

## BAB II

### TINJAUAN PUSTAKA

#### 1.1 Penelitian Terdahulu/Terkait

Penelitian terkait digunakan sebagai bahan pertimbangan yang memiliki kaitan dengan penelitian yang di lakukan saat ini. Terdapat beberapa penelitian terkait yang menjadi bahan referensi antara lain penelitian yang dilakukan oleh Nuraini Ika Pratiwi Kalingara, Oktariani Nurul Pratiwi, dan Hilman Dwi Anggana dengan judul Analisis Sentimen *Review Customer Terhadap Layanan Ekspedisi Jne Dan J&T Express Menggunakan Metode Naïve Bayes*.

Penelitian yang ke dua yaitu dari Rara Ayu Pusputa dengan judul penelitian Analisis sentiment Terhadap *Review E-Commerce Pada Google Play Store Menggunakan Metode Naïve Bayes classifier (NBC) Dengan Seleksi Fitur Informasi Gain (IG)*.

Penelitian Ke tiga yang dilakukan oleh Muhammad Farid El Firdaus, Nurfaizah, dan Sarmini dengan judul Analisis Sentimen Tokopedia Pada Ulasan di *Google Playstore Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor*.

Penelitian ke empat di lakukan oleh Mahardika Tania Nitami dan Herny Februariyanti dengan judul penelitian Analisis Sentimen Ulasan Ekspedisi J&T Express Menggunakan Algoritma *Naïve Bayes*.

Penelitian ke lima yang dilakukan oleh Muhammad Rezki, Desiana Nur Kholifah, Muhammad Faisal, Priyono, dan Rachmat Suryadithia dengan judul penelitian Analisis Review Pengguna *Google Meet dan Zoom Cloud Meeting Menggunakan Algoritma Naïve Bayes*.

Berikut ini adalah rincian penelitian-penelitian terdahulu yang mendasari penelitian sekarang.

**Tabel 2.1** Penelitian Terdahulu

No	Peneliti, Tahun	Judul Penelitian	Jumlah dan nama atribut	Preprocessing	Metodologi	Tempat penelitian	Hasil Penelitian
1	Muhammad Farid El Firdaus, Nurfaizah, Sarmini (2022)[5]	Analisis Sentimen Tokopedia pada ulasan di <i>Google Playstore</i> menggunakan algoritma	1. Ulasan	1. <i>Case Folding</i> 2. <i>Tokenizing</i> 3. <i>Stopword</i>	<i>Naïve Bayes, K-Nearest Neighbor</i>	Purwokerto, Indonesia	Algoritma Naïve Bayes: 75,30% Algoritma K-Nearest Neighbor : 86,09%

		<i>Naïve Bayes Classifier dan K-Nearest Neighbor</i>					
2	Rara Ayu Puspita (2022)[7]	Analisis sentimen terhadap <i>Review E-commerce</i> pada <i>Google play store</i> menggunakan metode <i>Naïve Bayes Classifier (NBC)</i> dengan seleksi <i>Fitur Information Gain (IG)</i>	1. Review 2. Rating 3. Komentari	1. <i>Spelling Normalization</i> 2. <i>Case Folding</i> 3. <i>Tokenizing</i> 4. <i>Filtering</i>	<i>Naïve Bayes Classifier (NBC)</i> , <i>Fitur Information Gain (IG)</i>	Semarang, Indonesia	<i>Naïve Bayes Classifier Tokopedia</i> : 50% <i>Naïve Bayes Classifier Shopee</i> : 40% <i>Fitur Information Gain Tokopedia</i> : 77,5% <i>Fitur Information Gain Shopee</i> : 65%
3	Nuraini Ika Pratiwi Kalingara, Okta Nurul Pratiwi, Hilman Dwi Anggana (2021)[8]	Analisis sentimen <i>Review customer</i> terhadap layanan ekspedisini <i>JNE dan JNT Express</i> menggunakan metode <i>Naïve Bayes</i>	1. Ketepatan pengiriman 2. Status paket 3. Estimasi waktu sampai	1. <i>Case Folding</i> 2. <i>Tokenization</i> 3. <i>Stopword Removal</i> 4. <i>Stemming</i>	<i>Naïve Bayes</i>	Bandung, Indonesia	<i>Naïve Bayes JNE</i> : 76% <i>Naïve Bayes JNT</i> : 75%
4	Muhammad Rezki, Desiana Nurkholifah, Muhammad Faisal,	Analisis review pengguna <i>Google Meet</i> dan <i>Zoom Cloud Meeting</i>	1. Ulasan	1. <i>Case Floding</i> 2. <i>Tokenizing</i> 3. <i>Filtering</i> 4. <i>Stemming</i>	<i>Naïve Bayes</i>	Jakarta Pusat, Indonesia	<i>Naïve Bayes + Feature Smote dan PSO Zoom Cloud</i> : 85,76% <i>Naïve Bayes</i>

	Priyono, Rachmad Suryadithia (2020)[9]	menggunakan Algoritma <i>Naïve Bayes</i>					+ <i>Feature Smote</i> dan <i>PSO Google Meet</i> : 79,33%
5	Mahardika Tania Nitami, Herny Februariyan ti (2022)[10]	Analisis sentimen ulasan ekspedisi JNT Express menggunakan algoritma <i>Naïve Bayes</i>	1. Ulasan	1. <i>Case Folding</i> 2. <i>Normalisasi</i> 3. <i>Tokenizing</i> 4. <i>Stopword Removal</i>	<i>Naïve Bayes</i>	Semarang, Indonesia	Presisi kelas Positif : 70% Presisi Kelas Negatif 95% Error rate : 13%
6	Caihua Liu, Muneera Ban , Didar Zowghi, Matthew Kearney (2021) [11]	<i>Analysing user reviews of inquiry-based learning apps in science education</i>	<i>User Reviews</i>	-	<i>Mobile Learning</i>	Guangdong Province,China	Simulasi, 87% dari total rating EarthViewer , dan 90% dari total rating Solar Walk.
7	Ziedhan Alifio Dieksona, Muhammad Rivyan Bagas Prakosoa, Muhammad Savio Qalby han Alifio Dieksona, Muhammad Rivyan Bagas Prakosoa, Muhammad	Sentiment analysis for customer review: Case study of Traveloka	1. Review 2. Rating 3. Comment	-	Logistic Regression, SVM, Naïve Bayes	Jakarta Pusat, Indonesia	Metode Naive Bayes sebesar 82,91%.

	Savio (2020)[12]						
8	Zoltan Geler, Miloš Savić, Brankica Bratić, Vladimir Kurbalija, Mirjana Ivanović & Weihui Dai (2021)[13]	Sentiment prediction based on analysis of customers assessments in food serving businesses	1. Meal quality 2. Service quality 3. Interior quality	-	evaluated using mean absolute error (MAE) and root mean square error (RMSE)	Republic of China	Kesimpulan nya, kata kunci diambil dari kata kunci pelanggan
9	Ting-Peng Liang, Xin Li, Chin-Tsung Yang & Mengyue Wang (2015)[14]	What in Consumer Reviews Affects the Sales of Mobile Apps: A Multifacet Sentiment Analysis Approach	1. Textual consumer reviews	-	MFSAA	Turkish	ulasan cocok untuk prediksi tiga kriteria kepuasan yang diamati: rasa makanan, layanan, dan lingkungan

10	Lei Zhang, Xuening Chu & Deyi Xue(2018)[15]	Identification of the to-be-improved product features based on online reviews for product redesign	1. Customer online reviews	-	Corpus-based measures	Canada	Keuntungan dari pendekatan baru ini dirangkum sebagai berikut: (1) Ulasan online pelanggan merupakan sumber informasi yang baik untuk analisis kebutuhan dan preferensi pelanggan. (2) Metode berbasis online-review yang diusulkan efektif untuk mengidentifikasi target fitur produk yang akan ditingkatkan dari seluruh kandidat fitur
----	---	--	----------------------------	---	-----------------------	--------	---

## 1.2 Analisis Review

*Review* adalah istilah yang berasal dari bahasa Inggris yang memiliki arti tinjauan atau ulasan. Secara pengertiannya, Review adalah kegiatan yang dilakukan seseorang untuk memberikan ulasan terhadap sesuatu hal yang sudah pernah mereka coba sebelumnya.

Biasanya, review akan membahas mengenai kelebihan dan kekurangan dari suatu produk atau jasa. Merivew pasti akan berpegang pada prinsip analisis dan fakta bukan hanya opini sembarangan. Review berupa ringkasan singkat, tetapi hal ini bisa sangat berguna untuk berbagai bidang, baik pemasaran hingga seni.[16]

### 1.3 Jasa Layanan

Dalam ilmu ekonomi, **jasa** atau **layanan** adalah aktivitas ekonomi yang melibatkan sejumlah interaksi dengan konsumen atau dengan barang-barang milik, tetapi tidak menghasilkan transfer kepemilikan. Menurut ahli Phillip Kotler Jasa adalah setiap tindakan atau unjuk kerja yang ditawarkan oleh salah satu pihak ke pihak lain yang secara prinsip intangibel dan tidak menyebabkan perpindahan kepemilikan apapun. Produksinya bisa terkait dan bisa juga tidak terikat pada suatu produk fisik.

Seringkali dikatakan bahwa jasa memiliki karakteristik unik yang membedakannya dari barang atau produk-produk manufaktur. Empat karakteristik yang paling sering dijumpai dalam jasa dan pembeda dari barang pada umumnya adalah (Payne, 2001:9):

- a. **Tidak berwujud** Jasa bersifat abstrak dan tidak berwujud, berarti jasa tidak dapat dilihat, dirasakan, dicicipi atau disentuh seperti yang dapat dirasakan dari suatu barang.
- b. **Heterogenitas** Jasa merupakan variabel non – standar dan sangat bervariasi. Artinya, karena jasa itu berupa suatu unjuk kerja, maka tidak ada hasil jasa yang sama walaupun dikerjakan oleh satu orang. Hal ini dikarenakan oleh interaksi manusia (karyawan dan konsumen) dengan segala perbedaan harapan dan persepsi yang menyertai interaksi tersebut.
- c. **Tidak dapat dipisahkan** Jasa umumnya dihasilkan dan dikonsumsi pada saat yang bersamaan, dengan partisipasi konsumen dalam proses tersebut. Berarti, konsumen harus berada di tempat jasa yang dimintanya, sehingga konsumen melihat dan bahkan ikut ambil bagian dalam proses produksi tersebut.
- d. **Tidak tahan lama** Jasa tidak mungkin disimpan dalam persediaan. Artinya, jasa tidak bisa disimpan, dijual kembali kepada orang lain, atau dikembalikan kepada produsen jasa di mana ia membeli jasa.

Sedangkan karakteristik jasa menurut Kotler dan Keller adalah sebagai berikut:

- a. **Tidak berwujud** Berbeda dari produk fisik, jasa tidak dapat dilihat, dirasa, diraba, didengar, atau dicium sebelum dibeli.
- b. **Tidak terpisahkan** Biasanya jasa dihasilkan dan dikonsumsi secara bersamaan. Hal ini tidak berlaku bagi barang-barang fisik yang diproduksi, disimpan sebagai persediaan, didistribusikan melalui banyak penjual, dan dikonsumsi kemudian. Jika seseorang

memberikan jasa tersebut, penyediannya adalah bagian dari jasa itu. Karena klien tersebut juga hadir pada saat jasa itu dihasilkan, interaksi penyediaan klien merupakan ciri khusus pemasaran jasa.

- c. **Bervariasi** Karena bergantung pada siapa memberikannya serta kapan dan dimana diberikan, jasa sangat bervariasi.
- d. **Tidak tahan lama** Jasa tidak dapat disimpan. Sifat jasa yang mudah rusak tersebut tidak akan menjadi masalah apabila permintaan tetap berjalan lancar. [17]

#### 1.4 Google Playstore

*Google Playstore*, sebelumnya **Android Market**, adalah layanan distribusi digital yang dioperasikan dan dikembangkan oleh Google. Ini berfungsi sebagai toko aplikasi resmi untuk sistem operasi Android, yang memungkinkan pengguna untuk menelusuri dan mengunduh aplikasi yang dikembangkan dengan Android software development kit (SDK) dan diterbitkan melalui Google. *Google Playstore* juga berfungsi sebagai toko media digital, yang menawarkan program musik, buku, film, dan televisi. Ini sebelumnya menawarkan perangkat keras Google untuk pembelian sampai diperkenalkannya pengecer perangkat keras online yang terpisah, Google Store, pada 11 Maret 2015, dan juga menawarkan publikasi berita dan majalah sebelum perbaikan Google News pada 15 Mei 2018.

Aplikasi tersedia melalui *Google Playstore* baik gratis atau dengan biaya. Mereka dapat diunduh langsung pada perangkat Android melalui aplikasi seluler Play Store atau dengan menyebarkan aplikasi ke perangkat dari situs web *Google Playstore*. Aplikasi yang mengeksploitasi kemampuan perangkat keras dari suatu perangkat dapat ditargetkan untuk pengguna perangkat dengan komponen perangkat keras tertentu, seperti sensor gerak (untuk game yang bergantung pada gerakan) atau kamera yang menghadap ke depan (untuk panggilan video online). *Google Playstore* store memiliki lebih dari 82 miliar unduhan aplikasi pada 2016 dan telah mencapai lebih dari 3,5 juta aplikasi yang diterbitkan pada 2017.<sup>[1]</sup> Ini telah menjadi subjek berbagai masalah mengenai keamanan, di mana perangkat lunak berbahaya telah disetujui dan diunggah ke toko dan diunduh oleh pengguna, dengan tingkat keparahan yang berbeda-beda.

*Google Playstore* diluncurkan pada 6 Maret 2012, menyatukan **Android Market**, **Google Music**, dan **Google eBookstore** di bawah satu merek, menandai perubahan dalam strategi distribusi digital Google. Layanan yang termasuk dalam Google Play adalah Google

Play Books, Google Play Games, Google Play Film & TV dan Google Play Music. Setelah merek-ulang, Google secara bertahap memperluas dukungan geografis untuk setiap layanan.[18]

### 1.5 Naïve Bayes

Naïve Bayes Naive Bayes merupakan sebuah metode dengan pendekatan 2 penafsiran yang sedikit berbeda. Menurut penjelasan Bayes, model ini menjelaskan seberapa jauh tingkat derajat kepercayaan pada suatu subjektif harus dapat merubah secara rasional ketika adanya suatu petunjuk atau tujuan baru. Sedangkan penjelasan frekuentis, model ini mengartikan bahwa representasi sebagai invers probabilitas melalui dua scenario kejadian . [19]

Naive merupakan dasar dari statistika Bayes dan dapat diterapkan dalam banyak bidang seperti sains, rekayasa, ilmu ekonomi, teori permainan kedokteran, hukum, dan lain sebagainya.

1. Menentukan Evidence Tunggal (E) dan hipotesa tunggal (H) seperti pada persamaan 1

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \quad (1)$$

Penjelasan:

$P(H|.E)$ , Kemungkinan hipotesaH terjadi jika evidence E terpenuhi

$P(E|.H)$  , Kemungkinan munculnya evidence E jika hipotesis H terjadi

$P(H.)$ , KemungkinanhipotesaH tanpa melihat evidence apapun

$P(E.)$ , Kemungkinan evidence E tanpa melihatapapun

2. Mencari Evidence tunggal (E) dan hipotesa ganda ( $H_1, H_2, \dots, H_n$ ) seperti pada persamaan 2

$$P(H_i|E) = \frac{P(E|H_i) \times P(H_i)}{\sum_{k=1}^n P(E|H_k) \times P(H_k)} \quad (2)$$

Penjelasan

$P(H_i|E)$ , Kemungkinan hipotesaHi benar terjadi jika diberikan evidence E

$P(E_i|.H)$ , Kemungkinan munculnya evidence E jika diketahui hipotesaHi Benar

$P(E_i|.H)$ , Kemungkinan Hi benar terjadi Hi

N, jumlah kemungkinan yang akan terjadi



3. Mencari Evidence ganda dan tunggal ganda seperti pada persamaan 3

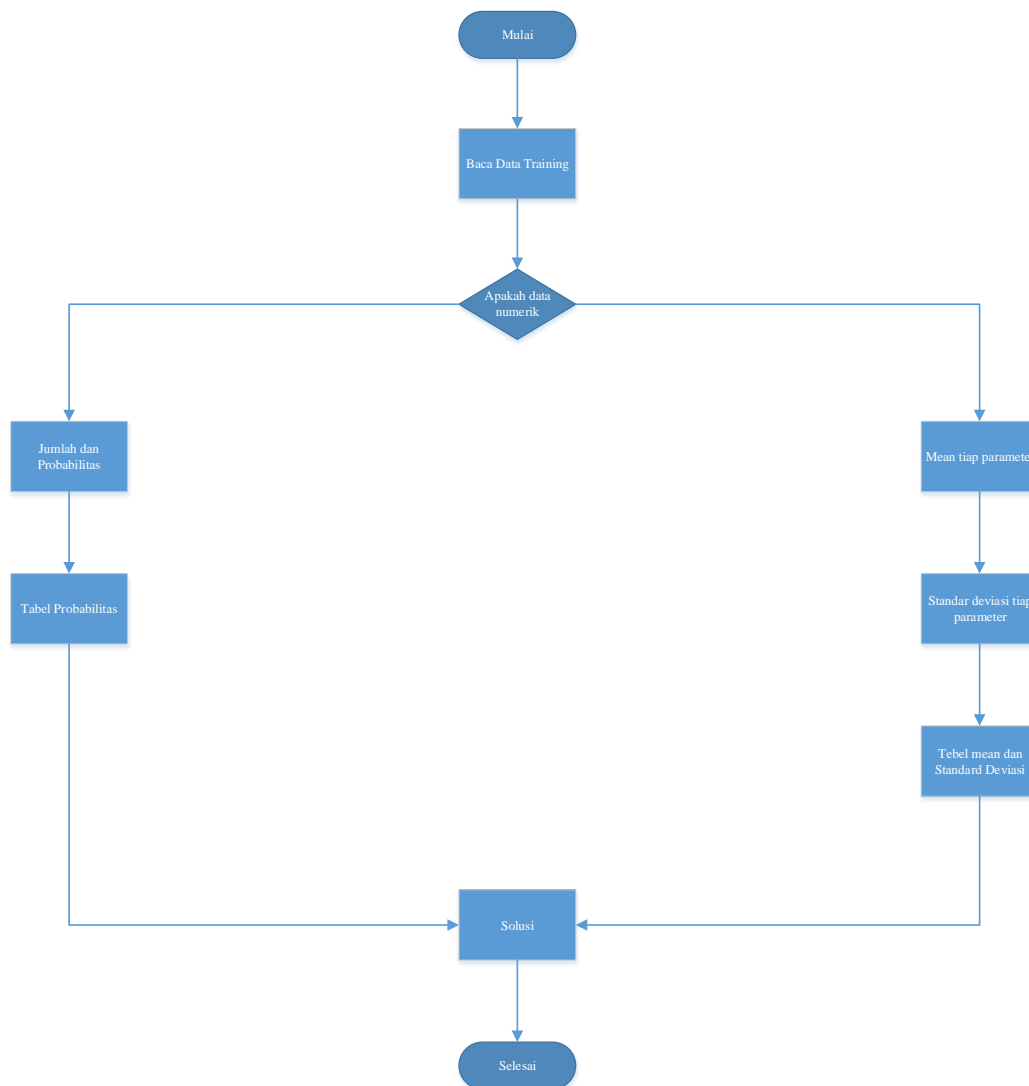
$$P(H_i|E_1 E_2 \dots E_m) = \frac{P(E_1|H_i) \times P(E_2|H_i) \times \dots \times P(E_m|H_i) \times P(H_i)}{\sum_{k=1}^n P(E_1 E_2 \dots E_m|H_k) \times P(H_k)} \quad (3)$$

Namun pada pengaplikasian yang dilakukan tidak akan mungkin terjadi dikarenakan harus memiliki pengetahuan semua kemungkinan bersyarat melalui semua kemungkinan yang dikombinasikan, maka rumus perhitungan tersebut diubah menjadi seperti persamaan 4.

$$P(H_i|E_1 E_2 \dots E_m) = \frac{P(E_1|H_i) \times P(E_2|H_i) \times \dots \times P(E_m|H_i) \times P(H_i)}{\sum_{k=1}^n P(E_1|H_k) \times P(E_2|H_k) \times \dots \times P(E_m|H_k) \times P(H_k)} \quad (4)$$

Langkah-langkah algoritma *Naïve Bayes* menurut Suntoro (2019), sebagai berikut: [20]

1. Siapkan dataset.
2. Hitung jumlah kelas pada data training.
3. Hitung jumlah kasus yang sama dengan kelas yang sama.
4. Kalikan semua hasil sesuai dengan data testing yang akan dicari kelasnya.
5. Bandingkan hasil perkelas, nilai tertinggi ditetapkan sebagai kelas baru



**Gambar 2.1** Flowchart *Naïve Bayes*

### 1.6 *K-Nearest Neighbor*

*K-Nearest Neighbor* (*K-NN*) merupakan salah satu algoritma yang digunakan dalam masalah pengklasifikasian. Prinsip kerja *K-NN* ialah mencari jarak terdekat antar data yang akan dievaluasi dengan tetangga terdekat dalam data pelatihan. Algoritma *K-Nearest Neighbor* (*K-NN*) adalah salah satu algoritma paling sederhana untuk memecahkan masalah klasifikasi dan sering menghasilkan hasil yang kompetitif dan signifikan. Untuk menghitung jarak menggunakan jarak Euclidean. [21]

Algoritma *K-Nearest Neighbor* sangat cocok untuk memperkirakan peluang apa yang akan terjadi selanjutnya menggunakan kasus-kasus yang sudah ada. Dengan metode *K-Nearest Neighbor* maka akan sangat cocok dalam pengambilan keputusan berdasarkan kemiripan dengan kasus-kasus terdahulunya.

Tahapan Metode K-Nearest Neighbor, Langkah yang digunakan dalam metode *K-Nearest Neighbor*:

- a. Tentukan parameter *K* (jumlah tetangga paling dekat).
- b. Hitung kuadrat jarak euclid masing – masing objek terhadap data sampel yang diberikan.
- c. Urutkan objek – objek kedalam kelompok yang memiliki jarak terkecil.
- d. Kumpulkan kategori *Y* (Klasifikasi nearest neighbor).
- e. Dengan kategori nearest neighbor yang paling banyak, maka dapat diprediksikan nilai query instance yang telah dihitung.

*K-Nearest Neighbor* dirumuskan sebagai berikut:

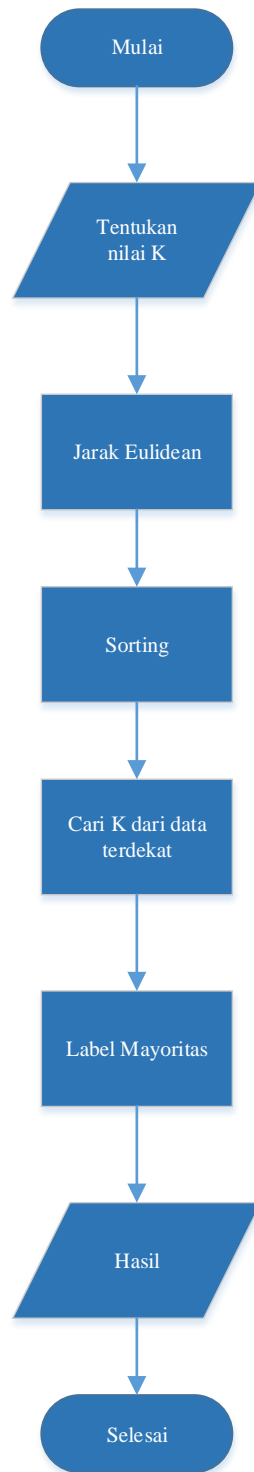
$$\cos(\theta_{ij}) = \frac{\sum_k (d_{ik} \cdot d_{jk})}{\sqrt{\sum_k d_{ik}^2} \sqrt{\sum_k d_{jk}^2}} \dots \dots \dots x(5)$$

Keterangan:

- $\cos(\theta_{ij})$  : *Similiarity K-Nearest Neighbor*
- $d_{ik}$  : bobot Data dokumen *Testing*
- $d_{jk}$  : bobot Data dokumen *Training*
- $d^2$  : Panjang Vektor dokume

Bobot Label pada Dataset

- 1 = Positif
- 2 = Netral
- 3 = Negatif



**Gambar 2.2** Flowchart *K-Nearest Neighbor*

## 1.7 Particle Swarm Optimization (PSO)

*Particle Swarm Optimization (PSO)* merupakan algoritma yang banyak digunakan dan dikembangkan dengan cepat karena penerapannya yang mudah dan hanya sedikit partikel yang perlu disesuaikan dan dikembangkan berdasarkan kecerdasan berkelompok (swarm intelligence) serta didasarkan pada studi perilaku pergerakan kawanan burung dan ikan (Risawati et al., 2020). [22]

*Particle Swarm Optimization* dirumuskan sebagai berikut (Taufiq, Studi, Informatika, Bangsa, & Sentimen, 2020).

$$V_i(t) = V_i(t-1) + c_1r_1[Xpbest_i - X_i(t)] + c_2r_2[Xgbest - X_i(t)]$$

$$X_i(t) = X_i(t-1) + V_i(t)$$

Keterangan Rumus :

$V_i(t)$  = Kecepatan partikel  $i$  saat iterasi  $t$

$X_i(t)$  = Posisi partikel  $i$  saat iterasi  $t$

$c_1$  dan  $c_2$  = Learning rates untuk kemampuan individu dan pengaruh sosial

$r_1$  dan  $r_2$  = Bilangan random yang berdistribusi uniformal dalam interval 0 dan 1

$Xpbest_i$  = Posisi terbaik partikel  $i$

$Xgbest$  = Posisi terbaik global