

BAB 2

DASAR TEORI

2.1 KAJIAN PUSTAKA

Beberapa penelitian sebelumnya yang menggunakan metode klasifikasi *Convolutional Neural Network* telah banyak dijadikan referensi dan ide dalam melakukan penelitian ini. Hal ini dilakukan untuk menghindari suatu duplikasi pada penelitian dan sebagai referensi serta landasan dalam proses sebelum penelitian. Adapun beberapa penelitian yang menggunakan metode klasifikasi *Convolutional Neural Network* atau pada rambu pembatas kecepatan antara lain sebagai berikut:

Penelitian yang dilakukan oleh Nur Fadila dan Rifki Kosasih pada tahun 2019 dengan judul “Klasifikasi Jenis Kendaraan Menggunakan Metode *Convolutional Neural Network* (CNN)”. Penelitian ini memiliki tujuan untuk pengenalan jenis-jenis kendaraan di jalan raya. Penelitian ini menggunakan metode *Convolutional Neural Network* (CNN). Data yang digunakan pada penelitian ini adalah 120 citra yang meliputi citra mobil, sepeda motor dan sepeda.. Dari penelitian ini menghasilkan akurasi yang tinggi sebesar 73,3% [17].

Kemudian pada tahun 2018, Erlyna Nour Arrofiqoh dan Harintaka melakukan penelitian dengan judul “Implementasi Metode *Convolutional Neural Network* Untuk Klasifikasi Tanaman Pada Citra Resolusi Tinggi” penelitian ini menggunakan metode klasifikasi yang sama yaitu CNN namun objek yang diklasifikasikan berbeda yaitu tanaman. Penelitian ini mencoba mengimplementasikan metode klasifikasi semantik otomatis yang dapat membedakan jenis tumbuhan sebagai alternatif pengenalan objek berdasarkan metode deep learning menggunakan *convolutional neural network* (CNN). Penelitian ini dilakukan dengan menggunakan 5 kelas jenis tanaman, yang meliputi kelas tanaman kelapa, padi, pisang, bawang merah, dan cabai dengan masing-masing kelas terdiri dari 100 dataset sehingga data keseluruhan yang digunakan berjumlah 500 data. Pengujian berhasil mengklasifikasikan 5 jenis tanaman secara otomatis dengan menghasilkan akurasi tinggi sebesar 82% [18].

Pada tahun 2020 dilakukan penelitian tentang pengenalan rambu pembatas kecepatan oleh Nugraheny Wahyu Try dan Fitri Utamingrum yang berjudul “Sistem Pengenalan Rambu Pembatas Kecepatan Menggunakan Metode *Histogram of Oriented Gradients* dan Klasifikasi *K-Nearest Neighbor* Berbasis Raspberry Pi”. Pada penelitian ini menggunakan metode *Histogram of Oriented Gradients* (HOG) dan *K-Nearest Neighbor* untuk mengenali rambu pembatas kecepatan maksimum 20km, 25km, 30km, 40km, 50km, minimum 20km dengan menggunakan 650 data latih dan 48 data uji. Pada penelitian ini mendapatkan hasil akurasi yang sangat tinggi antara 95,83% - 97,91% [19].

Pada tahun 2021 dilakukan penelitian oleh Candra Nur Ihsan dengan judul penelitian “Klasifikasi Data Radar Menggunakan Algoritma Convolutional Neuran Network (CNN) pada penelitian ini bertujuan untuk mengenali data pada radar cuaca yang tidak normal, menggunakan dataset dari sistem SANTANU dengan data latih sebanyak 800 data dan data uji sebanyak 50 data. Berdasarkan model yang dibuat mendapatkan hasil akurasi yang sangat tinggi sebesar 96,6% [14].

Masih di tahun 2020 dilakukan penelitian yang berjudul “Klasifikasi Aksara Jawa Dengan CNN” yang dilakukan oleh Edo Prasetyo N.A. Wijaya, Hendrawan Armanto dan Lukman Zaman bertujuan untuk membedakan aksara jawa. Pada penelitian ini menggunakan 20 huruf aksara tulisan tangan dengan jumlah 2470 dataset. Penelitian ini berhasil mengklasifikasikan aksara jawa dengan tingkat akurasi sebesar 70,22% [20].

Penelitian yang dilakukan oleh Hendry fonda, Yuda Irawan dan Anita Febriani pada tahun 2020 yang berjudul “Klasifikasi Batik Riau Dengan Menggunakan Convolutional Neural Networks (CNN)” . tujuan penelitian ini adalah untuk membedakan batik riau dengan batik dari daerah lain. Penelitian ini menggunakan 168 dataset batik diantaranya 68 merupakan gambar batik Riau dan 100 merupakan gambar yang bukan batik Riau. Masih dengan metode CNN klasifikasi yang dilakukannya mendapatkan hasil akurasi 65% [21].

Masih di tahun 2020 penelitian yang dilakukan oleh Mohammad Farid Naufal dengan judul “Analisis Perbandingan Algoritma SVM, KNN dan CNN

Untuk Klasifikasi Citra Cuaca”. Pada penelitian ini melakukan perbandingan kinerja metode SVM, KNN dan CNN pada citra cuaca sebanyak 4 kelas dengan total dataset sebanyak 1120 dataset. Berdasarkan hasil yang didapatkan berdasarkan peforma metode CNN mendapatkan akurasi yang paling baik dibandingkan dengan KNN dan SVM[22].

Dikuatkan oleh penelitian yang dilakukan oleh Bobby Kurniadi W., Hariyanto Prasetyo, Ghifari Ahmad L., Bagas Aditya Wibisono, dan Desta Sandya Prasvita pada tahun 2021 yang berjudul “Analisis Perbandingan Algoritma SVM dan CNN untuk Klasifikasi Buah”. Penelitian ini bertujuan untuk membandingkan hasil akurasi dari metode Suport Vektor Machine (SVM) dan *Convolutional Neural Network* (CNN) dengan menggunakan data dari kaggle sebanyak 17 jenis buah. Hasilnya menggunakan dua metode berbeda, SVM mendapatkan hasil akurasi 93,09% dan CNN mendapat hasil akurasi yang lebih tinggi dengan hasil akurasi 96,87. Berdasarkan penelitian ini metode CNN merupakan metode yang lebih baik dalam mengklasifikasikan citra gambar[23].

2.2 DASAR TEORI

2.2.1 Rambu Lalu Lintas

Menurut Kamus Besar Bahasa Indonesia (KBBI) kata rambu berarti tanda atau petunjuk[24]. Rambu dapat diartikan sebagai tanda untuk melarang, mengingatkan, menunjukkan dan memberitahukan tentang sesuatu makna untuk diikuti dan diketahui. Sedangkan berdasarkan Undang – Undang 22 Tahun 2009 tentang Lalu Lintas dan Angkutan Jalan menyatakan bahwa Rambu Lalu Lintas adalah bagian perlengkapan Jalan yang berupa lambang, huruf, angka, kalimat, dan atau perpaduan yang berfungsi sebagai peringatan, larangan, perintah, atau petunjuk bagi pengguna jalan[25]. Rambu lalu lintas sebagai alat pengendali lalu lintas yang berupa peringatan, larangan, perintah, dan petunjuk yang bertujuan untuk menjaga ketertiban dan kelancaran dalam berlalulintas bagi pengguna jalan. Terdapat beberapa jenis rambu lalu lintas yang ada di

jalan raya, masing-masing memiliki arti dan makna yang berbeda-beda, antara lain :

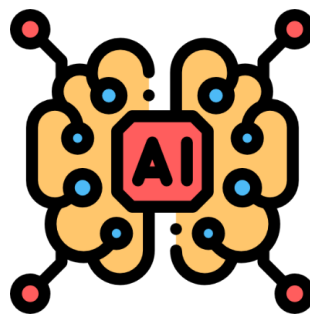
- a. Rambu Perintah, adalah rambu yang menyatakan perintah yang wajib ditaati oleh pengguna jalan, dimaksudkan untuk memberi petunjuk pendahuluan kepada pemakai jalan dan ditempatkan pada jarak yang layak sebelum titik kewajiban dimulai. Pada rambu ini, dasar palang rambu berwarna biru, sedangkan tulisan, angka, atau simbol pada rambu berwarna putih. Contohnya adalah rambu penanda tempat parkir atau jalur sepeda[26].
- b. Rambu Larangan, yaitu bentuk pengaturan yang jelas dan tegas. Sama halnya dengan rambu perintah, rambu larangan merupakan rambu yang berfungsi untuk melarang setiap pengguna jalan untuk melakukan suatu tindakan – tindakan di dalam ruang lalu lintas jalan, seperti dilarang parkir, dilarang berhenti dll [27].
- c. Rambu peringatan, Rambu peringatan digunakan untuk memberi peringatan kemungkinan ada bahaya di jalan atau tempat berbahaya pada jalan dan menginformasikan tentang sifat bahaya.
- d. Rambu anjuran, bentuk pengaturan yang merupakan himbauan yang boleh ataupun tidak dilakukan oleh pengguna jalan.
- e. Rambu anjuran, bentuk pengaturan yang merupakan himbauan yang boleh ataupun tidak dilakukan oleh pengguna jalan.
- f. Rambu petunjuk, merupakan rambu yang digunakan untuk memandu Pengguna Jalan saat melakukan perjalanan atau untuk memberikan informasi lain kepada Pengguna Jalan, seperti rambu petunjuk arah, jurusan, jarak, kondisi lokasi, tempat beribadah, rumah makan[28].



Gambar 2.1 Rambu Lalu Lintas[29]

2.2.2 *Artificial Intelligence*

Artificial Intelligence atau kecerdasan buatan adalah kecerdasan yang ditambahkan kepada suatu sistem yang bisa diatur dalam konteks ilmiah atau bisa juga disebut *Artificial Intelligence* atau hanya disingkat AI. Kecerdasan diciptakan dan dimasukkan kedalam suatu mesin atau komputer agar dapat melakukan pekerjaan seperti yang dilakukan oleh manusia[30]. *Artificial Intelligence* (AI) dirancang sama seperti manusia bahkan bisa melebihi manusia dalam mengolah data, menganalisis hingga memberikan keputusan dengan tingkat akurasi yang tinggi. *Artificial Intelligence* meliputi *Vision*, *Robotic*, *Speech*, Sistem Cerdas Tiruan, *Expert System* (Sistem Pakar), *Understanding*, dan bahasa Alami sebagai buatan yang cerdas dengan ketajaman dalam berpikir membekali dirinya dengan basis *Artificial Intelligence* (AI) akan bekerja memproses keseluruhan data dan memberikan hasil yang lebih sederhana dalam bentuk narasi maupun hasil yang kita butuhkan[31].

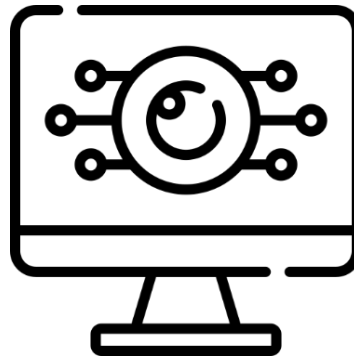


Gambar 2.2 *Artificial Intelligence*[32]

2.2.3 *Computer Vision*

Computer Vision adalah proses pembelajaran yang dilakukan oleh mesin dengan menganalisa gambar maupun video untuk memperoleh hasil sebagaimana yang bisa dilakukan manusia. Dengan kata lain, *Computer Vision* merupakan teknologi yang dapat meniru cara kerja visualisasi manusia[33]. *Computer Vision* mempunyai keterkaitan dengan beberapa bidang yaitu, *image processing* (pengolahan citra) dan

machine vision (visi mesin). Ada kesamaan yang signifikan dalam berbagai teknik dan aplikasi yang mencakup tiga bidang ini. Hal ini menunjukkan teknik dasar yang digunakan dan dikembangkan kurang lebih sama. Secara luas *Computer Vision* berhubungan dan diterapkan dengan bidang lain seperti *Artificial Intelligence* (AI), *robotika*, otomasi industri, pengolahan sinyal, optik fisik, *neurobiology*, dan lain-lain. *Computer Vision* sekarang ini telah sering digunakan untuk berbagai hal, contohnya saja mendeteksi wajah pada gambar (*face detection*), mengenali wajah (*facial expression recognition*) dan dalam prakteknya sering digunakan bersama dengan jaringan syaraf tiruan (*artificial neural network*)[34].



Gambar 2.3 *Computer Vision*[35]

2.2.4 Machine Learning

Machine Learning (ML) merupakan bidang studi yang fokus pada desain dan analisis algoritma sehingga memungkinkan komputer untuk dapat belajar. Menurut Samuel, ML berisi sebuah algoritma yang bersifat *generic* (umum) dimana algoritma tersebut dapat menghasilkan sesuatu yang menarik atau bermanfaat dari sejumlah data tanpa harus menulis kode yang spesifik. *Machine Learning* juga dapat diartikan sebuah komputer yang memiliki kemampuan belajar tanpa diprogram secara *eksplisit*. Program tersebut memanfaatkan data untuk membangun model dan mengambil keputusan berdasarkan model yang telah dibangun[36].

Machine Learning dirancang untuk secara otomatis menemukan prinsip dan keteraturan dengan pengalaman dari data yang dikumpulkan,

atau interaksi dengan *trial-and-error*. Secara garis besar *Machine Learning* dikelompokkan berdasarkan cara belajar sehingga dapat melakukan tugasnya. Pembagian *Machine Learning* berdasarkan klasifikasi dibagi menjadi tiga jenis yaitu :

- a. *Supervised Learning* : Dengan dataset pelatihan yang terdiri dari dari pasangan *input-dan-Output* berdasarkan sampel kategori yang diketahui. *Supervised Learning* bertujuan untuk membangun pemetaan dan hubungan fungsional antara *input* dan *Output* secara *implisit*. Umumnya, tugas-tugas pembelajaran yang diawasi meliputi klasifikasi dan regresi. Untuk klasifikasi, *Output* dari pasangan *input-dan-Output* dalam dataset pelatihan berkaitan dengan jumlah kategori diskrit yang terbatas untuk diberi label.
- b. *Unsupervised Learning*: Dibandingkan dengan *Supervised learning* dimana dataset adalah pasangan *input-dan-Output*, *Unsupervised Learning* tidak memiliki data *Output* untuk target pembelajaran selama proses pembelajaran. Pada *Unsupervised learning* hanya perlu mengumpulkan sampel yang sangat mirip.
- c. *Reinforcement Learning*: Berbeda dengan *Supervised Learning* dan *Unsupervised learning*, *Reinforcement Learning* tidak membutuhkan dataset pelatihan, Sinyal penguatan yang disediakan oleh lingkungan dalam *Reinforceent Learning* digunakan untuk mengevaluasi tindakan, hal ini bertujuan untuk menemukan tindakan yang sesuai[37].

Konsep dasar machine learning meliputi kemampuan suatu mesin dapat meningkatkan kecerdasan dalam belajar mandiri yang akan meningkatkan produktivitas manusia..

- a. Manfaat pembelajaran mesin dalam memprediksi.

Inti dari pembelajaran mesin ini adalah memprediksi sesuatu hal berdasarkan pola maupun faktor lain yang telah dilakukan training. Pengenalan ini dilakukan secara mudah tanpa menghabiskan banyak waktu.

b. Pembelajaran mesin membutuhkan pelatihan.

Untuk membuat suatu model maka diperlukan pemberian model pembelajaran mesin seperti apa hasil prediksi yang diinginkan.

c. Ketepatan >80% dianggap sukses.

Teknologi ini tidak dapat mengetahui dimana platform dapat mencapai akurasi 100% dengan mengidentifikasi suatu gambar. Aturan yang dikatakan ini adalah bahwa model dengan tingkat akurasi 80% adalah sebuah kesuksesan. Sebagaimana jika memikirkan betapa bergunanya untuk mengidentifikasi suatu gambar yang berjumlah 800.000 dengan benar, sementara “mungkin tidak” mendapatkan 200.000 gambar yang benar, maka masih menyimpan 80% dari waktu yang dimiliki. Hal itu merupakan perspektif nilai yang sangat besar[38].

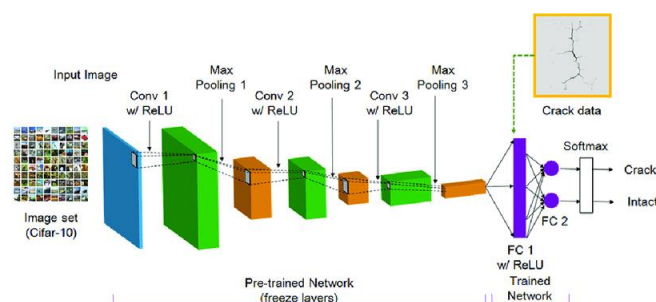


Gambar 2.4 *Machine Learning*[39]

2.2.5 Deep Learning

Pembelajaran mesin mendalam yang disebut dengan *Deep Learning* ialah salah satu pembelajaran dari berbagai macam metode pembelajaran mesin yang menggunakan *Convolutional Neural Networks* (CNN)[40]. Model *Deep Learning* dapat mempelajari komputasinya sendiri dengan menggunakan otaknya sendiri. *Deep Learning* dirancang untuk terus menganalisa data seperti pada otak manusia dalam mengambil keputusan. Agar kemampuan *Deep Learning* semakin mempuh maka *Deep Learning* menggunakan algoritma *artificial neural network* (ANN), yang terinspirasi dari jaringan biologis otak

manusia.[41]



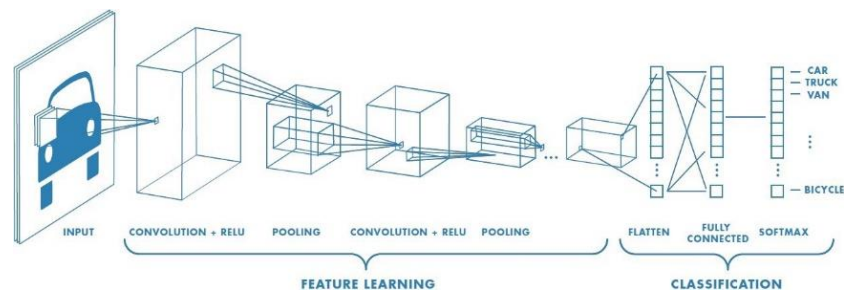
Gambar 2.5 Deep Learning Architecture[42]

2.2.6 Convolutional Neural Network

Convolutional Neural Network (CNN) adalah salah satu jenis neural network yang biasa digunakan untuk mengekstraksi fitur dari sebuah data gambar. CNN sangat mirip dengan jaringan saraf tiruan yang standar yang dapat divisualisasikan sebagai kumpulan neuron atau node atau unit yang disusun sebagai *graf asiklik* (graf yang tanpa adanya loop di dalamnya)[20]. CNN memiliki ciri khas yaitu terdapat lapisan tersembunyi atau *hidden layer* yang hanya terhubung ke subset neuron di lapisan sebelumnya. Karena konektivitas tersebut, CNN dapat mempelajari fitur secara *implisit*. Arsitektur CNN menghasilkan ekstraksi fitur hirarki yaitu *filter* yang dilatih untuk tujuan spesifik, misal pada lapisan pertama biasanya difokuskan pada identifikasi tepian atau *fluktuasi* warna, kemudian lapisan kedua biasanya lebih ke identifikasi bentuk, dan *filter* lapisan berikutnya biasanya lebih diarahkan untuk mempelajari bagian-bagian parsialisasi dari objek, baik yang terlihat sedikit atau sebagian maupun yang terlihat cukup banyak serta lapisan terakhir digunakan untuk mengidentifikasi objek[18].

Salah satu jenis *neural network* yang biasa digunakan pada data *image* adalah CNN. Karena dalamnya tingkat jaringan maka CNN termasuk dalam jenis *Deep Neural Network* dan sering digunakan dalam data citra. Ada dua metode yang dimiliki oleh CNN, yaitu klasifikasi menggunakan *feed forward* dan tahap pembelajaran menggunakan

backpropagation. Ada kesamaan struktur yang dimiliki CNN dengan *artificial neural network*. Pada klasifikasi citra, CNN mendapat masukan atau citra masukan untuk diproses dan diklasifikasi ke kategori tertentu. Perbedaan CNN dan ANN adalah pada arsitektur tambahan pada CNN yang dioptimisasi untuk fitur yang ada pada citra masukan.



Gambar 2.6 Arsitektur Algoritma CNN[41]

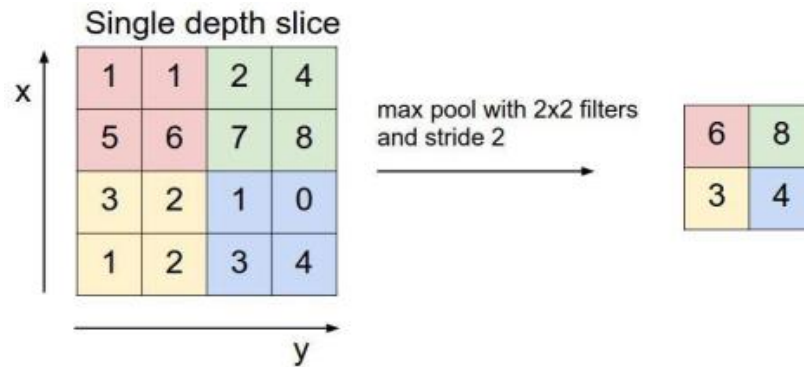
Ada beberapa komponen utama yang ada dalam CNN yaitu diantaranya adalah:

a. *Convolution Layer*

Convolution layer merupakan lapisan pertama pada CNN. Lapisan tersebut akan melakukan kovolusi citra masukan dengan *filter* yang telah didefinisikan tanpa merusak struktur citra awal. Fungsi *convolution layer* yaitu mengambil fitur pada citra yang akan digunakan untuk melatih model[43].

b. *Pooling Layer*

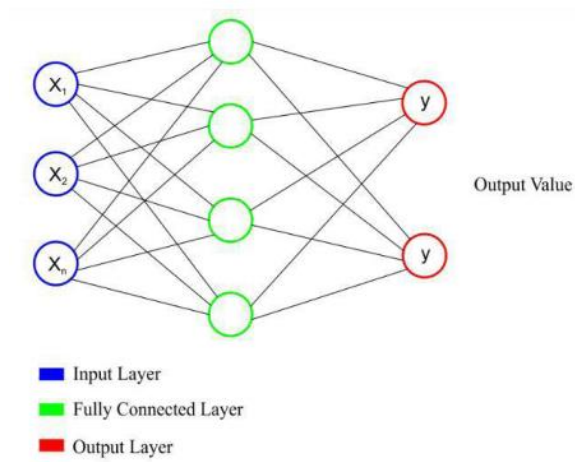
Pooling Layer berfungsi untuk menjaga ukuran data ketika *concolution*, yaitu dengan melakukan *downsampling* (prediksi sampel). *Pooling* dapat mereptesentasikan data menjadi lebih kecil, mudah dikelola, dan mudah mengontrol *overftting*. Proses *pooling* yang umum digunakan adalah *max pooling* yaitu memilih nilai maksimum dalam suatu area tertentu. Proses *pooling* lainnya yaitu *average pooling* yaitu menghitung nilai rata-rata dalam suatu area[44].



Gambar 2.7 Max Pooling Layer[41]

c. *Fully Connected Layer*

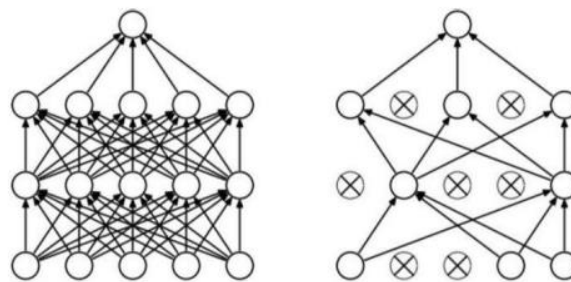
Layer tersebut merupakan *layer* yang biasanya dipakai dalam penerapan MLP dan bertujuan untuk melakukan transformasi pada dimensi data agar dapat diklasifikasikan secara *linear*. Disetiap neuron pada *convolution layer* perlu ditransformasi menjadi data satu dimensi terlebih dahulu sebelum dapat dimasukkan kedalam sebuah *Fully Connected Layer*. Dengan itu menyebabkan kehilangan data informasi spasialnya dan menjadi tidak *reversibel*. *Fully Connected Layer* hanya bisa diimplementasikan di akhir jaringan. *Convolution layer* dengan ukuran kernel 1 x 1 menjalankan fungsi yang sama dengan sebuah *Fully Connected Layer* namun dengan tetap mempertahankan karakter spasial dari data[45].



Gambar 2.8 Fully Connected Layer[46]

d. Dropout

Dropout adalah teknik regularisasi jaringan syaraf dimana beberapa neuron akan dipilih secara acak dan tidak dipakai selama pelatihan. Neuron-neuron ini dapat dibuang secara acak. Hal ini berarti bahwa kontribusi neuron yang dibuang akan diberhentikan sementara jaringan dan bobot baru juga tidak diterapkan pada neuron pada saat melakukan *backpropagation*. *Dropout* merupakan proses mencegah terjadinya *overfitting* dan juga mempercepat proses *learning*. *Dropout* mengacu kepada menghilangkan neuron yang berupa *hidden* maupun *layer* yang *visible* didalam jaringan. Dengan menghilangkan suatu neuron, berarti menghilangkannya sementara dari jaringan yang ada. Neuron yang akan dihilangkan akan dipilih secara acak. Setiap neuron akan diberikan probabilitas yang bernilai antara nol dan satu.



Gambar 2.9 Jaringan Syaraf Sebelum dan Sesudah Menggunakan Teknik *Dropout*[41]

2.2.7 Python

Python adalah bahasa pemrograman interpretatif yang berorientasi objek. pemrograman *imperatif* dan pemrograman *fungsional*. *Python* dapat dijalankan di berbagai platform sistem operasi, beberapa platform yang mendukung *Python* diantaranya Linux, Windows, Mac OS X, Java Virtual Machine. *Python* dikembangkan oleh Guido Van Rossum, programmer kelahiran Belanda, pada tahun 1991 di CWI, Amsterdam sebagai kelanjutan dari bahasa pemrograman ABC[47].



Gambar 2.10 Logo *Python Programming*[48]

Python mendukung *multi paradigma* pemrograman, namun tidak dibatasi pada pemrograman berorientasi objek, pemrograman *imperatif*, dan pemrograman *funksional*. Salah satu fitur yang tersedia pada *Python* adalah sebagai bahasa pemrograman dinamis yang dilengkapi dengan manajemen memori otomatis. Seperti halnya pada bahasa pemrograman dinamis lainnya, *Python* umumnya digunakan sebagai bahasa *Script* meski pada praktiknya penggunaan bahasa ini lebih luas mencakup konteks pemanfaatan yang umumnya tidak dilakukan dengan menggunakan bahasa *Script*[49]. Beberapa fitur dan kelebihan yang dimiliki *Python* adalah:

- a. Memiliki koleksi kepastakaan yang banyak. Artinya telah tersedia modul-modul siap pakai untuk berbagai keperluan, seperti pembuatan game hingga *Artificial Intelligence*.
- b. Memiliki struktur bahasa yang jelas, sederhana, dan mudah dipelajari.
- c. Berorientasi *prosedural* dan objek sekaligus *multi paradigma*.
- d. Memiliki pengelolaan memori otomatis *garbage collection*.
- e. Bersifat *modular* sehingga mudah dikembangkan dengan menciptakan modul-modul baru, baik dibangun dengan bahasa *Python* maupun C/C++[47].

Library pada *Python* merupakan gabungan dari sekumpulan package dan module dengan fungsionalitas yang sama dengan tujuan untuk memudahkan kalian dalam membuat suatu aplikasi, tanpa harus menulis banyak kode. Library pada *Python* merupakan sebutan untuk kode program tambahan yang digunakan dalam kebutuhan tertentu. *Python* mempunyai lebih dari 140.000 library yang dikembangkan

melalui open source project. Dengan menggunakan library pada *Python*, kamu akan dapat menghasilkan kode secara efisien dan menghemat waktu tanpa harus menulis seluruh skrip. Berikut ini beberapa contoh library yang umum digunakan oleh praktisi data, yaitu[50]:

a. TensorFlow

Tensorflow adalah pustaka pemrograman sumber terbuka di Python untuk membangun model klasifikasi citra. Perpustakaan ini dikembangkan oleh Tim Google Bran. Tensorflow terutama digunakan dalam tugas pembelajaran mesin. TensorFlow ditulis dalam C++, Python, CUDA. Kita dapat menggunakan CPU dan GPU untuk mengeksekusi kode di TensorFlow. Meskipun mengeksekusi kode dengan GPU jauh lebih cepat. TensorFlow memungkinkan Anda untuk melatih dan menerapkan model Anda dengan nyaman[51].



Gambar 2.11 Logo *Tensorflow*[52]

b. *Scikit-learn*

Scikit-learn adalah pembelajaran mesin yang paling komprehensif dan bersumber terbuka paket dengan Python. *Scikit-learn* merupakan sebuah paket pembelajaran mesin dalam bahasa pemrograman Python yang banyak digunakan dalam ilmu data. Paket *Scikit-learn* mencakup implementasi daftar lengkap metode pembelajaran mesin di bawah data terpadu dan konvensi prosedur pemodelan, menjadikannya alat yang nyaman bagi ahli statistik pendidikan dan perilaku. Karena pembelajaran mesin sering kali merupakan komponen yang lebih umum aplikasi (seperti layanan Web), diinginkan untuk dilengkapi dengan menggunakan yang sama bahasa pemrograman sebagai bagian

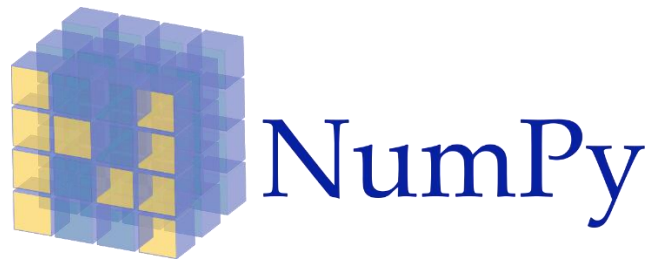
lain dari aplikasi untuk integrasi tanpa batas. Manfaat dari berbagai penerapan Python, Scikit-belajar menjadi sebuah paket yang semakin populer untuk aplikasi terkait pembelajaran mesin. Paket *Scikit-learn* mencakup empat topik utama yang terkait dengan pembelajaran mesin. Mereka adalah transformasi data, pembelajaran yang diawasi, pembelajaran tanpa pengawasan, dan evaluasi dan pemilihan model[53].



Gambar 2.12 Logo *Scikit Learn*[54]

c. Numpy

Numpy adalah pustaka Python untuk membantu dalam melakukan operasi matematika terutama ketika melakukan kalkulasi pada vector fungsi[55]. NumPy merupakan singkatan dari Numerical Python. NumPy merupakan salah satu library Python yang berfungsi untuk proses komputasi numerik. NumPy memiliki kemampuan untuk membuat objek N-dimensi array. Array merupakan sekumpulan variabel yang memiliki tipe data yang sama. Kelebihan dari NumPy Array adalah dapat memudahkan operasi komputasi pada data, cocok untuk melakukan akses secara acak, dan elemen array merupakan sebuah nilai yang independen sehingga penyimpanannya dianggap sangat efisien[56].

Gambar 2.13 Logo *Numpy*[57]

d. Keras

Keras adalah program aplikasi antarmuka (API) untuk membuat jaringan saraf, dikembangkan sesuai dengan prinsip panduan bahwa “mampu” untuk beralih dari ide ke hasil dengan penundaan sesedikit mungkin adalah kunci untuk melakukan penelitian yang baik”[58]Keras merupakan library jaringan syaraf tiruan tingkat tinggi yang ditulis dengan bahasa python dan mampu berjalan di atas TensorFlow, CNTK, atau Theano [5]. Library ini menyediakan fitur yang digunakan dengan fokus mempermudah pengembangan lebih dalam tentang *Deep Learning*[46].

Gambar 2.14 Logo *Keras*[59]

e. *PyTorch*

PyTorch adalah pustaka pembelajaran mesin sumber terbuka yang dibuat berdasarkan pustaka Torch. *PyTorch* digunakan untuk aplikasi seperti penglihatan komputer dan pemrosesan bahasa alami *PyTorch* adalah perangkat lunak bebas dan sumber terbuka yang dirilis di bawah lisensi BSD Modifikasi. Meskipun antarmuka Python lebih stabil dan merupakan fokus utama pengembangan, *PyTorch* juga memiliki antarmuka C++[60].



Gambar 2.15 Logo *Pytorch*[61]

f. Pandas

Pandas adalah pustaka Python yang penulis gunakan untuk membaca data, mengolah data dan normalisasi data. Pandas bisa membaca data csv dan kompatibel dengan Numpy untuk melakukan operasi aritmatika di dalamnya [55].



Gambar 2.16 Logo *Pandas*[62]

2.2.8 *Confusion Matrix*

Confusion Matrix adalah matriks untuk mengukur kinerja suatu klasifikasi dan berfungsi untuk menentukan persentase accuracy, precision dan recall[63]. *Confusion Matrix* memiliki 4 variabel yaitu TP (True Positive) berarti ketika sistem memprediksi *positive* dan hasilnya benar, TN (True Negative) berarti ketika sistem memprediksi negatif dan hasilnya benar, FP (False Positive) berarti ketika sistem memprediksi *positive* dan hasilnya salah, FN (False Negative) berarti ketika sistem memprediksi negatif dan hasilnya salah[64]. Pengukuran Kinerja dengan *Confusion Matrix* dapat menggunakan perhitungan sebagai berikut:

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+TN+FN)} \times 100\% \quad (2.1)$$

Keterangan :

TP = True *Positive*

TN = True Negative

FP = False *Positive*

FN = False Negative

		True Class	
		Positive	Negative
Predicted Class	Positive	true positives count (TP)	false negatives count (FP)
	Negative	false positives count (FN)	true negatives count (TN)

Gambar 2.17 *Confusion Matrix Table*[65]