JITET (Jurnal Informatika dan Teknik Elektro Terapan)
Vol. 13 No. 3, pISSN: 2303-0577 eISSN: 2830-7062 http://dx.doi.org/10.23960/jitet.v13i3.6594

ANALISIS PERILAKU PENGGUNA TERHADAP AKSES INTERNET DI PT CHIYODA INTERNATIONAL INDONESIA MENGGUNAKAN MACHINE LEARNING

Adi Agustani[1*], Herri Setiawan[2], Tasmi[3]

[1,2]Teknik Informatika, Universitas Indo Global Mandiri
[3]Sistem Komputer, Universitas Indo Global Mandiri
Jalan Jend. Sudirman Km.4 No. 62, Kota Palembang, Indonesia

Keywords:
Perilaku Pengguna;
Akses Internet;
Machine Learning;
K-Means;
Python.

Corespondent Email:
2022110017p@students.uigm.ac.id

Abstrak. Akses internet merupakan faktor krusial dalam mendukung kelancaran operasional perusahaan. PT Chiyoda International Indonesia, yang sangat bergantung pada infrastruktur TI dan konektivitas internet, menjadi objek dalam penelitian ini. Tujuan utama penelitian adalah menganalisis pola perilaku pengguna dalam mengakses internet melalui penerapan teknik *machine learning* menggunakan metode *K-Means* berbasis *Python*. Data yang dianalisis berasal dari lebih dari 10.000 entri aktivitas internet selama enam bulan, mencakup lima departemen utama. Hasil analisis *elbow plot* menunjukkan bahwa tiga *cluster* merupakan jumlah optimal, dengan penurunan nilai *inertia* sebesar 65% dibandingkan konfigurasi awal. Klasterisasi menunjukkan bahwa 40% pengguna tergolong dalam intensitas tinggi (rata-rata 4,5 jam/hari), 35% intensitas sedang (3 jam/hari), dan 25% intensitas rendah (1,5 jam/hari). Analisis visualisasi *cluster plot* mengungkap perbedaan signifikan antar departemen, terutama dalam jenis aplikasi yang diakses. Departemen dengan intensitas tertinggi cenderung menggunakan aplikasi non-produktif seperti media sosial dan *streaming*. Temuan ini menjadi dasar penting bagi perusahaan dalam mengoptimalkan alokasi *bandwidth* dan merumuskan kebijakan pengawasan akses. Penelitian ini diharapkan memberikan kontribusi strategis dalam pengelolaan sumber daya digital secara efisien di era transformasi teknologi.

Abstract. *Internet access is a critical factor in ensuring smooth company operations. PT Chiyoda International Indonesia, which heavily relies on IT infrastructure and internet connectivity, serves as the subject of this study. The primary objective is to analyze user behavior patterns in internet usage through the application of* machine learning *techniques using the* K-Means *method implemented in* Python*. The dataset comprises over 10,000 internet activity entries collected over six months from five major departments. The initial analysis using the* elbow plot *revealed that three clusters were optimal, with a 65% reduction in* inertia *compared to the initial configuration. Clustering results showed that 40% of users fall into the high-intensity cluster (average usage of 4.5 hours/day), 35% into the medium-intensity cluster (3 hours/day), and 25% into the low-intensity cluster (1.5 hours/day). The* cluster plot *analysis indicated significant behavioral differences across departments, particularly regarding application types and usage frequency. Departments with the highest usage levels tended to access non-productive applications such as social media and streaming services, potentially affecting work efficiency. These findings provide a strategic foundation for optimizing* bandwidth *allocation and implementing internet access control policies. This study offers valuable insights into corporate internet usage dynamics and*

contributes to effective digital resource management in the era of technological transformation.

1. PENDAHULUAN

Di era digital yang terus berkembang, akses internet telah menjadi komponen fundamental dalam mendukung kelancaran operasional perusahaan. PT Chiyoda International Indonesia, yang bergerak dalam sektor industri dengan ketergantungan tinggi pada teknologi informasi, menjadikan konektivitas internet sebagai infrastruktur vital [1]. Ketersediaan sambungan internet yang stabil, andal, dan efisien tidak hanya berfungsi sebagai penunjang aktivitas harian, tetapi juga berperan penting dalam menjaga produktivitas karyawan dan daya saing perusahaan di tingkat global [2].

Penggunaan internet yang tidak terkendali dapat mengakibatkan pemborosan sumber daya dan menimbulkan potensi risiko terhadap keamanan jaringan interna [3]-[4]. Oleh karena itu, pemahaman terhadap pola perilaku pengguna dalam mengakses internet sangat diperlukan untuk memastikan efektivitas dan efisiensi pemanfaatan infrastruktur digital [5]. Salah satu pendekatan yang relevan untuk menjawab tantangan ini adalah penerapan teknik *machine learning*, khususnya melalui implementasi bahasa pemrograman *Python* [7]. Teknik ini memungkinkan pengolahan data secara cerdas guna menghasilkan wawasan yang dapat digunakan untuk pengambilan keputusan berbasis data [8].

Machine learning merupakan cabang dari *artificial intelligence* yang memungkinkan sistem komputer mempelajari pola dan membuat prediksi atau keputusan tanpa diprogram secara eksplisit [9]. Dalam konteks industri, penerapan *machine learning* dan *data mining* telah terbukti efektif dalam memecahkan berbagai persoalan terkait analisis perilaku pengguna, efisiensi operasional, dan pengambilan keputusan berbasis prediksi [10]. Berbagai studi juga menunjukkan potensi besar *machine learning* dalam mendeteksi anomali serta mengidentifikasi kecenderungan pengguna yang dapat berimplikasi terhadap keamanan atau [11]-[12].

Salah satu metode *machine learning* yang umum digunakan untuk analisis perilaku adalah *K-Means clustering*. Metode ini bekerja tanpa label data (unsupervised learning) untuk mengelompokkan objek berdasarkan kemiripan fitur atau karakteristik [13]. Dalam lingkungan organisasi, metode ini dapat dimanfaatkan untuk mengelompokkan pengguna berdasarkan intensitas penggunaan internet, jenis aplikasi yang diakses, serta frekuensi interaksi jaringan. Analisis ini memberikan pemetaan perilaku yang dapat digunakan sebagai dasar perumusan kebijakan pengelolaan akses internet yang lebih terukur dan strategis [14].

Dengan pendekatan *K-Means*, data yang berasal dari aktivitas jaringan dapat dianalisis secara objektif untuk menghasilkan *cluster* perilaku pengguna. Setiap *cluster* mencerminkan karakteristik unik yang berkaitan dengan durasi penggunaan, aplikasi yang dominan diakses, dan kemungkinan keterkaitan dengan produktivitas atau risiko keamanan. Visualisasi hasil klasterisasi dapat mempermudah pihak manajemen TI dalam memahami pola besar penggunaan internet dan menetapkan prioritas dalam alokasi *bandwidth* maupun pengendalian akses.

Penelitian ini difokuskan pada analisis perilaku pengguna terhadap akses internet di PT Chiyoda International Indonesia dengan menerapkan metode *K-Means clustering* berbasis *machine learning*. Tujuan utama adalah mengidentifikasi pola penggunaan internet, mengevaluasi kecenderungan pengguna dalam memanfaatkan sumber daya digital, serta memberikan rekomendasi yang dapat menunjang efisiensi operasional dan peningkatan keamanan jaringan perusahaan. Pendekatan ini juga diharapkan berkontribusi dalam pengembangan praktik manajemen teknologi informasi berbasis analitik di lingkungan korporasi modern.

2. TINJAUAN PUSTAKA

2.1 Akses Internet dalam Lingkungan Korporasi

Akses internet menjadi infrastruktur utama dalam mendukung kegiatan operasional dan komunikasi digital di lingkungan perusahaan modern. Ketersediaan koneksi internet yang stabil, cepat, dan aman sangat penting untuk memastikan kelancaran kerja, kolaborasi tim, dan efisiensi proses bisnis. Dalam konteks manajemen TI, pemanfaatan internet secara efektif juga berdampak langsung pada produktivitas dan penghematan sumber daya [15]. Namun demikian, tingginya ketergantungan terhadap internet juga menimbulkan tantangan, terutama dalam hal pengelolaan akses yang tidak efisien, pemborosan bandwidth, serta potensi ancaman keamanan dari aktivitas pengguna yang tidak terkontrol [16].

2.2 Perilaku Pengguna dan Efisiensi Digital

Perilaku pengguna (user behaviour) dalam mengakses internet memegang peranan penting dalam menentukan efektivitas sistem informasi perusahaan. Studi tentang pola penggunaan, durasi akses, dan jenis aplikasi yang dibuka dapat membantu perusahaan dalam merancang kebijakan alokasi bandwidth dan pengendalian akses.

Pengguna yang terlalu banyak mengakses aplikasi non-produktif, seperti media sosial atau *streaming*, dapat mengganggu performa jaringan dan menurunkan efisiensi kerja secara keseluruhan [17]. Oleh karena itu, penting untuk melakukan pemantauan dan analisis perilaku pengguna secara sistematis menggunakan metode yang berbasis data.

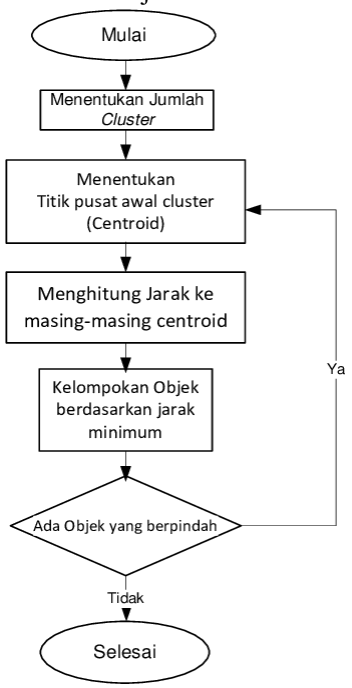
2.3 Konsep Machine Learning dalam Analisis Data

Machine learning merupakan cabang dari *artificial intelligence* yang fokus pada pengembangan algoritma untuk mempelajari pola dari data dan melakukan prediksi atau pengelompokan secara otomatis. Dalam konteks manajemen jaringan dan perilaku pengguna, *machine learning* memungkinkan analisis data berskala besar dengan efisiensi dan akurasi tinggi [18]. Teknik ini dapat digunakan untuk mendeteksi anomali, mengelompokkan pengguna berdasarkan intensitas akses, dan memprediksi potensi risiko terhadap sistem. Studi sebelumnya menunjukkan bahwa *machine learning* efektif dalam mendukung pengambilan keputusan berbasis data dalam berbagai sektor, termasuk industri manufaktur dan layanan TI [19].

2.4 Algoritma K-Means untuk Analisis Perilaku

Algoritma *K-Means* adalah salah satu metode *unsupervised learning* yang paling banyak digunakan untuk keperluan *clustering* atau pengelompokan data berdasarkan kemiripan antar objek. Dalam konteks analisis perilaku pengguna internet, *K-Means* dapat mengelompokkan pengguna ke dalam beberapa klaster berdasarkan pola penggunaan seperti durasi akses, jenis aplikasi, dan waktu interaksi [20]. Metode ini telah digunakan secara luas dalam eksplorasi data karena kesederhanaan implementasi dan kemampuannya mengidentifikasi struktur tersembunyi dalam data besar. Dengan bantuan *visualisasi cluster*, hasil analisis ini dapat memberikan gambaran yang informatif bagi manajemen perusahaan dalam pengambilan keputusan strategis [15].

Gambar di bawah ini menunjukkan alur kerja algoritma *K-Means* yang digunakan dalam proses klasterisasi data pada penelitian ini. Proses dimulai dengan menentukan jumlah klaster (*K*) yang akan digunakan sebagai dasar pengelompokan. Selanjutnya, titik pusat awal (*centroid*) ditentukan secara acak dari dataset, lalu algoritma menghitung jarak setiap objek ke seluruh centroid yang ada. Berdasarkan perhitungan tersebut, objek dikelompokkan ke dalam klaster yang memiliki jarak terdekat. Setelah pengelompokan awal selesai, centroid setiap klaster diperbarui berdasarkan rata-rata posisi objek-objek dalam klaster tersebut. Proses ini diulang terus menerus hingga tidak ada lagi objek yang berpindah klaster, yang menandakan bahwa model telah mencapai kondisi konvergen. Mekanisme ini memungkinkan sistem secara otomatis mengelompokkan perilaku pengguna internet berdasarkSan pola yang ditemukan dalam data historis, sehingga mendukung analisis yang lebih terstruktur dan objektif.



Gambar 1. Tahapan Algoritma *K-Means*

2.5 Relevansi terhadap Penelitian

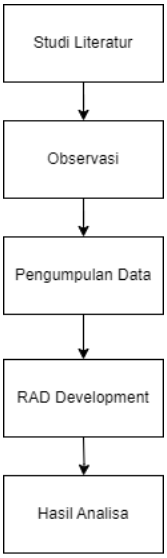
Penelitian ini memanfaatkan algoritma *K-Means* dalam kerangka *machine learning* berbasis *Python* untuk menganalisis perilaku pengguna terhadap akses internet di PT Chiyoda International Indonesia. Pendekatan ini memungkinkan perusahaan mengidentifikasi kelompok pengguna dengan intensitas akses berbeda serta memahami dampaknya terhadap performa jaringan dan produktivitas kerja. Dengan menganalisis pola akses internet secara kuantitatif, penelitian ini diharapkan mampu memberikan kontribusi nyata dalam pengelolaan sumber daya digital yang lebih efisien, aman, dan adaptif terhadap kebutuhan operasional perusahaan.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif berbasis *machine learning* dengan algoritma *K-Means* untuk menganalisis perilaku pengguna terhadap akses internet di PT Chiyoda International Indonesia. Pendekatan ini dipilih karena kemampuannya dalam mengelompokkan data berdasarkan pola tanpa memerlukan label sebelumnya. Di samping itu, metode *Rapid Application Development* (RAD) digunakan untuk membangun sistem analisis secara iteratif dan adaptif, memungkinkan pengembangan yang cepat

sesuai perubahan kebutuhan data dan umpan balik pengguna [19].

Rangkaian proses penelitian ini meliputi tahapan: studi literatur, observasi perilaku pengguna, pengumpulan dan preprocessing data, penerapan algoritma *K-Means*, serta visualisasi dan interpretasi hasil klasterisasi [19]- [21]. Rancangan umum tahapan penelitian disajikan pada Gambar 2.



Gambar 2. Tahapan Penelitian Analisis Perilaku Pengguna Internet

Data yang digunakan dikumpulkan dari sistem pencatatan aktivitas jaringan internal perusahaan. Informasi yang terekam mencakup waktu akses, identitas pengguna, departemen, volume data yang dikirim dan diterima (*sent/received*), serta aplikasi yang diakses. Data ini bersifat log historis dan dianonimkan untuk menjaga privasi serta keamanan informasi. Contoh format data yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Contoh Format Data Akses Internet

Date	Time	Level	Source	Username	Gender	Department	Sent/ Received	Application
13-May	07:29:57	notice	10.XXX.6.3	YosexxxxxxLukito	Male	Electrical	840B/288B	Jasper Viewer
13-May	07:29:57	notice	10.XXX.6.9	MirtxxxxxxlinImkotta	Male	Field Control	21.83KB/7.28KB	Microsoft.Portal
13-May	07:29:57	notice	10.XXX.6.9	MirtxxxxxxlinImkotta	Male	Field Control	1.04KB/7.88KB	Microsoft.365.Portal
13-May	07:29:57	notice	10.XXX.6.27	Ahmaxxxxxxillah	Male	Pre-Commissioning	1.14KB/3.27KB	Microsoft.WSER
13-May	07:29:57	notice	10.XXX.6.118	Udiaxxxxxx	Male	Piping & Ducting	730B/9.9KB	Microsoft.Portal
13-May	07:29:57	notice	10.XXX.5.167	VirgxxxxxxaloJr.	Male	Piping & Ducting	4.24KB/7.28KB	Microsoft.Portal
13-May	07:29:57	notice	10.XXX.7.158	JonaxxxxxxidiatmoJonaxxxxxxidiatmo	Male	Pre-Commissioning	252B/252B	Microsoft.365.Portal
13-May	07:29:57	notice	10.XXX.7.24	MahmxxxxxxidzuddinMahmxxxxxxidzuddin	Male	Piping (Tripatra)	41B/132B	Microsoft.Portal

Sebelum dilakukan analisis, data melalui tahap preprocessing, termasuk penghapusan duplikasi, penanganan nilai kosong, dan normalisasi fitur menggunakan metode *Min-Max Scaling* agar seluruh variabel berada pada skala yang seragam. Atribut yang digunakan antara lain: *traffic sent/received*, *application category*, *risk level*, *recency*, dan *monetary value*.

Selanjutnya, algoritma *K-Means* diterapkan untuk mengelompokkan pengguna berdasarkan kesamaan perilaku akses internet. Penentuan jumlah klaster optimal dilakukan menggunakan metode *elbow*, dilanjutkan dengan proses iteratif inisialisasi centroid dan pembaruan posisi hingga tercapai kondisi konvergensi [22]. Hasil klasterisasi divisualisasikan dalam bentuk grafik dua dimensi untuk mengamati penyebaran data antar kelompok.

Analisis ini juga didukung oleh kategorisasi parameter perilaku pengguna sebagaimana dirangkum pada Tabel 2, guna memberikan pemahaman yang lebih mdalam terhadap jenis aktivitas pengguna [23].

Tabel 2. Parameter Kategori Akses Internet

Parameter	Deskripsi
Serious Cyberloafing	Akses ilegal atau berisiko tinggi (misal: situs terlarang, perjudian)
Minor Cyberloafing	Akses hiburan ringan (misal: media sosial, belanja online, *streaming*)
Normal Usage	Akses aplikasi kerja (misal: Zoom, Outlook, Microsoft Office)

Spesifikasi perangkat keras dan perangkat lunak yang digunakan dalam analisis ditampilkan pada Tabel 3.

Tabel 3 Spesifikasi Lingkungan Penelitian

Komponen	Spesifikasi
Sistem Operasi	Windows 10 Pro 64-bit
Bahasa Pemrograman	Python 3.10
RAM	16 GB
Prosesor	Intel Core i7 / AMD Ryzen setara
Library	pandas, sklearn, matplotlib, io

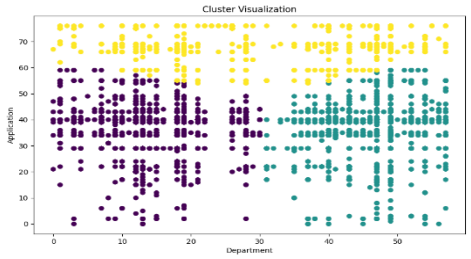
Pengembangan sistem dilakukan secara iteratif menggunakan metode RAD, mulai dari identifikasi kebutuhan, pengembangan prototipe, pengujian awal, hingga implementasi akhir. Hasil dari proses ini berupa sistem analitik berbasis [24] Python yang mampu mengelompokkan pengguna internet ke dalam klaster-klaster dengan karakteristik perilaku yang berbeda. Output sistem disajikan dalam bentuk dashboard visual dan laporan analisis untuk mendukung pengambilan keputusan strategis terkait

pengelolaan sumber daya internet dan kebijakan keamanan jaringan.

4. HASIL DAN PEMBAHASAN

4.1 Hasil

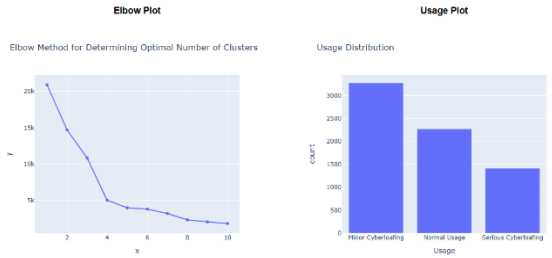
Penelitian ini menghasilkan klasifikasi perilaku pengguna internet di PT Chiyoda International Indonesia menggunakan pendekatan machine learning berbasis algoritma *K-Means Clustering*. Tahapan awal melibatkan pengumpulan data log firewall yang mencatat aktivitas digital pengguna secara rinci, seperti waktu akses, nama pengguna, alamat IP, aplikasi yang digunakan, dan volume data yang dikirim serta diterima. Proses ini memberikan fondasi penting untuk pemetaan perilaku pengguna jaringan.



Gambar 3. Visualisasi Klaster Berdasarkan Department dan Aplikasi

Visualisasi ini menunjukkan hasil *clustering* antara fitur departemen dan aplikasi. Titik-titik pada scatter plot merepresentasikan entri data, sementara warna yang berbeda menandakan klaster yang terbentuk. Pola penyebaran memperlihatkan bahwa pengguna dari departemen tertentu cenderung mengakses aplikasi-aplikasi spesifik, yang dapat membantu perusahaan dalam mengidentifikasi konsentrasi penggunaan teknologi dan potensi penyalahgunaan.

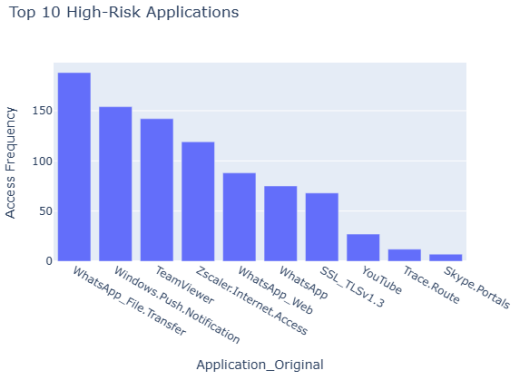
Data selanjutnya dianalisis untuk menentukan jumlah klaster optimal menggunakan metode elbow. Grafik yang dihasilkan menunjukkan penurunan tajam pada nilai WCSS (within-cluster sum of squares) hingga klaster ke-3, yang menandakan bahwa pembagian data ke dalam tiga klaster adalah yang paling representatif.



Gambar 4. Elbow Plot dan Distribusi Kategori Penggunaan

Diagram sebelah kiri menunjukkan grafik elbow, di mana nilai WCSS menurun tajam pada k = 3, menunjukkan jumlah klaster optimal. Diagram kanan menunjukkan jumlah data dalam masing-masing kategori penggunaan yang terbentuk setelah klasterisasi, yaitu *Minor Cyberloafing*, *Normal Usage*, dan *Serious Cyberloafing*. Kategori *Minor Cyberloafing* mendominasi jumlah data, menandakan bahwa aktivitas non-produktif ringan cukup umum di lingkungan kerja.

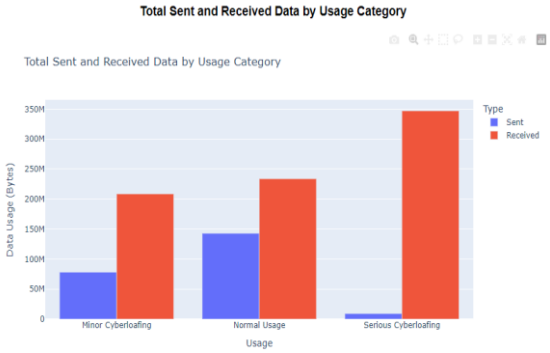
Untuk mendalami konten masing-masing klaster, dilakukan visualisasi terhadap frekuensi penggunaan aplikasi yang tergolong berisiko tinggi.



Gambar 5. Aplikasi dengan Risiko Tinggi (Top 10 High-Risk Applications)

Visualisasi ini menunjukkan sepuluh aplikasi dengan intensitas akses tertinggi yang dinilai berisiko, seperti *WhatsApp File Transfer*, *Windows Push Notification*, dan *TeamViewer*. Aplikasi-aplikasi ini memiliki potensi membuka jalur komunikasi atau transfer data yang tidak selalu dalam pengawasan sistem keamanan perusahaan, sehingga dapat mengancam integritas jaringan internal.

Analisis kemudian dilanjutkan dengan mengukur total data yang dikirim (sent) dan diterima (received) berdasarkan kategori aktivitas pengguna. Hasilnya menunjukkan pola penggunaan data yang bervariasi.



Gambar 6. Total Sent and Received Data berdasarkan Kategori Penggunaan

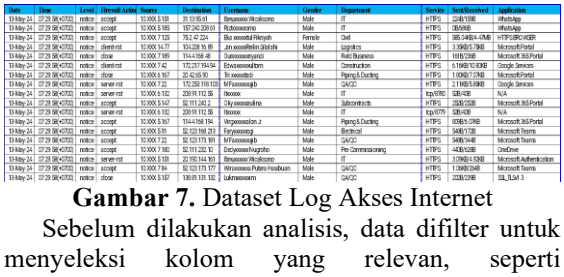
Grafik ini memperlihatkan bahwa *Serious Cyberloafing* menghasilkan data diterima tertinggi, yaitu lebih dari 345 juta byte, namun data terkirim sangat kecil. Sebaliknya, kategori *Normal Usage* menunjukkan keseimbangan dengan data diterima sekitar 230 juta byte dan data terkirim 140 juta byte, mencerminkan aktivitas kerja seperti konferensi video dan kolaborasi daring. Adapun *Minor Cyberloafing* menunjukkan nilai menengah, dengan pola konsumsi data cenderung pasif.

Melalui hasil-hasil visualisasi ini, penelitian berhasil mengungkap gambaran menyeluruh tentang bagaimana pola akses internet mencerminkan tingkat produktivitas dan risiko dalam suatu organisasi. Dengan demikian, temuan ini dapat dijadikan acuan untuk merancang strategi pengelolaan jaringan yang lebih adaptif dan berbasis bukti.

4.2 Pembahasan

Penelitian ini bertujuan untuk menganalisis perilaku pengguna dalam mengakses internet di lingkungan PT Chiyoda International Indonesia dengan memanfaatkan pendekatan machine learning, khususnya algoritma *K-Means Clustering*. Data yang dianalisis berupa log akses internet yang terekam secara otomatis oleh sistem firewall perusahaan. Log ini mencakup berbagai informasi penting seperti tanggal dan waktu akses, nama pengguna, alamat IP, departemen pengguna, aplikasi yang diakses, serta volume data yang dikirim dan diterima.

Gambar 7 memperlihatkan potongan dataset asli yang digunakan dalam penelitian. Data ini menunjukkan aktivitas pengguna secara rinci, memungkinkan penelusuran pola perilaku digital berdasarkan dimensi waktu, volume data, serta jenis aplikasi yang digunakan.



Gambar 7. Dataset Log Akses Internet

Sebelum dilakukan analisis, data difilter untuk menyeleksi kolom yang relevan, seperti 'Department', 'Sent/Received', dan 'Application'. Langkah selanjutnya adalah tahap preprocessing data. Preprocessing mencakup pemisahan kolom 'Sent/Received' menjadi kolom 'Sent' dan 'Received', penanganan nilai kosong dengan nilai default, serta konversi ukuran data ke format numerik. Selain itu, kategori seperti 'Department', 'Application', dan 'Usage' diubah ke format numerik menggunakan *Label Encoding*, agar dapat diproses oleh algoritma machine learning.

Setelah data diproses, dilakukan visualisasi hasil filtering sebagaimana ditampilkan pada Gambar 8. Visualisasi ini menegaskan bahwa data telah disusun dalam format yang konsisten dan siap untuk analisis lebih lanjut. Proses ini sangat penting untuk memastikan kualitas dan integritas data sebelum digunakan dalam klasterisasi.

0	18	66	1.590000e+02
1	18	66	6.860000e+02
2	7	29	4.687135e+06
3	21	40	5.888000e+03
4	13	34	2.360000e+02
5	10	22	1.105920e+04
6	40	40	7.239680e+03
7	47	22	5.990400e+03
8	18	76	4.000000e+01

Gambar 8. Hasil Preprocessing Data

Analisis klaster dilakukan menggunakan algoritma *K-Means*. Grafik elbow menunjukkan bahwa titik siku berada pada k = 3, sehingga model dilatih dengan jumlah klaster sebanyak tiga. Setelah model *K-Means* diterapkan, data dikelompokkan ke dalam tiga klaster. Hasil klasterisasi yang berupa tabel ditampilkan dalam Gambar 9. Setiap baris menunjukkan atribut pengguna, aplikasi, data yang diterima, dan klaster hasil prediksi.

Data Table

	Department	Application	Received	Usage	Cluster
0	18	66	1.590000e+02	0	2
1	18	66	6.860000e+02	0	2
2	7	29	4.687135e+06	0	0
3	21	40	5.888000e+03	0	0
4	13	34	2.360000e+02	0	0
5	10	22	1.105920e+04	0	0
6	40	40	7.239680e+03	0	1
7	47	22	5.990400e+03	0	1
8	18	76	4.000000e+01	0	2

Gambar 9. Hasil Cluster

Gambar data klasterisasi diatas yang ditampilkan menggambarkan hasil pengelompokan berdasarkan variabel departemen, aplikasi, jumlah data yang diterima (*received*), kategori penggunaan (*usage*), dan label klaster akhir. Setiap baris merepresentasikan satu entri log aktivitas internet yang telah dikodekan secara numerik, kemudian dikelompokkan oleh algoritma *K-Means* ke dalam salah satu dari tiga klaster. Misalnya, entri dengan *department* 18 dan aplikasi 66 muncul beberapa kali dan tergolong dalam klaster 2, menunjukkan adanya pola penggunaan yang konsisten dari departemen tersebut terhadap aplikasi tertentu. Variasi nilai *received* juga memperlihatkan bahwa meskipun aplikasi yang digunakan serupa, volume data yang diterima dapat berbeda secara signifikan, yang menunjukkan tingkat intensitas aktivitas yang tidak

seragam. Dengan demikian, tabel ini memberikan landasan kuantitatif yang kuat untuk mengidentifikasi karakteristik tiap klaster dan mendukung pemetaan perilaku digital lintas departemen dalam organisasi.

KESIMPULAN

Hasil analisis data menggunakan metode K-Means menunjukkan bahwa pola penggunaan internet di PT. Chiyoda International Indonesia dapat dikelompokkan ke dalam tiga klaster utama dengan karakteristik berbeda, yaitu berdasarkan intensitas akses data, jenis aplikasi yang digunakan, dan distribusi aktivitas antar departemen. Visualisasi seperti elbow plot dan cluster plot menunjukkan bahwa mayoritas aktivitas pengguna mengarah pada penggunaan aplikasi non-produktif, terutama dalam kategori *Serious Cyberloafing*, yang memiliki potensi risiko terhadap efisiensi dan keamanan jaringan perusahaan. Oleh karena itu, disarankan agar perusahaan mengoptimalkan alokasi sumber daya IT berdasarkan kebutuhan tiap departemen, memperketat pengawasan terhadap aktivitas digital yang mencurigakan, memberikan pelatihan terkait penggunaan aplikasi yang aman dan produktif, serta melakukan evaluasi berkala terhadap kebijakan penggunaan internet untuk memastikan kesesuaian dengan tujuan organisasi dan standar keamanan informasi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada seluruh pihak yang telah memberikan dukungan dalam penyusunan penelitian ini. Ucapan terima kasih disampaikan kepada PT. Chiyoda International Indonesia atas izin dan akses data yang diberikan, sehingga memungkinkan penelitian ini terlaksana dengan baik. Penulis juga berterima kasih kepada dosen pembimbing dan rekan-rekan yang telah memberikan masukan, arahan, serta motivasi selama proses penyusunan. Semoga hasil penelitian ini dapat memberikan kontribusi nyata bagi pengembangan analisis data berbasis machine learning dalam lingkungan organisasi serta bermanfaat untuk referensi keilmuan dan pengambilan kebijakan berbasis data.

DAFTAR PUSTAKA

[1] Al-Muhrami, M.A.S., Alawi, N.A., Alzubi, M., Al-Refaei, A.A.-A., "Affecting the Behavioural Intention to Use Electronic Banking Services Among Users in Yemen: Using an Extension of the Unified Theory of Acceptance and Use of Technology," in *2021 2nd International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, 2021, pp. 257–264, https://doi.org/10.1109/ICSCEE50312.2021.9497917

[2] Badillo, S., Banfai, B., Birzele, F., Davydov, I.I., Hutchinson, L., Kam-Thong, T., Siebourg-Polster, J., Steiert, B., Zhang, J.D., "An Introduction to Machine Learning," *Clinical Pharmacology & Therapeutics*, vol. 107, pp. 871–885, 2020, https://doi.org/10.1002/cpt.1796

[3] Dogan, A., Birant, D., "Machine learning and data mining in manufacturing," *Expert Systems with Applications*, vol. 166, p. 114060, 2021, https://doi.org/10.1016/J.ESWA.2020.114060

[4] Gani, A.G., "Sejarah dan Perkembangan Internet di Indonesia," *Jurnal Mitra Manajemen*, vol. 1, 2020

[5] Govender, P., Sivakumar, V., "Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019)," *Atmospheric Pollution Research*, vol. 11, pp. 40–56, 2020, https://doi.org/10.1016/J.APR.2019.09.009

[6] Hari Rachmawanto, E., Atika Sari, C., Pramono, H., Shinta Sari, W., "Visitor Prediction Decision Support System at Dieng Tourism Objects Using the K-Nearest Neighbor Method," *Journal of Applied Intelligent System*, 2022

[7] Heliyanti Susana, Nana Surana, Faturohman, Kaslani, "Penerapan Model Klasifikasi Metode Naive Bayes terhadap Penggunaan Akses Internet," *JURSISTEKNI*, vol. 7, pp. 25, 2021

[8] Islam, R., Patamsetti, V.V., Gadhi, A., Gondu, R.M., Bandaru, C.M., Kesani, S.C., Abiona, O., "Design and Analysis of a Network Traffic Analysis Tool: NetFlow Analyzer," *International Journal of Communications, Network and System Sciences*, vol. 16, pp. 21–29, 2023, https://doi.org/10.4236/ijcns.2023.162002

[9] Koushik, A.N.P., Manoj, M., Nezamuddin, N., "Machine learning applications in activity-travel behaviour research: a review," *Transport Reviews*, vol. 40, pp. 288–311, 2020, https://doi.org/10.1080/01441647.2019.1704307

[10] Mahmoud Abbasi, Amin Shaharaki, Amir Tahekordi, "Deep Learning for Network Traffic Monitoring and Analysis (NTMA): A Survey," *Elsevier*, pp. 19–41, 2021

[11] Moseley, B., Wang, J.R., "Approximation Bounds for Hierarchical Clustering: Average Linkage, Bisecting K-means, and Local Search," *Journal of Machine Learning Research*, 2023

[12] Muttaqin, M., "Internet Usage Behavior of the ICT Young Workforce in the Border Region,"

Journal Pekommas, vol. 4, p. 11, 2019, https://doi.org/10.30818/jpkm.2019.2040102

[13] Nagari, S.S., Inayati, L., "Implementation of Clustering Using K-Means Method to Determine Nutritional Status," *Jurnal Biometrika dan Kependudukan*, vol. 9, pp. 62–68, 2020, https://doi.org/10.20473/jbk.v9i1.2020.62-68

[14] Nur' Kamisa, Almira Devita P, Dian Novita, "Pengaruh Online Customer Review dan Online Customer Rating terhadap Kepercayaan Konsumen (Studi Kasus: Pengguna Shopee di Bandar Lampung)," *Journal of Economic and Business Research*, vol. 2, pp. 21–29, 2022

[15] Pratap Chandra Sen, Mahiranab Hajra, Mitradu Ghosh, "Supervised Classification Algorithms in Machine Learning: A Survey and Review," *Emerging Technology in Modeling and Graphics*, pp. 99–111, 2020, https://doi.org/10.1023/A:1014043630878

[16] Rabbani, M., Wang, Y.L., Khoshkangini, R., Jelodar, H., Zhao, R., Hu, P., "A hybrid machine learning approach for malicious behaviour detection and recognition in cloud computing," *Journal of Network and Computer Applications*, vol. 151, p. 102507, 2020, https://doi.org/10.1016/J.JNCA.2019.102507

[17] Ran, X., Zhou, X., Lei, M., Tepsan, W., Deng, W., "A novel K-means clustering algorithm with a noise algorithm for capturing urban hotspots," *Applied Sciences*, vol. 11, 2021, https://doi.org/10.3390/app112311202

[18] Retnoningsih, E., Pramudita, R., "Mengenal Machine Learning Dengan Teknik Supervised dan Unsupervised Learning Menggunakan Python," *BINA INSANI ICT JOURNAL*, vol. 7, pp. 156–165, 2020

[19] Sigit Hadi Prayogo, Dana Indra Sensue, "Analisis Usability pada Aplikasi Berbasis Web dengan Mengadopsi Model Kepuasan Pengguna (User Satisfaction)," *IJoICT*, 2020

[20] Sinaga, K.P., Yang, M.S., "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80716–80727, 2020, https://doi.org/10.1109/ACCESS.2020.2988796

[21] Sitorus, Z., "Penerapan Data Mining untuk Clustering Penduduk Miskin di Kota Tanjungbalai Menggunakan Metode Algoritma K-Means," *Journal of Science and Social Research*, 2024

[22] Syed, S., Singh, H., Thangaraju, S., Thangaraju, S.K., Eazreen Bakri, N., Yok Hwa, K., "The Impact of Cyberloafing on Employees' Job Performance: A Review of Literature," *Journal of Advances in Management Sciences & Information Systems*, vol. 6, pp. 16–28, 2020, https://doi.org/10.6000/2371-1647.2020.06.02

[23] van Engelen, J.E., Hoos, H.H., "A survey on semi-supervised learning," *Machine Learning*, vol. 109, pp. 373–440, 2020, https://doi.org/10.1007/s10994-019-05855-6

[24] Verbeeck, N., Caprioli, R.M., Van de Plas, R., "Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry," *Mass Spectrometry Reviews*, 2020, https://doi.org/10.1002/mas.21602