

## BAB 3 METODOLOGI PENELITIAN

### 3.1 Metodologi Penelitian

#### 3.1.1 Perumusan Masalah

Menggabungkan masalah dan informasi yang telah didapatkan sehingga dapat menghasilkan solusi dari permasalahan tersebut. Masalah yang ditemukan adalah bagaimana ulasan gim dapat menentukan opini dari gim tersebut, sehingga dilakukan analisis sentimen dari gim yang ada dalam toko digital Steam.

#### 3.1.2 Studi Pustaka

Penelitian dilakukan dengan mencari atau mendapatkan informasi dengan cara membaca buku, artikel, maupun *browsing* terhadap metode analisis sentimen dan perbandingannya, serta mengengai metode *Support Vector Machine* dan bagaimana dapat meningkatkan performanya dengan metode *Particle Swarm Optimization*.

#### 3.1.3 Pengumpulan Data

Pengumpulan data dilakukan menggunakan dataset ulasan dari toko digital *Steam*, karena merepresentasikan sebagian besar pemain komputer. Data tersebut dikumpulkan di Kaggle [13]. Data yang dipilih adalah ulasan dari gim *Baldur's Gate 3* Data yang digunakan merupakan versi setelah *patch 5*, tanggal 30 November 2023.

#### 3.1.4 Perancangan Sistem dan Implementasi

Setelah data dikumpulkan dilakukan analisa solusi dan perancangan sistem menggunakan *flowchart*. Kemudian, algoritma akan diimplementasikan pada data ulasan yang telah diolah, algoritma yang digunakan adalah *Support Vector Machine* dan *Particle Swarm Optimization*.

### 3.1.5 Uji Coba dan Evaluasi

Setelah data yang diperoleh hasilnya akan dilakukan evaluasi dimana data akan diuji akurasi serta presisi dari percobaan yang dilakukan. Evaluasi pada penelitian ini menggunakan *Confusion Matrix* dimana akan dicari nilai presisi, akurasi, *recall* serta f1-score untuk mengetahui kesesuaian algoritma yang telah dipakai

### 3.1.6 Penulisan Laporan

Setelah dilakukan ujicoba, menulis laporan dari hasil evaluasi tersebut dan hasil dari penelitian yang dilakukan, dan mencatat serta menyusun informasi tersebut menjadi hasil penelitian yang dapat diakui sehingga dapat membantu dalam penelitian selanjutnya.

## 3.2 Dataset

Ulasan-ulasan yang diambil berasal dari Kaggle [13]. Berikut adalah kriteria dari data ulasan dengan contoh data tabel pada Gambar 3.1.

- `recommendationid` = id unik ulasan gim.
- `language` = bahasa ulasan.
- `review` = isi ulasan.
- `timestamp_created` = waktu ulasan dibuat (format unix).
- `timestamp_updated` = waktu ulasan diubah.
- `voted_up` = hasil rekomendasi.
- `votes_up` = apakah ulasan membantu.
- `votes_funny` = apakah ulasan lucu.
- `written_during_early_access` = ulasan yang ditulis dalam masa awal/uji coba gim.
- `comment_count` = jumlah komentar.
- `steam_purchase` = pengguna yang membeli gim.

- `received_for_free` = pengguna yang mendapatkan gim secara gratis.

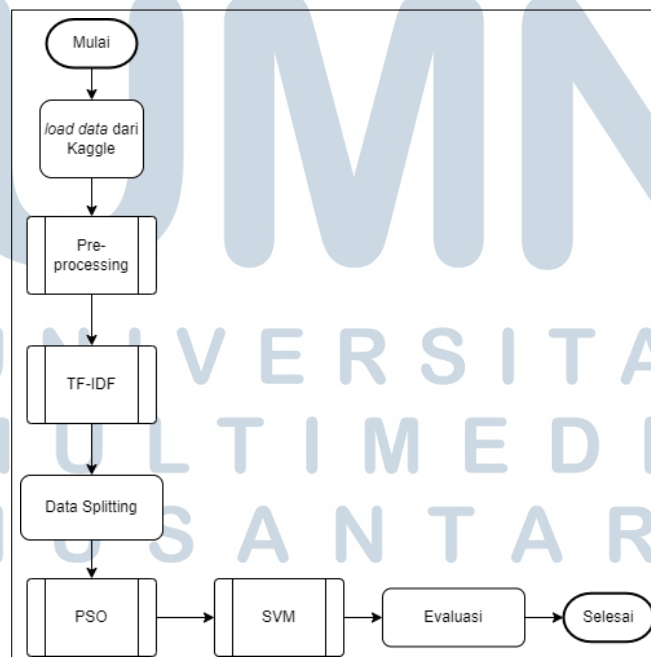
Kolom yang digunakan pada *training* adalah *review* dan *voted\_up*, sementara kolom *timestamp\_created*, *written\_during\_early\_access*, *steam\_purchase* dan *received\_for\_free* akan digunakan sebagai parameter untuk menyaring data yang digunakan.

Unnamed: 0	recommendationid	language	review	timestamp_created	timestamp_updated	voted_up	votes_up	votes_funny	weighted_vote_score	written_during_early_access	comment_count	steam_purchase	received_for_free	sentiment		
0	0	153560814	english	game hit right mark	1702542971	1702542971	True	0	0	0.0	0	False	0	True	False	positive
1	1	153560623	english	took like hour understand basic	1702542657	1702542657	True	0	0	0.0	0	False	0	True	False	positive
2	2	153560414	english	game play story first turn base rpg game base far	1702542275	1702542275	True	0	0	0.0	0	False	0	True	False	positive
3	3	153560343	english	game baby girl	1702542158	1702542158	True	0	0	0.0	0	False	0	True	False	positive
4	4	153560963	english	yes without rn doubt	1702541518	1702541518	True	0	0	0.0	0	False	0	True	False	positive

Gambar 3.1. Data ulasan

### 3.3 Perancangan Sistem

Langkah-langkah yang dilakukan dalam penyusunan penelitian dijabarkan dari awal sampai akhir. Langkah yang dilakukan adalah *load* data dari Kaggle [13], melakukan pemrosesan ulasan, pembobotan ulasan, ulasan dibagi menjadi data latih dan tes, setelah itu baru diimplementasikan algoritma PSO dan SVM dan akhirnya adalah evaluasi yang dapat dilihat pada Gambar 3.2.



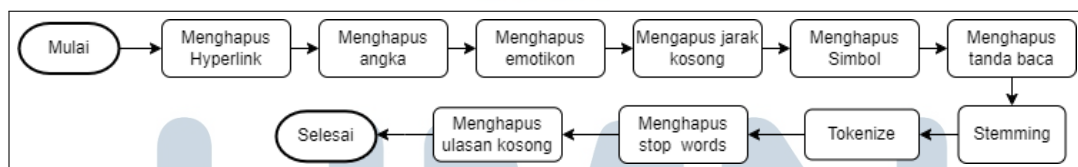
Gambar 3.2. Flowchart Sistem

### 3.3.1 Preprocessing

Pada tahap ini dilakukan untuk memudahkan pengolahan data yang akan dites, tahap ini berupa tahap sebagai berikut.

1. Cleaning, merupakan tahap untuk membersihkan kata-kata pada ulasan dari hal-hal yang tidak diperlukan seperti, simbol-simbol, emotikon, tanda baca dan tautan.
2. Case folding, mengubah semua kata ulasan menjadi huruf non-kapital (huruf kecil).
3. Tokenizing, memisahkan kalimat menjadi kata tunggal.
4. Filtering, melakukan penyaringan kata-kata umum yang sering digunakan.
5. Stemming, mengubah kata berimbuhan menjadi kata dasar.
6. Jika ada ulasan yang menjadi kosong akan dihapus.

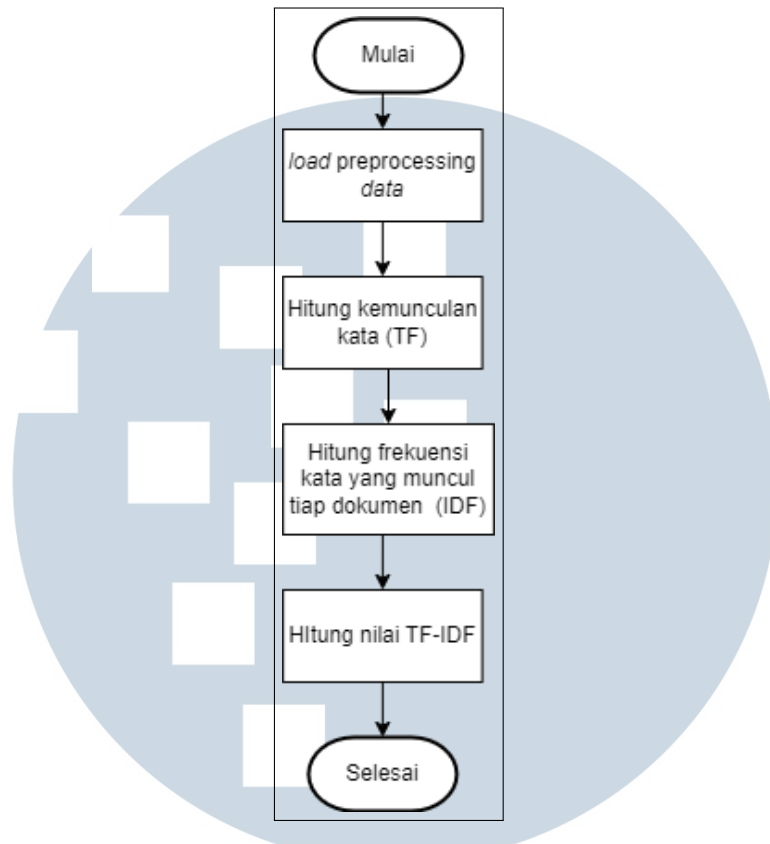
. Alur pengolahan ulasan dapat juga dilihat pada Gambar 3.3.



Gambar 3.3. Alur Preprocessing data

### 3.3.2 TF-IDF

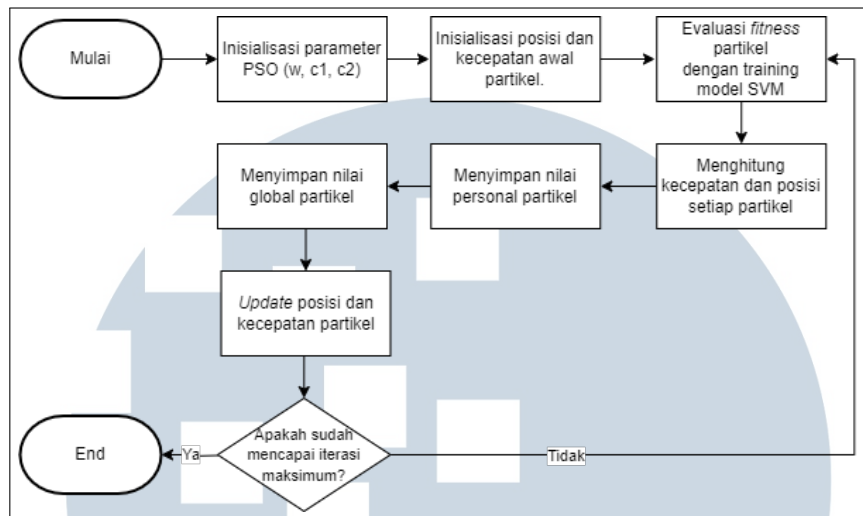
Data hasil pemrosesan akan dibobotkan dengan metode TF-IDF, selanjutnya ulasan akan dihitung istilah-istilah yang sering muncul dengan rumus 2.1 untuk perhitungan TF pada , selanjutnya penyebaran kata-kata dihitung seperti pada rumus 2.2, untuk mendapatkan bobot dari tiap istilah atau kata maka nilai menghitung nilai TF-IDF dengan rumus 2.3, langkah alur dapat dilihat pada gambar 3.4.



Gambar 3.4. Alur TF-IDF

### 3.3.3 Particle Swarm Optimization

Data yang sudah terbagi akan di optimisasi menggunakan metode PSO, Parameter  $w$ ,  $c1$  dan  $c2$  akan ditentukan, setelah itu menentukan populasi partikel, letak dan kecepatan pada partikel pertama-tama secara acak. Partikel tersebut akan dievaluasi dengan fungsi fitness yang akan menentukan tujuan dari partikel tersebut. Setelah itu, partikel akan dihitung kecepatan dan posisinya yang paling mendekati dengan tujuan tersebut, partikel yang paling mendekati akan dijadikan nilai *personal best* sementara populasi yang mendekati akan dijadikan nilai *global best*. Mengulang langkah tersebut hingga iterasi yang ditetapkan dicapai. Alur jalannya metode dapat dilihat pada Gambar 3.5.

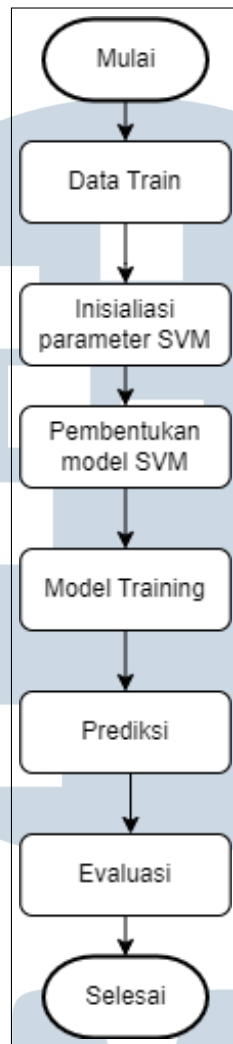


Gambar 3.5. Flowchart PSO

### 3.3.4 Support Vector Machine

Data ulasan dibagi menjadi dua yaitu data training dan data testing, lalu akan dibentuk model SVM. Data tersebut akan dilatih dan hasilnya akan diprediksi, jika hasilnya belum cukup tinggi maka akan dilakukan perubahan kernel pada model SVM dan membandingkan kernel mana yang memiliki prediksi akhir tinggi. Alur jalannya SVM dapat dilihat pada Gambar 3.6.





Gambar 3.6. Flowchart SVM

U I V I N  
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