

SKRIPSI

**SISTEM PREDIKSI *DWELLING TIME CONTAINERS* DI
TERMINAL PETI KEMAS (STUDI KASUS PELABUHAN
SOEKARNO HATTA MAKASSAR)**

Disusun dan diajukan oleh:

**TIARA YANIA IFANI LAKITA
D121 18 1022**



**PROGRAM STUDI SARJANA TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS HASANUDDIN
GOWA
2024**

LEMBAR PENGESAHAN SKRIPSI

**SISTEM PREDIKSI DWELLING TIME CONTAINERS DI
TERMINAL PETIKEMAS (STUDI KASUS PELABUHAN
SOEKARNO HATTA MAKASSAR)**

Disusun dan diajukan oleh

**Tiara Yania Ifani Lakita
D121181022**

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian
Studi Program Sarjana Program Studi Teknik Informatika
Fakultas Teknik Universitas Hasanuddin
Pada tanggal 7 Juni 2024
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

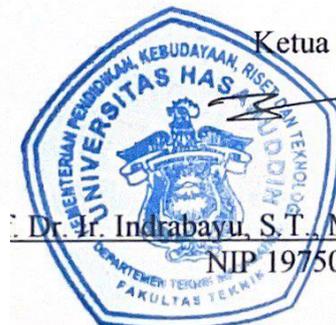
Pembimbing Utama,

Pembimbing Pendamping,


Prof. Dr. Ir. Indrabayu, S.T., M.T.,
M.Bus.Sys., IPM., ASEAN.Eng.
NIP 19750716 200212 1 004


A. Ais Prayogi Alimuddin,
S.T., M.Eng.
NIP 19830510 201404 1 001

Ketua Program Studi,




Dr. Ir. Indrabayu, S.T., M.T., M.Bus.Sys., IPM., ASEAN.Eng.
NIP 19750716 200212 1 004



PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini ;

Nama : Tiara Yania Ifani Lakita
NIM : D121181022
Program Studi : Teknik Informatika
Jenjang : S1

Menyatakan dengan ini bahwa karya tulisan saya berjudul

Sistem Prediksi Dwelling Time Containers di Terminal Petikemas (Studi Kasus
Pelabuhan Soekarno Hatta Makassar)

Adalah karya tulisan saya sendiri dan bukan merupakan pengambilan alihan tulisan orang lain dan bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri.

Semua informasi yang ditulis dalam skripsi yang berasal dari penulis lain telah diberi penghargaan, yakni dengan mengutip sumber dan tahun penerbitannya. Oleh karena itu semua tulisan dalam skripsi ini sepenuhnya menjadi tanggung jawab penulis. Apabila ada pihak manapun yang merasa ada kesamaan judul dan atau hasil temuan dalam skripsi ini, maka penulis siap untuk diklarifikasi dan mempertanggungjawabkan segala resiko.

Segala data dan informasi yang diperoleh selama proses pembuatan skripsi, yang akan dipublikasi oleh Penulis di masa depan harus mendapat persetujuan dari Dosen Pembimbing.

Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan isi skripsi ini hasil karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Gowa, 7 Juni 2024

Yang Menyatakan



Tiara Yania Ifani Lakita



ABSTRAK

TIARA YANIA IFANI LAKITA. *Sistem Prediksi Dwelling Time Containers di Terminal Petikemas (Studi Kasus Pelabuhan Soekarno Hatta Makassar)* (dibimbing oleh Indrabayu dan A. Ais Prayogi Alimuddin)

Kinerja logistik di Indonesia masih tertinggal dibandingkan dengan negara tetangga seperti Singapura, Malaysia, Thailand, Vietnam, dan Filipina. Indeks Kinerja Logistik (LPI) menunjukkan bahwa kinerja logistik Indonesia belum efisien. Data terbaru tahun 2023 menunjukkan bahwa Indonesia memiliki kinerja yang lebih buruk dibandingkan dengan negara-negara tersebut dalam semua kategori LPI. Salah satu masalah yang signifikan adalah *dwelling time*, yaitu lamanya waktu bongkar muat kontainer di pelabuhan. Rata-rata *dwelling time* di Terminal Petikemas Pelabuhan Soekarno Hatta Makassar melebihi standar pemerintah yang ditetapkan sebesar 3 hari. Hal ini mempengaruhi produktivitas perusahaan pengiriman barang melalui kontainer. Untuk mengatasi masalah ini, diperlukan pengujian variabel-variabel tertentu dalam pengembangan model sistem prediksi *dwelling time* kontainer di Terminal Petikemas (Studi Kasus: Pelabuhan Soekarno Hatta Makassar)", yang dapat membantu operator terminal dalam mengelola sumber daya secara efisien dan menerapkan kebijakan yang tepat. Variabel-variabel tersebut antara lain Pelabuhan Asal, Status Kontainer, Komoditas, Jenis Kontainer, Ukuran Kontainer, *Gross* dan *Dwelling Time* sebanyak 8.181 baris data. Metode prediksi yang digunakan dalam penelitian ini adalah ANN dan SVR dengan matriks perhitungan *error* menggunakan RMSE dan MAE. Hasil penelitian didapatkan RMSE 1,85 dan MAE 1,3 pada metode SVR serta RMSE 1,73 dan MAE 1,25 pada metode ANN. Hal ini menunjukkan bahwa kedua metode belum dapat memprediksi *dwelling time* dengan akurat dikarenakan kualitas data yang dimiliki.

Kata kunci: *Dwelling Time, Container, SVR, ANN*



ABSTRACT

TIARA YANIA IFANI LAKITA. *Prediction System for Dwelling Time of Containers at Container Terminal (Case Study: Soekarno Hatta Port in Makassar) (supervised by Indrabayu and A. Ais Prayogi Alimuddin)*

The logistics performance in Indonesia still lags behind neighboring countries such as Singapore, Malaysia, Thailand, Vietnam, and the Philippines. The Logistics Performance Index (LPI) indicates that Indonesia's logistics performance is not yet efficient. The latest data from 2023 shows that Indonesia performs worse than these countries in all LPI categories. One of the significant problem is dwelling time, whis is the length of time for loading and unloading containers at the port. The average dwelling time at the Makassar Soekarno Hatta Port Container Terminal exceeds the government standard set at 3 days. This affects the productivity of companies shipping goods via containers. To overcome this problem, it is necessary to test certain variables in developing a container dwelling time prediction system model at the Container Terminal (Case Study: Soekarno Hatta Port, Makassar)", which can help terminal operators manage resources efficiently and implement appropriate policies. These variables include Port of Origin, Container Status, Commodity, Container Type, Container Size, Gross and Dwelling Time totaling 8,181 rows of data. The prediction method used in this research is ANN and SVR with an error calculation matrix using RMSE and MAE. Results research obtained RMSE 1.85 and MAE 1.3 for the SVR method and RMSE 1.73 and MAE 1.25 for the ANN method. This shows that the two methods cannot predict dwelling time accurately due to the quality of the data they have.

Keywords: Dwelling Time, Container, SVR, ANN



DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	iv
PERNYATAAN KEASLIAN.....	iv
ABSTRAK	i
ABSTRACT	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR	vii
DAFTAR TABEL.....	viii
DAFTAR SINGKATAN DAN ARTI SIMBOL	ix
DAFTAR LAMPIRAN.....	x
KATA PENGANTAR	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian/Perancangan	3
1.4 Manfaat Penelitian/Perancangan	3
1.5 Ruang Lingkup/Asumsi perancangan.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Petikemas (Container)	5
2.2 <i>Dwelling Time</i>	5
2.3 Jenis Kontainer	6
2.4 Ukuran Kontainer	6
2.5 Status Kontainer	7
2.6 Komoditi	7
2.7 <i>Gross</i>	7
2.8 Pelabuhan Asal	8
2.9 <i>Knowledge Discovery in Database</i>	8
2.10 <i>Machine Learning</i>	16
2.11 Peramalan (<i>forecasting</i>).....	19
2.12 <i>Support Vector Regression</i>	20
2.13 <i>Artificial Neural Network</i>	25
2.14 RMSE	30
2.15 Penelitian Terkait.....	31
BAB 3 METODE PENELITIAN/PERANCANGAN	35
3.1 Tahapan Penelitian	35
3.2 Waktu dan Lokasi Penelitian	35
3.3 Instrumen Penelitian	35
3.4 Teknik Pengambilan Data	36
3.5 Perancangan Sistem.....	36
3.6 Prediksi/Peramalan Menggunakan <i>Support Vector Machine</i>	44
3.7 Prediksi/Peramalan Menggunakan <i>Artificial Neural Network</i>	48
BAB 4 HASIL DAN PEMBAHASAN.....	53
4.1 Dataset	53
4.2 Penerapan Algoritma <i>Support Vector Regression</i>	57
4.3 Penerapan Algoritma <i>Artificial Neural Network</i>	61
4.4 Pembahasan	64



BAB 5. KESIMPULAN DAN SARAN 70
5.1 Kesimpulan..... 70
5.2 Saran 71
DAFTAR PUSTAKA 72



DAFTAR GAMBAR

Gambar 1 Proses Knowledge Discovery in Database.....	9
Gambar 2 Map pengembangan produk App	17
Gambar 3 Hyperplane dan margin pada SVR.....	21
Gambar 4 Insensitive zone	22
Gambar 5 Ilustasi <i>hyperplane</i> pada dimensi yang lebih tinggi	23
Gambar 6 Jaringan Syaraf dengan lapisan tunggal	27
Gambar 7 Jaringan Syaraf Tiruan lapisan banyak	27
Gambar 8 Jaringan Saraf dengan lapisan kompetitif yang memiliki $-\eta$	28
Gambar 9 Struktur ANN	29
Gambar 10 Sampel data Inbound List.....	37
Gambar 11 Sampel data Terminal In & Out	37
Gambar 12 Informasi nilai null pada setiap feature	40
Gambar 13 Median gross untuk setiap ukuran kontainer.....	40
Gambar 14 Banyaknya data hasil data <i>cleaning</i>	42
Gambar 15 Banyaknya data hasil <i>outlier handling</i>	43
Gambar 16 Hasil <i>GridSearch</i> SVR.....	45
Gambar 17 <i>Bias term</i> dari metode SVR.....	46
Gambar 18 Informasi persebaran data (<i>data distribution</i>).....	54
Gambar 19 Hasil uji ANOVA dataset.....	57
Gambar 20 Hasil <i>tuning</i> parameter SVR	59
Gambar 21 Hasil perhitungan nilai error saat evaluasi model SVR	61
Gambar 22 Hasil perhitungan nilai error saat evaluasi model ANN	64



DAFTAR TABEL

Tabel 1 Rumus dari setiap tipe kernel pada SVR	14
Tabel 2 Rumus dari setiap tipe kernel pada SVR	24
Tabel 3 Data yang dipilih dari sumber data <i>Inbound List</i>	38
Tabel 4 Data yang dipilih dari sumber data <i>Inbound List</i>	38
Tabel 5 Hasil penggabungan dataset	39
Tabel 6 Hasil pembersihan <i>missing value</i>	41
Tabel 7 Hasil <i>data cleaning</i>	42
Tabel 8 Hasil <i>outlier handling</i>	43
Tabel 9 Hasil transformasi data	44
Tabel 10 <i>Support vectors</i> dari metode SVR	45
Tabel 11 Sampel data support vector, Lagrange multiplier, and target value	46
Tabel 12 Sampel data yang dipakai untuk melakukan prediksi	53
Tabel 13 Sampel data yang memiliki perbedaan pola kombinasi data	56
Tabel 14 Informasi korelasi data (<i>data correlation</i>)	56
Tabel 15 Pembagian data training dan data testing pada metode SVR	58
Tabel 16 Hasil <i>training</i> pada prediksi menggunakan metode SVR	59
Tabel 17 Hasil prediksi menggunakan metode SVR	60
Tabel 18 Hasil scaling data pada metode ANN	62
Tabel 19 Pembagian data <i>training</i> dan data testing pada metode ANN	62
Tabel 20 Hasil <i>training</i> pada prediksi dengan menggunakan metode ANN	63
Tabel 21 Hasil prediksi menggunakan metode ANN	63



DAFTAR SINGKATAN DAN ARTI SIMBOL

Lambang/Singkatan	Arti dan Keterangan
LPI	<i>Logistik Performance Indeks</i>
INSA	<i>Indonesian National Shipowners Association</i>
AIMP	<i>Artificial Intelligence and Multimedia Processing</i>
SVR	<i>Support Vector Regression</i>
ANN	<i>Artificial Neural Network</i>
ISO	<i>Internasional Standar Organization</i>
VGM	<i>Verified Gross Mass</i>
KDD	<i>Knowledge Discovery In Database</i>
ML	<i>Machine Learning</i>
ReLU	<i>Rectified Linear Unit</i>
NB	<i>Naïve Bayes</i>
CDT	<i>Container Dwelling Time</i>
ANOVA	<i>Analysis of Varians</i>
.csv	<i>Comma Separated Values File</i>
RBF	<i>Radial Basis Function</i>
C	Parameter reguralisasi
RMSE	<i>Root Mean Square Error</i>
MAE	<i>Mean Absolute Error</i>



DAFTAR LAMPIRAN

Lampiran 1 Dataset	76
Lampiran 2 <i>Source Code</i>	77
Lampiran 3 Dataset Hasil <i>Preprocessing</i>	94
Lampiran 4 Hasil Prediksi.....	95



KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Allah SWT atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan tugas akhir dengan judul “Sistem Prediksi *Dwelling Time Containers* di Terminal Petikemas (Studi Kasus Pelabuhan Soekarno Hatta Makassar)”.

Penelitian ini dilatarbelakangi pengamatan penulis terhadap beberapa ketidaksesuaian lamanya waktu singgah kontainer di Pelabuhan Soekarno-Hatta, yang mana belum sesuai dengan aturan dengan aturan dari pemerintah. Sehingga terjadi berbagai masalah logistik yang dapat menghambat jalannya perekonomian di jalur laut. Maka dari itu, penulis bermaksud untuk membuat model prediksi *dwelling time* kontainer di Pelabuhan Soekarno Hatta, dengan harapan dapat membantu operator terminal maupun pihak-pihak terkait untuk dapat mengelola kebijakan yang tepat untuk Pelabuhan yang bersangkutan.

Penulis menyadari bahwa penyusunan dan penulisan tugas akhir ini tidak dapat terselesaikan dengan baik tanpa adanya bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih banyak kepada :

1. Keluarga penulis, Bapak Abubakar A.J. Lakita dan Ibu Titik Priyati selaku kedua orang tua penulis, Tika Septiana Lakita dan Tisa Aprilia Ramadhani Lakita selaku kakak kandung penulis yang senantiasa memberikan doa dan dukungan yang tiada hentinya, serta selalu sabar dan tanpa kenal lelah dalam mendidik penulis selama ini.
2. Bapak Prof. Dr. Ir. Indrabayu, S.T., M.T., M.Bus.Sys., IPM, ASEAN. Eng. selaku pembimbing I dan Bapak A. Ais Prayogi Alimuddin, S.T., M.Eng. selaku pembimbing II, yang senantiasa menyediakan waktu, tenaga, pikiran dan perhatian yang luar biasa dalam mengarahkan penulis untuk menyelesaikan tugas akhir.
3. Segenap staf dan dosen Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin yang telah membantu untuk kelancaran penyelesaian tugas akhir.
4. Bapak Asrullah yang telah senantiasa membantu penulis selama pengambilan data penelitian di Komando Pusat Otoritas Terminal Petikemas Makassar.
5. Segenap keluarga AIMP *Research Group* Universitas Hasanuddin yang telah memberikan bantuan selama penelitian dan diskusi terkait penyusunan tugas akhir serta memberikan motivasi di masa-masa sulit penyelesaian skripsi.
6. Maulana Bagaswara Marsidi yang selalu memberi bantuan, nasihat dan dukungan yang tulus selama penyelesaian tugas akhir, serta Lilo yang selalu menemani dan menghibur dalam setiap kesempatan.
7. Segenap teman seperjuangan tugas akhir, Indah, Dee, Caca, Lisyia, Aqilah, Lisyia, Mage, kak Imad yang telah membantu memberikan pertolongan, semangat, serta hiburan di masa-masa sulit penyelesaian tugas akhir. Serta pihak-pihak lain yang tidak sempat disebutkan dan tanpa sadar telah menjadi inspirasi maupun membantu penulis dalam menyelesaikan tugas akhir.



BAB I PENDAHULUAN

1.1 Latar Belakang

Indonesia berada pada posisi dimana kinerja logistiknya masih tertinggal dari beberapa negara seperti Singapura, Malaysia, Thailand, Vietnam dan Filipina. Diukur dari komponen Logistik Performance Index (LPI), World Bank menyebutkan bahwa kinerja logistik Indonesia belum efisien. Berdasarkan data terakhir pada tahun 2023, dari keenam kategori dalam LPI, kinerja Indonesia lebih buruk dibandingkan kelima negara tersebut pada semua kategori. Dan dalam beberapa kategori seperti pengiriman internasional, kompetensi logistik, pelacakan dan penelusuran serta ketepatan waktu, Indonesia mengalami penurunan signifikan jika dibandingkan dengan tahun 2018.

Jika dilihat dari Logistics Performance Index (LPI), permasalahan utama tingginya biaya logistik nasional disebabkan dari masalah infrastruktur yang berkontribusi terhadap kelancaran barang di pelabuhan. Hal ini merupakan hambatan di bidang logistik Indonesia yang berdampak pada melemahnya daya saing nasional (Utami, 2015). *Indonesian National Shipowners Association (INSA)* menyebutkan kunci penurunan biaya logistik terdapat pada peningkatan efisiensi dan efektivitas distribusi barang sektor angkutan laut melalui pelabuhan. Sedangkan rendahnya daya saing infrastruktur, memberikan kontribusi terhadap kurang lancarnya arus distribusi barang. Kondisi infrastruktur di pelabuhan khususnya pada Pelabuhan Soekarno Hatta yang ada sekarang ini dinilai masih kurang memadai.

Lemahnya dukungan sektor logistik nasional tersebut memicu berbagai permasalahan dalam distribusi barang dikarenakan kurangnya efisiensi pelayanan kepabeanan serta infrastruktur terutama terkait masalah lamanya waktu bongkar muat barang di pelabuhan (*dwelling time*). *Dwelling time* merupakan ukuran waktu yang dibutuhkan kontainer sejak kontainer dibongkar dari waktu kapal sandar

) sampai dengan keluar dari kawasan pelabuhan (*gate out*). Selama ini, *dwelling time* di Terminal Petikemas Pelabuhan Soekarno Hatta Makassar mencapai 5-6 hari, yang mana belum memenuhi standar yang ditetapkan oleh



pemerintah yakni maksimal 3 hari. Sedangkan permasalahan perusahaan yang bergerak di bidang jasa pengiriman barang melalui kontainer adalah produktivitas proses bongkar dan proses muat kontainer di pelabuhan yang dipengaruhi proses *dwelling time*. Tingginya nilai *dwelling time* terjadi karena beberapa faktor salah satu yang paling dominan adalah ketidaksiapan pelabuhan dalam mengantisipasi arus barang dan masalah operasional lainnya.

Berkaitan dengan aturan pemerintah, permasalahan mengenai *dwelling time* ini perlu dikaji agar dapat mengevaluasi proses bongkar muat kontainer. Hal ini penting dilakukan agar tidak terjadi penumpukan kapal yang ingin berlabuh di pelabuhan Soekarno Hatta Makassar. Karena penumpukan kapal dapat menjadi reaksi berantai untuk kapal-kapal setelahnya sehingga membuat jalannya operasional dari seluruh pihak terkait menjadi tidak teratur.

Di sisi lain, terdapat berbagai parameter menyangkut kontainer yang diasumsikan dapat mempengaruhi *dwelling time* antara lain hari dan bulan keluar, pelabuhan asal, ukuran dan jenis kontainer, serta *yard occupation ratio* terminal. Maka dari itu diperlukan untuk menguji berbagai parameter-parameter yang ada untuk memprediksi *dwelling time*. Hal ini diharapkan dapat menjadi bahan masukan yang dapat membantu operator terminal untuk dapat mengelola sumber daya mereka secara efisien dan menerapkan kebijakan yang tepat pada waktu dan tempat yang sesuai. Untuk itu diusulkan judul “**Sistem Prediksi *Dwelling Time Containers* di Terminal Petikemas (Studi Kasus Pelabuhan Soekarno Hatta Makassar)**” sebagai bentuk evaluasi pengelolaan kontainer di Terminal Petikemas Pelabuhan Soekarno Hatta Makassar.

1.2 Rumusan Masalah

Berdasarkan pada latar belakang masalah di atas, maka perumusan masalah dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana tahapan prediksi *dwelling time* dengan metode-metode *forecasting* menggunakan parameter-parameter yang diolah dari Terminal Petikemas Pelabuhan Soekarno Hatta Makassar?



- b. Bagaimana perbandingan hasil prediksi *dwelling time* dengan metode-metode *forecasting* di Terminal Petikemas Pelabuhan Soekarno Hatta Makassar?
- c. Apakah parameter-parameter yang digunakan kompatibel untuk memprediksi *dwelling time* di Terminal Petikemas Pelabuhan Soekarno Hatta Makassar?

1.3 Tujuan Penelitian/Perancangan

Tujuan yang diharapkan dalam penelitian ini adalah sebagai berikut:

- a. Menjelaskan tahapan prediksi *dwelling time* dengan metode-metode *forecasting* menggunakan parameter-parameter yang diolah dari Terminal Petikemas Pelabuhan Soekarno Hatta Makassar.
- b. Menganalisa dan membandingkan hasil prediksi metode-metode *forecasting* menggunakan parameter-parameter yang diolah Terminal Petikemas Pelabuhan Soekarno Hatta Makassar .
- c. Menyimpulkan kompatibilitas parameter-parameter yang ada terhadap prediksi *dwelling time* di Terminal Petikemas Pelabuhan Soekarno Hatta Makassar.

1.4 Manfaat Penelitian/Perancangan

Dengan dilakukannya penelitian ini, diharapkan dapat memberikan manfaat sebagai berikut:

- a. Akademisi
 - i. Penelitian ini diharapkan dapat menjadi referensi bagi akademisi maupun peneliti berikutnya dalam melakukan penelitian yang terkait dengan tema yang diteliti.
- b. Perusahaan
 - i. Penelitian ini diharapkan dapat menjadi bahan masukan yang berkaitan dengan judul yang diteliti.
 - ii. Diharapkan dapat menjadi acuan bagi operator terminal dapat mengelola sumber daya mereka secara efisien dan menerapkan kebijakan yang tepat pada waktu dan tempat yang sesuai.



1.5 Ruang Lingkup/Asumsi perancangan

- a. Data yang digunakan adalah data Terminal Petikemas Makassar, menggunakan dengan data yang dimiliki oleh sistem terapan baru perekaman data-data kapal kontainer di terminal petikemas yang bersangkutan.
- b. Pembangunan sistem akan menggunakan metode SVR dan ANN, dan menghasilkan model yang akan dipakai untuk memprediksi *dwelling time*
- c. Variabel-variabel penelitian terdiri dari:
 - i. Ukuran Kontainer
 - ii. Tipe/Jenis Kontainer
 - iii. Status Kontainer
 - iv. Komoditas
 - v. Pelabuhan Asal
 - vi. *Gross*
 - vii. *Dwelling Time*.



BAB II TINJAUAN PUSTAKA

2.1 Petikemas (Container)

Petikemas adalah peti berbentuk empat persegi panjang yang dirancang khusus dengan ukuran tertentu terbuat dari besi maupun aluminium serta memiliki pintu di salah satu sisinya serta dapat digunakan berulang kali juga digunakan sebagai tempat untuk menyimpan sekaligus mengangkut muatan yang ada di dalamnya dan telah ditetapkan berdasarkan standar internasional (ISO).

Filosofi dibalik petikemas adalah membungkus atau membawa muatan dalam peti-peti yang sama dan membuat semua kendaraan dapat mengangkutnya sebagai satu kesatuan, baik kendaraan itu berupa kapal laut, kereta api, truk, atau angkutan lainnya, dan dapat membawa secara cepat, dan efisien atau bila mungkin dari pintu ke pintu (*door to door*) (Afiffudin, 2019).

2.2 *Dwelling Time*

Dwelling time memiliki beberapa pengertian berdasarkan berbagai sumber karena definisi dwell time sendiri dapat berasal dari aspek manapun khususnya aspek pengangkutan. *Dwelling time* merupakan gambaran umum kecepatan proses pelayanan logistik dalam Pelabuhan sejak dilakukan pembongkaran barang impor (*discharge*) sampai dengan pengeluaran barang impor dari pelabuhan (Ramiaji, 2017).

Menurut *World Bank* (2011), *dwelling time* adalah waktu yang dihitung mulai dari suatu petikemas dibongkar muat dari kapal sampai petikemas tersebut meninggalkan terminal pelabuhan melalui pintu utama. Sedangkan bea & cukai berpendapat bahwa *dwelling time* adalah waktu berapa lama petikemas barang impor ditimbun di tempat penimbunan sementara di pelabuhan sejak dibongkar dari kapal sampai dengan barang impor keluar dari tempat penimbunan sementara (Azizah, 2019).



2.3 Jenis Kontainer

Jenis kontainer adalah media kemas yang dibuat dengan menyesuaikan fungsi, ukuran serta kegunaannya. Menurut Edy Hidayat pada tahun 2009 dalam bukunya, beberapa jenis petikemas antara lain:

- a. *General Cargo* adalah petikemas yang dipakai untuk mengangkut muatan umum (*General Cargo*).
- b. *Thermal Container* adalah petikemas yang dilengkapi dengan pengaturan suhu untuk muatan tertentu.
- c. *Tank Container* adalah tangkai yang ditempatkan dalam kerangka petikemas yang dipergunakan untuk muatan cair (*bulk liquid*) maupun gas (*bulk gas*).
- d. *Dry Bulk Container* adalah kontainer umum yang dipergunakan khusus untuk mengangkut muatan curah (*bulk cargo*).
- e. *Platform Container* adalah petikemas yang terdiri dari lantai dasar.

2.4 Ukuran Kontainer

Menurut Edy Hidayat pada tahun 2009, agar pengoperasian petikemas dapat berjalan dengan baik, maka semua pihak yang terlibat harus menyetujui agar ukuran-ukuran petikemas harus sama dan sejenis serta mudah diangkut. Adapun Badan *Internasional Standar Organization* (ISO) telah menetapkan ukuran-ukuran dari petikemas sebagai berikut (Afiffudin, 2019):

- a. *Container 20' Dry Freight (20 Feet)*

Ukuran Luarnya : 20' [p] x 8' [l] x 8,6' [t] atau 6.058 x 2.438 x 2.591M.

Ukuran Dalamnya : 5.919 x 2.340 x 2.380 M.

Kapasitasnya : *Cubic Capacity* : 33 Cbm. Pay Load : 22.1 Ton.

- b. *Container 40' Dry Freight (40 Feet)*

Ukuran Luarnya : 40' [p] x 8' [l] x 8,6' [t] atau 12.192 x 2.438 x 2.591M.

Ukuran Dalamnya : 12.045 x 2.309 x 2.379 M.

Kapasitasnya : *Cubic Capacity* : 67.3 Cbm.

Pay Load : 27.396 Ton

- c. *Container 40' High Cube Dry*

Ukuran Luarnya : 40' [p] x 8' [l] x 9,6' [t] atau 12.192 x 2.438 x 2.926M.



Ukuran Dalamnya : 12.056 x 2.347 x 2.684 M.

Kapasitasnya : *Cubic Capacity* : 76 Cbm.

Pay Load : 29.6 Ton.

2.5 Status Kontainer

Status kontainer adalah status atau keadaan kontainer yang dimuat. Dalam hal ini ada 2 status kontainer yang umumnya didokumentasikan, yaitu (Anggia, 2019):

- a. *Full*, yaitu keadaan dimana kontainer terisi penuh dengan berbagai komoditas.
- b. *Empty*, yaitu keadaan kontainer yang kosong karena tidak ada komoditas yang dimuat. Biasanya kontainer ini dimuat di kapal untuk dikembalikan setelah pembongkaran dan tidak ada komoditas yang dimuat kembali. Adapun untuk bongkaran kontainer *empty* biasanya langsung dipindahkan oleh pemilik kontainer ke depo kontainer, karena menghindari biaya penimbunan di CY yang cukup tinggi dan progresif.

2.6 Komoditi

Menurut UU RI No. 10 Tahun 2011 tentang Perdagangan Berjangka Komoditi, komoditi adalah semua barang, jasa, hak dan kepentingan lainnya, dan setiap derivatif dari Komoditi, yang dapat diperdagangkan dan menjadi subjek Kontrak Berjangka, Kontrak Derivatif Syariah, dan/atau Kontrak Derivatif lainnya.

2.7 Gross

Berat Kotor Petikemas (*Gross Weight*) adalah jumlah keseluruhan berat kotor dari kemasan petikemas, yaitu berat kontainer ditambah dengan berat komoditi atau isi di dalamnya. Contoh berat barang 500 Kg dan berat kontainer 1000 Kg maka G.W : 1.500 Kg.

Adapun yang dinamakan Berat Kotor Petikemas Terverifikasi (*Verified Gross Mass/ VGM*) sebagaimana yang ditetapkan dalam Peraturan Menteri Perdagangan Republik Indonesia Nomor Pm 53 Tahun 2018 Tentang Kelaikan dan Berat Kotor Petikemas Terverifikasi. VGM adalah jumlah keseluruhan berat kotor dari kemasan petikemas yang diperoleh melalui salah satu



metode yang ditetapkan (Peraturan Menteri Perhubungan Republik Indonesia Nomor PM 25 Tahun 2022 tentang Kelaikan Petikemas dan Berat Kotor Petikemas Terverifikasi).

2.8 Pelabuhan Asal

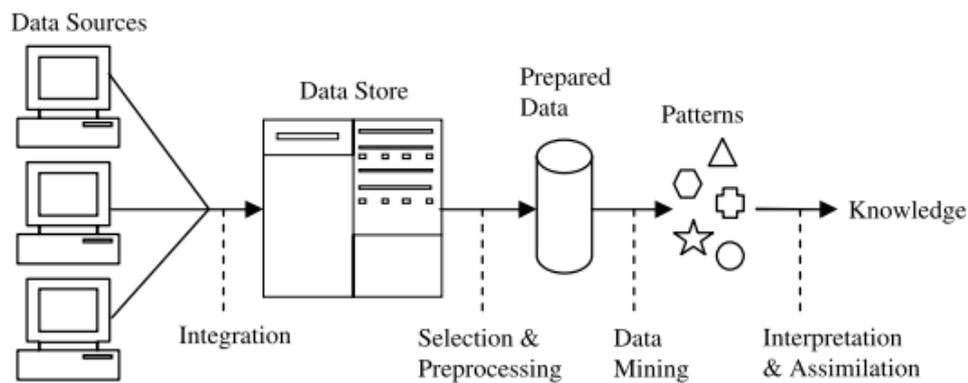
Pelabuhan adalah daerah perairan yang terlindung terhadap gelombang, yang dilengkapi dengan fasilitas terminal laut meliputi dermaga dimana kapal dapat bertambat untuk bongkar muat barang, kran-kran (*Cren*) untuk bongkar muat barang, gudang laut dan tempat-tempat menyimpan dimana kapal membongkar muatannya, dan gudang-gudang dimana barang-barang dapat disimpan dengan waktu yang lebih lama selama menunggu pengirim ke daerah tujuan atau pengapalan (Triatmojo, 2010).

Adapun pelabuhan asal atau *original port* atau *port of loading* yang disebutkan dalam penelitian ini adalah pelabuhan dimana kontainer pertama kali dimuat atau pelabuhan asal kontainer yang bersangkutan. Selain itu ada namanya *port of discharge* yang merupakan pelabuhan tempat kontainer yang bersangkutan dibongkar.

2.9 Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) merupakan metode untuk memperoleh pengetahuan dari *database* yang ada. Dalam *database* terdapat tabel - tabel yang saling berhubungan / berelasi. Hasil pengetahuan yang diperoleh dalam proses tersebut dapat digunakan sebagai basis pengetahuan (*knowledge base*) untuk keperluan pengambilan keputusan. Istilah *Knowledge Discovery in Database* (KDD) dan *data mining* seringkali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain, dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining* (Mardi, 2014).





Gambar 1 Proses Knowledge Discovery in Database

Sumber: Bramer, 2020

Jika mengacu pada Gambar 1, *Knowledge Discovery in Database* menyangkut seluruh proses ekstraksi pengetahuan, termasuk bagaimana data disimpan dan diakses, bagaimana menggunakan algoritma yang efisien dan terukur untuk menganalisis dataset besar, bagaimana menafsirkan dan memvisualisasikan hasil, dan bagaimana memodelkan dan mendukung interaksi antara manusia dan mesin. Ini juga menyangkut dukungan untuk mempelajari dan menganalisis domain aplikasi. Proses KDD secara garis besar dapat dijelaskan sebagai berikut.

2.9.1 Data Collection

Data Collection adalah proses mengumpulkan dan mengukur informasi tentang variabel yang ditargetkan dalam sistem yang mapan, yang kemudian memungkinkan seseorang untuk menjawab pertanyaan yang relevan dan mengevaluasi hasil. Pengumpulan Data pada penelitian ini adalah proses pengumpulan data dari Biro yang diperlukan sebagai dataset. Proses pengumpulan data ini dilakukan dengan datang langsung ke Biro dengan melakukan wawancara dan meminta data secara jamaah ke bagian administrasi (Yunitasari dkk., 2022).

2.9.2 Data Selection dan Preprocessing

Data yang sudah ada di *database* seringkali tidak semuanya dibutuhkan, maka dibutuhkan penyeleksian data untuk data yang benar-benar relevan dan an dalam proses selanjutnya. Data hasil seleksi yang akan digunakan oses *data mining*, disimpan dalam suatu berkas terpisah dari basis data nal (Wahyudi dkk., 2022).



Sedangkan *preprocessing* adalah proses perubahan bentuk data yang tidak terstruktur menjadi data yang terstruktur agar sesuai kebutuhan untuk diproses selanjutnya. Data yang telah melewati *preprocessing* adalah data yang lebih terstruktur (Veriane dkk., 2023).

Ada beberapa langkah yang dapat dilakukan dalam tahap *preprocessing*, di antaranya sebagai berikut (Assena, 2020):

- a. *Data Cleaning*, yaitu proses untuk mengisi *missing value* atau isian yang hilang atau kurang pas, mengoreksi data yang tidak konsisten, dan mendeteksi redundansi data yaitu penumpukan data atau duplikasi data
- b. *Data Integration*, yaitu suatu proses untuk menggabungkan data dari beberapa *file* sumber. *Data integration* ini hanya dilakukan apabila data yang akan diolah bersumber dari beberapa *file* sumber.
- c. *Data Transformation*, yaitu proses mengubah suatu data supaya mendapatkan data yang lebih berkualitas atau sesuai dengan kebutuhan. Dalam proses ini peneliti hanya akan melakukan proses *transformation* data pada *file* pertama dan *file* kedua.
- d. *Data Reduction*, yaitu proses untuk mengurangi atau mereduksi sejumlah data yang tidak dibutuhkan. *Data reduction* sangat berguna untuk mendapatkan atribut dan data yang akan digunakan sesuai keperluan.

2.9.3 Data Mining

Menurut *Gartner Group*, *data mining* adalah proses menemukan hubungan baru yang mempunyai arti, pola dan kebiasaan dengan memilah-milah sebagian besar data yang disimpan dalam media penyimpanan dengan menggunakan teknologi pengenalan pola seperti teknik statistik dan matematika (Larose dkk., 2005). Data yang diproses dengan teknik *data mining* akan menciptakan suatu pengetahuan ilmu baru yang bersumber dari data lama, hasil yang diperoleh dari pemrosesan data tersebut bisa digunakan untuk menentukan keputusan di masa depan (Santosa, 2007). *Data mining* merupakan bagian dari proses penemuan

uan dari basis data *Knowledge Discovery in Databases*.

ata mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dilakukan (Rosmini dkk., 2018), yaitu:



a. Deskripsi

Para peneliti dan analisis secara sederhana mencoba menemukan cara untuk menggambarkan pola dan kecenderungan trend yang tersembunyi dalam data. Deskripsi dari pola dan kecenderungan tersebut seringkali memberikan kemungkinan penjelasan untuk suatu pola tertentu.

b. Estimasi

Estimasi adalah keseluruhan proses yang memerlukan serta menggunakan *estimator* untuk menghasilkan sebuah *estimation* dari suatu parameter. Sedangkan menurut Kamus Besar Bahasa Indonesia (KBBI) Estimasi adalah perkiraan, penilaian atau pendapat. Ini menunjukkan bahwa istilah estimasi dapat digunakan secara umum untuk menyatakan perkiraan, penilaian atau pendapat mengenai sesuatu (Nainel dkk., 2020).

Teknik untuk melakukan estimasi terhadap sebuah data baru yang tidak memiliki keputusan berdasarkan histori data yang telah ada, dimana variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai dari variabel prediksi. Contohnya melakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk memprediksi kasus baru lainnya (Pradnyana dkk., 2018).

c. Prediksi

Prediksi memiliki kemiripan dengan klasifikasi dan estimasi, namun dalam prediksi nilai dari hasil akan ada di masa yang akan datang. Beberapa algoritma dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi (Santoso, 2007).

lasifikasi

Suatu teknik dengan melihat pada kelakuan dan atribut dari kelompok yang telah didefinisikan. Teknik ini dapat memberikan klasifikasi



pada data baru dengan memanipulasi data yang ada yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Salah satu contoh yang mudah dan populer adalah dengan decision tree yaitu salah satu teknik klasifikasi yang paling populer karena mudah untuk interpretasi seperti Algoritma C4.5, ID3 dan lain-lain. Contoh pemanfaatannya misalnya pada bidang akademik terkait klasifikasi siswa yang layak masuk ke dalam kelas unggulan atau akselerasi di sekolah tertentu (Pradnyana at al., 2018).

Dalam klasifikasi variabel, tujuan bersifat kategorik. Sebagai contoh penggolongan dataset terbagi ke dalam dua kelas (*binary*) ataupun lebih dari dua kelas (*multiclass*).

e. *Clustering*

Teknik untuk mengelompokkan data dan membentuk kelas objek-objek yang memiliki kemiripan. Klaster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam klaster lain. Proses klasterisasi berbeda dengan proses klasifikasi yaitu tidak adanya variabel target dalam pengklasteran. Proses klasterisasi mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), dimana kemiripan data dalam satu kelompok tinggi (maksimal) dan kemiripan data dengan data pada kelompok lain rendah (minimal). Contoh klasterisasi misalkan proses mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar (Pradnyana dkk., 2018).

Clustering merupakan teknik pengelompokan *record* data, pengamatan atau kasus dalam kelas yang memiliki kemiripan. *Cluster* adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record* lain dalam *cluster*.

f. Asosiasi



Teknik untuk mengenali kelakuan dari kejadian-kejadian khusus atau proses dimana hubungan asosiasi muncul pada setiap kejadian. Adapun teknik pemecahan masalah yang sering digunakan seperti Algoritma

Apriori. Contoh pemanfaatan Algoritma apriori yaitu pada bidang *marketing* ketika sebuah Minimarket melakukan tata letak produk yang dijual berdasarkan produk-produk mana yang paling sering dibeli konsumen, selain itu seperti tata letak buku yang dilakukan pustakawan di perpustakaan (Pradnyana dkk., 2018).

Mengidentifikasi hubungan antara berbagai peristiwa yang terjadi pada satu waktu. Dalam dunia bisnis lebih umum disebut sebagai analisis keranjang belanja.

Menurut Sihombing pada tahun 2012, *data mining* memiliki kelebihan sebagai alat analisis seperti :

- a. *Data mining* ini mampu menangani data-data dalam jumlah besar dan juga kompleks.
- b. *Data mining* juga dapat menangani data-data dengan berbagai macam tipe atribut.
- c. *Data mining* mampu mencari dan mengolah data secara otomatis. Disebut semi otomatis karena dalam beberapa teknik *data mining*, diperlukan parameter yang harus di *input* oleh *user* secara manual.
- d. *Data mining* dapat menggunakan pengalaman ataupun kesalahan terdahulu untuk meningkatkan kualitas dan hasil analisa sehingga mendapat hasil yang terbaik.

Menurut Sihombing pada tahun 2012, *data mining* memiliki kelemahan dalam pencarian data tidak mencari secara individualis, tetapi set individualis, atau dengan kata lain dikelompokkan dengan kriteria-kriteria tertentu (Tarigan dkk., 2022).

2.10 ANOVA Table

Analisis varians (*analysis of variance*) atau ANOVA adalah suatu metode analisis statistika yang termasuk ke dalam cabang statistika inferensi. Uji dalam menggunakan uji F karena dipakai untuk pengujian lebih dari 2 sampel. (*Analysis of variances*) digunakan untuk melakukan analisis komparasi abel. Teknik analisis komparatif dengan menggunakan tes “t” yakni mencari perbedaan yang signifikan dari dua buah mean hanya efektif bila



jumlah variabelnya dua. Untuk mengatasi hal tersebut ada teknik analisis komparatif yang lebih baik yaitu *Analysis of Variances* yang disingkat ANOVA. ANOVA digunakan untuk membandingkan rata-rata populasi bukan ragam populasi. Adapun asumsi dasar yang harus terpenuhi dalam analisis varian adalah (Santoso 2008):

1. Kenormalan

Distribusi data harus normal, agar data berdistribusi normal dapat ditempuh dengan cara memperbanyak jumlah sampel dalam kelompok.

2. Kesamaan variansi

Setiap kelompok hendaknya berasal dari populasi yang sama dengan variansi yang sama pula. Bila banyaknya sampel sama pada setiap kelompok maka kesamaan variansinya dapat diabaikan. Tapi bila banyak sampel pada masing-masing kelompok tidak sama maka kesamaan variansi populasi sangat diperlukan.

3. Pengamatan bebas

Sampel hendaknya diambil secara acak (*random*), sehingga setiap pengamatan merupakan informasi yang bebas.

Tabel 1 Contoh tabel ANOVA

	df	SS	MS	F	Significance F
Regression	1	2650706,899	2650707	1,606271	0,214454864
Residual	31	51156927,4	1650223	-	-
Total	32	53807634,3	-	-	-

Sumber: Budi, 2020

Tabel 1 menunjukkan bagaimana contoh penampakan tabel ANOVA. Tabel ANOVA (*Analysis of Variance*) menguji penerimaan (*acceptability*) model dari perspektif statistik dalam bentuk analisis sumber keragaman. ANOVA ini sering juga diterjemahkan sebagai analisis ragam (Budi, 2020).

Dari tabel ANOVA tersebut diungkapkan bahwa keragaman data aktual terikat (Jumlah penduduk miskin) bersumber dari model regresi dan dari Dalam pengertian sederhana untuk kasus ini adalah variasi (turun-naiknya atau kecilnya) jumlah penduduk miskin disebabkan oleh variasi dari luas



daerah (model regresi) serta dari faktor-faktor lainnya yang mempengaruhi jumlah penduduk miskin yang tidak dimasukkan dalam model regresi (residual).

Degree of Freedom (df) atau derajat bebas dari total adalah $n-1$, dimana n adalah banyaknya observasi. Dalam hal ini banyaknya observasi adalah 33 maka derajat bebas total adalah 32. Derajat bebas dari model regresi adalah 1, karena ada satu variabel bebas dalam model ini (luas daerah). Derajat bebas untuk residual adalah sisanya yaitu derajat bebas total – derajat bebas regresi = $33 - 1 = 32$.

Kolom SS (*Sum of Square*) atau jumlah kuadrat untuk *regression* diperoleh dari penjumlahan kuadrat dari prediksi variabel terikat (Jumlah penduduk miskin) dikurangi dengan nilai rata-rata jumlah penduduk miskin dari data sebenarnya. Jadi secara manual mencari terlebih dahulu rata-rata permintaan dari data asli. Kemudian masing-masing prediksi jumlah penduduk miskin dikurangi dengan rata-rata tersebut kemudian dikuadratkan. Selanjutnya, seluruh hasil perhitungan tersebut dijumlahkan. Pada data di atas nilai dari SS *regression* adalah 2650706,899.

Kolom SS untuk residual diperoleh dari jumlah pengkuadratan dari residual. Hasil perhitungan tersebut kemudian dijumlahkan. Berdasarkan output di atas maka nilai SS total adalah 53807634,3 Hasil ketika SS tersebut memiliki arti dimana apabila SS total yang diperoleh adalah adalah 53807634,3 yang memiliki arti, variasi dari jumlah penduduk miskin yang dikuadratkan adalah sebesar nilai tersebut. Bervariasinya jumlah penduduk miskin disebabkan oleh sebagian berasal dari variabel bebas (luas daerah) yaitu sebesar 2650706,899 (regresi). Kemudian sisanya sebesar 51156927,4 disebabkan oleh variabel lain yang juga mempengaruhi jumlah penduduk miskin tetapi tidak dimasukkan dalam model (residual).

Jika membandingkan (bagi) antara SS regresi dengan SS total, maka akan didapatkan proporsi dari total variasi jumlah penduduk miskin yang disebabkan oleh variasi luas daerah. Praktikan mencoba membagi antara nilai SS regresi dengan SS total yaitu $2650706,899 / 53807634,3 = 0,049$. hasil tersebut sama dengan hasil dari R^2 atau koefisien determinasi yang telah dibahas di atas.



anjutnya kolom berikutnya dari ANOVA adalah kolom MS (*Mean of* atau rata-rata jumlah kuadrat. Ini adalah hasil bagi antara kolom SS dengan n . Berdasarkan hasil di atas besarnya MS adalah regresi dan residual

berturut-turut adalah 2650707 dan 1650223. Dari perhitungan MS ini, selanjutnya dengan membagi antara MS Regresi dengan MS Residual didapatkan nilai F. Nilai F ini yang dikenal dengan F hitung dalam pengujian hipotesa dibandingkan dengan nilai F tabel. Jika F hitung $>$ F tabel, maka dapat dinyatakan bahwa secara simultan (bersama-sama) luas daerah berpengaruh signifikan terhadap jumlah penduduk miskin. Setelah melakukan pembagian antara MS regresi dengan residual didapatkan hasil nilai F hitung yaitu 1,606271 seperti pada hasil analisis. Selain itu, dapat juga membandingkan antara taraf nyata dengan p-value (dalam istilah Excel adalah *Significance F*). Jika taraf nyata $>$ dari *p-value* maka kesimpulannya sama dengan di atas.

2.11 *Machine Learning*

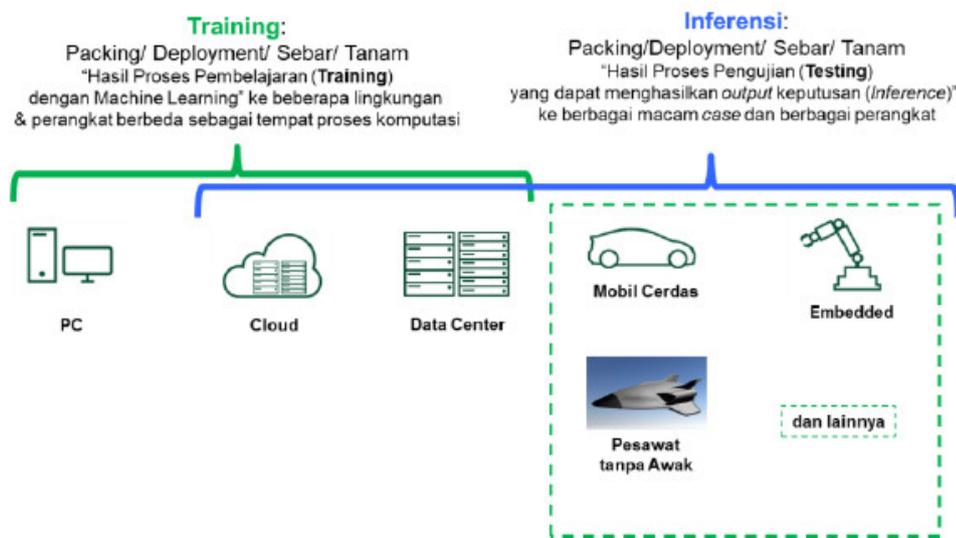
Machine Learning (ML) atau pembelajaran mesin merupakan pendekatan dalam AI yang banyak digunakan untuk menggantikan atau menirukan perilaku manusia untuk menyelesaikan masalah atau melakukan otomatisasi. Sesuai namanya, ML mencoba menirukan bagaimana proses manusia atau makhluk cerdas belajar dan menggeneralisasi.¹ Setidaknya ada dua aplikasi utama dalam ML yaitu, klasifikasi dan prediksi. Ciri khas dari ML adalah adanya proses pelatihan, pembelajaran, atau *training*. Oleh karena itu, ML membutuhkan data untuk dipelajari yang disebut sebagai data *training*. Klasifikasi adalah metode dalam ML yang digunakan oleh mesin untuk memilah atau mengklasifikasikan objek berdasarkan ciri tertentu sebagaimana manusia mencoba membedakan benda satu dengan yang lain. Sedangkan prediksi atau regresi digunakan oleh mesin untuk menerka keluaran dari suatu data masukan berdasarkan data yang sudah dipelajari dalam *training*. Metode ML yang paling populer yaitu Sistem Pengambil Keputusan, *Support Vector Machine* (SVM) dan Neural Network.

Machine Learning atau Mesin Pembelajaran adalah cabang dari AI yang fokus belajar dari data (learn from data), yaitu fokus pada pengembangan sistem yang mampu belajar secara “mandiri” tanpa harus berulang kali diprogram manusia. ML

gunakan Data yang valid sebagai bahan belajar (ketika proses *training*) digunakan ketika testing untuk hasil *output* yang optimal (Cholissodin,



Hasil pengembangan produk berbasis AI (*Machine Learning* dan Teknik lainnya misal dengan Optimasi, etc) ini harapannya dapat lebih memberikan kemudahan dan langsung dapat diterapkan di masyarakat luas atau bahkan masuk ke industri dalam skala nasional dan internasional. Gambar 2 berikut merupakan ilustrasi Map produk untuk *apply* ke masyarakat.



Gambar 2 Map pengembangan produk App

Sumber: Cholissodin, 2012

Penelitian terkini mengungkapkan bahwa machine learning terbagi menjadi tiga kategori: *Supervised Learning*, *Unsupervised Learning*, *Reinforcement Learning* (Somvanshi & Chavan, 2016). Berikut penjelasan kategori machine learning secara lebih detail.

2.11.1 Supervised Learning

Teknik yang digunakan oleh *supervised learning* adalah metode klasifikasi dimana kumpulan data sepenuhnya diberikan label untuk mengklasifikasikan kelas yang tidak dikenal (Thupae dkk., 2018).

Kumpulan sampel digunakan untuk meringkas karakteristik distribusi ukuran perilaku dalam setiap jenis aplikasi sehingga membentuk model perilaku dari data

(k., 2011). *Supervised learning* dikelompokkan lebih lanjut dalam masalah klasifikasi dan regresi. Masalah klasifikasi adalah ketika variabel *output* berbentuk kategori seperti merah atau biru atau penyakit dan tidak ada penyakit. Sedangkan



masalah regresi adalah ketika variabel *output* adalah nilai riil, seperti *dollar* atau berat (Brownlee, 2016).

2.11.2 Unsupervised Learning

Teknik yang digunakan oleh *unsupervised learning* sering disebut cluster dikarenakan tidak ada kebutuhan untuk pemberian label dalam kumpulan data dan hasilnya tidak mengidentifikasi contoh di kelas yang telah ditentukan.

Dalam jenis pembelajaran *unsupervised learning*, sistem disediakan dengan beberapa *input* sampel tetapi tidak ada *output* yang hadir. Karena tidak ada *output* yang diinginkan di sini kategorisasi dilakukan sehingga algoritma membedakan dengan benar antara kumpulan data. Ini adalah tugas mendefinisikan fungsi untuk menggambarkan struktur yang tersembunyi dari data yang tidak berlabel (Somvanshi & Chavan, 2016). *unsupervised learning* dikelompokkan lebih lanjut dalam masalah *clustering* dan asosiasi. Masalah pengelompokan (*clustering*) adalah tempat untuk menemukan pengelompokan yang melekat dalam data, seperti mengelompokkan pelanggan berdasarkan pada perilaku pembelian. Sedangkan masalah asosiasi adalah aturan yang menggambarkan sebagian besar data yang ada, seperti orang yang membeli A juga cenderung membeli B (Brownlee, 2016).

2.11.3 Reinforcement Learning

Reinforcement Learning biasanya berada antara *supervised learning* dan *unsupervised learning*, teknik ini bekerja dalam lingkungan yang dinamis dimana konsepnya harus menyelesaikan tujuan tanpa adanya pemberitahuan dari komputer secara eksplisit jika tujuan tersebut telah tercapai (Das & Nene, 2017).

Reinforcement learning berasal dari teori belajar hewan. Pembelajaran ini tidak memerlukan pengetahuan sebelumnya, dapat secara mandiri mendapatkan kebijakan opsional dengan pengetahuan yang diperoleh melalui coba-coba dan terus berinteraksi dengan lingkungan yang dinamis (Qiang & Zhongli, 2011).

Masalah *reinforcement learning* diselesaikan dengan mempelajari pengalaman melalui trial-and-error (Mahmud dkk., 2018). Algoritma reinforcement terkait dengan algoritma pemrograman dinamis yang sering digunakan menyelesaikan masalah optimisasi (Mitchell, 1997).



2.12 Peramalan (*forecasting*)

Peramalan merupakan suatu kegiatan untuk memperkirakan suatu kejadian apa yang akan terjadi pada masa yang akan datang. Peramalan dapat dikatakan sebagai awal dari sebuah proses pengambilan keputusan. Sebelum melakukan sebuah peramalan, hendaknya harus diketahui dahulu apa maksud dan tujuan dari melakukan peramalan. Pada dasarnya, peramalan memperkirakan suatu keadaan dimasa yang akan datang berdasarkan keadaan masa lalu dan sekarang yang diperlukan untuk menetapkan kapan suatu peristiwa akan terjadi, sehingga tindakan yang tepat dapat dilakukan. Contohnya adalah permintaan terhadap suatu produk pada periode waktu yang akan datang. Pada hakikatnya, sebuah peramalan hanya merupakan suatu perkiraan terhadap suatu objek. Akan tetapi, dengan menggunakan teknik peramalan yang tepat maka peramalan akan menjadi lebih dari sekedar perkiraan. Pada umumnya, hasil peramalan yang baik dapat dilihat dari kecilnya nilai kesalahan meramal atau *forecast error* yang dapat diukur dengan menggunakan *Mean Absolute Deviation*, *Mean Square Error*, dan *Mean Absolute Percentage Error* (Iswahyudi,2016).

Dua hal pokok yang harus diperhatikan dalam proses peramalan yang akurat dan bermanfaat :

- a. Pengumpulan data yang relevan berupa informasi yang dapat menghasilkan peramalan yang akurat.
- b. Pemilihan teknik peramalan yang tepat yang akan memanfaatkan informasi data yang diperoleh semaksimal mungkin.

Apabila dilihat sifat ramalan yang telah disusun, maka peramalan dapat dibedakan atas 2 macam, yaitu:

- a. Peramalan Kualitatif

Pada peramalan kualitatif, hasil peramalan sangat tergantung pada analis atau peramal yang menyusunnya. Hal ini dikarenakan hasil peramalan berdasarkan pemikiran yang bersifat *judgement* atau pendapat.

- b. Peramalan Kuantitatif

Pada peramalan kuantitatif, hasil peramalan yang dibuat sangat tergantung pada metode peramalan yang digunakan. Penggunaan metode berbeda akan menghasilkan hasil yang berbeda walaupun data yang



digunakan sama. Setiap metode yang digunakan harus mempertimbangkan tingkat keakuratan hasil ramalan sehingga tidak terjadi penyimpangan yang terlalu signifikan antara hasil ramalan dengan kenyataan yang terjadi di lapangan.

Metode peramalan kuantitatif dibedakan menjadi dua bagian, yaitu sebagai berikut:

- a). Metode peramalan yang didasarkan pada penggunaan pola hubungan antara variabel yang diperkirakan dengan variabel waktu yang merupakan deret waktu atau *Time Series*. Contohnya metode *smoothing*, regresi, dekomposisi
- b). Metode peramalan yang berdasarkan pada penggunaan analisa pola hubungan antara variabel yang diperkirakan dengan variabel lain yang mempengaruhinya, yang bukan waktu yang disebut Causal. Contoh metode regresi, korelasi, ekonometrik, *input-output*

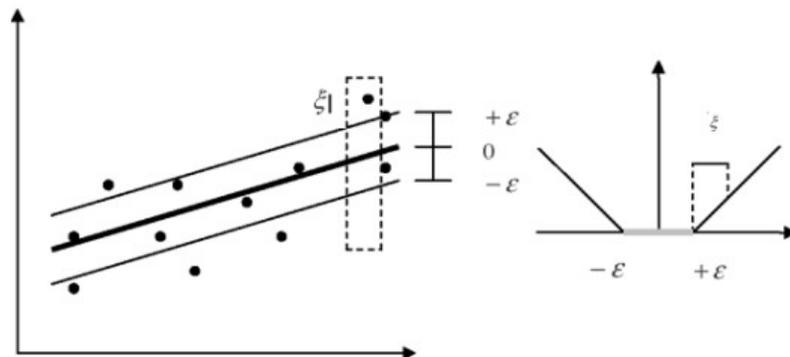
2.13 Support Vector Regression

Support Vector Regression (SVR) adalah metode pengembangan dari *Support Vector Machine* (SVM) yang digunakan pada permasalahan regresi. SVR termasuk dalam algoritma *supervised learning* yang digunakan untuk memprediksi nilai variabel kontinu. Perbedaan SVM dan SVR adalah SVM digunakan untuk mencari *hyperplane* (fungsi pemisah) yang terbaik di antara 2 objek yang tidak terbatas jumlahnya dengan cara memaksimalkan jarak (margin) antara dua objek yang berbeda sedangkan SVR digunakan untuk menemukan suatu fungsi yang memiliki deviasi paling besar ϵ dari target aktual y_i . Menurut Smola & Scholkopt pada tahun 2004, SVR bertujuan untuk menemukan fungsi $f(x)$ sebagai suatu *hyperplane* (garis pemisah) berupa fungsi regresi, pada semua input data yang memiliki deviasi paling besar ϵ dari target aktual y_i untuk semua data training dan membuat galat sekecil mungkin

Keunggulan SVR adalah kemampuan untuk mengatasi masalah data dengan trik kernel. Dalam SVR terdapat parameter *epsilon* (ϵ), *cost* (C), *gamma* (γ) yang ditentukan nilai optimalnya menggunakan metode *grid search*.



Grid search merupakan kombinasi parameter yang diujikan pada suatu model SVR untuk mencari nilai error dalam klasifikasi.



Gambar 3 Hyperplane dan margin pada SVR

Sumber: Smola & Scholkopt, 2004

Pada Gambar 3 garis hitam tebal merupakan *hyperplane* sedangkan untuk dua garis yang mengapitnya adalah *soft margin*. Jarak antara *hyperplane* dan *soft margin* adalah sebesar ε dan titik-titik yang berada pada $+\varepsilon$ sampai $-\varepsilon$ merupakan support vector, namun untuk titik yang melewati soft margin dibutuhkan adanya variabel slack ξ .

Ide dasar penggunaan metode SVR adalah misalkan terdapat n set data *training* $x_i y_i$, dengan $x_i \in R^d$ adalah vektor input dari data ke- i dimana $i=1,2,\dots,n$, dan d adalah dimensi dan y_i adalah nilai target. Persamaan fungsi regresi secara umum dapat ditulis persamaan (1) berikut (Smola & Scholkopt, 2004):

$$f(x) = w^T \varphi(x) + b \quad (1)$$

Keterangan :

$\varphi(x)$ = titik di dalam feature space F , hasil dari pemetaan x di dalam input space

$f(x)$ = fungsi regresi

w = vektor bobot, yang mempunyai dimensi 1

b = bias

x = input atau vektor input

Sedangkan koefisien w dan b mempunyai fungsi untuk meminimalkan fungsi perti pada persamaan (2) di bawah ini:

$$R = \min \frac{1}{2} \|w\|^2 + C \frac{1}{l} (\sum_1^l L_\varepsilon(y_i, f(x_i))) \quad (2)$$



Keterangan :

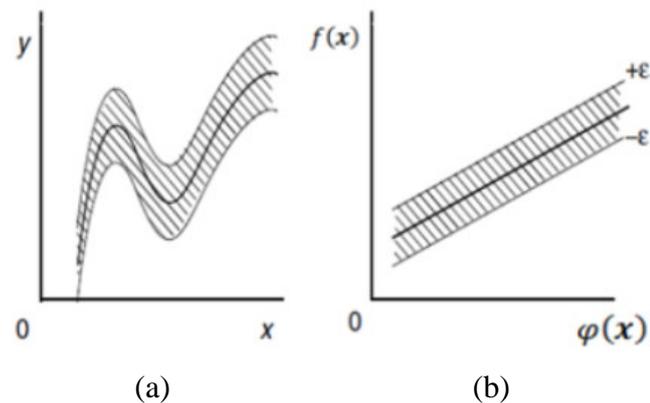
R = Risk Function (Fungsi Resiko)

L_ε = Loss Function yang bertipe ε -insentive loss function

$\|w\|$ = normalisasi w

ε = epsilon yaitu deviasi atau derajat toleransi terhadap error

C = nilai penalti jika deviasi lebih besar dari batas error



Gambar 4 Insensitive zone (a) original input space, dan (b) feature space

Sumber: Yasin dkk., 2016

Pada gambar 4, gambar Insensitive zone (a) original input space, dan (b) feature space $D(x, y) = \pm\varepsilon$ adalah jarak terjauh *support vector* dari *hyperplane*, kemudian disebut margin. Memaksimalkan margin akan meningkatkan probabilitas data ke dalam radius $\pm\varepsilon$. Jarak dari *hyperplane* $D(x, y)$ ke data (x, y) adalah $|D(x, y)|/\|w^*\|$, dimana:

$$w^* = (1 - w^T)^T \quad (3)$$

Diasumsikan bahwa jarak maksimum data terhadap *hyperplane* adalah δ , maka estimasi yang ideal akan terpenuhi dengan:

$$\begin{aligned} \frac{|D(x, y)|}{\|w^*\|} &\leq \delta \\ |D(x, y)| &\leq \delta \|w^*\| \\ \delta \|w^*\| &= \varepsilon \end{aligned} \quad (4)$$



h karena itu untuk memaksimalkan margin δ , diperlukan $\|w^*\|$ yang
t. Optimasi penyelesaian masalah dengan bentuk Quadratic Programming:

$$\min \frac{1}{2} \|w\|^2 \quad (5)$$

dengan syarat

$$y_i - w^T \varphi(x_i) - b \leq \varepsilon \text{ untuk } i = 1, \dots, l$$

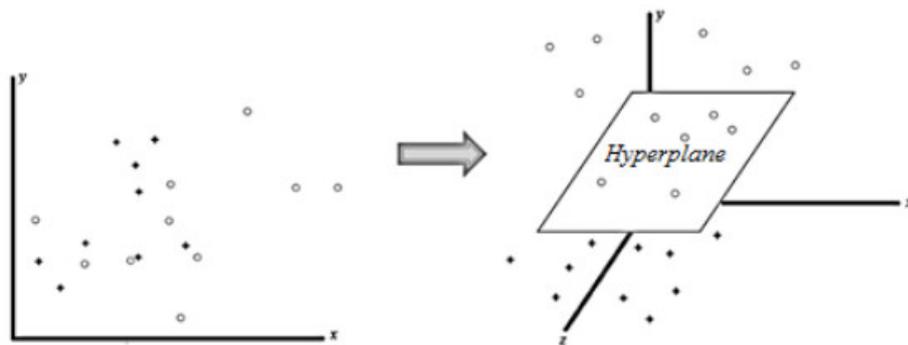
$$w^T \varphi(x_i) - y_i + b \leq \varepsilon \text{ untuk } i = 1, \dots, l$$

Faktor $\|w\|^2$ dinamakan regulasi. Meminimalkan $\|w\|^2$ akan membuat suatu fungsi setipis (flat) mungkin, sehingga bisa mengontrol kapasitas fungsi (function capacity).

Diasumsikan bahwa semua titik ada dalam rentang $f(x) \pm \varepsilon$ (feasible), dalam hal ketidaklayakan (*infesibility*), dimana ada beberapa titik yang mungkin keluar dari rentang $f(x) \pm \varepsilon$ maka ditambahkan variabel slack ξ dan ξ^* untuk mengatasi masalah pembatasan yang tidak layak (infeasible constraints) dalam problem optimasi (Santosa, 2007).

2.13.1 Fungsi Kernel

Menurut Furi, Jondri, & Saepudin pada tahun 2015 dalam dunia nyata permasalahan jarang yang bersifat linear dan kebanyakan bersifat non linear. Untuk menyelesaikan permasalahan non linear ini digunakan fungsi kernel. Untuk memecahkan masalah linear dalam ruang berdimensi tinggi, yang harus dilakukan adalah mengganti *inner product* (x_i dan x_j) dengan fungsi kernel yang sesuai.



Gambar 5 Ilustrasi *hyperplane* pada dimensi yang lebih tinggi

Sumber: Athoillah, 2015

Ilustrasi dari konsep kernel pada Gambar 5 di atas yaitu, gambar pada sisi kiri menunjukkan data dari dua kelas berbeda pada ruang input berdimensi dua yang at dipisahkan secara linear. Sedangkan pada gambar sebelah kanan terlihat *iture space* mampu memetakan setiap data pada ruang input ke ruang fitur



berdimensi lebih tinggi, sehingga kedua kelas data tersebut bisa dipisahkan secara linear oleh *hyperplane*.

Untuk membantu mengatasi permasalahan non linear pada dimensi tinggi yang harus dilakukan yaitu mengganti *inner product* (x_i dan x_j) dengan fungsi kernel. Karena kinerja dari metode SVR ditentukan oleh jenis fungsi kernel dan parameter yang digunakan. Adapun menurut Bhavsar & Mahesh pada tahun 2012, fungsi-fungsi kernel yang sering digunakan pada metode SVR adalah sebagai berikut.

Tabel 2 Rumus dari setiap tipe kernel pada SVR

No.	Tipe Kernel	Formula
1.	Linear	$K(x_i, x_j) = (x_i^T x_j)$
2.	Polynomial	$K(x_i, x_j) = (x_i, x_j + 1)^P$
3.	Radial Basis Function	$K(x_i, x_j) = \exp(-\gamma(x_i, x_j)^2)$

Sumber: Bhavsar & Mahesh, 2012

dengan:

x_i, x_j = Vektor dari dua dataset

P = Derajat polinomial

γ = Gamma

Pemilihan fungsi kernel yang tepat merupakan hal yang penting untuk menentukan *feature space*. Kernel linear merupakan fungsi kernel paling sederhana. Kernel linear biasanya digunakan pada data set yang datanya sudah terpisah secara linear. Kernel polinomial adalah sebuah persamaan yang terdiri dari variabel dan koefisien yang memiliki suku banyak dan sering digunakan jika data training dinormalisasi. Sedangkan kernel radial atau Gaussian merupakan kernel yang digunakan untuk menyelesaikan masalah pada data yang tidak terpisah secara linear.

2.13.2 Grid Search Optimization



masalah yang sering terjadi ketika menggunakan metode SVR adalah menentukan parameter model yang optimal. Salah satu cara optimasi yang

dapat digunakan untuk menentukan parameter terbaik pada metode SVR adalah optimasi *Grid Search* atau *Grid Search Optimization*. *Grid Search Optimization* ini akan membagi jangkauan parameter yang akan dioptimalkan ke dalam grid dan melintasi semua titik untuk mendapatkan parameter yang optimal (Yasin dkk, 2014). Dalam aplikasinya, *Grid Search* dipadukan dengan matriks kinerja dan diukur menggunakan *cross validation* pada data training. *Grid Search Optimization* akan melatih pasangan-pasangan parameter dan akan menentukan/memilih pasangan parameter yang optimal dengan akurasi terbaik atau rata-rata galat yang terkecil dari uji *cross validation Grid Search*.

Menurut Leidiyana pada tahun 2013, *cross validation* adalah pengujian standar yang dilakukan untuk memprediksi *error rate*. Data *training* dibagi secara *random* ke dalam beberapa bagian dengan perbandingan yang sama kemudian *error rate* dihitung bagian demi bagian, selanjutnya dihitung rata-rata seluruh *error rate* untuk mendapatkan *error rate* keseluruhan. Dalam *cross validation*, dikenal validasi *leave-one-out* (LOO). Dalam LOO, data dibagi kedalam 2 subset, subset 1 berisi N-1 data training dan satu data sisanya untuk testing (Santosa, 2007).

Salah satu metode *cross validation* yang umum digunakan adalah *k-fold validation*. Menurut Jiawei dkk pada tahun 2011, prosedur dari metode *cross validation* adalah sebagai berikut:

1. Membagi data menjadi k bagian dengan ukuran yang sama.
2. K-1 bagian dijadikan data *training* dan satu bagian dijadikan data testing.
3. Proses ini dilakukan sebanyak k pengulangan pada setiap kombinasi data *training* dan data testing. Nilai akurasi dari setiap iterasi dirata-ratakan untuk mendapatkan estimasi nilai akurasi akhir.

2.14 Artificial Neural Network

Artificial Neural Network (ANN) merupakan sebuah sistem cerdas yang digunakan untuk mengolah informasi yang merupakan perkembangan dari generalisasi model matematika. Prinsip kerja ANN terinspirasi dari prinsip kerja jaringan saraf (*neural network*) manusia. Para ilmuwan menciptakan algoritma yang bekerja menyerupai pola kerja saraf (neuron) tersebut, maka



digunakanlah nama *Artificial Neural Network*, atau dalam Bahasa dengan biasa disebut Jaringan Saraf Tiruan (JST).

Meniru cara kerja yang sama dengan otak, input atau masukan diumpamakan sebagai neuron, kemudian masukan-masukan tersebut dikalikan dengan suatu nilai (bobot) dan diolah dengan fungsi tertentu (fungsi aktivasi sehingga menghasilkan suatu keluaran (Stergiou dan Siganos, 2006). Secara kausal, ANN bekerja dengan mempresentasikan dinamika suatu sistem ke dalam suatu model matematika dengan cara menentukan arsitektur jaringan, metode pembelajaran jaringan, dan akurasi model jaringan. Dengan interpretasi tersebut akan diketahui pola perilaku sistem yang dapat digunakan untuk meramalkan keluaran tertentu pada periode ke depan.

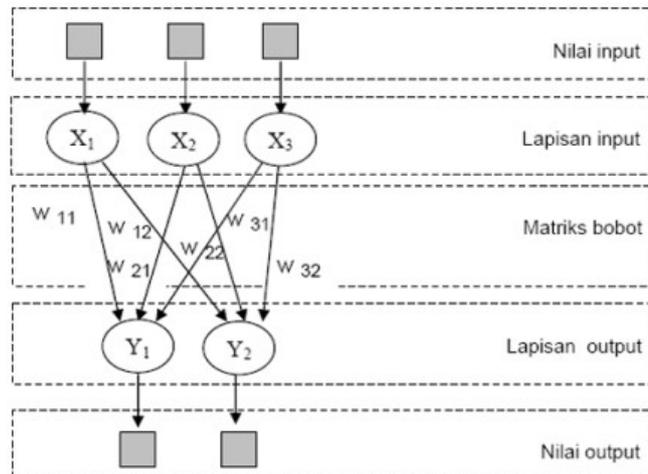
2.14.1 Arsitektur Artificial Neural Network

Pada ANN, neuron diasumsikan dapat dikelompokkan dalam layer seperti yang telah dijelaskan sebelumnya. Umumnya, terdapat 3 struktur ANN yaitu *single layer network* dan *multilayer network* (berdasarkan banyak lapisan) serta *competitive layer network*.

a. Jaringan dengan lapisan tunggal (*single layer network*)

Jaringan lapis tunggal terdiri dari satu layer dari bobot yang saling terhubung. Jaringan ini hanya menerima *input* atau informasi yang masuk kemudian secara langsung akan mengolahnya menjadi sinyal keluaran (*output*) tanpa harus melalui lapisan tersembunyi (*hidden layer*). Pada Gambar 6, terlihat neuron-neuron pada kedua lapisan saling berhubungan. Seberapa besar hubungan antara 2 neuron tersebut ditentukan oleh bobot yang bersesuaian. Semua unit *input* akan dihubungkan dengan setiap unit *output*.



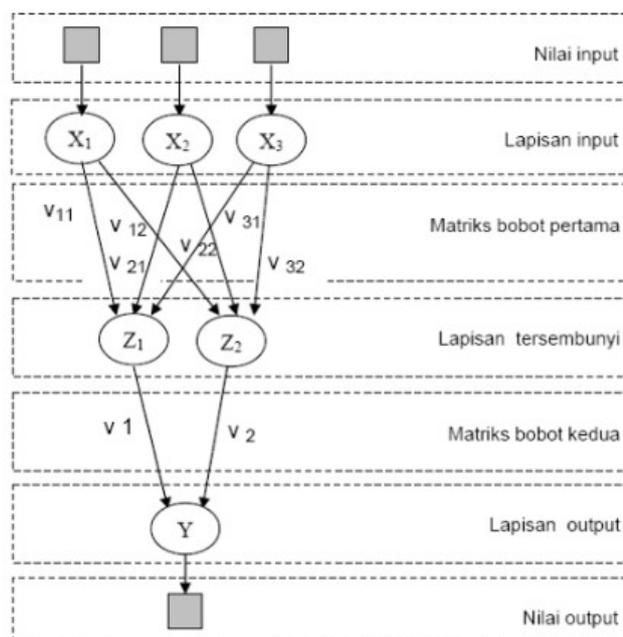


Gambar 6 Jaringan Saraf dengan lapisan tunggal

Sumber: Wuryandari & Afrianto, 2012

b. Jaringan dengan banyak lapisan (*multilayer network*)

ANN dengan tipe ini memiliki lebih dari satu layer yang diberi nama *hidden layer*. Gambar 7 menunjukkan jaringan dengan *multilayer*, dimana semua layer yang terletak di antara *input layer* dan *output layer* merupakan *hidden layer*, karena jumlah *hidden layer* yang bisa lebih dari satu layer. ANN lapis jamak ini menggunakan fungsi aktivasi nonlinear yang dapat digunakan untuk menyelesaikan berbagai tipe permasalahan yang lebih rumit dan kompleks daripada lapisan tunggal

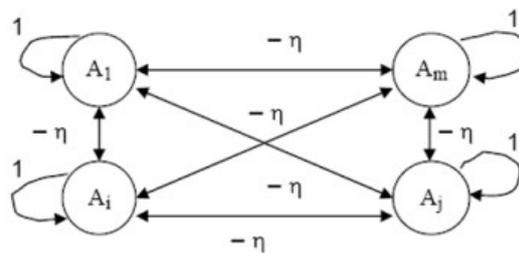


Gambar 7 Jaringan Saraf Tiruan lapisan banyak

Sumber: Wuryandari & Afrianto, 2012

c. Jaringan dengan lapisan kompetitif (*competitive layer network*)

Pada jaringan ini sekumpulan neuron bersaing untuk mendapatkan hak menjadi aktif. Umumnya hubungan antar neuron pada lapisan kompetitif ini tidak diperlihatkan pada diagram arsitektur. Gambar 8 berikut menunjukkan salah satu contoh arsitektur jaringan dengan lapisan kompetitif yang memiliki bobot $-\eta$.



Gambar 8 Jaringan Saraf dengan lapisan kompetitif yang memiliki $-\eta$

Sumber: Wuryandari & Afrianto, 2012

2.14.2 Fungsi Pelatihan

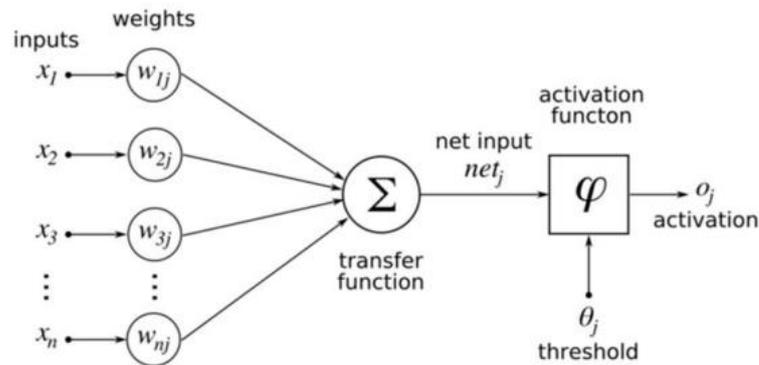
Fokus utama dalam peramalan menggunakan metode ANN adalah untuk melatih *network* untuk dapat memahami *pattern* nilai yang dimiliki dari sebuah kasus berdasarkan data historis untuk dapat memprediksi nilai yang akan datang. Fungsi pelatihan berfungsi sebagai algoritma yang digunakan pada ANN untuk melatih *network* dalam memahami *pattern* tersebut. Terdapat banyak fungsi pelatihan yang tersedia dan dapat digunakan, hal ini menjadikan sebagian orang bingung untuk menentukan fungsi pelatihan apa yang paling baik untuk digunakan. Namun berdasarkan uji coba yang dilakukan MATLAB didapatkan dua fungsi pelatihan yang dipilih dalam penelitian ini. Kedua fungsi tersebut adalah Lavenberg-Marquardt (LM), BFGS Quasi-Newton (BFG). Kedua fungsi pelatihan diatas telah diuji dan dibuktikan lebih baik dari fungsi lainnya oleh MATLAB

memecahkan kasus dengan tipe permasalahan baik *function approximation* maupun *pattern recognition*.



2.14.3 Fungsi Aktivasi

Fungsi aktivasi merupakan fungsi yang digunakan untuk mengolah inputan informasi. Sebagai gambaran, fungsi aktivasi dapat dilihat pada Gambar 9 di bawah ini.



Gambar 9 Struktur ANN

Sumber: Wuryandari & Afrianto, 2012

Dapat dilihat pada struktur ANN pada Gambar 9 di atas, fungsi transfer berfungsi mengubah sejumlah inputan ($x_1, x_2, x_3, \dots, x_n$) yang memiliki bobot ($W_{1j}, W_{2j}, \dots, W_{nj}$) dan bias menjadi nilai input (net_j) bagi fungsi aktivasi. Selanjutnya fungsi aktivasi memproses nilai input untuk dibandingkan dengan threshold yang ditentukan dan mengaktivasi nilai menjadi output (o_j). Beberapa fungsi aktivasi yang digunakan pada ANN adalah:

Selain Fungsi Tanda, terdapat beberapa fungsi lain yang praktis untuk digunakan oleh neuron, yaitu:

a. *Step Function*

Fungsi aktivasi *Step* dan *Sign* disebut juga *hard limit function*, sering digunakan oleh neuron untuk tugas-tugas pengambilan keputusan (*decision making*) dalam pengklasifikasian dan pengenalan pola.

b. *Linear Function*

Fungsi aktivasi *Linear* menyediakan *output* yang setara dengan *input* neuron yang diberikan bobot. Biasanya digunakan untuk aproksimasi linear.

$$f(x) = x \quad (6)$$

sigmoid Function



Fungsi *Sigmoid* mentransformasi *input*, yang bisa memiliki nilai di antara positif dan negatif tak terhingga, menjadi nilai yang dapat dinalar diantara 0 dan 1, biasanya untuk jaringan *back-propagation*.

1. *Binary Sigmoid Function*

$$f(x) = \frac{1}{1+e^{(-kx)}} \quad (7)$$

2. *Bipolar Sigmoid Function*

$$f(x) = \frac{1-e^{(-kx)}}{1+e^{(-kx)}} \quad (8)$$

d. *Fungsi Rectified Linear Unit (ReLU)*

Rectified Linear Unit (ReLU) function memiliki kelebihan dalam Network yang diinisiasi secara *random*, hanya 50% dari *hidden layer* yang akan di aktivasi. ReLU Function dirumuskan sebagai berikut.

$$f(x) = \max(0, x) \quad (9)$$

e. *Softmax*

Softmax Funtion atau disebut juga *softmax regression* adalah bentuk *logistic regression* yang *input value*-nya di normalisasi ke dalam *probability distribution*, dirumuskan sebagai,

$$f_i(\vec{x}) = \frac{e^{(x_i)}}{\sum_{j=1}^J e^{(x_j)}} \text{ untuk } i = 1, \dots, J \quad (10)$$

2.15 RMSE

Setiap hasil dari proses prediksi (biasanya untuk istilah perkiraan jangka pendek) atau peramalan (biasanya untuk istilah perkiraan jangka panjang) diperlukan proses evaluasi untuk mengetahui performa dari model peramalan yang diterapkan (Cholissodin dkk., 2020).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - A_1)^2} \quad (11)$$

Ket:



nilai *root mean square error*
 aktual
 hasil prediksi

n = jumlah data

2.16 Penelitian Terkait

Berikut beberapa penelitian terkait terhadap penelitian yang dilakukan:

2.16.1 Estimating the Determinant Factors of Container Dwell Seaports

Pada penelitian Nadereh dkk tahun 2012 ini, dilakukan identifikasi faktor penentu CDT (*container dwelling time*) dan menggambarkan alat komputasi yang sesuai untuk memperkirakan CDT berdasarkan serangkaian faktor tertentu dimana operator terminal biasanya mengumpulkan data yang relevan. Artikel ini membandingkan kinerja tiga algoritma *data mining* untuk memperkirakan CDT yaitu Naive Bayes, decision tree, dan NB-decision tree hybrid. Dengan menggunakan model dengan kinerja terbaik, data terminal sampel digunakan untuk mengukur bagaimana perubahan faktor penentu CDT berdampak pada CDT, kapasitas pekarangan, dan pendapatan terminal. Hasilnya mengungkapkan bahwa dampak dari perubahan faktor penentu CDT dapat cukup besar untuk mempengaruhi kapasitas terminal dan pendapatan yang diperoleh dari biaya kelebihan waktu berlabuh. Algoritma yang berkinerja terbaik adalah *decision tree*, yang kemudian hasilnya digunakan untuk mengukur dampak perubahan penentu CDT pada CDT. Meskipun tidak ada kesimpulan umum yang dapat diturunkan mengenai pentingnya faktor apa pun dan dampaknya terhadap CDT, pendekatan generik yang dikembangkan dapat digunakan dalam kombinasi dengan data dari terminal tertentu untuk membantu menemukan cara potensial untuk mengelola CDT secara efektif dan menentukan dampak yang diantisipasi pada terminal. Kapasitas pekarangan dan pendapatan. Penelitian ini menyediakan alat yang berguna bagi anggota komunitas pelabuhan, perdagangan, dan transportasi untuk mengevaluasi kebijakan yang tepat guna meningkatkan pengoperasian fasilitas yang merupakan mata rantai penting dalam rantai pasokan dan sangat penting bagi ekonomi global. Penelitian di masa depan dapat dibangun berdasarkan pekerjaan

memprediksi CDT menggunakan kumpulan data yang lebih komprehensif tentang faktor penentu CDT, seperti informasi pengirim dan penerima



2.16.2 Development of Models Predicting Dwell Time of Import Containers in Port Container Terminals – An Artificial Neural Networks Application

Pada penelitian Ioanna dkk tahun 2015 ini, tujuan umumnya adalah untuk mengusulkan pengembangan kerangka metodologis yang menggabungkan berbagai faktor yang mempengaruhi *Dwell Time (DT)* petikemas di terminal petikemas. Terminal petikemas dianggap sebagai elemen kunci dari rantai logistik karena merupakan penghubung antara moda transportasi laut dan pedalaman. Peramalan beban kerja sangat penting dalam hal kedatangan truk untuk menghindari kemacetan dan kelancaran integrasi terminal petikemas dalam rantai pasokan. Operator terminal, cenderung membuat keputusan penumpukan berdasarkan sebagian besar faktor seperti berat, ukuran, dan jenis petikemas. Metodologi yang disarankan memerlukan pengumpulan data agregat dan penerapan *Artificial Neural Networks (ANN)* untuk mengidentifikasi faktor penentu *Dwell Time (DT)*. Hasil pertama dari ANN menunjukkan bahwa faktor terpenting yang mempengaruhi akurasi model secara signifikan adalah sebagai berikut: ukuran dan jenis petikemas, hari dan bulan pembongkaran petikemas, pelabuhan asal kapal dan komoditas yang diangkut. Dari perbandingan model Jaringan Saraf Tiruan yang berbeda, hasilnya dapat disimpulkan bahwa semakin banyak informasi yang tersedia semakin baik akurasi model. Oleh karena itu, operator terminal disarankan untuk mengumpulkan sebanyak mungkin informasi tentang setiap petikemas yang mereka tangani agar dapat digunakan untuk pengembangan model prediktif.

2.16.3 Analisis Prediksi Dwell Time Of Import Containers Menggunakan Artificial Neural Network Pada Terminal Petikemas Semarang

Penelitian Sakinah pada tahun 2019 ini bertujuan untuk mengidentifikasi dan menjelaskan faktor faktor berpengaruh terhadap *dwell time* dan cara menangani *dwell time* yang dipengaruhi *month of discharge*, *day of discharge*, *container's size*, *container's type*, *yard occupancy* dengan di Terminal Petikemas Semarang. Metode yang digunakan dalam penelitian ini adalah metode *artificial neural network (ANN)*. Metode ini mengadopsi prinsip kerja otak manusia dalam memecahkan dan mampu menghasilkan hasil prediksi yang akurat melalui iannya dalam meminimalkan error pada model. Menurut Sarakolaei dkk. in 2012, ANN dapat memberikan hasil lebih baik, saat data dan variabel



independen dan dependen tidak ada hubungan matematis. Hasil dari penelitian ini menunjukkan bahwa faktor yang paling penting mempengaruhi secara signifikan akurasi model adalah sebagai berikut: *day of discharge, container's type, month of discharge, container's size, yard occupancy*.

2.16.4 Prediksi Waktu Sandar Kapal di Pelabuhan Batu Ampar, Kota Batam, Provinsi Kepulauan Riau

Penelitian Candra dkk. tahun 2019 ini melakukan prediksi waktu sandar kapal (*dwelling time*) di pelabuhan Batu Ampar dengan bantuan *data mining*. Pelabuhan Batu Ampar merupakan pelabuhan barang terbesar di Kota Batam yang memiliki lalu lintas terbaik untuk kegiatan ekspor dan kegiatan penting. Waktu tunggu masih menjadi masalah dalam layanan pelabuhan. Waktu tunggu merupakan salah satu indikator efisiensi pengelolaan pelabuhan. Rata-rata waktu tunggu pelabuhan Batu Ampar untuk kegiatan bongkar pada Triwulan I tahun 2015 adalah 7 hari, sedangkan kegiatan bongkar muat adalah 5 hari. Hal inilah yang membuat pelabuhan Batu Ampar masih banyak dikeluhkan sehingga mengakibatkan banyaknya antrian kapal. Untuk itu diperlukan analisis untuk menghasilkan model yang dapat memberikan gambaran waktu tunggu di pelabuhan dan mengevaluasi model analitik yang dibangun. Analisis data sekunder pelabuhan Batu Ampar menggunakan *data mining*. Metode *data mining* dilakukan dengan menggunakan metode pembelajaran terbimbing yaitu regresi berganda dan pohon keputusan. Tujuan umum dari regresi berganda adalah untuk lebih jauh tentang hubungan antara variabel independen atau prediktor dan variabel atau kriteria dependen. Pohon keputusan yang digunakan untuk mengekstrak data port menggunakan klasifikasi ini. Klasifikasi pohon keputusan dapat menemukan data yang mengandung kelas objek yang terkumpul dengan baik, sehingga kelas dapat diinterpretasikan secara lengkap dalam konteks teori substantif. Dua metode evaluasi model dilakukan untuk dua hasil pemodelan yang dibangun. Uji Analisis Varians (ANOVA) digunakan untuk mengevaluasi model regresi berganda, sedangkan untuk model pohon keputusan dievaluasi dengan matriks konfusi. Hasil data menunjukkan lama kapal bongkar/muat dengan tiga variabel jenis , bendera, dan volume. Dengan menggunakan regresi berganda, dihasilkan prediksi waktu berlabuh kapal. Hasil evaluasi, model dibuat signifikan.



Dengan tingkat kepercayaan 95% model prediksi yang dibuat akan mewakili nilai sebenarnya. Untuk pohon keputusan, model evaluasi yang dibuat sudah sesuai dengan presisi sebesar 84,50%.

