

BAB II

LANDASAN TEORI

2.1 Work From Home, Positive and Negative Impacts

Work From Home (WFH) didefinisikan sebagai pengaturan kerja jarak jauh yang memfasilitasi orang untuk bekerja secara fleksibel dari lokasi mana pun selain dari kantor secara konvensional [8]. WFH dikenal juga sebagai *teleworking/telecommuting*, atau *remote working*, tetapi walaupun begitu terdapat perbedaan antara *teleworking* dan *remote working*, dimana *remote working* berarti karyawan tinggal di luar wilayah geografis tempat kerja perusahaan sedangkan *teleworking/telecommuting* menyiratkan bahwa karyawan mungkin melakukan beberapa pekerjaan di tempat lain [7], namun untuk tujuan penelitian ini, istilah *working from home* atau WFH akan digunakan sebagai pilihan untuk merujuk pada kerja virtual atau bekerja jauh dari kantor.

WFH mampu mengubah dinamika dan lingkungan tempat kerja, yang mungkin dapat menguntungkan atau merugikan karyawan serta perusahaan. *Work From Home* saat ini juga menjadi salah satu alternatif yang memberikan peluang untuk lebih melindungi masyarakat setempat agar tidak tertular COVID-19.

Penelitian [9] [10] menyatakan bahwa WFH dicirikan dan mempunyai kelebihan yaitu fleksibilitas lebih besar, *work-life balance* lebih baik, pekerjaan tanpa gangguan, fokus yang lebih besar hingga mampu meningkatkan motivasi dan inovasi, tetapi penelitian lain [11] mengungkapkan bahwa WFH cenderung mengakibatkan penurunan motivasi kerja karena terdapat pola pikir bahwa “persepsi rumah sebagai tempat istirahat”, serta keterbatasan komunikasi hingga lebih banyak distraksi kerja saat di rumah, oleh karena itu penelitian ini akan memberikan *insight* baru mengenai kemungkinan WFH untuk tetap diterapkan khususnya pada GORP Kompas Gramedia Palmerah berdasarkan preferensi karyawan dan faktor apa saja yang cukup berpengaruh.

2.2 Employee Productivity

Penelitian [11] menyebutkan bahwa WFH memberikan dampak positif dan negatif dalam produktivitas karyawan. Keberagaman dampak ini dipengaruhi juga

oleh beberapa faktor yaitu faktor individual seperti keterampilan digital dan motivasi kerja, faktor kelompok seperti kepemimpinan dan *teamwork*, faktor organisasi seperti manajemen kinerja dan *support top level management* [2]. Produktivitas dapat dicerminkan dalam hasil kualitatif yang diukur berdasarkan tingkat fokus karyawan, *work-life balance*, dan inovasi serta kreatifitas karyawan, sedangkan hasil kuantitatif diukur berdasarkan kecepatan *problem solving*, banyaknya pekerjaan yang mampu diselesaikan hingga *cost* yang digunakan [2].

Hasil penelitian [12][13][14] menyatakan bahwa WFH membawa dampak positif sehingga terdapat peningkatan produktivitas selama WFH berdasarkan survey di dua negara berbeda yaitu Jepang dan America, sedangkan penelitian [11][15] menyebutkan bahwa WFH membawa dampak negatif pada produktivitas kerja sehingga tidak terdapat perubahan produktivitas, bahkan cenderung menurunkan produktivitas karyawan selama WFH. Penelitian [16] menyebutkan bahwa produktivitas pekerja selama pandemi berbeda-beda tergantung pengalaman WFH karyawan sebelum adanya pandemi COVID-19, sehingga berdasarkan beberapa jurnal sebelumnya, faktor produktivitas perlu diteliti lebih lanjut apakah mampu mempengaruhi kelanjutan adopsi WFH pada Kompas Gramedia Palmerah.

2.3 Employee Performance

Work From Home mempunyai dampak pada *employee performance* dengan indikator pendukungnya seperti *work rules*, *environment* hingga *communication*. Dampak positif ini terlihat dari penelitian [17] yang menyebutkan bahwa WFH membawa pengaruh yang cukup signifikan terhadap *employee performance* dan *top-level management support* karena efektivitas dan efisiensi WFH tidak dapat dipisahkan dari monitoring dan evaluasi oleh *top-level management*. Penelitian [18] juga menyebutkan bahwa *employee performance* dipengaruhi oleh tiga faktor utama yaitu individu (visi, pengetahuan, skill), lingkungan dan *job demand* (tugas dan scope yang diberikan pada karyawan). Berdasarkan dua penelitian sebelumnya, sudah terbukti bahwa WFH cukup mempengaruhi *employee performance*, maka dari itu penelitian kali ini juga akan meneliti lebih lanjut apakah faktor *performance* mampu mempengaruhi kelanjutan adopsi WFH pada Kompas Gramedia Palmerah.

2.4 Work Improvement

Kebijakan pemerintah mengenai pemberlakuan WFH selama pandemi COVID-19 di Indonesia ternyata membawa dampak positif dalam peningkatan kerja karyawan seperti meningkatkan motivasi kerja karyawan, produktivitas hingga membuat karyawan mampu menjalani *work-life balance* dalam bekerja [17], sehingga penelitian kali ini akan melihat apakah peningkatan kerja tersebut mampu menjadi faktor pengaruh untuk melanjutkan kebijakan WFH pada PT Kompas Gramedia Palmerah setelah pandemi COVID-19 berakhir.

2.5 Work Environment

Work environment atau lingkungan kerja dianggap penting saat WFH, karena berdasarkan penelitian [18], *work environment* saat bekerja dari rumah mampu menciptakan lingkungan kerja lebih fleksibel karena karyawan bebas mengatur jadwal kerja dan lingkungan kerja sesuai dengan kebutuhan dan kepribadian individu mereka. Penelitian tersebut juga menyebutkan bahwa lingkungan kerja yang baik dan nyaman akan membawa dampak positif dan membantu karyawan untuk meningkatkan *performance* kerja mereka, maka dari itu penelitian kali ini akan dilakukan untuk melihat apakah lingkungan kerja cukup berpengaruh dan menjadi salah satu faktor adopsi kebijakan WFH dapat terus dilanjutkan setelah pandemi COVID-19 pada PT Kompas Gramedia Palmerah.

2.6 Communication & Support

Komunikasi merupakan cara penyampaian data atau informasi dari satu pihak ke pihak lain, sehingga komunikasi merupakan aspek penting dari proses manajemen karena jika terdapat *misscommunication* dapat berakibat fatal. Komunikasi dikatakan efektif jika terarah, dan berdampak positif bagi dua belah pihak [19]. Semenjak penerapan WFH, komunikasi dilakukan secara *remote* dengan *tools* komunikasi virtual. Perubahan metode komunikasi ini menjadi tantangan bagi perusahaan karena dalam berkomunikasi secara virtual, karyawan merasa rapat secara *online* menjadi tidak nyaman dan sedikit terlalu personal karena kamera terlalu *close up* [20], sehingga peran pemimpin sangat dibutuhkan dalam memberikan pengaruh terhadap kinerja karyawan [20].

Masalah lain yang timbul saat berkomunikasi *online* adalah terhambat karena koneksi internet yang buruk serta menjaga energi untuk tetap positif selama WFH [19]. Berdasarkan penelitian [20], dukungan dan motivasi secara emosional dan sosial dari supervisi sangat mempengaruhi *performance* dan kepuasan karyawan, sehingga pada penelitian kali ini akan dilihat apakah faktor komunikasi dan support cukup berpengaruh terhadap kelanjutan adopsi WFH berdasarkan preferensi karyawan GORP PT Kompas Gramedia Palmerah.

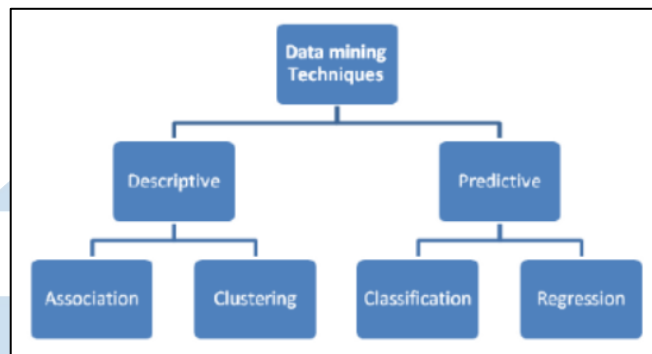
2.7 Continuous Adoption Intention

Continuous Adoption Intention atau niat adopsi berkelanjutan mengacu pada sejauh mana seorang individu berniat untuk menerima, melakukan dan mengadopsi tindakan tertentu secara terus-menerus dan konsisten [6]. Niat adopsi juga didorong oleh beberapa instrumental dan pertimbangan lain hingga seseorang memiliki niat positif yang berkelanjutan terhadap perilaku yang sama.

2.8 Data Mining

2.8.1 Data Mining Technique

Data mining merupakan teknik atau proses untuk mengekstrak data dan menganalisa data dari berbagai bidang seperti *database*, *machine learning*, statistik, *pattern recognition*, *neural networks*, *artificial intelligence*, hingga *data visualization* [21]. Aplikasi dari *data mining technique* sangat luas dan dapat digunakan di semua bidang operasi bisnis, *research scientific*, hingga *decision making*. *Data mining tools* akan memprediksi perilaku data dan membantu mengambil keputusan yang dibantu oleh *knowledge* tertentu. Tujuan akhir dari *data mining* adalah mengoptimalkan keputusan dan perilaku bisnis berdasarkan hasil analisis yang diperoleh untuk meningkatkan bisnis tertentu [22]. Teknik *data mining* dibagi lagi menjadi dua kelompok yaitu *descriptive analytics* dan *predictive analytics* sebagai berikut.



Gambar 2. 1 Data Mining Techniques [21]

2.8.2 Data Mining Process

Terdapat beberapa proses yang terlibat dalam implementasi *data mining technique* yaitu [23]:

1. *Business Research*: pemahaman yang lengkap mengenai tujuan perusahaan, sumber daya yang tersedia, dan skenario yang sejalan dengan *requirement* yang ada harus dilakukan sebelum memulai implementasi *data mining technique*.
2. *Data Quality Checks*: data yang dikumpulkan dari berbagai sumber perlu diperiksa dan dicocokkan untuk memastikan tidak ada hambatan dalam proses integrasi data.
3. *Data Cleansing*: *data cleansing* atau pembersihan data biasa dilakukan dalam *pre-processing model* untuk membersihkan data dari *noise*, *redundant data*, *missing values* hingga *outliers*.
4. *Data Transformation*: tahap dari transformasi data yang bisa dilakukan dengan beberapa pilihan yaitu *data smoothing*, *data summary*, *data generalization*, *data normalization* hingga *data attribute construction*.
5. *Data Modeling*: terakhir, untuk identifikasi pola data yang lebih baik maka diperlukan *modelling* data yang diimplementasikan dalam dataset dengan algoritma tertentu.

2.8.3 CRISP-DM

CRISP-DM (*Cross Industry Standard Process for Data Mining*) merupakan *process model* yang dianggap sebagai prinsip panduan paling

relevan untuk proses *analytics* [24]. CRISP-DM biasanya digunakan untuk membantu menerjemahkan masalah bisnis ke dalam tugas dan proses *analytics*, menyarankan transformasi data dan teknik *analytics* yang tepat serta memberikan saran untuk mengevaluasi hasil dari model yang sudah dikembangkan [24]. CRISP-DM merupakan *framework data mining* yang memiliki siklus hidup terbagi dalam enam fase pengembangan yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* dan *deployment*.

Pada Tabel 2.1 merupakan penjelasan *framework* atau *life-cycle* CRISP-DM dalam implementasi *data mining*.

Tabel 2.1 Tahapan Framework CRISP-DM [24]

Phase	Description
<i>Business Understanding</i>	Tahapan pertama saat situasi atau tujuan bisnis harus dinilai untuk mendapatkan gambaran tentang <i>requirement</i> yang dibutuhkan. Penentuan tujuan data mining merupakan salah satu aspek terpenting dalam fase ini.
<i>Data Understanding</i>	Tahap kedua untuk mengumpulkan data, mengeksplorasi, memahami serta memeriksa kualitas data yang telah didapatkan.
<i>Data Preparation</i>	Tahap ketiga adalah persiapan data. Persiapan dan pemilihan data harus dilakukan dengan menetapkan kriteria dan kualitas data yang dapat ditangani oleh model klasifikasi dengan cara <i>cleansing data</i> .
<i>Modeling</i>	Tahap keempat adalah tahap pmodelan data yang terdiri dari pemilihan teknik pemodelan klasifikasi yang sesuai dengan tujuan, <i>business requirement</i> serta jenis data yang didapatkan.
<i>Evaluation</i>	Tahap kelima adalah tahap evaluasi. Hasil dari proses pengembangan moel akan dilakukan evaluasi atau pemeriksaan terhadap kesesuaian tujuan bisnis yang telah ditetapkan. Evaluasi juga bisa dibantu dengan beberapa <i>performance measurement</i> seperti penggunaan <i>confusion matrix</i> untuk menghitung nilai akurasi, precision dan recall dari model yang sudah dikembangkan.
<i>Deployment</i>	Tahapan terakhir adalah deployment, atau fase penyebaran serta penerapan seluruh hasil yang ada. Deployment bisa berupa laporan akhir atau saran dan sugesti yang bisa diberikan untuk mendukung <i>decision making</i> dari <i>top-level management</i> .

2.9 Predictive Analytics

2.9.1 Pengertian Predictive Analytics

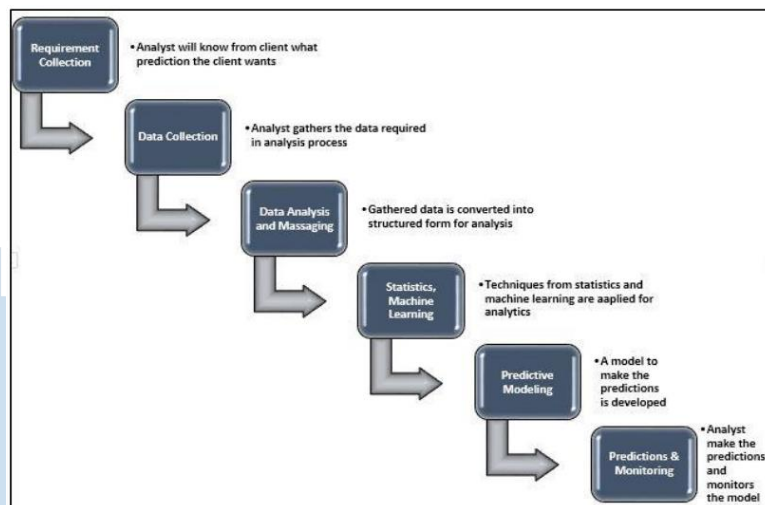
Predictive analytics merupakan kategori analisa data yang bertujuan untuk memprediksi hasil di masa depan berdasarkan data saat ini atau masa lalu (*historical data*) dan teknik analisis seperti *statistical modelling*, *data mining* hingga *machine learning* [25]. Analisa ini menggunakan hubungan yang dikonfirmasi antara *explanatory variables* dan *criterion variables* dari masa lalu untuk memprediksi hasil di masa depan. Tujuan dari *predictive analytics* sendiri yaitu menghasilkan informasi yang relevan, mempunyai keputusan yang lebih cerdas dan dapat digunakan untuk memprediksi peristiwa masa depan dalam data yang sangat besar [26].

Predictive analytics dapat diterapkan dalam *data mining* untuk memprediksi beberapa peristiwa masa depan pada beberapa sektor yaitu medis, bisnis, pendidikan hingga deteksi kejahatan. Aplikasi dari analisa ini nantinya akan menyatukan teknologi informasi, *business modelling process* dan prediksi untuk masa depan.

2.9.2 Predictive Analytics Process

Predictive analytics melibatkan beberapa langkah dimana data akan dianalisa dan diprediksi berdasarkan data saat ini atau data di masa lalu (*historical data*). Proses prediksi bukanlah satu langkah untuk membuat prediksi dan keputusan di masa depan, melainkan proses langkah demi langkah yang melibatkan banyak proses dari pengumpulan kebutuhan, pengembangan model hingga pemantauan untuk melihat kebenaran model sehingga dapat dijadikan patokan dalam proses pengambilan keputusan. Berdasarkan Gambar 2.2 merupakan proses tahapan dari *predictive analytics*.

UNIVERSITAS
MULTIMEDIA
NUSANTARA



Gambar 2. 2 Predictive Analytics Process [25]

1. Requirement Collection

Proses pertama dalam *predictive analytics* adalah mengetahui *requirement* yang dibutuhkan, karena tujuan dan *knowledge* sangat dibutuhkan dalam pengembangan model prediksi.

2. Data Collection

Proses selanjutnya adalah proses *data collection* atau pengumpulan data yang diperlukan dalam pengembangan model prediktif. Data yang diambil mungkin sudah mempunyai struktur yang jelas (*structured data*) atau bahkan data mentah yang tidak terstruktur (*unstructured data*).

3. Data Analysis and Massaging

Proses ketiga adalah analisa data dan *preparation* sebelum digunakan dalam pengembangan model prediktif. Data tidak terstruktur (*unstructured data*) nantinya akan diubah menjadi data terstruktur berdasarkan format yang sesuai, karena efektivitas model prediktif sepenuhnya tergantung pada kualitas data.

4. Statistics, Machine Learning

Proses *predictive analytics* menggunakan beberapa teknik *statistics*, *data mining* dan *machine learning*, seperti *neural networks*, *decision tree*, *support vector machine* dan model lain.

Semua model *predictive analytics* didasarkan pada teknik statistik

dan *machine learning*, oleh karena itu penerapan konsep *machine learning* sering kali digunakan dalam pengembangan model prediktif karena *machine learning* mempunyai keunggulan lebih dibandingkan dengan teknik statistik konvensional.

5. *Predictive Modeling*

Tahap dimana model prediktif mulai dikembangkan berdasarkan teknik statistik dan *machine learning* dengan *requirement* dan dataset yang sesuai. Jika proses pengembangan model sudah selesai, maka model akan dilakukan pengujian dengan memeriksa validitas model dengan melihat seberapa besar tingkat *error* dan akurasi yang dihasilkan oleh model tersebut.

6. *Prediction and Monitoring*

Tahap terakhir adalah melakukan prediksi dan pengambilan keputusan berdasarkan hasil model. Model akan terus dipantau secara konsisten untuk memastikan apakah model masih memberikan hasil yang benar dan membuat prediksi yang akurat.

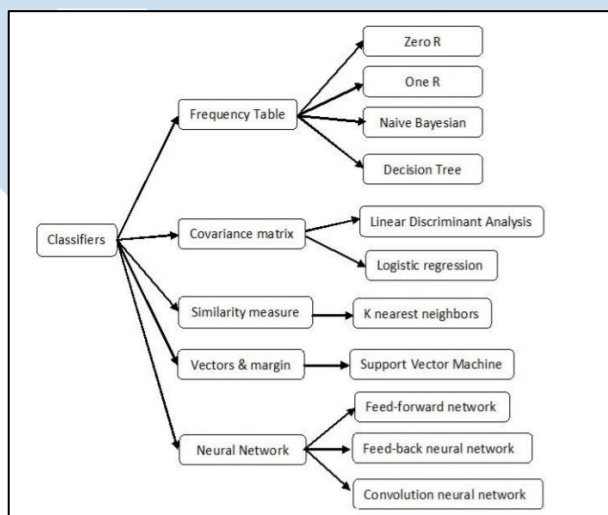
2.9.3 Predictive Analytics Models

Predictive analytics models didefinisikan sebagai model yang melakukan prediksi dan akan belajar berdasarkan pengalaman untuk membuat prediksi di masa depan. Keakuratan hasil dengan model yang dikembangkan tergantung dengan tingkat analisis data yang dilakukan. Pendekatan dan teknik yang digunakan untuk melakukan *predictive analytics* dapat dipisahkan menjadi dua teknik yang berbeda yaitu teknik regresi dan teknik *machine learning* [26]. Berikut merupakan *predictive analytics model* yang paling populer digunakan, yaitu *Regression Techniques*, *Linear Regression Model*, *Logistic Regression*, *Classification Model*, *Clustering Model*, *Forecast Model*, *Time Series Model*.

2.10 Classification Modelling

Classification modeling merupakan model yang dibangun berdasarkan pendekatan *data mining* yang digunakan untuk menghasilkan prediksi dari suatu

kelompok tertentu dan seringkali dibangun dengan teknik *machine learning* [27]. *Supervised machine learning techniques* biasa digunakan untuk *classification modelling* dan *regression*, karena *supervised techniques* mencoba untuk menemukan hubungan dengan *input attributes (independent variables)* dan *target attribute (dependent variable)* [27]. Pada *supervised techniques* seluruh dataset dibagi menjadi dua bagian yaitu data untuk *training* dimana *classifier* akan belajar dari data tersebut dan data yang tersisa akan digunakan untuk *testing* akurasi model yang dibangun [22]. *Supervised learning classifiers* dibagi lagi menjadi lima *group classification algorithms* berdasarkan *frequency table*, *covariance matrix*, *similarity measure*, *vectors & margin* dan *neural network*. Gambar 2.3 merupakan algoritma yang dapat digunakan berdasarkan masing-masing *group* tersebut.



Gambar 2. 3. *Classification Algorithms in Supervised Machine Learning*

2.11 Label Encoding

Label Encoding merupakan salah satu teknik pada *categorical encoding* yang digunakan untuk menangani variabel *categorical* khususnya *ordinal* menjadi *numerical value* [28]. Teknik ini dapat diimplementasikan dengan *library sklearn* yaitu *LabelEncoder*. Penggunaan *label encoding* akan lebih efektif saat value dari suatu *attribute* mempunyai prioritas yang berbeda [29], contohnya pada penelitian ini terdapat *attributes 'number of meetings'* dengan urutan *value* yaitu *1 meetings*, *2-3 meetings*, dan *>3 meetings*. *Value* tersebut dapat diurutkan berdasarkan tingkatan dari jumlah *meetings* paling sedikit hingga paling banyak, maka dari itu

attribute number of meetings lebih cocok diimplementasikan Label Encoding untuk merubah tipe data *categorical* menjadi *numeric*.

2.12 One Hot Encoding

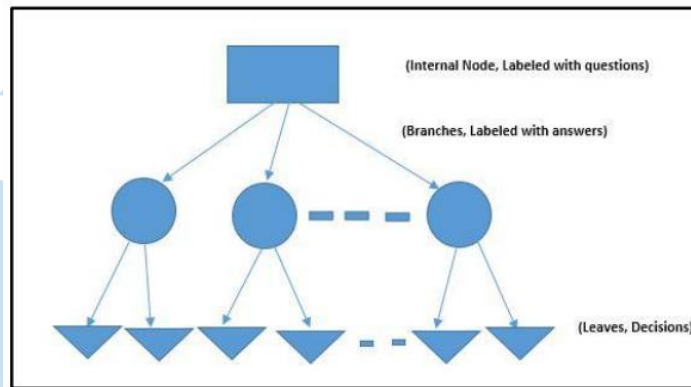
One Hot Encoding juga termasuk dalam salah satu teknik *categorical encoding* yang menangani variabel *categorical* dengan cara membuat *attributes* baru berdasarkan jumlah *unique values* dalam *attributes* tersebut [28]. Proses ini merupakan proses untuk membuat *attributes dummy* yang direpresentasikan oleh *binary vector* [28]. Teknik ini dapat diimplementasikan dengan *library sklearn* yaitu *OneHotEncoder*. Penggunaan *one hot encoding* lebih cocok diterapkan pada *categorical* yang tidak mempunyai prioritas atau tingkatan tertentu [29], contohnya pada penelitian ini terdapat *attributes* '*division*' dengan *value corporate controller, finance and accounting, merchandising and distribution* dan *system and IT*. *Value* dari *attributes* tersebut tidak dapat dilihat tingkatan prioritas karena semua divisi dianggap sama, maka dari itu *one hot encoding* lebih cocok diimplementasikan pada *attribute* ini dan tiap *values* dari *attributes* tersebut akan dirubah kedalam empat kolom baru sesuai jumlah *value* nya dengan bentuk *binary vector*.

2.13 Algoritma Klasifikasi yang Digunakan

2.13.1 Decision Tree

Decision tree merupakan salah satu algoritma klasifikasi yang membangun model dalam bentuk *hierarchical structure (tree)*. *Decision tree* dikembangkan melalui proses inkremental langkah demi langkah dengan memecah dataset menjadi lebih kecil [30]. Pada proses akhir, *decision tree* akan menghasilkan *tree* (pohon) dengan tiga tipe pemisahan berupa cabang (*nodes*) yaitu *root node/internal node*, *decision nodes/branches* dan *leaf nodes* (simpul daun). *Internal nodes* biasanya akan dilabeli dengan pertanyaan yang akan dicari, *Decision nodes* biasanya akan memiliki dua atau lebih cabang dimana setiap cabang mewakili pilihan dari beberapa alternatif, sedangkan *leaf nodes* mewakili klasifikasi yang sudah diberikan label dengan keputusan berdasarkan masalah atau dataset yang ada [25]. Kriteria pemisahan pohon yang biasa digunakan adalah *information gain*, *gini index*, dan *gain ratio*. *Decision tree* dapat digunakan

untuk data kategorikal dan numerik. Berikut merupakan gambaran dari *decision tree structure*.



Gambar 2. 4 *Decision Tree Structure*

Decision tree mempunyai beberapa kelebihan dan kekurangan dalam penggunaannya. Kelebihan *decision tree algorithm* diantaranya adalah mudah untuk menafsirkan *decision rules*, mudah untuk menggabungkan beberapa lapisan data numerik dan kategorikal sehingga tidak perlu memilih data lagi dan kuat dalam hal *outlier* pada *training data* [30]. Kekurangan *decision tree algorithm* diantaranya adalah *decision tree* cenderung bekerja lebih baik jika terdapat beberapa atribut data yang sangat relevan, tetapi akan memberikan hasil cukup buruk saat digunakan pada dataset yang cukup kompleks karena *decision tree* mempunyai limit untuk merespon variabel, sehingga tidak mungkin untuk memprediksi diluar batas minimum dan maksimum variabel tersebut dalam *training data* [30].

2.13.2 Random Forest

Random forest merupakan salah satu model *machine learning* yaitu algoritma klasifikasi yang dikembangkan berdasarkan konstruksi beberapa pengembangan *decision trees* dengan *training* dan *testing* data untuk menghasilkan prediksi rata-rata dari semua *decision trees* yang terlibat [31]. *Random forest* menggunakan berbagai metode *exploratory data analysis* (EDA) dan menghasilkan skor akurasi yang baik. Kemampuan untuk memproses kumpulan data yang besar dengan spasial yang lebih tinggi

merupakan salah satu manfaat dari *random forest*, yaitu memproses banyak *variable input* dan mengidentifikasi semua *variable* yang signifikan serta menyortir pentingnya *variable* tersebut yang selanjutnya akan dianggap sebagai fitur [31].

Random forest juga mempunyai kelebihan dan kekurangan dalam penggunaannya. Kelebihan *random forest algorithm* adalah memberikan prediksi yang akurat tanpa limit *variable*, dapat mengukur pentingnya setiap fitur pada *training data set*, dan kedekatan antara sampel juga dapat diukur dengan *training data set* [30]. Kekurangan *random forest algorithm* adalah untuk kategorikal *variable* dengan jumlah yang berbeda, *random forest* bias dalam mendukung *variable* tersebut dalam level yang lebih banyak [30].

2.13.3 Naïve Bayes

Naïve Bayes *classifier* merupakan salah satu model klasifikasi yang handal, cepat, akurat dan berfokus pada probabilitas bersyarat. Model ini dikembangkan berdasarkan *Bayesian Theorem* dan cocok untuk input berdimensi tinggi karena teori ini mengasumsikan fitur dan atribut independen [32]. Model *naïve bayesian* mudah dibuat tanpa estimasi parameter yang rumit sehingga membuatnya sangat berguna untuk sekumpulan data yang sangat besar [22].

Naïve Bayes mempunyai beberapa kelebihan dan kekurangan dalam penggunaannya. Kelebihan dari model ini yaitu hanya membutuhkan data *training* lebih sedikit untuk memperkirakan parameter klasifikasi yang diperlukan [32]. Kekurangan dari model ini adalah perlu mengetahui *prior probability* yang didasarkan pada asumsi, dan prosesnya membutuhkan banyak pilihan subjektif di antara model asumsi, tetapi untuk mengatasi masalah atau kekurangan tersebut dapat dilakukan dengan teknik *smoothing* [32].

2.14 Confusion Matrix

Confusion Matrix merupakan *summary* hasil prediksi untuk melihat akurasi model pada masalah klasifikasi [33]. *Matrix* ini biasanya digunakan untuk melakukan evaluasi terhadap model klasifikasi yang sudah dikembangkan. Jumlah

prediksi yang benar atau salah nantinya akan dirangkum dengan nilai hitungan dan dipecah oleh masing-masing kelas. Output yang dihasilkan biasanya berbentuk tabel atau *plot confusion matrix* seperti pada Gambar 2.5.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 2. 5 *Confusion Matrix* [34]

- TN (*True Negative*): merupakan jumlah prediksi negative yang berhasil diprediksi dan benar negatif (aktual negatif, prediksi negatif).
- TP (*True Positive*): merupakan jumlah prediksi positive yang berhasil diprediksi dan benar positive (aktual positive, prediksi positif).
- FP (*False Positive*): merupakan positif palsu atau jumlah contoh negatif aktual yang diklasifikasikan sebagai positif (aktual positif, prediksi negatif).
- FN (*False Negative*): merupakan negatif palsu atau jumlah contoh positif aktual yang diklasifikasikan sebagai negatif (aktual negatif, prediksi positif).

Salah satu *matrix* yang paling umum digunakan saat melakukan pengembangan model klasifikasi adalah pengukuran akurasi model. Keakuratan model dapat dihitung berdasarkan Gambar 2.6.

$$Accuracy = \frac{TN + TP}{TN + FP + FN + TP}$$

Gambar 2. 6 *Accuracy Formula* [33]

2.15 Python

Python merupakan *high level programming language* yang bersifat *open source* dan cukup mudah untuk dipelajari dibandingkan dengan bahasa lain seperti C, C#, Java Script, Java dll [35]. Python sudah banyak digunakan secara luas untuk melakukan prediksi, *scripting* dan pengolahan *machine learning* lainnya dengan

beragam library yang dapat dijalankan di berbagai *platform*. Python juga memberikan lebih banyak fleksibilitas pada user untuk mengimplementasikan berbagai algoritma *machine learning*, hingga lebih dari seribu variasi algoritma [36].

2.16 Penelitian Terdahulu

Tabel 2.2 menunjukkan beberapa penelitian dahulu yang berkaitan dengan prediksi kelanjutan adopsi Work From Home pada perusahaan dngan teknik atau algoritma tertentu yang dapat digunakan untuk menunjang penelitian ini.

Tabel 2.2 Penelitian Terdahulu

No	Jurnal	Penulis	Metode / Algoritma	Hasil dan Simpulan
1.	A Comparative Study of Work From Home vs Work From Office: Preference of Women Employees in IT Industry, Design Engineering, vol. 1, no. 7, pp. 5763–5775, 2021. [37]	Feleen Christy J, Vibha, Narmada Kalgi, Arokiaraj David, M. Maheswari & M.Christopher	Pengumpulan data: kuesioner 200 responden wanita di sektor IT. Metode: Uji hipotesa berdasarkan hasil Chi-Squared test	Berdasarkan uji hipotesa yang telah dilakukan, hasil menunjukkan bahwa terdapat satu hipotesa yang ditolak yaitu “Usia karyawan wanita tidak mempengaruhi pilihan cara kerja (WFH/WFO)”, tetapi secara kesimpulan dari penelitian tersebut adalah responden lebih menyukai model WFH jika mereka mempunyai lingkungan kerja yang bagus di rumah dan mendapatkan fleksibilitas kerja serta motivasi yang positif dari <i>top-level management</i> .
2.	Employee Productivity Modelling on a Work From Home Scenario During the Covid-19 Pandemic: A Case Study Using Classification	Dilhari Attygalle, Geethanadee Abhayawardana	Pengumpulan data: kuesioner 60 karyawan di Sri Lanka Algoritma: Decision Tree & analisa deskriptif berupa visualisasi	<i>Classification model</i> dengan <i>decision tree</i> merupakan model terbaik dengan tingkat akurasi 88% untuk melakukan klasifikasi pada produktivitas karyawan selama WFH. Klasifikasi tersebut menghasilkan 2 kelompok yang berbeda

	Trees, <i>J. Bus. Manag. Sci.</i> , vol. 9, no. 3, pp. 92–100, 2021 [4]			yaitu 2 <i>department</i> tergolong pada “ <i>positive productivity</i> ” sedangkan 3 lainnya merupakan “ <i>non-positive productivity</i> ”
3.	A Comparative Machine Learning Study on IT Sector Edge Nearer to Working From Home (WFH) Contract Category for Improving Productivity,” <i>J. Artif. Intell. Capsul. Networks</i> , vol. 2, no. 4, pp. 217–225, 2021 [5]	Dr. Akey Sungheetha, Dr. Rajesh Sharma R	Pengumpulan data: survey kuesioner 3 jenis masyarakat pada 3 bidang berbeda Algoritma: random forest, decision tree, naïve bayes	Model <i>Decision Tree</i> mampu memprediksi peningkatan produktivitas selama WFH paling baik dengan akurasi sebesar 74%. Analisa prediksi peningkatan produktivitas selama WFH di berbagai sektor • IT = 70% • Manufaktur = 25-30% • Pertanian = 10% Dari hasil tersebut, disimpulkan bahwa praktek kerja WFH terbukti efektif hanya untuk sektor IT, sedangkan sektor lain kurang efektif untuk dilakukan sehingga WFO Optional dapat diterapkan untuk sektor lain.
4.	Continuance Adoption of Working from Home after the COVID-19 Outbreak: Empirical Evidence from Saudi Arabia,” <i>J. Asian Financ. Econ. Bus.</i> , vol. 8, no. 7, pp. 67–78, 2021 [6]	Salem Mohamed Ahmed, Nor Khalil MD	Teori: Continuous Adoption Intention untuk melihat apakah adopsi WFH dapat diteruskan pasca COVID-19 Metode: uji hipotesa	Hasil yang didapatkan dari penelitian tersebut adalah sikap positif karyawan terhadap kelanjutan WFH diukur dalam hal otonomi pekerjaan dan keseimbangan kehidupan kerja. Penelitian ini juga menemukan bahwa terdapat hubungan yang cukup terlihat antara pengendalian dan niat pada adopsi WFH pasca COVID-19.

6.	Implementation of Data Mining Technique for Performance of WFH and WFO Agents Using the K-Means Method Case Study Study of PT Infomedia Telkom Consumer Profiling Services,” <i>Budapest Int. Res. Exact Sci. J.</i> , vol. 3, no. 2, pp. 117–125, Apr. 2021 [38]	Jaja, Nandi Priatna, Tazkia Salsabila Ardan	Pengumpulan data: database operasional Call Centre Outbound PT Infomedia Telkom Algoritma: K-Means	Didapatkan 3 pola clustering • <i>Cluster_0 = Low</i> • <i>Cluster_1 = High</i> • <i>Cluster_2 = Medium</i> Hasil <i>performance</i> dari ke 3 <i>clustering</i> tersebut selanjutnya akan dibuatkan saran untuk dilakukan beberapa program seperti Cluster_0 (Low) untuk Coaching, Cluster_1 (High) = Rewarding, Cluster_2 (Medium) = Refreshment"
----	---	---	---	---

Berdasarkan Tabel 2.2, terdapat enam penelitian terdahulu yang dapat disimpulkan bahwa preferensi karyawan untuk terus menerapkan WFH membawa dampak yang cukup positif dan efektif untuk diterapkan hanya di beberapa bidang tertentu, selain itu didapatkan bahwa model dengan algoritma *decision tree* terbukti memberikan hasil yang baik dalam melakukan klasifikasi dan prediksi kelanjutan WFH. Penelitian ini akan berfokus untuk memprediksi kelanjutan WFH berdasarkan faktor-faktor tertentu dengan perbandingan algoritma *decision tree*, *random forest* dan *naïve bayes* berdasarkan penelitian [5] dan membuktikan bahwa model algoritma *decision tree* mempunyai performa terbaik berdasarkan penelitian [4][5].

Berikut beberapa perbandingan penelitian terdahulu dengan penelitian ini:

1. Berfokus pada perbandingan penggunaan algoritma klasifikasi yaitu *decision tree*, *random forest* dan *naïve bayes*, serta penggunaan teknik *data mining* dengan *framework* CRISP-DM dimana penelitian pada Tabel 2.2 tidak menggunakan *framework* CRISP-DM dan lebih banyak melakukan penelitian dengan satu algoritma atau uji hipotesa.
2. Menggunakan lima faktor utama yang dijadikan sebagai faktor pertimbangan untuk prediksi kelanjutan WFH yaitu *work improvement*, *employee productivity*, *employee performance*, *work environment* dan

communication & support, dimana penelitian pada Tabel 2.2 hanya berfokus pada salah satu faktor saja seperti *productivity* atau *performance*.

3. Sumber data yang digunakan diambil langsung berdasarkan hasil survey kuesioner karyawan GORP Kompas Gramedia Palmerah, dimana pertanyaan dibuat berdasarkan penelitian [4][6][7].

