BAB II
TINJAUAN PUSTAKA

BAB II

TINJAUAN PUSTAKA

II.1. Penelitian Terkait

Pada penelitian yang dilakukan oleh (Ridho Lubis et al., n.d.) yang berjudul **Detection of HOG Features on Tuberculosis X-Ray Results Using SVM and KNN** menggunakan metode *Support Vector Machine* dan *K-Nearest Neighbors* menghasilkan Fitur HOG ini dapat membantu dalam deteksi gambar pada jenis Tuberkulosis berdasarkan sinar-X dalam deteksi HOG pada gambar sinar-X Tuberkulosis dengan KNN untuk gambar positif mencapai akurasi sebesar 77,95%, sementara untuk gambar negatif mencapai akurasi sebesar 78,65%. Sementara itu, hasil deteksi HOG pada gambar sinar-X Tuberkulosis menggunakan SVM untuk gambar positif mencapai akurasi sebesar 65,75%, sedangkan untuk gambar negatif mencapai akurasi sebesar 79,39%.

Pada penelitian yang dilakukan oleh (Balasubramanian et al, 2021). yang berjudul **Robust retinal blood vessel segmentation using convolutional neural network and support vector machine** menggunakan metode *convolutional neural network dan support vector machine* menghasilkan indeks kappa rata-rata sebesar 88,88%, spesifisitas sebesar 98,345%, akurasi sebesar 97,775%, dan sensitivitas sebesar 97,88%. Dibandingkan dengan karya yang sudah ada, teknik yang diusulkan hampir meningkatkan akurasi klasifikasi sebesar 2-4%

Pada penelitian yang dilakukan oleh (Padimi, 2022). yang berjudul **Optimizing Predictions of Brain Stroke Using Machine Learning**

menggunakan metode Machine Learning menghasilkan Pada regresi logistik, kami telah menguji pada 20% dari data pengujian. Setelah pengujian, model memiliki kinerja keseluruhan sebesar 77,7%. Dalam ROC, model mencapai AUC sebesar 0,86. Pada pohon keputusan, kami telah menguji pada 20% dari data pengujian. Setelah pengujian, model memiliki kinerja keseluruhan sebesar 77,7%. Dalam ROC, model mencapai AUC sebesar 0,913. Pada random forest, kami telah menguji pada 20% dari data pengujian. Setelah pengujian, model memiliki kinerja keseluruhan sebesar 94,4%. Dalam ROC, model mencapai AUC sebesar 0,992. Pada KNN, kami telah menguji pada 20% dari data pengujian. Setelah pengujian, model memiliki kinerja keseluruhan sebesar 93,3%. Dalam ROC, model mencapai AUC sebesar 0,971. Pada naïve bayes, kami telah menguji pada 20% dari data pengujian. Setelah pengujian, model memiliki kinerja keseluruhan sebesar 80,3%. Dalam ROC, model mencapai AUC sebesar 0,86.

Pada penelitian yang dilakukan oleh (Abd Aziz et al., 2021) yang berjudul **Detection of Collaterals from Cone-Beam CT Images in Stroke** menggunakan metode *Support Vector Machine*, *K-Nearest Neighbors* dan *one-beam computed tomography* menghasilkan akurasi rata-rata, SVM memberikan kinerja terbaik dalam mengklasifikasikan fitur GLCM dengan akurasi 99,92%. Random forest diidentifikasi sebagai klasifikasi terbaik untuk fitur momen invarian, mencapai akurasi 78,04%. Terakhir, KNN ditemukan lebih cocok untuk fitur Bentuk, Fitur GLCM memiliki akurasi tertinggi di antara tiga jenis fitur untuk modalitas gambar CBCT. Fitur GLCM mencapai akurasi antara 97,41% hingga 99,92%, sedangkan

fitur momen invarian dan Bentuk memiliki akurasi yang jauh lebih rendah, sekitar 77% dengan yang tertinggi adalah 79,33%.

Pada penelitian yang dilakukan oleh (Wanayumini et al., 2022)yang berjudul menggunakan metode *Multi Layer Perceptron (MLP) dan Support Vector Machine (SVM)* menghasilkan bahwa, dalam klasifikasi metode Multi Layer Perceptron(MLP) dengan fungsi aktivasi Logistic dan fungsi optimisasi memberikan nilai accuracy, precision dan recall terbaik dibandingkan *Support Vector Machine* yaitu sebesar 97.7%

Pada penelitian yang dilakukan oleh (Pokorny et al., 2023a) yang berjudul **On the Role of Training Data for SVM-Based Microwave Brain Stroke Detection and Classification** menggunakan metode Support Vector Machine (SVM) menghasilkan Data sintetis yang dihasilkan digunakan untuk menguji empat hipotesis berbeda, yang berkaitan dengan parameter yang sesuai untuk set data latihan, rentang frekuensi yang sesuai, jumlah titik frekuensi yang diperlukan, serta tingkat variasi subjek untuk mencapai akurasi klasifikasi SVM tertinggi. Hasil penelitian menunjukkan bahwa algoritma SVM mampu mendeteksi keberadaan stroke dan mengklasifikasikannya (yaitu iskemik atau hemoragik), bahkan ketika dilatih dengan data satu frekuensi serta kekurangannya antara lain Keterbatasan dalam representasi data nyata, Ketergantungan pada parameter sistem, Keterbatasan pada SVM serta Validitas eksternal

Pada penelitian yang dilakukan oleh (Jabber et al., 2020)yang berjudul **An Intelligent System for Classification of Brain Tumours with GLCM and Back Propagation Neural Network** menggunakan metode *Back Propagation Neural*

Network menghasilkan Klasifikasi tumor secara komputerisasi menggunakan MRI diajukan di mana fitur-fitur diekstraksi menggunakan Gray Level Co-occurrence Matrices (GLCM) dan klasifikasi menggunakan BPNN. Akurasi sebesar 94% dicapai dengan metodologi yang diajukan

Pada penelitian yang dilakukan oleh (Nanglia et al., 2020) yang berjudul **Lung cancer classification using feed forward back propagation neural network for CT images** menggunakan metode *Back Propagation Neural Network* menghasilkan ekanisme pelatihan menggunakan 200 gambar kanker dan metode yang diajukan menghasilkan akurasi klasifikasi sebesar 96% dan sensitivitas sebesar 94,7%.

Pada penelitian yang dilakukan oleh (Wu et al., 2020) yang berjudul **Back-propagation Artificial Neural Network for Early Diabetic Retinopathy Detection Based On A Priori Knowledge** menggunakan metode *Back Propagation Neural Network* menghasilkan bahwa BP-ANN dengan pengetahuan a priori dapat mencapai hasil deteksi yang lebih baik. Selain itu, hasil kami juga sebanding dengan algoritma state-of-the-art lain yang dilaporkan dan Kami juga mengembangkan antarmuka pengguna grafis dari BP-ANN yang diusulkan kami untuk membantu dalam skrining DR.

Pada penelitian yang dilakukan oleh (Tarigan dkk. 2017) yang berjudul **"*Genetic Algorithm* pada data Plat Nomor Kendaraan"** menggunakan metode *Genetic Algorithm dan Backpropagation Neural Network* menghasilkan *Backpropagation Neural Network* dengan optimasi *Genetic Algorithm* memiliki

Waktu konvergensi = 36.83% dan Akurasi = 1.35% lebih cepat dari *Backpropagation Neural Network* yang tidak di optimasi.

II.2. Machine Learning

M*achine Learning* adalah aplikasi atau bagian dari kecerdasan buatan yang membuat sistem memiliki kemapuan belajar secara otomatis dan meningkatkan kemampuannya belajar secara otomatis dan meningkatkan kemampuan berdasarkan pengalaman tanpa program komputer dan ditulis secara statis. *Machine learning* juga didefenisikan sebagai algoritma yang bertujuan menemukan dan mengaplikasikan pola-pola di dalam data(Hao,2018).

Berdasarkan paparan diatas *machine learning* dapat didefinisikan sebagai suatu algoritma atau program komputer yang data membuat sistem menjadi cerdas dengan mempelajari data-data yang tersedia dimana algoritma atau program tersebut tidak didefenisikan eksplisit (Purba Daru Kusuma,2020).

II.2.1. Pengeritan *Machine Learning*

Machine Learning sering dijelaskan sebagai pembelajaran tanpa pengawasan manusia atau dari pengalaman. Ini mencakup studi dan pengembangan algoritma yang dapat meramalkan dan belajar dari data. Cara kerja algoritma-algoritma tersebut adalah dengan menggunakan input untuk membuat model. Salah satu contohnya adalah penggunaan data untuk membentuk keputusan dan prediksi, daripada hanya mengikuti instruksi program yang sudah ditetapkan secara kaku. (Baharuddin, Azis, and Hasanuddin 2019). *Machine Learning* merupakan praktik

pemrograman komputer untuk belajar dari data dan program dengan mudah dapat menentukan apakah diberikan penting atau "spam" (Russell, 2018) *Machine Learning* diciptakan oleh Samuel pada tahun 1959, istilah *Machine Learning* diberikan untuk bidang studi komputer yang mampu belajar tanpa diprogram secara eksplisit, dengan kata lain, *machine Learning* dapat memperoleh pengetahuan langsung dari data dan belajar memecahkan masalah. *Machine Learning* tidak lama lagi akan mempengaruhi komunitas statistik (Bruce Ratner, 2017). *Big data* memerlukan metode otomatis analisis data, itulah yang disediakan oleh *Machine learning*. Secara khusus, mendefinisikan *Machine learning* sebagai kumpulan metode yang dapat secara otomatis mendeteksi pola dalam data dan kemudian menggunakan pola yang terungkap untuk memprediksi data di masa depan, atau melakukan jenis pengambilan keputusan lain dalam situasi ketidakpastian (Murphy, n.d.).

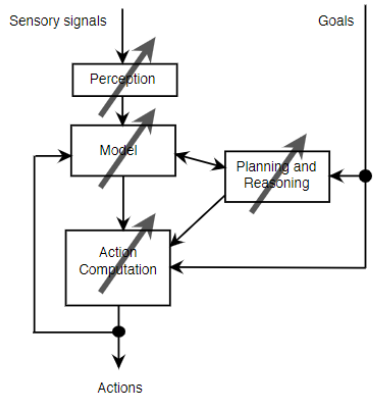
Menurut (Nilsson, 1998) Machine Learning adalah media pembelajaran, seperti kecerdasan, mencakup rentang proses yang begitu luas sehingga sulit untuk didefinisikan dengan tepat. Definisi dalam kamus mencakup frasa seperti "mendapatkan pengetahuan, pemahaman, atau keterampilan melalui studi, instruksi, atau pengalaman" dan "modifikasi kecenderungan perilaku melalui pengalaman." Ahli zoologi dan psikologi mempelajari pembelajaran pada hewan dan manusia. Dalam buku ini, kita berfokus pada pembelajaran pada mesin. Terdapat beberapa kesamaan antara pembelajaran pada hewan dan mesin. Tentu saja, banyak teknik dalam pembelajaran mesin berasal dari upaya psikolog untuk memperjelas teori-teori mereka tentang pembelajaran hewan dan manusia melalui

model komputasi. Kemungkinan besar, konsep dan teknik yang sedang dieksplorasi oleh peneliti dalam pembelajaran mesin juga dapat menerangi beberapa aspek pembelajaran biologis.

Sejauh berkaitan dengan mesin, kita dapat mengatakan, secara umum, bahwa mesin belajar ketika struktur, program, atau data-nya berubah (berdasarkan masukan atau sebagai respons terhadap informasi eksternal) sedemikian rupa sehingga kinerja masa depan yang diharapkan meningkat. Beberapa perubahan ini, seperti penambahan catatan ke dalam basis data, tergolong dalam disiplin lain dan tidak selalu lebih dipahami karena disebut sebagai pembelajaran. Tetapi, misalnya, ketika kinerja mesin pengenalan ucapan meningkat setelah mendengar beberapa contoh ucapan seseorang, kita merasa sangat berhak mengatakan bahwa mesin tersebut telah belajar.

Pembelajaran mesin biasanya mengacu pada perubahan pada sistem yang melakukan tugas-tugas terkait kecerdasan buatan (AI). Tugas-tugas tersebut melibatkan pengenalan, diagnosis, perencanaan, kontrol robot, prediksi, dll. "Perubahan" tersebut dapat berupa peningkatan pada sistem yang sudah berjalan atau sintesisab initio dari sistem baru. Untuk sedikit lebih spesifik, kami menunjukkan arsitektur dari suatu sistem AI yang khas. Kami menunjukkan arsitektur dari "Penyedia" AI yang khas dalam Gambar II.1. Penyedia ini menerima dan memodelkan lingkungannya serta menghitung tindakan yang sesuai, mungkin dengan mengantisipasi efeknya. Perubahan yang terjadi pada salah satu komponen yang ditampilkan dalam gambar tersebut dapat dianggap sebagai pembelajaran. Mekanisme pembelajaran yang berbeda dapat digunakan tergantung pada

subsistem yang mengalami perubahan. Kami akan mempelajari beberapa metode pembelajaran yang berbeda dalam buku ini.



Gambar II.1. ***Arsitektur Artificial Intelligence***

Berdasarkan Penjelasan diatas *machine learning* dapat didefinisikan sebagai suatu algoritma atau program komputer yang data membuat sistem menjadi cerdas dengan mempelajari data-data yang tersedia dimana algoritma atau program tersebut tidak didefenisikan eksplisit (Purba Daru Kusuma,2020).

II.2.2. Klasifikasi *Machine Learning*

Machine learning berdasarkam cara belajarnya dibagai menjadi tiga kelompok yaitu (Ibnu Daqiqil Id,2021) :

1. *Supervised Learning*

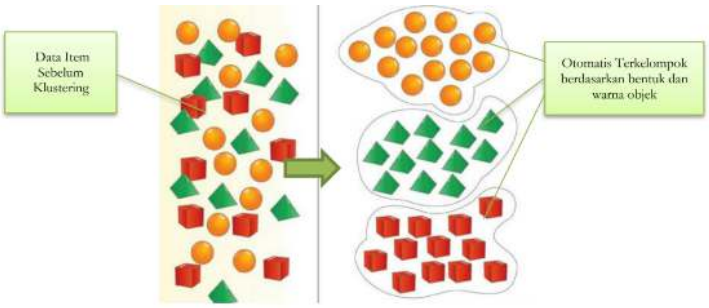
   Supervised leaning adalah pembelajaran terarah/terawasi, komputer atau mesin akan mempelajari data training yang berisi label. sifat dan variasi data latihan serta parameter sistem mempengaruhi akurasi klasifikasi yang dicapai (Pokorny et al., 2023b) dan (Wanayumini, 2022). Selain klasifikasi , *regresi* juga dimasukkan sebagai pembelajaran terarah. Adapun contoh-contoh algoritma

yang termasuk ke dalam kategori *supervised* adalah Linear *Regressuin (Regresi Linear), Logistic Regression (Regresi Logistik), Linear Discriminant Aalysis,k-Nearest Neighbors, Support Vector Machiner (SVMs), Radmom Forest, dan Neurol Network*

2. *Unsepervised Learning*

Menurut (Wanayumini, 2022). *Unsuvervised learning* adalah kebalikan *supervised learning* dimana proses pembelajaran dilakukan untuk menemukan pola-pola didalam data. Secara sistematis, *un supervised learning* terjadi ketika kita memiliki sejumlah data masukan (X) dan tanpa variablel *output* yang berhubungan. Masalah *Unsupervise learning* dibagi menjadi dua jenis yaitu asosiasi dan *clustering*.

Gambar II.2 mengilustrasikan poin ini dengan dua jenis item yang berbeda. Bergantung pada informasi, produk-produk ini akan dibagi menjadi beberapa kategori. Komputer hanya mengetahui fitur warna dan bentuk yang akan digunakan untuk mengidentifikasi item-item ini. Tanpa pengetahuan sebelumnya, komputer akan membagi item-item ini menjadi dua kelompok menggunakan algoritma pengelompokan. Algoritma melihat konten setiap item dan mencoba mengelompokkannya ke dalam beberapa kelompok. Algoritma seperti k-means dan peta pengorganisasian diri digunakan dalam teknik pengelompokan.



Gambar II.2. Ilustrasi proses Klasifikasi data

3. *Reunforcement Learning*

Reunforcement Learning berdasarkan cara kerjanya dapat dikelompokkan menjadi dua yaitu :

a. *Intance-based learning* (atau sering disebut memory-based laearning) adalah sebuah kelompok algoritma ML yang bekerja dengan membandingkan data testing dengan data yang telah dipelari pada proses traning. Algoritma tidak membuat sebuah generasi tetapi kea rah perbadingann dengan data yang disimpan dimemori.

b. *Model based learning* adalah kebalikan dari instance-based dimana menggunakan memori untuk melakukan pemecahan masalah, algoritma ini membuat sebuah model yang berisifat generik.

Berdasarkan Penjelasan diatas dapat disimpulkan Klasifikasi merupakan proses pembelajaran sebuah fungsi atau model terhadap sekumpulan data latih, sehingga model tersebut dapat digunakan untuk memprediksi klasifikasi dari data uji (Rosnelly R et al., 2021)

II.3. Image Processing

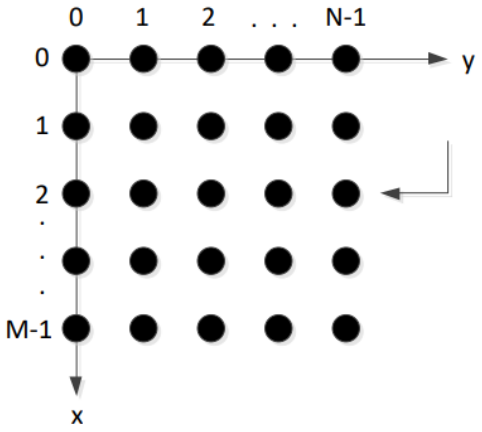
image processing adalah cabang ilmu komputer yang berkaitan dengan manipulasi digital atau analisis citra untuk memperoleh informasi yang berguna atau mengubah citra menjadi bentuk yang lebih baik atau lebih mudah dipahami. Image processing melibatkan pemrosesan digital dan analisis citra, termasuk ekstraksi fitur, pemrosesan histogram, segmentasi, restorasi citra, deteksi objek, pengenalan pola, dan banyak lagi, Proses image processing dimulai dengan mengambil citra digital menggunakan perangkat seperti kamera atau scanner(Nurdinawati et al., 2021). Citra yang diperoleh kemudian dapat disubmit ke berbagai operasi dan teknik pengolahan citra untuk mendapatkan informasi yang diinginkan atau untuk meningkatkan kualitas citra. Beberapa contoh aplikasi image processing termasuk pengolahan medis (seperti diagnosis penyakit melalui gambar medis), pengolahan citra satelit (untuk analisis permukaan bumi), pengolahan citra forensik (identifikasi sidik jari atau wajah), pengolahan citra di industri (pemeriksaan kualitas produk), dan banyak lagi(Desiani et al., n.d.). Tujuan utama dari image processing adalah untuk meningkatkan pemahaman dan ekstraksi informasi dari citra digital, baik secara otomatis maupun melalui intervensi manusia. Penggunaan image processing dapat memberikan manfaat besar dalam berbagai bidang, termasuk kedokteran, industri, keamanan, transportasi, robotika, dan banyak lagi (Brigham et al., 2015).

II.4. Citra Digital

Citra adalah suatu gambaran atau kemiripan dari suatu objek. Citra analog tidak dapat diproses secara komputerisasi, agar dapat diproses maka citra analog diubah menjadi citra digital. Citra digital adalah citra yang dapat diolah oleh komputer tetapi jika citra yang dihasilkan dari peralatan digital bisa langsung diolah oleh komputer (Andono et al., 2017).

II.4.1. Referensi Citra Digital

Telah diketahui bahwa hasil sampling dan kuantitas dari sebuah citra adalah bilang real yang membentuk matriks M baris dan N kolom. Ini berarti ukuran citra adalah M x N (Andono et al., 2017). Penggunaan optimasi *Pixel* berguna untuk deteksi objek, dan segmentasi nilai pixel dianggap faktor yang signifikan (Wanayumini, 2022). Secara umum sistem koordinat yang digunakan untuk mewakili citra adalah seperti gambar II.3.

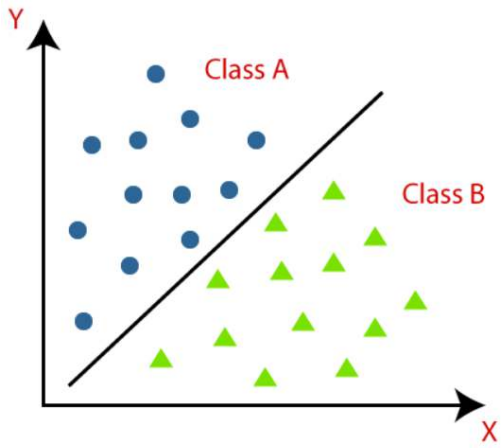


Gambar II.3. Sistem Koordinat yang dipergunakan untuk mewakili citra.

(Andono et al., 2017).

II.5. Support Vector Machine (SVM).

Support Vector Machine(SVM) adalah salah satu metode klasifikasi yang Tujuannya adalah untuk membuat batas keputusan antara dua kelas yang memungkinkan prediksi label dari satu atau lebih vektor fitur. Batas keputusan ini, yang dikenal sebagai hiperplane, diorientasikan sedemikian rupa sehingga jaraknya sejauh mungkin dari titik data terdekat dari setiap kelas. Titik-titik terdekat ini disebut support vector. Diberikan dataset pelatihan yang diberi label (Huang et al., 2018). Klasifikasi adalah proses pengelompokan, artinya mengumpulkan benda/entitas yang sama serta memisahkan benda/entitas yang tidak sama. Tujuan utama dari klasifikasi adalah untuk memprediksi sifat dari suatu item atau data berdasarkan kelas barang yang tersedia. (Oluwaseun & Chaubey, 2019). yang diulistrasikan pada gambar II.2. Secara umum klasifikas dapat diilustrasikan seperti gambar II.2. di bawah ini.



Gambar II.4. Klasifikasi dalam Machine Leaning. (Revathy et al., 2021)

Gambar (b) menunjukkan hyperlane dengan margin terbesar, sedangkan Gambar (a) menunjukkan beberapa pilihan hyperlane yang mungkin untuk dataset tersebut. Semua hyperlane pada Gambar (a) dapat digunakan, tetapi hyperlane dengan margin terbesar akan meningkatkan generalisasi metode klasifikasi. (Neneng *et. al.*, 2016).

Hasil dari pendekatan ini, yang berakar dalam teori pembelajaran statistik, memiliki potensi untuk melampaui hasil dari pendekatan lainnya. SVM berfungsi dengan baik pada dataset berdimensi tinggi juga. Bahkan pada SVM yang menggunakan kernel, diperlukan pemetaan data asli dari dimensi asalnya ke dimensi yang relatif lebih tinggi. SVM beroperasi secara berbeda dari ANN, karena tidak semua data digunakan dalam proses pelatihan. Pada *Support Vector Machine* (SVM), model yang digunakan untuk klasifikasi yang dipelajari terbentuk dari subset data sajaSelain itu, SVM berbeda dengan *Nearest Neighbor*, yang mencatat semua data latih untuk tujuan prediksi. SVM hanya menyimpan sebagian kecil dari set pelatihan untuk digunakan dalam membuat prediksi. Karena tidak semua data latih dipertimbangkan untuk berpartisipasi dalam setiap iterasi pelatihan, SVM memiliki keunggulan ini. Teknik ini juga dikenal sebagai Support Vector Machine karena kontribusi dari data, atau Support Vectors (SV). (Neneng *et. al.*, 2016).

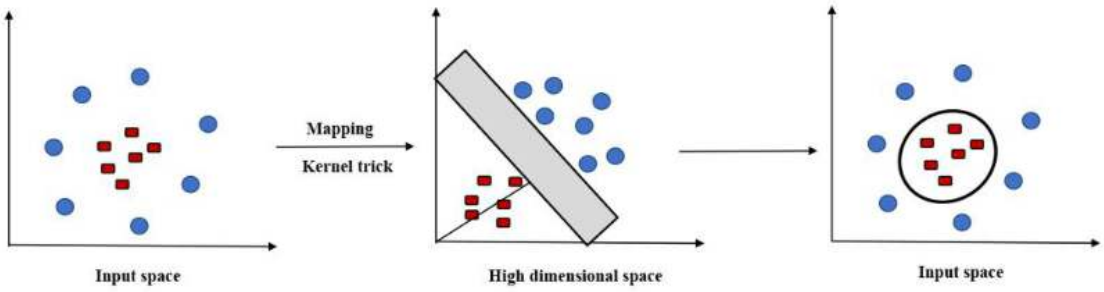
Pengklasifikasi SVM memiliki beberapa keunggulan, seperti kemampuan generalisasi yang lebih baik untuk menghindari overfitting dan masalah regresi, serta kemampuan untuk mengatasi data non-linear secara signifikan dengan menggunakan trik kernel. Jika data yang dikumpulkan tidak seimbang, SVM merupakan pilihan yang baik untuk klasifikasi karena dapat dengan mudah

menyelesaikan masalah data yang tidak seimbang. Selain itu, SVM secara efektif mengatasi masalah ukuran hasil ganda dengan mengembangkan batas kesalahan klasifikasi yang lebih santai (Balasubramanian et al, 2021)

Klasifikasi SVM adalah bagian penting dari pendekatan karena keseluruhan proses bergantung pada klasifikasi yang dilakukan melalui algoritma ini*, Support Vector Machine*, disingkat SVM dapat digunakan untuk tugas regresi dan klasifikasi. Akan tetapi ini banyak digunakan dalam tujuan klasifikasi.(Revathy et al., 2021).

1. *Support Vector* (SV)

Karush-Kuhn-Tucker (KKT) conditions adalah kondisi-kondisi yang menyatakan bahwa solusi yang diperoleh adalah optimal. atau pada solusi tersebut variabel dan kendala dari bentuk dual bertemu. Adalah sbb: (Neneng *et. al*., 2016)



Gambar II.5. *Support Vector Machine Linier Model*(Huang et al., 2018)

$a_n \geq 0$ .....(II.1)

$t_n y(x_n) - 1\xi_n \geq 0$ .....(II.2)

$a_n\{t_n y(x_n) - 1 + \xi_n\} = 0$ .....(II.3)

$\mu_n \geq 0$ .....(II.4)

$\xi_n \geq 0$ .....(II.5)

$\mu_n \xi_n = 0$ .....(II.6)

Support vectors adalah data training dengan nilai an > 0. maka dari II.3

$t_n y\ (x_n) = 1 - \ \xi_n$ .....(II.7)

Dimana :

$t_n y$ = Maximum Margin

x_n = Parameter Bobot x_i

ξ_n = Paramater nilainya.

Dimana :

a. Jika an < C. maka berdasarkan hasil turunan bahwa an = C – μn maka μn > 0. Selanjutnya. berdasarkan II.7 $\mu_n \xi_n$= 0 maka ξ = 0. Dengan kata lain adalah data training yang terletak pada hyperplane.
b. Jika an = C. maka μn = 0. Dari II.6 maka $\xi_n \neq 0$. Dengan kata lain adalah data training yang terletak didalam margin. baik yang terklasifikasi dengan benar ($\xi_n \leq 1$) atau salah ($\xi_n > 1$).
c. Selanjutnya. b dapat dicari dengan cara sbb:

$$t_n y\ (x_n) = 1$$

$$t_n\Big(\sum_{m \in S} a *_m\ t_m\ k\ (x_n, x_m) = 1 \rightarrow \frac{1}{N_s} \sum_{n \in S} \Big(t_n - \sum_{m \in S} a *_m t_m k\ (x_n, x_m)\Big)$$

i. Misal diberikan data z. maka prediksi dari data tersebut ditentukan berdasarkan fungsi y(x). yaitu:

a. Jika y(z) ≥ 0 maka z adalah kelas C_1 (t = +1)

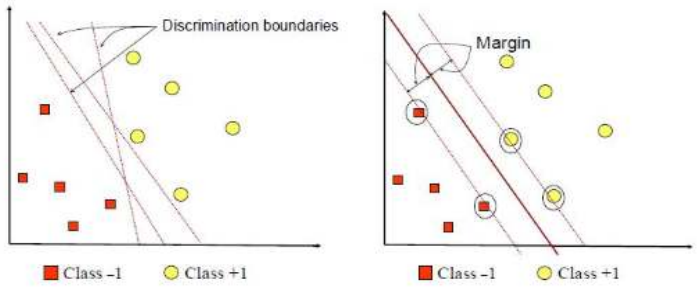
Jika dimana S adalah himpunana indeks dari support vectors. dan N adalah jumlah semua support vector. Misal diberikan data z. maka prediksi dari data tersebut ditentukan berdasarkan fungsi y(x). yaitu:

- Jika $y(z) \geq 0$ maka z adalah kelas C_1 $(t = +1)$
- Jika $y(z) < 0$ maka z adalah kelas C_2 $(t = -1)$.

Menurut (Drajama. 2017) ada 2 kasus dalam *Support Vector Machine*(SVM) yaitu:

1. Kasus data yang tepisah secara linear

   Kasus data yang terpisah secara linear adalah kasus yang mempermudah seperti yang terlihat pada gambar berikut :



(a) (b)

Gambar II.6. SVM Berusaha menemukan *Hyperlane* pemisah (Drajana, 2017)

Gambar 2.6 memperlihatkan beberapa pattern yang merupakan dari duah buah *class* +1 dan -1 yang mempunyai tupel pelatihan 2-D. *Pattern* yang bergabung pada *class* + 1 disimbolkan dengan lingkaran berwarna kuning. Upaya untuk mengidentifikasi hyperplane yang memisahkan dua kelompok dapat diartikan sebagai masalah klasifikasi. Dengan membandingkan margin hyperplane dengan pola terdekat dalam setiap kelas, hyperplane pemisah optimal antara kedua kelas dapat

ditentukan. Support vector adalah pola terdekat ini. Hyperplane terbaik, yang terletak di tengah-tengah kedua kelas, ditunjukkan oleh garis solid pada Gambar 2.9 di sebelah kanan. Support vector ditandai oleh titik merah dan kuning di dalam lingkaran hitam. Informasi yang ada ditunjukkan sebagai berikut: x€ R^d sedangkan label masing-masing dinotsikan $y_{i \in} \{-1, +1\}$ untuk i=1. 2. …. n yang mana n adalah banyaknya data. Diasumsikan kedua class dapat terpisah secara sempurna oleh *hyperplane* berdimensi d. yang didefenisikan pada persamaan 2.11.

$$W.x + b = 0 \quad \text{.....(II.8)}$$

Pattern x_1 yang terdapat pada class -1 dapat dirumuskan sebagai pattern yang memenuhi persamaan 2.12.

$$W.x + b \leq -1 \quad \text{.....(II.9)}$$

Pattern x_i yang terdapat pada class -1 dapat dirumuskan dengan persamaan 2.13.

$$W.x + b \geq 1 \quad \text{.....(II.10)}$$

Dimana :

W = *Vector* bobot

x = nilai masukan atribut

b = bias

Dengan memaksimalkan jarak antara *hyperplane* dan titik terdekatnya, margin terbesar dapat ditentukan. yaitu 1 / || W ||. Masalah ini dapat dirumuskan sebagai masalah quadratic programming (QP), di mana tujuannya adalah untuk menemukan titik minimum yang dinyatakan dalam persamaan sambil memperhatikan kondisi yang harus dipenuhi pada persamaan 2.14 dan persamaan 2.15 berikut.

$min_w = r(w) = \frac{1}{2} || W ||^2$ .....(II. 11)

$y_i (x_i . w + b) - 1 \geq 0$ .....(II.12)

Dimana :

x_i = Data input

y_i = Keluaran dari data x_i. w n. b

w. b = Paramater yang dicari nilainya.

Problem ini dipecahkan dengan teknik komputasi. diantaranya *lagrange multiplier tang* dinyatakan pada persamaan II.6.

L (w. b. a) = $\frac{1}{2 || w ||^2} - \sum_{i=1}^{l} \propto_i \ y_i (x_i . w + b) - 1$ dg t = 1,2 ...l (II.13)

Dimana a_1 adalah *lagrange multiplier* yang bernilai 0 atau $a_1 \geq 0$. nilai optimal dari persamaan 2.16 dapat dihitung dengan meminimalkan L terhadap w dan b dan memaksimalkan L terhadap a_l. Dengan memperhatikan sifat bahwa pada titik *gradient* L=0 persamaan 2.15 dapat dimodifikasi sebagai masksimasi *problem* yang hanya mengandung a_l sebagimana terlihat pada persamaan 2.17 dan 2.18.

$\sum_{i=1}^{l} a1 - \frac{1}{2} \sum_{i,j=0}^{l} = 1 a_i \ a_j \ y_i \ y_j x_i \ x_j$ .....(II.14)

$\sum_{i}^{l} a_1 \ y_i = 0 \ untuk \ a_i \geq 0 (i = 1,2, \dots l)$ .....(II.15)

Dengan demikian. maka akan diperoleh a_1 yang kebanyakan bernilai positif yang disebut sebagai *support Vector*.

2. Kasus Data yang tidak terpisah secara linear

Kasus data yang tidak terpisah secara linear diasumsikan bahwa *class* pada input space tidak dapat terpisah secara sempurna. Hal ini menyebabkan constraint pada persamaan 2.13 tidak dapat terpenuhi. Sehingga Optimasi tidak dapat

dilakukan. Untuk mengatasi masalah ini SVM dirumuskan ulang dengan memperkenalkan teknik softmargin. Dalam *softmargin* persamaan 2.15 dimodifikasi dengan menggunakan *slack* variabel sehingga pada persamaan 2.19.

$y_i(x_i w + b) \geq 1 - \varepsilon_i$ ..... (II.16)

Dengan demikian persamaan 2.17 diubah menjadi persamaan 2.20

$mtn_w = \tau(w) = \frac{1}{2|| w ||^2} + C \sum_{i=1}^{l} \varepsilon_i$ ..... (II.17)

Fitur C digunakan untuk mengontrol *tradeoff* antara margin dan kesalahan klasifikasi ε.

a. *Kernel Trick* dan *Non-Linear Classification* pada SVM

Pada umumnya masalah yang terjadi dalam dunia nyata jarang bersifat linear separable. Kebanyakan bersifat *non-linear*. SVM dimodifikasi dengan memasukkan fungsi kernel. Dalam *non linear* SVM. pertama-tama data x dipetakan oleh fungsi ϕ (x) keruang vektor yang berdimensi lebih tinggi. Pada ruang vektor yang baru ini. hyperlane yang memisahkan kedua class tersebut dapat dikonstruksikan. Hal ini sejalan dengan teori cover yang menyatakan "Jika suatu transformasi bersifat non linear dan dimensi dari *feature space* cukup tinggi. maka data pada input *space* dapat dipetakan ke *feature space* yang baru. dimana *pattern-pattern* tersebut pada probabilitas tinggi dapat dipisahka secara linear".

Pemetaan ini dilakukan dengan menjaga topologi data. dalam artian dua data yang berjarak dekat pada input *space* akan berjarak dekat juga pada *feature* space. sebaliknya dua data yang berjarak jauh pada input *space* akan juga berjarak jauh pada *feature space*. Selanjutnya proses pembelajaran pada SVM dalam menemukan titik-titik *support vektor*. hanya bergantung pada dot *product* dari data

yang sudah ditransformasikan pada ruang baru yang berdimensi lebih tinggi. yaitu $\phi(x_i).\phi(x_j)$. Karena umumnya transformasi ϕ ini tidak diketahui. dan sangat sulit untuk difahami secara mudah. maka perhitungan dot product tersebut sesuai teori Mercer dapat digantikan dengan *kernel* yang terlihat pada persamaan II.11.

$K(x_i\, x_j) = \phi(x_j).\phi(x_j)$ .....(II.18)

dimana :

x_i = *Support Vector*

Beberapa kernel yang terdapat pada SVM meliputi: (Prasetyo. 2014).

1. Polinominal Derajat h

   Ketika dataset pelatihan dinormalisasi, trik kernel *polinomial* sangat efektif untuk menyelesaikan masalah klasifikasi. Persamaan II.19 memberikan ungkapan untuk trik kernel ini.

$K(x_i\, x_j) = (x_i\, x_j + 1)^{h}$ ..... (II.19)

2. *Radial Basis Function*

   Kernel trick radial basis *function* merupakan *kernel* yang paling banyak digunakan untuk menangani masalah klasifikasi, terutama untuk dataset yang tidak terpisah secara *linear*. Keunggulan akurasi pelatihan dan akurasi prediksi menjadi alasan utamanya. Persamaan 2.23 menyajikan ungkapan untuk kernel *radial basis function.*

$K(x_i, x_j) = e^{-||x_i - x_i||\ 2/2a^2}$ ..... (II.20)

Dimana: X_i dan X_j = Pasangan dua data training.

3. *Sigmoid* (tangen hiperbolik)

Kernel sigmoid merupakan *kernel trick*. *Support Vektor Machine* yang merupakan pengembangan dari jaringan saraf tiruan. dimana *kernel* ini dinyatakan dengan persamaan II.21.

K (x_i, x_j) = tan $(kx_i . x_j - \delta)$ ..... (II.21)

Kernel trick memberikan berbagai keuntungan karena menghilangkan kebutuhan bagi pengguna untuk memahami bentuk fungsi non-linier guna menentukan vektor dukungan selama proses pembelajaran SVM. Sebaliknya, pengguna hanya perlu mengetahui fungsi trik kernel yang sedang digunakan. Fungsi kernel basis radial (RBF) adalah yang paling efektif dalam klasifikasi, terutama ketika berurusan dengan data yang tidak dapat dipisahkan secara linier. Selain masalah data yang tidak dapat dipisahkan secara linier, masalah umum lain yang dihadapi dalam penerapan teknik klasifikasi *machine learning* seperti *Support Vector Machines* adalah masalah dimensionalitas dataset, yang juga disebut sebagai "kutukan dimensionalitas." Data akan tersebar merata dalam ruang yang dihuni saat dimensionalitas meningkat.

Manfaat dari pengurangan dimensi:

1. Mencegah terjadinya efek dari dimensionalitas.
2. Mengurangi jumlah waktu dan memori yang dibutuhkan oleh machine learning.
3. Membuat data lebih mudah divisualisasikan.
4. Membantu untuk mengurangi fitur-fitur yang tidak relevan

Teknik pengurangan dimensionalitas data diantaranya adalah *principal component analysis* (PCA). standar deviasi. *zero-mean. min-max normalization*. dan lain-lain.

a. Standar Deviasi

Standar deviasi juga kadang-kadang disebut sebagai *varians*, adalah suatu teknik yang digunakan untuk menemukan variasi dalam suatu dataset dan untuk mengurangi jumlah dimensi dalam suatu dataset sambil tetap mempertahankan unit pengukuran yang sama dengan data asli. Singkatnya, deviasi standar mengukur sejauh mana nilai-nilai dalam suatu dataset tersebar. Jarak rata-rata antara satu set titik dan mean adalah cara lain untuk mendefinisikannya. Akar kuadrat dari jumlah kuadrat selisih antara setiap titik data dan mean dataset, dibagi oleh total jumlah titik data, adalah standar deviasi. Persamaan II.22 dapat digunakan untuk menuliskan standar deviasi.

$$S = \sum \sqrt{\frac{|x_i - x|^2}{n-1}} \quad \text{(II.22)}$$

Dimana :

x = Rataan hitung

x_i = Input data

N = Jumlah data

s = Standar deviasi.

II.6. Backpropagation Neural network

Pelatihan pada *neural network* membutuhkan 3 operasi yang berurutan: *forward propagation, backpropagation,* dan memperbarui bobot. Dalam *forward propagation,* data dipropagasikan dari *input layer* menuju *output layer.* Setiap *neuron* akan menjumlahkan total bobot dari *input neuron* yang terhubung dari *layer* sebelumnya, dan kemudian menambahkan dengan sebuah *bias* yang dapat dilihat pada persamaan 2.23 sebagai berikut (Park dan Suh 2022) :

$u_j^l = \sum_{i=1}^{N} w_{ij}^l x_{ij}^l + Bias^1$ ..... (II.23)

Dimana :

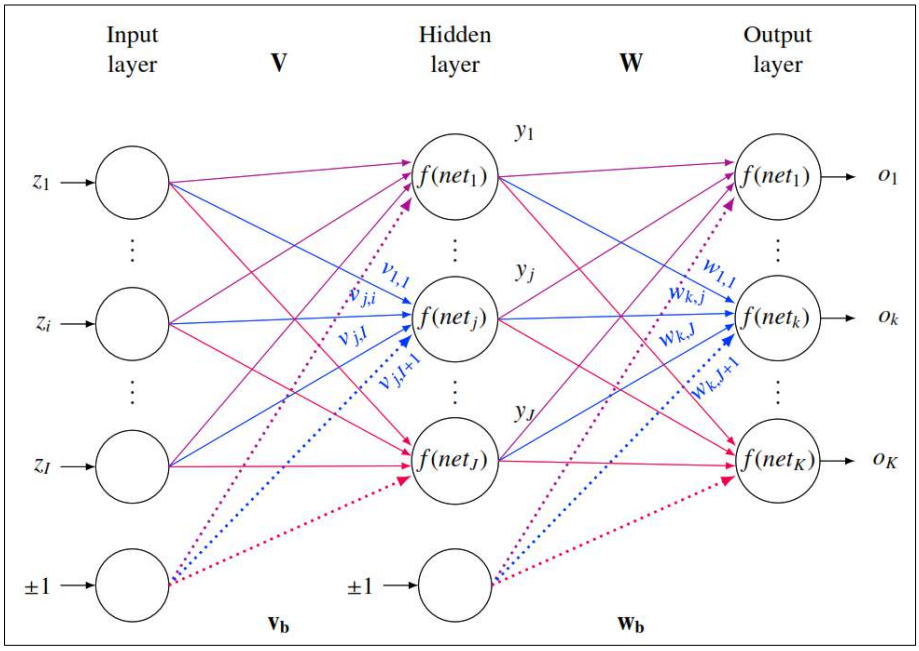
u_j^l = Luaran dari Neuron ke j dalam lapisan j

$\sum_{i=1}^{N} w_{ij}^l$ = jumlah bobot yang dikalikan dengan keluaran neuron sebelumnya dan ditambahkan dengan bias pada neuron ke-j.

x_{ij}^l = vektor bobot yang menghubungkan neuron sebelumnya dengan neuron ke-j.

$Bias^1$ = bias neuron ke-j.

Backpropagation sangat populer pada algoritma *supervised learning,* arsitektur dari *multilayer perceptron* dapat dilihat pada **Gambar II.5** sebagai berikut (Sapkal dan Kulkarni 2018) :



Gambar II.7. Arsitektur *Multilayer Perceptron*

(Sumber : Sapkal dan Kulkarni 2018)

Backpropagation digunakan untuk menyesuaikan bobot (w_{ij}^{l}) dengan menghitung turunannya. Turunan pada *output layer* dapat dilihat pada persamaan II.18 sebagai berikut (Park dan Suh 2022) :

$$\frac{dE}{dy_z^o} = y_z^o - t_z \quad \text{..... (II.24)}$$

Hasil perhitungan berbeda antara hasil *forward propagation* (y_z^o) dengan target *output* (t_z). Turunan dari *error* dengan perhitungan bobot dapat dilihat pada persamaan 2.5 sebagai berikut (Park dan Suh 2022) :

$$\frac{dE}{dy_k^l} = \sum_z w_{kz}^{l+1} \frac{dE}{dy_z^{l+1}} \quad \text{..... (II.25)}$$

Bobot disesuaikan berdasarkan hitungan *error* dalam *backpropagation*. Pertama, Δw_{ij}^{l} dihitung dengan mengalikan hasil dari *backpropagation* δ_i^l dengan

hasil dari *forward propagation* y_j^{l-1} pada persamaan 2. . Bobot w_{ij}^l kemudian diperbarui seperti pada persamaan 2.6 dan 2.7 dimana *learning rate* akan ditentukan pada proses pelatihan. Proses ini akan diulangi terhadap semua bobot (Park dan Suh 2022) :

$$\Delta w_{ij}^l = \delta_i^l y_j^{l-1} \quad \text{..... (II.26)}$$

$$w_{ij}^l = w_{ij}^l - \Delta w_{ij}^l \eta \quad \text{..... (II.27)}$$

II.7. Fungsi Aktivasi

Fungsi aktivasi membantu dalam pelatihan agar menjadi lebih baik, proses pelatihan dan kemampuan generalisasi menjadi lebih baik. Fungsi akivasi mengendalikan setiap unit didalam *layer* dengan bekerja pada bobot dan *bias*. Proses ini menggabungkan *non linear* dalam *output* dari setiap *neuron*. Bobot diupdate berdasarkan *error output.* Proses dari *backpropagation* mendukung fungsi aktivasi dengan memberikan *gradient* (Singh dkk. 2023).

Output dari pelatihan akan bergerak melalui fungsi aktivasi yang ditentukan untuk melewati *layer* selanjutnya. Kebanyakan fungsi aktivasi yang digunakan adalah *Rectified Linear Unit (ReLU), Tanh,* dan *Sigmoid*. Rumus *ReLU* dan hitungan turunan fungsi aktivasi *ReLU* dapat dilihat pada persamaan 2.8 dan persamaan 2.9 sebagai berikut (Park dan Suh 2022) :

$$y_j^l = max(0, u_j^l) \quad \text{..... (II.28)}$$

$$\delta_k^l = \frac{dE}{du_k^l} = \frac{dE}{dy_k^l}\frac{dy_k^l}{du_k^l} \quad \text{..... (II.29)}$$

Fungsi perpindahan *linear* sederhana yang menarik dalam konteks *multilayer neural network* dapat dilihat pada persamaan 2.10 sebagai berikut (Oostwal, Straat, dan Biehl 2021) :

$$g(x) = max(0, x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases} \quad \text{..... (II.30)}$$

Fungsi aktivasi *sigmoid* terutama digunakan untuk klasifikasi multi-klasifikasi karena memetakan nilai *input* dalam rentang (0,1) yang pada dasarnya probabilitasnya milik *class* yang diberikan yang dapat dilihat pada persamaan II.25 sebagai berikut (Elfwing, Uchibe, dan Doya 2018; Singh dkk. 2023):

$$\sigma(x) = \left(\frac{1}{1+\exp(-x)}\right) \quad \text{..... (II.31)}$$

II.8. Mean Square Error (MSE)

Loss Function pada model dapat dinyatakan dengan *Mean Square Error (MSE)* pada persamaan 2.12 sebagai berikut (H.-Y. Qu dkk. 2023) :

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(yi^{predict} - yi^{data})^2 \quad \text{..... (II.32)}$$

II.9. Stochastic Gradient Descent (SGD)

Pada bidang *neural network, stochastic gradient descent* sering digunakan sebagai metode yang efektif untuk mempercepat hasil konvergensi. Menghasilkan *gradient* baru dari gradient sebelumnya dimana metode ini diadopsi dari banyak algoritma optimisasi yang ada (Song, Pons, dan Yen 2021).

Stochastic Gradient Descent : x_t menjadi sebuah contoh acak dari *data set* pada sebuah iterasi *t*, $J_\theta(x_t)$ adalah fungsi objektif yang mengevaluasi pada data x_t

dengan parameter θ, $g_t = \nabla_\theta J_\theta(x_t)$ adalah *gradient*, dan α adalah *learning rate* (Ilboudo, Kobayashi, dan Sugimoto 2022). Algoritma *SGD* mengupdate θ_{t-1} to θ_t melalui persamaan 2.13 sebagai berikut (Robbins dan Monro 1951; Ilboudo, Kobayashi, dan Sugimoto 2022) :

$\theta_t = \theta_{t-1} - \alpha g_t$ ..... (II.33)

II.10. CT Scan

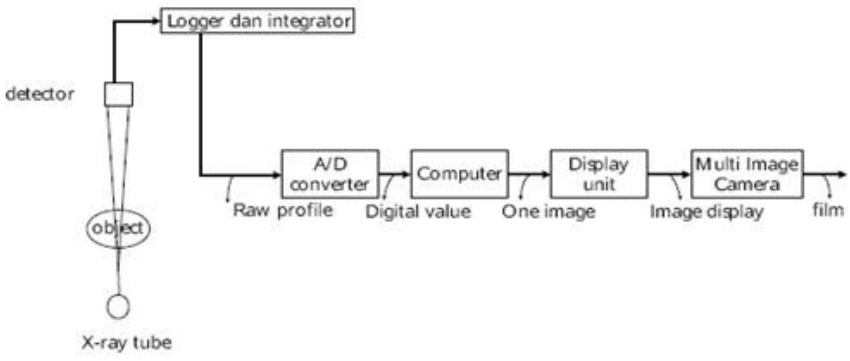
Computed Tomography (CT) Scan merupakan salah satu teknologi di bidang kesehatan yang memungkinkan dokter untuk melihat tubuh manusia dengan lebih detail. Teknologi ini memungkinkan pengambilan gambar tiga dimensi yang akurat dari organ dan jaringan dalam tubuh manusia (Wahyuni & Amalia, 2022b) . Teknologi CT Scan pertama kali dikembangkan pada tahun 1972 oleh Godfrey Hounsfield dan Allan Cormack, yang kemudian memenangkan hadiah Nobel di bidang Fisiologi dan Kedokteran pada tahun 1979. Sejak saat itu, teknologi ini terus berkembang dan semakin canggih dalam penggunaannya, terutama dalam mencari penyakit seperti kanker dan penyakit lainnya. Contoh Hasil CT Scan dapat dilihat pada **Gambar II.6** sebagai berikut (Wahyuni & Amalia, 2022b)

Gambar II.8. Gambar Hasil CT scan

II.10.1. Mobile/Portable CT Scan CT scan

Rumah sakit dapat lebih baik memanfaatkan peralatan CT Scan dengan menerapkan teknologi CT *portabel*. Ini memiliki dua manfaat utama: Pertama, alur kerja CT Scan yang ditingkatkan di rumah sakit mempercepat perawatan pasien non-ICU sekaligus meningkatkan kualitas perawatan pasien. Kedua, terdapat keuntungan finansial yang terkait dengan mesin CT portabel karena penggunaan peralatan yang lebih tinggi. Gambar II.7 menunjukkan contoh dan alur kerja CT scan sebagai berikut. (Wahyuni & Amalia, 2022b)

Gambar II.9. Foto mobile CT Scan



Gambar II.10. Alur prinsip kerja CT scan

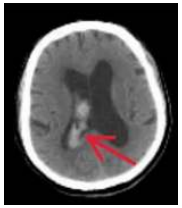
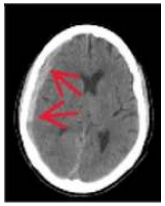
II.11. Stroke

Stroke, juga dikenal sebagai serangan otak, terjadi ketika pasokan darah ke otak terhenti (admin & Padimi, 2022).World Health Organization (WHO) mendefinisikan stroke sebagai suatu kondisi dimana ditemukan defisit neurologis fokal dan global dengan tanda klinis yang dapat memburuk dan berlangsung selama kurang lebih 24 jam serta menyebabkan kematian.. (Solikhun et al., 2023) Intracranial hemorrhage adalah kondisi di mana terjadi pendarahan di dalam tengkorak, yang melibatkan ruang intrakranial seperti otak, pembuluh darah di otak, atau jaringan sekitarnya. Pendarahan ini dapat disebabkan oleh berbagai faktor, termasuk trauma kepala, pecahnya pembuluh darah di otak (misalnya, aneurisma), hipertensi, atau gangguan pembekuan darah. 2 jenis stroke: Stroke iskemik (iStroke) yang disebabkan oleh penyumbatan pembuluh darah oleh bekuan darah dan stroke hemoragik (hStroke) yang disebabkan oleh pendarahan di dalam tengkorak, juga skenario tanpa stroke (Pokorny et al., 2023b)

Intracranial hemorrhage merujuk pada pendarahan yang terjadi di dalam tengkorak, yang merupakan masalah kesehatan yang serius dan membutuhkan pengobatan medis yang cepat dan seringkali intensif. Misalnya, pendarahan intrakranial menyumbang sekitar 10% dari kasus stroke di Dunia, di mana stroke merupakan penyebab kematian terbesar kelima.

Ada lima subtipe pendarahan: Intraparenkimal, Intraventrikular, Subarachnoid, Subdural, dan Epidural (dapat lihat pada table II.1). Pasien dapat mengalami lebih dari satu jenis pendarahan, yang mungkin muncul dalam gambar yang sama.

Tabel II.1 Tipe Pendarahan pada otak

	Intraparenkimal	Intraventrikular	Subarachnoid	Subdural	Epidural
Lokasi	Bagian dalam Kepaka	bagian dalam ventrikel	antara arachnoid dan pia mater	antara dura dan arachnoid	antara dura dan Tengkorak
Image					
Mekanisme	Tekanan darah tinggi, trauma, Ateri ke vena, kelainan pada struktur tubuh serta tumor dll	dapat dikaitkan dengan perdarahan, peradangan dan subarachnoid	pecahnya Ateri ke vena atau kelainan pada struktur tubuh atau trauma	Trauma	trauma atau setelah operasi
Sumber	Arteri atau Vena	Arteri atau Vena	Trauma Arteri	Vena	Arteri
Bentuk	Berbentuk bulat di sisi kanan atau kiri Otak besar	menyesuaikan diri dengan bentuk ventrikel pada selapot tengah otak	melacak sepanjang sulkus dan celah	Pendarahan pada sisi samping kepala	Pendarahan Besar pada sisi kepala Pada Otak besar
Gejala	akut (sakit kepala tiba-tiba, mual, muntah)	akut (sakit kepala tiba-tiba, mual, muntah	akut (sakit kepala terus menerus)	Sakit Kepala yang Memungkinkan bertambah bahaya (semangkin memburuk)	Benturan terhadap kepala

II.12. Penelitian yang Relevan

Tabel II.2 Penelitian Relevan

No	Penelitian	Metode	Tujuan	Kekurangan	Manfaat
1	*Comparison of Neural Network Training Functions for Hematoma Classification in Brain CT Images*	*Backpropagation Neural Network, Artificial neural network*	Menemukan fungsi pelatihan terbaik untuk klasifikasi *hematoma* otak, Untuk evaluasi performa fungsi pelatihan, parameter yang digunakan adalah akurasi pengenalan, kecepatan pelatihan, dan kebenaran.	Kesulitan utama dalam mengadopsi ANN adalah menemukan kombinasi yang paling sesuai dari fungsi pembelajaran, transfer, dan pelatihan untuk tugas klasifikasi	Algoritma ini mengukur kesalahan output, menghitung gradien kesalahan dengan menyesuaikan bobot secara menurun
2	*An Intelligent System for Classification of Brain Tumours with GLCM and Back Propagation Neural Network*	*Backpropagation Neural Network*	Membandingkan dan Menentukan model mana yang terbaik dari dataset yang sama	Kurangnya pengembangan/ penelitian terkait prihal permasalahan pada	Karya yang diusulkan dengan pencapaian akurasi sekitar 94% menunjukkan bahwa ini adalah metode terbaik untuk klasifikasi otomatis tumor otak. Karya yang diusulkan dibandingkan dengan akurasi dari berbagai jenis klasifier seperti ANFIS, PNN, CNN, dan SVM.

No	Penelitian	Metode	Tujuan	Kekurangan	Manfaat
3	*Detection of Collaterals from Cone-Beam CT Images in Stroke*	*cone-beam computed tomography, support vector machine (SVM); K-nearest neighbors (KNN)*	Memperbaiki kualitas pelayanan, akurasi waktu pemindaian yang cepat	Masih Membutuhkan algoritma lain, masih terdapat noise, menerapkan banyak Jenis ekstraksi ciri	Perbaikan kualitas citra, dan ekstraksi karakteristik citra dengan hasil akurasi klasifikasi yang menjanjikan di atas 90% dan mampu mendeteksi pembuluh kolateral dan non-kolateral dari gamba
4	*On the Role of Training Data for SVM-Based Microwave BrainStroke Detection and Classification*	*SVM; brain stroke; microwave devices; numerical mode*	Data sintetis yang dihasilkan digunakan untuk menguji empat hipotesis berbeda, yang berkaitan dengan parameter yang sesuai untuk set data latihan, rentang frekuensi yang sesuai, jumlah titik frekuensi yang diperlukan, serta tingkat variasi subjek untuk mencapai akurasi klasifikasi SVM tertinggi.	Keterbatasan dalam representasi data nyata, Ketergantungan pada parameter sistem, Keterbatasan pada SVM, Validitas eksternal	Hasil penelitian menunjukkan bahwa algoritma SVM mampu mendeteksi keberadaan stroke dan mengklasifikasikannya (yaitu iskemik atau hemoragik), bahkan ketika dilatih dengan data satu frekuensi.

No	Penelitian	Metode	Tujuan	Kekurangan	Manfaat
5	*A hybrid mental health prediction model using Support Vector Machine, Multilayer Perceptron, and Random Forest algorithms*	*Support Vector Machine, Multilayer Perceptron, and Random Forest.*	Memprediksi tahapan kecemasan pra-klinis, termasuk kecemasan minimal, kecemasan ringan, kecemasan sedang, kecemasan berat, dan kecemasan sangat berat	Diperlukan sinkronisasi prediksi pembelajaran mesin dengan tahapan kecemasan mental pra-klinis untuk memberikan rekomendasi kesehatan mental yang disesuaikan, dengan mempertimbangkan konteks budaya pelaporan kesehatan mental di berbagai wilayah	Penelitian ini bertujuan untuk mendukung para profesional kesehatan mental dalam membuat keputusan yang akurat dan tepat waktu
6	*Lung cancer classification using feed forward back propagation neural network for CT images*	*forward back propagation, genetic algorithm and fitur speed up robust feature (SURF)*	mengoptimalkan proses klasifikasi	Membutuhkan waktu yang lama dalam pemprosesan Tomografi terkomputasi dalam mennganalsisi penyakit kanker	Mekanisme pelatihan menggunakan 200 gambar kanker dan metode yang diusulkan menghasilkan akurasi klasifikasi sebesar 96% dan sensitivitas sebesar 94,7%. Makalah ini juga membahas modifikasi masa depan yang mungkin dilakukan dalam karya yang disajikan.

No	Penelitian	Metode	Tujuan	Kekurangan	Manfaat
7	*Performance Evaluation of SVM and K-Nearest Neighbor Algorithm over Medical Data set*	*Support Vector Machine and K-Nearest Neighbor Algorithm*	mengklasifikasikan data dan mendapatkan prediksi (mencari pola tersembunyi) untuk target yang diinginkan	Terdapat masalah dalam menentukan algoritma data mining yang paling cocok untuk data tersebut,	menerapkan model-model yang berbeda pada satu set data dan membandingkan kinerja mereka untuk memilih yang terbaik
8	Implementasi Sistem Pendeteksi Myocardial Ischemia menggunakan Metode Support Vector Machine (SVM)	*Support Vector Machine, Artifical Neural Network, KNearest Neighbors*	Mengklasifikasi Myocardial Ischemia dan Normal dalam mendeteksi tingginya kadar kolestrol darah	Perlunya kolaborasi Serta kombinasi Penelitian lain Seperti pemanfaatan mikrokontroler seperti Arduino Uno, dll	penelitian ini dilaksanakan dengan membangun Sistem Pendeteksi Myocardial Ischemia menggunakan Metode Support Vector Machine untuk mempermudah akses pelayanan diagnosis myocardial ischemia.
9	Measuring the Accuracy of SVM with Varying Kernel Function for Classification of Indonesian Wayang on Images	*Support Vector Machine*	Mengukur Akurasi SVM dengan Berbagai Fungsi Kernel untuk Klasifikasi Wayang Indonesia	Banyak metode yang digunakan dalam klasifikasi, terutama dalam pengolahan gambar.	Ekstraksi Fitur melakukan proses yang terjadi dalam bagian ini adalah untuk "mengkode" dari gambar menjadi fitur dalam bentuk angka yang mewakili gambar sehingga media gambar yang akan diklasifikasikan dapat dengan mudah mendapatkan pola dalam pengenalan gambar

No	Penelitian	Metode	Tujuan	Kekurangan	Manfaat
10	Detection of HOG Features on Tuberculosis X-Ray Results Using SVM and KNN	*Support Vector Machine and K-Nearest Neighbor*	untuk mendapatkan deteksi yang akurat dengan melakukan ekstraksi fitur agar gambar dapat dikenali melalui komputasi	Perlu dilakukan penelitian serta perbandingan menggunakan Metode, algoritma lainnya guna meningkatkan akurasi penelitian	membantu dalam deteksi gambar pada jenis Tuberkulosis berdasarkan sinar-X dalam deteksi HOG pada gambar sinar-X Tuberkulosis dengan KNN untuk gambar positif mencapai akurasi sebesar 77,95%, sementara untuk gambar negatif mencapai akurasi sebesar 78,65%. Sementara itu, hasil deteksi HOG pada gambar sinar-X Tuberkulosis menggunakan SVM untuk gambar positif mencapai akurasi sebesar 65,75%, sedangkan untuk gambar negatif mencapai akurasi sebesar 79,39%.
11	*Komparasi Metode Multi Layer Perceptron (MLP) dan Support Vector Machine (SVM) untuk Klasifikasi Kanker Payudara*	*Multi Layer Perceptron (MLP) dan Support Vector Machine (SVM)*	Mengembangkan pola dalam Machine learning untuk dapat membantu proses Pemeriksaan kesehatan, efektifitas serta akurasi	Keterbatasan dalam data set yg di sediakan, dan masih bersifat data sekunder	Adapun hasil yang didapatkan menunjukan bahwa, dalam klasifikasi metode Multi Layer Perceptron(MLP) dengan fungsi aktivasi Logisticdan fungsi optimisasi Adammemberikan nilai accuracy, precision dan recallterbaik dibandingkan Support Vector Machine yaitu sebesar 97.7%