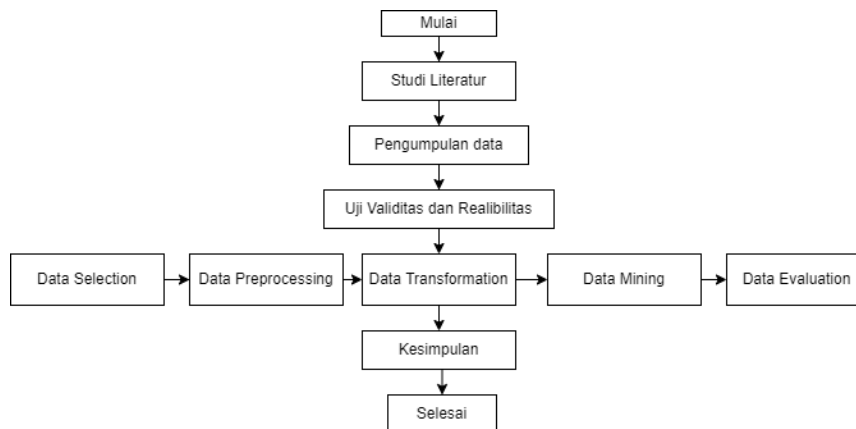


## BAB III METODOLOGI PENELITIAN

### 3.1 Diagram Alir Penelitian/Proses Penelitian

Subab ini menjelaskan tahapan penelitian yang akan dilakukan. Proses yang akan terjadi dimulai dari tahap studi literatur, pengumpulan data, uji validitas dan reabilitas, *Knowledge Discovery Database* (KDD) hingga kesimpulan. Berikut ini merupakan gambaran umum dari tahapan-tahapan yang akan dilakukan dalam penelitian.



**Gambar 3. 1 Diagram Alir Penelitian**

Gambar 3.1 merupakan tahapan proses pada penelitian ini. Berikut merupakan penjelasan dari tahapan pada diagram alir penelitian:

#### 3.1.1 Studi Literatur

Studi Literatur dilakukan dengan membaca sumber ilmiah seperti buku, artikel jurnal, dan tesis yang terkait dengan topik atau pertanyaan penelitian ini. Hal ini dilakukan untuk mengetahui kaitannya penelitian ini dengan pengetahuan yang sudah ada.

#### 3.1.2 Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan menyebarkan Google Form berisi kuisioner *Depression Anxiety and Stress Scale* (DASS-21) dengan menggunakan *Convenience Sampling*. *Convenience sampling* adalah teknik pengambilan sampel berdasarkan kemudahan peneliti, yaitu yang ditemui peneliti

secara kebetulan, dipandang cocok, serta bersedia menjadi sumber data dan sesuai dengan kriteria yang ditentukan oleh peneliti. Oleh karena itu, responden pada penelitian ini adalah pengguna *Facebook* yang bersedia mengisi kuesioner yang diberikan. Kuisisioner disebar melalui grup *WhatsApp*, grup *Facebook*, *Instagram*, dan responden penelitian ini meliputi pengguna *Facebook* seluruh Indonesia.

### 3.1.3 Uji Validitas dan reliabilitas

Secara teknik untuk melihat baik tidaknya alat pengukur penelitian, dapat diketahui dengan melihat kriteria utamanya, yaitu (a) validitas, (b) reabilitas atau kehandalan [8]. Teknik mengetahui suatu instrumen valid atau tidak dapat dilakukan dengan menghitung koefisien korelasi uji instrumen dengan uji kriteria atau product moment.

$$r_{xy} = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{(n(\sum x_i^2) - (\sum x_i)^2)(n(\sum y_i^2) - (\sum y_i)^2)}} \quad (3.1)$$

Hasil perhitungan koefisien korelasi kemudian dibandingkan dengan nilai-nilai tabel. Jika  $r$  hitung  $>$  atau  $=$   $r$  tabel maka pertanyaan dikatakan valid. Jika  $r$  hitung  $<$   $r$  tabel maka pertanyaan dikatakan tidak valid. Pada penelitian ini digunakan 193 data hasil kuesioner untuk dihitung nilai validitas dan reabilitasnya. Pada pengujian reabilitas instrumen menggunakan perhitungan koefisien *Alfa Cronbach* dengan persamaan 3.1. Hasil perhitungan koefisien alfa kemudian dibandingkan dengan nilai nilai tabel. Jika alfa  $>$  atau  $=$   $r$  tabel maka instrumen penelitian dikatakan *reliable*. Jika nilai alfa  $<$   $r$  tabel maka instrumen penelitian dikatakan tidak *reliable*.

$$r_i = \frac{k}{(k-1)} \left\{ 1 - \frac{\sum s_{i^2}}{s_{i^2}} \right\} \quad (3.2)$$

### 3.1.4 Knowledge Discovery Database

Dalam penelitian ini, Penulis menggunakan metode atau tahapan *Knowledge Discovery Database*. Berikut merupakan tahapan-tahapan proses yang ada pada *Knowledge Discovery Database*:

#### A. Data Selection

*Data Selection* didefinisikan sebagai proses penentuan jenis dan sumber data yang sesuai. Selain itu, pada tahap ini juga ditentukan tujuan dari proses *Knowledge Discovery Database*, termasuk mencari data apa saja yang tersedia dan mengintegrasikan semua data untuk proses *Knowledge Discovery* menjadi satu *dataset*. Pada tahap ini dilakukan pemilihan data, data yang digunakan dalam penelitian ini berasal dari hasil Google Formulir yang telah disebar menggunakan Kuisisioner *Depression Anxiety Stress and Scale 21* (DASS-21).

### **B. Data Preprocessing**

*Data Preprocessing* adalah *data cleaning* atau pembersihan data dari *noise*. Pada tahap ini peneliti membersihkan hasil kuisisioner yang kurang lengkap, menghilangkan missing value atau memperbaiki kesalahan data yang tidak sesuai dengan format *dataset* penelitian agar tidak terjadi redundansi atau penumpukan data yang mengakibatkan data menjadi tidak konsisten.[50]

$$\text{Hasil kuesioner DASS Total} = (\sum \text{sub item}) \times 2 \quad (3.3)$$

Nilai total setiap item dihitung dengan melakukan penambahan ke semua sub-item dan kemudian dikalikan dua. Setelah didapatkan nilai total setiap item, langkah selanjutnya adalah membandingkan nilai setiap item, dan dipilih nilai item tertinggi untuk dijadikan label.

### **C. Data Transformation**

*Data Transformation* ialah proses mengubah format, struktur, atau nilai data. Data yang didapatkan dari kuisisioner disesuaikan dengan format data untuk proses *data mining*. Transformasi data dilakukan untuk memudahkan peneliti dalam menganalisis dengan software pendukung teknik *data mining*. Yaitu dilakukan dengan perubahan data pada input data numerik menjadi kategorik agar dapat dikategorikan sehingga data dapat diproses dalam *data mining*.

### **D. Data mining**

*Data mining* merupakan proses mengungkap pola dan informasi berharga lainnya dari kumpulan data besar. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery Database* secara keseluruhan. Proses klasifikasi pada penelitian ini dengan menggunakan metode

*Support Vector Machine*. Hasil dari proses *data mining* ini adalah klasifikasi dari depresi kecemasan dan *stress*.

### **E. Data Evaluation**

*Data Evaluation* didefinisikan sebagai mengidentifikasi pola yang meningkat mewakili pengetahuan berdasarkan ukuran yang diberikan. Pola informasi (*knowledge*) yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang dapat dimengerti oleh pihak yang berkepentingan. Dalam proses ini dilakukan evaluasi performansi menggunakan *confusion matrix* meliputi *accuracy*, *precision*, *f1*, dan *recall*.

1. *Accuracy* merupakan perbandingan kasus yang terklasifikasi benar dengan jumlah seluruh kasus.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.4)$$

2. *Recall* menunjukkan berapa persen data kategori positif yang terklasifikasi dengan benar oleh sistem.

$$Recall = \frac{TP}{TP+FN} \quad (3.5)$$

3. *Precision* mengevaluasi kemampuan sistem untuk menemukan peringkat yang paling relevan dan didefinisikan sebagai presentase dokumen yang di retrieve dan benar-benar relevan terhadap query.

$$Precision = \frac{TP}{TP+FP} \quad (3.6)$$

4. *F1* secara representasi mengindikasikan bahwa model klasifikasi yang digunakan mempunyai *precision* dan *recall* yang baik atau sebaliknya.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.7)$$

### **3.1.5 Kesimpulan**

Dari hasil uji coba yang telah dilakukan dalam penelitian ini didapatkan hasil klasifikasi depresi kecemasan dan *stress*. Setelah hasil klasifikasi didapatkan maka dilakukan evaluasi performansi terhadap metode *Support Vector Machine*.