

## BAB II

### LANDASAN TEORI

#### 2.1. Penelitian Terdahulu

Penelitian terdahulu bertujuan untuk mendapatkan bahan perbandingan dan juga acuan dalam melakukan penelitian. Selain itu juga untuk menghindari anggapan kesamaan dengan penelitian ini, maka dalam kajian pustaka ini peneliti mencantumkan hasil-hasil penelitian terdahulu sebagai berikut ini :

Tabel II.1 Hasil Penelitian yang Relevan

No	Judul	Nama & Tahun	Metode	Hasil
1.	<i>Students' Adaptability Level Prediction in Online Education using Machine Learning Approaches</i>	[3]	<i>Metode yang digunakan adalah Machine Learning</i>	Model <i>Random Forest</i> memiliki nilai akurasi terbaik yaitu 89.63% bila dibandingkan dengan model lainnya yaitu <i>Decision Tree</i> 87.56%, <i>Naïve Bayes</i> 70.95%, <i>K-Nearest Neighbour (KNN)</i> 76,348%, <i>Support Vector Machine (SVM)</i> 66.80%, dan <i>Artificial Neural Network (ANN)</i> 82.99%.
2.	<i>Rule Induction with CN2 Some Recent Improvements</i>	[4]		1. Bagaimana algoritma CN2 dapat digunakan untuk <i>generate rules</i> yang berurutan

				<p>maupun yang acak,</p> <p>2. Dari perubahan yang dilakukan telah meningkatkan performa dari algoritma CN2.</p>
3.	<p><i>Analisis Sentimen Pembelajaran Daring Menggunakan Metode Naïve Bayes, KNN, dan Decision Tree</i></p>	[5]	<p><i>Metode yang digunakan adalah Naïve Bayes, KNN, dan Decision Tree</i></p>	<p><i>Accuracy</i> dari tiga metode yang digunakan yaitu <i>Naïve Bayes</i> 55,49%, <i>KNN</i> 61,47%, dan <i>Decision Tree</i> 61.92%. Hasil <i>Precision</i> dari ketiga metode tersebut adalah <i>Naïve Bayes</i> 46.45%, <i>KNN</i> 53.90%, dan <i>Decision Tree</i> 73.63%. Dan hasil <i>Recall</i> dari ke tiga metode tersebut adalah <i>Naïve Bayes</i> 58.91%, <i>KNN</i> 40.07%, dan <i>Decision Tree</i> 11.42%. Maka dapat disimpulkan dari hasil tersebut bahwa metode <i>Decision Tree</i> memiliki tingkat <i>accuracy</i> yang paling tinggi dari kedua metode lainnya sebesar 61.92% <i>accuracy</i>, 73.63% <i>precision</i>, dan 11,42% <i>recall</i>.</p>
4.	<p><i>Memotivasi Belajar Dengan Menggunakan</i></p>	[6]	<p><i>Metode yang digunakan adalah</i></p>	<p>a. Motivasi belajar</p>

	<i>E-Learning</i>		<i>deskriptif</i>	<p>merupakan segala dorongan yang mampu memberikan energi dan semangat untuk melakukan Tindakan belajar demi mencapai arah dan tujuan yang diinginkan. Motivasi ini perlu dimiliki peserta didik agar tujuan pembelajaran dapat tercapai. Indikator motivasi intrinsik dapat dilihat melalui adanya hasrat dan keinginan berhasil, dorongan dan kebutuhan dalam belajar, serta harapan dan cita-cita masa depan. Sedangkan motivasi ekstrinsik dapat dilihat dari adanya penghargaan dalam belajar, kegiatan yang menarik dalam belajar dan lingkungan belajar yang kondusif. Dengan berbagai kelebihan yang</p>
--	-------------------	--	-------------------	--

				<p>ditawarkan, e-learning merupakan salah satu media pembelajaran yang dapat digunakan untuk memotivasi peserta didik.</p> <p>b. E-learning merupakan pembelajaran yang menggunakan fasilitas kecanggihan TIK yang memberikan kemudahan bagi pengguna untuk belajar mandiri, dimana saja, kapan saja, sehingga mendorong mereka senang belajar dan berupaya meningkatkan kompetensi belajarnya.</p> <p>c. Untuk memotivasi peserta didik baik secara intrinsic maupun ekstrinsik, penggunaan e-learning dalam proses pembelajaran harus memperhatikan prinsip-prinsip penggunaan serta kelebihan</p>
--	--	--	--	--

				<p>dan kekurangan e-learning.</p> <p>d. Proses pembelajaran dengan menggunakan e-learning hendaknya tidak menempatkan peserta didik hanya sebagai “pendengar” atau “penonton” saja, melainkan juga mendorong partisipasi aktif dari peserta didik untuk berinteraksi, berdialog, bekerja sama, berbagai dan membangun pengetahuan Bersama.</p> <p>e. Dalam menggunakan e-learning guru selain harus kreatif dan inovatif juga harus memiliki sikap kritis dalam memilih bahan pembelajaran, beretika baik dalam memanfaatkan bahan tersebut, menghindari penggunaan gambar-gambar atau audio yang kurang relevan</p>
--	--	--	--	--

				<p>dengan materi pembelajaran, mendorong partisipasi aktif dari peserta didik, memberikan perhatian dan menyediakan waktu lebih terhadap peserta didik yang mengalami kesulitan dalam menggunakan e-learning, professional serta memiliki motivasi untuk terus belajar dan meningkatkan pengetahuan dan keterampilan.</p>
5.	<p><i>Analisa Pemanfaatan Teknologi Informasi dalam Pembelajaran Jarak Jauh di Tengah Pandemi Virus Corona Covid-19</i></p>	[7]	<p><i>Metode yang digunakan adalah Observasi, Studi Pustaka, dan Wawancara</i></p>	<p>Bahwa pemanfaatan teknologi informasi dalam pembelajaran jarak jauh pada kampus Universitas Bina Sarana Informatika Cabang Salemba 22 Jakarta Pusat, kelas 12.4A.07 ditengah pademi virus corona covid-19 berdasarkan bahasan diatas bisa disimpulkan bahwa peranan teknologi informasi sangat membantu dalam</p>

				<p>proses pembelajaran jarak jauh ditengah pandemi virus corona covid-19 sehingga semua proses belajar mengajar dapat berjalan dengan baik dan berdasarkan hasil penelitian ini bisa disimpulkan meskipun Negara Indonesia sedang berjuang menghadapi pandemi virus corona covid-19, semua proses belajar mengajar bisa tetap dilakukan dengan baik berkat bantuan kemajuan teknologi informasi yang sudah sangat maju saat ini, internet bisa menghubungkan dosen dan mahasiswa melalui laman <i>e-learning</i>, <i>whatsapp</i>, <i>google class</i>, aplikasi <i>zoom</i> maupun <i>youtube</i>.</p>
6.	<p><i>Pembelajaran Daring pada Masa Pandemi Covid-19: Implementasinya pada Sekolah Menengah Pertama</i></p>	[8]	<p><i>Metode yang digunakan dalam penelitian ini adalah metode Analisis Data</i></p>	<p>MTs Miftahul Ulum, lamongan melakukan pembelajaran daring dengan memanfaatkan <i>whatsapp</i> karena dinilai lebih efektif karena</p>

				<p>mudah digunakan oleh siswa dengan biaya penggunaan <i>data</i> internet yang cenderung lebih murah, bila dibandingkan dengan <i>platform</i> aplikasi berbasis <i>online</i> lainnya. Namun kendala terbesar dalam melaksanakan pembelajaran daring adalah turunnya motivasi dan adanya kesedangan capaian belajar antar anak, terutama ketika guru hanya memberikan penugasan tanpa memberikan penguatan materi dan mendesain pembelajaran yang bermakna.</p>
7.	<p><i>Problematika Pembelajaran Daring di Masa Pandemi Covid-19 di SMA Dharma Praja Denpasar</i></p>	[9]	<p><i>Metode yang digunakan dalam penelitian ini adalah metode Kualitatif, Kuantitatif, dan R&amp;D</i></p>	<p>Dalam proses pembelajaran daring guru dapat membuat banyak variasi model pembelajaran, dimana dalam pembawaan guru harus terlihat gembira, membuat suasana yang ceria dan dapat menarik minat siswa untuk mengikuti pembelajaran yang guru tersebut bawakan. Bila</p>



				<p>pembelajaran daring dilakukan melalui <i>whatsapp</i>, guru jangan hanya memberikan tugas kepada siswa, tetapi juga memberikan materi untuk bisa digunakan siswa sebagai bahan belajar.</p> <p>Keberhasilan seorang guru dalam melakukan pembelajaran daring pada situasi pandemi Covid-19 ini adalah kemampuan guru dalam berinovasi merancang, dan meramu materi, metode pembelajaran, dan aplikasi apa yang sesuai dengan materi dan metode yang akan digunakan. Dimana kreatifitas adalah kunci utama bagi seorang guru untuk dapat memotifasi siswa tetap semangat dalam belajar secara daring dan tidak menjadi beban psikis bagi siswa.</p>
8.	<i>Kontribusi Pembelajaran E-Learning Melalui Media Whattsapp Terhadap Motivasi</i>	[10]	<i>Metode penelitian yang digunakan adalah Analisis Data</i>	Pembelajaran e-learning melalui media whatsapp berkontribusi terhadap motivasi

	<p><i>Belajar Siswa di MI Badrissalam Tahun 2020</i></p>			<p>belajar peserta didik MI Badrissalam pada tahun 2020. Dengan hasil uji regresi linier sederhana yaitu harga a sebesar 59,996 dan harga b sebesar 0,252. Hasil korelasi <math>r_{hitung} 0,204502 &gt; r_{tabel} 0,1986</math> yang menunjukkan adanya hubungan yang positif dan signifikan antara kedua variabel dengan tingkat hubungan rendah antara kedua variabel yaitu 0,2045. Kemudian dilakukan uji regresi linear sederhana dengan menggunakan uji t, diperoleh <math>t_{hitung}</math> sebesar 2,02553 dan <math>t_{tabel}</math> dengan taraf signifikansi 5% yaitu 1,9885. Karena <math>t_{hitung} 2,02533 &gt; t_{tabel} 1,9885</math> dapat ditarik kesimpulan bahwa pembelajaran e-learning melalui media WhatsApp berpengaruh terhadap motivasi belajar siswa. Pembelajaran e-learning melalui media whatsapp</p>
--	--	--	--	---

				berkontribusi rendah terhadap motivasi belajar peserta didik di MI Badrissalam pada tahun 2020. Dengan hasil pengujian diperoleh besar koefisien determinasinya adalah 0,04182 atau 4,1%. Jadi dapat diartikan bahwa pembelajaran e-learning melalui media whatsapp memberikan kontribusi yang rendah yaitu sebesar 4,2% terhadap motivasi belajar siswa dan 95,9% ditentukan faktor lain seperti lingkungan belajar, faktor keluarga, media belajar lain, faktor kecerdasan siswa, serta beberapa faktor lain baik dari dalam diri siswa maupun dari luar.
9.	<i>Membandingkan Teknik Data Mining untuk Memprediksi Prestasi Akademik Mahasiswa</i>	[11]	<i>Metode yang digunakan dalam penelitian ini adalah Data Mining</i>	kegiatan untuk memprediksi kinerja mahasiswa lebih cenderung berguna untuk membantu pendidik dan juga peserta didik dalam meningkatkan proses pembelajaran dan juga pengajaran. Sedangkan untuk

				teknik prediksi yang paling sering digunakan dalam melakukan prediksi ini adalah metode klasifikasi dalam <i>data mining</i> .
--	--	--	--	--

Tabel diatas berisikan tentang hasil penelitian yang relevan dengan penelitian yang sedang dilakukan, dimana relevansi dari penelitian yang diuraikan pada tabel diatas adalah kesamaan tentang pembahasan pembelajaran *online* dan juga perbandingan beberapa teknik *data mining*. Sedangkan untuk perbedaan dengan penelitian ini adalah pada penelitian ini lebih ditekankan penggunaan empat teknik *data mining* untuk dapat menentukan faktor apa saja yang mempengaruhi dalam penyelenggaraan pembelajaran *online*.

## 2.2. Pembelajaran Online *E-learning*

Pembelajaran *online* pada dasarnya adalah pembelajaran jarak jauh (PJJ). Sistem pembelajaran jarak jauh merupakan sistem yang sudah ada sejak pertengahan abad 18. Sejak awal, pembelajaran jarak jauh selalu menggunakan teknologi untuk pelaksanaan pembelajarannya, mulai dari teknologi paling sederhana hingga yang terkini. Secara singkat, sejarah perkembangan pembelajaran jarak jauh dapat dikelompokkan berdasarkan teknologi dominan yang digunakannya [2].

Menurut Munir Teknologi pembelajaran dapat dikelompokkan menjadi dua bagian yaitu *technology based learning* dan *technology based web learning*. *Technology based learning* terdiri dari dua teknologi informasi yaitu *audio information technologies*, seperti radio, telpon, *audio tape*, atau *voice mail*, dan *video information technologies*, seperti *video tape*, *video text*, atau *video messaging*. Dalam pelaksanaan pembelajaran terjadi kombinasi dari kedua teknologi tersebut yaitu *audio/data*, *video/data*, atau *audio/video*. Sedangkan, *technology based web learning* pada dasarnya adalah *data information technologies* seperti *internet* yang di dalamnya terdapat *e-mail*, *bulletin board*, atau *tele-collaboration*. Komunikasi antara pembelajar dan pengajar agar bisa berjalan dengan baik dan efektif, maka diperlukan interaksi yang aktif dari keduanya yang bisa dilaksanakan

melalui cara langsung (*synchronous*) atau cara tidak langsung, (*asynchronous*), misalnya pesan direkam dahulu sebelum digunakan [1].

*E-learning* sering disebut juga sebagai pembelajaran *online* atau *online course*. Pembelajaran *online* dalam pelaksanaannya memanfaatkan dukungan jasa teknologi, khususnya teknologi informasi dan komunikasi, seperti komputer, telpon, *audio*, *video*, transmisi satelit, dan sebagainya. Pembelajaran *online* ini memungkinkan untuk menyelenggarakan pendidikan jarak jauh yang bisa menjangkau lebih banyak orang dan berbagai tempat sampai daerah terpencil atau pedalaman sekalipun yang membutuhkan pendidikan [1].

### **2.3. Kaggle**

*Kaggle* adalah layanan tambahan dari Google LLC, yang merupakan sebuah *online foundation* yang diperuntukkan untuk dapat digunakan oleh investigator *data* dan para profesional dalam *software engineering*. *Kaggle* memberdayakan para pelanggannya untuk dapat menemukan dan mengirimkan *dataset*, penelitian dan *build models* yang berhubungan dengan ruang lingkup *data science*, hasil kerja dengan banyak *data investigators* dan profesional dalam bidang AI lainnya, dan bertujuan untuk menjawab tantangan *data science*.

*Kaggle* memulai kompetisi AI mulai dari tahun 2010 dan sekarang juga mulai memberikan *data interface* terbuka, sebuah *workbench data science* yang berbasis *cloud*, dan pembelajaran *Artificial Intelligence* [12].

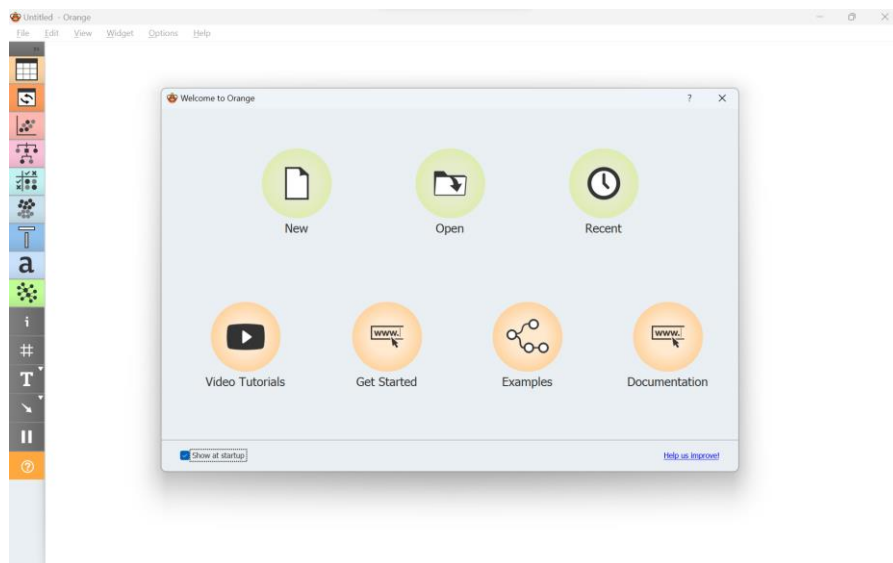
### **2.4. Orange**

*Orange* adalah sebuah *data mining dan visualization toolbox* yang diperuntukkan untuk pemula maupun ahli. Untuk dapat mengeksplorasi *data* dengan *Orange*, tidak diperlukan pemrograman ataupun pengetahuan matematika yang mendalam [13].

*Orange* adalah sebuah *software* untuk melakukan visualisasi dan analisis *data* yang bersifat *open source* untuk pemula maupun ahli. Dimana *data* dapat di tambang (*mined*) melalui *visual programming* ataupun *Python scripting*. Dimana *Orange* juga

menyertakan komponen untuk semua tugas penambangan *data* (*data mining*), visualisasi, dan *machine learning* yang populer[14].

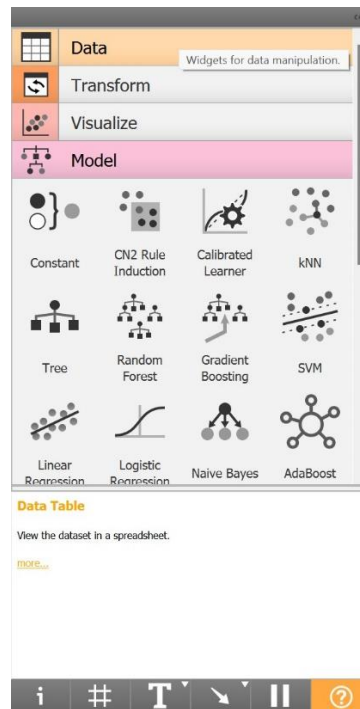
*Orange* adalah sebuah *data mining tool* yang dapat digunakan untuk melakukan *visual programming* dan analisis *data* yang lebih eksploratif, yang dapat di tulis menggunakan *Python*. *Orange* memiliki banyak komponen yang dikenal dengan *widgets*. *Orange* juga mendukung *macOS*, *Windows*, dan *Linux* [15].



Gambar II.1. Tampilan awal *Orange Data Mining Tools*

### 2.4.1. Orange Widgets

*Orange widgets* memberikan *graphical UI* kedalam penggunaan *Orange's data mining* dan metode AI. *Widgets* ini menjalankan entri informasi dan *preprocessing widgets*, klasifikasi, regresi, *order strategies*, dan *clustering*. Pilihan *widgets* untuk penilaian model dan hasil dari persepsi penilaian, seperti *widgets* untuk mentransfer mode ke PMML (*Predictive Model Markup Model*). *Widgets* merepresentasikan informasi melalui *tokens* yang mana beralih dari *widget* pengirim ke penerima. Sebagai contoh, informasi [12].



Gambar II.2. Tampilan *Widgets* pada *Orange Data Mining*

## 2.5. Data Mining

*Data mining* adalah sebuah proses untuk menemukan korelasi baru yang berarti, pola, dan tren dengan menyaring sejumlah besar *data* yang disimpan dalam *repositories*, menggunakan *pattern recognition technologies* serta teknik statistik dan juga teknik matematika [16].

*Data mining* adalah sebuah *science*, *art*, dan juga *technology* untuk mengeksplorasi kumpulan dari sebuah *data* yang besar dan juga kompleks untuk menemukan pola yang berguna [17].

*Data mining* merujuk pada proses ekstraksi atau “*mining*” pengetahuan dari sebuah kumpulan *data* yang besar. Dimana istilah tersebut sebenarnya keliru. Perlu diingat bahwasannya penambangan sebuah emas dari batu dan pasir tetap disebut sebagai penambangan emas bukannya disebut penambangan batu atau pasir. Seharusnya *data mining* seharusnya lebih tepat disebut sebagai “*knowledge mining from data*,” yang sayangnya sebutan tersebut sangat Panjang. “*Knowledge mining*,” adalah istilah yang lebih pendek, namun tidak mencerminkan penekanan pada *mining* kumpulan *data* yang besar [18].

## 2.6. Cross Validation Method

Konsep dasar dari metode *Cross Validation* (CV) adalah menjalankan sebuah model menggunakan *k-1 folds* yang dijalankan sebagai *data training*. Selanjutnya, hasil dari model tersebut divalidasi pada bagian *data* yang tersisa, yang tidak lain adalah *data testing*. Adanya validasi dapat digunakan untuk mengetahui seberapa akurat model klasifikasi yang dihasilkan [19].

## 2.7. KNN

*K-Nearest Neighbor* adalah algoritma klasifikasi *data* yang berdasarkan *data training* dengan jarak yang paling dekat dengan objek [5].

$$d(x, y) = \sqrt{\sum xi - yi} \tag{1}$$

## 2.8. Decision Tree

*Decision Tree* adalah algoritma klasifikasi *data* yang menggunakan pohon keputusan. Algoritma *Decision Tree* memiliki dua input yaitu *tarin sample* dan *sample*. *Training samples* merupakan contoh *data* dipakai dalam membangun *tree* yang telah diproses keasliannya, kemudian *sample* digunakan untuk klasifikasi *data* [5].

$$Entropy (A) = \sum -pi \log_2 p_i \tag{2}$$

## 2.9. CN2 Induction Rule

*CN2 Induction Rule* pada model ini harus melakukan evaluasi *rule* yang telah ditemukan untuk dapat memutuskan mana yang terbaik. Aturan kualitas yang memungkinkan adalah akurasi dalam *data* yang telah dilakukan pelatihan. Pada pengujian ini, model *CN-2 Rule* memiliki nilai akurasi yang cukup baik [20].



Algoritma *CN2 Induction Rule* didesain untuk induksi yang efisien dari aturan bentuk yang sederhana dan dapat dipahami “*if...then...*”, bahkan dalam *domain* dimana *noise* mungkin ada [4].

## 2.10. Random Forest

*Random Forest* adalah metode *ensemble* dalam hal *learning* yang digunakan untuk klasifikasi, regresi, dan *task* lainnya. Kinerja *Random Forest* diadaptasi dari *decision tree*, dengan setiap *tree*, subset atribut diambil secara acak dari atribut terbaik untuk dipilih secara *split*. Akhir model dari *random forest* didasarkan pada hasil dari keseluruhan subset *tree* yang telah dikembangkan [20].

$$\bar{r}(X, D_n) = E_{\theta}[X, \theta, D_n] \quad (3)$$

## 2.11. Klasifikasi

Klasifikasi merupakan teknik dalam *data mining* untuk mengelompokkan *data* berdasarkan keterikatan *data* terhadap *data* sampel [20].

Klasifikasi merupakan salah satu metode analisis yang terdapat dalam *data mining*. Klasifikasi dapat memetakan *data* ke suatu kelas/grup yang telah di tentukan sebelumnya. Hal ini disebut dengan *supervised learning* karena sebelum melakukan klasifikasi, kelas dari setiap obserbasi telah ditentukan terlebih dahulu [19].

Klasifikasi *data mining* adalah penempatan objek-objek ke salah satu dari beberapa kategori yang telah ditetapkan sebelumnya. Klasifikasi banyak digunakan untuk memprediksi kelas pada suatu label tertentu, yaitu dengan mengklasifikasi *data* (membangun model) berdasarkan *training set* dan nilai-nilai (label kelas) dalam mengklasifikasikan atribut tertentu dan menggunakannya dalam mengklasifikasin *data* yang baru [21].

Klasifikasi (*Classification*) adalah teknik *data mining* klasik berdasarkan *Machine Learning*. Pada dasarnya, klasifikasi digunakan untuk mengklasifikasikan setiap *item* dalam satu set *data* menjadi salah satu dari sekumpulan kelas atau kelompok yang telah ditentukan sebelumnya. Metode klasifikasi menggunakan teknik matematika seperti pohon keputusan, pemerograman linier, jaringan saraf dan statistik [11].

## 2.12. Elemen Pengukuran

*Confusion matrix* digunakan untuk mengindikasikan property dari sebuah *classification rule*. Yang di dalamnya terdapat jumlah elemen yang telah di klasifikasikan dengan benar maupun tidak benar dari setiap *class* [17].

Sebuah *confusion matrix* meringkas performa klasifikasi dari sebuah *classifier* yang terhubung pada sebuah *data test*, ini adalah sebuah *matrix* dua dimensi yang di kelompokkan dalam satu dimensi oleh sebuah *true class* dari sebuah objek dan juga terkadang oleh *class* yang di tentukan oleh *classifier* [22].

Pada Tabel dibawah ini menampilkan contoh dari *confusion matrix* untuk pengklasifikasian tiga kelas, dengan kelas A, B, dan C. Dimana untuk menghitung *confusion matrix* dengan tiga *class* maka kita harus menghitung nilai *True Positive* (TP) dimana pada tabel dibawah adalah AA, BB, dan CC kemudian dibagi dengan jumlah nilai *data* yaitu

$$AA + AB + AC + BA + BB + BC + CA + CB + CC \tag{4}$$

Tabel II.2. Contoh *Confusion Matrix* dengan 3 *Class*

Actual class	Assigned class		
	A	B	C
A	AA	AB	AC
B	BA	BB	BC
C	CA	CB	CC

$$Accuracy = \frac{AA + BB + CC}{AA + AB + AC + BA + BB + BC + CA + CB + CC} \tag{5}$$

Dapat juga persamaan ini disederhanakan menjadi seperti dibawah ini :

$$Accuracy = \frac{TP}{Jumlah\ Data}$$

(6)

Tabel II.3. Pengujian *confusion matrix*

<i>Actual class</i>	<i>Assigned Class</i>	
	<i>Positive</i>	<i>Negative</i>
<i>Positive</i>	TP	FN
<i>Negative</i>	FP	TN

Dalam beberapa kasus *confusion matrix* biasa menggunakan 2 kelas yaitu *Positive Class* dan juga *Negative Class* dalam kasus konteks ini ke empat *cells* dari *matrix* di desain menjadi *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN), seperti yang dapat dilihat pada Tabel 2 [22].

$$Precision = \frac{TP}{TP + FP}$$

(7)

$$Recall = \frac{TP}{TP + FN}$$

(8)

$$F1 = 2 \times \frac{precision \times recall}{precision + recall}$$

(9)