

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Sebelumnya

Penulis menggunakan 20 artikel penelitian sebagai referensi dalam penelitian ini, yang terbagi menjadi dua jenis topik penelitian, yaitu:

1. Penelitian mengenai kinerja algoritma SVM

Artikel yang digunakan sebagai referensi pada topik ini adalah penelitian yang membahas mengenai kinerja algoritma SVM pada masalah penelitian yang berbeda.

Penelitian pertama adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma C4.5, *random forest*, dan *naïve bayes* pada masalah klasifikasi nilai peserta Jogja International Scout Camp 2020 (JISC2020). Hasil dari penelitian ini adalah algoritma SVM memiliki akurasi yang paling tinggi dibandingkan algoritma lainnya (Azhari et al., 2021).

Penelitian kedua adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma *decision tree* dan K-NN pada masalah penentuan jurusan di SMK Siti Banun. Hasil dari penelitian ini adalah algoritma SVM memiliki akurasi yang paling tinggi dibandingkan algoritma lainnya (Sinaga et al., 2021).

Penelitian ketiga adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma *naïve bayes* pada masalah klasifikasi ras

kucing. Hasil dari penelitian ini adalah algoritma SVM memiliki akurasi yang lebih baik dibandingkan algoritma *naïve bayes* (Kusuma et al., 2022).

Penelitian keempat adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma *naïve bayes* dengan seleksi fitur menggunakan *particle swarm optimization* (PSO) pada masalah analisis sentimen *e-wallet review*. Hasil dari penelitian ini adalah algoritma *naïve bayes* memiliki akurasi yang lebih baik dibandingkan algoritma SVM (Putri et al., 2020).

Penelitian kelima adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma *logistic regression* dan K-NN pada masalah klasifikasi bunga Iris. Hasil dari penelitian ini adalah algoritma SVM memiliki akurasi yang paling tinggi dibandingkan algoritma lainnya (Pinto et al., 2018).

Penelitian keenam adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma *decision tree*, *naïve bayes*, dan *logistic regression* pada masalah klasifikasi citra biji kopi. Hasil dari penelitian ini adalah algoritma SVM dan *decision tree* memiliki akurasi yang paling tinggi dibandingkan algoritma lainnya (Waliyansyah & Hasbulla, 2021).

Penelitian ketujuh adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma *decision tree*, *naïve bayes*, dan K-NN pada masalah klasifikasi beberapa dataset UCI. Hasil dari penelitian ini adalah algoritma *decision tree* memiliki akurasi yang paling tinggi dibandingkan algoritma lainnya (Untoro et al., 2020).

Penelitian kedelapan adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma ANFIS pada masalah klasifikasi parasit

malaria. Hasil dari penelitian ini adalah algoritma SVM memiliki akurasi yang lebih baik dibandingkan algoritma ANFIS (Hartati et al., 2019).

Penelitian kesembilan adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma *backpropagation* pada masalah klasifikasi jenis tubuh wanita muda. Hasil dari penelitian ini adalah algoritma SVM memiliki akurasi yang lebih baik dibandingkan algoritma *backpropagation* (Zhang, 2022).

Penelitian kesepuluh adalah penelitian mengenai kinerja algoritma SVM dibandingkan dengan algoritma *random forest* pada masalah pengenalan emosi berdasarkan suara. Hasil dari penelitian ini adalah algoritma *random forest* memiliki akurasi yang lebih baik dibandingkan algoritma SVM (Zhang, 2022).

Rangkuman dari hasil penelitian-penelitian ini ditunjukkan pada Tabel II.1.

Tabel II.1 Penelitian Kinerja Algoritma SVM

Penulis	Judul	Topik	Hasil
(Azhari et al., 2021)	Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Bayes	Komparasi C4.5, <i>random forest</i> , SVM dan <i>naïve bayes</i>	Akurasi C4.5 = 89.67%
			Akurasi <i>random forest</i> = 83.33%
			Akurasi SVM = 95%
			Akurasi <i>naïve bayes</i> = 86.67%
(Sinaga et al., 2021)	Komparasi Metode Decision Tree, KNN, dan SVM Untuk Menentukan Jurusan Di SMK	Komparasi <i>decision tree</i> , K-NN, dan SVM	Akurasi <i>decision tree</i> = 78%
			Akurasi K-NN = 84%
			Akurasi SVM = 89%

Penulis	Judul	Topik	Hasil
(Kusuma et al., 2022)	Komparasi Algoritma Support Vector Machine Dan Naive Bayes Pada Klasifikasi Ras Kucing	Komparasi SVM dan <i>naïve bayes</i>	Akurasi SVM = 88.4%
			Akurasi <i>naïve bayes</i> = 79.5%
(Putri et al., 2020)	Comparison of Naive Bayes Algorithm and Support Vector Machine using PSO Feature Selection for Sentiment Analysis on E-Wallet Review	Komparasi SVM dan <i>naïve bayes</i>	Akurasi SVM = 91.3%
			Akurasi <i>naïve bayes</i> = 93.1%
(Pinto et al., 2018)	Iris Flower Species Identification Using Machine Learning Approach	Komparasi SVM, <i>logistic regression</i> , dan K-NN	Akurasi SVM = 96%
			Akurasi <i>logistic regression</i> = 91%
			Akurasi K-NN = 93%
(Waliyansyah & Hasbulla, 2021)	Comparison of Tree Method, Support Vector Machine, Naïve Bayes, and Logistic Regression on Coffee Bean Image	Komparasi kinerja <i>decision tree</i> , SVM, <i>naïve bayes</i> , dan <i>logistic regression</i>	Akurasi <i>decision tree</i> = 98.3%
			Akurasi SVM = 98.3%
			Akurasi <i>naïve bayes</i> = 98.1%
			Akurasi <i>logistic regression</i> = 96.6%
(Untoro et al., 2020)	Evaluation of Decision Tree, K-NN, Naive Bayes and SVM with MWMOTE on UCI Dataset	Komparasi kinerja <i>decision tree</i> , <i>naïve bayes</i> , K-NN, dan SVM	Akurasi <i>decision tree</i> = 96.3%
			Akurasi <i>naïve bayes</i> = 78.74%
			Akurasi K-NN = 92.95%
			Akurasi SVM = 82%
(Hartati et al., 2019)	Performance of SVM and ANFIS for classification of malaria parasite and its life-cycle-stages in blood smear	Komparasi kinerja SVM dan ANFIS	Akurasi SVM = 99.2%
			Akurasi ANFIS = 88.5%
(Zhang, 2022)	Comparing SVM with BP to Identify the Young Females' Body Type	Komparasi kinerja SVM dan <i>backpropagation</i>	Akurasi SVM = 99%
			Akurasi <i>backpropagation</i> = 90.17%

Penulis	Judul	Topik	Hasil
(Wincy Pon Annal et al., 2022)	Emotion Recognition from Speech using SVM and Random Forest Classifier	Komparasi kinerja SVM dan <i>random forest</i>	Akurasi SVM = 76%
			Akurasi <i>random forest</i> = 85%

2. Penelitian mengenai kinerja fungsi kernel SVM

Artikel yang digunakan sebagai referensi pada topik ini adalah penelitian yang membahas mengenai kinerja fungsi kernel RBF, *polynomial*, dan *linear* pada masalah penelitian yang berbeda.

Penelitian pertama adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah pendeteksian *online spam review*. Hasil dari penelitian ini adalah fungsi kernel RBF memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (N & V, 2021).

Penelitian kedua adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah pendeteksian intrusi pada *cloud environment*. Hasil dari penelitian ini adalah fungsi kernel RBF memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (Nirmalajyothi et al., 2020).

Penelitian ketiga adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah pendeteksian *scintillation* yang disebabkan oleh ketidakteraturan kerapatan elektron dalam plasma ionosfer. Hasil dari penelitian ini adalah fungsi kernel *linear* memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (Savas & Dovic, 2019).

Penelitian keempat adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah analisis faktor resiko hipertensi. Hasil dari penelitian ini adalah

fungsi kernel *linear* memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (Azies & Trishnanti, 2019).

Penelitian kelima adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah klasifikasi Indeks Pembangunan Manusia (HDI). Hasil dari penelitian ini adalah fungsi kernel RBF memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (Azies et al., 2019).

Penelitian keenam adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah klasifikasi persalinan. Hasil dari penelitian ini adalah fungsi kernel *linear* memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (Intan, 2019).

Penelitian ketujuh adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah identifikasi serangan DDoS. Hasil dari penelitian ini adalah fungsi kernel RBF memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (Perwira & Prapcoyo, 2021).

Penelitian kedelapan adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah pengenalan wajah (dataset ORL), pengenalan *scene* (dataset MIT), deteksi tumor (dataset *Brain Tumor*), dan pengenalan digit (dataset MNIST). Hasil dari penelitian ini adalah fungsi kernel RBF memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (Kancherla et al., 2019).

Penelitian kesembilan adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah klasifikasi Indikator Kinerja Utama (KPI) dari jaringan selular di Nigeria. Hasil dari penelitian ini adalah fungsi kernel *linear* memiliki RMSE yang paling rendah dibandingkan fungsi kernel lainnya (Philip et al., 2019).

Penelitian kesepuluh adalah penelitian mengenai kinerja fungsi kernel SVM pada masalah pendeteksian tumor otak menggunakan citra MRI. Hasil dari penelitian ini adalah fungsi kernel *polynomial* memiliki akurasi yang paling tinggi dibandingkan fungsi kernel lainnya (Naser et al., 2021).

Rangkuman dari hasil penelitian-penelitian ini ditunjukkan pada Tabel II.2.

Tabel II.2 Penelitian Kinerja Fungsi Kernel SVM

Penulis	Judul	Topik	Hasil
(N & V, 2021)	Ensemble of Multiple Kernel SVM Classifiers for Detection of Online Spam Reviews	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	Akurasi RBF = 89.47%
			Akurasi <i>polynomial</i> = 82.48%
			Akurasi <i>linear</i> = 84.93%
(Nirmalajyothi et al., 2020)	Performance of Various SVM Kernels for Intrusion Detection of Cloud Environment	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	Akurasi RBF = 88.69%
			Akurasi <i>polynomial</i> = 50.99%
			Akurasi <i>linear</i> = 80.06%
(Savas & Dosis, 2019)	The impact of different kernel functions on the performance of scintillation detection based on support vector machines	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	Akurasi RBF = 86.13%
			Akurasi <i>polynomial</i> = 86.15%
			Akurasi <i>linear</i> = 90.98%
(Azies & Trishnanti, 2019)	Perbandingan Support Vector Machine Method (SVM) Linear, Radial Basis Function (RBF) dan Kernel Polinomial Dalam Analisis Faktor Risiko Hipertensi	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	Akurasi RBF = 94.41%
			Akurasi <i>polynomial</i> = 98.88%
			Akurasi <i>linear</i> = 100%
Penulis	Judul	Topik	Hasil
(Azies et al., 2019)	Comparison of Kernel Support Vector	Kinerja kernel RBF,	Akurasi RBF = 98.1%

	Machine (SVM) in Classification of Human Development Index (HDI)	<i>polynomial</i> , dan <i>linear</i>	Akurasi <i>polynomial</i> = 82.5%
			Akurasi <i>linear</i> = 95.1%
(Intan, 2019)	Comparison of Kernel Function on Support Vector Machine in Classification of Childbirth	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	Akurasi RBF = 82%
			Akurasi <i>polynomial</i> = 66%
			Akurasi <i>linear</i> = 83%
(Perwira & Prapcoyo, 2021)	Software Defined Network : The Comparison of SVM kernel on DDoS Detection	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	Akurasi RBF = 99.42%
			Akurasi <i>polynomial</i> = 98.89%
			Akurasi <i>linear</i> = 99.02%
(Kancherla et al., 2019)	Effect of different kernels on the performance of an SVM based classification	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	Akurasi RBF = 92.62%
			Akurasi <i>polynomial</i> = 84.78%
			Akurasi <i>linear</i> = 89.85%
(Philip et al., 2019)	Comparative Analysis and Investigations of Various SVM Kernels Using Cellular Network KPI Data	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	RMSE RBF = 1.6351
			RMSE <i>polynomial</i> = 2.3886
			RMSE <i>linear</i> = 0.0019
(Naser et al., 2021)	SVM Kernels comparison for brain tumor diagnosis using MRI	Komparasi kinerja kernel RBF, <i>polynomial</i> , dan <i>linear</i>	Akurasi RBF = 76%
			Akurasi <i>polynomial</i> = 79%
			Akurasi <i>linear</i> = 75.5%

3. Penelitian mengenai parameter fungsi kernel SVM

Artikel yang digunakan sebagai referensi pada topik ini adalah penelitian yang membahas mengenai penentuan parameter d dan γ fungsi kernel RBF, *polynomial*, dan *linear* pada masalah penelitian yang berbeda.

Penelitian pertama adalah penelitian mengenai parameter optimal pada fungsi kernel RBF dan *polynomial* dengan variasi nilai parameter γ dan d yang berbeda. Hasil dari penelitian ini adalah parameter optimal bagi fungsi kernel RBF adalah $\gamma = 0.0454$, dan parameter optimal bagi fungsi kernel *polynomial* adalah $\gamma = 0.0454$, $d = 3$ (Al-Mejibli et al., 2020).

Penelitian kedua adalah penelitian mengenai parameter optimal pada fungsi kernel RBF, *polynomial* dan *linear* dengan variasi nilai parameter C dan γ yang berbeda. Hasil dari penelitian ini adalah parameter optimal bagi fungsi kernel RBF adalah $C = 16$, $\gamma = 0.25$, parameter optimal bagi fungsi kernel *polynomial* adalah $C = 4$, $\gamma = 0.25$, dan parameter optimal bagi fungsi kernel *linear* adalah $C = 16$ (Asrol et al., 2021).

Penelitian ketiga adalah penelitian mengenai parameter optimal pada fungsi kernel RBF dan *linear* dengan variasi nilai parameter C dan γ yang berbeda. Hasil dari penelitian ini adalah parameter optimal bagi fungsi kernel RBF adalah $C = 75$, $\gamma = 0.01$ dan parameter optimal bagi fungsi kernel *linear* adalah $C = -1$ (Pannakkong et al., 2022).

Penelitian keempat adalah penelitian mengenai parameter optimal pada fungsi kernel RBF, *polynomial* dan *linear* dengan variasi nilai parameter C , γ , dan d yang berbeda. Hasil dari penelitian ini adalah parameter optimal bagi fungsi kernel RBF adalah $C = 1$, $\gamma = 0.14286$, parameter optimal bagi fungsi kernel *polynomial* adalah $C = 1$, $\gamma = 0.14286$, $d = 3$, dan parameter optimal bagi fungsi kernel *linear* adalah $C = 1$ (Johra, 2018).

Penelitian kelima adalah penelitian mengenai parameter optimal pada fungsi kernel RBF, *polynomial* dan *linear* dengan variasi nilai parameter C , γ , dan d yang berbeda. Hasil dari penelitian ini adalah parameter optimal bagi fungsi kernel RBF adalah $C = 100$, $\gamma = 0.03$, parameter optimal bagi fungsi kernel *polynomial* adalah $C = 700$, $\gamma = 1$, $d = 1$, dan parameter optimal bagi fungsi kernel *linear* adalah $C = 7$ (Alida & Mustikasari, 2020).

Rangkuman dari hasil penelitian-penelitian ini ditunjukkan pada Tabel II.3.

Tabel II.3 Penelitian Parameter Fungsi Kernel SVM

Penulis	Judul	Fungsi Kernel	Parameter Optimal
(Al-Mejibli et al., 2020)	The effect of gamma value on support vector machine performance with different kernels	RBF	$\gamma = 0.0454$
		<i>Polynomial</i>	$\gamma = 0.0454$ $d = 3$
(Asrol et al., 2021)	Support Vector Machine with K-fold Validation to Improve the Industry's Sustainability Performance Classification	RBF	$C = 16$ $\gamma = 0.25$
		<i>Polynomial</i>	$C = 4$ $\gamma = 0.25$
		<i>Linear</i>	$C = 16$
(Pannakkong et al., 2022)	Hyperparameter Tuning of Machine Learning Algorithms Using Response Surface Methodology: A Case Study of ANN, SVM, and DBN	RBF	$C = 75$ $\gamma = 0.001$
		<i>Linear</i>	$C = -1$
(Johra, 2018)	Penerapan Non-Linier Support Vector Machine pada Penggunaan Alat Kontrasepsi di Provinsi Maluku Utara	RBF	$C = 1$ $\gamma = 0.14286$
		<i>Polynomial</i>	$C = 1$ $\gamma = 0.14286$ $d = 3$
		<i>Linear</i>	$C = 1$
(Alida & Mustikasari, 2020)	Rupiah Exchange Prediction of US Dollar Using Linear, Polynomial, and Radial Basis Function Kernel in Support Vector Regression	RBF	$C = 100$ $\gamma = 0.03$
		<i>Polynomial</i>	$C = 700$ $\gamma = 1$ $d = 1$
		<i>Linear</i>	$C = 7$

2.2 Penyuluh Agama

Islam sebagai agama dakwah, selalu mendorong pemeluknya untuk senantiasa aktif dalam berdakwah untuk mengajak manusia lain dapat menerima Islam sebagai pedoman hidupnya, serta melaksanakan amar ma'ruf dan *nahi munkar*. Penyuluh agama Islam merupakan seseorang yang berperanan penting sebagai pelaksana dakwah (da'i) yang memiliki tugas melakukan kegiatan bimbingan dan penyuluhan agama Islam, khususnya bagi komunitas umat muslim. Peran penting penyuluh agama Islam ini sangatlah strategis dalam hal meningkatkan kualitas kehidupan beragama maupun kesejahteraan umat Islam, sehingga penyuluh agama Islam dituntut untuk dapat menguraikan persoalan yang dialami oleh masyarakat untuk kemudian dicari solusinya dalam konteks dakwah (Ilham, 2018).

Peranan penyuluh agama Islam semakin hari semakin berat, seiring dengan penyimpangan-penyimpangan yang terjadi di masyarakat, sehingga harus ada suatu strategi untuk mempersiapkan para penyuluh agama Islam dalam melakukan dakwah dan tugas-tugasnya yang lain. Keberhasilan seorang penyuluh agama Islam dalam melaksanakan dakwah di dalam masyarakat sangat berkaitan dengan terciptanya suatu komunitas yang mampu membimbing dan menyuluh ajaran-ajaran Islam (Musliamin, 2019).

Kementrian Agama sebagai salah satu aparaturnya, memiliki tugas penting dalam membina para penyuluh agama Islam, dalam perannya untuk membangun iklim keagamaan yang kondusif bagi perkembangan masyarakat yang dinamis, progresif, toleran dan damai. Untuk mencapai keberhasilan kegiatan

penyuluhan agama Islam di dalam masyarakat, Kementrian Agama dituntut untuk menghasilkan penyuluh agama Islam yang fungsional dan memiliki tingkat keterampilan, kemampuan, dan kecakapan yang memadai sehingga proses kegiatan bimbingan dan penyuluhan agama dapat berjalan sistematis, berhasil guna, dan berdaya guna. Bidang Penamas di tingkat Kanwil Kemenag atau tingkat Kandepag ditunjuk sebagai pihak yang berkompeten langsung mengampu program penyuluhan, merencanakan program serta mengorganisir sumber daya penyuluh agama Islam (Fahrurrozi & Munir, 2021).

Secara garis besar, penyuluh agama Islam memiliki tiga fungsi yang berkaitan dengan tugas pokoknya, adapun fungsi tersebut adalah :

1. Fungsi informatif dan edukatif

Penyuluh agama Islam diharapkan dapat berfungsi sebagai dai dalam mendakwahkan ajaran agama Islam dengan sebaik-baiknya sesuai dengan sunnah Nabi dan tuntutan Alqur'an.

2. Fungsi konsultatif

Penyuluh agama Islam diharapkan dapat ikut berperan serta dalam memikirkan dan memecahkan problema serta masalah yang dihadapi umat pemeluk agama Islam, baik yang bersifat pribadi, keluarga, atau persoalan di dalam masyarakat umum.

3. Fungsi advokatif

Penyuluh agama Islam diharapkan dapat memiliki tanggung jawab secara moral dan sosial dalam menjalankan kegiatan bimbingan dan pembelaan terhadap umat binaannya, meliputi akidah, ibadah, dan akhlak.

2.3 *Machine Learning*

Machine learning merupakan sebuah bidang ilmu yang mengadopsi aplikasi komputer dan algoritma matematika dan melakukan pembelajaran terhadap data untuk menghasilkan sebuah keputusan prediksi di masa yang akan datang (Roihan et al., 2019). *Machine learning* memiliki kemampuan untuk melakukan atau memfasilitasi berbagai tugas yang mencakup pengolahan sejumlah besar data input untuk dilatih sehingga menghasilkan sebuah model. Dalam *machine learning*, terdapat beberapa kelompok tugas yang dapat diselesaikannya, seperti klasifikasi (*classification*), pengelompokan (*prediction*), dan prediksi (*prediction*) (Pannakkong et al., 2022).

Sebagai bagian dari *machine learning*, klasifikasi dapat digunakan sebagai alternatif pemecahan masalah dengan dasar bahwa dengan menggunakan klasifikasi, sekumpulan data yang belum diketahui pola-pola di dalamnya dapat dilatih sehingga menghasilkan sebuah keputusan prediksi yang dapat digunakan untuk data-data berikutnya (Sadikin et al., 2020).

Klasifikasi menggunakan teknik *supervised learning* dalam proses pembelajarannya, dengan cara menganalisis sekumpulan data latih dan memetakan hasilnya ke dalam sebuah fungsi yang dapat digunakan untuk memprediksi data-data berikutnya. Data latih inilah yang digunakan algoritma klasifikasi untuk menentukan batasan-batasan pada keputusan yang dihasilkan, berdasarkan hasil pengamatan yang diperoleh dari sekumpulan data tersebut (Ghosh et al., 2019).

Proses klasifikasi terdiri dari tiga tahapan: tahapan pengembangan model, dimana model dibuat untuk menyelesaikan masalah pengklasifikasian atribut atau

kelas; tahapan penerapan model untuk menentukan atribut atau kelas data pengujian yang belum diketahui; tahapan evaluasi model sebelumnya menggunakan parameter ukur tertentu untuk menentukan apakah model tersebut layak digunakan atau tidak (Untoro et al., 2020).

Dalam prosesnya, klasifikasi menggunakan sebuah *classifier* seperti model *decision tree*, pengklasifikasi berdasarkan *naive bayes* dan model *support vector machine* (Han, 2021).

2.4 *Support Vector Machine*

Support vector machine (SVM) merupakan algoritma pengenalan pola dan analisa data yang dapat bekerja pada masalah linier maupun non-linier dengan cara mencari sebuah *hyperplane* yang memisahkan data menjadi dua kelas dengan menggunakan persamaan II.1, dimana x_i merupakan data masukan, dan y_i merupakan output dari x_i , serta w merupakan parameter klasifikasi (Thenata, 2021).

$$\min \frac{1}{2} \|w\|^2, \text{subjected to } y_i(wx_i + b) \geq 1, i = 1, \dots, \lambda \dots\dots\dots \text{II.1}$$

SVM merupakan salah satu metode *supervised learning* yang dapat mengelola ruang fitur dalam ukuran besar dan mempunyai kemampuan generalisasi yang tinggi. Namun, algoritma ini relatif lebih kompleks untuk diimplementasikan dan membutuhkan waktu yang lama serta membutuhkan pemakaian memori yang cukup besar selama tahap pelatihan dan tahap klasifikasi (Risawati et al., 2020).

Secara umum, langkah-langkah dalam proses klasifikasi menggunakan algoritma SVM adalah sebagai berikut (Kurniawan et al., 2020):

4. Ubah dataset menjadi bentuk vektor masukan $x_i \in R^N$, dengan $i = 1, 2, \dots, N$
5. Berikan label $y_i \in \{+1, -1\}$ pada seluruh data, dengan label +1 dan -1 masing-masing mewakili kelas positif dan negatif.
6. Hitung fungsi keputusan yang menyatakan hubungan antara *input-output* dengan menggunakan persamaan (II.2), Di mana w menyatakan vektor bobot, b adalah parameter bias, f merupakan kelas prediksi:

$$f(x_i) = w^T \varphi(x_i) + b \dots\dots\dots \text{II.2}$$

7. Jika nilai $f(x_i) = 0$, maka data pada tersebut merupakan *support vector*
8. Untuk $i = 1, 2, \dots, N$, berlaku aturan pada persamaan (II.3):

$$w^T \varphi(x_i) + b \begin{cases} > 0, \text{ untuk } y_i = 1 \\ < 0, \text{ untuk } y_i = -1 \end{cases} \dots\dots\dots \text{II.3}$$

9. Persamaan (2) dan (3) dapat disederhanakan menjadi persamaan (II.4), yang merupakan fungsi keputusan SVM:

$$y_i(w^T \varphi(x_i) + b) \geq 1 \dots\dots\dots \text{II.4}$$

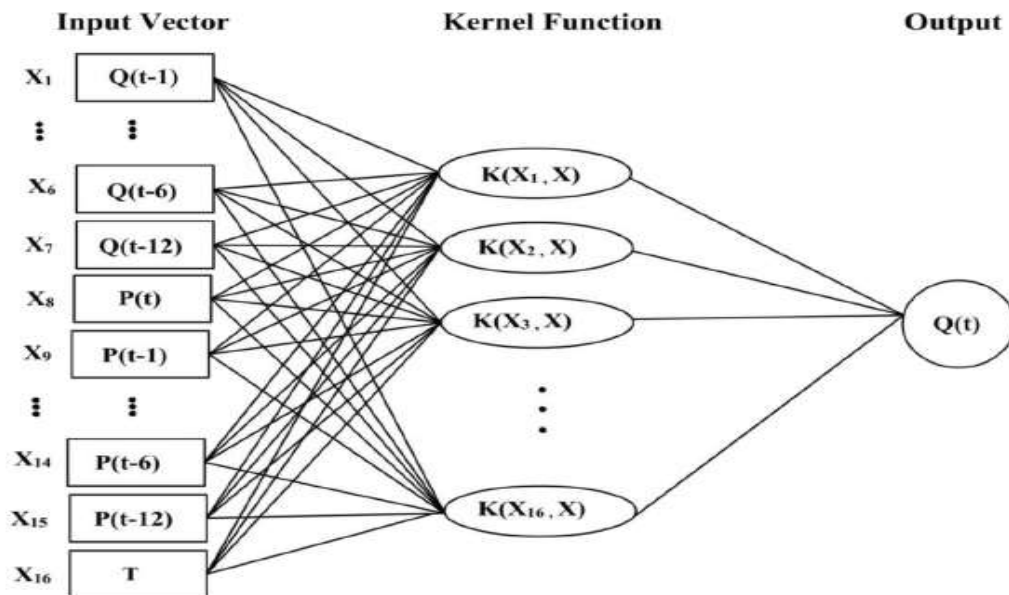
10. Diperoleh persamaan (II.5) yang merupakan persamaan *hyperplane* SVM:

$$w^T \varphi(x_i) + b = c, -1 < c < 1 \dots\dots\dots \text{II.5}$$

11. Hitung nilai optimal dari *margin hyperplane* dengan persamaan (II.1).

Untuk mengatasi masalah non-linier, algoritma SVM menggunakan fungsi kernel (*kernel trick*), yang mentransformasikan vector inputan ke dimensi yang lebih tinggi, sehingga pada proses perhitungannya nilai $\varphi(x_i)$ akan digantikan dengan nilai yang dihasilkan oleh transformasi menggunakan fungsi kernel yang digunakan. Dalam algoritma SVM, dikenal tiga fungsi kernel yang umum digunakan di dalam masalah klasifikasi, yaitu *gaussian radial basis function* (RBF), *polynomial*, dan *linear* (Azies & Trishnanti, 2019).

Secara umum, arsitektur dari SVM dapat dilihat pada Gambar II.1.



Gambar II.1 Arsitektur SVM