum fitur untuk mengembalikan nilai anomali dari setiap sampel. Didasarkan pada prinsip isolasi, di mana anomali diidentifikasi sebagai titik data yang lebih mudah diisolasi dibandingkan dengan data normal, Isolation Forest adalah salah satu algoritma terbaru yang menunjukkan potensi besar dalam deteksi anomali. Isolation Forest bekerja dengan membangun banyak pohon isolasi dan menggunakan sifat pohon-pohon ini untuk mengukur skor anomali. Kemampuan algoritma ini untuk menyesuaikan diri dengan berbagai jenis data dan sangat efektif dalam menangani kumpulan data yang besar (Zulfikar et al., 2023);(Ness et al., 2025);(Triana, 2024);(Sri Lakshmi et al., 2023).

Penelitian ini juga akan membahas masalah yang dihadapi saat menggunakan algoritma-algoritma tersebut, seperti masalah data yang tidak seimbang dan pentingnya preprocessing yang tepat. Dengan memahami kekuatan dan kelemahan setiap algoritma, diharapkan dapat ditemukan strategi terbaik untuk digunakan dalam sistem keamanan jaringan.

Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan untuk mengembangkan metode deteksi anomali yang lebih efisien dan efektif. Oleh karena itu, organisasi dapat lebih siap menghadapi ancaman siber yang terus muncul, melindungi data dan informasi penting, dan menjaga kepercayaan pengguna terhadap sistem yang mereka gunakan.

Studi "Deteksi Anomali Menggunakan Hutan Isolasi untuk Belanja Barang Persediaan Konsumsi pada Satuan Kerja Kepolisian Republik Indonesia" oleh Ahmad Zulfikar, Farhan Ariq Rahmani, dan Nurul Azizah menunjukkan bahwa model Hutan Isolasi dapat mendeteksi kesalahan dalam data transaksi akun belanja persediaan (521811). Hasil uji coba yang dilakukan terhadap berbagai nilai pencemaran parameter menunjukkan bahwa nilai 0,3% adalah yang terbaik untuk menemukan kesalahan. Meskipun tidak ada bukti langsung yang menunjukkan penipuan, informasi penelitian ini dapat digunakan untuk menambahkan fitur ke dalam sistem otorisasi pengeluaran satuan kerja sebagai sinyal awal (Zulfikar et al., 2023). Ini akan mirip dengan cara sistem keuangan SAKTI digunakan. Selain itu, penelitian yang ditulis oleh Niko Suwaryo, Ismasari Nawangsih, dan Sri Rejeki dengan judul "Deteksi Serangan pada Sistem Deteksi Intrusion (IDS) untuk Klasifikasi Serangan dengan Algoritma Naive Bayes, C.45, dan K-NN dalam Meminimalisasi Risiko terhadap Pengguna" menemukan bahwa algoritma KNN, Naive Bayes, dan C.45 memiliki tingkat recall, ketepatan, dan akurasi yang tinggi. Algoritma C.45 memiliki akurasi 97,80% (Suwaryo et al., 2014).

Dalam penelitiannya, "Implementasi Algoritma Decision Tree untuk Klasifikasi Produk Laris", Asmaul Husnah Nasrullah menunjukkan bahwa model klasifikasi Algoritma Decision Tree C4.5 berhasil melakukan klasifikasi produk laris dengan akurasi 90% dan nilai AUC 0.709. Dengan nilai ini, model ini masuk dalam kategori yang cukup baik untuk menilai pola penjualan dan stok produk UD (Nasrullah, 2021). Cipta Karya Gorontalo. Selain itu, penelitian yang dilakukan oleh Muhammad Iqbal Manalu dan Fatur Padilla Hutabarat dari Universitas Sumatera Utara berjudul "Deteksi Anomaly Lalu Lintas Jaringan dengan Metode Decision Tree" menemukan bahwa algoritma Decision Tree memiliki akurasi di atas 90% dalam mendeteksi anomali lalu lintas jaringan. Karena akurasi ini, algoritma ini dapat digunakan untuk mengembangkan alat atau aplikasi yang dapat melacak lalu lintas jaringan serta mencari tahu apa yang terjadi di dalamnya (Manalu et al., 2024).

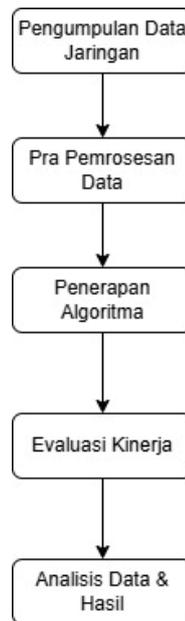
Studi yang berjudul "Analisis Perbandingan Deteksi Anomaly Lalu Lintas dengan Metode Naive Bayes dan DBSCAN" dilakukan pada tahun 2020 di Universitas Pamulang oleh Rino Bahtiar, Moch. Dwi Sakti T, Aris Setiawan, dan Perani Rosyani. Studi ini membandingkan tingkat akurasi kedua metode untuk mendeteksi anomali lalu lintas jaringan. Hasil penelitian menunjukkan bahwa berdasarkan distribusi sampel data grafik dan Radviz, Naive Bayes memiliki nilai probabilitas antara 0,1 hingga 0,8. Sementara itu, metode DBSCAN menunjukkan hasil yang lebih berbeda, dengan tingkat akurasi pada 30.000 data pertama hampir mencapai 1,5%, dengan TP sebesar 15.365 dan FN (nilai negatif negatif) sebesar 14.635.

METODE

Dalam penelitian ini menerapkan pendekatan kuantitatif melalui desain eksperimen, dan melakukan deteksi serangan atau anomali trafik pada trafik jaringan dan mencari teknik yang efektif untuk menemukan anomali trafik. Studi ini menggunakan algoritma Naive Bayes, Decision Tree, dan Isolation Forest untuk menguji anomali trafik pada jaringan komputer.

Tahap Penelitian

Penelitian ini dilakukan melalui berbagai tahap untuk menemukan anomali. Berikut tahap-tahap yang di lakukan:



Gambar 1
Flowchart Prosesi deteksi anomali

HASIL DAN PEMBAHASAN

Pertumbuhan ekonomi adalah indikator yang biasanya digunakan untuk menjadi tolak ukur sejauh mana kegiatan ekonomi suatu negara membawa pendapatan tambahan bagi masyarakat selama periode waktu tertentu. Pertumbuhan ekonomi adalah perkembangan kegiatan ekonomi yang menghasilkan peningkatan jumlah produk dan jasa yang dihasilkan dan peningkatan kekayaan.

Investasi mengacu pada kegiatan di mana investasi dilakukan dalam berbagai kegiatan ekonomi (produksi) dengan harapan menerima manfaat (keuntungan) di masa depan. Dalam pekerjaan ini, investasi dibagi menjadi dua bagian, yaitu perkembangan realisasi penanaman modal dalam negeri (PMDN) dan PMA (penanaman modal asing) yang dinyatakan dalam persentase. Untuk mengetahui besarnya Investasi tahun 2010-2020 dapat dilihat pada tabel berikut:

Tabel 1
Jumlah Investasi Di Indonesia Tahun 2010-2021

Tahun	proyek	Investasi						Jumlah Investasi	%	Data Log Investasi
		PMDN (Miliar Rupiah)	%	proyek	PMA (Miliar Rupiah)	%				

2010	875	60.626,30		3076	145787,27		206.413,57		5,31
2011	1313	76.000,70	25,36	4342	176.594,77	21,13	252.595,47	22,37	5,40
2012	1210	92.182,00	21,25	4579	237.540,65	34,51	329.722,65	30,53	5,52
2013	2129	128.150,60	39,02	9612	348.818,71	46,85	476.969,31	44,66	5,68
2014	1652	156.126,30	21,83	8885	354.909,47	1,75	511.035,77	7,14	5,71
2015	5100	179.465,90	14,95	17738	403.861,00	13,79	583.326,90	14,15	5,77
2016	7511	216.230,80	20,49	25321	389.161,65	-3,64	605.392,45	3,78	5,78
2017	8838	262.350,50	21,33	26257	436.784,81	12,24	699.135,31	15,48	5,84
2018	10815	328.604,90	25,25	21972	424.407,70	-2,83	753.012,60	7,71	5,88
2019	30451	386.498,40	17,62	30345	392.130,53	-7,61	778.628,93	3,40	5,89
2020	96623	413.535,50	7,00	56726	404.338,16	3,11	817.873,66	5,04	5,91
2021	106002	447.063,60	8,11	27271	443.667,45	9,37	890.731,05	8,91	5,95

Sumber: BPS Indonesia, Bank Indonesia, BKPM, data diolah, 2021

Berdasarkan tabel 1 di atas menunjukkan bahwa Investasi di Indonesia dari tahun 2010-2021 mengalami fluktuatif dimana pada tahun 2010 Investasi di Indonesia sebesar 206.413,57, sedangkan pada tahun 2011-2021 Investasi di Indonesia terus mengalami kenaikan sebesar 890.731,05. Hal ini terjadi karena jumlah penanaman Modal Asing dan Penanaman Modal Dalam Negeri mengalami peningkatan sehingga dapat meningkatnya pertumbuhan ekonomi dan mendorong peningkatan investasi yang masuk ke Indonesia (Astuti, 2018).

Tenaga Kerja (X2)

Tenaga Kerja adalah dalah penduduk berusia kerja (15 tahun ke atas), yang sedang bekerja atau sedang aktif mencari pekerjaan, yang masih mau dan mampu bekerja. Tenaga kerja merupakan faktor produksi yang sangat penting bagi suatu negara, selain faktor alam dan faktor modal. Meskipun negara memiliki sumber daya alam dan modal yang besar, namun tetap membutuhkan tenaga kerja sebagai faktor produksi. Tenaga kerja, modal dan sumber daya alam merupakan faktor produksi yang tidak hanya berperan penting dalam meningkatkan volume produksi, tetapi juga dapat memberikan kontribusi terhadap pertumbuhan pendapatan nasional. Untuk mengetahui besarnya Tenaga Kerja Di Indonesia tahun 2010-2021 dapat dilihat pada tabel berikut:

Tabel 2
Jumlah Tenaga Kerja Di Indonesia tahun 2011-2020
Tenaga Kerja (Ribu Jiwa)

Tahun	Angkatan Kerja Berkerja	Jumlah Pengangguran	Angkatan Kerja	Bukan Angkatan Kerja	Penduduk Berumur 15 Tahun Keatas	Jumlah Penduduk	Data Log Tenaga Kerja
2010	108.207.767	8.319.779	116.527.546	55.542.793	172.070.339	237.641	8,03
2011	107.416.309	8.681.392	116.097.701	57.754.016	173.851.727	241.991	8,03
2012	112.504.868	7.344.866	119.849.734	57.024.098	176.873.832	245.425	8,05
2013	112.761.072	7.410.913	120.171.985	59.795.358	179.967.361	248.818	8,05
2014	114.628.026	7.244.905	121.872.931	61.119.273	182.992.204	252.165	8,06
2015	114.819.199	7.560.822	122.380.021	63.720.896	186.100.917	255.462	8,06
2016	118.411.973	7.031.755	125.443.728	63.652.974	189.096.722	258.705	8,07
2017	121.022.423	7.040.323	128.062.746	64.016.670	192.079.416	261.891	8,08
2018	126.282.186	7.073.385	133.355.571	64.770.982	198.126.533	265.015	8,10
2019	128.755.271	7.104.424	135.859.695	65.325.319	201.185.014	266.912	8,11
2020	128.454.184	9.767.754	138.221.938	65.750.522	203.972.460	270.203	8,11
2021	131.050.523	9.102.052	140.152.575	66.555.724	206.708.299	272.682	8,12

Sumber: BPS Indonesia, Bank Indonesia, BKPM, data diolah, 2021

Berdasarkan tabel 2 menunjukkan bahwa Tenaga Kerja di Indonesia dari tahun ke tahun selalu berfluktuatif, hal ini dapat dilihat dari tahun awal penelitian Pada tahun 2010 sebesar 108.207.767. Pada tahun 2011-2020 mengalami peningkatan sebesar 128.454.184. Demikian pada tahun 2021 mengalami peningkatan sebesar 131.050.523. Jumlah tenaga kerja di Indonesia selama periode 2010-2021 berfluktuatif karena terjadinya dari seluruh penduduk berusia kerja (15-64) di suatu wilayah yang siap dan mau bekerja. Sementara yang lainnya lebih memilih untuk melakukan aktifitas lain seperti sekolah, mengurus rumah tangga, dan lainnya. Tinggi rendahnya angka TPAK yang dihasilkan tergantung pada pentingnya pilihan kegiatan penduduk yang bekerja, termasuk pekerjaan, sekolah, rumah tangga atau kegiatan lainnya.

Tingkat Inflasi (X3)

Inflasi adalah kecenderungan naiknya harga barang dan jasa pada secara umum dan terus menerus. Kenaikan harga dari satu atau dua barang saja tidak disebut inflasi. Kecuali jika kenaikan itu meluas ke sebagian besar harga barang-barang lain. Kondisi ledakan terus menerus juga harus diperhatikan. Ketika harga barang dan jasa naik di negara itu, inflasi meningkat. Kenaikan harga barang dan jasa menyebabkan penurunan nilai uang. Karena kenaikan harga barang-barang ini, seperti

musiman, menjelang hari-hari besar, atau yang terjadi sekali saja. Untuk mengetahui besarnya Inflasi Di Indonesia tahun 2010-2021 dapat dilihat pada tabel berikut:

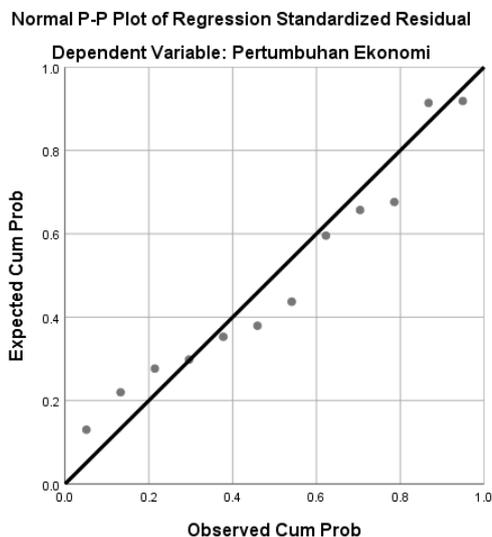
Tabel 3
Jumlah Tingkat Inflasi Di Indonesia tahun 2010-2021

Tahun	Tingkat Inflasi	Indeks Harga Konsumen	Data Log Tingkat
	Persentase (%)	Jumlah	Inflasi
2010	6,96	125,17	0,84
2011	3,79	129,91	0,58
2012	4,30	135,49	0,63
2013	8,38	146,84	0,92
2014	8,36	119,00	0,92
2015	3,35	122,99	0,53
2016	3,02	126,71	0,48
2017	3,61	131,28	0,56
2018	3,13	135,39	0,50
2019	2,72	139,07	0,43
2020	1,68	105,68	0,23
2021	1,87	107,66	0,27

Sumber: BPS Indonesia, Bank Indonesia, BKPM, data diolah, 2021

Berdasarkan tabel 3 menunjukkan bahwa tingkat inflasi dan indeks harga konsumen di Indonesia cenderung mengalami fluktuasi sejak tahun 2010-2021. Adanya kenaikan inflasi tahun 2010-2014 cukup berpengaruh terhadap kondisi perekonomian di Indonesia. Selain itu, adanya penurunan nilai mata uang juga memicu kenaikan inflasi. Pada tahun 2010-2013 mengalami kenaikan indeks harga konsumen sebesar 146.84. Disebabkan adanya barang dan jasa yang dikonsumsi oleh rumah tangga melebihi kapasitas yang diproduksi sehingga harga yang ditentukan meningkat maka terjadi kenaikan inflasi. Pada tahun 2014 mengalami Penurunan indeks harga konsumen sebesar 119.00. Disebabkan oleh penurunan terhadap harga barang dan jasa yang dapat dikonsumsi oleh rumah tangga hal ini terjadi karena inflasi. Pada tahun 2015-2019 mengalami Kenaikan indeks harga konsumen sebesar 139.07. Pada tahun 2020 mengalami Kenaikan indeks harga konsumen sebesar 105.68. Disebabkan oleh penurunan terhadap harga barang dan jasa yang dapat dikonsumsi oleh rumah tangga hal ini terjadi karena inflasi cenderung mengalami penurunan sejak 2015 hingga 2021 yang merupakan dampak dari berbagai kebijakan yang dilakukan pemerintah Indonesia. tingkat inflasi tertinggi terjadi pada tahun 2014 sebesar 8.36% dan tingkat inflasi terendah pada tahun 2020 sebesar 1,68%. Sedangkan Pada tahun 2021 mengalami kenaikan indeks harga konsumen sebesar 107.66.

Uji normalitas dengan grafik normal P-Plot membentuk garis lurus diagonal, setelah itu data di plotting dan dibandingkan dengan garis diagonal. Jika distribusinya normal, garis yang menggambarkan data sebenarnya mengikuti garis diagonal, seperti yang ditunjukkan di bawah ini:



Gambar 1
Grafik Normal P-Plot

Gambar 1 terlihat bahwa grafik Normal Probability Plot, menunjukkan bahwa data menyebar di sekitar garis diagonal dan mengikuti arah garis diagonal dan menunjukkan pola distribusi normal, sehingga dapat disimpulkan bahwa asumsi normalitas terpenuhi dan dapat digunakan untuk memprediksi Pertumbuhan Ekonomi di Indonesia berdasarkan variabel bebasnya. Untuk lebih jelasnya harus dilakukan dengan pengujian *kolmogorov-smirnov*. Hasil uji normalitas *kolmogorov-smirnov* dapat dilihat pada tabel berikut :

Tabel 4
Hasil Uji Kolmogorov-Smirnov
One-Sample Kolmogorov-Smirnov Test

	Unstandardized Residual
N	12
Mean	0,0000000

Normal Parameters ^{a,b}	Std. Deviation	0,00438673
Most Extreme Differences	Absolute	0,162
	Positive	0,162
	Negative	-0,123
Test Statistic		0,162
Asymp. Sig. (2-tailed)		.200 ^{c,d}

Sumber: Output SPSS 25 (data diolah 2022)

Berdasarkan tabel 4 diatas menunjukkan bahwa data tersebut berdistribusi normal. Pengujian normalitas data dengan menggunakan uji kolmogorov-smirnov ini dengan kriteria jika masing-masing variabel $>0,05$ maka dari hasil uji diatas dapat dilihat nilai signifikasinya $0,200$ atau lebih besar dari $0,05$ maka data tersebut dinyatakan berdistribusi normal atau layak digunakan dalam penelitian.

Uji Multikolinieritas

Tujuan dari pengujian ini adalah untuk menguji apakah ada korelasi antara variabel-variabel independen dalam suatu model regresi. Berdasarkan aturan *variance inflation factor* (VIF) dan *tolerance*, dimana data tersebut dinyatakan tidak Multikolenieritas apabila tolerancinya lebih besar dari 0.1 dan VIFnya kurang dari 10 . Adapun hasil uji multikolinieritas dapat dilihat tabel 5 berikut:

Tabel 5
Uji Multikolenieritas

Model	Collinearity Statistics	
	Tolerance	VIF
1 (Constant)		
Investasi	0,110	9,085
Tenaga Kerja	0,057	17,500
Inflasi	0,387	2,582
Covid 19	0,251	3,992

Sumber: Output SPSS 23 (data diolah 2022)

Pada tabel 5 di atas dapat dijelaskan bahwa tidak terjadi gejala multikolenieritas antara masing-masing variabel independen dalam model regresi yaitu dengan melihat VIF dan nilai tolerance. Berdasarkan tabel di atas maka dapat diketahui nilai VIF untuk Investasi sebesar $9,085 < 10$ dan nilai tolerance sebesar $0,110 > 10$ sehingga Investasi dinyatakan tidak terjadi multikolenieritas. Sedangkan nilai VIF untuk Tenaga Kerja sebesar $17,500 < 10$ dan nilai tolerance sebesar $0,057 > 10$ sehingga Tenaga Kerja dinyatakan tidak terjadi multikolenieritas. Nilai VIF untuk Inflasi sebesar $2,582 < 10$ dan nilai tolerance sebesar $0,387 > 10$ sehingga Inflasi dinyatakan tidak terjadi multikolenieritas. Sedangkan Nilai VIF untuk covid 19 sebesar $3,992 < 10$ dan nilai tolerance sebesar $0,251 > 10$ sehingga Inflasi dinyatakan tidak terjadi multikolenieritas.

Uji Autokorelasi

Salah satu metode analisis untuk mendeteksi ada tidaknya autokorelasi adalah uji nilai Durbin-Watson (DW test). Jika nilai DW lebih besar dari batas atas (du) dan lebih kecil dari jumlah variabel

bebas maka dapat disimpulkan tidak terjadi autokorelasi. Adapun hasil uji autokorelasi dapat dilihat pada tabel 6 berikut:

Tabel 6
Uji Autokorelasi
Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.998 ^a	0,996	0,994	0,00550	2,173

Sumber: Output SPSS 25 (data diolah 2022)

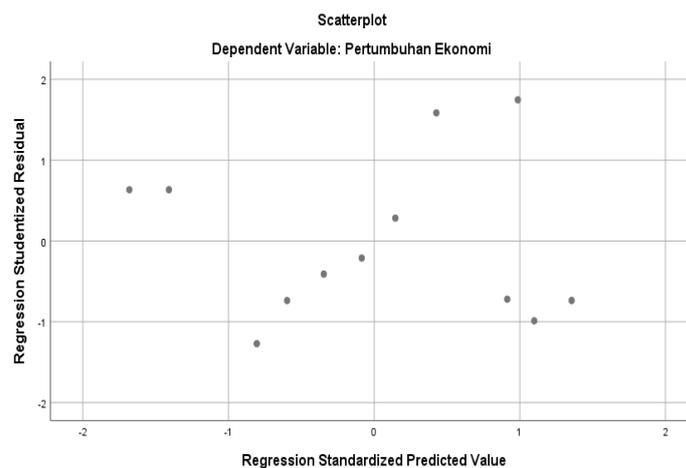
Berdasarkan tabel 6 diatas menunjukkan bahwa Dari hasil analisis dengan menggunakan SPSS diatas dapat diketahui bahwa nilai Durbin Watson menunjukkan angka 2,173. Nilai DL dan DU di dapatkan dengan melihat tabel Durbin Watson dengan N= 12, K=4. Nilai DL sebesar 0,512 dan nilai DU sebesar 2,177 dan kurang dari 4-2,177 ($DU < DW \leq 4 - DU$) maka dapat disimpulkan bahwa tidak ada Autokolerasi positif maupun negatif atau dapat disimpulkan tidak terdapat Autokolerasi dalam penelitian ini.

Uji Heteroksedastisitas

Grafik scartterplot an ntara nilai prediksi variabel dependen, yaitu SRESID, dan residual ZPRED, dimana sumbu y adalah y yang telah diprediksi dan sumbu x adalah residual (y prediksi – y sesungguhnya) yang telah di-studentized. Ada atau tidak adanya heteroskedastisitas dapat dideteksi sebagai berikut:

1. Jika terdapat suatu pola tertentu, seperti titik-titik yang membentuk suatu pola teratur tertentu, hal ini menunjukkan adanya heteroskedastisitas.
2. Jika tidak ada pola yang jelas dan titik-titik berada di atas dan di bawah angka 0 pada sumbu y, maka tidak terjadi heteroskedastisitas.

Adapun hasil gambar uji heteroksedastisitas menggunakan SPSS versi 23, dapat dilihat pada gambar 2 berikut:



Gambar 2
Uji Heteroksedastisitas
Sumber: Output SPSS 25 (data diolah 2022)

Berdasarkan gambar 2 Scatterplot tersebut, terlihat titik-titik menyebar secara acak dan tidak membentuk suatu pola tertentu yang jelas, serta tersebar baik diatas maupun di bawah angka 0 pada sumbu Y. Hal ini berarti tidak terjadi heteroskedastisitas dalam model regresi, sehingga pengaruh variabel pertumbuhan ekonomi dapat diprediksi dengan menggunakan model regresi berdasarkan masukan dari variabel bebas.

Analisis Regresi Linear Berganda

Persamaan regresi dapat dilihat dari tabel hasil uji coefisient berdasarkan hasil SPSS versi 25 terhadap ketiga variabel independen Investasi, Tenaga Kerja Inflasi, dan Covid 19 terhadap variabel dependen Pertumbuhan Ekonomi di Indonesia ditunjukkan pada tabel 7.

Tabel 7
Analisis Regresi Linear Berganda

Model	Unstandardized Coefficients	
	B	Std. Error
1 (Constant)	-5,039	1,663
Investasi	0,141	0,024
Tenaga Kerja	1,388	0,221
Inflasi	-0,026	0,012
Covid 19	-0,012	0,007

Berdasarkan dapat dituliskan dalam penelitian ini sebagai berikut:

tabel 7 diatas maka persamaan regresi

$$\text{LogY} = -5,039 + 0,141 \text{ LogX}_1 + 1,388 \text{ LogX}_2 - 0,026 \text{ LogX}_3 - 0,012 \text{ X}_4 + e.$$

Dari persamaan regresi berganda diatas dapat dilihat sebagai berikut:

Nilai koefisien sebesar -5,039 jika variabel investasi (X1), tenaga kerja (X2), Inflasi (X3) dan Covid 19 (X4) konstan atau X= 0, maka pertumbuhan ekonomi sebesar -4,284

1. Nilai koefisien = 0,141 artinya variabel tenaga kerja, inflasi dan covid 19 konstan. Dan variabel investasi meningkat sebesar 1 Rupiah, kemudian pertumbuhan ekonomi meningkat sebesar 0,141 Rupiah. Koefisien bernilai positif yang berarti terdapat hubungan positif antara investasi dengan pertumbuhan ekonomi, karena semakin besar investasi maka semakin besar pula pertumbuhan ekonomi.
2. Nilai koefisien = 1,388 artinya variabel investasi, inflasi dan covid 19 konstan. Dan variabel tenaga kerja meningkat sebesar 1 Rupiah, kemudian pertumbuhan ekonomi meningkat sebesar 1,388 Rupiah. Koefisien positif berarti terdapat hubungan positif antara penyerapan tenaga kerja dengan pertumbuhan ekonomi, karena bertambahnya jumlah pekerja menurunkan tingkat pengangguran sehingga menaikkan tingkat produksi, yang dapat menyebabkan peningkatan pertumbuhan ekonomi ke tingkat output.
3. Nilai koefisien = -0,026 artinya variabel investasi, tenaga kerja dan covid 19 konstan. Dan variabel inflasi meningkat sebesar 1 persen, kemudian pertumbuhan ekonomi meningkat sebesar -0,026 persen. Koefisiennya negatif, yang berarti ada hubungan negatif antara inflasi dan pertumbuhan ekonomi, karena inflasi berakselerasi dengan cepat, jumlah uang beredar dan tingkat konsumsi domestik naik. Jadi pertumbuhan ekonomi mengalami penurunan terhadap tingkat produksi barang atau jasa yang dihasilkan.

4. Nilai koefisien = -0,012 artinya variabel investasi, tenaga kerja dan inflasi konstan. Dan variabel Covid 19 meningkat sebesar 1, kemudian pertumbuhan ekonomi meningkat sebesar -0,012. Koefisiennya negatif, yang berarti ada hubungan negatif antara Covid 19 dan pertumbuhan ekonomi.

Pengaruh Investasi Terhadap Pertumbuhan Ekonomi

Dari hasil pengolahan data tersebut menggunakan aplikasi SPSS menunjukkan bahwa nilai koefisien variabel investasi sebesar 0.141 dengan nilai signifikan lebih kecil dari 0,05 ($0,001 < 0,01$), artinya jika investasi meningkat sebesar satu rupiah maka pertumbuhan ekonomi akan meningkat sebesar 0.141 persen. Sehingga variabel investasi berpengaruh positif terhadap pertumbuhan ekonomi Indonesia. Dari nilai koefisien regresinya, variabel investasi menunjukkan angka yang relatif kecil, ini mengindikasikan bahwa kontribusi investasi sebagai penggerak pertumbuhan ekonomi Indonesia masih belum optimal. Hal ini dikarenakan iklim investasi yang masih kurang kondusif, seperti regulasi yang tidak konsisten, pelayanan yang belum maksimal, dan kurangnya jaminan keamanan komersial yang tidak dilaksanakan dengan baik untuk pertumbuhan ekonomi. Dengan cara ini, investor berpikir lebih hati-hati tentang keuntungan dan pengembalian modal yang diinvestasikan.

Jika beberapa faktor yang mempengaruhi keputusan investasi, seperti keuntungan masa depan, tingkat keamanan, keadaan politik, dan regulasi telah sesuai dengan yang diharapkan, tercipta iklim investasi yang baik, menjadi daya tarik tersendiri bagi investor, dan mereka yakin untuk menanamkan modalnya. Semakin banyak investor menginvestasikan dananya, semakin banyak sumber modal yang tersedia untuk membangun dan menumbuhkan perekonomian negara.

Investasi adalah penanaman modal atau pembentukan modal dapat diartikan sebagai pengeluaran atau perbelanjaan atas penanaman modal atau usaha untuk membeli barang modal dan peralatan produksi untuk meningkatkan kemampuan memproduksi barang dan jasa yang tersedia dalam perekonomian. Model investasi negara yang berperan sebagai bentuk modal bagi pembangunan negara untuk mencapai berbagai tujuan pembangunan dapat dikelompokkan menjadi dua, yaitu model investasi swasta dan model investasi negara. Investasi negara memainkan peran penting dalam membentuk model pembangunan negara untuk meningkatkan produksi.

Sebelum pandemi Covid-19, kondisi ekonomi global terus menunjukkan pertumbuhan positif. Bahkan sebelum Covid-19, ada beberapa ancaman yang mengintai ekonomi global, yaitu ketegangan geopolitik antara Amerika Serikat dan Iran, perang dagang antara Amerika Serikat dan Uni Eropa yang dipicu oleh EU-Green Deal, perang dagang antara Amerika Serikat dan Uni Eropa yang dipicu oleh green deal UE. Amerika Serikat, Cina dan urusan Brexit yang belum selesai. Namun secara umum, kondisi perekonomian global masih baik dan potensi investasi masih baik sebelum adanya pandemi Covid-19.

Di Indonesia, virus Covid-19 pertama kali terdeteksi pada awal atau pertengahan Maret. Setelah virus ditemukan, tren IHSG turun. Karena saat itu ada masalah Covid-19 yang mulai menyebar dari Wuhan hingga Jepang, Korea dan Singapura yang paling dekat dengan Indonesia. Sehingga penurunan ini menyebabkan IHSG kita jatuh di bawah 4000. Tentu penurunan ini tidak lepas dari pandangan investor yang melihat bahwa pemerintah Indonesia saat itu tidak serius menangani Covid-19, sehingga ketika krisis kesehatan terjadi dan perasaan itu ada, makanya investor lebih suka mengeluarkan dananya dari pasar modal, sehingga pasti akan menyebabkan harga saham turun.

Perusahaan sangat mempengaruhi pergerakan pasar modal dalam hal investasi. Saat PSBB terjadi, banyak perusahaan yang kolaps. Jika kita lihat hari ini, emiten yang berperan di sektor pariwisata semuanya negatif. Sehingga jika dicermati, pandemi COVID-19 tidak hanya berdampak

pada aspek keuangan perusahaan tetapi juga berdampak pada aspek riil dan fundamental. Maka wajar jika harga saham turun atau bahkan saat ini harga saham tidak berkinerja sebaik sebelum pandemi.

Sementara banyak perusahaan tidak dapat bertahan dalam kondisi saat ini, perusahaan telekomunikasi sebenarnya telah meningkat selama pandemi ini. Di masa pandemi ini Telkomsel, XL, Indosat meraup keuntungan luar biasa karena penggunaan internet untuk work from home (WFH) dan belajar di rumah meningkat. Dan beberapa perusahaan makanan dan minuman seperti Indofood sudah cukup sukses karena meskipun pandemi melanda, keuntungan perusahaan mereka terus bertambah karena perusahaan Indofood memproduksi kebutuhan pokok yang dibutuhkan saat ini.

Hal tersebut sejalan dengan penelitian terdahulu (Sianipar, 2019) bahwa variabel Investasi berpengaruh signifikan dan bertanda positif terhadap PDB, artinya dengan adanya investasi meningkat, PDB meningkat. Hal ini sejalan dengan teori Harrod Domar bahwa semakin tinggi nilai investasi maka semakin cepat pertumbuhan ekonomi dapat dilihat dari pembentukan PDB dan teori percepatan bahwa dalam jangka panjang pertumbuhan investasi meningkatkan pendapatan nasional.

Pengaruh Tenaga Kerja Terhadap Pertumbuhan Ekonomi

Dari hasil pengolahan data tersebut menggunakan aplikasi SPSS menunjukkan bahwa nilai koefisien variabel Tenaga Kerja sebesar 1,388 dengan nilai signifikan lebih kecil dari 0,05 ($0,000 > 0,01$), artinya jika tenaga kerja meningkat sebesar satu orang maka pertumbuhan ekonomi akan meningkat sebesar 1,388. Variabel tenaga kerja berhubungan positif dan signifikan terhadap pertumbuhan ekonomi. Peningkatan jumlah pekerja meningkatkan tingkat produksi barang dan jasa, sehingga pertumbuhan ekonomi meningkat. Hasil penelitian ini menyimpulkan bahwa jumlah tenaga kerja cenderung dapat meningkatkan angka pertumbuhan ekonomi. Oleh karena itu perlu diupayakan pemenuhan sumber daya manusia yang berkualitas sebagai modal pembangunan, penyerapan tenaga kerja di sektor informal disebabkan karena tingkat keterampilan yang relatif rendah dibandingkan dengan sektor formal.

Penyebaran pandemi COVID-19 telah memaksa pemerintah untuk menerapkan penguncian regional dan pembatasan sosial secara ekstensif. Akibatnya, kebijakan ini mengganggu aktivitas ekonomi dan sosial, yang pada gilirannya mentransmisikan gangguan ke seluruh ekonomi, termasuk gangguan di pasar tenaga kerja dan pendapatan yang lebih rendah bagi pekerja di seluruh wilayah.

Gangguan ekonomi yang disebabkan oleh kebijakan penahanan regional yang diterapkan untuk mencegah penyebaran virus telah menyebabkan banyak bisnis tutup dan bangkrut, yang menyebabkan PHK dan PHK massal, terutama di sektor-sektor yang paling terkena dampak. Kebijakan ketenagakerjaan selama pandemi COVID-19 fokus pada mendukung sektor bisnis/pengusaha, pekerja dan penciptaan lapangan kerja. Langkah-langkah ini juga mencakup dukungan keuangan dari beberapa lembaga dan otoritas untuk mendukung perusahaan, rumah tangga, dan pekerja yang terkena dampak pandemi, serta langkah-langkah dari otoritas untuk membantu pekerja yang menganggur memulai bisnis dan mendukung pelatihan. Untuk pencari kerja dan untuk menyediakan pekerjaan darurat.

Pemerintah telah mengambil langkah-langkah yang tepat untuk mengurangi dampak pandemi pada kehidupan kerja. Upaya tersebut dilakukan dengan memberikan paket stimulus ekonomi kepada dunia usaha, pembayaran pajak penghasilan kepada pekerja, jaring pengaman sosial melalui program kesejahteraan pegawai formal dan informal, Program Kartu Prakerja, perluasan program industri padat karya dan perlindungan TKI. Pemerintah juga melaksanakan reformasi kehidupan kerja melalui undang-undang ketenagakerjaan, mendorong masuknya investasi, serta memastikan peningkatan perlindungan dan kesejahteraan karyawan. Dewan juga berfokus pada pengembangan kualitas sumber daya manusia sebagai salah satu bidang prioritas kehidupan kerja.

Hal tersebut tidak sejalan dengan penelitian terdahulu (Sianipar, 2019) bahwa variabel tenaga kerja berpengaruh negatif dan tidak signifikan terhadap pertumbuhan ekonomi. Kualitas pekerjaan masih rendah, sehingga mereka hanya bekerja di sektor yang berpenghasilan rendah. Selain itu, peningkatan jumlah pekerja informal melemahkan pertumbuhan ekonomi. Pelatihan di tempat kerja dapat ditingkatkan oleh negara, tetapi juga oleh sektor swasta. Magang dapat berhubungan dengan menjahit, kecantikan dan banyak lagi. Selain itu, untuk menciptakan tenaga kerja yang kompeten, profesional, dan berdaya saing, diperlukan inovasi-inovasi yang sangat kreatif. setiap tahun meningkatnya permintaan pekerjaan mengurangi pengangguran dan meningkatkan kesejahteraan masyarakat. Kebijakan lain dari pemerintah untuk menambah jumlah pegawai dapat mendorong arus informasi tentang lowongan pekerjaan yang terbuka ke seluruh pelosok tanah air sehingga masyarakat dapat dengan mudah mencari pekerjaan yang sesuai dengan keahliannya.

Pengaruh Tingkat Inflasi Terhadap Pertumbuhan Ekonomi

Dari hasil pengolahan data tersebut menggunakan aplikasi SPSS menunjukkan bahwa nilai koefisien variabel inflasi sebesar -0.026 dengan nilai signifikan lebih kecil dari 0,05 ($0,061 < 0,05$), artinya jika inflasi meningkat sebesar satu persen (1%) maka pertumbuhan ekonomi akan menurun sebesar -0.026 persen (*Ceteris Paribus*). Sehingga variabel Inflasi berpengaruh negatif terhadap pertumbuhan ekonomi. Dalam penelitian ini, pertumbuhan ekonomi yang di proxy kan dengan nilai PDB atas dasar harga konstan memperlihatkan hubungan yang negatif, dimana ketika inflasi meningkat, PDB Indonesia cenderung meningkat, namun data menunjukkan bahwa pertumbuhan ekonomi melambat. Pertumbuhan ekonomi adalah pertumbuhan barang dan jasa selama kurun waktu tertentu, yang biasanya satu tahun, pertumbuhan ekonomi ini lebih banyak dipengaruhi oleh teknologi, dimana teknologi merupakan salah satu faktor terpenting dalam produksi barang dan jasa. Dengan kata lain, orang lebih mudah membuat suatu barang karena biaya produksinya rendah, tingkat produksi yang rendah juga menyebabkan inflasi yang rendah karena sedikitnya uang yang beredar untuk membuat barang tersebut.

Penyebab inflasi selama pandemi adalah inflasi alamiah, yaitu inflasi disebabkan oleh sebab-sebab alamiah, di mana manusia tidak mempengaruhinya (dalam arti pencegahan), inflasi adalah inflasi yang disebabkan oleh penurunan total supply atau kenaikan inflasi. Total pasokan tuntutan Karena penurunan tingkat produksi. Jika virus ini tidak segera mereda, dikhawatirkan perekonomian lokal akan terpuruk akibat pekerjaan yang terdampak pandemi COVID-19 akibat beberapa sektor dan profesi terdampak virus corona.

Dampak dari adanya inflasi menyebabkan perekonomian masyarakat terpuruk karena harga barang-barang yang memenuhi kebutuhan meningkat akibat terbatasnya kegiatan akibat rumitnya kegiatan produksi sehingga sulit bagi masyarakat untuk memenuhi kebutuhannya. Banyak masyarakat yang kehilangan pekerjaan atau mata pencaharian karena virus COVID-19, penurunan besar dalam upah dalam skala besar, sumber daya alam terbatas, tetapi permintaan terus meningkat, sehingga orang harus berpikir dan bekerja keras untuk memenuhi kebutuhan mereka.

Hal tersebut tidak sejalan dengan penelitian terdahulu (Sianipar, 2019) bahwa variabel inflasi berpengaruh signifikan dan bertanda negatif terhadap pertumbuhan ekonomi, yang artinya ketika inflasi menurun maka akan meningkatkan pertumbuhan ekonomi di Indonesia. Hal ini karena ketika inflasi terjadi menyebabkan harga barang dan jasa rumah tangga naik. Hasil penelitian ini sejalan dengan (Sepriani & Hulu, 2021) menjelaskan bahwa inflasi biasanya mengakibatkan penurunan daya beli masyarakat, karena pada kenyataannya tingkat pendapatan mereka juga menurun.

Hal tersebut tidak sejalan dengan penelitian terdahulu (Rizal et al., 2019) sebelumnya bahwa inflasi berpengaruh negatif dan signifikan terhadap pertumbuhan ekonomi di Indonesia. Ketika inflasi meningkat, pertumbuhan ekonomi melambat, dan sebaliknya, ketika inflasi melambat, pertumbuhan ekonomi meningkat.

Pengaruh Covid 19 Terhadap Pertumbuhan Ekonomi

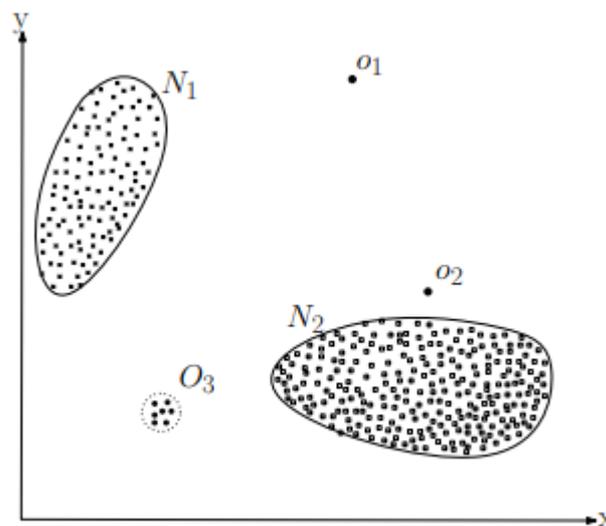
Dari hasil pengolahan data tersebut menggunakan aplikasi SPSS menunjukkan bahwa nilai koefisien variabel Covid 19 sebesar -0.012 dengan nilai signifikan lebih besar dari 0,05 ($0,149 < 0,05$), artinya jika Covid 19 meningkat sebesar satu persen (1%) maka pertumbuhan ekonomi akan menurun sebesar -0.012 persen (*Ceteris Paribus*). Sehingga variabel Covid 19 berpengaruh negatif terhadap pertumbuhan ekonomi.

World Health Organization (WHO) menyatakan Covid-19 sebagai pandemi global pada 9 Maret 2020. Setelah itu, pemerintah Indonesia juga mulai melakukan pencegahan penyebaran Covid-19. Pemerintah juga memutuskan untuk menahan penyebaran epidemi dengan menyediakan anggaran untuk sektor kesehatan, mengingat kondisi ekonomi di Indonesia saat itu. Pandemi tidak hanya berdampak pada tubuh manusia, tetapi juga berdampak pada semua bidang seperti ekonomi, kesehatan, sosial, budaya, transportasi.

Kapan kedatangan Covid-19 dimulai di Indonesia, yaitu pada triwulan I-2020, pertumbuhan ekonomi nasional dari sisi produk domestik bruto (PDB) masih positif - 2,97%. Total PDB atas dasar harga berlaku untuk triwulan II-2020 adalah Rp3.687,7 triliun, tetapi dihitung atas dasar harga konstan (tahun dasar 2010) adalah Rp2.589,6 triliun. Kemudian, dampak pandemi Covid-19 mulai terasa pada triwulan II-2020, yaitu mengalami penurunan atau penurunan sebesar 5,32 persen dibandingkan triwulan II-2019. Selain itu, juga mengalami penurunan pada triwulan III dan IV. dibandingkan dengan tahun 2019 sebesar 3,49% dan 2,19% (BPS 2020). Hal yang sama terjadi tidak hanya di Indonesia, tetapi juga di Amerika Serikat, Italia, Jerman dan Korea Selatan.

Pola data yang tidak sesuai dengan konsep perilaku normal disebut anomaly. Anomali dalam kumpulan data dua dimensi sederhana ditunjukkan pada Gambar 2. Data ini memiliki dua wilayah normal, N1 dan N2, karena sebagian besar pengamatan terjadi di kedua wilayah ini. Titik-titik di luar wilayah tersebut, seperti titik o1 dan o2, dan titik-titik di wilayah o2 dan O3, adalah anomali (Chandola et al., 2009).

Ada banyak alasan anomali data dapat terjadi, seperti aktivitas ilegal, seperti penipuan kartu kredit, intrusi dunia maya, aktivitas teroris, atau kerusakan sistem. Namun, karakteristik umum dari setiap alasan tersebut menarik bagi analisis. "Kemenarikan", atau keterkaitan anomali dengan dunia nyata, adalah ciri utama dalam mendeteksi anomali.



Gambar 2
Contoh anomaly sederhana dalam kumpulan data dua dimensi (sumber: <https://jmp.kemenkeu.go.id>)

Naive Bayes

Naive Bayes yang diciptakan oleh ilmuwan Inggris Thomas Bayes adalah cara untuk mengklasifikasikan data dengan metode probabilitas dan statistik. Metode ini berisi tentang peluang masa depan berdasarkan pengalaman masa lalu. Kemudian, teori Naive Bayes, yang menganggap kondisi antar atribut menjadi saling bebas, digunakan untuk menggabungkan teorema ini. Metode Naive Bayes untuk klasifikasi menunjukkan bahwa karakteristik tertentu dari sebuah kelas tidak terkait dengan karakteristik kelas lain (Oleh et al., 2022)(Bahtiar et al., 2020).

Berikut persamaan Teorema Bayes:

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}$$

Keterangan:

- P(H|E) : Probabilitas hipotesis H di berikan bukti E
- P(E|H) : Probabilitas bukti E di berikan hipotesis H
- P(H) : Probabilitas awal hipotesis H
- P(E) : Probabilitas total bukti E

Decision Tree

Decision Tree adalah algoritme pembelajaran mesin yang populer karena kemampuan untuk melakukan klasifikasi dan regresi dengan aturan keputusan yang mudah dipahami. Kelebihan utama algoritme ini adalah kemampuan untuk menangani data yang besar dan kompleks dengan waktu komputasi yang relatif cepat, algoritme ini membangun model pohon keputusan dari data pelatihan dan

menggunakan aturan yang dihasilkannya untuk mengkategorikan data baru. Untuk menyelesaikan masalah ini, penelitian ini mengembangkan dan mengevaluasi model pohon keputusan yang efisien yang dapat mendeteksi anomali lalu lintas jaringan dan meningkatkan keamanan dan kinerja jaringan komputer (At Thooriqoh et al., 2022).

Isolation Forest

Metode pengelompokan data yang dikenal sebagai Hutan Isolasi bertujuan untuk menemukan anomali dengan mengisolasi titik-titik data yang dianggap menyimpang atau tidak sesuai. Dalam metode ini, data dipisahkan secara rekursif dengan menggunakan pohon keputusan, di mana setiap pemisahan dilakukan secara acak pada setiap fitur. Titik data yang lebih cepat terisolasi atau membutuhkan lebih banyak pemisahan biasanya dianggap sebagai anomali. Karena waktu komputasi yang lebih cepat dibandingkan dengan metode lain yang berfokus pada deteksi anomali, metode ini sangat efektif pada dataset besar dan mampu mengidentifikasi outlier dengan efisien. (Zulfikar et al., 2023) (Triana, 2024) (Sri Lakshmi et al., 2023)

Pengujian

Pada penelitian ini, untuk deteksi anomali trafik pada jaringan komputer, kami menggunakan algoritma Naive Bayes, Decision Tree, dan Isolation Forest. Naive Bayes dan Decision Tree adalah dua algoritma pembelajaran mesin yang digunakan untuk problem klasifikasi, sedangkan Isolation Forest merupakan model deteksi anomali yang mengisolasi titik-titik data yang terlihat sebagai anomali. Ketiga metode ini difusi untuk meningkatkan akurasi dan efisiensi dalam identifikasi anomali pada jaringan komputer.

Dalam melakukan penelitian ada tahap-tahap yang dilakukan dalam mendeteksi anomali trafik pada Jaringan komputer yang meliputi:

a. Pengumpulan Data

Firewall membuat log trafik jaringan untuk mengumpulkan data. Log ini mencatat berbagai aktivitas jaringan, seperti protokol, alamat IP, port, waktu akses, ukuran paket, dan status koneksi. Selain trafik yang diizinkan oleh firewall, data ini menunjukkan aktivitas normal dan anomali. Trafik normal dan anomali kemudian dibedakan dengan memproses data ini. Selanjutnya, model deteksi anomali dilatih dan diuji dengan menggunakan Naive Bayes, Decision Tree, dan Isolation Forest. Pengambilan data ini memanfaatkan Paramiko untuk mengakses Log trafik Jaringan yang tersimpan pada server Jaringan komputer.

b. Pra Pemrosesan Data Jaringan

Pembersihan Data : Langkah ini dilakukan karena data yang diterima merupakan data kotor, penghapusan data yang tidak relevan, penanganan nilai yang hilang (nilai yang tidak ada), dan normalisasi fitur dilakukan untuk memastikan data dalam format yang dapat dianalisis. Karena data yang diterima merupakan data kotor, data kotor dianggap sebagai data yang tidak akurat, tidak konsisten, atau tidak lengkap karena kesalahan yang terjadi dalam kumpulan data.

Pembagian data : Pada langkah ini, data yang akan digunakan untuk analisis dibuat menjadi dua bagian data pelatihan dan data pengujian.

1. Penerapan Algoritma

Pada penerapannya menggunakan tiga algoritma pembelajaran mesin yaitu: Naive Bayes, Decision Tree dan Isolation Forest untuk mengembangkan metodologi untuk mendeteksi anomali traffic pada jaringan komputer. Ketika Naive Bayes, Decision Tree, dan Isolation Forest diimplementasikan pada deteksi anomali trafik jaringan, setiap algoritma ini dimanfaatkan dengan prinsip kekuatannya. Masing-masing algoritma mampu menganalisis pola trafik yang normal serta

mengetahui kegiatan tindakan yang menduga mencurigakan. Naive Bayes mengklasifikasi trafik berdasarkan kemungkinan terjadi fitur terkait, Decision Tree digunakan untuk pembuatan keputusan berdasarkan aturan yang dipahami dengan mudah. Sementara Isolation Forest mampu secara efektif mendeteksi anomali sebab algoritma memisahkan data berbeda dari pola umum melalui pembentukan pohon yang mengisolasi titik data tersebut. Kombinasi ketiga algoritma ini meningkatkan keandalan dan akurasi sistem deteksi anomali, yang memungkinkan identifikasi dini terhadap ancaman dan serangan yang mungkin terjadi pada jaringan komputer.

Naïve Bayes

Gambar 3
Hasil data Naïve Bayes

Metode Analisis	Precision	Recall	F1-Score	Support
Kelas 0.0	0.51	1.00	0.67	42
Kelas 1.0	0.00	0.00	0.00	41

Decision tree

Gambar 4
Hasil data Decision Tree

Metode Analisis	Precision	Recall	F1-Score	Support
Kelas 0.0	0.51	1.00	0.67	42
Kelas 1.0	0.00	0.00	0.00	41

Isolation Forest

Isolation Forest detected 42 anomalies.

Dalam data yang dianalisis, Model Hutan Isolasi menemukan 42 anomali. Deteksi ini menunjukkan adanya entri atau pola yang dianggap menyimpang atau tidak biasa dibandingkan dengan data normal. Anomali-anomali ini dapat dikaitkan dengan aktivitas yang mencurigakan, kesalahan data, atau fenomena lain yang memerlukan analisis lebih lanjut untuk mengetahui penyebab dan konsekuensi mereka secara menyeluruh.

Penerapan Algoritma

Pada penerapannya menggunakan tiga algoritma pembelajaran mesin yaitu: Naive Bayes, Decision Tree dan Isolation Forest untuk mengembangkan metodologi untuk mendeteksi anomali traffic pada jaringan komputer. Ketika Naive Bayes, Decision Tree, dan Isolation Forest diimplementasikan pada deteksi anomali trafik jaringan, setiap algoritma ini dimanfaatkan dengan prinsip kekuatannya. Masing-masing algoritma mampu menganalisis pola trafik yang normal serta mengetahui kegiatan tindakan yang menduga mencurigakan. Naive Bayes mengklasifikasi trafik berdasarkan kemungkinan terjadi fitur terkait, Decision Tree digunakan untuk pembuatan keputusan berdasarkan aturan yang dipahami dengan mudah. Sementara Isolation Forest mampu secara efektif mendeteksi anomali sebab algoritma memisahkan data berbeda dari pola umum melalui pembentukan

pohon yang mengisolasi titik data tersebut. Kombinasi ketiga algoritma ini meningkatkan keandalan dan akurasi sistem deteksi anomali, yang memungkinkan identifikasi dini terhadap ancaman dan serangan yang mungkin terjadi pada jaringan komputer.

Evaluasi kinerja

1. Naïve Bayes

Akurasi: 51% (sangat rendah; hanya mencapai sedikit di atas tebakan acak di dataset biner). Macro Average Precision, Recall, F1-score: 0.25, 0.50, 0.34 (menunjukkan ketimpangan dalam performa model dengan dua kelas). Bias Kelas model memberikan Recall 1.00 untuk kelas 0 yang juga merupakan mayoritas kelas tetapi Precision 0.51, menunjukkan tingkat error yang jauh lebih tinggi. Performa pada Kelas 1 Precision 0.00, Recall 0.00, F1-score 0.00 (model benar-benar tidak pernah mendeteksi kelas 1).

Kinerja model ini tidak cocok untuk penanganan dataset ketidakseimbangan. Kemungkinan alasan kesulitan adalah karena distribusi data kedua kelas menjadi sangat tidak seimbang atau kemampuan pembelajaran Naïve Bayes tidak mencukupi pola data yang kompleks. Diperlukan metode resampling data, baik oversampling atau undersampling atau penggunaan model yang m atas untuk dataset ketidakseimbangan.

2. Decision Tree

Akurasi: 51% (salah satunya sama dengan Naïve Bayes, yaitu tidak perlu dimasukkan). Macro Average Precision, Recall, F1-score: 0,25, 0,50, 0,34 (persis sama dengan Naïve Bayes). Bias kelas juga cenderung membuat model dengan Recall 1.00 untuk kelas mayoritas 0 tetapi Precision 51%. Performa pada Kelas 1 Precision 0.00, Recall 0.00, F1-score 0.00 (model benar-benar tidak pernah mendeteksi kelas 1).

Kejadian salah satu dari Algoritma lain, model ini condong ke kelas mayoritas dan sebagian besar akan melihat kelas 0. Mirip dengan Naïve Bayes, model ini overfitting ke kelas mayoritas memperhatikan nol dari kelas minoritas. Diperlukan pemotongan pohon dalam proses pemotongan pohon, meningkatkan model dengan pengaturan hyperparameter atau menggunakan backpack seperti Random Forest dengan ensemble methods.

Isolation Forest Anomaly Detection

Deteksi Anomali beberapa dari entri semacam ini diputer sebagai anomali. Untuk kasus ini, ada beberapa entri seperti baris 6, 7, dan 10. Itu perlu melacak bahwa model melihat pola tidak normal dalam data. Analisis Pola jika entri yang ditampilkan sebagai detektor, entri anomali dilakukan oleh IP address atau port berbasis, memerlukan analisis lebih lanjut apakah ada potensi terjadi perluasan atau menyarankan lalu lintas negatif sering gejala. Isolation Forest mendeteksi beberapa pola asing, tetapi perlu verifikasi cari tahu apakah entri yang diterapkan hanya mudah karena suatu lalu lintas liar. Rekomendasi perbaikan Mengatasi ketidakseimbangan data. Preference klasifikasi model. Selesai dengan model Random Forest dengan gradasi pita. Preference menggunakan ambang batas: ambang batas yang optimal untuk Naïve Bayes tutor.

Improving anomaly detection

Analisis address IP dan port mana yang paling sering dijelaskan dalam anomaly untuk mendeteksi ormasi kejahatan meyakini. Metode gabungan Isolation Forest dengan cara lainnya, misalnya dengan Metode Pendorong Keuntungan Faktor Parameter Merugikan.

Analisa

Kinerja model keseluruhan

1. Akurasi data model adalah 51%, yang dimana menunjukkan bahwa data model benar hanya 51% dari keseluruhan prediksi.

2. Macro average untuk presicion, recall dan F1-score masing-masing adalah 0.25, 0.50, dan 0.34. Hal ini menunjukkan adanya ketidak seimbangan performa antara kedua kelas.

Decision Tree

Kinerja model keseluruhan

5. Akurasi model dari data adalah 51%, membuktikan model hanya mampu memprediksi 51% dengan benar dari keseluruhan data.
6. Macro average untuk presicion, recall dan F1-score nilai masing-masing adalah 0.25, 0.50, dan 0.34, yang menunjukkan adanya ketidak seimbangan performa pada kedua kelas.

Model muncul pada bias ke arah kelas 0.0 dengan mengolah semua contoh sebagai bagian dari kelas itu dan tidak bisa mengenali pola yang berkaitan pada kelas 1.0. Ini adalah salah satu tanda bahwa model berusaha keras dalam mengapresiasi pola kompleks yang memisahkan kelas 1.0, mungkin karena alasan data yang tidak seimbang atau karena keterbatasan kemampuan model dalam mengolah ciri kedua kelas secara komprehensif.

Selanjutnya, distribusi data harus diperiksa untuk mengetahui apakah ketidakseimbangan data berdampak pada kinerja model. Metode seperti oversampling pada kelas minoritas atau undersampling pada kelas mayoritas dapat digunakan untuk mengatasi ketidakseimbangan dalam data. Selain itu, mencoba algoritma yang lebih kompleks, seperti Random Forest atau Gradient Boosting, atau pendekatan yang mengoptimalkan bobot berdasarkan distribusi kelas, dapat meningkatkan kemampuan model untuk menangani ketidakseimbangan dalam data.

KESIMPULAN

Penelitian ini mengevaluasi tiga algoritma Naive Bayes, Decision Tree, dan Isolation Forest dalam mendeteksi anomali pada trafik jaringan komputer. Hasilnya menunjukkan bahwa Naive Bayes dan Decision Tree memiliki akurasi rendah (51%) dengan bias terhadap trafik normal serta kesulitan dalam mendeteksi pola anomali karena ketidakseimbangan data. Sementara itu, Isolation Forest lebih efektif, berhasil mengidentifikasi 42 entri anomali tanpa bias terhadap distribusi kelas. Tantangan utama yang ditemukan adalah ketidakseimbangan data dan keterbatasan algoritma klasifikasi dalam menggeneralisasi pola kompleks. Untuk meningkatkan kinerja, disarankan penggunaan teknik oversampling seperti SMOTE, pengoptimalan model dengan algoritma lebih kompleks seperti Random Forest atau Gradient Boosting, serta validasi tambahan menggunakan data baru atau simulasi trafik. Pendekatan hybrid yang menggabungkan ketiga algoritma juga direkomendasikan untuk meningkatkan akurasi dan keandalan sistem deteksi anomali dalam mengidentifikasi ancaman keamanan jaringan.

DAFTAR PUSTAKA

- Anwar, K. (2022). Analisa sentimen Pengguna Instagram Di Indonesia Pada Review Smartphone Menggunakan Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 2(4), 148–155.
- Astuti, P. W. (2018). *Analisis pengaruh investasi terhadap pertumbuhan ekonomi (Studi pada 33 Provinsi di Indonesia)*. Universitas Brawijaya.
- At Thooriqoh, H., Naufal Azzmi, M. H., Ari Tofan, Y., & Mazharuddin Shiddiqi, A. (2022). Malicious Traffic Detection in Dns Infrastructure Using Decision Tree Algorithm. *JUTI: Jurnal Ilmiah*
-

Teknologi Informasi, 45–52.

- Bahtiar, R., Sakti, M. D., Setiawan, A., & Rosyani, P. (2020). Analisis Perbandingan Detection Traffic Anomaly dengan Metode Naive Bayes dan DBSCAN. *Jurnal Kreativitas Mahasiswa Informatika*, 1(2), 99–103.
- Beno, J., Silen, A. ., & Yanti, M. (2022). No 主観的健康感を中心とした在宅高齢者における健康関連指標に関する共分散構造分析Title. *Braz Dent J.*, 33(1), 1–12.
- Chandola, V., BANERJEE, A., & KUMAR, V. (2009). Survey of Anomaly Detection. *ACM Computing Survey (CSUR)*, 41(3), 1–72.
- Manalu, M. I., Hutabarat, F. P., Utara, U. S., Islam, U., & Sumatera, N. (2024). *Network Traffic Anomaly Detection Using the Decision Tree Method 1 1. 01(01)*, 37–44.
- Nasrullah, A. H. (2021). Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris. *Jurnal Ilmiah Ilmu Komputer*, 7(2), 45–51. <https://doi.org/10.35329/jiik.v7i2.203>
- Ness, S., Eswarakrishnan, V., Sridharan, H., Shinde, V., Janapareddy, N. V. P., & Dhanawat, V. (2025). Anomaly Detection in Network Traffic using Advanced Machine Learning Techniques. *IEEE Access*, 13(January), 16133–16149. <https://doi.org/10.1109/ACCESS.2025.3526988>
- Oleh, D., Persyaratan, U., Dan, P., Tugas, P., Sebagai, A., Proses, A., & Strata, S. (2022). *Klasifikasi Anomali Trafik Pada Ids Menggunakan Algoritma Naïve Baiyes Dan Random Forest Proposal Tugas Akhir*.
- Pertiwi, K. M. D., Oktavia, V. R., & Maulana, R. F. (2023). Deteksi Botnet pada Jaringan DNS secara Virtual menggunakan Decision Tree. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 9(3), 446. <https://doi.org/10.26418/jp.v9i3.70193>
- Prasetyo, A., Affandi, L., & Arpandi, D. (2018). Implementasi metode naive bayes untuk intrusion detection system (ids). *Jurnal Informatika Polinema*, 4(4), 280.
- Rizal, A., Zulham, T., & Asmawati, A. (2019). Analisis pengaruh pertumbuhan ekonomi, inflasi, dan suku bunga terhadap kredit macet di Indonesia. *Jurnal Ekonomi Dan Kebijakan Publik Indonesia*, 6(1), 1–16.
- Sepriani, L., & Hulu, Y. J. (2021). Pengaruh Investasi Penanaman Modal Terhadap Pertumbuhan Ekonomi Melalui Peningkatan Produk Domestik Bruto di Indonesia. *Wacana Ekonomi (Jurnal Ekonomi, Bisnis Dan Akuntansi)*, 20(2), 77–83.
- Sianipar, Y. L. (2019). *Pengaruh Inflasi, Investasi, Nilai Tukar, Dan Tenaga Kerja Terhadap Pertumbuhan Ekonomi Indonesia*. Universitas Brawijaya.
- Sri Lakshmi, M., Rajavikram, G., Dattatreya, V., Swarna Jyothi, B., Patil, S., & Bhavsingh, M. (2023). Evaluating the Isolation Forest Method for Anomaly Detection in Software-Defined Networking Security. *Journal of Electrical Systems*, 19(4), 279–297. <https://doi.org/10.52783/jes.639>
-

Suwaryo, N., Nawangsih, I., & Rejeki, S. (2014). Deteksi Serangan Pada Intrusion Detection System (Ids) Untuk Klasifikasi Serangan Dengan Algoritma Naïve Bayes, C.45 Dan K-Nn Dalam Meminimalisasi Resiko Terhadap Pengguna. *Jurnal Sistem Informasi Universitas Suryadarma*, 8(2). <https://doi.org/10.35968/jsi.v8i2.732>

Triana, O. (2024). Deteksi Anomali Jaringan Menggunakan Algoritma Isolation Forest. *Jurnal Dunia Data*, 1(5), 1–18.

Zulfikar, A., Rahmani, F. A., Azizah, N., Perbendaharaan, D. J., Keuangan, K., & Pinang, P. (2023). Deteksi Anomali Menggunakan Isolation Forest Belanja Barang Persediaan Konsumsi Pada Satuan Kerja Kepolisian Republik Indonesia. *Jurnal Manajemen Perbendaharaan*, 4(1), 1–15. <https://doi.org/10.33105/jmp.v4i1.435>



© 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).