

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pandemi Covid-19 yang pertama kali muncul di Wuhan, Tiongkok pada Desember 2019 telah menyebar ke seluruh dunia dan menjadi pandemi global. Virus ini menyebar dengan sangat cepat dan telah menginfeksi jutaan orang di seluruh dunia, serta menyebabkan kematian bagi ribuan orang [1]. Covid-19 memiliki dampak yang luas pada sektor kesehatan, sosial, ekonomi, dan lingkungan hidup [2]. Banyak negara yang harus menerapkan *lock down* dan pembatasan sosial untuk meminimalkan penyebaran virus dan mengendalikan pandemi ini. Namun, dampak pandemi ini tidak hanya terasa pada negara yang terkena dampak secara langsung, namun juga negara-negara lain yang bergantung pada perdagangan internasional dan pariwisata. Pandemi ini telah menyebabkan kerugian ekonomi yang besar dan meningkatkan angka pengangguran di seluruh dunia [3]. Oleh karena itu, diperlukan upaya global untuk mengatasi pandemi Covid-19 dan meminimalkan dampaknya pada seluruh aspek kehidupan manusia [4].

Pandemi Covid-19 tersebut telah memaksa banyak orang untuk beradaptasi dengan situasi baru yang membuat berbagai aktivitas dilakukan secara daring seperti konsep bekerja dari rumah atau *Work From Home* (WFH) untuk membatasi mobilitas masyarakat dan mengurangi penyebaran Covid-19 di lingkungan kerja [5]. Dampak yang terjadi atas diberlakukannya hal tersebut adalah meningkatnya pemanfaatan teknologi untuk menggantikan aktivitas tatap muka seperti belajar, bekerja, dan bahkan berbelanja. Seiring dengan meningkatnya penggunaan teknologi daring, banyak perusahaan dan individu yang menemukan bahwa bekerja dari rumah memberikan pengalaman yang positif bagi karyawan. Studi pada tahun 2020 membuktikan bahwa WFH memberikan keuntungan seperti fleksibilitas dan keseimbangan kerja-hidup

yang lebih baik, sehingga pandemi telah membuka peluang baru dalam memanfaatkan teknologi daring dalam bekerja [6]. Oleh karena itu, konsep WFH dapat dipertimbangkan oleh perusahaan untuk tetap diaplikasikan setelah pandemi.

Konsep WFH juga diterapkan di negara Indonesia sebagai salah satu negara yang juga terdampak pandemi Covid-19. Saat ini, Indonesia mulai mengalami perubahan dari status pandemi menjadi endemi Covid-19 sejak kebijakan terkait PPKM (Pemberlakuan Pembatasan Kegiatan Masyarakat) dicabut oleh Presiden Joko Widodo pada 30 Desember 2022. Endemi menjadikan Covid-19 sebagai suatu keadaan yang biasa timbul dalam suatu populasi seperti penyakit malaria dan demam berdarah [7]. Hal itu dikarenakan atas melandainya kasus positif Covid-19 yang dialami masyarakat Indonesia. Perubahan status tersebut berpengaruh terhadap kebijakan pemerintah terkait peraturan pembatasan kegiatan masyarakat. Pembatasan kegiatan dalam dunia kerja mengalami perubahan ke arah yang lebih bebas membuat karyawan dapat lebih fleksibel untuk melakukan pekerjaan di dalam kantor (*Work From Office*) ataupun di rumah (*Work From Home*) yang juga dipengaruhi oleh kebijakan perusahaan. PT XYZ menjadi salah satu perusahaan yang perlu menetapkan kebijakan bekerja secara WFO atau WFH tersebut.

PT XYZ merupakan perusahaan penyedia layanan telekomunikasi yang sudah ada selama lebih dari 20 tahun. Dengan memiliki lebih dari 5000 karyawan per tahun 2022, tentu perusahaan harus memberikan berbagai kebijakan baru untuk menyesuaikan keberlangsungan proses kerja karyawan di kondisi endemi tersebut. Salah satu program yang dimiliki PT XYZ adalah *training* untuk setiap karyawannya. Investasi pada pelatihan dan pengembangan karyawan dapat meningkatkan produktivitas, kinerja, serta memperkuat budaya perusahaan [8]. *Training* karyawan tersebut menjadi program dengan kebijakan tata cara pelaksanaan yang berubah dalam menyesuaikan kondisi Indonesia menghadapi Covid-19. *Training* karyawan dapat dilakukan dengan dua cara yaitu secara *offline* melalui berbagai kelas

pelatihan di suatu tempat atau secara *online* melalui *platform* pembelajaran *online* sebagai wujud digitalisasi.

Berdasarkan studi yang dilakukan oleh McKinsey, salah satu perusahaan konsultan terkemuka menyatakan bahwa digitalisasi akan berpengaruh terhadap perekonomian negara termasuk berbagai sektor industri dan perusahaan yang ada di negara tersebut [9]. Penerapan teknologi juga dilakukan oleh PT XYZ terutama dalam mengatur prosedur *training* karyawan. Sebelum pandemi, kegiatan *training* karyawan di PT XYZ dilakukan dengan membuat beberapa kelas untuk menghadirkan pengajar (*trainer*) untuk membahas topik tertentu namun hal tersebut tidak dapat dilakukan seiring status Indonesia masih pandemi sehingga harus dilakukan secara daring dari rumah. Oleh karena itu PT XYZ melakukan kontrak dengan beberapa penyedia layanan *online learning* untuk kegiatan *training* seperti Udemy, Percipio, dan juga LinkedIn Learning. Saat ini, kegiatan *online training* pada PT XYZ dilakukan dengan LinkedIn Learning dan terbagi menjadi 2 jenis yaitu *assigned learning* dengan materi yang ditentukan oleh perusahaan serta *self-learning* dengan materi pilihan sesuai preferensi karyawan masing-masing.

Dalam kondisi Indonesia terkini yang sudah membaik dari status pandemi menjadi endemi [10], PT XYZ masih menerapkan kegiatan *training* secara *online* melalui akses terhadap layanan *online learning*. Hal tersebut dilakukan karena dirasa lebih efektif dan efisien dalam meningkatkan produktivitas karyawan dibandingkan dengan membuat beberapa kelas secara *offline*. Namun, terdapat permasalahan atas kebijakan tersebut yaitu kegiatan *training* karyawan belum dapat diketahui atau dianalisis saat ini berbeda dengan kegiatan *training offline* yang dapat melakukan pendataan terkait karyawan yang menjadi peserta di tiap kelasnya. Hal tersebut bermanfaat untuk memberikan *assigned learning* yang tepat dan sesuai dengan preferensi karyawan seperti dalam hal topik pembelajaran karyawan. Pemberian *assigned learning* haruslah tepat yang bertujuan untuk memberikan kumpulan *course training* yang selain dapat melatih kompetensi yang diperlukan

perusahaan namun juga sesuai dengan preferensi masing-masing karyawan. Preferensi tersebut dapat diketahui yang diperoleh terkait topik pembelajaran saat karyawan melakukan *self-learning*. Untuk menjawab permasalahan tersebut maka perlu untuk dilakukan pembuatan visualisasi data *training* karyawan yang dapat memetakan aktivitas karyawan dalam melakukan *online training* baik *self-learning* maupun *assigned learning*. Selain itu perlu untuk dibuat klasifikasi berdasarkan *self-learning* untuk memudahkan dalam pemberian *assigned learning* yang sesuai dengan preferensi karyawan (*personalized learning*).

Klasifikasi data sebagai salah satu luaran dari topik penelitian ini telah menjadi bagian yang penting dari bidang kecerdasan buatan dan telah berkembang pesat dalam beberapa tahun terakhir. Klasifikasi dapat dilakukan untuk berbagai tujuan dengan tipe data yang berbeda seperti gambar dan teks. Contoh klasifikasi dengan gambar dapat dilakukan untuk memetakan dan inventarisasi suatu lahan basah [11]. Contoh klasifikasi dengan teks sebagai data juga dapat dilakukan seperti dalam dunia perbankan untuk memprediksi tingkat kecurangan dalam penggunaan kartu kredit [12]. Klasifikasi dengan beragam tipe data tersebut dapat dilakukan dengan *machine learning* beserta dengan beragam algoritma di dalamnya yang dapat dibandingkan kualitasnya dalam menemukan algoritma yang terbaik untuk model *machine learning* dalam permasalahan klasifikasi data [13]. Beragam algoritma dalam *machine learning* untuk klasifikasi seperti *decision tree* dan *random forest* sering digunakan karena menghasilkan akurasi yang sangat baik [12]-[14]. Salah satu penelitian berhasil menemukan bahwa *random forest* memiliki akurasi yang terbaik dalam penelitian tingkat fatalnya kecelakaan di jalan dibandingkan dengan algoritma *logistic regression* dan *decision tree* serta analisis prediktor-prediktor gangguan perilaku tidur gerak mata cepat pada pasien parkinson dibandingkan dengan algoritma *decision tree* [13], [15]. Pada penelitian lain, *decision tree* memiliki akurasi yang baik bahkan ketika dibandingkan dengan *random forest* yakni dalam kasus analisis sentimen di YouTube dan prediksi

efisiensi operasional bank [14], [16]. Klasifikasi data dan pembangunan model *machine learning* sebagai bagian dari proses data *mining* dapat diaplikasikan menggunakan metode yang paling populer dalam proyek data *mining* yaitu CRISP-DM [17]. CRISP-DM telah ditetapkan secara *de-facto* dalam berbagai proyek atau penelitian terkait data *mining* selama lebih dari 2 dekade [18]. CRISP-DM telah menyelesaikan kasus seperti klasifikasi hingga dilakukan visualisasi data [19]-[21]. CRISP-DM memiliki enam tahapan lengkap mulai dari mengenali masalah hingga penerapan solusi yang telah dibangun melalui pembentukan model dan evaluasinya [19].

Oleh karena itu, algoritma *random forest* dan *decision tree* dengan menggunakan metode CRISP-DM merupakan langkah yang dilakukan dalam menjawab salah satu permasalahan penelitian ini yaitu untuk membuat model klasifikasi dalam pemberian *assigned learning* karyawan yang bersifat *personalized learning* berdasarkan topik pembelajaran. Adapun objek penelitian terkait kegiatan *training* karyawan serta dua buah hasil yang ingin dicapai yaitu visualisasi data dan model *machine learning* menjadi kebaruan penelitian dari seluruh penelitian terdahulu dalam penelitian ini. Visualisasi data dan model *machine learning* yang dibentuk merupakan bagian dari proyek data *mining* dalam penelitian ini. Metode CRISP-DM diaplikasikan dalam menentukan tahapan yang dilakukan agar dapat terciptanya visualisasi data dan model *machine learning* yang dapat menjawab permasalahan perusahaan. Pemilihan metode CRISP-DM dilakukan atas keberadaannya yang populer dan telah menjadi standar secara *de-facto* dalam proyek data *mining* [17], [18].

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dibahas tersebut, maka rumusan masalah penelitian ini adalah sebagai berikut:

- 1) Bagaimana hasil visualisasi data aktivitas *online training* karyawan PT XYZ?

- 2) Bagaimana hasil model *machine learning* dengan algoritma *decision tree* untuk *personalized learning*?
- 3) Bagaimana hasil model *machine learning* dengan algoritma *random forest* untuk *personalized learning*?
- 4) Bagaimana hasil perbandingan akurasi, *precision*, *recall*, dan *F1 score* model *machine learning* yang terbaik antara dengan algoritma *decision tree* dan *random forest* untuk *personalized learning*?

### 1.3 Batasan Masalah

Terdapat batasan dalam penelitian ini yaitu sebagai berikut:

- 1) Data yang digunakan dalam penelitian ini untuk pembuatan *dashboard* visualisasi data *training* karyawan berasal dari penyedia *platform training* karyawan yang kemudian diintegrasikan ke dalam data *warehouse* perusahaan. Namun, karena struktur data *warehouse* masih dalam tahap migrasi, PT XYZ masih melakukan penarikan data secara manual dari penyedia *platform* tersebut.
- 2) Penelitian ini menghasilkan model *machine learning* yang dibuat hanya menggunakan data historis dari periode Juli 2022 hingga Februari 2023 yang sesuai dengan periode awal penggunaan *platform* pembelajaran terbaru yaitu LinkedIn Learning oleh PT XYZ.
- 3) Penelitian ini menggunakan metode data *mining* CRISP-DM dengan tahap *deployment* dilakukan hanya untuk *dashboard* visualisasi aktivitas *training* karyawan, sementara pembangunan model *machine learning* menjadi sumber daya perusahaan terlebih dahulu sebelum digunakan.

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

## 1.4 Tujuan dan Manfaat Penelitian

### 1.4.1 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dibahas tersebut, maka tujuan penelitian ini adalah:

- 1) Membangun aplikasi *business intelligence* di PT XYZ yang memetakan aktivitas karyawan dalam melakukan *online training* baik *self-learning* maupun *assigned learning*.
- 2) Membangun model *machine learning* dengan algoritma *decision tree* untuk *personalized learning*.
- 3) Membangun model *machine learning* dengan algoritma *random forest* untuk *personalized learning*.
- 4) Membandingkan hasil akurasi, *precision*, *recall*, *F1 score* model *machine learning* terbaik antara dengan algoritma *decision tree* dan *random forest* untuk *personalized learning*.

### 1.4.2 Manfaat Penelitian

Manfaat penelitian yang dihasilkan dalam penelitian ini adalah sebagai berikut:

- 1) Memberikan hasil berupa visualisasi data kepada PT XYZ yang menjelaskan bahwa data aktivitas *training* karyawan secara daring melalui *platform* pembelajaran *online* dapat diproses menjadi informasi baru sebagai fakta pendukung untuk keperluan pengambilan keputusan dalam perusahaan.
- 2) Memberikan kemudahan bagi perusahaan dalam merekomendasikan pelatihan yang cocok bagi karyawan dengan memanfaatkan model *machine learning* yang dapat mempertimbangkan preferensi karyawan dalam hal topik pembelajaran.

## 1.5 Sistematika Penulisan

- BAB I**                    **PENDAHULUAN**  
Bab ini menjelaskan latar belakang dilakukannya penelitian ini, rumusan masalah yang akan diselesaikan, batasan masalah dalam penelitian, tujuan dan manfaat yang ingin dicapai, dan sistematika penulisan laporan.
- BAB II**                    **LANDASAN TEORI**  
Bab ini menjelaskan landasan teori yang digunakan dalam penelitian ini. Terbagi menjadi beberapa subbab yaitu tinjauan teori topik, pengertian algoritma yang digunakan, pengertian *tools* yang digunakan, dan penelitian terdahulu.
- BAB III**                    **METODOLOGI PENELITIAN**  
Bab ini menjelaskan terkait metode penelitian yang digunakan sekaligus memaparkan data yang telah diperoleh berikut dengan teknik analisis data tersebut.
- BAB IV**                    **ANALISIS DAN HASIL PENELITIAN**  
Bab ini menjelaskan tentang hasil pembuatan model *machine learning* dan *dashboard* visualisasi data sebagai hasil proyek data *mining* yang diaplikasikan dengan metode yang digunakan dalam penelitian ini.
- BAB V**                    **SIMPULAN DAN SARAN**  
Bab ini berisi kesimpulan pada keempat bab sebelumnya sekaligus saran untuk penelitian selanjutnya.

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