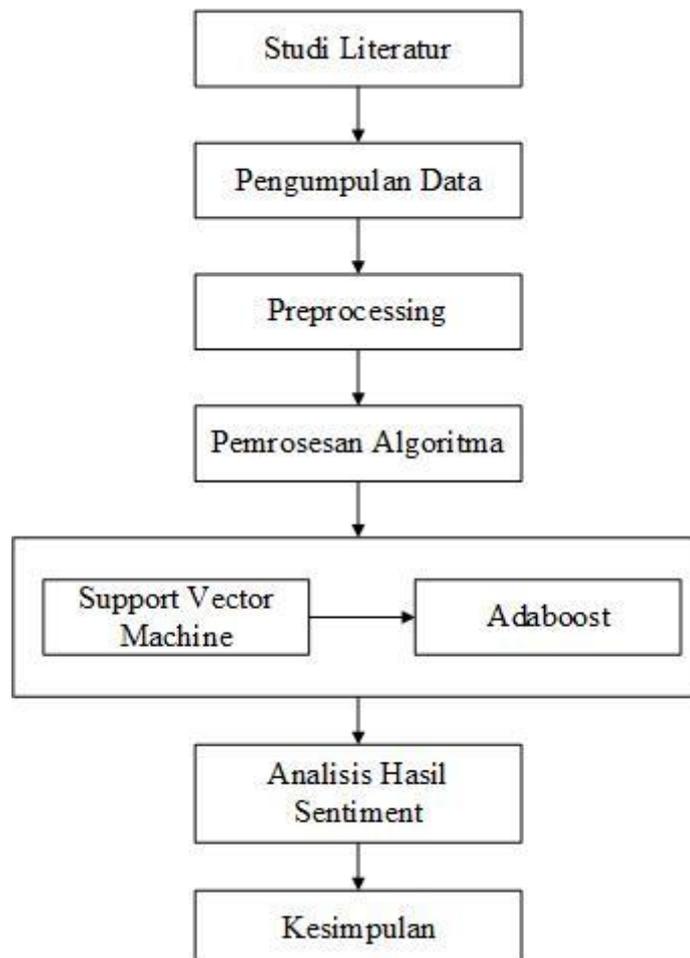


BAB III

METODOLOGI PENELITIAN

3.1. Metodologi Penelitian

Pada bab ini akan membahas langkah-langkah dari proses penelitian yang akan dikerjakan, dalam melakukan analisa dan mencari sebuah dataset dalam melakukan scraping untuk memudahkan penelitian dan dapat berjalan dengan sistematis dan memenuhi tujuan yang diinginkan. Pembuatan alur penelitian merupakan sebagai tahapan dalam melakukan penelitian ini. Berikut adalah alur penelitian yang akan dilakukan dalam penelitian ini.



Gambar 3. 1. Alur Penelitian

3.2. Studi Literatur

Tahapan yang pertamakali dilakukan yaitu studi literatur atau studi awal yang merupakan tahap yang pertama kali dilakukan dalam penelitian. Pada tahap ini, peneliti melakukan observasi atau 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 dalam melakukan sentiment analisis. Dalam studi literatur ini juga dilakukan pencarian terhadap sumber-sumber teori yang relevan dengan topik penelitian dari berbagai sumber seperti buku, penelitian terkait yang ada dalam suatu jurnal, internet, dan lain sebagainya yang mendukung proses penelitian dengan tujuan untuk memperkuat permasalahan.

3.3. Pengumpulan Data

Data yang digunakan pada penelitian ini adalah berasal dari kolom ulasan aplikasi Game Mobile Legend yang ada di website Google Play Store. Google Play Store adalah aplikasi resmi milik Google untuk perangkat yang menggunakan sistem baik operasi Android maupun Web. Bagi pengguna bisa mendapatkan aplikasi Tokopedia dapat dengan cara mengunduh aplikasi Game Mobile Legend di Google Play Store. Pada Google Play Store juga terdapat kolom ulasan yang berisi ulasan para pengguna. Banyaknya dan beragam ulasan tersebut dapat dijadikan sebagai bahan penelitian di bidang data mining. Cara untuk mendapatkan ulasan tersebut dilakukan proses yang di namakan *Web Scraping*

3.4. Preprocessing

Tahap ini dilakukan untuk mengolah data mentah agar data terbebas dari kesalahan. Banyaknya data ulasan yang didapat banyak yang menggunakan kata yang tidak berstruktur seperti singkatan, emotikon, simbol dang angka hal ini diperlukan teknik preprocessing. Preprocessing berguna untuk mengekstrak informasi dari ulasan, mengubah kata-kata yang tidak terstruktur itu menjadi bentuk standar. Tahap ini terdiri dari beberapa proses, diantaranya:

a. Case Folding

Pada tahapan Case folding adalah pemrosesan untuk mengubah kata menjadi bentuk yang sama, misalnya huruf kecil atau besar. Pada data ini peneliti mengubah semua kata menjadi huruf kecil atau disebut lowercase.

b. Penhapusan Simbol Angka dan Emoji

Menghapus simbol, angka, string ASCII, dan tanda baca. Pesan ulasan biasanya berisi simbol, angka, dan tanda baca. Semua ini akan dihapus menggunakan sintaks ekspresi reguler.

c. Lemmatization

Pada tahapan Lemmatization adalah pemrosesan transformasi untuk menemukan bentuk normalisasi satu kata. Mengubah kata menjadi dalam bentuk kata dasarnya agar lebih mudah inti dari kata pada ulasan seperti yang ada pada dataset penelitian.

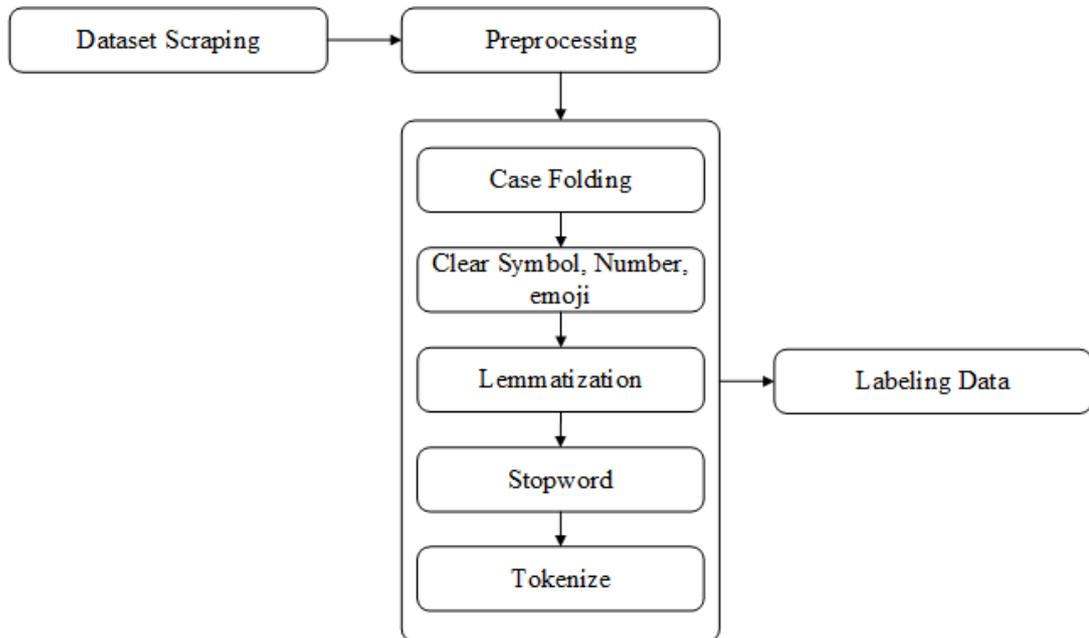
d. Stopword

Tahapan pemrosesan Stopword adalah untuk menghilangkan kata-kata umum dan sering yang tidak memiliki pengaruh signifikan dalam kalimat. Dalam tugas pra-pemrosesan ini seperti menghapus stopwords pada data ulasan berdasarkan daftar kata stopwords yang berisi kata-kata seperti yang ada pada dataset.

e. Tokenize

Tahapan Tokenizer bertujuan untuk membagi kalimat menjadi beberapa bagian. Tokenisasi dapat dibentuk dalam kata-kata, frasa atau elemen bermakna lainnya[30]. Pada prinsipnya proses ini adalah memisahkan setiap kata yang menyusun suatu dokumen. umumnya setiap kata terpisahkan dengan kata yang lain oleh karakter spasi, sehingga proses tokenisasi mengandalkan karakter spasi pada dokumen untuk melakukan pemisahan kata. Setelah melalui proses tokenisasi maka kalimat tersebut menjadi sekumpulan array yang setiap selnya berisi kata-kata yang ada pada kalimat tersebut.

Setelah tahapan preprocessing maka dibuat labeling data sebagai pengolahan data untuk melakukan modeling pada sentiment analisis dengan menggunakan *Support Vector Machine* dan *juga Adaboost*. Berikut adalah gambar pemrosesan preprocessing dalam penelitian ini.

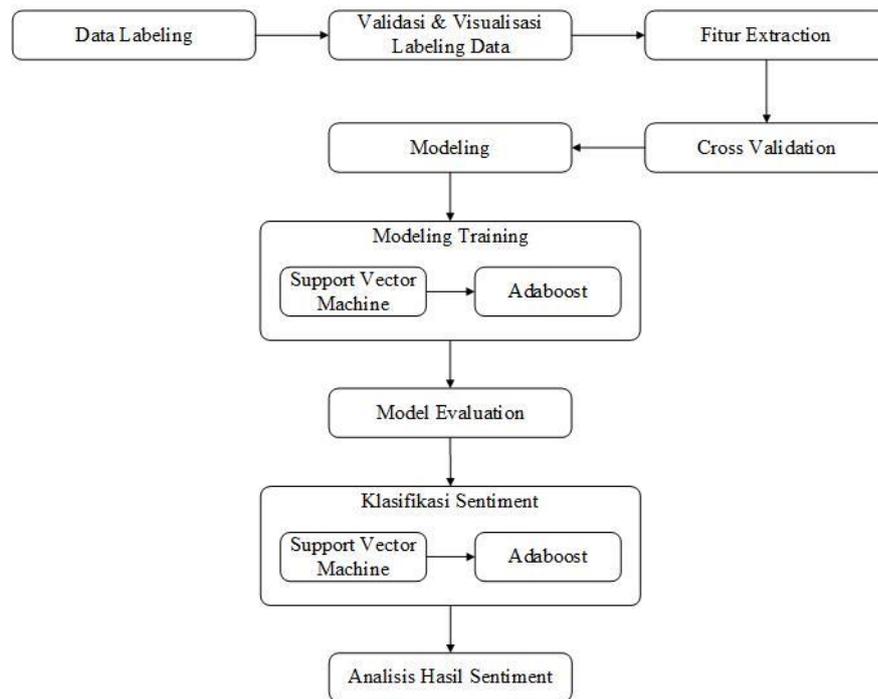


Gambar 3. 2. Tahapan Preprocessing

3.5. Pemrosesan Algoritma

Pada tahapan pemrosesan algoritma setelah dilakukan preprocessing pada dataset yang telah dilakukan dan juga pelabelan data hingga label data tersebut dilakukan validasi pelabelan untuk menentukan feature extraction sehingga dapat dilakukan cross validation untuk melakukan pemrosesan agar data dengan menggunakan *Support Vector Machine* dan juga *adaboost*. Pada tahap ini, dilakukan pembangunan model atau *modeling*. *Model training* diawali dengan melatih *Support Vector Machine*, pelatihan SVM dilakukan berdasarkan rumus dari SVM itu sendiri. Kemudian *Output* dari *Support Vector Machine* berupa hasil klasifikasi data. Kemudian meningkatkan *weak learner Support Vector Machine* dengan *Adaboost*.

Setelah model dilatih, model akan dievaluasi dengan dua sudut pandang, yaitu dari sudut pandang label positif, dan negatif agar dapat diketahui kemampuan pengklasifikasi dalam mengelompokkan masing-masing label. Evaluasi menggunakan *recall* dengan rumus, *precision* berdasarkan rumus, *f-measure* (2.16), dan *accuracy*. Hasil evaluasi dari SVM dengan *Adaboost* akan dibandingkan. Berikut adalah tahapan pemrosesan dalam melakukan sentiment analisis dengan menggunakan *Support Vector Machine* dan *adaboost*.



Gambar 3. 3. Tahapan Modeling dengan SVM dan Adaboost

3.6. Analisis Hasil Sentiment

Setelah model selesai dibuat, model digunakan untuk analisis sentimen. Seluruh data digunakan untuk analisis sentimen. Pada saat menganalisis hasil sentimen, dipaparkan jumlah sentimen positif, dan negatif serta pemaparan dari hasil Confusion Matrix untuk mengukur kinerja yang memiliki beberapa parameter dengan Parameter yang digunakan adalah *TP* (true positive), *FN* (false negative), *TN* (true negative), dan *FP* (false positive) pada data uji.

3.7. Alat dan Bahan

Penelitian ini menggunakan perangkat keras Laptop ASUS dengan Processor Intel(R) Core(TM) i3-7020U CPU @ 2.30GHz 1.30 GHz, RAM 4,00 GB, sedangkan perangkat lunak yang digunakan Microsoft Excel dan RapidMiner Studio untuk pengolahan data dan bahan yang digunakan adalah scraping data komentar masyarakat pada aplikasi game mobile legend di google play store.