

BAB II

LANDASAN TEORI

2.1 Penelitian Terdahulu

Tabel 2. 1 Penelitian Terdahulu

Judul Jurnal	Nama Jurnal	Penulis/Tahun	Hasil	Kesimpulan
<i>SENTIMENT ANALYSIS OF LEAGUE OF LEGENDS: WILD RIFT REVIEWS ON GOOGLE PLAY USING NAÏVE BAYES CLASSIFIER</i> [15]	Jurnal Ilmiah KURSOR	Nur Sabilly, Khadijah, Fajar Agung Nugroho (2023)	Penelitian ini akan mencoba dua metode yaitu dengan metode BOW (<i>Bag Of Words</i>) dan TF-IDF(<i>Term Frequency –Inverse Document Frequency</i>) dengan menggunakan data <i>training</i> sebanyak 200 sampel, menghasilkan metode BOW memiliki performa lebih baik dengan nilai akurasi, <i>precision</i> , <i>recall</i> , dan <i>F1-Score</i> masing-masing sebesar 79.3%, 82.10%, 83.50%, 82.81%.	Penelitian ini menyimpulkan bahwa proses transformasi kata-kata non-standar menjadi kata-kata standar meningkatkan kinerja model algoritma NBC (<i>Naive Bayes Classifier</i>) pada kedua ekstraksi fitur, yaitu BOW (<i>Bag of Words</i>) dan TF-IDF. Kedua, proses <i>stemming</i> pada <i>dataset</i> pra-pemrosesan umumnya menghasilkan kinerja klasifikasi algoritma NBC yang lebih rendah dalam sebagian besar eksperimen, baik ketika menggunakan fitur TF-IDF maupun BOW. Serta, Hasil pengujian menunjukkan bahwa ekstraksi fitur BOW menghasilkan kinerja model algoritma NBC yang lebih baik dibandingkan dengan ekstraksi fitur TF-IDF.

<p>ANALISIS SENTIMEN TERHADAP PENGGUNAAN APLIKASI GAME ONLINE PUBG MOBILE MENGGUNAKAN ALGORITMA NAIVE BAYES [16]</p>	<p>JATI(Jurnal Mahasiswa Teknik Informatika)</p>	<p>Muhammad Haikal, Martanto, Umi Hayati</p>	<p>Penelitian ini akan melakukan analisis sentimen untuk aplikasi PUBG Mobile dengan menggunakan algoritma Naïve-Bayes. Data ulasan akan diambil dari platform Google Playstore sebanyak 2000 data ulasan dengan metode <i>scraping</i>. Hasil dari penelitian ini berupa akurasi tertinggi yang didapatkan oleh algoritma Naïve-Bayes sebesar 70% dengan presisi 82%, <i>recall</i> 61%, dan <i>f1-score</i> 53%</p>	<p>Kesimpulan pada penelitian ini ialah bahwa akurasi tertinggi pada algoritma Naïve-Bayes dengan rasio 90:10 sebesar 70%. Hasil <i>confusion matrix</i> menampilkan performa model yang kurang seimbang antar kelasnya. Penelitian ini juga menemukan mayoritas sentimen negatif pada ulasan mengenai masalah teknik seperti <i>bug</i>, <i>lag</i>, dan <i>crash</i>.</p>
<p>Perbandingan Algoritma Naïve Bayes, KNN, dan Decision Tree terhadap Ulasan Aplikasi Threads dan Twitter [11].</p>	<p>KLIK: Kajian Ilmiah Informatika dan Komputer</p>	<p>Muhammad Iqbal, Ade Davy Wiranata, Rayhan Suwito, Ridha Faiz Ananda (2023).</p>	<p>Penelitian ini menggunakan 899 data opini pada aplikasi Twitter dan 638 data opini untuk aplikasi Threads, ditunjukkan dengan pembagian data rasio 80:20, ialah untuk algoritma Naïve-Bayes mendapatkan akurasi 66,41% (Threads) dan 85.56% (Twitter), algoritma Decision Tree mendapatkan akurasi 65.52% (Threads) dan 72.78% (Twitter), dan algoritma KNN mendapatkan akurasi 62.50%(Thread) dan 71.57% (Twitter).</p>	<p>Hasil kesimpulan dari penelitian ini ialah bahwa algoritma Naïve-Bayes merupakan algoritma terbaik dengan akurasi tertinggi untuk aplikasi Threads yaitu 66,41%, <i>precision</i> sebesar 64.86%, dan <i>recall</i> sebesar 78.85% sedangkan akurasi untuk aplikasi Twitter ialah 85.56%, <i>precision</i> 84.69%, dan <i>recall</i> 88.30%.</p>
<p>Analisis Sentimen Ulasan Aplikasi Dana dengan Metode <i>Random Forest</i> [12]</p>	<p>Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer</p>	<p>Fanka Angelina Larasati, Dian Eka Ratnawati, Buce Trias Hanggara (2022)</p>	<p>Penelitian ini menggunakan 1354 data dari hasil <i>scraping</i> di Google Playstore yang dibagi menjadi kelas positif sebanyak 577 data, kelas negatif sebanyak, 279 data dan kelas netral sebanyak 509. Menggunakan data <i>training</i> masing-masing 250 data maka</p>	<p>Hasil penelitian ini menyimpulkan bahwa dengan pembagian data sebanyak 250 data masing-masing merupakan pembagian paling tertingi untuk menghasilkan nilai akurasi, <i>recall</i>, <i>precision</i>,</p>

			<p>nilai <i>precision</i> sebesar 84%, <i>recall</i> sebesar 84%, <i>F1-Score</i> sebesar 84% dan nilai akurasi juga menjadi 84%.</p>	<p>dan <i>F1-Score</i> didapati hasil tertinggi yaitu 84%. Serta dari hasil kata yang sering muncul, DANA memerlukan untuk meningkatkan layanan <i>customer service</i> dan juga kemampuan dalam <i>top up & transfer</i> pada aplikasinya.</p>
<p>Klasifikasi Ulasan Konsumen Menggunakan Random Forest dan SMOTE [10]</p>	<p>Journal of System and Computer Engineering (JSCE)</p>	<p>Nurul Istiqamah, Muhammad Rijal (2024)</p>	<p>Penelitian ini menggunakan 14999 data yang disimpan dalam format tabel .xlsx. Proses pembagian data akan dibagi menjadi 70:30. Penelitian ini akan mengkombinasikan teknik SMOTE dan RUS pada model algoritma Random Forest. Model Random Forest menghasilkan nilai akurasi sebesar 73%, nilai <i>recall</i> 73%, dan nilai presisi 98%. Model Random Forest yang menggunakan SMOTE menghasilkan akurasi 72%, nilai <i>recall</i> 75%, dan nilai presisi 89%. Selanjutnya, untuk model Random Forest dengan SMOTE dan RUS menghasilkan nilai akurasi sebesar 75%, nilai <i>recall</i> 78%, dan nilai presisi 89%. Pada pengujian kedua dengan menerapkan parameter <i>max_features</i> berjumlah 8000 pada model Random Forest dengan SMOTE dan RUS menghasilkan peningkatan akurasi menjadi 77%.</p>	<p>Kesimpulan yang didapat pada penelitian ini ialah bahwa teknik SMOTE dan RUS dapat meningkatkan kinerja model kurang lebih sekitar 3% dan model dapat mengenali data yang awalnya minor. Hasil akurasi tertinggi yang diperoleh sebesar 77%.</p>

<p><i>Sentiment Analysis on the Impact of MBKM on Student Organizations Using Supervised Learning with Smote to Handle Data Imbalance</i> [13]</p>	<p>Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi</p>	<p>Lailatul Cahyaningrum, Ardytha Luthfiarta, Mufida Rahayu (2024)</p>	<p>Pengumpulan data akan menggunakan 3 cara yaitu metode kuesioner Google Form, <i>crawling data</i> Twitter, dan <i>crawling data</i> Youtube menghasilkan total data sebanyak 1770 data dan menggunakan rasio 90:10 dalam proses pembagian data. Penelitian ini akan menggunakan metode SMOTE untuk menyeimbangkan kelas data pada tiga algoritma <i>stemming</i> yaitu sastrawi, Nazief & Adriani, dan Arifin Sentiono. Pada penelitian ini model Naïve Bayes yang telah dilakukan proses SMOTE pada tiga jenis algoritma <i>stemming</i> ini mendapatkan hasil akurasi 75% - 82%, pada model SVM mendapatkan nilai akurasi 80%-91% sedangkan pada model Decision Tree mendapatkan nilai akurasi 61% - 77%.</p>	<p>Kesimpulan pada penelitian ini adalah bahwa algoritma SVM merupakan algoritma terbaik dari kedua algoritma tersebut dengan nilai akurasi tertinggi sebesar 91% dengan rasio data 90:10, sedangkan untuk model Naïve Bayes dan Decision Tree dengan rasio yang sama hanya mendapatkan nilai akurasi tertinggi masing-masing sebesar 82%, dan 77%.</p>
<p><i>COVID-19 Vaccines Related User's Response Categorization Using Machine Learning Techniques</i> [14]</p>	<p><i>Computation</i></p>	<p>Ahmed Shahzad, Bushra Zafar, Nouman Ali, Uzma Jamil, Abdulaziz Jarallah Alghadhban, Muhammad Assam, Nivin A. Ghamry, Elsayed Tag Eldin (2022)</p>	<p>Penelitian ini akan melakukan analisis sentimen pada setiap tipe vaksin berdasarkan <i>tweets</i> di aplikasi Twitter. Algoritma ML yang digunakan pada penelitian ini adalah Random Forest, Naïve Bayes, Decision Tree, Logistic Regression, dan SVM. Rasio pembagian data latih dan data uji pada pengujian ini adalah 70:30. Penelitian ini menghasilkan akurasi untuk algoritma Random Forest sebesar 74.42% - 83.61%, untuk algoritma Naïve Bayes mendapatkan</p>	<p>Hasil penelitian ini menyimpulkan bahwa algoritma Decision Tree menjadi algoritma dengan tingkat akurasi terbaik dengan jumlah sebesar 93,87% dengan <i>dataset</i> Sinopharm, dengan <i>dataset</i> Sinovac 92,8%, dengan <i>dataset</i> Pfizer 91,07%, dengan <i>dataset</i> AstraZeneca 90,94%, dan untuk <i>dataset</i> Moderna sebesar 88,01%.</p>

			hasil akurasi sebesar 71.02% - 75.28%, selanjutnya untuk algoritma Decision Tree mendapatkan akurasi sebesar 88.01% - 93.87%, algoritma Logistic Regression mendapatkan akurasi sebesar 78.72% - 86.48%, dan terakhir algoritma SVM mendapatkan akurasi sebesar 81.21% - 87.67%.	
Support Vector Machine VS Information Gain: Analisis Sentimen Cyber bullying di Twitter Indonesia [17].	ULTIMA InfoSys	Christevan Destitus, Wella, Suryasari (2020)	Data pada penelitian ini akan diambil dengan periode waktu 1 April 2020 hingga 28 Mei 2020, dengan setiap kata yang ditarik ialah sebanyak 30 <i>tweet</i> . Metode algoritma <i>Support Vector Machine</i> , prediksi ini akan menggunakan 5 parameter, yaitu epsilon, variabel lambda, maksimum iterasi, kostanta gamma dan nilai C. Pengujian ini menghasilkan param nilai C memiliki nilai akurasi paling tertinggi sebanyak 87%. Hasil pengujian pada <i>information gain</i> menghasilkan nilai akurasi sebesar 86%, <i>precision</i> 81%, <i>recall</i> 95% dan <i>f-measure</i> 87%.	Berdasarkan hasil penelitian yang telah dilakukan menghasilkan bahwa nilai akurasi yang didapatkan dengan menggunakan metode <i>Support Vector Machine</i> ialah sebesar 80%, <i>precision</i> 75.1%, <i>recall</i> 96%, dan <i>f-measure</i> 85% dan pada pengujian <i>threshold</i> mendapatkan akurasi sebesar 86%.
Sentiment Analysis about Indonesian Lawyers Club Television Program Using K-Nearest Neighbor, Naïve Bayes Classifier, and Decision Tree [18].	IJNMT(International Journal of New Media Technology)	Nico Nathanael Wilim, Raymond Sunardi Oetama (2021)	Penelitian ini akan menggunakan data <i>tweets</i> dari tahun 2018 hingga 2019 dengan kata kunci “IndonesianLaywersClub”, “ILCTvone”, dan “ILC” dengan setiap bulannya sebanyak 30 <i>tweets</i> . Hasil penelitian dengan sentimen manual menghasilkan sentimen positif tahun 2018 ialah 194 sedangkan tahun 2019	Hasil penelitian ini menyimpulkan bahwa sentimen positif untuk acara ILC dari tahun 2018 hingga 2019 menurun drastis sedangkan acara Mata Najwa mengalami peningkatan. Algoritma yang menghasilkan tingkat akurasi

			adalah 96. Total sentimen negatif tahun 2018 ialah 166 sedangkan 2019 ialah 264. Hasil penelitian ini menghasilkan algoritma Naïve Bayes yang terbaik pada tahun 2018 sedangkan KNN merupakan yang terbaik untuk tahun 2019. Hal ini mengartikan bahwa tidak ada algoritma yang selalu menjadi terbaik.	terbaik ialah KNN sebesar 76.94% dan tidak ada algoritma yang menunjukkan performa terbaik, Naïve-Bayes algoritma terbaik untuk 2018 namun KNN merupakan algoritma terbaik 2019.
PENERAPAN METODE RANDOM FOREST DALAM MENGANALISIS SENTIMEN PENGGUNA APLIKASI CAPCUT DI GOOGLE PLAY STORE [19].	JATI (Jurnal Mahasiswa Teknik Informatika)	Ayu Sagita, Ahmad Faqih, Gifthera Dwilestari, Bambang Siswoyo, Denni Pratama (2023)	Penelitian ini menggunakan model algoritma Random Forest. Pengumpulan data akan dilakukan dari aplikasi Google Playstore untuk aplikasi CapCut dengan jumlah data sebanyak 5000 ulasan dengan menggunakan teknik <i>scrapping</i> . Pembagian data akan dibagi menjadi rasio 90:10 yaitu 90% data latih dan 10% data uji. Selanjutnya, data ulasan akan dilakukan proses <i>pre-processing</i> seperti, <i>labeling, cleaning, case folding, tokenization, stopword, dan stemming</i> . Selanjutnya, akan dilakukan proses pembobotan dengan teknik TF-IDF dan setelah itu dapat di implementasikan ke dalam model Random Forest. Hasil <i>Confusion Matrix</i> pada penelitian ini adalah sebanyak 166 data <i>True Positive</i> dan 209 data <i>True Negative</i> . Hasil dari <i>classification Report</i> pada penelitian ini mendapatkan akurasi sebesar 86%, presisi 89%, <i>recall</i> sebesar	Hasil penelitian ini mendapatkan akurasi cukup tinggi untuk model Random Forest sebesar 86%, presisi 89%, <i>recall</i> sebesar 81%, dan <i>f-1 score</i> sebesar 85%. Hal tersebut membuktikan bahwa model Random Forest dapat menjadi pilihan yang cukup efektif dengan kelebihanannya dalam mengurangi resiko <i>overfitting</i> dan data yang tidak lengkap.

			81%, dan <i>f-1 score</i> sebesar 85%.	
ALGORITMA DECISION TREE UNTUK ANALISIS SENTIMEN PUBLIC TERHADAP MARKETPLACE DI INDONESIA [20].	NARATIF : Jurnal Ilmiah Nasional Riset Aplikasi dan Teknik Informatika	Azril Tazidan Octa.N, Ma'mun Hasbullah, Miftahul Rizal, Muhamad Fauzan Rajab, Nova Agustina (2023)	Proses penelitian ini akan menggunakan data dari aplikasi Twitter mengenai opini atau komentar public tentang <i>marketplace</i> di Indonesia. Pengumpulan data ini akan menggunakan API Twitter dengan menerapkan kata kunci tertentu untuk melakukan identifikasi ulasan <i>tweets</i> yang berkaitan dengan tema <i>marketplace</i> . Selanjutnya, data yang telah dikumpulkan dapat dilakukan proses <i>labeling</i> dengan membagi tiga kelas yaitu positif, negatif, dan netral. Penelitian ini akan menggunakan model algoritma Decision Tree dengan dua metode yaitu <i>Cross Validation</i> dan <i>Split Validation</i> . Pengujian pada penelitian ini mendapatkan nilai akurasi tertinggi sebesar 70.27% dengan metode <i>cross validation</i> sedangkan dengan metode <i>split validation</i> mendapatkan nilai akurasi sebesar 66,95%.	Hasil dari penelitian ini mendapatkan bahwa algoritma Decision Tree mendapatkan akurasi tertinggi sebesar 70.27% dengan menggunakan metode <i>cross validation</i> . Namun, apabila menggunakan metode <i>split validation</i> akurasi menjadi menurun sebesar 66.95%. Hal ini dapat disimpulkan bahwa metode <i>cross validation</i> lebih unggul dibandingkan metode <i>split validation</i> .

Berdasarkan penelitian terdahulu yang telah melakukan analisis sentimen untuk berbagai aplikasi seperti pada tabel 2.1. Penelitian ini mempunyai kesamaan pada algoritma yang digunakan seperti beberapa penelitian terdahulu [10], [11], [12], [13], [14], [18], [19], dan [20] yaitu algoritma Decision Tree dan Random Forest. Penelitian ini juga mempunyai kesamaan dalam melakukan implementasi teknik *oversampling* SMOTE pada model seperti penelitian terdahulu [10] dan [13].

Persamaan lainnya berada pada letak objek yang dimana masih dalam satu area yaitu aplikasi *game*, seperti penelitian terdahulu [15] dan [16].

Perbedaan penelitian ini dengan penelitian terdahulu ialah penelitian ini akan menganalisis bagaimana sentimen pemain *game online* untuk aplikasi MLBB dan LOL serta menganalisis kelebihan dan kekurangan yang ada pada aplikasi MLBB dan LOL berdasarkan ulasan para pemain. Penelitian ini juga akan menggunakan algoritma Decision Tree dan Random Forest dan menguji kinerja model dengan menggunakan teknik SMOTE untuk mendapat model dengan akurasi terbaik. Penelitian terdahulu [10], [11], [12], [13], dan [14] akan menjadi penelitian acuan dari penelitian ini. Penelitian terdahulu tersebut akan dibandingkan tingkat nilai akurasi dengan penelitian ini berdasarkan model yang telah dibuat. Hasil perbandingan akan ditampilkan di bagian analisis dan pembahasan pada tabel 4.5.

2.2 Teori tentang Topik Skripsi

2.1.1 Game Online

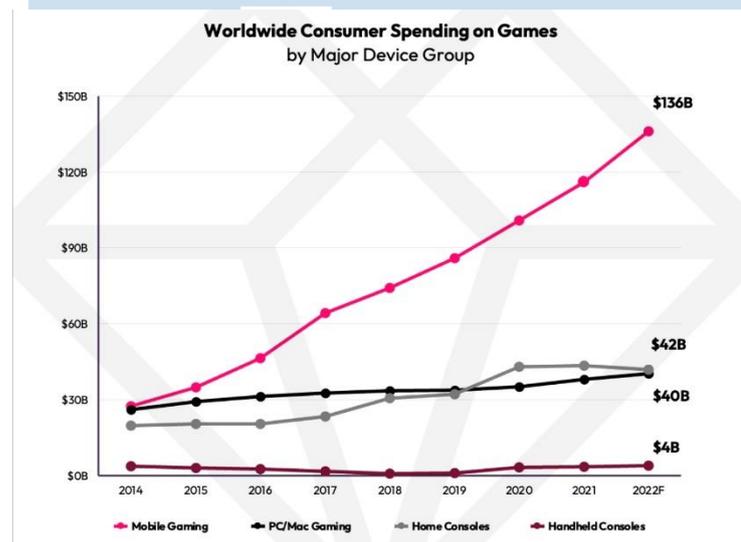
Game Online merupakan sebuah *game* atau permainan elektronik yang dapat diakses dengan menggunakan jaringan internet sehingga pemain dapat melakukan interaksi antar pemain lainnya [21]. Permainan daring atau *game online* merupakan suatu jenis hiburan permainan elektronik yang menggunakan jaringan internet atau komputer sebagai penyedia media, *game online* ini dapat dikatakan sebagai sebuah kemajuan teknologi dibandingkan hanya sebagai *genre* hiburan saja. Pada permainan *game online* ini terdapat suatu mekanisme sistem yang dapat menghubungkan pemain dengan pemain lainnya dengan pola tertentu pada suatu permainan [22].

2.1.2 Perkembangan Game Online

Permainan elektronik atau *game* pada awalnya dikembangkan di industri militer dimana internet diciptakan [23]. Sebelum terciptanya *game online* umumnya masyarakat memainkan permainan tanpa menggunakan jaringan internet atau *offline games*. Perkembangan permainan daring atau *online games* tidak lepas dari kemajuan teknologi dan jaringan internet komputer. Peningkatan signifikan pada *game online* tersebut adalah suatu pantulan pada

cepatnya perkembangan jaringan komputer yang pada awalnya hanya berskala kecil hingga menjadi sebuah jaringan internet dan makin terus berkembang hingga saat ini [23].

Pada era sekarang dengan berkembangnya teknologi telepon genggam menjadi telepon pintar atau *smartphone* tidak menutup kemungkinan industri *game online* untuk dapat masuk ke dalam *platform* tersebut [24].



Gambar 2. 1 Pendapatan *Games* Pada Setiap *Platform* [25]

Pada gambar 2.1 menunjukkan laporan bagaimana peningkatan pada setiap *platform* dalam kategori *game* dari tahun 2014 hingga 2022. Tercatat bahwa 61% *Mobile Games* mewakili dari jumlah keseluruhan pendapatan pada pasar *games* [25]. Hal ini menunjukkan bahwa industri *game online* di sektor *mobile* akan terus meningkat dari tahun ke tahun.

Berkembang pesatnya industri *game* tentunya, menciptakan berbagai jenis *genre* pada setiap *game* tersebut [26]. Beberapa *genre* pada *games* tersebut di antara lain:

1. RPG (*Role-Playing Games*)

Merupakan permainan dimana pemain dapat menjadi atau mengambil suatu peran dalam karakter pada *game* tersebut.

2. *Survival Games*

Merupakan permainan dimana pemain harus dapat bertahan hidup dengan waktu yang lama, namun pemain juga akan diberikan beberapa kondisi yang sulit untuk dijadikan tantangan tersendiri di dalam *game* tersebut.

3. *Fighting Games*

Merupakan permainan dimana pemain harus dapat menghadapi dan mengalahkan lawan baik itu AI (*Artificial Inteligents*) ataupun manusia asli, genre ini dapat melatih keterampilan pemain dalam segi pengetahuan kombo dan juga refleks [26].

4. *Horror Games*

Merupakan permainan yang hampir mirip dengan jenis *Survival Games*, namun bedanya pemain harus dapat bertahan dengan keadaan yang mengerikan disertai dengan latar belakang yang gelap dan horor.

5. *Multiplayer Online Battle Arena* (MOBA)

Merupakan permainan yang memerlukan kekompakan antara pemain satu dengan lainnya dalam menghadapi lawan, pemain dan timnya harus dapat menghancurkan pertahanan musuh untuk dapat memenangkan permainan tersebut. *Mobile Legends Bang-Bang* dan *League of Legends* merupakan *games* yang termasuk dalam jenis MOBA.

2.1.3 Game “*Mobile Legend: Bang-Bang*”

Mobile Legends: Bang-Bang atau MLBB merupakan sebuah *game Multiplayer Online Battle Arena* (MOBA) dari perusahaan di Shanghai, China yang dikembangkan oleh pengembang Moonton dan *game* ini diluncurkan pada tahun 2016 secara global di *platform* Android.

Mobile Legends: Bang Bang

Moonton
Pembelian dlm aplikasi

Play the 5v5 MOBA game on mobile with players worldwide.



4,1★
34,6 jt ulasan

500 jt+
Hasil download

Remaja

Instal di Windows

Instal pada perangkat lain

Google Play Game beta diperlukan untuk menginstal game ini di Windows. Dengan mendownload versi beta dan game, Anda menyetujui [Persyaratan Layanan Google](#) dan [Persyaratan Layanan Google Play](#). [Pelajari lebih lanjut](#).

Gambar 2. 2 *Mobile Legends: Bang Bang*

MLBB telah memiliki jumlah pemain hingga jutaan di dunia, seperti pada gambar 2.2 yang menampilkan jumlah unduhan *game* MLBB mencapai 500 juta unduhan di *Playstore* dan mempunyai bintang atau *rating* 4,1 dari 34,6 juta ulasan hingga tahun 2024 ini. MLBB merupakan permainan yang membutuhkan jaringan internet dikarenakan permainan ini mengkoneksikan pemain ke seluruh pemain di dunia. Pada permainan ini pemain akan tergabung dalam sebuah tim yang berisikan lima anggota dan harus bertarung dengan tim lainnya dengan tujuan menghancurkan markas lawan [27].

Gaya permainan pada *game* MLBB lebih mengutamakan kerjasama antar pemain untuk mencapai objektif bersama. Pemain dapat memilih *hero* atau karakter yang masing-masing karakter tersebut mempunyai kemampuan yang berbeda, dan pemain dapat melakukan kerja sama dalam membentuk strategi yang efektif untuk menghadapi lawan. Karakter atau *hero* pada permainan ini memiliki tingkat kesulitan yang beragam serta mempunyai kemampuan yang berbeda. Pemain dapat memilih berbagai karakter sesuai dengan bagaimana gaya bermain mereka, serta pemain dapat melengkapi peran mereka dalam permainan.

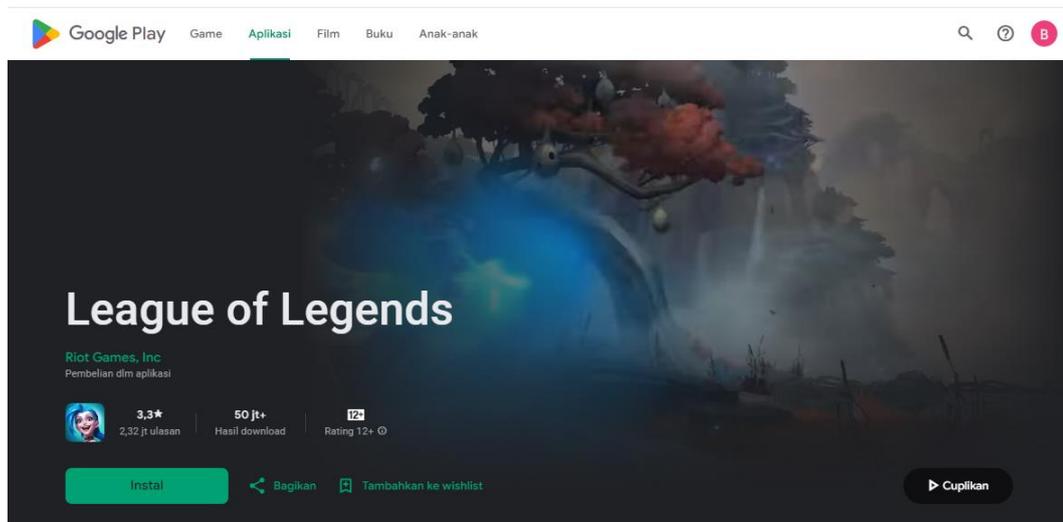
MLBB juga memiliki sistem poin untuk meningkatkan peringkat pemain, apabila pemain bermain dan memenangkan *game* tersebut maka peringkat pemain tersebut dapat meningkat dan dapat dijadikan sebagai pengukuran tingkat kemampuan pemain. Selaku pengembang, Moonton pastinya akan terus

melakukan pembaharuan pada *game* tersebut dengan menambahkan beberapa konten baru, penyesuaian pada karakter, dan gaya permainan serta memperbaiki *bugs*. Hal ini bertujuan dalam membuat permainan tersebut menjadi lebih nyaman dan menarik [27].

Komunitas pada *game* MLBB sangatlah besar dengan total unduhan 500 juta pada Google Playstore tercatat sekitar 81 juta pemain aktif yang masih bermain permainan tersebut hingga juli 2023 [28]. Pertumbuhan yang terus meningkat mencerminkan konsistensi pada *game* ini. Pengembang juga terus meningkatkan komunitas pada *game* ini dengan mengadakan beberapa turnamen dan berbagai *event*.

2.1.4 Game “League of Legends”

League of Legends atau LOL merupakan permainan setipe dengan MLBB yaitu MOBA yang dikembangkan oleh perusahaan RIOT Games berasal dari Los Angeles, Amerika Serikat [29]. Pada awalnya permainan LOL ini merupakan *game* yang hadir pada *platform* PC dan diadaptasi ke dalam *platform mobile* [29].



Gambar 2. 3 *League of Legends*

LOL resmi diluncurkan ke *platform* IOS dan Android pada tahun 2020 dan hingga saat ini LOL telah diunduh sebanyak 50 juta kali di situs Google

Playstore dan mempunyai bintang atau *rating* 3.4 dari 2.29 juta ulasan terhitung sejak Januari 2024 seperti pada gambar 2.3 yang diambil dari potongan situs Google Playstore.

Gaya permainan pada *game* LOL memberikan pengalaman bermain MOBA secara detail di perangkat seluler mereka. Pemain akan dibagi menjadi 5V5 dan mempunyai tujuan untuk saling berhadapan untuk menghancurkan markas dan “*Nexus*” lawan. Karakter atau dikenal dengan “*Champions*” dapat dipilih oleh pemain sebelum permainan dimulai. Setiap karakter mempunyai kemampuan dan tingkat kesusahan yang berbeda-beda serta pemilihan karakter dapat mempengaruhi strategi dalam permainan yang akan berlangsung.

Seperti pada *game* MOBA lainnya, LOL juga menerapkan sistem peringkat pada pemainnya. Setiap pertandingan dengan jenis “*Ranked Match*” akan menentukan peringkat setiap pemain, kemenangan pada setiap pertandingan akan meningkatkan peringkat pemain ke lebih atas. Penambahan konten-konten dalam *game* selalu dilakukan oleh pengembang, serta *bugs* yang ditemukan oleh pemain dapat dilaporkan kepada pihak pengembang untuk diperbaiki pada pembaharuan selanjutnya.

Komunitas untuk permainan LOL ini cukup besar terhitung hingga akhir Desember 2023 jumlah pemain aktif pada *game* tersebut mencapai 15,6 juta pemain [30]. Interaksi para pemain dalam permainan atau *in-game* menjadi sebuah pengalaman antar pemain untuk lebih dekat. Para pemain LOL juga dapat berinteraksi di luar dari *game* tersebut seperti di forum resmi, Discord, ataupun media sosial resminya.

2.1.5 Sentimen Analisis

Konsep sentimen analisis merupakan sebuah riset yang dilakukan dengan menggunakan data dari pernyataan, opini-opini, sentimen ataupun perasaan yang diekspresikan secara tertulis atau tekstual [31]. Proses analisis ini akan melakukan pengelompokan pada teks-teks sentimen yang telah terkumpul. Kalimat atau pendapat yang telah dikumpulkan dapat menghasilkan sebuah nilai yang mengartikan pendapat atau opini dengan nilai positif atau negatif.

Sentimen analisis dapat membantu suatu bisnis perusahaan dalam meningkatkan produk layanan mereka, beberapa manfaat yang diberikan ialah [32]:

1. Memberikan pandangan yang objektif
Proses bisnis dapat menghindarkan sebuah *bias* pribadi terkait dengan tanggapan negatif konsumen dengan menggunakan proses sentimen analisis. Akibatnya, perusahaan dapat menghasilkan hasil yang objektif dan konsisten dalam menganalisis opini konsumen.
2. Mengembangkan layanan produk lebih baik
Proses sentimen analisis akan membantu organisasi dalam meningkatkan layanan produk mereka, sesuai dengan *feedback* dari para konsumen yang spesifik dan positif.
3. Analisis dalam skala yang luas
Bisnis akan terus mendapatkan data informasi dalam jumlah yang besar dan tidak terstruktur dari *email*, survei, atau *feedback* produk. Sentimen analisis memungkinkan perusahaan dalam menghasilkan perasaan atau emosional konsumen dengan bentuk data tekstual [33].

2.1.5.1 Metode Analisis Sentimen

Metode sentimen analisis dapat dipergunakan untuk memahami dan melakukan evaluasi pada sentimen, tanggapan, opini, atau perasaan yang dikelompokkan dalam bentuk teks. Beberapa metode yang digunakan pada sentimen analisis melibatkan langkah berikut [34]:

1. Pengumpulan data
Informasi data yang didapatkan berasal dari berbagai sumber, seperti ulasan produk, *platform* sosial media, atau forum resmi. Data yang dikumpulkan dapat mencakup komentar, opini ataupun wawancara.
2. Pra-pemrosesan data dan pembersihan
Pada tahap ini akan melibatkan pembersihan pada informasi data yang tidak terkait atau relevan, seperti karakter tertentu, tautan, atau *noise* data. Setelah itu, dapat dilakukan pra-pemrosesan teks, seperti *stemming*, pembersihan kata penghubung, dan normalisasi data.

3. Pembobotan teks dan Tokenisasi

Kata pada teks akan dibagi-bagi menjadi sebuah token untuk di analisis lebih lanjut. Setiap token atau kata akan diberikan nilai bobot berdasarkan kepentingan pada konteks sentimen analisis. Proses ini dapat melibatkan teknik pembobotan yaitu TF-IDF (*Term Frequency-Inverse Document Frequency*).

4. Klasifikasi dengan *Machine Learning*

Proses selanjutnya ialah dengan menggunakan metode klasifikasi pada ML, digunakan untuk mengelompokkan dan mengidentifikasi sentimen dengan otomatis. Model yang dibuat dapat menggunakan informasi data yang telah dilabel sebelumnya.

5. Evaluasi

Tahap selanjutnya ialah melakukan evaluasi pada model yang telah terklasifikasi dengan menggunakan parameter seperti, *Accuracy*, *Precision*, dan *Recall*.

6. Visualisasi model

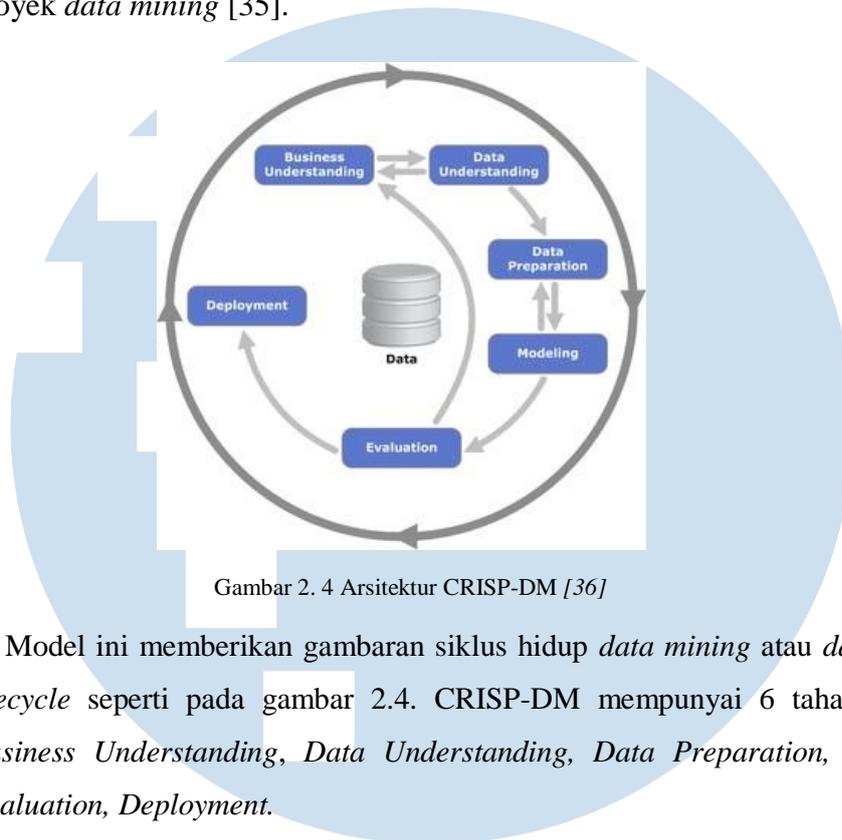
Hasil model dari sentimen analisis tersebut dapat ditampilkan dalam bentuk visual grafik, histogram atau *heatmap*. Tahap ini dapat membantu untuk lebih mudah memahami pendistribusian tren dan sentimen pada model.

2.3 Teori tentang Framework / Algoritma yang digunakan

2.2.1 Metode CRISP-DM

CRISP-DM atau *Cross-Industry Standard Process for Data Mining* merupakan sebuah model dari proses *data mining* yang bermanfaat dalam menyelesaikan masalah proyek *data mining* di kalangan industri. Proses CRISP-DM terdiri dari metodologi dan model proses untuk *data mining*, yang

memberikan pendekatan komprehensif kepada siapa pun untuk melakukan proyek *data mining* [35].



Gambar 2. 4 Arsitektur CRISP-DM [36]

Model ini memberikan gambaran siklus hidup *data mining* atau *data mining lifecycle* seperti pada gambar 2.4. CRISP-DM mempunyai 6 tahapan yaitu, *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, *Deployment*.

1. *Business Understanding*

Tahap pertama ini mungkin merupakan tahap paling penting dari setiap proyek *data mining*, ini mengarah pada pemahaman tujuan dan kebutuhan penelitian dari sudut pandang bisnis dan merubah pemahaman ini menjadi jawaban dari masalah *data mining* [37]. Pada tahap ini diperlukannya pemahaman mengenai latar belakang dan tujuan dari proses bisnis yang berhubungan dengan kasus yang dibahas.

a. Menentukan Tujuan Bisnis

Menentukan objektif bisnis pada penelitian yang dilakukan dengan mengenali pola/*pattern* pada kasus sehingga dapat menghasilkan sebuah prediksi.

b. Memahami Tujuan dari *Data Mining*

Memahami tujuan dari *data mining* pada kasus tersebut.

2. *Data Understanding*

Merupakan tahap kedua pada metode CRISP-DM yang bertujuan untuk menentukan secara tepat data apa yang akan dianalisis, mengidentifikasi kualitas data yang tersedia, dan membangun hubungan antar data. Oleh sebab itu tujuan utama dari tahap ini ialah untuk mempersiapkan kemungkinan kekurangan yang akan terjadi, melakukan pendekatan, dan identifikasi sub-kategoris. Pada kasus tertentu, pemahaman data akan dimulai dengan mengumpulkan data awal sehingga dapat diidentifikasi masalah pada data dan menentukan hipotesis awal pada data.

a. Mengumpulkan Data Awal

Pengumpulan data dapat dilakukan dengan beberapa cara seperti, observasi, kuesioner, ataupun wawancara. Data yang terkumpul dapat menjadi data awal yang menunjang penelitian mengenai kasus ini.

b. Mendeskripsikan Data

Bertujuan untuk melakukan analisis terhadap data awal serta memahami data tersebut sehingga dapat diolah pada tahap selanjutnya.

c. Eksplorasi Data

Bertujuan untuk mengolah dan melakukan eksplorasi pada data yang telah dideskripsikan serta melakukan analisis pada data tersebut.

d. Melakukan Verifikasi Data

Data yang telah diolah akan di verifikasi akan kualitasnya.

3. *Data Preparation*

Pada tahap ini mencakup pada semua kegiatan yang diperlukan dalam menyusun *dataset* terakhir untuk di analisis yang dibuat dari data mentah. Proses persiapan data berkemungkinan akan diulang berkali-kali dan tanpa adanya urutan yang ditentukan. Pada tahap ini dapat melakukan rancangan pada *dataset* yang sesuai dengan tujuan *data mining* dan tabel-tabel pada *dataset* yang memiliki keterkaitan dengan

objek penelitian yang akan dipilih. Selanjutnya data/tabel yang telah dipilih akan dibangun dengan menggunakan *query*. Tabel-tabel yang memiliki informasi yang berbeda mengenai objek yang sama akan digabungkan ke dalam *dataset* baru yang telah disiapkan pada proses awal. Terakhir dapat melakukan pembersihan data pada tabel sehingga tabel tersebut jadi lebih sederhana, hal ini akan membuat tabel lebih difokuskan pada data yang memiliki hubungan.

4. *Modeling*

Pada tahap ini teknik pemodelan akan dipilih dan diterapkan. Fase ini melibatkan penerapan dari berbagai algoritma *Machine Learning* (Regresi, Klasifikasi, Klustering) atau *Statiscal Methods*. Beberapa algoritma memiliki persyaratan khusus pada bentuk data, oleh sebab itu seringkali perlu kembali ke tahap *Data Preparation*. Pada tahap ini, peneliti dapat memilih algoritma yang cocok dalam memprediksi objek penelitian, lalu peneliti dapat membangun model sesuai dengan kriteria-kriteria yang tersedia.

5. *Evaluation*

Model-model yang telah dibuat harus menjalani fase evaluasi dan verifikasi secara menyeluruh terhadap langkah-langkah pembuatannya untuk memastikan bahwa model tersebut sesuai dengan tujuan. Tujuan dasar dari fase ini ialah untuk memastikan bahwa semua masalah telah ditangani secara benar. Apabila pola yang telah dihasilkan masih belum selesai maka diperlukannya analisis lebih lanjut sehingga dapat menghasilkan perbaikan yang diharapkan.

6. *Deployment*

Setelah model dibuat, diuji, dan dievaluasi pada validasi data, tahap *deployment* dapat diterapkan dengan berupa pembuatan laporan atau mengimplementasikan proses *data mining* pada *department* lain.

2.2.2 SMOTE

SMOTE atau *Synthetic Minority Over-sampling Technique* merupakan sebuah metode yang digunakan dalam melakukan penyeimbangan data kelas yang tidak seimbang [38]. Teknik ini bekerja dengan membuat sebuah sampel baru dari kelas minoritas dengan cara mengambil contoh dari kelas minoritas yang sudah ada dan membuat sampel serupa namun tidak identik. Nilai selisih vektor pada kelas minoritas dan nilai *nearest neighbor* pada kelas mayoritas akan diambil oleh SMOTE pada algoritma yang diproses. Selanjutnya, nilai tersebut akan dikalikan dengan angka yang acak antara 1 hingga 0. Hasil dari perhitungan tersebut akan di tambahkan ke dalam vektor fitur baru sehingga mendapatkan nilai vektor baru [39].

$$V_{new} = V_i + (M_i - V_i) \times P$$

Rumus 2. 1 Rumus SMOTE

Keterangan:

V_i = nilai vektor pada kelas minoritas

M_i = nilai *nearest neighbors*

P = nilai acak antar 1 hingga 0

2.2.3 Decision Tree

Decision Tree adalah salah satu metode klasifikasi yang paling menarik yang melibatkan konstruksi pohon keputusan, yang terdiri dari *node* keputusan yang dihubungkan oleh cabang dari *node root* ke *node daun (end)* [40]. Algoritma ini adalah struktur yang dapat digunakan untuk membagi sejumlah besar kumpulan data menjadi data yang lebih kecil dengan menerapkan seperangkat aturan keputusan di dalamnya. Atribut *node* keputusan akan diuji dan setiap hasilnya akan menghasilkan cabang, setiap cabang akan diarahkan ke *node* lainnya atau *end node* lain untuk mengambil keputusan [40]. Decision Tree merupakan struktur *flowchart* dimana setiap *node* merepresentasikan nilai

atribut dan setiap cabang merepresentasikan hasil pengujian, dan setiap daun merepresentasikan sebuah kelas atau distribusi kelas.

Untuk membedakan Decision Tree dengan algoritma klasifikasi lainnya, dapat dilihat melalui cara Decision Tree menyelesaikan masalah dimana ia melakukannya dengan menggunakan visualisasi seperti akar pohon, dimana Decision Tree memiliki beberapa terminologi di dalamnya, yaitu:

- a. *Root Node*: Ini mewakili populasi atau sampel lengkap, yang kemudian dipecah menjadi dua atau lebih kelompok yang homogen.
- b. *Splitting*: Ini adalah pembagian *node* menjadi dua atau lebih *sub-node*.
- c. *Decision Node*: *decision node* terbentuk ketika sebuah *sub-node* terbagi menjadi beberapa *sub-node*.
- d. *Leaf/Terminal Node*: *Leaf* atau *Terminal node* adalah *node* yang tidak membelah.
- e. *Pruning*: adalah proses menghilangkan *sub-node* dari sebuah *node* keputusan. Proses pemisahan dapat digambarkan sebagai kebalikan dari *Splitting*.
- f. *Branch/Sub-Tree*: *Branch* atau *sub-tree* adalah bagian dari keseluruhan pohon.
- g. *Parent dan Child Node*: *parent node* dari *sub-node* adalah *node* yang dibagi menjadi *sub-node*, sedangkan *sub-node* adalah anak-anak dari *parent node*.

Langkah dalam membuat algoritma *Decision Tree* sebagai berikut:

- a. Mengubah data (tabel) menjadi model pohon. Pada langkah ini kita menentukan atribut yang dipilih mulai dari *root*, *branch* hingga keputusan. Kita dapat menggunakan perhitungan *gainratio* setiap kriteria dengan data sampel untuk mencari atribut. Berikut ini adalah langkah-langkah *Gainratio*:

$$Gainratio(S, A) = \frac{Gain(S, A)}{SplitInformation(S, A)}$$

Rumus 2. 2 Rumus *Gainratio*

- *Gain* > Total Jumlah Informasi
 - *SplitInformation* > Atribut dengan banyak contoh.
- b. Mengubah model pohon menjadi *rule*. Rumus untuk menghasilkan aturan didefinisikan sebagai berikut:

IF premise THEN Conclucions

Node root dan *branch* akan menjadi premis dari *rule*, sedangkan *node* daun akan menjadi bagian dari kesimpulan (solusi). Setiap premis yang memuat satu atribut akan dihubungkan dengan disjungsi, sedangkan premis yang memiliki kelanjutan dari premis di cabang berikutnya akan dihubungkan dengan konjungsi.

- c. Sederhanakan *rule* (pemangkasan), langkah-langkahnya dilakukan sebagai berikut:
- Buat tabel distribusi dengan memulai semua nilai kejadian di setiap *rule*.
 - Menghitung tingkat kriteria independensi dalam suatu *rule*, antara atribut dan atribut target.
 - Menghilangkan kriteria yang dianggap tidak perlu, seperti tingkat independensi yang tinggi.

Berikut ini adalah beberapa manfaat dari *Decision Tree*:

- a. Bekerja dengan data kontinu dan kategorikal.
- b. Memiliki kemampuan untuk menghasilkan berbagai *output*.
- c. Dapat memahami temuan yang tepat, dan ketergantungan pohon dapat diukur dan dipercaya.

- d. Anda dapat mempelajari data, mengungkap variabel yang relevan, dan menemukan hubungan antara beberapa variabel. Menggunakan strategi ini dalam waktu yang sangat singkat untuk menyempurnakan variabel target dan membangun fitur baru.
- e. Mudah dipahami dan dikomunikasikan kepada orang lain.
- f. Berguna untuk membersihkan data dibandingkan dengan pendekatan lain, tidak memakan waktu lama karena angka yang hilang dan *outlier* memiliki pengaruh yang kecil setelah titik tertentu.
- g. Korelasi non-linier antar fitur tidak berpengaruh terhadap efisiensi dan efektivitas pohon keputusan.
- h. Mempersiapkan data tidak memakan waktu lama karena tidak memerlukan penggantian *missing value*, normalisasi data, atau lainnya.

2.2.4 Random Forest

Random Forest adalah metode klasifikasi yang menggunakan metode Decision Tree dalam menentukan klasifikasinya dengan memilih atribut acak pada setiap *node*. Suara terbanyak dari Decision Tree yang dikembalikan akan digunakan sebagai prosedur klasifikasi. Sebuah atribut acak yang dibawa dapat digunakan untuk membuat sebuah Random Forest. Sebuah Decision Tree dapat dihasilkan dengan menggunakan metode CART (*Classification and Regression Tree*) [41]. Decision Tree yang telah mencapai ukuran maksimum tidak akan di sederhanakan (*Pruning*) sehingga akan menghasilkan Random Forest dari kumpulan *trees* [42].

Random Forest adalah metode klasifikasi yang menggunakan sekumpulan Decision Tree terstruktur dengan vektor independen acak yang didistribusikan secara identik dan setiap Decision Tree memberikan suara unit untuk kelas paling populer pada *input X* [42].

Berikut ini adalah beberapa manfaat dari Random Forest:

1. Akurasi yang baik.
2. Relatif kuat terhadap *outlier* dan *noise*.

3. Lebih cepat dari *Bagging* dan *Boosting*.
4. Sederhana dan mudah untuk diparalelkan

Berikut ini adalah prosedur untuk membuat Random Forest:

1. Sepertiga dari sampel ditinggalkan ketika sampel *bootstrap* diproduksi dengan mengambil data sampel dan mengganti setiap Decision Tree.
2. Data OOB mengacu pada contoh yang telah ditinggalkan (*Out of Bags*).
3. Ada data di setiap Decision Tree di *Forest*. OOB digunakan untuk menghitung ketidakakuratan setiap Decision Tree.
4. Random Forest juga dapat menghasilkan tingkat signifikansi dan estimasi yang bervariasi. Estimasi adalah teknik untuk menghilangkan dan mengganti nilai serta *outlier* yang hilang [42].

2.4 Teori tentang Tools / Software yang digunakan

2.3.1 Python

Python ditemukan oleh Guido Van Rossum pada tahun 1990 di Belanda.

Python mempunyai beberapa keunggulan seperti berikut [43]:

1. *Readability*

Python memberikan *source code* yang sangat sederhana sehingga mudah untuk ditulis, diingat, serta dipakai ulang. Hal tersebut memberikan kemudahan dalam pengembangan aplikasi atau sistem dimulai dari tahap koding, *testing*, revisi, atau *error*.

2. Efisien

Python menyediakan *library* atau pustaka yang cukup lengkap, *Python code* lebih sederhana apabila dibandingkan dengan kode pada bahasa pemrograman lainnya.

3. Multifungsi

Jika menggunakan Python pengguna dapat membuat situs, robotika, aplikasi jaringan, hingga pembuatan kecerdasan buatan (AI). Python menyediakan banyak modul yang dapat dipakai dalam pengembangan aplikasi.

4. Interoperabilitas

Program-program Python dapat berinteraksi dengan bahasa pemrograman lainnya dan juga dapat dioperasikan pada hampir semua operasi sistem seperti Windows, Linux, Unix, atau MacOS.

2.3.2 Google Colaboratory

Google Collaboratory merupakan alat data analisis dan *Machine Learning* yang memungkinkan pengguna untuk menggabungkan kode Python dan *rich text* yang dapat dieksekusi bersamaan dengan *charts*, gambar, HTML, LaTeX dan lainnya menjadi satu dokumen yang tersimpan di *Google Drive* [44]. *Google Colab* memungkinkan pengguna untuk menjalankan kode Python tanpa harus melakukan proses instalasi ataupun setup lain. Beberapa keunggulan dari *Google Colab* yaitu:

1. *Libraries* yang sangat banyak

Pustaka atau *Library* pada *platform* ini menyediakan pustaka *Machine Learning* yang telah diunduh sebelumnya.

2. Penyimpanan *Cloud*

Google Colab menyediakan fitur penyimpanan *cloud* sehingga pengguna dapat mengakses file yang telah disimpan pada Google Drive dengan perangkat apapun.

3. Kolaborasi

Fitur kolaborasi berguna dalam pembuatan kode secara bersama sama oleh pengguna lainnya.

U M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A