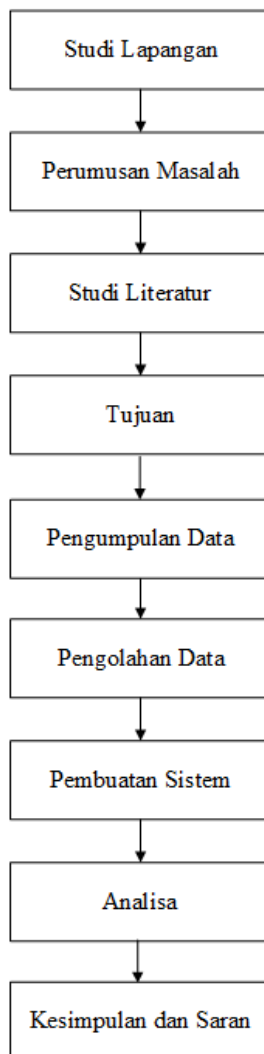


## BAB III METODE PENELITIAN

Penelitian dilakukan dengan menggunakan metode kuantitatif, dimana penyelesaian dari masalah penelitian dilakukan dengan pengambilan data yang kemudian diukur dan dianalisa sesuai dengan hasil pengujian dari penelitian ini.

### 3.1 Alur Penelitian



**Gambar 3. 1 Diagram Alur Penelitian**

#### 3.1.1 Studi Lapangan

Salah satu cara untuk memperoleh data dengan melakukan pengamatan langsung terhadap objek yang akan diteliti untuk mendapatkan informasi mengenai apa yang akan direncanakan.

### **3.1.2 Perumusan Masalah**

Perumusan masalah dibuat dengan mempertimbangkan beberapa hal yaitu dari segi metode, hasil penelitian, dan analisis penelitian yang akan dilakukan. Tujuan dibuatnya perumusan masalah dalam skripsi ini yaitu untuk mengetahui bagaimana sistem yang seharusnya dibuat untuk memprediksi penyakit paru-paru menggunakan 10 algoritma *Machine Learning* dengan data riil dari rumah sakit Cut Meuthia.

### **3.1.3 Studi Literatur**

Studi literature ini bertujuan untuk memprediksi penyakit paru-paru menggunakan 10 algoritma *Machine Learning* yang diimplementasikan dengan menggunakan data yang berasal dari rumah sakit Cut Meuthia. Pada tahap awal studi, akan dilakukan penelusuran dan analisis terhadap literatur yang relevan mengenai prediksi penyakit paru-paru menggunakan metode *Machine Learning*. Studi ini akan melibatkan tinjauan terhadap artikel jurnal, buku, makalah konferensi, dan sumber informasi terpercaya lainnya yang berkaitan dengan topik tersebut.

### **3.1.4 Pengumpulan Data**

Setelah mendapatkan pemahaman yang cukup dari studi literatur, langkah selanjutnya adalah mengumpulkan data riil dari rumah sakit Cut Meuthia yang berkaitan dengan penyakit paru-paru. Data tersebut akan mencakup berbagai atribut seperti usia, jenis kelamin, riwayat kesehatan, gejala klinis, dan hasil tes medis terkait penyakit paru-paru.

### **3.1.5 Pengolahan Data**

Setelah data dikumpulkan, selanjutnya dilakukan pengolahan data. Setelah data terkumpul, data tersebut akan dipreproses dan dipersiapkan untuk analisis lebih lanjut. Langkah ini termasuk membersihkan data, mengisi nilai yang hilang, dan mengubah format data jika diperlukan. Proses ini penting untuk memastikan kualitas data yang digunakan dalam pemodelan *Machine Learning*.

### 3.1.6 Pembuatan Sistem

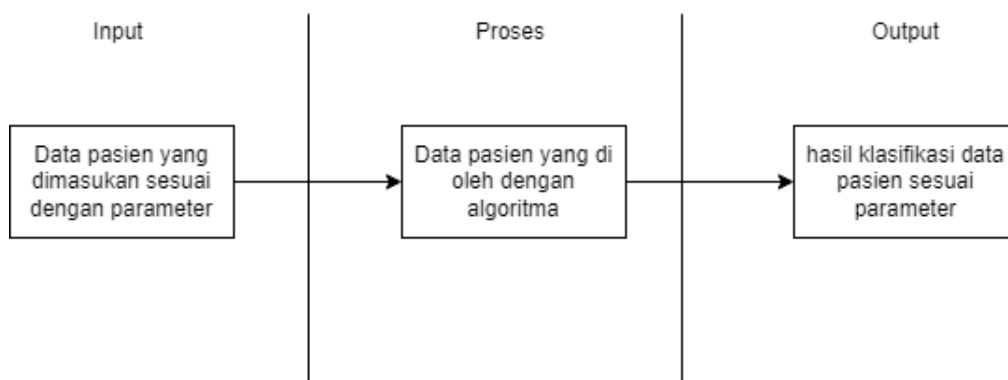
Setelah data terkumpul, data tersebut akan dipreproses dan dipersiapkan untuk analisis lebih lanjut. Langkah ini termasuk membersihkan data, mengisi nilai yang hilang, dan mengubah format data jika diperlukan. Proses ini penting untuk memastikan kualitas data yang digunakan dalam pemodelan Machine Learning. Hasil dari setiap algoritma akan dievaluasi menggunakan metrik evaluasi yang relevan, seperti akurasi, *presisi*, *recall*, dan *F1-score*. Analisis ini akan memberikan pemahaman yang lebih baik tentang kinerja masing-masing algoritma dalam konteks prediksi penyakit paru-paru.

### 3.1.7 Analisa

Dalam akhir studi literature ini, akan disajikan rangkuman temuan dan kesimpulan yang diambil dari hasil analisis data menggunakan 10 algoritma *Machine Learning*. Kesimpulan tersebut dapat mencakup algoritma yang memberikan performa terbaik dalam memprediksi penyakit paru-paru berdasarkan data riil dari rumah sakit Cut Meuthia. Studi ini dapat memberikan wawasan penting bagi pengembangan metode prediksi penyakit paru-paru menggunakan Machine Learning dalam konteks rumah sakit tersebut.

## 3.2 Pemodelan Sistem

Pemodelan sistem ini dibuat sebagai awal dari perencanaan *software* yang akan dirancang.



**Gambar 3. 2 Blok Diagram Perancangan Sistem**

Pada Gambar 3.2 menjelaskan alur diagram perancangan sistem yang terdiri dari tiga bagian yaitu perancangan *input*, proses, dan *output* dan pada sistem kerja alat mempunyai fungsi yang berbeda-beda. Diagram *input* diatas menentukan apa

saja data yang akan dimasukkan yaitu: nama, jenis kelamin, usia, keluhan pasien, suhu, *respiration rate*, saturasi oksigen, tekanan darah, dan *heart rate*. Pada bagian proses, data pasien dikumpulkan lalu diklasifikasi menggunakan metode dan algoritma yang sudah ditentukan, dan hasil outputnya merupakan hasil untuk mendiagnosis pasien apakah terdeteksi penyakit paru-paru.

### 3.3 Alat yang Digunakan

#### 3.3.1 Hardware

Pada penelitian digunakan *hardware* berupa satu buah laptop untuk menjalankan *software* simulasi, berikut ini spesifikasi dari *hardware* yang digunakan

**Tabel 3. 1 Spesifikasi Laptop**

Sistem Operasi	Windows 10 Pro 64-bit
<i>Processor</i>	Intel® Core™ i7-6700HQ CPU @ 2.60GHz
<i>RAM</i>	16 GB
<i>Storage</i>	500 GB

#### 3.3.2 Software

Selain *hardware*, digunakan juga *software* sebagai penunjang dari penelitian ini. *Software* yang digunakan adalah WEKA merupakan sebuah perangkat lunak yang menerapkan berbagai algoritma *machine learning* untuk melakukan beberapa proses yang berkaitan dengan sistem temu kembali informasi atau data *mining*. Beberapa fitur unggulan yang dimiliki oleh WEKA yaitu *classification*, *regression clustering*, *association rules*, *visualization*, dan *data preprocecing*.

#### 3.3.3 Dataset

Dalam penelitian ini, data dikumpulkan dari Rumah Sakit Cut Meuthia, sebanyak 300 *records*, dan parameter data yang digunakan dalam penelitian ini ada pada Tabel 3.2.

**Tabel 3. 2 Format Data Asli**

No	Parameter	Format
1	Batuk (bt)	<i>Text</i>

No	Parameter	Format
2	Batuk berdahak (bh)	<i>Text</i>
3	Batuk berdarah (br)	<i>Text</i>
4	Sesak (ss)	<i>Text</i>
5	Jenis kelamin	<i>Binary ( pria dan wanita )</i>
6	Usia	Numerik
7	Diagnosis	<i>Text</i>
8	Suhu	Numerik
9	<i>Repository rate</i>	Numerik
10	Saturasi oksigen	Numerik
11	Tekanan darah	Numerik
12	<i>Heart rate</i>	Numerik
13	Diagnosis penyakit	<i>Text</i>

Untuk tujuan analisis, format data yang berbentuk teks harus diubah ke biner, integer atau code sehingga data dapat diproses oleh algoritma. Pertama, parameter “gejala” harus dikonversi ke data biner. Ada gejala utama yang berhubungan dengan penyakit paru – paru. yaitu, batuk (bt), batuk berdahak (bh), batuk berdarah (br), sesak (ss).

Parameter harus dikonversi ke data biner untuk diproses oleh algoritma menjadi Tabel 3.3 yaitu hasil konversi data dari rumah sakit cut meuthia dengan format parameter yang disesuaikan dengan algoritma

**Tabel 3. 3 Pre-Processing Data**

No	Parameter	Format Asli	Format Setelah Pre-Processing
1	Jenis kelamin	<i>Binary</i>	<i>Binary</i>
2	Usia	Numerik	Numerik
3	Batuk (bt)	<i>Text</i>	<i>Binary</i>
4	Batuk berdahak (bh)	<i>Text</i>	<i>Binary</i>
5	Batuk berdarah (br)	<i>Text</i>	<i>Binary</i>
6	Sesak (ss)	<i>Text</i>	<i>Binary</i>

No	Parameter	Format Asli	Format Setelah Pre- <i>Processing</i>
7	Suhu	Numerik	Numerik
8	<i>Repository rate</i>	Numerik	Numerik
9	Saturasi oksigen	Numerik	Numerik
10	Tekanan darah	Numerik	Numerik
11	<i>Heart rate</i>	Numerik	Numerik
12	Diagnosis penyakit	<i>Text</i>	<i>Binary</i>

### 3.3.4 Data Mining

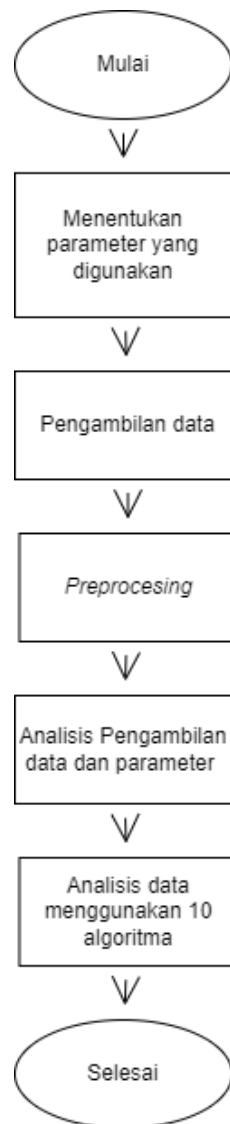
Untuk analisis *data mining* dalam penelitian ini, digunakan sebuah *tools* yang bernama WEKA alat ini sangat populer untuk analisis *data mining*, dengan lisensi yang gratis. Dalam penelitian ini menggunakan sepuluh algoritma yang sudah disebutkan pada dasar teori yang akan diuji dengan *tools* WEKA. Metodologi yang digunakan dalam WEKA untuk setiap algoritma terdapat pada Tabel 3.4

**Tabel 3. 4 Metode WEKA Untuk Analisis**

No	Algoritma	Klasifier
1.	Naive Bayes	Cross-validation folds = 10
2.	Logistic Regression	Maxlts = 1 Ridge 1.0E-8
3.	KNN	k = 9 DistanceWeighting = no NNSearchAlgorithm = LinearNN Search
4.	K-Star	globalBlend = 20 numDecimalPlaces = 2 missingMode = Average column entropy curves batchsize = 100
5.	AdaBosstM1	Classifier = Decision Stump numlterations = 0 seed = 1

No	Algoritma	Klasifier
		weightThreshold = 100
6.	Bagging Algoritma	bagSizePercent = 100 classifier = REPTree numliterations = 10 seed=1
7.	OneR	batchSize = 100 minBucketSize = 6 numdecilmalplaces = 2
8.	RandomForest	numTrees = 100 maxdepth = 0 seed = 1
9.	SVM (SMO)	SVMType = C-SVC Degree = 3 Gamma = 0 Epsilon = 0.001 KernalType = radial basis function
10.	MLP	LearningRate = 0.3 Momentum = 0.2 Seed = 0 TrainingTime = 500

### 3.4 Flowchart Penelitian



**Gambar 3. 3 Flowchart Penelitian**

Berikut adalah penjelasan skema flowchart penelitian dari tahap-tahap yang disebutkan:

1. Mulai: Tahap ini adalah awal dari penelitian, di mana penelitian dimulai dengan merumuskan tujuan, permasalahan, dan lingkup penelitian.
2. Menentukan Parameter yang Digunakan: Pada tahap ini, peneliti menentukan parameter yang akan digunakan dalam analisis dan prediksi penyakit paru-paru. Parameter tersebut mungkin mencakup gejala, suhu tubuh, laju pernapasan, saturasi oksigen, tekanan darah, detak jantung, jenis



kelamin, dan usia.

3. Pengambilan Data: Peneliti mengumpulkan data kesehatan pasien dari Rumah Sakit Cut Meuthia untuk digunakan dalam penelitian ini. Data ini mencakup informasi tentang parameter-parameter yang telah ditentukan sebelumnya.
4. Preprocessing: Pada tahap ini, data yang telah dikumpulkan akan melalui proses preprocessing untuk membersihkan data dari noise atau data yang tidak valid, melakukan normalisasi jika diperlukan, dan melakukan seleksi fitur untuk mengurangi dimensi data jika perlu.
5. Analisis Pengambilan Data dan Parameter: Pada tahap ini, peneliti akan menganalisis hubungan antara data kesehatan yang telah dikumpulkan dengan parameter-parameter yang telah ditentukan sebelumnya. Tujuan dari analisis ini adalah untuk memahami lebih jauh tentang karakteristik data yang akan digunakan dalam proses prediksi.
6. Analisis Data menggunakan 10 Algoritma: Tahap ini merupakan inti dari penelitian, di mana data yang telah dipreprocess akan digunakan untuk analisis prediksi penyakit paru-paru menggunakan 10 algoritma machine learning yang berbeda. Algoritma-algoritma ini mungkin termasuk Naïve Bayes, KNN, SVM, dan lain-lain. Hasil analisis dan performa masing-masing algoritma akan dievaluasi dan dibandingkan untuk memilih algoritma yang paling tepat dalam memprediksi penyakit paru-paru berdasarkan akurasi dan kinerjanya.
7. Selesai: Tahap ini menandakan berakhirnya penelitian. Pada tahap ini, peneliti menyimpulkan hasil penelitian, memberikan kesimpulan, dan memberikan rekomendasi berdasarkan temuan penelitian yang telah dilakukan.

Dengan menggunakan skema flowchart ini, penelitian dapat dijalankan secara sistematis dan mengikuti alur logis dari tahap awal hingga akhir penelitian.

### **3.5 Skenario Pengujian Sistem**

Pengujian sistem pada penelitian tugas akhir ini sangat diperlukan, hal ini bertujuan untuk melihat kerja dari sistem yang dibuat atau yang telah dirancang apakah dapat bekerja dengan baik atau tidak sesuai dengan fungsi yang diharapkan

penulis. Pengujian ini dilakukan dengan cara, menguji tingkat keberhasilan dan peforma dari upaya penelitian yang bertujuan untuk membandingkan efektivitas 10 *algoritma machine learning* dalam memprediksi penyakit paru-paru menggunakan data riil yang diperoleh dari Rumah Sakit Cut Meuthia. Pengujian sistem ini didesain dengan tujuan utama untuk memperoleh pemahaman yang lebih mendalam tentang keandalan dan akurasi setiap algoritma dalam mengidentifikasi penyakit paru-paru.

**Tabel 3. 5 Skema Pengujian Sistem**

No	Algoritma	Klasifier	Hasil
1.	Naive Bayes	Cross-validation folds = 10	Akurasi = 78,67 % Kecepatan = 0.01 m/s
2.	Logistic Regression	Maxlts = 1 Ridge 1.0E-8	Akurasi = 72.33 % Kecepatan = 0.01 m/s
3.	KNN	k = 9 DistanceWeighting = no NNSearchAlgorithm = LinearNN Search	Akurasi = 77.67 % Kecepatan = 0 m/s
4.	K-Star	globalBlend = 20 numDecimalPlaces = 2 missingMode = Average column entropy curves batchsize = 100	Akurasi = 71.67 % Kecepatan = 0 m/s
5.	AdaBosstM1	Classifier = Decision Stump numlterations = 0 seed = 1 weightThreshold = 100	Akurasi = 74.67 % Kecepatan = 0.04 m/s
6.	Bagging Algoritma	bagSizePercent = 100 classifier = REPTree numliterations = 10 seed=1	Akurasi = 70 % Kecepatan = 0.06 m/s
7.	OneR	batchSize = 100 minBucketSize = 6	Akurasi = 68.67% Kecepatan = 0 m/s

No	Algoritma	Klasifier	Hasil
		numdecimalplaces = 2	
8.	RandomForest	numTrees = 100 maxdepth = 0 seed =1	Akurasi = 72.67 % Kecepatan = 0.13 m/s
9.	SVM (SMO)	SVMType = C-SVC Degree = 3 Gamma = 0 Epsilon = 0.001 KernalType = radial basis function	Akurasi = 75 % Kecepatan = 0.07 m/s
10.	MLP	LearningRate = 0.3 Momentum = 0.2 Seed = 0 TrainingTime = 500	Akurasi =74.33 % Kecepatan = 2.75 m/s