

**ANALISIS DALAM PREDIKSI POLA MAP BERMAIN PADA
PERUSAHAAN *SERVER GHOSTCAP* MENGGUNAKAN
METODE *DECISION TREE***

SKRIPSI

OLEH :

Tody Yoshua Arcta
178160119



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MEDAN AREA
MEDAN
2024**

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Document Accepted 20/2/25

Access From (repository.uma.ac.id)20/2/25

**ANALISIS DALAM PREDIKSI POLA MAP BERMAIN PADA
PERUSAHAAN *SERVER GHOSTCAP* MENGGUNAKAN
METODE *DECISION TREE***

SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana di Fakultas
Teknik Universitas Medan Area

Oleh:

Tody Yoshua Arcta

178160119



**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MEDAN AREA
MEDAN
2024**

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

Document Accepted 20/2/25

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Access From (repository.uma.ac.id)20/2/25

HALAMAN PENGESAHAN

Judul Skripsi : analisis dalam prediksi pola *MAP* bermain pada perusahaan
server GhostCap menggunakan metode *decision tree*
Nama : Tody Yoshua Arcta
NPM : 178160119
Program Studi : Sarjana (S1) Teknik Informatika
Fakultas : Teknik

Disetujui Oleh

Pembimbing



Rizki Muliono, S.Kom, M.Kom.
NIDN : 0109038902

Diketahui oleh :

Dekan Fakultas Teknik

Ketua Prodi Terknik Informatika



M. Asa Subarso, S.T, M.T
NIDN : 0102027402



Rizki Muliono, S.Kom, M.Kom
NIDN : 0109038902

Tanggal Lulus : 10 Juli 2024

HALAMAN PERNYATAAN

Saya menyatakan dengan sungguh-sungguh bahwa tugas akhir ini adalah hasil penelitian, pemikiran dan presentasi asli saya sendiri. Saya tidak mencantumkan tanpa pengakuan bahan yang telah diterbitkan atau ditulis oleh orang lain sebelumnya, atau sebagai bahan yang telah diajukan untuk gelar diploma di Universitas Medan Area atau perguruan tinggi lainnya.

Apabila dikemudian hari terdapat kejanggalaan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik sesuai dengan ketentuan yang berlaku di Universitas Medan Area.

Demikian pernyataan ini saya buat.



**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS**

Sebagai civitas akademika Universitas Medan Area, saya yang bertanda tangan di bawah ini:

Nama : Tody Yoshua Arcta
NPM : 178160119
Fakultas : Teknik
Studi : Informatika
Jenis Karya : Tugas Akhir

Demi pengembangan ilmu pengetahuan, setuju untuk memberikan kepada Universitas Medan **Hak Bebas Royalti Non-eksklusif** (*Non-exclusve Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

Analisis Dalam Prediksi Pola *MAP* Bermain Pada Perusahaan *Server GhostCap* Menggunakan Metode *Decision Tree*

Bersama dengan perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti yang bersifat non-eksklusif ini, Universitas Medan Area berhak menyimpan, mengalihkan media/format, mengelola dalam bentuk database, memelihara dan mempublikasikan tugas akhir/tesis/skripsi saya selama saya tetap menyebut nama saya sebagai pencipta/penulis dan sebagai pemilik hak cipta. Demikian Suratpernyataan ini saya buat dengan sebenarnya.

Medan, 05 Juni 2024
Yang menyatakan


(Tody Yoshua Arcta)

ABSTRAK

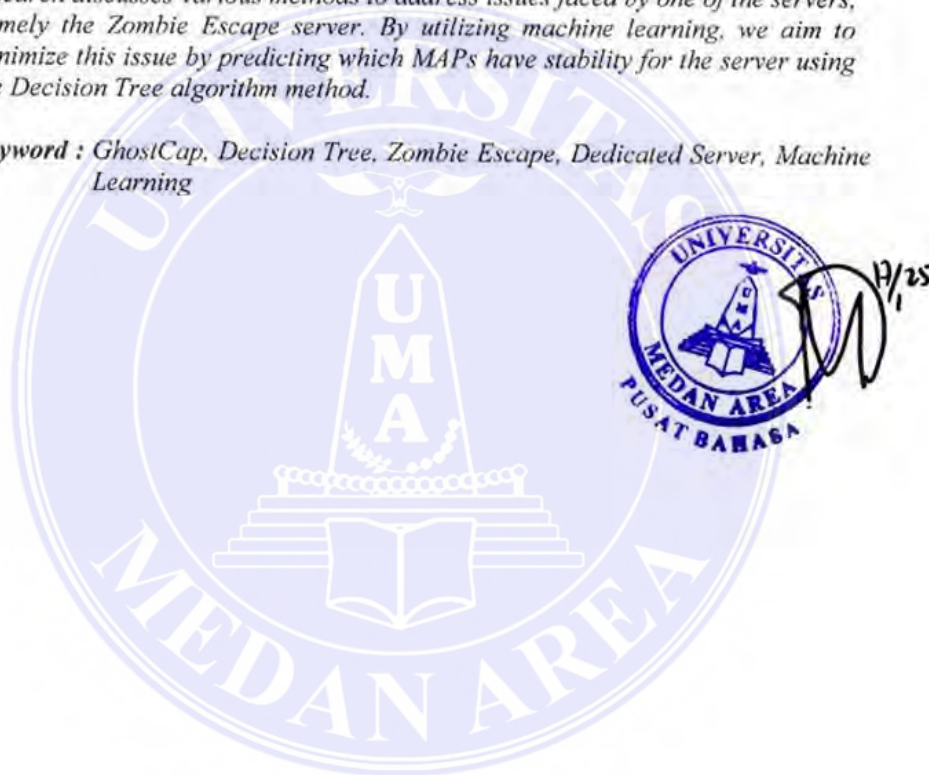
Perusahaan *GhostCap* didirikan oleh Pedrotski (Michael Steven Pedrotti) yang dimulai pada tahun 2015 awalnya merupakan *gaming community* yang hanya pada *game Planetside 2* dan ingin memperluas jangkauannya yakni membuka *server* di bagian belahan dunia. *GhostCap* merupakan perusahaan komunitas *game* yang dapat menampung berbagai pilihan *server* yang tersedia, *GhostCap gaming* berfokus untuk menyatukan para gamer *OCE* (wilayah grafis oceania) dan *SEA* (wilayah South east asia). *MAP* (Peta) yang terdapat pada *server GhostCap zombie escape server* berjumlah sekitar 4954 *zombie escape MAPs* dengan banyaknya jumlah *MAP* tersebut maka cara yang sangat mudah untuk memahami *user behavior* yakni menggunakan metode algoritma *Decision Tree machine learning*. Namun adanya batasan masalah yang terdapat pada perusahaan *GhostCap* yakni penurunan user setiap pergantian *MAP* (tertentu dengan layout difficulty *MAP* yang berbeda menyebabkan berkurangnya pemain pada *MAP* tertentu, Dengan penelitian ini akan membahas berbagai metode untuk menanggulangi masalah yang terdapat pada salah satu *server* yakni *server zombie escape* dengan menggunakan metode *machine learning* kita dapat meminimalisir permasalahan ini dengan memprediksi *MAP* apa yang memiliki stabilitas untuk *server* tersebut menggunakan metode algoritma *Decision tree*.

Kata Kunci : *GhostCap, Decision Tree, Zombie Escape, Dedicated Server, Machine Learning*

ABSTRACT

GhostCap Company was founded by Pedrotski (Michael Steven Pedrotti) in 2015. Initially, it was a gaming community focused solely on the game Planetside 2, aiming to expand its reach by opening servers across different parts of the world. GhostCap is a gaming community company that can accommodate various available server options. GhostCap Gaming focuses on uniting gamers from OCE (Oceania region) and SEA (Southeast Asia region). The MAPs (maps) available on the GhostCap Zombie Escape server total around 4,954 Zombie Escape MAPs. With this vast number of MAPs, understanding user behavior can be simplified by using the Decision Tree machine learning algorithm. However, the limitations faced by GhostCap include a decline in users with each MAP switch, especially due to differences in MAP layout difficulty that cause a decrease in players on certain MAPs. This research discusses various methods to address issues faced by one of the servers, namely the Zombie Escape server. By utilizing machine learning, we aim to minimize this issue by predicting which MAPs have stability for the server using the Decision Tree algorithm method.

Keyword : *GhostCap, Decision Tree, Zombie Escape, Dedicated Server, Machine Learning*



RIWAYAT HIDUP

Penulis dilahirkan di kota Rantau Prapat pada tanggal 03 maret 1999 kec. Labuhan batu, anak ke tiga dari 2 bersaudara dari ayah OKTO HUMALA SIAHAAN dan ibu TITIN SUPRIATIN. Penulis menyelesaikan pendidikan sekolah dasar di sekolah dasar METHODIST 2 RANTAU PRAPAT, Kecamatan, Rantau Utara, kab Labuhanbatu. Pada tahun 2011. Pada tahun yang sama penulis melanjutkan pendidikan Sekolah Menengah Pertama (SMP) pada Sekolah Menengah Pertama (SMP) Negeri 2 Rantau Utara, Kecamatan, Rantau Utara, kab. Labuhanbatu selama 3 tahun penuh dan selesai pada tahun 2014. Penulis melanjutkan pendidikan selanjutnya pada Sekolah Menengah Atas (SMA) Negeri 1 Rantau Utara, Kecamatan, Rantau Utara, kab Labuhanbatu pada tahun 2014 dan lulus pada tahun 2017.

Pada tahun 2017 penulis kembali melanjutkan pendidikan pada perguruan tinggi swasta, tepatnya pada Universitas Medan Area (UMA) Fakultas Teknik pada program studi Informatika.



KATA PENGANTAR

Puji syukur dipanjatkan kehadirat Tuhan yang maha esa, oleh anugerah-Nya yang melimpah, yang masih memberikan kesehatan dan kesempatan sehingga dapat diselesaikan skripsi yang berjudul “ Analisis dalam prediksi pola MAP bermain pada perusahaan *server GhostCap* menggunakan metode *Decision Tree*. ”. Penelitian ini merupakan sebagai salah satu syarat dalam menyusun skripsi pada Program Strata 1 (S1) Teknik Informatika di Universitas Medan Area.


Penelitian ini disusun berdasarkan data dan ilmu yang dimiliki selama melakukan perkuliahan. Dengan telah selesainya Skripsi ini ucapan terima kasih ditujukan kepada kedua orang tua, ayah handa tercinta Okto Humala Siahaan dan ibu tercinta Titin Supriatin atas kesabaran, ketabahan serta ketulusan hati memberikan dorongan moril maupun material serta do'a yang tiada henti-hentinya kepada peneliti. Ucapan terimakasih juga ditujukan untuk pihak-pihak yang telah mengambil bagian dalam penyusunan penelitian ini kepada :

1. Kepada kedua orang tua yang selalu memberikan dukungan semangat dan perjuangan selama masa perkuliahan.
2. Bapak Prof. Dr. Dadan Ramdan, M.Eng, M.Sc, selaku Rektor Universitas Medan Area yang telah memberikan fasilitas kepada mahasiswa/I untuk dapat belajar dengan baik di Universitas Medan Area.
3. Bapak Dr. Eng., Supriatno, S.T, M.T., selaku Dekan Fakultas Teknik di Universitas Medan Area.
4. Bapak Rizki Muliono, S.Kom, M.Kom, selaku ketua Program Studi Teknik Informatika Universitas Medan Area dan sebagai pembimbing I yang telah memberikan arahan dan bimbingan yang sangat berguna kepada penulis dalam menyelesaikan penelitian ini.
5. Seluruh Bapak / Ibu Dosen yang telah memberikan ilmu pengetahuan kepada mahasiswa/I di Universitas Medan Area.
6. Seluruh Staff Universitas Medan Area.
7. Pedrotski (Michael Steven Pedrotti) selaku pemilik perusahaan *GhostCap*

Terima kasih kepada seluruh teman-teman seperjuangan di Universitas Medan Area yang selalu memberikan dukungan serta motivasi. Dan ucapan terima kasih disampaikan kepada pihak-pihak lain yang tidak dapat disebutkan satu per satu yang secara langsung maupun tidak langsung telah membantu dalam menyelesaikan skripsi ini. Tidak ada penghargaan yang berarti dalam membalas jasa-jasa tersebut selain doa dan ucapan terima kasih yang tulus dan ikhlas agar kebaikan yang Bapak, Ibu, Saudara/i kiranya Tuhan YME yang membalasnya di dalam pekerjaan Bapak, Ibu, Saudara/i. Dalam penyusunan skripsi ini masih banyak terdapat kesalahan dan kekurangan. Oleh karena itu, peneliti mengharapkan kritik dan saran yang membangun dari berbagai pihak.

Semoga skripsi ini dapat bermanfaat bagi para pembaca dan semua pihak khususnya bidang manajemen perusahaan. Demikian diakhiri dengan sangat bersyukur dan kiranya apa yang telah dituliskan ini dapat berguna kepada pihak manapun yang membutuhkannya.

Medan, 10 Februari 2024


Tody Yoshua Arcta
178160119

DAFTAR ISI

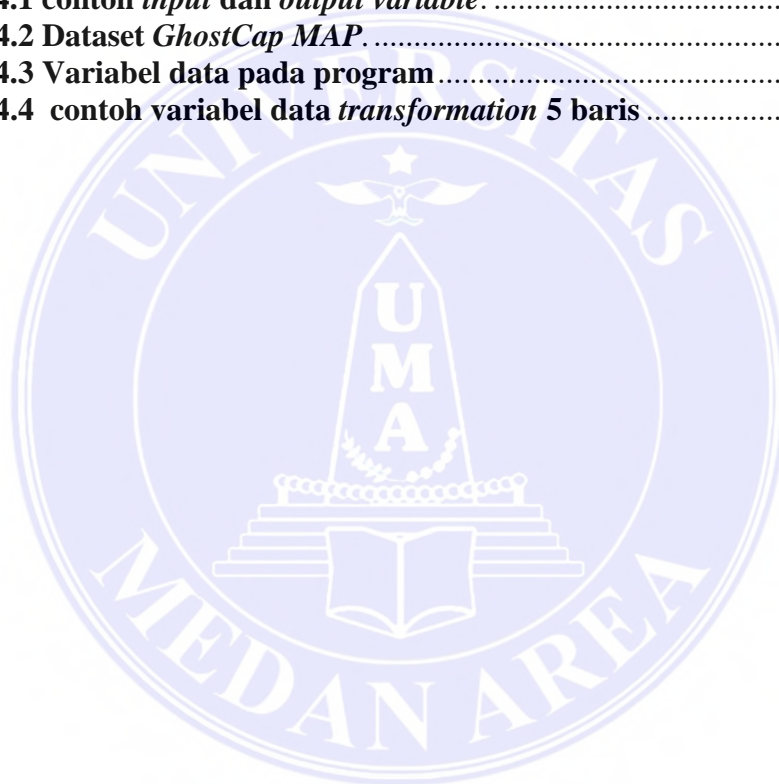
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	i
ABSTRAK	ii
ABSTRACT	iv
RIWAYAT HIDUP	iv
KATA PENGANTAR.....	v
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
BAB II LANDASAN TEORI	5
2.1. Analisa	5
2.2. Alat Pendukung Analisa.....	6
2.2.1. <i>Google Collab</i>	6
2.2.2. <i>Python</i>	6
2.2.3. <i>Microsoft Excell</i>	7
2.2.4. <i>Decision Tree Classifier</i>	8
2.2.5. <i>Extract, Transform, Load (ETL)</i>	11
2.2.6. <i>Exploratory Data Analysis (EDA)</i>	13
2.3. Unified Modeling Language	15
2.4. Penelitian Terdahulu	15
BAB III METODE PENELITIAN.....	21
3.1. Metode Analisa Penelitian	21
3.2. Data Primer	22
3.3. Raw Dataset	23
3.4. Data Cleaning.....	25
3.5. Menentukan Output dan Input	27
3.6. <i>Encoding</i>	28
3.7. <i>Data Splitting</i>	29
3.8. Model Creation dan Fitting	29
3.8.1. Model Creation.....	30
3.8.2. <i>Fitting</i>	32
3.8.3. Implementasi <i>decision tree</i>	32
3.8.4. Contoh Penyelesaian perhitungan manual <i>Decision Tree</i>	37
4.0. Finalisasi dan Remodeling	39

BAB 4 HASIL DAN PEMBAHASAN.....	40
4.1. Hasil	40
4.1.1. Pengumpulan Data	41
4.1.2. Data Cleaning.....	42
4.1.3. Pengujian.....	46
BAB 5 KESIMPULAN DAN SARAN.....	52
5.1. Kesimpulan	52
5.2. Saran.....	52
DAFTAR PUSTAKA.....	53



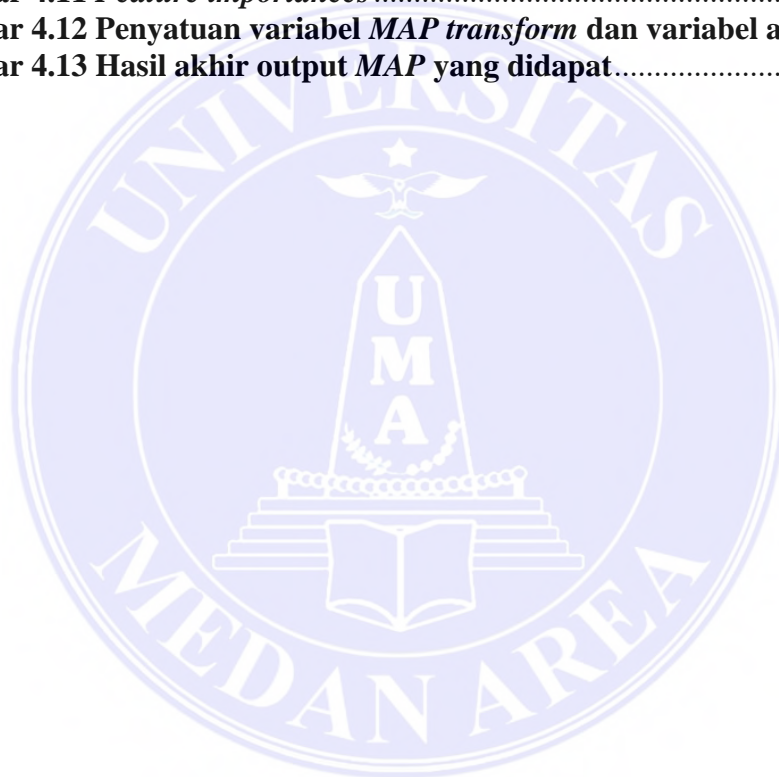
DAFTAR TABEL

Tabel 1.1 Penelitian terdahulu	13
Tabel 3.1 Metode Analisa Penelitian	16
Tabel 3.2 Dataset	18
Tabel 3.3 Raw Dataset types	19
Tabel 3.4 Raw Dataset types transformation	19
Tabel 3.5 Raw Dataset types transform	19
Tabel 3.6 Raw Dataset types transform classified	20
Tabel 3.7 Tabel Variabel I/O	21
Table 3.8 Contoh data untuk perhitungan manual	25
Tabel 3.9 contoh hasil pada table 3.8.	27
Tabel 4.0 format <i>Input</i> dan <i>output table variable</i>.	39
Tabel 4.1 contoh <i>input</i> dan <i>output variable</i>.	39
Tabel 4.2 Dataset <i>GhostCap MAP</i>.	41
Tabel 4.3 Variabel data pada program	43
Tabel 4.4 contoh variabel data <i>transformation 5 baris</i>	46



DAFTAR GAMBAR

Gambar 3.2 gambar model <i>creation pipeline</i>	34
Gambar 4.1 <i>Total minus data type</i>	42
Gambar 4.2 <i>Data type MAP</i>	42
Gambar 4.3 <i>Preprocessing data</i>	43
Gambar 4.4 <i>Transformasi data total minute group</i>	43
Gambar 4.5 variabel <i>MAP</i> sebelum di <i>transformasi</i>	44
Gambar 4.6 Variabel <i>MAP</i> sesudah di <i>transformasi</i>	45
Gambar 4.7 <i>split data menjadi x dan y variable</i>	46
Gambar 4.8 <i>Normalisasi data scaling min max</i>	47
Gambar 4.9 Tes akurasi <i>decision tree</i>	48
Gambar 4.10 Hasil akurasi <i>Decision Tree</i>	48
Gambar 4.11 <i>Feature importances</i>	49
Gambar 4.12 <i>Penyatuan variabel MAP transform dan variabel awal</i>	50
Gambar 4.13 Hasil akhir output <i>MAP</i> yang didapat	51



BAB I

PENDAHULUAN

1.1 Latar Belakang

Perusahaan *GhostCap* didirikan oleh Pedrotski (Michael Steven Pedrotti) yang dimulai pada tahun 2015 awalnya merupakan gaming community yang hanya pada *game* Planetside 2 dan ingin memperluas jangkauannya yakni membuka *server* di bagian belahan dunia. *GhostCap* merupakan perusahaan komunitas *server* yang dapat menampung berbagai pilihan *server* yang tersedia, *GhostCap* gaming berfokus untuk menyatukan para *server* OCE (wilayah grafis *oceania*) dan SEA (wilayah *South east asia*) (FloppyChicken45, 2021) .

Pada perusahaan *server GhostCap*, terdapat kendala terkait ketidakstabilan dalam mencari beberapa pemain yang diinginkan pada jam tertentu. Keadaan ini menyebabkan ketidakstabilan dalam menetapkan pemain yang berpartisipasi di *server GhostCap*. Untuk mengatasi masalah ini, peneliti bertujuan untuk menganalisis dan memahami struktur serta masalah inti yang terjadi pada perusahaan *server GhostCap*.

Metode yang digunakan dalam penelitian ini adalah analisis data dengan menerapkan algoritma *decision tree*. Dengan menggunakan metode ini, diharapkan dapat ditemukan solusi yang efektif untuk meningkatkan stabilitas *server GhostCap* dalam mencari dan menetapkan pemain pada jam tertentu.

Hasil artikel dan jurnal terdahulu yang menggunakan *Decision Tree* sebagai algoritma penentu untuk memprediksi hasil yang akurat dan tepat yang berjudul “*Passenger data analysis of Titanic using machine learning approach in the context of chances of surviving the disaster*“ peneliti mencoba metode yang sama

dan melakukan survey data yang telah diteliti pada tanggal 9 oktober 2022 terdapatnya sejumlah ketidak seimbangannya pemain yang bermain pada jam-jam tertentu yakni memiliki beberapa faktor yang menyebabkan hal tersebut, salah satu *server* yang sangat berpengaruh pada masalah ini merupakan salah satu *server GhostCap* yakni *server zombie escape* yang memiliki kompleksitas dalam berbagai hal maupun di sisi user ataupun *server manager*. *MAP* (peta) yang terdapat pada *server GhostCap zombie escape server* berjumlah sekitar 4954 *zombie escape MAP* dengan banyaknya jumlah *MAP* tersebut maka cara yang sangat mudah untuk memahami *user behavior* yakni menggunakan metode algoritma *Decision Tree machine learning*.

Penelitian ini akan membahas berbagai metode untuk menanggulangi masalah yang terdapat pada salah satu *server* yakni *server zombie escape* dengan menggunakan metode *machine learning* kita dapat meminimalisir permasalahan ini dengan memprediksi *MAP* apa yang memiliki stabilitas untuk *server* tersebut menggunakan metode algoritma *Decision tree*.

1.2 Rumusan Masalah

Adapun masalah yang akan dibahas di dalam sistem pendukung keputusan ini adalah:

1. Bagaimana memprediksi pola *MAP* bermain pada perusahaan *dedicated server GhostCap* dan menerapkan metode *Decision Tree* ?

1.3 Batasan Masalah

Agar pembahasan skripsi ini tidak menyimpang dari yang telah dirumuskan, maka diperlukan batasan-batasan. Batasan-batasan dalam penelitian ini adalah :

1. Kriteria-kriteria yang menjadi prioritas variable adalah date, start hour, start minute, *MAP*, total hour, total minute dan max players.
2. Hanya membahas tentang analisa prediksi pola *MAP*.
3. Metode yang digunakan adalah metode Decision Tree.

1.4 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan. Adapun tujuan dari penelitian ini :

1. Untuk Merancang kriteria yang di pakai dalam *Decision Support System* analisa pola *MAP* pada perusahaan *GhostCap* menggunakan metode *Decision Tree*.
2. Untuk mengimpelentasikan metode *Decision Support System* pada analisa pola *MAP* pada perusahaan *GhostCap* menggunakan metode *Decision Tree*.

1.5 Manfaat Penelitian

Dalam penelitian ini memiliki beberapa manfaat. Adapun manfaat dari penelitian ini yaitu:

1. Meminimalisir terjadinya pemilihan *MAP* yang tidak cocok untuk dimainkan pada kejadian tertentu.
2. Memaksimalkan peforma *server* dalam *Seeding Player* untuk memiliki statistik yang stabil dan akurat.
3. Meningkatkan rating *server* pada *website* rank external seperti *Gametracker*, *BattleMetrics* atau *website* lainnya dan mengoptimalkan kinerja *SEO* (*Search engine optimization*).
4. Meningkatkan kinerja *server* untuk tetap dapat bersaing pada *server* kompetitor lainnya.

5. Diharapkan untuk menambah wawasan dan referensi para peneliti kedepannya dan juga bisa digunakan untuk berbagai perusahaan yang ingin melanjutkan atau mengembangkan penelitian ini.



BAB II

LANDASAN TEORI

2.1 Analisa

Proses analisis dan perancangan sistem memainkan peran yang sangat penting dalam pengembangan perangkat lunak. Tahap ini merupakan langkah awal yang bertujuan untuk memahami kebutuhan pengguna dan menentukan permasalahan yang dapat diatasi melalui sistem yang akan dibangun. Pada tahap ini, tim pengembang bekerja sama dengan pemangku kepentingan untuk mengumpulkan informasi yang diperlukan, melakukan wawancara, dan melakukan analisis mendalam terkait dengan proses bisnis yang terlibat. Kebutuhan dan permasalahan telah diidentifikasi, tahap selanjutnya adalah perancangan sistem. Pada tahap ini, tim pengembang akan merumuskan gambaran mengenai sistem yang akan dibuat. Hal ini meliputi merancang arsitektur sistem, mengidentifikasi komponen dan modul yang diperlukan, serta merencanakan interaksi antara komponen-komponen tersebut. Pada tahap perancangan juga mencakup pemilihan desain dan struktur yang akan digunakan. Tim pengembang harus mempertimbangkan berbagai faktor seperti kehandalan, performa, skalabilitas, dan kebutuhan pengguna dalam menentukan desain yang optimal. Dalam hal ini, menggunakan standar bahasa seperti UML dapat membantu dalam menggambarkan desain sistem secara visual, sehingga memudahkan pemahaman dan komunikasi antara anggota tim pengembang. Penelitian melakukan analisis yang komprehensif dan perancangan yang baik pada tahap awal pengembangan sistem ini menjadi landasan yang kuat untuk memastikan sistem yang dibangun

sesuai dengan kebutuhan pengguna, efisien dalam operasinya, dan mampu mengatasi permasalahan yang ada (Adiguna, 2018).

2.2 Alat Pendukung Analisa

Aplikasi merupakan sarana tools untuk mendukung atau mempermudah sebuah karya ilmiah atau pekerjaan dengan lebih fleksibel, akurat, dan cepat. Dengan adanya aplikasi alat pendukung tersebut terbagi atas berbagai macam jenis alat pendukung yakni sebagai berikut :

2.2.1 Google Collab

Google Collab merupakan sebuah tools *web based cloud machines* produk dari *Google internal research* yang dapat mempermudah users untuk menganalisis data dan pemrograman dengan berbasis Bahasa pemrograman *Python*, dan juga keuntungan dalam menggunakan *Google Collab* yakni juga dapatkan keuntungan seperti GPU secara gratis, Fleksibel, dapat berkolaborasi, terintegritasi. *Google Colaboratory* atau *Colab* adalah layanan cloud berbasis *Notebook Jupyter* yang digunakan untuk penelitian dan pendidikan pembelajaran mesin. Layanan ini memberikan akses gratis ke GPU yang kuat dan runtime yang dikonfigurasi sepenuhnya untuk pembelajaran mendalam (Tiago Carnerio, 2018).

2.2.2. Python

Python adalah bahasa pemrograman tingkat tinggi yang sangat mampu dalam menjalankan berbagai instruksi serbaguna secara langsung melalui metode interpretatif, menggunakan Pemrograman Berorientasi Objek dan semantik dinamis untuk meningkatkan keterbacaan sintaksnya. Sebagai bahasa pemrograman tingkat tinggi, *Python* dikenal karena kemudahan pembelajarannya, berkat manajemen memori otomatisnya. Dengan *Python*, pengembang dapat

dengan mudah memahami dasar-dasar dan segera mulai membuat program sederhana, karena bahasa ini menekankan keterbacaan kode dan menyediakan manajemen memori otomatis, mengurangi kesalahan dan berdasarkan data yang diperoleh pengembang dapat memahami pada logika inti program mereka. Semantik dinamis *Python* juga memberikan kontribusi pada daya tariknya, memungkinkan pengembang menulis kode dengan sintaks yang jelas dan mudah dipahami. Selain itu, *Python* memiliki ekosistem yang luas dan aktif, menawarkan banyak perpustakaan dan modul yang tersedia secara gratis untuk membantu pengembang membangun aplikasi dengan cepat dan efisien. Dengan kesimpulan, keterbacaan, dan dukungan yang luas, *Python* telah menjadi salah satu bahasa pemrograman yang paling banyak digunakan dalam berbagai domain, termasuk pengembangan web, analisis data, kecerdasan buatan, dan pengembangan perangkat lunak skala besar. Fleksibilitas dan sifat pengguna yang ramah membuat *Python* menjadi pilihan yang kuat bagi para pengembang yang ingin menciptakan solusi perangkat lunak yang efektif dan efisien (Ilham, 2020) .

2.2.3. Microsoft Excell

Microsoft Excel berperan sebagai aplikasi yang dapat mengelola data secara otomatis melalui berbagai bentuk, seperti rumus, perhitungan dasar, pengolahan data, pembuatan tabel, dan grafik, serta manajemen data. Fungsi utama dari Microsoft Excel mencakup pembuatan daftar laporan keuangan, daftar nilai, dan daftar hadir. Selain itu, aplikasi ini dapat melakukan operasi perhitungan seperti perkalian, penjumlahan, pengurangan, pembagian, dan perhitungan rerata. Penggunaan Microsoft Excel tidak hanya memungkinkan penyajian data dengan lebih cepat, sistematis, dan akurat, tetapi juga memungkinkan untuk melakukan

perhitungan secara otomatis menggunakan rumus dan logika tertentu. Hal ini sejalan dengan praktik-praktik yang umumnya dilakukan dalam pengelolaan data dan analisis di berbagai konteks penggunaan aplikasi ini (Ogearti, 2020) .

2.2.4. Decision Tree Classifier

Decision Tree Classifier merupakan sebuah metode untuk memprediksi kejadian mendatang menggunakan metode algoritma tersendiri dalam bentuk struktur akar pohon yang akan terus memecah menjadi himpunan jika banyak variable yang terbentuk maka semakin banyak himpunan tersebut yang akan terus menerus bertahap dan berkembang untuk menemukan hasil akhir dari proses tersebut (Chauhan, 2022) .

Berikut berbagai penjelasan mengenai algoritma Decision Tree yang akan diterapkan pada penelitian ini.

Cara menghitung entropy :

$$\text{Entropy (S)} = \sum_{i=1}^n - P_i * \log_2 p_i \dots\dots\dots(2.1)$$

Keterangan :

S = Himpunan Kasus

A = Fitur

n = Jumlah partisi

pi = Proposisi dari Si terhadap S

Cara menghitung nilai Gain :

$$\text{Gain (S,A)} = \text{Entropy (S)} - \sum_{i=1}^n \frac{|S_i|}{|S|} \text{entropy (Si)} \dots\dots\dots(2.2)$$

Keterangan :

S = Himpunan

A = Atribut

N = Jumlah partisi atribut A

(S_i) = Jumlah kasus pada partisi ke $-i$

(S) = jumlah Kasus dalam S

Decision Tree menarik karena alasan berikut :

1. *Decision Tree* adalah suatu metode analisis yang relatif sederhana dan mudah dipahami. Dalam *Decision Tree*, data dipecah menjadi serangkaian keputusan dan kondisi yang berurutan, sehingga memungkinkan untuk melakukan klasifikasi atau prediksi terhadap suatu data. Keunggulan dari pohon keputusan adalah kemampuannya untuk memberikan gambaran visual yang jelas dan intuitif tentang proses pengambilan keputusan.
2. Kelebihan *Decision Tree* adalah kemampuannya untuk dengan cepat diubah menjadi sekumpulan prinsip yang dapat digunakan dalam pengambilan keputusan produksi. Dengan menganalisis cabang-cabang dan simpul-simpul dalam *Decision Tree*, kita dapat mengidentifikasi pola-pola dan aturan-aturan yang penting dalam pengambilan keputusan produksi. Hal ini memungkinkan kita untuk mengambil keputusan yang tepat dan efisien dalam operasionalisasi produksi.
3. *Decision Tree* mampu mengklasifikasikan baik hasil yang bersifat kategori maupun numerik. Namun, perlu diperhatikan bahwa atribut yang dihasilkan dalam pohon keputusan harus bersifat kategori. Hal ini dikarenakan struktur pohon keputusan berdasarkan pemisahan berdasarkan aturan biner, di mana setiap cabang hanya memiliki dua pilihan, yaitu "ya" atau "tidak". Oleh karena itu, atribut numerik perlu diubah menjadi kategori atau kelompok yang sesuai agar dapat digunakan dalam *Decision Tree*.

4. Keunikan *Decision Tree* adalah tidak memerlukan hipotesis priori untuk mengevaluasi kelebihan dari hasil yang didapat. Dalam banyak metode analisis lainnya, hipotesis awal diperlukan untuk membandingkan dan menguji model yang telah dibangun. Namun, dalam *Decision Tree*, hasilnya dievaluasi berdasarkan perbandingan kesalahan klasifikasi pada *Dataset training*. pohon keputusan memberikan pendekatan yang lebih intuitif dan dapat digunakan tanpa memerlukan asumsi sebelumnya.

Kekurangan yang terdapat dari *Decision Tree*, di sisi lain, adalah :

1. Mekanisme *Decision Tree* yang optimal dapat terhambat oleh berbagai faktor dan dalam beberapa situasi, dapat mengakibatkan keputusan yang tidak tepat atau kurang efektif. Hal ini dapat disebabkan oleh kurangnya informasi yang akurat atau lengkap, ketidaktepatan dalam analisis data, pengaruh emosi atau bias personal, serta tekanan waktu yang membatasi proses *Decision Tree*. Ketika mekanisme *Decision Tree* tidak berfungsi secara optimal, konsekuensinya adalah kemungkinan terjadinya keputusan yang tidak efisien, tidak menguntungkan, atau bahkan merugikan. Oleh karena itu, penting untuk mengidentifikasi dan mengatasi faktor-faktor yang dapat mengganggu proses *Decision Tree* guna mencapai hasil yang lebih baik dan lebih akurat.
2. *Decision Tree* merupakan model terstruktur yang terdiri dari banyak lapisan, bercabang seperti cabang-cabang sebuah pohon. Setiap lapisan mewakili atribut atau fitur khusus yang mempengaruhi proses pengambilan keputusan. Lapisan-lapisan yang banyak dalam *Decision Tree* memungkinkan eksplorasi berbagai kemungkinan dan hasil potensial, menjadikannya alat yang menarik dan serbaguna dalam analisis data dan pemodelan prediktif. *Decision Tree*

digunakan secara luas dalam bidang seperti pembelajaran mesin, penambangan data, dan inteligensi bisnis, di mana mereka membantu dalam memahami hubungan kompleks dan mengambil keputusan berdasarkan informasi yang tersedia.

3. Seiring bertambahnya jumlah sampel pelatihan, kompleksitas perhitungan dari pohon keputusan juga dapat meningkat. Pada dataset yang lebih besar, pohon keputusan perlu memproses dan menganalisis jumlah informasi yang lebih besar, yang dapat menghasilkan perhitungan yang lebih rumit dan waktu pemrosesan yang lebih lama. Kompleksitas yang meningkat disebabkan oleh peningkatan jumlah cabang dan simpul yang perlu dievaluasi selama pembangunan dan penelusuran pohon keputusan. Oleh karena itu, penting untuk mempertimbangkan keseimbangan antara ukuran set pelatihan dan sumber daya komputasi yang diperlukan untuk membangun dan menggunakan model pohon keputusan secara efektif (Charbuty, 2021) .

2.2.5. *Extract, Transform, Load (ETL)*

Proses ini melibatkan proses *Extract-Transfer-Load* (ETL) dalam mengkurasi dan menyelaraskan data karakterisasi yang sudah ada, serta membangun sistem baru untuk mengumpulkan metadata dari deposit. Proses ETL diimplementasikan menggunakan kombinasi bahasa pemrograman C, *Python*, dan Ruby, dan bekerja dalam tiga tahap. Tahap pertama mencari data sampel yang sesuai pola penamaan berkas dalam gudang data. Tahap kedua memproses dan mengubah format data menggunakan seperangkat perangkat lunak yang dibangun khusus untuk mengekstrak data dari berbagai format, termasuk format yang bersifat properti. Pada tahap ketiga, data yang sudah diselaraskan dibersihkan dan

digabungkan sehingga data terbaru disimpan dalam database untuk akses yang mudah. Untuk membantu pengumpulan metadata yang efisien pada tahap pertama, dibangun antarmuka pengguna grafis (*Graphical User Interfaces/GUIs*) pada komputer yang terhubung pada instrumen sintesis dan karakterisasi. Selain itu, GUI ini juga sangat penting untuk mengumpulkan "resep sintesis" yang diperlukan untuk penambangan dan analisis data (Zakutayev, 2018) .

1. *Extract*

Proses *Extract* merupakan tahap awal dalam pemetaan data yang berguna untuk memilah dan mengidentifikasi berbagai sumber data yang akan diintegrasikan ke dalam sistem. Tahap ini melibatkan proses ekstraksi data dari berbagai sumber, yang dapat berupa file yang terformat, *query*, atau *endpoint*.

2. *Transform*

Transform adalah sebuah proses untuk pengolahan data setelah berhasil diekstrak dari suatu sumber. Pengolahan data ini bertujuan untuk membersihkan data agar dapat disesuaikan pada format seperti tipe data, manipulasi tanggal, huruf kapital, dan lain-lain yang sudah ditetapkan pada saat perancangan ERD

3. *Load*

Load adalah proses terakhir dalam rangkaian ETL (*Extract, Transform, Load*) yang bertujuan untuk memasukkan data ke dalam database yang sudah dirancang sehingga sesuai pada aliran data yang diinginkan (satrio, 2021) .

2.2.6. *Exploratory Data Analysis (EDA)*

Terutama, analisis data eksploratori adalah pendekatan untuk melihat apa yang dapat dikomunikasikan oleh data di luar tugas pemodelan formal atau pengujian hipotesis. EDA membantu menganalisis set data untuk merangkum karakteristik statistik mereka fokus pada empat aspek utama, seperti ukuran kecenderungan sentral (termasuk mean, modus, dan median), ukuran penyebaran (termasuk standar deviasi dan varians), bentuk distribusi, dan keberadaan pencilan. Paragraf berikutnya, Peneliti menyajikan deskripsi tentang aspek-aspek utama EDA ini. Teknik ini akan dijelaskan dibawah berikut :

1. Eksplorasi Data

Eksplorasi Data adalah tahap pertama dalam analisis data. Di sini kita dapat mengetahui tentang konten dari set data dan karakteristiknya. Ini memberi informasi tentang ukuran data. Kita dapat menemukan nilai yang hilang dalam data. Kita juga dapat menemukan hubungan yang ada di antara data tersebut. Visualisasi data dilakukan menggunakan data tabel dan memahami karakteristiknya.

2. *Data Cleaning*

Data Cleaning adalah proses untuk mendeteksi data yang rusak, menghapus bagian data yang tidak relevan, dan mengganti data yang benar. Proses sebenarnya dari pembersihan data adalah untuk menghapus kesalahan dan memvalidasi data. Data dapat dicek lintas untuk menghilangkan kesalahan. Masalah dapat dipecahkan memvalidasi data.

3. Pembangunan Model

Kita dapat menggunakan model statistik atau model pembelajaran mesin untuk menggambarkan variabel dan cara kerja variabel tersebut. Model dapat berupa model terarah (supervised) atau model tanpa pengawasan (unsupervised). Peneliti dapat menggunakan model klasifikasi, regresi untuk mendapatkan hasil output. Peneliti dapat memvisualisasikan hasil menggunakan model tersebut. Setelah itu, Peneliti harus mengevaluasi model tersebut.

4. Menyajikan Hasil

Kita dapat memvisualisasikan jumlah data yang kompleks menggunakan grafik, diagram, dan tabel. Otak manusia dapat memproses informasi menggunakan grafik dan diagram. Ini merupakan cara yang mudah untuk menyampaikan konsep. Hal ini dapat mengidentifikasi area yang perlu diperbaiki. Selain itu, grafik juga dapat menjelaskan faktor baik.

Tujuan dari EDA (Exploratory Data Analysis) dapat dirangkum sebagai berikut:

1. Maksimalkan wawasan tentang database/memahami struktur database;
2. Visualisasikan hubungan potensial (arah dan besaran) antara variabel paparan dan hasil;
3. Deteksi nilai outlier dan anomali (nilai yang signifikan berbeda dari pengamatan lainnya);
4. Mengembangkan model parsimonius (model prediktif atau eksplanatif yang berperforma sedikit variabel paparan) atau pemilihan awal model yang tepat;
5. Ekstraksi dan pembuatan variabel yang relevan secara klinis.

Metode EDA dapat diklasifikasikan sebagai berikut:

6. Metode grafis atau non-grafis

Metode univariat (hanya satu variabel, paparan atau hasil) atau multivariat (beberapa variabel paparan sendiri atau variabel hasil) (Sahoo, 2019) .

2.3 Unified Modeling Language

UML (*Unified Modeling Language*) telah menjadi salah satu bahasa standar yang luas digunakan di dunia industri untuk mendefinisikan kebutuhan, melakukan analisis dan desain, serta menggambarkan arsitektur dalam pemrograman berorientasi objek. UML memberikan representasi visual yang jelas dan terstruktur untuk mengkomunikasikan ide dan konsep dalam pengembangan perangkat lunak. menggunakan notasi yang konsisten, UML memungkinkan para profesional IT (*Information and Technology*) untuk menggambarkan hubungan antara komponen sistem, alur kerja, dan interaksi antara objek-objek yang terlibat. Hal ini memungkinkan tim pengembangan untuk memahami kompleksitas sistem secara lebih baik, berkolaborasi secara efisien, dan menghasilkan solusi yang tepat dan terstruktur. Dalam industri, UML telah menjadi alat yang sangat berharga dalam pengembangan perangkat lunak yang kompleks dan memfasilitasi komunikasi yang efektif antara tim pengembang, analis, dan pemangku kepentingan lainnya (Putra, 2019) .

2.4 Penelitian Terdahulu

Berikut adalah rangkuman hasil penelitian terdahulu yang memiliki keterkaitan dengan penelitian ini dalam memprediksi data.

Tabel 1.1 Penelitian terdahulu

No.	Nama Penulis	Metode	Hasil Penelitian
1.	Md Arfinul	Decision Tree	Pada training data set , modifikasi

	Haque et al		<p>telah dilakukan untuk mengisi nilai yang kosong dan hal error lain yang berkaitan untuk memenuhi sebuah fitur <i>engineering technique</i>. Dan machine learning algoritma di implementasikan untuk <i>classifier</i> dan <i>clustering techniques</i>. Bisa diketahui bahwa, penumpang yang bepergian pada keluarga yang sedikit merupakan keluarga berukuran dari 2 ke 4 selamat daripada yang tiada. Sebuah kematian diduga lebih besar daripada penumpang yang bepergian secara sendiri dan seorang yang bepergian pada keluarga besar. Dan, klasifikasi pada 2 kelas telah diuji menggunakan <i>decision tree</i> dimana yang selamat telah dianalisa dari beerbagai level. Juga, persentasi yang akurat telah dipantau yakni 86%.status selamat juga di analisa menggunakan teknik clustering. Jauh sebelum hari H, fitur engineering juga telah di uji lebih akurat untuk mendapat pengetahuan</p>
--	-------------	--	--

			<p>dari data titanic. Dan juga , klasifikasi lainnya dan clustering algoritma bisa di aplikasikan untuk berkembang pada nilai akurasi. Penyelidikan data analisis juga telah dilakukan pada data set penumpang titanic pada penelitian ini. Penelitian ini akan diperluas atau diaplikasi menggunakan algoritma ini pada kehidupan sehari hari dataset seperti kejadian alami data atau covid-19 dataset untuk analisa yang lebih dalam. (Arfinul, 2021)</p>
2.	Aradea, Satriyo A., Ariyan Z., Yuliana A.	Decision Tree	<p>Pada penentuan pola data PMB penelitian sudi kasus dikasifikasikan menggunakan model Decision Tree menggunakan algoritma ID3. Terdapat enam atribut data PMB pada pembahasan studi kasus ini yang diklasifikasikan, dan setelah atribut atribut tersebut diolah pola penentu yang paling berpengaruh dalam status data final hanya terdapat tiga atribut saja yaitu: skor ujian masuk pada jurusan SMA, pilihat prioritas studi</p>

			<p>program dan jurusan SMA. Tingkat keakuratan yakni 100% pada studikamus yang telah dibahas. Perlu diketahui bahwa penelitian ini perlu dilanjutkan untuk menemukan status akhir data yang diperlukan, seperti hasil atau skortest masuk yang telah dikelompokkan pada atribut tersebut, menjadi atribut baru yang di komposisi, misalkan skor nilai tes psikolog, skor nilai kemampuan kuantitatif, nilai kemampuan verbal, nilai skor logika, nilai skor tes lainnya. (Aradea, 2011)</p>
3.	Yudhistira Arie Wijaya, Agus Bahtiar, Kaslani, Nining R	Decision Tree	<p>Pada jumlah dataset 65532 model klasifikasi yang diterapkan yaitu decision tree menghasilkan nilai <i>recall</i> tertinggi yakni sebesar 100% dan tingkat akurasi sebesar 99.82%. dari nilai tersebut dapat diketahui bahwa metode algoritma <i>Decision Tree</i> dapat berkembang daripada penelitian sebelumnya yang menggunakan metode <i>Support Vector Machine</i> yang</p>

			memiliki tingkat keakurasian sebesar 98.5%. (Yudhistira Arie Wijaya, 2021).
4.	Rizky Haqmanullah Pambudi, Budi Darma Setiawan, Indiran	Penerapan Algoritma C4.5	Pendidikan dalam kehidupan suatu negara memegang peranan yang sangat penting untuk menjamin kelangsungan hidup negara dan bangsa. Berdasarkan penelitian yang telah dilakukan oleh Paulo Cortez dan Silvia bahwa selama satu dekade terakhir tingkat pendidikan di Portugis telah meningkat. Akan tetapi statistik menunjukkan bahwa tingkat pendidikan Portugis menempati peringkat terbawah karena banyaknya siswa putus sekolah. Penyebab siswa putus sekolah di Portugal disebabkan kegagalan siswa menyelesaikan beberapa bidang studi yaitu bidang studi matematika dan Bahasa Portugis (Cortez & Silva, 2008).
5.	Aldi Saputra,	Algoritma	Perguruan tinggi memiliki kewajiban

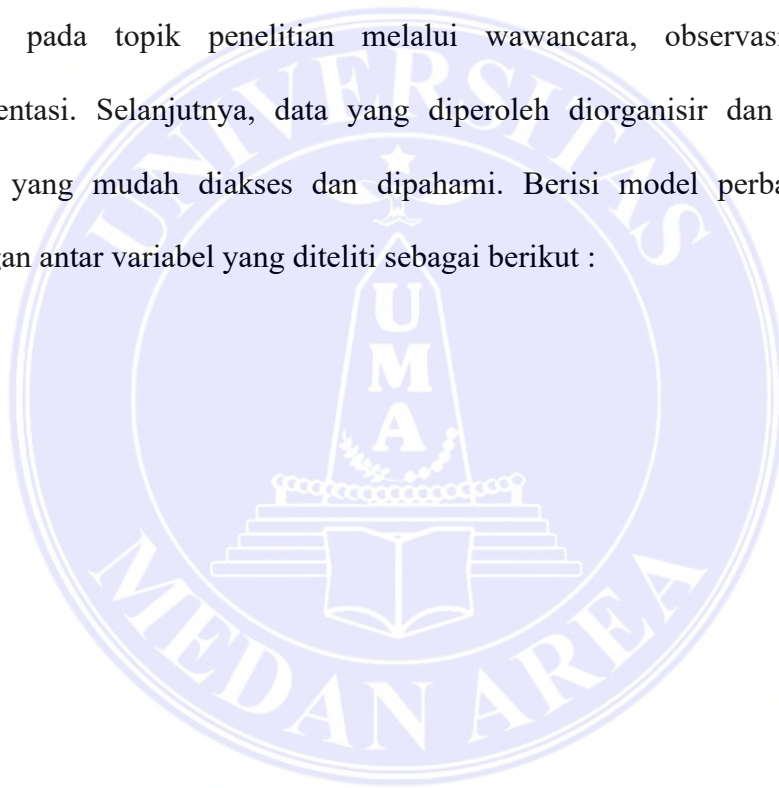
	<p>Triyani Arita Fitri, Karpen, Susanti</p>	<p>C4.5</p>	<p>dalam mengontrol mahasiswanya untuk menghasilkan lulusan yang berkualitas. Politeknik Kampar merupakan salah satu perguruan tinggi di bidang sawit yang ada di Indonesia (Fenty kurnia oktorina & Andri nofiar, 2021). Dalam hal ini, politeknik kampar memiliki tujuan dan target yang mana setiap lulusan dari perguruan tersebut menghasilkan lulusan terbaik dalam bidang industri yang dibuktikan dari nilai akademik yang di peroleh dari data tamatan lulusan yang sebelumnya (Aryadi, 2020).</p>
--	---	-------------	---

BAB III

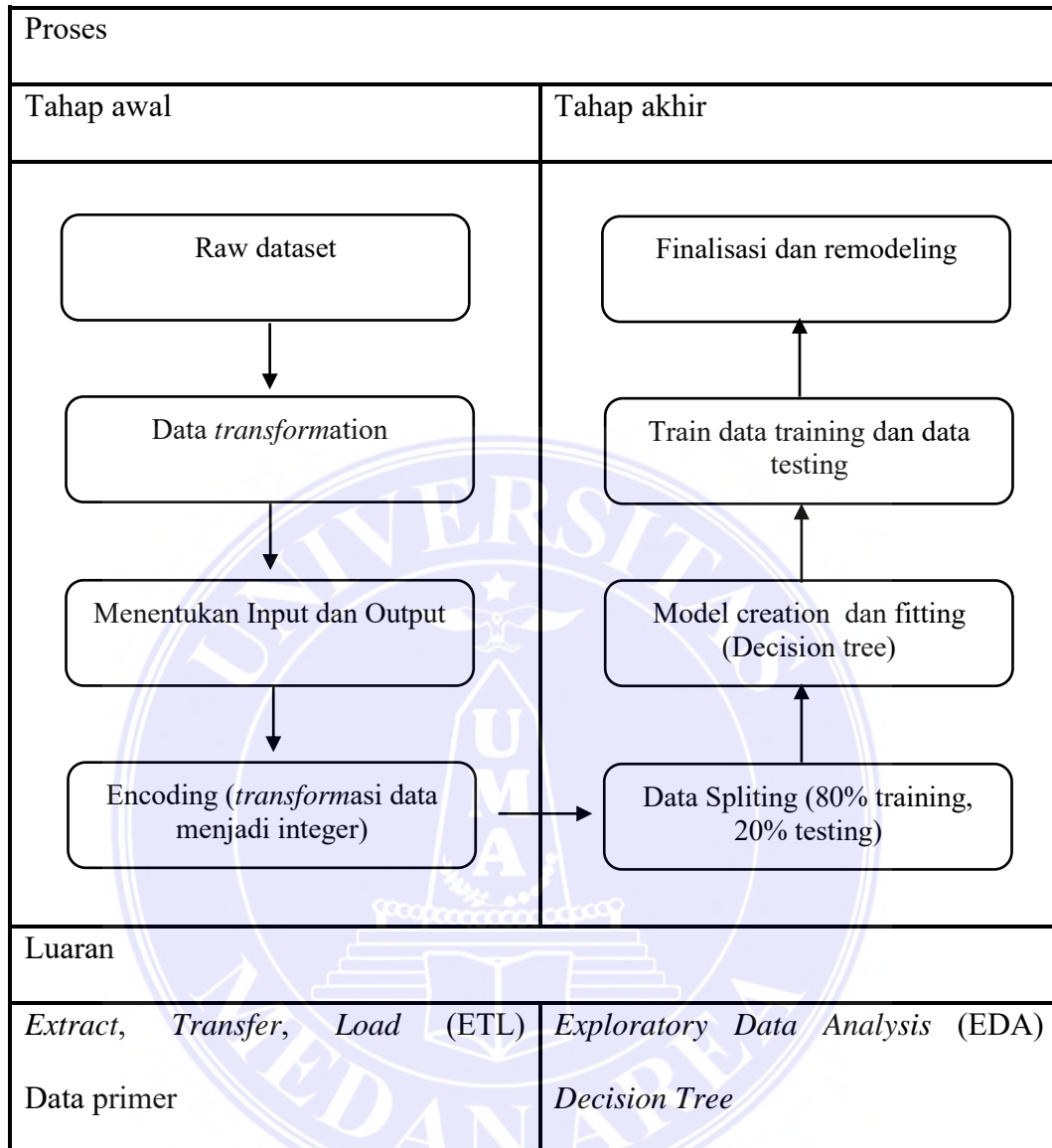
METODE PENELITIAN

3.1 Metode Analisa Penelitian

Metode Analisis Penelitian terdiri dari beberapa tahapan, yaitu pengumpulan data, pengorganisasian data, analisis data, dan interpretasi data. Tahapan pengumpulan data dilakukan pada cara mengumpulkan informasi yang relevan pada topik penelitian melalui wawancara, observasi, atau studi dokumentasi. Selanjutnya, data yang diperoleh diorganisir dan diatur dalam bentuk yang mudah diakses dan dipahami. Berisi model perbandingan atau hubungan antar variabel yang diteliti sebagai berikut :



Tabel 3.1 Metode Analisa Penelitian



3.2 Data Primer

Data primer merujuk pada data yang diperoleh secara langsung oleh peneliti dari lapangan sebagai objek penulisan. (Satrio, 2021) Data primer pada penelitian ini diambil dari perusahaan *server GhostCap* secara langsung oleh peneliti.

Penelitian ini mengumpulkan data menggunakan tiga teknik pengumpulan data, yaitu studi lapangan, studi literatur, dan wawancara.

1. Studi lapangan dapat dilakukan melalui tinjauan langsung ke perusahaan *GhostCap*. Fokus tinjauan adalah data-data yang terkait pada topik penelitian ini, yaitu raw dataset *MAP* pada *server Zombie Escape GhostCap*.
2. Studi literatur dilakukan untuk mempelajari buku dan jurnal yang membahas integrasi data ETL (*Extract, Transform, Load*) untuk mendapatkan raw dataset *MAP* pada server *GhostCap*.
3. Wawancara dilakukan kepada pemilik perusahaan *server GhostCap* untuk mendapatkan informasi yang diperlukan dalam membangun sistem untuk menangani masalah yang terjadi pada proses integrasi data. Hasil analisis menunjukkan bahwa terdapat masalah stabilitas pengguna saat pergantian tipe *MAP* tertentu *server game* tersebut.

3.3 Raw Dataset

Dalam penelitian ini, data dikumpulkan menggunakan teknik *Extract, Transform, Load* (ETL) melalui *Google Collaboratory*. Data yang dikumpulkan adalah data history *MAP* dari *server GhostCap* yang mencakup periode waktu mulai dari tanggal 30 April 2022 hingga 16 Oktober 2022, yang dihasilkan dari Warehouse *GhostCap*. Dalam pelatihan model, variabel data yang digunakan adalah *_date*, *start_hour*, *start_minute*, *MAP*, *total_hour*, *total_minutes*, dan *max_players*. Sumber data yang digunakan adalah *server GhostCap* pada periode tanggal 30 April 2022 hingga 16 Oktober 2022. Penulis menggunakan dataset mentah pada *server Zombie Escape GhostCap* yang telah disediakan oleh data warehouse *GhostCap*, dan berikut adalah keterangannya :

Tabel 3.2 Dataset

no	date	start hour	start min	map	total hours	total min	max players
1	4/30/2022	20	36	ze_aookz	1	106	64
2	4/30/2022	22	27	ze_greec	0	10	64
3	4/30/2022	22	42	ze_santa	1	75	64
4	5/1/2022	4	0	ze_666_	4	251	8
5	5/1/2022	12	30	ze_apoll	0	20	28
6	5/1/2022	19	27	ze_atix :	0	32	62
7	5/1/2022	10	49	ze_atix 1	0	45	42
8	5/1/2022	18	6	ze_best_	1	76	64
9	5/1/2022	1	34	ze_boate	0	25	28
10	5/1/2022	21	35	ze_borec	0	25	63
11	5/1/2022	7	22	ze_copy_	0	0	2
12	5/1/2022	9	38	ze_cv lc	0	45	20
13	5/1/2022	23	6	ze_deade	0	50	64
14	5/1/2022	22	5	ze_ffvi	0	55	64
15	5/1/2022	11	39	ze_gods_	0	20	38
16	5/1/2022	5	46	ze_grey_	0	50	3
17	5/1/2022	3	55	ze_honka	0	0	5
18	5/1/2022	10	29	ze_lotr 1	0	15	30
19	5/1/2022	15	43	ze_lotr i	0	35	60
20	5/1/2022	3	30	ze_milita	0	20	5
21	5/1/2022	6	41	ze_mojor	0	35	3
22	5/1/2022	12	5	ze_poker	0	20	28
23	5/1/2022	8	17	ze_ramp_	0	25	7
24	5/1/2022	0	3	ze_santa	0	55	64
25	5/1/2022	13	40	ze_shroo	0	55	60
26	5/1/2022	2	45	ze_silent	0	40	7
27	5/1/2022	2	4	ze_skill_	0	35	13
28	5/1/2022	20	14	ze_starw	1	75	64
29	5/1/2022	8	48	ze_suchi	0	45	8
30	5/1/2022	17	40	ze_surf_	0	20	63
31	5/1/2022	1	4	ze_surf_	0	25	29
32	5/1/2022	12	55	ze_templ	0	40	49
33	5/1/2022	14	41	ze_the n	0	55	64
34	5/1/2022	16	23	ze_ut200	1	71	64
35	5/2/2022	5	15	ze_666_	8	505	42
36	5/2/2022	7	48	ze_apert	0	45	4
37	5/2/2022	10	29	ze_artika	0	45	29
38	5/2/2022	3	23	ze_backr	0	35	5
39	5/2/2022	9	13	ze_bathr	0	45	12
...
4953	10/16/2022	2	40	zm lila	0	50	5

Jumlah dataset yang diperoleh pada *server GhostCap* yakni berjumlah 4953

jumlah data yang diperoleh dari tanggal 30 April 2022 hingga 16 Oktober 2022.

3.4 Data Cleaning

Berikut merupakan variable basis dataset yang terdapat pada perusahaan *GhostCap* yang akan di implementasikan menggunakan metode *Decision Tree* yaitu :

1. Berikut merupakan table variable penentu pada raw Dataset *GhostCap* yang belum di *transformasi* sebagai berikut:

Tabel 3.3 Raw Dataset types

Field	Type
<i>_date</i>	varchar
<i>start_hour</i>	integer
<i>start_minute</i>	integer
<i>MAP</i>	varchar
<i>total_hour</i>	integer
<i>total_minutes</i>	integer
<i>max_players</i>	integer

Berikut contoh dataset yang belum di *transformasi* :

Tabel 3.4 Raw Dataset types transformation

<i>_date</i>	<i>start_hour</i>	<i>start_minute</i>	<i>map</i>	<i>total_hours</i>	<i>total_minutes</i>	<i>max_players</i>
4/30/2022	20	36	ze_aooka_v3_t6	1	106	64
4/30/2022	22	27	ze_greece_escape_p	0	10	64
4/30/2022	22	42	ze_santassination_v3	1	75	64
5/1/2022	4	0	ze_666_crazy_escape_v2_4	4	251	8
5/1/2022	12	30	ze_apollo_p2	0	20	28

2. Berikut merupakan table variable penentu pada raw Dataset *GhostCap* yang sudah di *transformasi* tahap pertama sebagai berikut:

Tabel 3.5 *Raw Dataset types transform*

Variable	Type
<i>month</i>	integer
<i>day</i>	integer
<i>start_hour</i>	integer
<i>Total_minutes</i>	integer
<i>max_players</i>	integer
<i>MAP</i>	varchar

Contoh dataset yang sudah di *transformasi* seperti yang telah kita lihat bahwa ada beberapa data yang sudah di *transformasi* yakni

- a.) table *_date* di *transformasi* dan dipecah menjadi *day* dan *month* yang terdapat pada value date
- b.) table *total_hour* dan *total_minutes_groups* di kombinasi menjadi total minutes yang terbagi menjadi 3 value yakni short, normal, dan long maximum values yang terekam pada variabel total hour dan total minutes jika 15> menit dikelompokkan menjadi tipe short, 25> menit dikelompokkan menjadi normal dan 60> menit dikelompokkan menjadi long .

Tabel 3.6 Raw Dataset types transform classified

	A	B	C	D	E	F	G	H
1	_month	_day	start_hour	total_minutes_groups	max_playe	map		
2	5	1	4	long	8	ze_666_crazy_escape_v2_4		
3	5	2	5	long	42	ze_666_crazy_escape_v2_4		
4	5	3	2	long	5	ze_666_crazy_escape_v2_4		
5	5	4	3	long	44	ze_666_crazy_escape_v2_4		
6	5	5	4	long	62	ze_666_crazy_escape_v2_4		
7	5	6	0	long	22	ze_666_crazy_escape_v2_4		
8	5	7	4	long	54	ze_666_crazy_escape_v2_4		
9	5	8	4	short	1	ze_666_crazy_escape_v2_4		
10	5	11	1	long	45	ze_666_crazy_escape_v2_4		
11	5	12	4	long	2	ze_666_crazy_escape_v2_4		
12	5	13	4	long	34	ze_666_crazy_escape_v2_4		
...
4951	9	20	3	long	19	zm_supermercadochino_v8_2		
4952	9	27	2	long	14	zm_supermercadochino_v8_2		
4953	10	1	3	normal	14	zm_supermercadochino_v8_2		
4954	10	12	2	long	2	zm_supermercadochino_v8_2		

3.5 Menentukan Output dan Input

Penelitian sebelumnya bertujuan untuk mengukur interaksi antara variabel input dan output menggunakan metode statistik tertentu. Keterkaitan antara variabel input dan output dapat ditentukan saat memperhatikan faktor-faktor lain yang mempengaruhinya, seperti pengaruh variabel luar atau faktor lingkungan. Oleh karena itu, dalam penelitian tersebut juga dilakukan analisis regresi untuk memperkirakan pengaruh variabel luar terhadap variabel output. Dalam upaya untuk mencapai jumlah variabel yang optimal, penelitian tersebut juga menggunakan teknik seleksi variabel untuk mengevaluasi pentingnya masing-masing variabel input dalam memprediksi variabel output. Selain itu, penelitian tersebut juga mengukur hubungan antara variabel input dan output secara dinamis, saat melihat perubahan nilai variabel input dan output dari waktu ke waktu.

Penelitian ini terdapat variabel penentu untuk menentukan hasil akhir dan input yang diperlukan untuk menjalankan sebuah program yang akan di

implementasikan menggunakan decision tree nantinya, berikut merupakan variabel input yakni *_month*, *_day*, *start_hour*, *total_minutes_groups*, dan *max_players* dan variabel sebagai outputnya yakni *MAP* untuk menentukan hasil akhir dari prediksi dataset tersebut.

Tabel 3.7 Tabel Variabel I/O

Variable	Type
<i>_month</i>	input
<i>_day</i>	input
<i>start_hour</i>	input
<i>total_minutes</i>	input
<i>max_players</i>	input
<i>MAP</i>	output

3.6 *Encoding*

Encoding atau sering juga disebut *transformasi* data merupakan sebuah proses yang sangat penting dalam pengolahan data. Proses *Encoding* dilakukan sebelum kita melakukan data *Splitting*, dimana tujuannya adalah untuk mengubah data mentah menjadi bentuk yang dapat diproses oleh mesin. Hal ini dilakukan untuk memastikan bahwa setiap jenis data memiliki representasi yang tepat dan sistem dapat membaca dan memahami jenis data tersebut. Proses *Encoding* juga dapat membantu dalam memperbaiki kualitas data, seperti membuang nilai yang hilang atau tidak valid. Dalam hal hasil akhir yang akan diprediksi, proses *Encoding* dapat memainkan peran yang sangat penting, karena hasil prediksi yang akurat tergantung pada representasi data yang baik. Oleh karena itu, pemilihan teknik

Encoding yang tepat sangatlah penting untuk mencapai hasil prediksi yang optimal.

3.7 Data Splitting

Dataset asli akan kita *Split* menjadi dua bagian yaitu *Data Training* dan *Data Testing*. *Data Training* akan digunakan untuk melatih algoritma pada mesin, sedangkan *Data Testing* digunakan untuk menguji seberapa baik algoritma tersebut bekerja pada data yang belum pernah dilihat sebelumnya. Biasanya, rasio pembagian *Data Training* dan *Data Testing* adalah 80:20, dimana 80% digunakan untuk *Data Training* dan 20% untuk *Data Testing*. Pembagian data ini dilakukan untuk mempermudah dan mempercepat proses testing jika dataset yang digunakan memiliki jumlah data yang sangat besar. Metode ini sangat umum digunakan oleh data scientist untuk menguji berbagai jenis algoritma dan model yang ingin diimplementasikan pada dataset besar. Pada penggunaan pembagian data ini, waktu yang diperlukan untuk melatih model dan menguji performanya pada *Data Testing* dapat diminimalkan sehingga memungkinkan pengembangan model yang lebih cepat dan efektif pada big data.

3.8 Model Creation dan Fitting

Penelitian dataset pada tabel 3.4.13 bahwa peneliti menggunakan metode analisa berbasis EDA (Exploratory Data Analysis) Menggunakan EDA sangat penting karena dapat membantu dalam menjawab pertanyaan-pertanyaan dasar terkait data, seperti apakah data memiliki distribusi normal atau tidak, apakah terdapat outlier pada data, atau bagaimana hubungan antara variabel yang satu variabel yang lain. Selain itu, EDA juga dapat membantu dalam mengidentifikasi

masalah yang ada pada data, seperti data yang hilang atau tidak valid. Berikut merupakan langkah-langkah model creation dan fitting :

3.8.1. Model Creation

Model creation pipeline adalah serangkaian proses atau tahapan yang dilakukan untuk membangun sebuah model machine learning secara sistematis dan terstruktur. Pipeline ini dapat membantu untuk mempercepat proses pembuatan model machine learning, meningkatkan efisiensi, dan juga memungkinkan untuk mengotomatiskan beberapa tahapan dalam proses tersebut.

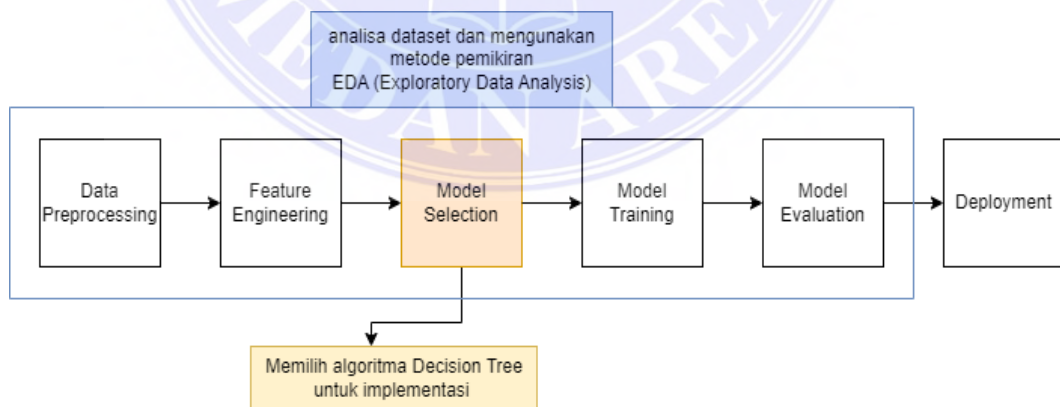
pipeline umumnya terdiri dari beberapa tahapan, di antaranya adalah:

1. Data preprocessing: Tahap ini meliputi proses pemrosesan data seperti menghapus atau mengisi data yang hilang, melakukan normalisasi data, dan mengubah format data agar dapat diolah oleh algoritma machine learning.
2. Feature engineering: Tahap ini melibatkan pemilihan dan ekstraksi fitur-fitur atau variabel-variabel yang relevan dari data. Tujuannya adalah untuk menghasilkan fitur-fitur yang paling berkorelasi pada hasil prediksi yang ingin dicapai.
3. Model selection: Tahap ini melibatkan pemilihan jenis model machine learning yang paling sesuai pada data dan tujuan prediksi. Jenis model yang dipilih akan berbeda-beda tergantung pada jenis data, skala data, dan tujuan prediksi.
4. Model training: Tahap ini melibatkan pelatihan model machine learning pada data yang telah diproses dan dipilih fitur-fiturnya. Tujuannya adalah

untuk mengoptimalkan parameter-model sehingga dapat menghasilkan prediksi yang akurat.

5. Model evaluation: Tahap ini melibatkan evaluasi kinerja model machine learning. Tujuannya adalah untuk memeriksa apakah model yang dibuat mampu menghasilkan prediksi yang akurat dan dapat diandalkan.
6. *Deployment*: Tahap ini melibatkan penerapan model *machine learning* ke dalam produksi. Tujuannya adalah agar model dapat digunakan dalam situasi dunia nyata dan mampu menghasilkan prediksi yang akurat.

Pada *pipeline* ini, proses pembuatan model *machine learning* dapat menjadi lebih terstruktur, efisien, dan mudah di-maintain. Selain itu, pipeline ini dapat memastikan bahwa semua tahapan yang diperlukan dalam pembuatan model *machine learning* dilakukan secara benar dan konsisten, sehingga dapat menghasilkan model yang berkualitas tinggi dan dapat diandalkan. Berikut merupakan contoh diagram model *creation pipeline* :



Gambar 3.2 gambar model creation pipeline

3.8.2. *Fitting*

Fitting pada dataset machine learning adalah proses penyesuaian model *machine learning* dengan *Data Training*. Dalam proses *fitting*, model machine learning akan mempelajari pola-pola dalam *Data Training* dan menyesuaikan parameter-parameter pada model agar dapat memberikan prediksi yang sesuai pada *Data Training*. Proses *fitting* biasanya dilakukan menggunakan algoritma optimisasi seperti *Decision Tree* untuk menemukan parameter optimal pada model. Setelah proses *fitting* selesai, model machine learning akan siap digunakan untuk melakukan prediksi pada data yang belum pernah dilihat sebelumnya (*Data Testing*).

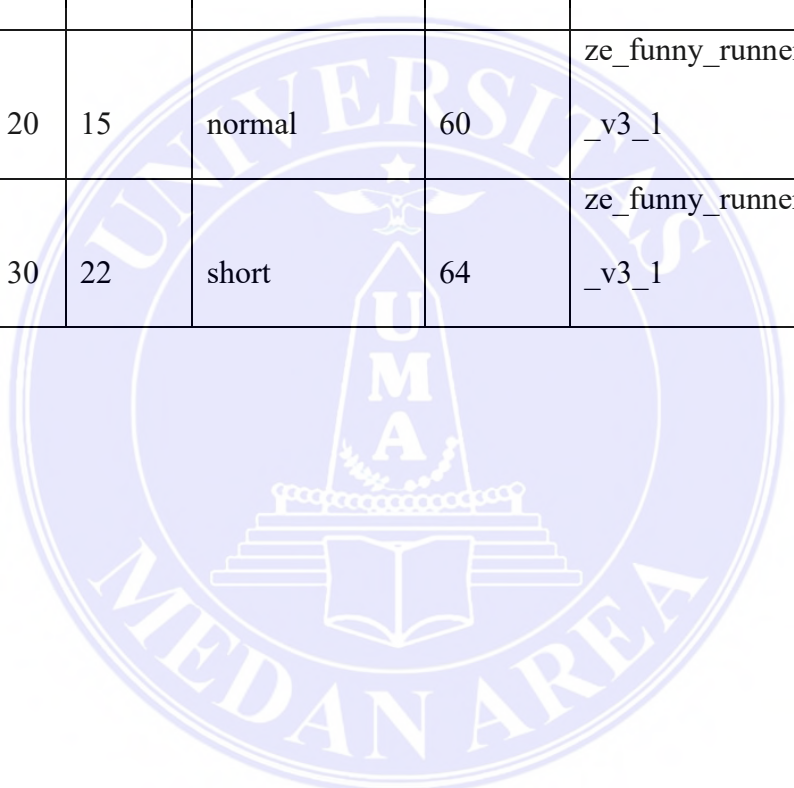
3.8.3. Implementasi *decision tree*

Bagian ini terdapat data yang telah di olah dan dianalisa oleh peneliti secara manual menggunakan rumus-rumus dari metode *Decision Tree* yang dipakai oleh penulis.

Table 3.8 Contoh data untuk perhitungan manual

_mo nth	_d ay	start_h our	total_minutes_ groups	max_pla yers	map	PLAY
Apr il	13	4	long	8	ze_666_crazy_es cape_v2_4	YA
Mei	20	22	normal	50	ze_666_crazy_es cape_v2_4	TIDAK
Juni	30	15	short	61	ze_666_crazy_es cape_v2_4	YA
Apr	13	23	long	33	ze_ffvii_mako_re	TIDAK

il					actor_v5_3_v5	
Mei	20	23	normal	50	ze_ffvii_mako_re actor_v5_3_v5	YA
Juni	30	3	short	5	ze_ffvii_mako_re actor_v5_3_v5	YA
Apr il	13	19	long	64	ze_funny_runner _v3_1	TIDAK
Mei	20	15	normal	60	ze_funny_runner _v3_1	YA
Juni	30	22	short	64	ze_funny_runner _v3_1	TIDAK



Tabel 3.9 contoh hasil pada table 3.8

NODE		Keterangan	jumlah kasus	s1 YA	s2 TIDAK	Entropy	Gain
1	total		9	5	4	0,99107606	
	_month						0,202771893
		APRIL	3	1	2	0,918295834	
		MEI	3	2	1	0,918295834	
		JUNI	3	3	1	0,528320834	
	_day						0,072780226
		20	3	2	1	0,918295834	
		13	3	1	2	0,918295834	
		30	3	2	1	0,918295834	
	_start hour						0,768853838
		3	1	1	0	0	

		4	1	1	0	0	
		15	2	2	0	0	
		19	1	0	1	0	
		22	2	0	2	0	
		23	2	1	1	1	
	_total minute group						0,072780226
		long	3	1	2	0,918295834	
		short	3	2	1	0,918295834	
		normal	3	2	1	0,918295834	
	_max player						0,768853838
		5	1	1	0	0	
		8	1	1	0	0	
		33	1	0	1	0	

		50	2	1	1	1	
		60	1	1	0	0	
		61	1	1	0	0	
		64	2	0	2	0	
	_map						0,072780226
		ze_666_crazy_escape_v2_4	3	2	1	0,918295834	
		ze_ffvii_mako_reactor_v5_3_v5	3	2	1	0,918295834	
		ze_funny_runner_v3_1	3	1	2	0,918295834	

3.8.4 Contoh Penyelesaian perhitungan manual *Decision Tree*

Cara menghitung entropy :

$$\text{Entropy (S)} = \sum_{i=1}^n - P_i * \log_2 p_i \dots\dots\dots(2.1)$$

Keterangan :

S = Himpunan Kasus

A = Fitur

n = Jumlah partisi

pi = Proposisi dari Si terhadap S

Penyelesaian entropy April

$$\begin{aligned} \text{Entropy (S)} &= (- 1/3 * \log_2 1/3) + (-2/3 * \log_2 2/3) \\ &= \log_2 ((1/3)^{-1 1/3}) + \log_2 ((2/3)^{-2/3}) \\ &= \log_2 ((1/3)^{1/3} * ((2/3)^{-2/3}) \\ &= \log_2 ((1/3^{-1/3} * 2^{-2/3}/3^{-2/3}) \\ &= \log_2 (2^{-2/3}/3^{-1}) \\ &= \log_2 (3/2^{2/3}) \\ &= \log_2 (3) - 2/3 \\ &= \underline{0,918296} \end{aligned}$$

Cara menghitung nilai Gain :

$$\text{Gain (S,A)} = \text{Entropy (S)} - \sum_{i=1}^n \frac{|S_i|}{|S|} \text{entropy (Si)}\dots\dots\dots(2.2)$$

Keterangan :

S = Himpunan

A = Atribut

N = Jumlah partisi atribut A

(Si) = Jumlah kasus pada partisi ke –i

(S) = jumlas Kasus dalam S

Penyelesaian gain Month

$$\begin{aligned}
 \text{Gain bulan} &= (0,991076) - (3/9 * 0,918296) - (3/9 * 0.918296) - (3/9 * \\
 &\quad 0.528321) \\
 &= (0,991076) - (1/3 * 0,991076) - (1/3 * 0,991076) - (1/3 * \\
 &\quad 0.528321) \\
 &= 247769/250,000 - (1/3 * 114787/125,000) - (1/3 * 114787/125,000) \\
 &\quad - (1/3 * 0.528321/100,000) \\
 &= 247769/250,000 - 114787/125,000 - 114787/125,000 - \\
 &\quad 176107/100,000 \\
 &= 121663/600,000 \\
 &= \underline{0.202772}
 \end{aligned}$$

3.9 Train data training dan data testing

Penelitian ini hasil dari *data splitting* pada tabel Tabel 3.3.10 total data menjadi 80% *Data Training* dan 20% *Data Testing*. Teknik ini guna di gunakan pada banyak penelitian terdahulu guna untuk mempersingkat waktu dalam memprediksi pada *big data* yang memiliki ratusan juta data yang akan nantinya di implementasikan menggunakan model algoritma tertentu 20% data tessting digunakan untuk memprediksi dataset yang akan di analisis apakah modeling algoritma tersebut cocok untuk di implementasikan atau tidak, teknik ini sangat efesien untuk testing implementasi modeling algoritma tanpa memakan waktu banyak yang alhasil jenis modeling algoritma tersebut tidak cocok atau memiliki akurasi yang tidak memuaskan.

4.0 Finalisasi dan Remodeling

Tahap terakhir ini merupakan hasil dari analisis dataset yang telah di analisa berikut merupakan *prediction value* yang akan di terapkan kepada dataset *GhostCap* ini menggunakan variabel berikut :

Tabel 4.0 format Input dan output table variabel

Input Variable				
_month	_day	Start_hour	total_minutes_groups	max_players
Output	<i>_MAP</i>			
Hasil				

Tabel 4.1 contoh input dan output variable

Input Variable				
5	7	1	long	52
Output	ze_ffvii_mako_reactor_v5_3_v5			
Hasil				

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan analisis *Dataset GhostCap* menggunakan algoritma *Decision Tree*, dapat disimpulkan bahwa:

1. Analisis *Dataset* tersebut terdapat jumlah record sebanyak 4953 record.
2. Variabel berpengaruh pada tingkat prediksi *MAP* merupakan month, day, start hour, total minute group dan max player.

5.2 Saran

Dari hasil penelitian yang dilakukan, terdapat beberapa saran yang dapat diimplementasikan, antara lain:

1. Untuk mendapatkan keakuratan yang konsisten pada model, diperlukan data yang memiliki variabel dan jumlah data yang besar.
2. Pada hasil penelitian ini, Perusahaan *GhostCap* dapat memanfaatkan hasil prediksi untuk melakukan evaluasi terhadap pola permainan pada setiap *MAP*. Hal ini akan membantu dalam menjaga konsistensi pergantian setiap *MAP* dan meningkatkan pengalaman bermain.
3. Harapannya, penelitian ini dapat dikembangkan lebih lanjut menjadi sebuah sistem atau aplikasi yang memudahkan dalam pemanfaatannya.

Demikian, hasil penelitian ini dapat lebih secara praktis digunakan dan memberikan manfaat yang lebih luas.

DAFTAR PUSTAKA

Adiguna, A. R. (2018). Analisis dan perancangan sistem informasi manajemen gudang pada PT Mitra Pinasthika Mulia Surabaya. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 612-621.

Abstrak dari doi:e-ISSN: 2548-964X.

Charbuty, B. &. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 20-28.

<https://www.jastt.org/index.php/jasttpath/article/download/65/24> [13 november 2023].

Chauhan, N. S. (2022). Decision Tree Algorithm, Explained. *Decision Tree Algorithm, Explained*.

<https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html> [13 november 2023].

FloppyChicken45. (2021). *GhostCap* About Us.

<https://gyo.gg/clan/GhostCap-gaming1605974021/#:~:text=GhostCap%20Gaming%20is%20an%20all,to%20the%20hottest%20AAA%20titles>. [13 november 2023].

Hendra Nusa Putra, S. M. (2018). Implementasi Diagram UML (Unified Modelling Language) dalam Perancangan Aplikasi Data Pasien Rawat Inap pada Puskesmas Lubuk Buaya. *Implementasi Diagram UML (Unified Modelling Language) dalam Perancangan Aplikasi Data Pasien Rawat Inap pada Puskesmas Lubuk Buaya*, 71.

<https://polgan.ac.id/jurnal/index.php/sinkron/article/view/130/69> [13 november 2023].

HJ, S. (t.thn.). Experimental design and analysis. *Experimental design and analysis*.

<http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf> [13 november 2023].

Ilham, M. (2020, march 26). Pengertian *Python*, Fungsi, dan Kekurangan.

<https://materibelajar.co.id/pengertian-python/> [13 november 2023].

Pengertian Microsoft Excel dan Fungsinya. (2022). *Pengertian Microsoft Excel dan Fungsinya*. Diambil kembali dari

<https://puskom.uma.ac.id/2022/01/pengertian-microsoft-excel-dan-fungsinya/#:~:text=Pengertian%20Microsoft%20Excel%20adalah%20sebuah,dan%20kolom%20untuk%20mengeksekusi%20perintah.>

[13 november 2023].

Petri, G. S. (2021). Bandit Modeling of *MAP* Selection in Counter-Strike: Global Offensive. *Bandit Modeling of MAP Selection in Counter-Strike: Global Offensive*, 2106.08888.

<https://doi.org/10.48550/arXiv.2106.08888> [13 november 2023].

Putra, D. W. (2019). Unified modelling language (uml) dalam perancangan sistem informasi permohonan pembayaran restitusi sppd. *Jurnal Teknoif*, 32-39.

<https://teknoif.itp.ac.id/index.php/teknoif/article/view/57> [13 november 2023].

Sahoo, K. S. (2019). *Exploratory Data Analysis using Python. Exploratory data analysis using Python., Volume-8(Issue-12)*, 4727-4728.

https://www.researchgate.net/profile/Dr-SubhenduPani/publication/337146539_IJITEE/links/5dc70b124585151435fb427e/IJITEE.pdf [diakses 12 oktober 2023].

Satrio, Y. D. 2021. penerapan ETL (*extract transform load*) untuk integrasi data simlitabmas pada aplikasi pengelolaan penelitian dan pengabdian masyarakat universitas dinamika. penerapan ETL (*extract transform load*) untuk integrasi data simlitabmas pada aplikasi pengelolaan penelitian dan pengabdian masyarakat universitas dinamika,21-22.

<https://repository.dinamika.ac.id/id/eprint/5890/1/16410100072-2021-UNIVERSITASDINAMIKA.pdf> [diakses 11 oktober 2023].

-----2021. penerapan ETL (*extract transform load*) untuk integrasi data simlitabmas pada aplikasi pengelolaan penelitian dan pengabdian masyarakat universitas dinamika. penerapan ETL (*extract transform load*) untuk integrasi data simlitabmas pada aplikasi pengelolaan penelitian dan pengabdian masyarakat universitas dinamika, 25.

<https://repository.dinamika.ac.id/id/eprint/5890/1/16410100072-2021-UNIVERSITASDINAMIKA.pdf> [diakses 11 oktober 2023].

tiago carnerio, r. v.-B. (07 October 2018). Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. *Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications*, 61677 - 61685.

<https://ieeexplore.ieee.org/document/8485684> [diakses 15 oktober 2023].

Zakutayev, A. W. (2018). An open experimental database for exploring inorganic materials. *Scientific data*, 10.

<https://www.nature.com/articles/sdata201853> [diakses 17 oktober 2023].



LAMPIRAN

source code

```
# Import Libraries

import numpy as np
import pandas as pd
import seaborn as sns

gcg = pd.read_csv('GCG_Input.csv',sep=';')
gcg.head(10)

gcg['total_minutes_groups'].unique()

gcg['map'].unique()

gcg_transformed = gcg.copy()
gcg_transformed.head()

from sklearn.preprocessing import LabelEncoder
LE = LabelEncoder()

# gcg_transformed.iloc[:,3] = LE.fit_transform(gcg_transformed.iloc[:,3].values)
# gcg_transformed.iloc[:,5] = LE.fit_transform(gcg_transformed.iloc[:,5].values)

gcg['total_minutes_groups'].unique()

gcg_transformed['total_minutes_groups'].unique()

gcg_transformed['map'].unique()

gcg_transformed.head(5)

# Encode total_minutes_groups
total_minutes_mapping = {'long': 3, 'normal': 2, 'short': 1}
gcg_transformed['total_minutes_groups_encoded'] =
gcg_transformed['total_minutes_groups'].map(total_minutes_mapping)

# Encode map column
gcg_transformed = gcg_transformed.sort_values(by='max_players',
ascending=False)
unique_maps = gcg_transformed['map'].unique()
map_mapping = {map_name: idx + 1 for idx, map_name in
enumerate(unique_maps)}
gcg_transformed['map_encoded'] = gcg_transformed['map'].map(map_mapping)
```

```
gcg_transformed.head(5)

filtered_df = gcg_transformed[['_month', '_day', 'start_hour', 'max_players',
'total_minutes_groups_encoded', 'map_encoded']]
filtered_df.head(5)

filtered_df['map_encoded'].unique()

# Split the data into independent 'X' and dependent 'Y' variables

x = filtered_df.iloc[:,0:5].values
y = filtered_df.iloc[:,5].values

# Split the dataset into 80% training and 20% testing
from sklearn.model_selection import train_test_split

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2,
random_state=0)

# scale the data

from sklearn.preprocessing import StandardScaler

sc = StandardScaler()

x_train = sc.fit_transform(x_train)
x_test = sc.fit_transform(x_test)

# create a function with many machine learning models

def models(x_train, y_train):

    # use Decision Tree
    from sklearn.tree import DecisionTreeClassifier
    tree = DecisionTreeClassifier(criterion='entropy', random_state=0)
    tree.fit(x_train, y_train)

# Print the Training Accuracy for each model

print('[1]Decision Tree Training Accuracy: ', tree.score(x_train,y_train))

return log, knn, svc_lin, svc_rbf, gauss, tree, forest

from google.colab import drive
drive.mount('/content/drive')
```

```
# get and train all of the models

model = models(x_train, y_train)

# get feature importance

forest = model[5]
importances = pd.DataFrame({'feature': gcg_transformed.iloc[:,0:5].columns,
'importance': np.round(forest.feature_importances_,3)})
importances = importances.sort_values('importance',
ascending=False).set_index('feature')
importances

def get_map(_num):
    reverse_map_mapping = {v: k for k, v in map_mapping.items()}
    return reverse_map_mapping.get(_num, "Map not found")

gcg_transformed.head(20)

# prediction variable
# [month,day,starthour,total minute group,maxplayer]
# Max value dibawahah sesuai dengan variabel penentu diatas.
# [1~12, 1~30 , 0 ~ 24 , 0~64, 1-3 ]

# _month    _day  start_hour  max_players  total_minutes_groups_encoded
# map_encoded

my_data = [[5,30,1,60,3]]

from sklearn.preprocessing import StandardScaler
sc = StandardScaler()

fitted = sc.fit(x_train)
my_data_scaled =fitted.transform(my_data)
# print(my_data_scaled)

# Print prediction of my prediction using Decision Tree Classifier

pred = model[5].predict(my_data_scaled)

print(get_map(int(pred)))
```



UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estetika/Jalan PBSI Nomor 1 Medan (061) 7366878, 7360168, 7364348, 7366781, Faks:(061) 7366998 Medan 20223
Kampus II : Jalan Gellibudu Nomor 79 / Jalan Sei Garuda Nomor 70 A, (061) 8225802, Faks: (061) 8225321 Medan 20123
Website: www.teknik.uma.ac.id E-mail: univ_medan@uma.ac.id

Nomor : 246/FT.6/01.10/V/2024
Lamp : -
Hal : Perpanjang SK Pembimbing Tugas Akhir

17 Mei 2024

Yth. Pembimbing Tugas Akhir
Rizki Muliono, S.Kom, M.Kom
di
Tempat

Dengan hormat,
Sehubungan telah berakhirnya waktu masa berlaku SK pembimbing nomor 484/FT.6/01.10/VI/2023 tertanggal 26 Juni 2023 maka perlu diterbitkan kembali SK Pembimbing Skripsi baru atas nama mahasiswa berikut :

Nama : Tody Yoshua Arcta
N P M : 178160119
Jurusan : Teknik informatika

Oleh karena itu kami mengharapkan kesediaan saudara :

Rizki Muliono, S.Kom, M.Kom (Sebagai Pembimbing)

Adapun Tugas Akhir Skripsi berjudul :

"Analisis dalam Prediksi Pola Map Bernain pada Perusahaan Server Ghostcap menggunakan Metode Algoritma Decision Tree"

SK Pembimbing ini berlaku selama enam bulan terhitung sejak SK ini diterbitkan. Jika proses pembimbing melebihi batas waktu yang telah ditetapkan, SK ini dapat ditinjau ulang.

Demikian kami sampaikan, atas kesediaan saudara diucapkan terima kasih.

Dekan,


Dr. Eng. Supriatno, ST, MT



UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK

Kampus 1 : Jalan Kalam Nomor 1 Medan, Estate Jalan PBSI Nomor 1 ☎ (061) 736678, 7360168, 7364346, 7366781, Fax (061) 7366868 Medan 20223
Kampus 2 : Jalan Seiatabadi Nomor 75 / Jalan Sili Sempu Nomor 70 A, ☎ (061) 8225802, Fax (061) 8226331 Medan 20122
Website: www.teknik.uma.ac.id E-mail: umv_medan@uma.ac.id

Nomor : 618 / FT.6/01.10/VIII/2023

11 Agustus 2023

Lamp : -

H a l : Penelitian Dan Pengambilan Data Tugas Akhir

Yth. Pimpinan Ghost Cup
Toronto St, Toronto, On M4B 0A3
Di
Kanada

Dengan hormat,

Kami mohon kesediaan Bapak/Ibu berkenan untuk memberikan izin dan kesempatan kepada mahasiswa kami tersebut dibawah ini :

NO	N A M A	N P M	PRODI
1	Tody Yoshua Arcta	178160119	Teknik informatika

Untuk melaksanakan Penelitian dan Pengambilan Data Tugas Akhir pada perusahaan/Instansi yang Bapak/Ibu Pimpin.

Perlu kami jelaskan bahwa Pengambilan Data tersebut adalah semata-mata untuk tujuan ilmiah dan Skripsi yang merupakan salah satu syarat bagi mahasiswa tersebut untuk mengikuti ujian sarjana pada Fakultas Teknik Universitas Medan Area dan tidak untuk dipublikasikan, dengan judul penelitian :

Analisis dalam Prediksi Pola Map Bernain pada Perusahaan Server Ghostcup menggunakan Metode Algoritma Decision Tree

Atas perhatian dan kerja sama yang baik diucapkan terima kasih.


Dekan,
Dr. Rahmad Syah, S. Kom, M. Kom

Tembusan :
1. Ka. BAMA1
2. Mahasiswa
3. File



SURAT KETERANGAN

Nomor : 03/GCG/S/K.3-2024

Yang bertanda tangan dibawah ini :

Nama : Michael S. Pedrotti
Jabatan : Direktur Utama
Unit Kerja : PT. GhostCap Toronto St. Toronto, On M4B OA3
Kanada

Mnerangkan dengan sesungguhnya bahwa mahasiswa "Universitas Medan Area" tersebut dibawah ini :

No	NAMA	NIM	PROGRAM STUDI
1.	TODY YOSHUA ARCTA	178160119	TEKNIK INFORMATIKA

telah selesai melaksanakan Penelitian di kantor PT. GhostCap daerah Toronto St. Toronto, On M4B OA3 Kanada Terhitung mulai tanggal 11 agustus 2023 sampai dengan tanggal 25 oktober 2023.

Demikian surat keterangan ini dibuat untuk dipergunakan sebagai mestinya

Direktur utama GhostCap
Toronto St, Toronto



turnitin Similarity Report ID: oid:29477:39411358

PAPER NAME	AUTHOR
Skripsi Tody bab3.doc	Tody yoshua arcta
WORD COUNT	CHARACTER COUNT
6613 Words	43500 Characters
PAGE COUNT	FILE SIZE
36 Pages	653.0KB
SUBMISSION DATE	REPORT DATE
Jul 22, 2023 10:57 AM GMT+7	Jul 22, 2023 10:58 AM GMT+7

- **21% Overall Similarity**
The combined total of all matches, including overlapping sources, for each database.
 - 20% Internet database
 - 5% Publications database
 - Crossref database
 - Crossref Posted Content database
 - 13% Submitted Works database
- **Excluded from Similarity Report**
 - Small Matches (Less than 10 words)