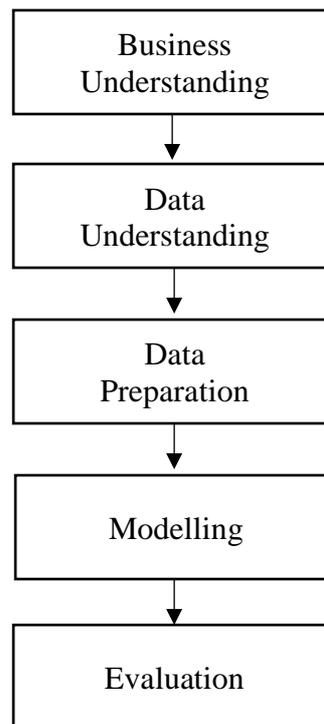


## BAB III METODOLOGI PENELITIAN

### 3.1 Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini metode Cross Industry Standard for Data Mining (CRISP-DM). CRISP-DM merupakan metode yang menggunakan model proses pengembangan data yang banyak digunakan para ahli untuk memecahkan masalah [34]. Adapun tahapan CRISP-DM, yaitu *Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation*). Dalam penelitian ini hanya memerlukan tahapan sebagai berikut :



*Gambar 3.1 Alur dalam tahapan Penelitian*

Tahapan pada gambar 3.1 . Adalah proses data mining pada penelitian ini :

1. Tahap Pemahaman bisnis juga dapat disebut fase pemahaman penelitian yang meliputi tujuan dan persyaratan proyek dengan jelas dalam hal bisnis atau unit penelitian secara keseluruhan, menterjemahkan tujuan dan batasan ke dalam perumusan definisi masalah data mining, menyiapkan strategi awal untuk mencapai tujuan tersebut.
2. Tahap pemahaman data dengan melakukan pengumpulan data menggunakan analisis data eksplorasi untuk membiasakan diri dengan data dan menemukan wawasan awal,

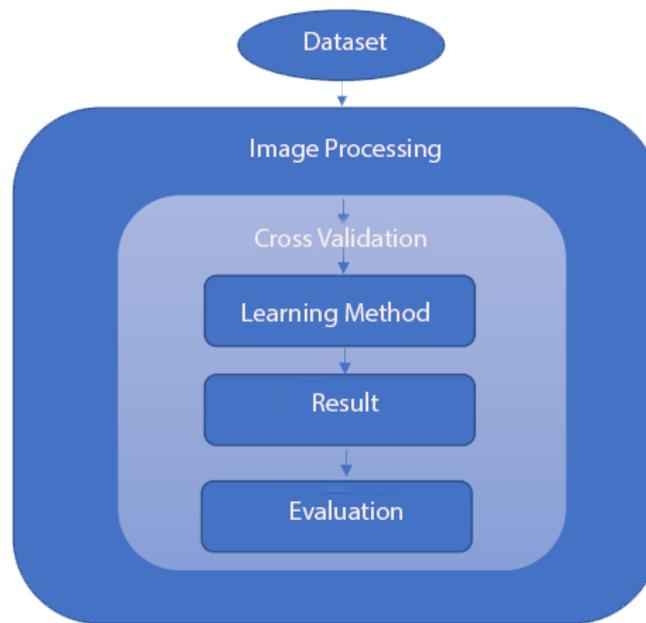
mengevaluasi kualitas data, apabila diinginkan pilih subset menarik yang mungkin berisi pola yang dapat ditindaklanjuti.

3. Tahap persiapan data adalah mempersiapkan data mentah awal kumpulan data akhir yang akan digunakan untuk semua tahapan berikutnya. Pada tahapan ini memilih kasus dan variabel yang diinginkan, menganalisis yang sesuai untuk analisis. Melakukan transformasi pada variabel tertentu, jika diperlukan, membersihkan data mentah sehingga siap untuk alat pemodelan. Data penelitian ini terdiri dari 4200 gambar rotgen paru-paru manusia, 6 Atribut dan target class adalah katagori dengan 2 label yaitu normal/tuberculosis. Table 3.1 menjelaskan atribut-atribut yang digunakan pada penelitian ini :

Tabel 3.1 Deskripsi Atribut Data

<b>ATRIBUT</b>	<b>DESKRIPSI</b>	<b>NILAI</b>
Katagori	Target Class	Normal/Tuberculosis
Nama File	Merupakan nama yang membedakan dari masing-masing data yang diawali dengan label katagori dan lanjutkan dengan angka	(Katagori-(numerik))
Lokasi File	Merupakan letak dimana data tersebut di simpan pada direktori penyimpanan	(Katagori\Nama File)
Ukuran	Bilangan yang menunjukkan satuan ukuran dari data	Kilobyte
Lebar	Skala ukuran data dari kanan ke kiri atau sebaliknya	Centimeter
Tinggi	Skala ukuran data dari atas ke bawah atau sebaliknya	Centimeter

4. Tahap pemodelan, Tahapan ini dilakukan dengan memilih dan menerapkan teknik pemodelan yang sesuai. Kalibrasi pengaturan model untuk mengoptimalkan hasil, seringkali, beberapa teknik berbeda dapat digunakan untuk masalah data mining yang sama, apabila diperlukan mengulangi kembali ke tahapan persiapan data untuk membawa bentuk data sesuai dengan persyaratan khusus dari teknik data mining tertentu. Tahap pemodelan pada penelitian menggunakan 2 metode Logistic Regression (LR) dan Support Vector Machine (SVM). Dibawah ini adalah alur pemodelan :



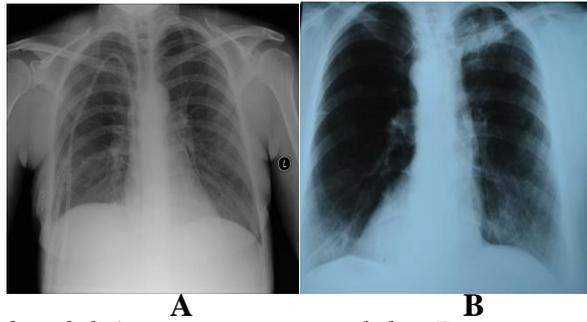
*Gambar 3.2 Bagan Pemodelan*

pada gambar 3.2 adalah rencana pemodelan yang akan dilakukan dimulai dari mengubah dataset gambar rotgen menjadi data numerik yang memiliki beberapa atribut. Selanjutnya data akan di klasifikasikan menggunakan 2 metode yaitu Logistic Regression dan Support Vector Machine.

5. Tahap evaluasi adalah mengevaluasi satu atau lebih model yang disampaikan dalam tahap pemodelan untuk kualitas dan efektivitas sebelum menerapkannya untuk digunakan di lapangan.. Pada tahap ini yang akan dilakukan adalah menentukan model sebenarnya mencapai tujuan yang ditetapkan untuk itu pada tahapan pertama, menetapkan apakah beberapa aspek penting dari masalah bisnis atau penelitian belum diperhitungkan secara memadai., mengambil keputusan tentang penggunaan hasil data mining.

### **3.2 Sampel citra rontgen paru-paru TBC dan paru-paru normal**

Gambar rontgen paru-paru yang digunakan pada penelitian ini bersifat public atau open source. Terbagi dalam 2 folder yaitu gambar rontgen paru-paru normal sebanyak 3500 data dan gambar rontgen paru-paru positif tuberculosis sebanyak 700 data.



**A** **B**  
*Gambar 3.3 A paru-paru normal dan B paru-paru TBC*

Pada gambar 3.3 sampel gambar A adalah hasil rontgen paru-paru normal dan gambar B adalah hasil rontgen paru-paru pasien TBC. Yang membedakan antar gambar rontgen paru-paru normal dengan gambar paru-paru TBC pada umumnya yaitu terdapat bercak putih atau flek putih berukuran kecil yang dalam dunia Kesehatan sering disebut infiltrate.

### **3.3 Alat dan Bahan**

Penelitian ini menggunakan perangkat keras Laptop *Acer Aspire Intel Celeron N4000* up to 2,6 Ghz, sedangkan perangkat lunak yang digunakan *Microsoft Word*, *Mendeley* dan *Orange* aplikasi yang digunakan untuk pengolahan data.