Analisis Interaksi Pemain dalam Komunitas *Grand Theft Auto V Roleplay* (GTAV RP) untuk Klasifikasi *Soft Skill* Menggunakan Metode Data Mining

1st Dito Prasetyo
Program Studi Teknik Informatika
Universitas Telkom Purwokerto
Purwokerto, Indonesia
ditoprasetyo@student.telkomuniversity.ac.id

2nd Iqsyahiro Kresna A
Program Studi Teknik Informatika
Universitas Telkom Purwokerto
Purwokerto, Indonesia
hiroka@telkomuniversity.ac.id

3rd Ipam Fuaddina Adam
Program Studi Teknik Informatika
Universitas Telkom Purwokerto
Purwokerto, Indonesia
ipamya@telkomuniversity.ac.id

***Abstrak* — Komunitas Grand Theft Auto V Roleplay (GTAV RP) menyediakan lingkungan virtual yang rumit di mana para pemain memiliki peran profesional yang berbeda yang mewakili keterampilan lunak, termasuk komunikasi, kerja tim, kepemimpinan, tanggung jawab, dan pengendalian emosi. Hingga saat ini, belum ada metode sistematis untuk mengidentifikasi dan mengkategorikan keterampilan lunak yang ditunjukkan oleh pemain dalam permainan, berdasarkan aktivitas roleplay. Tujuan utama dari penelitian ini adalah untuk menganalisis log aktivitas pemain dan mengklasifikasikan keterampilan lunak dari perspektif penambangan data dengan algoritma Naive bayes dan Support vector machine (SVM). Ekstraksi data dari log aktivitas pemain dari server NUSAV Roleplay dilakukan dengan pra-pemrosesan teks, ekstraksi fitur berdasarkan TF-IDF, dan klasifikasi. Hasilnya mengindikasikan, Naive bayes mencapai akurasi 88% dan Linear SVM mencapai 91,33%, sementara Linear SVM lebih unggul dalam mengklasifikasikan lima kategori keterampilan lunak. Data yang dikumpulkan dari studi ini juga akan memberikan informasi untuk analisis perilaku berdasarkan interaksi digital dan membantu manajer komunitas mengidentifikasi kemungkinan keterampilan lunak yang dimiliki oleh pemain.**

Kata kunci*— *Grand Theft Auto V Roleplay (GTAV RP), soft skill, data mining, naive bayes, klasifikas

I. PENDAHULUAN

Perkembangan teknologi informasi mengakibatkan meningkatnya jumlah data yang dihasilkan dari berbagai aktivitas digital. Kondisi tersebut mendorong pemanfaatan data mining sebagai metode dalam menemukan pola dan informasi krusial dari himpunan data berukuran besar, termasuk pada analisis perilaku digital di platform seperti e-commerce, media sosial, dan game online [1]. Dalam konteks analisis perilaku, data mining dapat dimanfaatkan untuk mengidentifikasi soft skill individu berdasarkan pola interaksi digital yang dilakukan pengguna. Soft skill seperti komunikasi, kerja sama tim, penyelesaian masalah, dan kepemimpinan merupakan kemampuan non-teknis yang memiliki peran penting dalam dunia profesional. Penelitian sebelumnya mengindikasikan, teknik data mining dan natural language processing (NLP) mampu dipergunakan sebagai cara mengidentifikasi soft skill melalui data tekstual secara otomatis [2]. Salah satu sumber data yang memiliki potensi untuk dianalisis yaitu komunitas Grand Theft Auto V Roleplay (GTAV RP), di mana pemain melakukan interaksi sosial layaknya kehidupan nyata dan seluruh aktivitas pemain tersimpan dalam bentuk data log [3].

II. KAJIAN TEORI

A. Game Grand Theft Auto V & Roleplay

1. *Grand Theft Auto V (GTA V) – Rockstar Games*

Grand Theft Auto V merupakan permainan bergenre *actionadventure* yang didesain oleh *Rockstar North* dan dipublikasikan oleh *Rockstar Games* pada tahun 2013 [14]. Permainan ini menjadi seri ketujuh dalam *Franchise Grand Theft Auto* dan awalnya dirilis untuk *platform PlayStation* 3 serta Xbox 360 sebel akhirnya tersedia pada *PlayStation* 4, *Xbox One*, dan *Windows*.



GAMBAR 1
(Logo *Grand Theft Auto V Roleplay* (GTA V RP))



GAMBAR 1
(*Gameplay Grand Theft Auto V Roleplay* (GTA V RP))

2. *FiveM* Sebagai *Platform* Modifikasi
   FiveM merupakan sebuah *multiplayer modification framework* untuk *Grand Theft Auto V* yang dikembangkan oleh *CitizenFX Collective* dan berada di bawah *platform* Cfx.re. *Platform* ini memungkinkan pemain untuk bermain pada server khusus (*dedicated server*) yang dapat dimodifikasi secara bebas dan terpisah dari server resmi GTA *Online* milik *Rockstar Games* [16]. *FiveM* mulai dikembangkan sejak tahun 2014 dan pada tahun 2023 resmi diakuisisi oleh *Rockstar Games*.



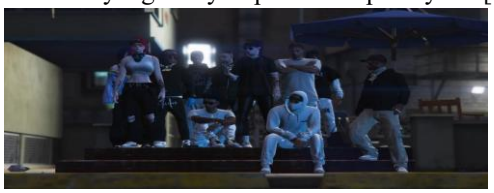
GAMBAR 3
(Logo *FiveM*)

3. Perbedaan *Grand Theft Auto V* dan *Grand Theft Auto V Roleplay*
   Grand Theft Auto V yang dikembangkan oleh *Rockstar Games* dan *Grand Theft Auto V Roleplay* yang berjalan di *platform FiveM* memiliki perbedaan fundamental dalam hal tujuan permainan, mekanik gameplay, dan interaksi pemain [17]. Perbedaan-perbedaan tersebut bisa disimak melalui Tabel 1 berikut:

TABEL 1
(Perbedaan GTA V dan GTA V RP)

Aspek	GTA V Original (Rockstar Games)	GTA V Roleplay (FiveM)
Platform	GTA Online resmi dari Rockstar Games	FiveM (*modifikasi multiplayer*)
Tujuan Permainan	Menyelesaikan misi, *heist*, dan aktivitas kriminal untuk memperoleh uang dan ekspansi bisnis	Memainkan peran karakter dengan identitas, pekerjaan, dan kehidupan sosial yang realistis
Fokus Gameplay	Berorientasi pada aksi seperti menembak, mengemudi, misi *heist,* dan kompetisi PvP	Berfokus pada *roleplay,* interaksi sosial, simulasi kehidupan, dan pengembangan karakter

4. *Grand Theft Auto V Roleplay* (GTAV RP)
   Grand Theft Auto V Roleplay atau GTA V RP merupakan mode permainan *multiplayer* berbasis *FiveM* yang mengubah konsep permainan asli *Grand Theft Auto V* menjadi pengalaman *roleplay* yang lebih realistis dan interaktif. Berbeda dengan GTA V original yang berfokus pada misi serta aksi kriminal, GTAV RP memungkinkan pemain untuk menciptakan dan memerankan karakter dengan identitas, pekerjaan, serta latar kehidupan masing-masing dalam sebuah dunia virtual yang menyerupai kehidupan nyatA [18].



GAMBAR 4
(*Grand Theft Auto V Roleplay* (GTA V RP))

5. *Command Roleplay* (*/me* dan */do*)
   Command roleplay merupakan perintah berbasis teks yang dipergunakan oleh pemain dalam *server roleplay* untuk menggambarkan tindakan, ekspresi, maupun kondisi karakter yang tidak dapat ditampilkan secara langsung melalui animasi bawaan permainan [19].

B. Konsep *Data Mining*
   Data mining ialah sebuah prosedur yang dapat digunakan menelusuri informasi yang berguna dari himpunan data yang berskala besar. Melalui pemanfaatan teknik statistik dan kecerdasan buatan data mining dapat secara efisien memproses pencarian informasi yang penting dari data yang besar[20][5]



GAMBAR 5
(Konsep *Data Mining*)

C. Algoritma *Naive Bayes*

1. Naive bayes ialah algoritma klasifikasi berbasis probabilistik yang kerap dipergunakan pada bidang klasifikasi teks, diagnosis, dan sistem prediksi perilaku [5]. Hal ini membuktikan bahwa algoritma ini Gambar 2.2 Gameplay Grand Theft Auto V Roleplay (GTA V RP) efektif dalam mengenali pola yang tersembunyi dari interaksi manusia dan memiliki potensi untuk dapat digunakan dalam klasifikasi soft skill berbasis data interaksi, yang sebagaimana dilakukan dalam penelitian ini [22]
2. Dengan menghitung probabilitas dari masing masing kelas algoritma Naive bayes dapat bekerja berdasarkan distribusi frekuensi dari fitur-fitur input yang diamati. Kesederhanaan dalam penerapan, kecepatan pemrosesan, dan efektivitas dengan menangani data yang berkategori membuat algoritma ini mempunyai kelebihan yang sangat efektif.

D. Algoritma *Support Vector Machine* (*SVM*)

SVM termasuk metode klasifikasi yang tergolong kategori supervised learning. Algoritma ini beroperasi melalui pencarian batas pemisah terbaik atau hyperplane guna memisahkan data dari kelas yang berbeda [23].

Secara matematis, fungsi keputusan pada SVM bisa diformulasikan berupa:

$f(x) = w \cdot x + b$ (2.1)

Catatan:

a. w merupakan vektor bobot
b. x merupakan vektor fitur dari data masukan
c. b merupakan nilai bias

Apabila nilai f(x) lebih besar atau sama dengan nol, maka data akan diklasifikasikan ke dalam kelas positif. Sebaliknya, apabila nilai f(x) kurang dari nol, maka data akan diklasifikasikan ke dalam kelas negatif.

E. Text Preprocessing

Text preprocessing merupakan tahapan awal pada proses text mining yang ditujukan guna membersihkan serta mengonversi data teks mentah ke bentuk yang lebih sistematis yang menjadikannya bisa diolah oleh algoritma machine learning.

TF-IDF

TF-IDF (*Term Frequency–Inverse Document Frequency*) ialah teknik ekstraksi fitur guna memeriksa krusialnya sebuah kata pada sebuah dokumen di keseluruhan himpunan dokumen (korpus). Pendekatan ini mempergunakan dua parameter untuk ekstraksi data: *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). Perhitungan *Term Frequency* (TF) dapat mempergunakan persamaan berikut:

TF(t,d) = f(t,d) / Σf(t',d)

Catatan:

a. f(t,d) = jumlah kemunculan term t pada dokumen d
b. Σf(t',d) = total seluruh term pada dokumen d

Sementara itu, nilai Inverse Document Frequency (IDF) dihitung dengan persamaan:

IDF(t) = log(N / df(t))

Catatan:

a. N = jumlah seluruh dokumen dalam corpus
b. df(t) = jumlah dokumen yang mengandung term t

Setelah nilai TF dan IDF diperoleh, maka nilai akhir TF-IDF dihitung dengan mengalikan kedua komponen tersebut:

TF-IDF(t,d) = TF(t,d) × IDF(t)

Tabel 2 Contoh Perhitungan TF-IDF

Dataset

Dokumen 1 : "memimpin briefing operasi tim"
Dokumen 2 : "membantu rekan angkat kotak"
Dokumen 3 : "memimpin tim operasi lapangan"

TABEL 2
(Contoh Perhitungan *TF-IDF* Memimpin)

Komponen	Rumus	Perhitungan	Hasil
TF	$f(t, d)$ / $\Sigma f(t', d)$	1/4	0.25
DF	Jumlah dokumen dengan *term*	Dokumen 1, Dokumen 3	2
IDF	$log(N / df(t))$	$log(3 / 2)$	0.176
TF-IDF	TF × IDF	0.25 × 0.176	0.044

F. *Soft Skill* dan Kategorinya

1. *Soft skill* merupakan keterampilan nonteknis yang berkaitan dengan kemampuan interpersonal dan karakter individu dalam berinteraksi dengan orang lain. Keterampilan ini meliputi komptensi komunikasi, kerja sama tim, kepemimpinan, kompetensi adaptasi, serta pemecahan masalah yang berperan penting dalam mendukung keberhasilan seseorang di lingkungan *professional.*
2. Kategori *Soft Skill* dalam Penelitian

   Penelitian ini mengklasifikasikan *soft skill* pemain *GTA V Roleplay* ke dalam lima kategori utama berdasarkan interaksi dan aktivitas mereka dalam game. Kelima kategori tersebut adalah *Communication* (Komunikasi), *Teamwork* (Kerja Sama Tim), *Leadership* (Kepemimpinan), *Responsibility* (Tanggung Jawab), dan *Emotional Control* (Pengendalian Emosi).

G. Perbandingan *Naive Bayes* dan *Support Vector Machine* (*SVM*)

Metode klasifikasi dalam data mining memiliki berbagai pendekatan yang dapat digunakan dalam pengolahan data teks, di antaranya ialah algoritma Naive bayes dan SVM [30]. Kedua metode ini tergolong berkategori supervised learning yang kerap dipergunakan pada proses klasifikasi teks. *Naive bayes* dikenal sebagai metode probabilistik yang sederhana, sedangkan SVM merupakan metode yang berbasis pada pencarian hyperplane maksimal guna memisahkan data antarkelas [22].

H. Karakter Pemain dalam *Roleplay* dan Hubungan dengan *Soft Skill*

Setiap pemain dalam komunitas *Grand Theft Auto V Roleplay* (GTAV RP) memilih dan mengembangkan karakter dengan latar belakang dan kepribadian berbeda, mulai dari tokoh penegak hukum seperti polisi atau pemadam kebakaran, hingga karakter dengan peran bisnis, medis, atau bahkan kriminal [28]. Karakter-karakter ini tidak sebatas mencerminkan gaya bermain, namun pula menuntut keterampilan sosial tertentu. Misalnya, seorang pemain yang berperan sebagai kepala kepolisian harus menunjukkan kemampuan kepemimpinan, komunikasi taktis, dan penyelesaian masalah secara cepat [13].

III. METODE

Metode penelitian ini menggunakan subjek berupa pemain aktif dan pengelola server pada komunitas Grand Theft Auto V Roleplay (GTAV RP) yang tergabung dalam server NUSAV Roleplay. Pemilihan subjek dilakukan secara purposive berdasarkan tingkat keterlibatan dan aktivitas mereka dalam permainan roleplay. Subjek penelitian terdiri atas pemain aktif yang secara rutin menjalankan aktivitas roleplay serta menggunakan perintah interaksi seperti /me dan /do dalam permainan, serta admin atau pengelola server yang memberikan izin penggunaan data log aktivitas dan berperan sebagai narasumber dalam proses wawancara untuk validasi hasil sistem. Adapun objek penelitian berupa data log aktivitas pemain yang tercatat secara otomatis oleh sistem server NUSAV Roleplay, meliputi penggunaan perintah

roleplay /me dan /do yang merepresentasikan tindakan dan deskripsi karakter pemain, interaksi antar pemain yang terekam dalam sistem log server, serta keterlibatan pemain dalam berbagai event komunitas yang diselenggarakan di server. Data log tersebut digunakan sebagai dasar untuk mengidentifikasi dan mengklasifikasikan kemampuan soft skill pemain ke dalam kategori Communication, Teamwork, Leadership, Responsibility, dan Emotional Control menggunakan metode data mining. Dalam pelaksanaannya, penelitian ini didukung oleh berbagai alat dan bahan yang menunjang proses pengumpulan, pengolahan, serta implementasi sistem klasifikasi soft skill. Perangkat yang digunakan mencakup perangkat lunak untuk pengolahan data dan pengembangan sistem, serta perangkat keras yang mendukung proses komputasi dan pengujian. Sementara itu, bahan penelitian terdiri atas data log aktivitas pemain dalam format CSV yang memuat perintah /me dan /do beserta timestamp dan identitas pemain, hasil wawancara dengan admin server NUSAV Roleplay sebagai data pendukung validasi sistem, serta hasil kuesioner yang diperoleh dari pemain aktif server NUSAV Roleplay untuk mengevaluasi persepsi pengguna terhadap sistem yang dikembangkan.

Penelitian ini diselenggarakan dengan terstruktur mengikuti alur yang telah dirancang agar setiap tahapan berjalan terarah dan terstruktur. Diagram alir penelitian bisa disimak melalui Bagian ini menjelaskan tahapantahapan penelitian yang terdapat pada diagram alir penelitian secara lebih rinci. Uraian difokuskan pada proses teknis dan metodologis yang dilaksanakan dalam studi, dimulai dari tahap pengumpulan data aktivitas pemain hingga implementasi sistem klasifikasi soft skill berbasis web. Setiap tahapan dirancang secara sistematis agar mampu menghasilkan model klasifikasi yang akurat dalam mengidentifikasi kemampuan soft skill pemain berdasarkan aktivitas roleplay yang dilakukan pada server NUSAV Roleplay. Tahap pertama penelitian adalah pengumpulan data log aktivitas pemain. Proses ini dilakukan dengan mengambil data secara langsung dari sistem server NUSAV Roleplay setelah memperoleh izin resmi dari pihak pengelola server sebagai bentuk persetujuan penggunaan data untuk kepentingan penelitian akademik. Data log yang digunakan merupakan catatan aktivitas pemain yang tersimpan secara otomatis oleh sistem server selama proses roleplay berlangsung. Data tersebut memuat berbagai aktivitas pemain, seperti penggunaan perintah roleplay, interaksi dengan pemain lain, serta partisipasi dalam berbagai kegiatan komunitas yang diselenggarakan di dalam server. Seluruh data yang diperoleh kemudian digunakan sebagai sumber utama dalam proses analisis dan klasifikasi soft skill pemain.

Selain pengumpulan data berupa log aktivitas, penelitian ini juga menggunakan metode wawancara dengan pengelola server NUSAV Roleplay. Wawancara bertujuan memperoleh pemahaman lebih mendalam mengenai pola interaksi pemain serta memvalidasi hasil klasifikasi yang dihasilkan sistem. Pelaksanaannya dilakukan secara semiterstruktur melalui platform Discord dengan narasumber dari kalangan admin server. Pemilihan narasumber menggunakan teknik purposive sampling berdasarkan pengalaman dan keterlibatan mereka dalam pengelolaan komunitas roleplay. Data hasil wawancara digunakan sebagai pendukung untuk memperkuat interpretasi penelitian.

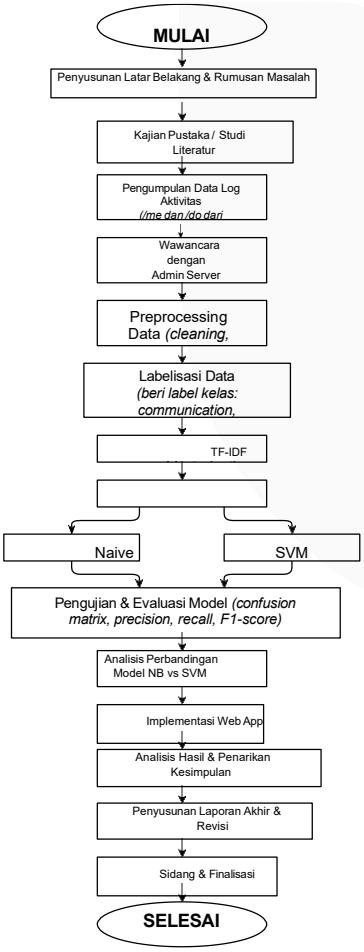
Setelah pengumpulan data selesai, dilakukan tahap preprocessing untuk membersihkan dan menyiapkan data sebelum proses klasifikasi. Tahap ini meliputi pembersihan data, normalisasi teks, penghapusan karakter khusus dan stopwords, serta penyeragaman format agar data siap digunakan pada tahap analisis berikutnya.

Tahapan selanjutnya adalah labelisasi data. Pada tahap ini setiap data log aktivitas pemain diberikan label soft skill berdasarkan isi aktivitas yang terekam dalam sistem. Proses pelabelan dilakukan dengan mengacu pada indikator soft skill yang telah ditentukan, yaitu Communication, Teamwork, Leadership, Responsibility, dan Emotional Control. Label yang diberikan berfungsi sebagai data target (target class) dalam proses pelatihan model machine learning sehingga algoritma dapat mempelajari pola-pola aktivitas yang merepresentasikan masing-masing kategori soft skill. Setelah data berhasil diberikan label, tahap berikutnya adalah transformasi data teks ke dalam bentuk numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Tahap ini diperlukan karena algoritma machine learning tidak dapat memproses data teks secara langsung. TF-IDF bekerja dengan memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen serta tingkat keunikannya terhadap keseluruhan dokumen. Semakin sering suatu kata muncul dalam dokumen tertentu dan semakin jarang muncul pada dokumen lain, maka bobot kata tersebut akan semakin tinggi. Sebagai contoh, pada aktivitas pemain yang berisi kalimat “memberikan arahan kepada anggota tim”, kata “arahan” dan “tim” akan memperoleh bobot tertentu berdasarkan perhitungan TF dan IDF yang selanjutnya digunakan sebagai representasi numerik dalam proses klasifikasi.

Tahap pembangunan model klasifikasi dilakukan setelah data berhasil ditransformasikan ke dalam bentuk matriks TF-IDF. Pada penelitian ini digunakan dua algoritma klasifikasi, yaitu Naive Bayes dan Support Vector Machine (SVM). Algoritma Naive Bayes merupakan metode klasifikasi yang bekerja berdasarkan Teorema Bayes dengan menghitung probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan kemunculan fitur-fitur yang terdapat pada data tersebut. Dalam penelitian ini, probabilitas kelas ditentukan berdasarkan kemunculan kata-kata pada log aktivitas pemain. Algoritma ini menghitung peluang setiap data termasuk ke dalam masing-masing kategori soft skill, kemudian memilih kelas dengan nilai probabilitas tertinggi sebagai hasil klasifikasi. Sebagai ilustrasi, apabila hasil perhitungan menunjukkan bahwa probabilitas kelas Leadership lebih tinggi dibandingkan kategori lainnya, maka aktivitas pemain tersebut akan diklasifikasikan sebagai Leadership. Selain Naive Bayes, penelitian ini juga menerapkan algoritma Support Vector Machine (SVM). SVM merupakan algoritma klasifikasi yang bekerja dengan mencari hyperplane optimal yang mampu memisahkan data antar kelas secara maksimal. Pada penelitian ini digunakan kernel linear karena data hasil TFIDF memiliki dimensi fitur yang tinggi sehingga lebih sesuai diproses menggunakan pendekatan linear. Fungsi keputusan SVM dihitung berdasarkan kombinasi antara vektor bobot, vektor fitur, dan nilai bias. Karena penelitian ini memiliki lima kategori soft skill, maka strategi klasifikasi yang digunakan adalah One-vs-Rest (OvR), yaitu membangun satu model klasifikasi untuk setiap kategori. Hasil akhir ditentukan berdasarkan model yang

menghasilkan nilai fungsi keputusan tertinggi sehingga data dapat dimasukkan ke dalam kategori soft skill yang paling sesuai.

Setelah proses pelatihan model selesai dilakukan, tahap berikutnya adalah pengujian dan evaluasi model. Tahap ini bertujuan untuk mengetahui tingkat performa algoritma Naive Bayes dan SVM dalam mengklasifikasikan soft skill pemain berdasarkan data log aktivitas roleplay. Evaluasi dilakukan menggunakan confusion matrix yang menampilkan perbandingan antara hasil prediksi model dan data aktual pada setiap kategori soft skill. Confusion matrix digunakan untuk mengetahui jumlah prediksi yang benar maupun salah sehingga dapat memberikan gambaran yang jelas mengenai kemampuan model dalam melakukan klasifikasi. Berdasarkan hasil confusion matrix, dilakukan perhitungan beberapa metrik evaluasi untuk mengukur performa model secara lebih komprehensif. Metrik yang digunakan dalam penelitian ini meliputi accuracy, precision, recall, dan F1-score. Accuracy digunakan untuk mengukur tingkat ketepatan klasifikasi secara keseluruhan. Precision digunakan untuk mengetahui seberapa tepat model dalam memberikan prediksi terhadap suatu kelas tertentu. Recall digunakan untuk mengukur kemampuan model dalam menemukan seluruh data yang benar-benar termasuk dalam suatu kelas. Sementara itu, F1-score merupakan kombinasi harmonis antara precision dan recall yang memberikan gambaran keseimbangan performa model.



GAMBAR 6
(*Flowchart* Diagram Alir Penelitian)

Tahap berikutnya adalah analisis perbandingan model. Analisis ini dilakukan dengan membandingkan performa algoritma Naive Bayes dan SVM berdasarkan nilai accuracy, precision, recall, dan F1-score yang diperoleh pada tahap evaluasi. Selain itu, perbandingan juga dilakukan terhadap aspek kecepatan komputasi dan kompleksitas model. Naive Bayes umumnya memiliki proses komputasi yang lebih sederhana dan cepat karena menggunakan pendekatan probabilistik, sedangkan SVM memiliki tingkat kompleksitas yang lebih tinggi namun sering kali menghasilkan performa klasifikasi yang lebih baik pada data berdimensi tinggi. Hasil perbandingan ini digunakan untuk menentukan algoritma yang paling sesuai dalam proses klasifikasi soft skill pemain pada server NUSAV Roleplay. Tahap selanjutnya adalah analisis kebutuhan sistem web application yang akan digunakan sebagai media implementasi model klasifikasi. Analisis kebutuhan dilakukan untuk mengidentifikasi fungsi-fungsi yang harus tersedia pada sistem agar dapat memenuhi kebutuhan pengguna. Kebutuhan fungsional sistem meliputi kemampuan memilih server yang tersedia pada halaman utama, melakukan pencarian dan analisis soft skill berdasarkan nama pemain, menampilkan hasil klasifikasi soft skill dominan, menampilkan daftar log aktivitas beserta kategori soft skill pada setiap aktivitas, menyediakan halaman database untuk melihat riwayat analisis seluruh pemain, serta menyimpan hasil analisis secara otomatis ke dalam basis data. Selain kebutuhan fungsional, sistem juga dirancang untuk memenuhi kebutuhan nonfungsional berupa performa yang cepat, kemudahan penggunaan antarmuka, ketersediaan akses melalui *browser* tanpa instalasi tambahan, serta keamanan data agar informasi pemain hanya digunakan untuk kepentingan penelitian. Sistem yang dikembangkan pada penelitian ini dirancang untuk digunakan oleh dua jenis pengguna, yaitu pemain dan admin *server*. Pemain dapat menggunakan sistem untuk mengetahui kecenderungan soft skill yang dimiliki berdasarkan aktivitas *roleplay* yang dilakukan selama bermain. Sementara itu, admin server dapat memanfaatkan sistem sebagai alat bantu evaluasi pemain secara objektif dan berbasis data. Informasi yang dihasilkan oleh sistem dapat digunakan untuk mendukung proses pengambilan keputusan dalam komunitas, seperti penentuan peran pemain, pembentukan tim.

IV. HASIL DAN PEMBAHASAN

A. Hasil

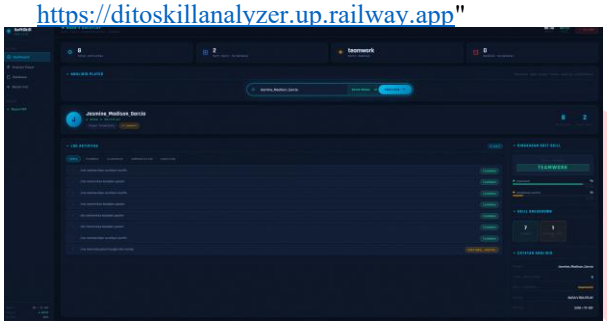
1. Implementasi Sistem

   Di tahapan ini, sistem yang telah dirancang berhasil diimplementasikan dalam bentuk aplikasi berbasis web yang dipergunakan sebagai cara menganalisis *soft skill* pemain *GTA V Roleplay*. Sistem ini mengintegrasikan model klasifikasi dengan antarmuka pengguna, sehingga proses analisis dapat dilakukan secara langsung melalui tampilan yang interaktif dan mudah digunakan.



GAMBAR 7
(Tampilan *Landingpage*)

"Sistem telah berhasil di-hosting secara online dan dapat diakses melalui URL: https://ditoskillanalyzer.up.railway.app"



GAMBAR 7
(Tampilan Halaman *Dasboard*)

Selanjutnya, sistem menyediakan fitur analisis player yang memungkinkan pengguna memasukkan nama pemain untuk dianalisis. Hasil analisis kemudian ditampilkan dalam bentuk daftar aktivitas beserta kategori soft skill yang berhasil diidentifikasi oleh sistem, sebagaimana tersaji melalui Gambar 8.



GAMBAR 8
(Tampilan Fitur Analisis *Player*)

Di samping hal tersebut, sistem juga menyediakan halaman database yang berfungsi untuk menyimpan serta menampilkan riwayat hasil analisis yang telah dilakukan sebelumnya secara terstruktur, sebagaimana tersaji melalui Gambar 9.



GAMBAR 9
(Tampilan Halaman *Database*)

Secara keseluruhan, sistem yang didesain bisa beroperasi secara optimal dan mampu menampilkan hasil analisis secara cepat serta mudah dipahami oleh pengguna.

2. Arsitektur Sistem

Arsitektur sistem penelitian ini mempergunakan konsep client server, di mana pengguna berinteraksi melalui browser dan seluruh permintaan diproses oleh backend berbasis *Flask.*

a) Implementasi *Backend*

Implementasi *backend* pada sistem ini dirancang mempergunakan *framework flask* yang berbasis bahasa pemrograman *Python*.

b) Hasil pengujian

Di tahapan ini dilakukan pengujian terhadap dua algoritma klasifikasi, yaitu *Naive Bayes* dan *Support vector Machine* (SVM), guna membandingkan kinerja masing-masing metode dalam proses klasifikasi *soft skill* pemain berdasarkan data log aktivitas pada server NUSAV *Roleplay*.

c) Confusion Matrix

Confusion matrix ialah tabel yang dipergunakan sebagai cara mengevaluasi performa model klasifikasi melalui cara memperbandingkan hasil prediksi terhadap label aktual pada data uji.

d) Hasil Evaluasi Model

Evaluasi kinerja model dilakukan mempergunakan metrik precision, recall, dan F1-score pada tiap-tiap kelas soft skill. Hasil pengujian dari kedua algoritma kemudian ditampilkan secara berdampingan pada Tabel 3.

TABEL 3
(Hasil Evaluasi Model)

Kelas	Precision (NB)	Precision (SVM)	Recall (NB)	Recall (SVM)	F1-score (NB)	F1-score (SVM)
Communication	100%	100%	85.7%	85.7%	92.3%	92.3%
Emotional Control	100%	83.3%	70.0%	100%	82.3%	90.9%
Leadership	100%	100%	80.0%	80.0%	88.9%	88.9%
Responsibility	87.5%	87.5%	100%	100%	93.3%	93.3%
Teamwork	75.0%	100%	100%	91.7%	85.7%	95.7%

e) Akurasi Keseluruhan Model

Performa model dievaluasi menggunakan metrik akurasi untuk mengetahui ketepatan klasifikasi data uji. Hasil akurasi kedua metode ditampilkan pada Tabel 4.

TABEL 4
(Akurasi Keseluruhan Model)

MODEL	*ACCURACY*
Naive Bayes	87.80%
Support Vector Machine (SVM)	92.68%

Berdasarkan Tabel 4.2, metode SVM memperoleh nilai akurasi keseluruhan sebesar 92.68%, lebih tinggi dibanding metode Naive bayes yang memperoleh akurasi sebesar 87.80%. Perbedaan nilai akurasi sebesar 4.88% menunjukkan SVM memiliki kemampuan generalisasi yang lebih baik terhadap data baru yang belum pernah diproses sebelumnya.

f) Analisis Perbandingan

Hasil pengujian menunjukkan metode SVM berperforma lebih unggul dibanding Naive bayes

dalam mengklasifikasikan soft skill pemain berdasarkan data log aktivitas. Dapat dilihat dari aspek berikut.

Akurasi yang Lebih Tinggi

Metode SVM memperoleh tingkat akurasi sebesar 92.68%, sedangkan Naive bayes memperoleh akurasi sebesar 87.80%. Hasil tersebut mengindikasikan, SVM memiliki kemampuan yang lebih baik dalam memprediksikan kelas secara keseluruhan.

Performa yang lebih konsisten pada data kompleks

SVM memperlihatkan performa yang lebih stabil pada kelas dengan pola data yang lebih kompleks, terutama pada kelas *Emotional Control* dan *Teamwork*. Pada kedua kelas tersebut, SVM menghasilkan nilai *F1-score* yang lebih tinggi dibanding *Naive bayes.*

Kemampuan menangani data berdimensi tinggi

Metode SVM mampu menangani data berdimensi tinggi dengan baik karena dapat menemukan *hyperplane* optimal untuk memisahkan kelas-kelas yang memiliki karakteristik serupa.

B. Pembahasan

Pada bagian ini dilaksanakan bahasan pada hasil implementasi dan pengujian sistem yang telah dipaparkan pada subbab sebelumnya. Pembahasan difokuskan pada interpretasi hasil klasifikasi serta makna yang dapat diambil dari performa *model* yang dipergunakan.

1. Analisis Hasil

   Mengacu temuan implementasi dan pengujian sistem yang telah dilakukan, sistem klasifikasi *soft skill* pemain berbasis data log aktivitas pada komunitas GTA V *Roleplay* mampu bekerja dengan baik dalam mengidentifikasi kategori *soft skill* dari setiap aktivitas pemain. Sistem ini memanfaatkan pendekatan *text mining* dengan tahapan *preprocessing*, transformasi TF-IDF, serta penerapan *model* klasifikasi *Naive bayes* dan SVM.

2. Keterbatasan Sistem

   Dalam pengembangan sistem klasifikasi soft skill berbasis data log aktivitas pemain GTA V Roleplay, terdapat beberapa keterbatasan yang perlu diperhatikan sebagai bahan evaluasi dan pengembangan lebih lanjut. Pertama, performa sistem sangat bergantung pada kualitas dan jumlah dataset yang dipergunakan dalam proses pelatihan model.

   Kedua, karakteristik data berupa teks log aktivitas pemain yang tidak terstruktur menyebabkan adanya ambiguitas dalam proses klasifikasi. Ketiga, sistem belum sepenuhnya memahami konteks percakapan secara menyeluruh.

V. KESIMPULAN

Mengacu temuan penelitian yang telah dilakukan, bisa diambil simpulan sistem klasifikasi soft skill pemain pada komunitas Grand Theft Auto V Roleplay (GTAV RP) berhasil dikembangkan dengan memanfaatkan pendekatan data mining yang berfokus pada analisis data log aktivitas pemain. Sistem yang dirancang mampu mengolah data teks yang berasal dari interaksi pemain, seperti penggunaan perintah /me dan /do, melalui beberapa tahapan, yaitu preprocessing, pembobotan menggunakan TF-IDF, serta proses klasifikasi dengan algoritma *Naive bayes* dan *Support Vector Machine* (SVM).

REFERENSI

[1] Y. Yang and H. Yu, "E-commerce user behavior analysis based on data mining," *2025 3rd International Conference on Internet of Things and Cloud Computing Technology (IoTCCT 2025), September 2729, 2025, Haikou, China*, vol. 1, 2025, doi: 10.1145/3776865.

[2] C. Panzaru and A. Grama, "Towards explicit soft skills labelling in ESCO through semantic NLP analysis," *J. Labour Mark. Res.*, vol. 59, no. 1, Dec. 2025, doi: 10.1186/s12651-025-00409-x.

[3] V. Dumblekar, J. Dubey, and U. Dhar, "An Empirical Study of the Constituent Factors of Player Behaviour," *Simul. Gaming*, vol. 56, no. 1, pp. 117–136, Feb. 2025, doi: 10.1177/10468781241297283.

[4] E. Venkatesan, G. Lecturer, and V. Thangavel, "Comparative study of naïve Bayes and SVM algorithms for text mining using natural language processing," *International Journal of Cloud Computing and Database Management*, vol. 6, no. 2, pp. 90–92, Jul. 2025, doi: 10.33545/27075907.2025.v6.i2b.111.

[5] C. Wu and C. H. Chang, "How multiplayer online games can yield positive effects on individual gamers, gaming companies, and society as a whole," *Humanit. Soc. Sci. Commun.*, vol. 12, no. 1, Dec. 2025, doi: 10.1057/s41599-025-04690-6.

[6] R. Mulla *et al.*, "Predicting Player Churn in the Gaming Industry: A Machine Learning Framework for Enhanced Retention Strategies," *J. Curr. Sci. Technol.*, vol. 15, no. 2, Apr. 2025, doi: 10.59796/jcst.V15N2.2025.103.

[7] E. A. Romero-Mendez, P. C. Santana-Mancilla, M. Garcia-Ruiz, O. A. Montesinos-López, and L. E. Anido-Rifón, "The Use of Deep Learning to Improve Player Engagement in a Video Game through a Dynamic Difficulty Adjustment Based on Skills Classification," *Applied Sciences (Switzerland)*, vol. 13, no. 14, Jul. 2023, doi: 10.3390/app13148249.

[8] W. Lu, J. Griffin, T. D. Sadler, J. Laffey, and S. P. Goggins, "Game-Based Learning Prediction Model Construction: Toward Validated Stealth Assessment Implementation," *Journal of Learning Analytics*, vol. 12, no. 1, pp. 293–321, Mar. 2025, doi: 10.18608/jla.2025.8105.

[9] A. Marengo, A. Pagano, and V. Santamato, "A machine learning framework for soft skills assessment: Leveraging serious games in higher education," *Computers and Education: Artificial Intelligence*, vol. 9, Dec. 2025, doi: 10.1016/j.caeai.2025.100469.

[10] S. Fareri, N. Melluso, F. Chiarello, and G. Fantoni, "SkillNER: Mining and mapping soft skills from any text," *Expert Syst. Appl.*, vol. 184, Dec. 2021, doi: 10.1016/j.eswa.2021.115544.

[11] V. Arinal and B. S. Purnomo, "Optimasi Metode Decision Tree Menggunakan Particle Swarm Optimization Untuk Analisis Sentimen Review Game GTA V Roleplay," *Jurnal Sains dan Teknologi*, vol. 5,

no. 1, pp. 457–461, 2023, doi: 10.55338/saintek.v5i1.1371.
[12] A. S. Hidayat, J. H. Agnes Silvia, and J. A. Wempi, "Perilaku Konsumtif Subkultur Pemuda Pada Game Grand Theft Auto Roleplay 'Executive Roleplay,'" *Wacana: Jurnal Ilmiah Ilmu Komunikasi*, pp. 226–240, Dec. 2024, doi: 10.32509/wacana.v23i2.3708.
[13] E. M. Ghozail and A. Asma Researcher, "Predicting Soft Skills from Behavioral Data: A Framework for HR and Crisis Management Prédire les soft skills à partir de données comportementales : un cadre conceptuel pour la GRH et la gestion de crise", [Online]. Available: www.revue-isg.com
[14] V. Nakhipova *et al.*, "Use of the Naive Bayes Classifier Algorithm in Machine Learning for Student Performance Prediction," *International Journal of Information and Education Technology*, vol. 14, no. 1, pp. 92–98, 2024, doi: 10.18178/ijiet.2024.14.1.2028.
[15] "URN_NBN_fi_jyu-201508042607".
[16] D. Lee, D. Lin, C.-P. Bezemer, and A. E. Hassan, "Building the perfect game -- an empirical study of game modifications," Mar. 2021, doi: 10.1007/s10664-019-09783-w.
[17] K. Martínez, "The Development Of Online Narrative In Video Games The Player's Agency in Grand Theft Auto V Roleplay," 2024.
[18] K. MARTÍNEZ, "The Development Of Online Narrative In Video Games The Player's Agency in Grand Theft Auto V Roleplay," *VISUAL Review. International Visual Culture Review / Revista Internacional de Cultura*, vol. 16, no. 7, pp. 153–163, Nov. 2024, doi: 10.62161/revvisual.v16.5387.
[19] L. H. Li, "Intersubjectivity and co-constructed framings: students' role-play talks in online English-speaking sessions," *Front. Psychol.*, vol. 15, 2024, doi: 10.3389/fpsyg.2024.1499192.
[20] P. Washington *et al.*, "Challenges and Opportunities for Machine Learning Classification of Behavior and Mental State from Images."
[21] A. Saifudin, "Metode Data Mining Untuk Seleksi Calon Mahasiswa Pada Penerimaan Mahasiswa Baru Di Universitas Pamulang," 2018, doi: 10.24853/jurtek.10.1.25-36.
[22] S. Gandotra, Y. Kumar, N. Modi, J. Choi, J. Shafi, and M. F. Ijaz, "Comprehensive analysis of artificial intelligence techniques for gynaecological cancer: symptoms identification, prognosis and prediction," *Artif. Intell. Rev.*, vol. 57, no. 8, Aug. 2024, doi: 10.1007/s10462-024-10872-6.
[23] N. D. Firdaus and R. R. Suryono, "Perbandingan Algoritma Naive Bayes dan SVM dalam Analisis Sentimen Pengguna AI di Platform X," *Technology and Science (BITS)*, vol. 6, no. 4, 2025, doi: 10.47065/bits.v6i4.7081.
[24] I. Alakbarova, "Analysis of Human Behavior and Interests Basedon Text Data," *International Journal of Education and Management Engineering*, vol. 15, no. 1, pp. 1–9, Feb. 2025, doi: 10.5815/ijeme.2025.01.01.
[25] L. Zhang, "Features extraction based on Naive Bayes algorithm and TF-IDF for news classification," *PLoS One*, vol. 20, no. 7 July, Jul. 2025, doi: 10.1371/journal.pone.0327347.
[26] N. Z. B. Jannah and K. Kusnawi, "Comparison of Naïve Bayes and SVM in Sentiment Analysis of Product Reviews on Marketplaces," *Sinkron*, vol. 8, no. 2, pp. 727–733, Mar. 2024, doi: 10.33395/sinkron.v8i2.13559.
[27] M. Ardana, R. Mayasari, and I. Maulana, "Bulletin Of Computer Science Research Perbandingan Naïve Bayes Dan Svm Untuk Analisis Sentimen Ulasan Kompas.Id Pada Data Tidak Seimbang," *Media Online)*, vol. 6, no. 1, pp. 399–408, 2025, doi: 10.47065/bulletincsr.v6i1.929.
[28] D. Zare, L. Fernandez-Sanz, V. Pospelova, and I. López-Baldominos, "NLP and Text Mining for Enriching IT Professional Skills Frameworks," *Applied Sciences (Switzerland)*, vol. 15, no. 17, Sep. 2025, doi: 10.3390/app15179634.
[29] G. T. de Macedo, A. Fontão, and B. Gadelha, "Building soft skills through a role-play based approach for Requirements Engineering remote education," *Journal of the Brazilian Computer Society*, vol. 30, no. 1, pp. 1–16, Mar. 2024, doi: 10.5753/jbcs.2024.3071.
[30] D. Ogaga and A. Olalere, "Evaluation and Comparison of SVM, Deep Learning, and Naïve Bayes Performances for Natural Language Processing Text Classification Task," Nov. 23, 2023. doi: 10.20944/preprints202311.1462.v1.
[31] V. Mulyadi, R. A. Yuana, and P. Hatta, "A Comparative Analysis of Naive Bayes and Support Vector Machine in Multi-Class Text Classification," *International Journal of Computing and Digital Systems*, vol. 2024, no. 1, pp. 1–10, doi: 10.12785/ijcds/Xxxxxx.
[32] U. U. Tolga Gümüş Authors Ahmet Gümüş Asli Diyadin Lenger Gülnaz Şengül Güneş Murat Çan Sinan Güneş Furkan Yuksel, *specific studies in management, strategy, economy and social sciences edited by*. 2023. [Online]. Available: www.iksadyayinevi.com