PREDIKSI *SERVICE LEVEL AGREEMENT* WAKTU TEMPUH KIRIMAN BARANG DAN DOKUMEN PADA PT POS INDONESIA DENGAN METODE *RANDOM FOREST HYBRID*

THESIS

Oleh:
MUHAMMAD ISA ANSORI
NIM. 210605220002

PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2023

PREDIKSI *SERVICE LEVEL AGREEMENT* WAKTU TEMPUH KIRIMAN BARANG DAN DOKUMEN PADA PT POS INDONESIA DENGAN METODE *RANDOM FOREST HYBRID*

THESIS

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Magister Komputer (M.Kom)

Oleh:
MUHAMMAD ISA ANSORI
NIM. 210605220002

PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2023

PREDIKSI *SERVICE LEVEL AGREEMENT* WAKTU TEMPUH KIRIMAN BARANG DAN DOKUMEN PADA PT POS INDONESIA DENGAN METODE *RANDOM FOREST HYBRID*

THESIS

Diajukan kepada:
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Magister Komputer (M.Kom)

Oleh:
MUHAMMAD ISA ANSORI
NIM. 210605220002

PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2023

PREDIKSI *SERVICE LEVEL AGREEMENT* WAKTU TEMPUH KIRIMAN BARANG DAN DOKUMEN PADA PT POS INDONESIA DENGAN METODE *RANDOM FOREST HYBRID*

THESIS

Oleh:
MUHAMMAD ISA ANSORI
NIM. 210605220002

Telah diperiksa dan disetujui untuk di uji:
Tanggal 12 Mei 2023

Pembimbing I,	Pembimbing II,
Dr. Ririen Kusumawati, M.Kom	Dr. M. Amin Hariyadi, M.T
NIP. 19720309 200501 2 002	NIP. 19670018 200501 1 001

Mengetahui,
Kepala Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

EDIKSI *SERVICE LEVEL AGREEMENT* WAKTU TEMPUH KIRIMAN BARANG DAN DOKUMEN PADA PT POS INDONESIA DENGAN METODE *RANDOM FOREST HYBRID*

THESIS

Oleh:
MUHAMMAD ISA ANSORI
NIM. 210605220002

Telah Dipertahankan di Depan Dewan Penguji Thesis
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Magister Komputer (M.Kom)
Tanggal 12 Mei 2023

Susunan Dewan Penguji			Tanda Tangan
Penguji Utama	:	Dr. Sri Harini, M. Si NIP 19731014200112 2 002	( )
Ketua Penguji	:	Dr. Irwan Budi Santoso, M.Kom NIP. 19770103 201101 1 004	( )
Sekretaris Penguji	:	Dr. Ririen Kusumawati, M.Kom NIP. 19720309 200501 2 002	( )
Anggota Penguji	:	Dr. M. Amin Hariyadi, M.T NIP. 19670018 200501 1 001	( )

Mengetahui dan Mengesahkan,
Kepala Program Studi Magister Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan dibawah ini:

Nama : Muhammad Isa Ansori

NIM : 210605220002

Program Studi : Magister Informatika

Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Thesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan Thesis ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 12 Mei 2023
Yang membuat pernyataan,

Muhammad Isa Ansori
NIM. 210605220002

MOTTO

"Buatlah sesuatu yang sulit menjadi mudah, dengan kemudahan itu akan membuka kreatifitas baru yang tidak terbatas"

PERSEMBAHAN

Dengan mengucapkan syukur Alhamdulillahi rabbil alamin, Thesis ini saya persembahkan untuk :

1. Seluruh Keluarga tercinta yang selalu memberikan doa dan semangat.
2. Seluruh Civitas Akademika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang telah memberikan kesempatan untuk menambah ilmu teknologi dan agama.
3. Seluruh Pimpinan dan karyawan PT Pos Indonesia (Persero) yang memberikan dukungan data dan kesempatan belajar.
4. Seluruh Civitas Akademika Politeknik Negeri Malang yang memberikan dorongan untuk meningkatkan kemampuan bidang teknologi.
5. Seluruh rekan-rekan mahasiswa Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang semua angkatan atas kerjasamanya selama ini.
6. Bapak, ibu, saudara dan rekan-rekan sekalian yang tidak bisa saya sebutkan satu persatu dalam mendukung Thesis ini hingga bisa diselesaikan.

KATA PENGANTAR

Assalamu'alaikum Wr. Wb

Syukur *Alhamdulillah* penulis haturkan kehadirat Allah SWT yang telah melimpahkan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan studi di Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang sekaligus menyelesaikan Thesis ini dengan baik.

Selanjutnya penulis haturkan ucapan terima kasih seiring do'a dan harapan *jazakumullah ahsanal jaza'* kepada semua pihak yang telah membantu terselesaikannya Thesis ini. Ucapan terima kasih ini penulis sampaikan kepada:

1. Ibu Dr. Ririen Kusumawati, M.Kom dan Bapak Dr. M. Amin Hariyadi, M.T, selaku dosen pembimbing Thesis, yang telah banyak memberikan pengarahan dan pengalaman yang berharga.
2. Segenap civitas akademika Program Studi Magister Informatika, terutama seluruh Bapak / Ibu dosen, terima kasih atas segenap ilmu dan bimbingannya.
3. Keluarga tercinta yang senantiasa memberikan do'a dan semangat
4. Semua rekan-rekan seperjuangan yang ikut mendukung dan membantu.

Penulis menyadari bahwa dalam penyusunan Thesis ini masih terdapat kekurangan dan penulis berharap semoga Thesis ini bisa memberikan manfaat kepada para pembaca khususnya bagi penulis secara pribadi. *Amiinn Yaa Rabbal Alamin*.

Wasalamu'alaikum Wr. Wb

Malang, 12 Mei 2023

Penulis

DAFTAR ISI

HALAMAN JUDUL i

HALAMAN PENGAJUAN ii

HALAMAN PERSETUJUAN iii

HALAMAN PENGESAHAN iv

HALAMAN PERNYATAAN v

MOTTO vi

HALAMAN PERSEMBAHAN vii

KATA PENGANTAR viii

DAFTAR ISI ix

DAFTAR GAMBAR xii

DAFTAR TABEL xiii

ABSTRAK xv

ABSTRACT xvi

مستخلص البحث xvii

BAB I PENDAHULUAN 1

1.1 Latar Belakang 1

1.2 Pernyataan Masalah 3

1.3 Tujuan Penelitian 4

1.4 Manfaat Penelitian 4

1.5 Batasan Masalah 5

BAB II STUDI PUSTAKA 6

2.1 Prediksi Waktu Tempuh 6

2.2 Service Level Agreement (SLA) 13

2.3 Kerangka Teori 14
BAB III METODOLOGI PENELITIAN 17
3.1 Desain Penelitian 17
3.2 *Data Collection* 18
3.2.1 Business Process Understanding 19
3.2.2 Data Understanding 20
3.2.3 Data Preparation 21
3.3 *Design System* 24
3.4 *Experiment* 25
3.5 Pengukuran Akurasi Prediksi 26
3.5.1 MAE (Mean Absolute Error) 26
3.5.2 MAPE (Mean Absolute Perecentage Error) 26
BAB IV METODE RANDOM FOREST 28
4.1 Desain 28
4.2 Ujicoba 30
4.2.1 Strategi Pertama 31
4.2.2 Strategi Kedua 37
4.3 Komparasi performa ujicoba strategi 1 dan 2 41
BAB V METODE SUPPORT VECTOR REGRESSION 43
5.1 Desain 43
5.2 Ujicoba 45
5.2.1 Strategi Pertama 46
5.2.2 Strategi Kedua 49
5.3 Komparasi performa ujicoba strategi 1 dan 2 52
BAB VI METODE RANDOM FOREST HYBRID 54
6.1 Desain 54
6.2 Ujicoba 55
6.2.1 Non Hybrid 55
6.2.2 Random Forest Hybrid 56
6.3 Komparasi performa ujicoba non hybrid dan RF hybrid 58

BAB VII PEMBAHASAN 59

7.1 Komparasi dan Performa Metode 59

7.1.1 Performa dan Komparasi metode RF dan SVR 59

7.1.2 Performa dan Komparasi metode RF Hybrid 63

7.2 Implementasi Hasil 68

7.3 Prediksi Service Level Agreement Kiriman Barang dan Dokumen dalam pandangan Islam 70

BAB VIII KESIMPULAN 73

8.1 Kesimpulan 73

8.2 Saran 75

DAFTAR PUSTAKA 76

DAFTAR GAMBAR

Gambar 2.1 Kerangka Teori 14
Gambar 3.1 Desain Penelitian 18
Gambar 3.2 Proses Bisnis Pengiriman Barang/Dokumen
(Sumber: Bagian Operasi PT Pos Indonesia, 2022) 20
Gambar 3.3 Capture data yang diolah 23
Gambar 3.4 Design System 24
Gambar 4.1 Flowchart proses prediksi FR 28
Gambar 4.2 Random Forest Diagram 30
Gambar 4.3 RF Plot Actual & Prediction strategi pertama 35
Gambar 4.4 RF Average Actual & Prediction strategi pertama 36
Gambar 4.5 Variable Importances 38
Gambar 4.6 RF Plot Actual & Prediction strategi kedua 40
Gambar 4.7 RF Average Actual & Prediction strategi kedua 41
Gambar 5.1 Flowchart proses prediksi SVR 43
Gambar 5.2 Cara Kerja SVR (Sethi 2020) 44
Gambar 5.3 Visualisasi SVR Strategi Pertama 49
Gambar 5.4 Visualisasi SVR Strategi Kedua 52
Gambar 6.1 Flowchart FR Hybrid (Sumber: data diolah peneliti) 54
Gambar 7.1 Komparasi Performa MAPE RF dan SVR 62
Gambar 7.2 Komparasi Performa MAE RF dan SVR 62
Gambar 7.3 Komparasi Performa MAPE RF, SVR, RF Hybrid 65
Gambar 7.4 Komparasi Performa MAE RF, SVR, FR Hybrid 65
Gambar 7.5 Komparasi nilai rata-rata RF, SVR, FR Hybrid 66
Gambar 7.6 Komparasi nilai MAX RF, SVR, FR Hybrid 67
Gambar 7.7 Komparasi nilai MIN RF, SVR, FR Hybrid 67

DAFTAR TABEL

Tabel 2.1 Daftar Jurnal 15
Tabel 3.1 Detil Dataset 22
Tabel 3.2 Variabel Dependen dan Indepenen 23
Tabel 3.3 Data Set Implementasi 25
Tabel 4.1 Hasil Prediksi RF 500 strategi pertama 32
Tabel 4.2 Hasil Prediksi RF 1000 strategi pertama 33
Tabel 4.3 Hasil Prediksi RF 1500 strategi pertama 33
Tabel 4.4 Performa metode RF strategi pertama 34
Tabel 4.5 Hasil Prediksi RF 500 strategi kedua 38
Tabel 4.6 Hasil Prediksi RF 1000 strategi kedua 39
Tabel 4.7 Hasil Prediksi RF 1500 strategi kedua 39
Tabel 4.8 Performa metode RF strategi kedua 40
Tabel 4.9 Komparasi Performa Srategi Pertama dan Kedua 42
Tabel 5.1 Hasil Prediksi SVR RBF strategi pertama 47
Tabel 5.2 Hasil Prediksi SVR Polynomial strategi pertama 47
Tabel 5.3 Performa metode SVR strategi pertama 48
Tabel 5.4 Hasil Prediksi SVR RBF strategi kedua 50
Tabel 5.5 Hasil Prediksi SVR Polynomial strategi kedua 51
Tabel 5.6 Performa metode SVR strategi kedua 51
Tabel 5.7 Komparasi Performa Srategi Pertama dan Kedua SVR 53
Tabel 6.1 Hasil Prediksi Non Hybrid 56
Tabel 6.2 Performa non hybrid 56
Tabel 6.3 Hasil Prediksi RF Hybrid 57
Tabel 6.4 Peforma RF Hybrid 57
Tabel 6.5 Performa Non Hybrid dan RF Hybrid 58
Tabel 7.1 Komparasi performa metode RF dan SVR 61
Tabel 7.2 Hasil perbandingan performa metode RF Hybrid 64
Tabel 7.3 Komparsi Performa Selisih Rata-rata, Max, Min RF dan RF Hybrid 66

Tabel 7.4 Komposisi Data Kiriman berdasar Tujuan..... 68
Tabel 7.5 Performa Nilai Rata-rata, Max, Min berdasar tujuan..... 69

ABSTRAK

Ansori, Muhammad Isa, 2023, Prediksi *Service Level Agreement* Waktu Tempuh Kiriman Barang dan Dokumen Pada PT Pos Indonesia Dengan Metode *Random Forest Hybrid,* Program Magister Informatika Universitas Islam Negeri Maulana Malik Ibrahim, Pembimbing: (1) Ririen Kusumawati, M.Kom (2) Dr. M. Amin Hariyadi, M.T

Kata Kunci : Prediksi, Random Forest, Service Level Agreement, Hybrid

Pekembangan bisnis pengiriman barang dan dokumen saat ini berkembang dengan pesat, terutama pada pengiriman barang online. Dengan berkembangnya industri ini tentunya tingkat persaingan antar pengguna jasa semakin tajam. Ketepatan waktu pengiriman menjadi alasan utama pelanggan untuk memilih jasa pengiriman yang terpercaya. PT Pos Indonesia sebagai Badan Usaha Milik Negara yang bergerak pada bidang jasa pengiriman barang dan dokumen harus meningkatkan kualitasnya untuk menghadapi kompetisi yang semakin sengit. Dengan menetapkan SLA *(Service Level Agreement)* untuk waktu tempuh kiriman, sebagai bukti perjanjian dengan *customer* dengan memberikan kepastian penyampaian kiriman. Namun dalam prakteknya masih terdapat beberapa keterlambatan dikarenakan penetapan SLA yang kurang tepat khususnya untuk pengiriman tujuan Pulau Kalimantan, Sulawesi, Maluku dan Papua. Dengan menggunakan metode *Random Forest* (RF) *Hybrid* untuk melakukan prediksi SLA yang lebih tepat. RF *Hybrid* terlebih dahulu dikelola dengan 2 (dua) metode terpilih yaitu metode *Random Forest* dan metode *Support Vector Regression* (SVR). Untuk nilai akurasi RF dengan pengukuran prediksi MAE dan MAPE mendapatkan nilai pengkuran akurasi MAE 0,6968 MAPE 15,7976, untuk SVR mendapatkan nilai pengukuran akurasi MAE 1,0407 MAPE 24,3927. Metode RF lebih baik dibandingkan dengan SVR, oleh karena itu metode RF dibuat menjadi RF *Hybrid* dengan melakukan perhitungan prediksi melalui nilai rata-rata dari kedua metode sebelumnya yaitu metode RF dan SVR. Dari perhitungan tersebut diapat nilai pengukuran akurasi untuk RF *Hybrid* MAE 0,7254 dan MAPE 16,77. Dari hasil pengujian tersebut RF dan RF *Hybrid* hampir memperoleh nilai yang sama dan masuk dalam kategora nilai yang baik dimana MAE berada pada kisaran 0 dan MAPE pada kisaran 10-20, walaupun metode RF masih nilai akurasi yang paling baik. Dengan hasil analisa pengukuran prediksi SLA ini diharapkan mampu memberikan dampak pada penetapan SLA di PT Pos Indonesia untuk menjaga kepercayaan dengan *customer*.

ABSTRACT

Ansori, Muhammad Isa, 2023, Prediction of Service Level Agreement Travel Time for Goods and Documents at PT Pos Indonesia Using Random Forest Hybrid Method, Master Study in Computer Science, State Islamic University Maulana Malik Ibrahim, Advisors: (1) Ririen Kusumawati, M.Kom (2) Dr. M. Amin Hariyadi, M.T

Keywords : Prediction, Random Forest, Service Level Agreement, Hybrid

The development of the goods and document delivery business is currently growing rapidly, especially in online goods delivery. With the development of this industry, of course, the level of competition between service users is getting sharper. Timely delivery is the main reason for customers to choose a trusted delivery service. PT Pos Indonesia as a State-Owned Enterprise engaged in the field of goods and document delivery services must improve its quality to face increasingly fierce competition. By establishing an SLA (Service Level Agreement) for the delivery time, as evidence of an agreement with the customer by providing certainty about the delivery of the shipment. However, in practice there are still some delays due to inaccurate SLA determination, especially for shipments to Kalimantan, Sulawesi, Maluku and Papua. By using the Hybrid Random Forest (RF) method to make more precise SLA predictions. RF Hybrid is first managed with 2 (two) selected methods, namely the Random Forest method and the Support Vector Regression (SVR) method. For the RF accuracy value with MAE and MAPE prediction measurements, the MAE accuracy measurement value is 0.6968 MAPE 15.7976, for SVR, the MAE accuracy measurement value is 1.0407 MAPE 24.3927. The RF method is better than the SVR, therefore the RF method is made into RF Hybrid by calculating the prediction through the average value of the two previous methods, namely the RF and SVR methods. From these calculations, the accuracy measurement values for RF Hybrid MAE are 0.7254 and MAPE are 16.77. From the test results, RF and RF Hybrid almost got the same score and were included in the good value category where MAE is in the range of 0 and MAPE is in the range of 10-20, although the RF method is still the best accuracy value. With the results of the SLA prediction measurement analysis, it is hoped that it will have an impact on the SLA determination at PT Pos Indonesia to maintain trust with customers.

مستخلص البحث

أنصاري، محمد عيسى، ٢٠٢٣ . توقع اتفاقية مستوى الخدمة (*Service Level Agreement*)، وقت السفر للبضائع والوثائق في شركة محدودة البريد الإندونيسي بطريقة الغابات الهجينة العشوائية (*Random Forest Hybrid*). بحث جامعي. برنامج الماجستير قسم الهندسة والمعلوماتية، الجامعة الإسلامية الحكومية مولانا مالك إبراهيم بمالانج. المشرفة: (١) ريرين كوسوماواتي، الماجستير (٢) المشرف: الدكتور أمين هريادي، الماجستير.

الكلمات المفتاحية: التوقع، الغابات الهجينة، اتفاقية مستوى الخدمة، العشوائية

تطوير أعمال تسليم البضائع والمستندات حاليا نموا سريعا، لاسيما في تسليم البضائع عبر الإنترنت. مع تطور هذه الصناعة، أصبح مستوى المنافسة بين مستخدم الخدمة أكثر حدة. التسليم في الوقت المناسب هو السبب الرئيسي للعملاء لاختيار خدمة توصيل موثوقة. شركة محدودة البريد الإندونيسي كمؤسسة مملوكة للدولة تعمل في مجال خدمات تسليم البضائع والمستندات، يجب عليها تحسين جودتها لمواجهة المنافسة الشرسة المتزايدة. من خلال إنشاء اتفاقية مستوى الخدمة (*SLA*) لوقت التسليم، كدليل على اتفاقية مع العميل من خلال توفير اليقين بشأن تسليم الشحنة. ومع ذلك، من الناحية العملية، لا تزال هناك بعض التأخيرات بسبب التحديد غير الدقيق لجيش تحرير السودان، خاصة بالنسبة للشحنات إلى كاليمانتان وسولاويزي ومالوكو وبابوا. باستخدام طريقة (*Random Forest Hybrid*) لعمل تنبؤات (*SLA*) أكثر دقة. تتم إدارة الغابات الهجينة العشوائية (*Random Forest Hybrid*) أولا باستخدام طريقتين، وهما طريقة الغابات الهجينة (*Random Forest*) ودعم الإنحدار المتجه (*Support Vector Regression*). للحصول على قيم دقة التردد اللأسلكي مع قياسات التنبؤ MAE و MAPE احصل على قيمة قياس دقة MAE تبلغ ٦٩٦٨و٠ MAPE ٧٩٧٦و١٥، بالنسبة إلى SVR، تحصل على قيمة قياس دقة MAE تبلغ ٠٤٠٧و١ MAPE ٣٩٢٧و٢٤. طريقة RF أفضل من طريقة SVR لذلك يتم تحويل طريقة التردد اللاسلكي إلى RF Hybrid من خلال حساب التنبؤ من خلال متوسط قيمة الطريقتين السابقتين، وهما طريقتا RF و SVR. من هذه الحسابات، قيم قياس الدقة لـ RF Hybrid MAE هي ٧٢٥٤و٠ و MAPE هي ٧٧و١٦. من نتائج الاختبار، حصل RF و RF Hybrid تقريبا على نفس النتيجة وتم تضمينهما في فئة القيمة الجيدة حيث يقع MAE في نطاق ٠ و MAPE في النطاق من ١٠ إلى ٢٠، على الرغم من أن طريقة RF لا تزال أفضل دقة قيمة. مع نتائج تحليل قياس التنبؤ باتفاقية مستوى الخدمة، من المأمول أن يكون لها تأثير على تحديد اتفاقية مستوى الخدمة في شركة محدودة البريد الإندونيسي للحفاظ على الثقة مع العملاء.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kepercayaan dan kepuasan pelanggan pada bisnis kurir atau pengiriman barang dan dokumen salah satunya di tentukan oleh ketepatan dan kecepatan penyampaian kiriman dimaksud. Oleh karena itu di setiap tujuan kiriman telah ditetapkan berapa lama kiriman tersebut sampai hingga tujuan. Adanya target waktu pengiriman ini merupakan perjanjian penyedia jasa ke pelanggan dalam bentuk *Service Level Agreement* (SLA) yang tertuang dalam setiap *Resi* kiriman dalam bentuk informasi estimasi waktu kiriman.(Hiles 2016)

SLA di PT Pos Indonesia (Persero) untuk kiriman paket atau dokumen sudah ditetapkan SLA yang dicantumkan pada resi pengiriman dengan jumlah waktu tempuh sesuai dengan produk dan tujuan kiriman yang dipilih oleh pelanggan berupa estimasi waktu tempuh kiriman. Produk Pos Kilat Khusus/Pos Reguler adalah layanan dengan jumlah transaksi cukup tinggi dengan variasi SLA yang cukup beragam antara 2-14 hari. Penetapan SLA ini dihitung dari barang atau dokumen diterima diloket hingga barang atau dokumen tersebut sampai ke penerima dengan memperhitungkan jumlah *hub* yang dilewati.

Namun pada kenyataannya terdapat beberapa keterlambatan untuk SLA tujuan tertentu khususnya wilayah pulau Kalimantan, Sulawesi, Maluku dan

Papua. Moda transportasi yang digunakan untuk kiriman paket dan dokumen tersebut sudah ditetapkan jadwal keberangkatan dan kedatangannya baik melalui jalur darat, laut dan udara. Dalam hal ini ketidaktepatan SLA untuk tujuan dimaksud bukan berdasarkan pengaruh moda transportasi ataupun jumlah *volumetrik* kiriman, akan tetapi penetapan SLA belum dilakukan pengkajian ulang sehingga terdapat beberapa kendala ketidaktepatan penyampaian kiriman yang sudah dijanjikan.

Ketepatan SLA ini selaras dengan konsep Islam, bahwa transaksi pengiriman di PT Pos Indonesia menurut hukum islam sesuai dengan *akad wakalah* yang disertai adanya upah *(ujrah)* atau lengkapnya akad *wakalah bil ujrah* sesuai dengan ketentuan yang dan menjalankan pross pengiriman sesuai amanah yang dan tanggung jawab yang diberikan. Jadi transaksi pengiriman PT Pos telah berjalan sesuai amanah dan tanggung jawab yang baik maka transaksi pengiriman barang atau dokumen di PT Pos Indonesia telah sesuai dengan akad *wakalah bil ujrah*.

Adapun yang dijadikan dasar hukum *wakalah* adalah Firman Allah QS. Al-kahfi (18): 19

وَكَذَٰلِكَ بَعَثْنَٰهُمْ لِيَتَسَآءَلُواْ بَيْنَهُمْ ۚ قَالَ قَآئِلٌ مِّنْهُمْ كَمْ لَبِثْتُمْ ۖ قَالُواْ لَبِثْنَا يَوْمًا أَوْ
بَعْضَ يَوْمٍ ۚ قَالُواْ رَبُّكُمْ أَعْلَمُ بِمَا لَبِثْتُمْ فَٱبْعَثُوٓاْ أَحَدَكُم بِوَرِقِكُمْ هَٰذِهِۦٓ إِلَى ٱلْمَدِينَةِ
فَلْيَنظُرْ أَيُّهَآ أَزْكَىٰ طَعَامًا فَلْيَأْتِكُم بِرِزْقٍ مِّنْهُ وَلْيَتَلَطَّفْ وَلَا يُشْعِرَنَّ بِكُمْ أَحَدًا

"Dan demikianlah Kami bangunkan mereka agar mereka saling bertanya di antara mereka sendiri. Berkatalah salah seorang di antara mereka: Sudah berapa lamakah kamu berada (disini?)". Mereka menjawab: "Kita berada (disini) sehari atau setengah hari". Berkata (yang lain lagi): "Tuhan kamu lebih mengetahui berapa lamanya kamu berada (di sini). Maka suruhlah salah seorang di antara kamu untuk pergi ke kota dengan membawa uang perakmu ini, dan hendaklah dia lihat manakah makanan yang lebih baik, maka hendaklah ia membawa makanan itu untukmu, dan hendaklah ia berlaku lemah-lembut dan janganlah sekali-kali menceritakan halmu kepada seorangpun" (Qs: Al-Kahfi : 19).

Ayat ini menerangkan bahwa perginya salah seorang dari ketiga ash-habul kahfi yang bertindak untuk dan atas nama rekan-rekannya sebagai wakil mereka dalam memilih dan membeli makanan. Dapat diartikan pula bahwa ayat ini mengandung pesan bahwa boleh melakukan perjanjian kerja atau mewakilkan suatu pekerjaan kepada orang lain yang memang dapat dipercaya dan telah ahli dalam bidang pekerjaan tersebut. Oleh karena itu dalam mengemban Amanah PT Pos Indonesia wajib memberikan ketepatan penyampaikan kiriman sesuai SLA yang diamanahkan.

Berkenaan dengan hal tersebut perlu dilakukan revisi atau analisa terhadap penetapan SLA dengan metode *machine learning* (Heidy 2020) agar didapat prakiraan SLA yang lebih tepat. Dengan *machine learning* diharapkan mampu memprediksi SLA yang terbaik, yang berdampak terhadap kepercayaan pelanggan terhadap layanan yang diberikan. Oleh karena itu penulis melakukan penelitian terhadap peningkatan ketepatan SLA ini dengan melakukan prediksi waktu tempuh dengan metode *Random Forest (RF)* secara *hybrid*.

1.2 Pernyataan Masalah

Penentuan SLA dengan menghitung waktu tempuh kiriman adalah hal yang paling penting sebagai dasar perjanjian kiriman terhadap pelanggan. Berdasarkan hal tesebut peneliti melakukan penelitian untuk menghitung prediksi waktu tempuh kiriman agar penentuan SLA lebih tepat dengan mengolah data yang sudah ada.

Berikut pernyataan masalah yang akan dibahas oleh peneliti :

1. Bagiamana mendapatkan model prediksi SLA yang optimal dalam menentukan waktu tempuh kiriman.
2. Bagaimana hasil prediksi SLA dapat memberikan ukuran waktu tempuh kiriman yang lebih tepat.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Menganalisa model prediksi SLA dengan pengolahan data pengiriman yang paling berpengaruh untuk mendapatkan hasil yang optimal.
2. Mendapatkan hasil prediksi SLA yang optimal untuk dijadikan penetapan SLA yang baru.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Untuk mengetahui prediksi mutu waktu tempuh kirimannya.
2. Pencegahan dini jika terdapat mutu ketepatan kiriman yang cenderung terlambat untuk segera dilakukan perbaikan.
3. Bahan kajian untuk menganalisa mutu waktu tempuh kiriman dengan kompetitor.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:.

1. Obyek produk layanan yang diteliti adalah Pos Kilat Khusus atau Pos Reguler yang dikirim dari Malang ke wilayah regional 6 meliputi wilayah Kalimantan, Sulawesi, Maluku dan Papua.
2. Data yang digunakan data transaksi bulan Januari 2022 sampai dengan Desember 2022.

BAB II

STUDI PUSTAKA

2.1 Prediksi Waktu Tempuh

Prediksi waktu tempuh kiriman, dibahas dalam beberapa jurnal diantaranya yaitu:

Vanajakshi *and* Rilett (2007), menjelaskan bahwa sebagian besar sistem transportasi perkotaan di Amerika Utara dilengkapi dengan pengawasan lalu lintas sistem yang memberikan informasi lalu lintas waktu nyata ke lalu lintas pusat manajemen. Informasi seperti lalu lintas saat ini maupun masa depan diprediksi berdasarkan data real time untuk memberi tahu pengemudi seperti waktu perjalanan. Makalah ini menyelidiki kegunaan SVM untuk prediksi jangka pendek dari waktu perjalanan. Sebuah perbandingan dilakukan antara kinerja SVM dan ANN, historic method and real time method. Model ANN yang digunakan adalah *multi layer feed forward neural network* dan model SVM yang digunakan adalah support vector regression dengan radial basis kernel function. Perbandingan antara ANN dan SVM menunjukkan bahwa kinerja SVM dan ANN dapat dibandingkan satu sama lainnya. SVM menjadi pilihan yang lebih baik untuk jangka pendek prediksi waktu perjalanan, jika jumlah data pelatihan lebih sedikit, atau ketika data pelatihan memiliki lebih banyak variasi dibandingkan dengan data pengujian. Juga, ditemukan bahwa pengaruh dari jumlah data latih yang digunakan lebih banyak pada ANN metode dari pada metode SVM. Secara keseluruhan, ditemukan bahwa SVM adalah

alternatif yang layak untuk ANN untuk prediksi jangka pendek masalah ketika jumlah data kurang.

Quah *et al.* (2019), dalam penelitiannya menjelaskan bahwa dengan meningkatnya tren belanja online mendesak kebutuhan yang terus menerus meningkatkan pengalaman pengguna dan pengiriman tepat waktu barang adalah salah satu bidang utama. Di dalam penelitian ini mengeksplorasi *machine learning* dalam memprediksi keterlambatan pengiriman barang pada jasa kurir nasional Malaysia bernama Poslaju. Model prediksi juga memungkinkan visualisasi tren pengiriman untuk Poslaju Malaysia. Hasil penelitian menunjukkan bahwa *decision tree* dan metode KNN (dengan nilai K lebih tinggi) memiliki ukuran presisi dan perolehan yang lebih baik dibandingkan dengan metode Naïve Bayes.

Wu, Ho, *and* Lee (2004), dalam penelitiannya menyampaikan bawah waktu perjalanan adalah ukuran mendasar dalam transportasi. Prediksi waktu tempuh yang akurat juga sangat penting untuk pengembangan sistem *Smart Transportasion* dan *advanced traveler information systems*. Dalam penelitian ini, menerapkan *Support Vector Regression* (SVR) untuk prediksi waktu perjalanan dan membandingkan hasilnya dengan metode prediksi waktu perjalanan menggunakan data trafik Jalan Raya. Diyakini bahwa SVR akan bekerja dengan baik untuk analisis deret waktu, dibandingkan dengan prediktor dasar lainnya, hasil penelitiannya menunjukkan prediktor SVR dapat secara signifikan mengurangi kesalahan rata-rata relatif dan kesalahan *root-mean-squared* dari waktu perjalanan yang diprediksi. Hasil percobaan prediksi waktu

perjalanan dalam jarak pada jam sibuk seperti yang diharapkan, prediktor rata-rata historis tidak dapat mencerminkan pola lalu lintas yang sangat berbeda dari rata-rata sebelumnya dan prediktor saat ini biasanya lebih lambat untuk mencerminkan perubahan pola lalu lintas. Prediktor SVR bekerja sangat baik dalam eksperimennya, hasilnya menunjukkan *RME* dan *RMSE* berbeda prediktor untuk jarak perjalanan yang berbeda di semua titik data set pengujian. Dalam penelitian ini menunjukkan bahwa prediktor SVR mengurangi keduanya *RME* dan *RMSE* menjadi berkurang dari setengah yang dicapai oleh waktu saat ini dan prediktor rata-rata historis untuk semua jarak yang berbeda. Untuk alasan ini, peneliti secara khusus memeriksa pengujian titik data di mana kesalahan prediksi dari setiap prediktor lebih besar dari atau sama dengan 5%. Sedangkan dengan SVR, prediktor tidak hanya meningkatkan kinerja keseluruhan, tetapi juga secara signifikan mengurangi kesalahan prediksi untuk kasus-kasus di mana ada kesalahan prediksi yang lebih buruk di salah satu prediktor.

Bristy, Durjoy, and Hasan (2021), tujuan dari penelitian ini adalah untuk menentukan fleksibilitas eCommerce, karena telah menjadi keuntungan bagi perekonomian negara mana pun dalam beberapa tahun terakhir . Fleksibilitas ditentukan oleh berbagai faktor, tetapi pengiriman produk tepat waktu sangat penting bagi pelanggan dan pemasok. Makalah ini membuat model prediktif yang akan bekerja melalui kumpulan data lengkap dari perusahaan berbasis online mana pun dan hasil dari prediksi, otoritas akan mengambil langkah yang tepat untuk memperbaiki kondisi tersebut. Penelitian mengembangkan model prediksi menggunakan algoritma machine learning dan

menerapkan beberapa teknik data mining. Penelitian memperoleh hasil terbaik untuk support vector machine dengan akurasi prediksi sebesar 69% daripada pengklasifikasian random forest dengan hasil klasifikasi sekitar 65%.

Kisgyörgy *and* Rilett (2002), Sistem Manajemen Lalu Lintas Lanjutan San Antonio, Texas, yang disebut Sistem *TransGuide* menggunakan sistem sensor yang dipasang di 26 mil jalan raya untuk memasukkan data ke jaringan komputer berkecepatan tinggi.. Tujuan dari penelitian ini adalah membangun model prediksi waktu tempuh jalan tol secara real-time jaringan San Antonio berdasarkan informasi yang dikumpulkan oleh sensor loop dan sistem GPS. Prediksi waktu perjalanan dari model tersebut dapat menjadi dasar dari sistem manajemen lalu lintas selanjutnya juga digunakan oleh sistem informasi wisatawan. Kekokohan dan keakuratan model merupakan fitur yang penting, karena sistem manajemen lalu lintas bergantung pada penerimaan dan kepatuhan pengemudi menjadi efektif. Penelitian ini pertama-tama mengkaji penggunaan Modular Neural Networks (MNN) untuk meramalkan beberapa periode fitur rekayasa lalu lintas, seperti kecepatan, okupansi, volume dan kemudian menentukan waktu tempuh yang diharapkan berdasarkan nilai prediksi ini. Kedua, waktu perjalanan beberapa periode diprediksi langsung dari data loop dengan MNN. Model diuji dan dilatih pada waktu perjalanan aktual dari San Antonio, dikumpulkan oleh sistem data GPS. Kemudian hasil dari kedua model tersebut dibandingkan satu sama lain dan dengan hasil waktu tempuh standar model prediksi.

Hudzaifah *and* Rismayadi (2021), dalam penelitiannya melakukan peramalan terhadap kondisi lalu lintas, dimana dalam waktu tertentu kepadatan lalu lintas terjadi begitu saja sesuai dengan tingkat kesibukan pengguna jalan, disamping itu kondisi cuaca adalah faktor dari kepada lalu lintas. Dengan menggunakan metode peramalan time series decomposition dapat memberikan informasi tentang kepadatan arus lalu lintas yang akan dilewati. Variabel waktu sebagai variabel yang akan diolah dengan variabel lainnya untuk menentukan hasil dari prakiraan dengan menganalisa empat pola yaitu polo siklis, horizontal, trend dan seasonal. Data waktu tempuh dan kondisi cuaca merupakan data yang akan dihitung dengan metode time series decomposition. Hasil dari penelitian ini diterapkan pada aplikasi mobile berbasis android agar dapat digunakan secara mobile terutama pada saat berkendara. Hasil dari penelitian ini dengan menggunakan RMSE nilai errornya sebesar 3.80%, hal ini membuktikan bahwa metode time series dicomposition dapat digunakan sebagai metode peramalan.

Rio *et al*. (2022), dalam penelitiannya menjelaskan bahwa transportasi laut adalah transportasi yang paling efektif digunakan oleh masyarakat Indonesia. Salah satu transportasi laut yang digunakan untuk mengangkut barang, ataupun orang adalah kapal motor (KM). Pada artikel ini dibahas prediksi jarak tempuh harian salah satu kapal motor, yaitu kapal motor Sangiang dengan rute dari Bitung ke Ternate. Prediksi jarak tempuh harian (mile) menggunakan metode supervised machine learning dengan tiga variabel bebas yaitu 'Waktu Jangkar (menit)', 'Kecepatan (knot/jam)', dan 'Waktu

Berlayar (menit)'. Dari beberapa metode yang digunakan maka diperoleh model multiple regression sebagai model terbaik untuk memprediksi jarak tempuh kapal motor Sangiang.

Florio (2021), dalam penelitiannya menyajikan kerangka pembelajaran mesin untuk mengoptimalkan rute pengiriman last-mile. Bertentangan dengan kebanyakan pengaturan masalah perutean kendaraan di mana tujuan fungsi didefinisikan dengan jelas, dalam pengaturan dunia nyata yang dipertimbangkan dalam tantangan, tujuan tidak ditentukan secara eksplisit dan harus disimpulkan dari data. Memanfaatkan teknik dari pembelajaran mesin dan klasik heuristik masalah penjual keliling, makalah ini mengusulkan algoritme "kumpulkan dan pilih" untuk meresepkan kualitas tinggi urutan pengiriman last-mile. Pada fase pooling, dilakukan eksploitasi pengetahuan struktural yang diperoleh dari data, seperti sebagai daerah masuk dan keluar umum yang diamati dalam rute pelatihan. Pada fase seleksi, memprediksi skor urutan kandidat dengan model regresi berdimensi tinggi, pra-terlatih, dan teregulasi. Nilai model prediksi, yang mencakup sejumlah besar variabel prediktor seperti durasi urutan, kepatuhan dengan time windows, earliness, lateness, dan kesamaan struktural dengan data pelatihan, menampilkan prediksi yang baik akurasi dengan out-of-sample mean squared error sebesar 0,00121. Secara keseluruhan, kerangka kerja ini mampu meresepkan rute pengiriman kompetitif, yang diukur dengan skor rata-rata di luar sampel sebesar 0,0474 di beberapa kumpulan data pengembangan. Selain itu, kerangka hanya membutuhkan sebagian kecil dari anggaran komputasi maksimum

diperbolehkan dalam tantangan: kurang dari 40 menit diperlukan untuk melatih model prediksi skor, dan seterusnya rata-rata kurang dari dua detik diperlukan untuk menentukan urutan kunjungan yang diberikan rute uji.

Pambudi *et al*. (2020), dalam penelitian menjelaskan bahwa dalam industri bisnis logistik ketepatan waktu penyampaian barang kiriman adalah nilai yang paling penting bagi kepercayaan pelanggan untuk terus menggunakan jasa pengiriman barang dari perusahaan tersebut. Terkumpulnya data-data transaksi yang besar dengan kemajuan teknologi dapat digunakna oleh suatu perusahaan untuk melakukan analisa prediksi dari jasanya dalam hal ini jasa pengiriman barang. Dengan menggunakan metodologi peneliitian CRISP-DM (Cross Industry Standard Process for Data Mining) dapat dilakukan proses machine learning untuk mengetahui nilai akurasi dari prediksi yang dilakukan. Dalam penelitiannya hasil prediksi yang dilakukan dengan menggunakan Random Forest mendapatkan nilai akurai 76,6% metode regresi logistik 73,81% dan meode ANN 72,84%, Metode random forest mempunyai nilai yang lebih baik dibandingkan dengan regresi logistik dan ANN. Penelitian ini tentunya dibutuhkan pengembangan dengan menggunakan beberapa metode lainnya untuk menghasilkan nilai akurasi yang lebih optimal.

Megantara *et al.* (2022), dalam penelitian menyampaikan bahwa di Kota Bandung telah diluncurkan Bus Rapid Transit untuk menghindari kemacetan lalu lintas terutama pada jam sibuk sekolah dan berangkat kerja begitu juga pulang sekolah dan bekerja. Keberadaan Bus Rapid Transit masih belum optimal dikarenakan tingkat kedatangan antar halte yang tidak tepat, hal

ini dikarenakan bahwa Bus tesebut masih menggunakan jalur yang sama dengan kendaraan lainnya, lain hal nya dengan Bus Trans Jakarata yang mempunyai jalur sendiri sehingga waktu kedatangan dan keberangkatan dapat dijadwalkan dengan tepat. Kondisi ini membuat peneliti tertantang untuk melakukan penelitian terhadap Bus Rapid Transit dengan menggukan mikrokomputer untuk data tracking. Dari data yang dihasilkna tersebuat nanti diolah dengan menggunakan machine learning untuk memprikdi waktu keberangkatan dan kedatangan. Metode yang digunakan adalan metode Random Forest dan XGBoost, dari kedua metode ini nanti akan dipilih mana metode yang paling baik tingkat akurasi peramalannya. Untuk Random Forset dengan nilai random state 102 mendapatkan akurasi cukup tinggi yaitu sebesar 98% dengan nilai error MAE sebesar 0,95, MSE 33,63 dan RMSE 5,80, dengan demikian Random Forest digunakan sebagai metode yang paling optimal dalam penelitian ini.

2.2 Service Level Agreement (SLA)

Merupakan perjanjian antara konsumen dan penyedia jasa dalam penggunaan jasanya agar dapat jadikan ukuran service yang dilakukan. Wustenhoff (2002). Dalam SLA berisi beberapa ketetapan-ketetapan yang disepakati konsumen dan penyedia dalam bentuk surat perjanjian dan atau ketentuan tertulis lainnya.

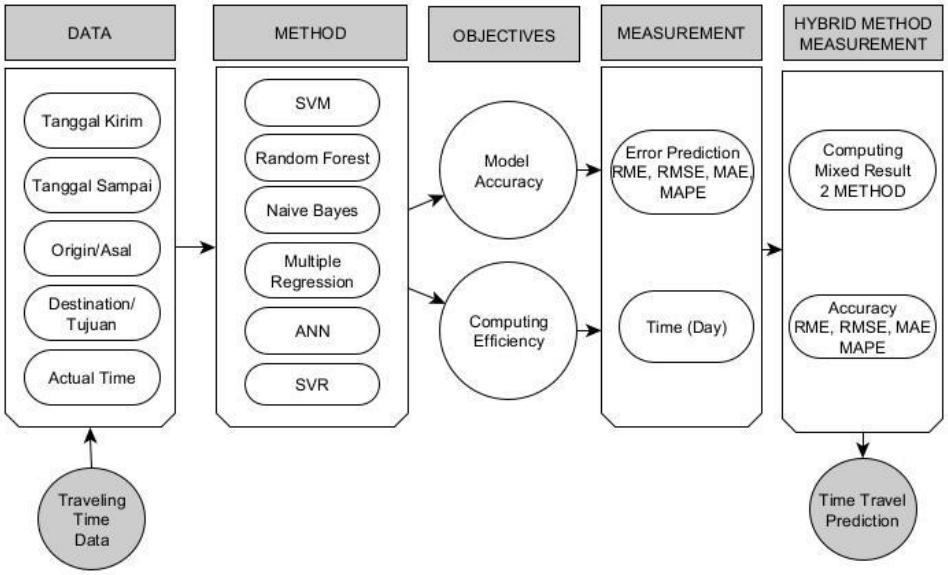
Menurut Wustenhoff, terdapat 5 aspek utama agar SLA dapat dikatakan baik, yaitu; 1. Jasa yang disediakan apakah cukup menjanjikan.

2. Dalam mewujudkan janjinya apa yang dilakukan oleh penyedia jasa. 3. Terdapat ketentuan yang jelas siapa yang mengukur jasa dimaksud. 4. Bagaimana jika jasa yang dijanjikan gagal dan apa yang akan dilakukan. 5. Upaya yang dilakukan jika terdapat SLA yang berubah sewaktu-waktu.

Dalam penelitian ini PT Pos Indonesia sebagai penyedia jasa dan konsumen atau pelanggannya adalah para pengguna jasa layanan kiriman yang membutuhkan kepastian dan ketepatan kirimannya.

2.3 Kerangka Teori

Dalam melakkan penelitian tentang prediksi kiriman perlu megacu pada jurnal-jurnal sebelumnya yang telah melakukan seabagai kerangka teori seperti terlihat pada Gambar 2.1 :



Gambar 2.1 Kerangka Teori

Kemudian langkah selanjutnya adalah mencoba beberapa metode yang telah dilakukan pada penelitian sebelumnyadalam melakukan prediksi pengiriman. Adapun beberapa penelitian atau jurnal yang menjadi acuan dalam penelitian ini dapat dilihat pada Tabel 2.1 :

Tabel 2.1 Daftar Jurnal

No	Nama Penulis & Tahun	Metode Penelitian	Judul Jurnal	Hasil Penelitian
1	Vanajakshi *et al.* (2007)	Support Vector Machine (SVM)	Support Vector Machine Technique for the Short Term Prediction of Travel Time	85.2%
2	Jo Wei Quah *et al*.(2019)	• Decision Tree (DT) • K-Nearest Neighbors (KNN) • Naïve Bayes (NB)	Timing-of-Delivery Prediction Model to Visualize Delivery Trends for Pos Laju Malaysia by Machine Learning Techniques	• 73% • 72% • .67%
3	Chun-Hsin Wu, *et al* (2004),	Support Vector Regression (SVR)	Travel-Time Prediction With Support Vector Regression	RME 1.21 RMSE 1.63
4	Anika Hossain Bristy (2021)	Support Vector Machine (SVM)	E-commerce shipping: Prediction of on time delivery of products using Data Mining	65%
5	Kisgyörgy and Rilett (2002)	Advanced Neural Network (ANN)	Travel Time Prediction By ANN	• 73%
6	Muhammad Hudzaifah, Ali Akbar Rismayadi (2021)	Time Series Decomposition	Peramalan Arus Lalu Lintas Berdasarkan Waktu Tempuh Dan Cuaca Menggunakan Metode Time Series Decomposition	RMSE 3.80%.
7	Afrioni Roma Rio, Berton Maruli Siahaan (2022)	• Multiple Regression • Lasso Regression • KNN Regression	Jarak Tempuh Kapal Motor Sangiang Menggunakan Supervised Machine Learning	RMSE • 9,93 • 9,8 • 39,6
8	Florio *et all.* (2021)	Sequence Selection	A Machine Learning Framework for Last-Mile Delivery Optimization	71%

No	Nama Penulis & Tahun	Metode Penelitian	Judul Jurnal	Hasil Penelitian
9	Liu *et al*. (2021)	• RF • ANN • Regresi Logistik	Prediksi Status Pengiriman Barang Menggunakan Metode Machine Learning	• 76.6% • 73,8% • 72,8%
10	Rendy Munadi, Sussi (2022)	RF	Prediksi Waktu Tempuh Bus Trans Metro Bandung dengan Internet Of Things dan Metode Machine Learning	98%

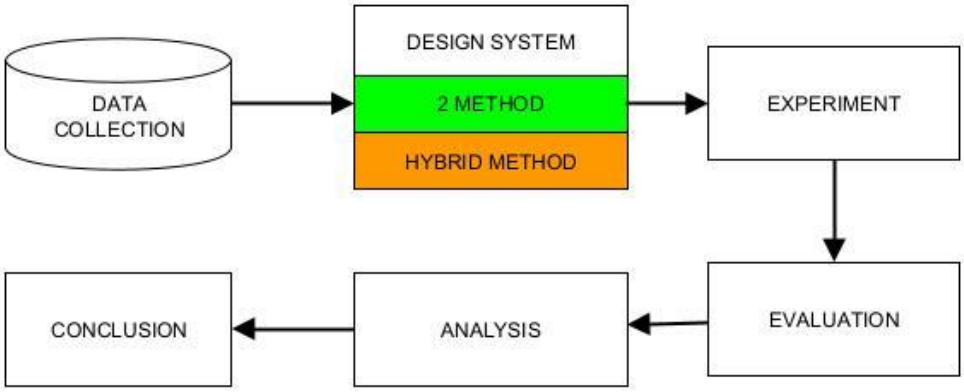
Berdasarkan Tabel 2.1 pada penelitian sebelumnya menunjukan bahwa algoritma *Random Forest (RF)* dan *Support Vector Regression (SVR)* memiliki performa yang terbaik diantara metode lain. Hal ini dibuktikan dengan nilai akurasi 98% untuk RF dan *error prediction* rendah 1,2% untuk SVR. *Hybrid Method* menggunakan RF karena mempunyai kemiripan dengan obyek yang diteliti serta hasil yang optimal,

BAB III

METODOLOGI PENELITIAN

3.1 Desain Penelitian

Beberapa tahapan dalam penelitian ini adalah melakukan beberapa tahapan secara urut sehingga proses desain hingga implementasi dapat dilakukan dimulai dari hal yang mendasar hingga hasil yang optimal. Tahapan dimulai dengan *Data Collection* yang awali dengan memahami *business process* untuk mengetahui dan memahami proses bisnis yang ada, dilanjutkan pemahaman data (*data understanding)* dan persiapan data *(data preparation)*. Setelah pengumpulan data kemudian dilakukan desain sistem, dilanjutkan *experiment* model dengan algoritma *random forest* (RF) dan *support vector regression* (SVR) dari hasil kedua algoritma tersebut akan dilakukan perhitungan secara *hybrid* pada algortima RF dengan mebandingkan rata rata hasil dari metode RF dan SVR. Selanjutnya adalah melihat hasil evaluasi dari *hybrid* RF, kemudian dilakukan analisa terhadap hasil. Terakhir dilakukan penarikan kesimpulan. Alur proses desain penelitian dapat dilihat pada Gambar 3.1:



Gambar 3.1 Desain Penelitian

Pada Gambar 3.1 Desain Penelitian menjelaskan desain rancangan penelitian yang dimulai dari memahami *Business Process* dilanjutkan *Data Collection* dengan menjelaskan dataset dan attribut yang digunakan, *Design System* meliputi persiapan data yang didahului dengan pemahaman proses bisnis dan pemahaman data serta penjelasan Algoritma RF dan SVR serta RF *hybrid*, selanjutnya dilakukan *Experiment* dengan melakukan beberapa skenario Algortima *hybrid* yang digunakan. Proses terakhir dilakukan, *Evaluation, Analysis, Conclusion.*

3.2 *Data Collection*

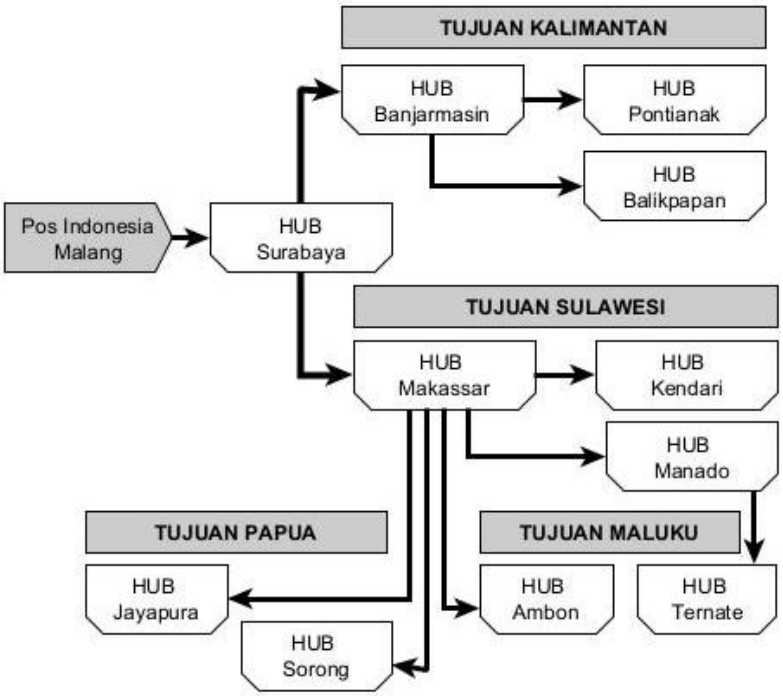
Data collection merupakan tahapan yang penting dalam penelitian ini, oleh karena itu dilakukan beberapa tahapan agar data yang akan diolah lebih tepat. Oleh karena itu dibutuhkan beberapa tahapan untuk mendapatkan kebutuhan data yang tepat diolah dimulai dari pemahaman tentang proses bisnis (*business process undesrtanding)* dari obyek yang akan diteliti pemahaman data yang akan dikumpulkan *(data understanding)* dan kesiapan

data yang akan diolah *(data preparation).* Dengan demikian *data collection* dapat didapat secara tepat untuk dilaksanakan proses selanjutnya.

3.2.1 Business Process Understanding

Tahap ini merupakan tahap awal sebelum mempersiapkan data yang akan diolah yaitu memahami proses bisnis yang ada. Pengiriman barang dan dokumen dari PT Pos Indonesia Cabang Utama Malang untuk tujuan Kalimantan, Sulawesi, Maluku dan Papua menjadi pilihan dalam penelitian ini. Pemilihan wilayah itu didasari dari prosentase penyampain kiriman yang belum memenuhi target dibanding dengan wilayah lain.

Proses pengiriman barang atau dokumen dilakukan beberapa *hub* (kiriman dikumpulkan di kantor tertentu yang ditunjuk sebagai *hub* untuk di distribusikan ke kota tujuan lainnya) kecuali kiriman lokal atau kiriman yang dilalui oleh alat transportasi yang tersedia. Untuk memahami alur pengiriman barang dan dokumen berikut dijelakan pada Gambar 3.2



Gambar 3.2 Proses Bisnis Pengiriman Barang/Dokumen
(Sumber: Bagian Operasi PT Pos Indonesia, 2022)

Pada gambar 3.2 dijelaskan bahwa terhadap kota tujuan melalui beberapa hub yang ditentukan. Proses pengiriman tidak berdasar *point to point* melain melalui mekanisme pengumpulan kiriman dalam satu *hub* yang telah ditetapkan untuk diteruskan ke beberapa *hub* hingga sampai tujuan.

3.2.2 Data Understanding

Setelah mengetahui proses bisnis berikutnya adalah memahami data yang terkait dengan proses bisnisnya. Foster Provost & Tom Fawcett, 2013. Proses pengiriman melalui *hub* dalam gambar 3.2 berpengaruh terhadap perhitungan SLA dan jarak yang sudah ditetapkan. Disamping itu jumlah *hub* yang dilalui juga berpengaruh untuk waktu tempuh pada tiap tujuan.

Terdapat data yang dimungkinkan berpengaruh terhadap implementasi model yang diterapkan nanti disamping jarak dan jumlah hub yaitu berat kiriman, apakah berat kiriman juga berpengaruh signifikan terhadap perhitungan prediksi yang dilakukan.

3.2.3 Data Preparation

Dalam tahapan ini diperlukan pemahaman variabel mana yang dinggan perlu dan penting serta variabel atau atribut mana yang tidak penting atau tidak berpengaruh dengan memahami data terlebih dahulu.

Penelitian ini menggunakan Data Primer. Data Primer didapat dari dashboard monitoring kiriman PT Pos Indonesia pada halaman web resmi bagian operasi (*https://board.mile.app*) untuk melakukan pengambilan data operasi kiriman barang atau dokumen sebagai variabel internal, dan data dari perhitungan waktu transportasi antar hub sebagai variabel dan jarak sebagai data external.

Data awalnya berbentuk CSV *(Comma Separated Values)* kemudian diolah dalam bentuk *Spreedsheet* untuk memudahkan proses berikutnya. Sebelum melakukan prediksi dengan algoritma pembelajaran mesin, data mentah akan diolah terlebih dahulu supaya siap untuk digunakan. Ada dua tahap, yaitu memuat dataset dan melakukan praproses data supaya matang. Praproses data disini akan melihat apakah data memiliki *missing value* atau tidak. Jika terdapat *missing value* maka baris data tersebut akan dihapus. Data yang tidak memiliki korelasi juga akan di keluarkan dari dataset. Selain itu

normalisasi diterapkan supaya data memiliki rentang nilai yang seragam. Data juga akan dibagi berdasarkan presentasi tertentu untuk digunakan sebagai data latih dan data uji. Data variabel internal dan external digabungkan menjadi satu untuk memudahkan proses pengolahan data. Berikut dataset yang akan digunakan pada penelitian ini ditunjukkan pada Tabel 3.1 :

Tabel 3.1 Detil Dataset

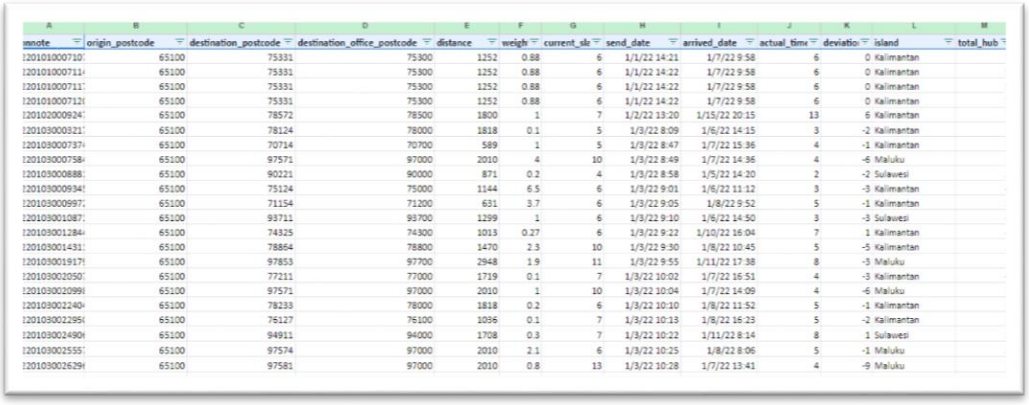
No	Atribut	Tipe Data	Keterangan/Identifikasi
1	Connote	Text	Bukti pengiriman
2	origin_postcode	Text	Kodepos Kantor Kirim
3	destination_postcode	Text	Kodepos Tujuan
4	destination_office_postcode	Text	Kodepos Kantor Tujuan
5	distance	Numeric	Perhitungan jarak kantor kirim ke kantor tujuan berdasar hub (Variabel External)
6	weight	Numeric	Berat Kiriman
7	current_sla	Numeric	SLA yang ditetapkan sebelumnya
8	send_date	Date	Tanggal pengiriman
9	arrived_date	Date	Tanggal kiriman sudah diterima
10	actual_time	Numeric	Realisasi waktu sebenarnya
11	deviation	Numeric	Selisih waktu antara SLA Day dan Actual time
12	island	Text	Nama Pulau
13	total_hub	Numeric	Perhitungan jumlah transportasi yang dilalui melalui hub (Variabel External)

Dari tabel 3.1 terdapat 13 attribut yang dibagi beberapa kelompok variabel dependen dan indenpenden, Sugiyono (2018) pada tabel 3.2 :

Tabel 3.2 Variabel Dependen dan Indepenen

No	Atribut	Variabel
1	Connote	independen
2	origin_postcode	independen
3	destination_postcode	independen
4	destination_office_postcode	independen
5	distance	independen
6	weight	independen
7	current_sla	independen
8	send_date	independen
9	arrived_date	independen
10	actual_time	dependen
11	deviation	dependen
12	island	independen
13	total_hub	independen

Adapun jumlah data yang diolah sebesar 18.769 data untuk kiriman tujuan pulau Kalimantan, Sulawesi, Maluku dan Papua yang dikirim dari Malang untuk layanan Pos Kilat Khusus/Pos Reguler dari bulan Januari 2022 hingga Desember 2022 atau 1 Tahun. (sumber : data dashboard operasi Pos Indonesia, 2022). Berikut beberapa data yang di *capture* dari data Excell pada Gambar 3.3



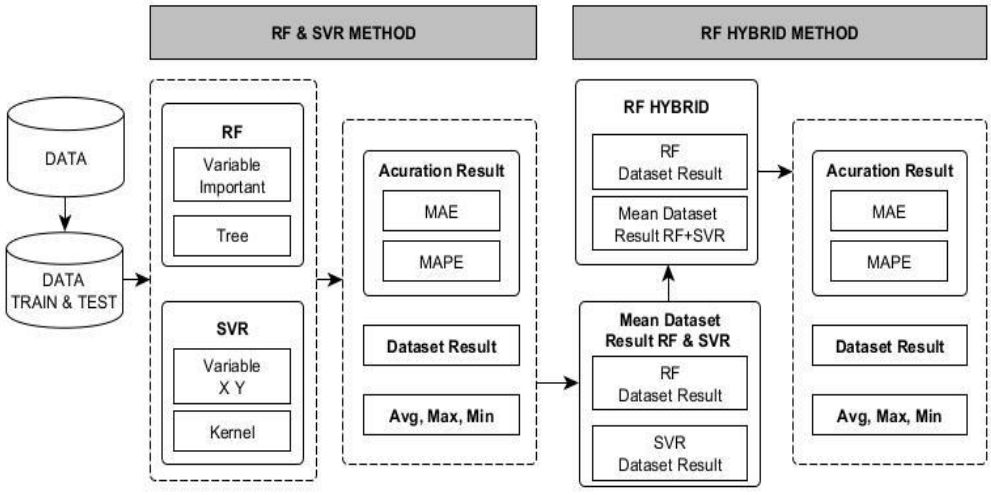
A	B	C	D	E	F	G	H	I	J	K	L	M
nnote	origin_postcode	destination_postcode	destination_office_postcode	distance	weigh	current_sla	send_date	arrived_date	actual_time	deviation	island	total_hub
201010007107	65100	75331	75300	1252	0.88	6	1/1/22 14:21	1/7/22 9:58	6	0	Kalimantan	
201010007114	65100	75331	75300	1252	0.88	6	1/1/22 14:22	1/7/22 9:58	6	0	Kalimantan	
201010007117	65100	75331	75300	1252	0.88	6	1/1/22 14:22	1/7/22 9:58	6	0	Kalimantan	
201010007126	65100	75331	75300	1252	0.88	6	1/1/22 14:22	1/7/22 9:58	6	0	Kalimantan	
201020009247	65100	78572	78500	1800	1	7	1/2/22 13:20	1/15/22 20:15	13	6	Kalimantan	
201030003217	65100	78124	78000	1818	0.1	5	1/3/22 8:09	1/6/22 14:15	3	-2	Kalimantan	
201030007374	65100	70714	70700	589	1	5	1/3/22 8:47	1/7/22 15:36	4	-1	Kalimantan	
201030007584	65100	97571	97000	2010	4	10	1/3/22 8:49	1/7/22 14:36	4	-6	Maluku	
201030008887	65100	90221	90000	871	0.2	4	1/3/22 8:58	1/5/22 14:20	2	-2	Sulawesi	
201030009347	65100	75124	75000	1144	6.5	6	1/3/22 9:01	1/6/22 11:12	3	-3	Kalimantan	
201030009977	65100	71154	71200	631	3.7	6	1/3/22 9:05	1/8/22 9:52	5	-1	Kalimantan	
201030010877	65100	93711	93700	1299	1	6	1/3/22 9:10	1/6/22 14:50	3	-3	Sulawesi	
201030012844	65100	74325	74300	1013	0.27	6	1/3/22 9:22	1/10/22 16:04	7	1	Kalimantan	
201030014317	65100	78864	78800	1470	2.3	10	1/3/22 9:30	1/8/22 10:45	5	-5	Kalimantan	
201030019177	65100	97853	97700	2948	1.9	11	1/3/22 9:55	1/11/22 17:38	8	-3	Maluku	
201030020507	65100	77211	77000	1719	0.1	7	1/3/22 10:02	1/7/22 16:51	4	-3	Kalimantan	
201030020998	65100	97571	97000	2010	1	10	1/3/22 10:04	1/7/22 14:09	4	-6	Maluku	
201030022404	65100	78233	78000	1818	0.2	6	1/3/22 10:10	1/8/22 11:52	5	-1	Kalimantan	
201030022956	65100	76127	76100	1036	0.1	7	1/3/22 10:13	1/8/22 16:23	5	-2	Kalimantan	
201030024906	65100	94911	94000	1708	0.3	7	1/3/22 10:22	1/11/22 8:14	8	1	Sulawesi	
201030025557	65100	97574	97000	2010	2.1	6	1/3/22 10:25	1/8/22 8:06	5	-1	Maluku	
201030026298	65100	97581	97000	2010	0.8	13	1/3/22 10:28	1/7/22 13:41	4	-9	Maluku	

Gambar 3.3 Capture data yang diolah

Dengan memperhatikan Table 3.1, Table 3.2 dan Gambar 3.3 memperlihatkan 13 atribut, jenis data serta contoh *capture* tampilan data, selanjutnya ditentukan attribut target dengan tipe data numerik yaitu atribut *actual_time* untuk dibandingkan dengan dengan *current_sla* untuk menghasilkan prediksi waktu tempuh dengan menggunakan Algortima RF dan SVR dan dilakukan *comparasi* secara *hybrid* hasil antara alogritma RF dan SVR dengan RF sebagai metode acuan yang digunakan untuk perhitungan *hybrid*.

3.3 *Design System*

Design system akan menjelaskan bagaimana sistem melakukuan prediksi SLA terhadap kiriman barang atau dokumen dengan menggunakan Algoritma *Random Forest* (RF) dan *Support Vector Regression* (SVR) sebagai metode awal, kemudian hasil dari kedua algortima tersebut di *mixing* untuk secara *hybrid* dengan Algoritma RF. Alur proses *design system* dalam penelitian ini dapat dilihat pada Gambar 3.4 :



Gambar 3.4 Design System

Pada gambar 3.4 tahapan awal yang dilakukan melakukan *filtering* beberapa data yang dapat mempengaruhi perhitungan prediksi waktu tempuh. Proses terbagi 3 metode, yang pertama metode RF dan SVR metode *hybrid*. Untuk metode pertama dengan melakukan prediksi RF dan SVR. Hasil dari 2 metode tersebut dilakukan perhitungan rata-rata sebagai dasar data yang akan diolah dengan *hybrid method* dengan metode acuan menggunakan RF.

3.4 *Experiment*

Pada tahap *experiment* digunakan *tool* yang bernama *Google Collab* sebagai pilihan, karena *Google Collab* berbasis *phyton, web based* tanpa melakukan instalasi pada PC/Laptop serta mudah dipresentasikan dengan dukungan aplikasi dan data terkumpul jadi satu pada *Google Drive*.

Data *spreedsheet* yang akan diolah disimpan dalam *google drive* sebagai sumber data, kemudian *google collab* mengambil sumber data dari *url* tempat data tersimpan. Data set yang digunakan dalam *experiment* ini dipilih beberapa attribut sepert pada table 3.3:

Tabel 3.3 Data Set Implementasi

No	Attribut	Type	Variable
1	distance	Numeric	Independen
2	weight	Numeric	Independen
3	actual time	Numeric	Dependen
4	deviation	Numeric	Dependen
5	Tot Hub	Numeric	Independen
6	current_sla	Numeric	Independen

Pada implementasi atribut acuan yang digunakan adalah *actual time* yang digunakan sebagai label dan menentukan jumlah data training dan

testing berbanding 75:25 sebagai dasar sebelum dilakukan optimasi. Dalam implementasi penelitian ini dilakukan beberapa tahapan atau *step* untuk mempermudah implementasi prediksi. Tahapan ini merupakan tahapan inti dari proses implementasi keseluruhan, karena tahapan secara detail akan di jelaskan pada pembahasan dalam bab tersendiri.

3.5 Pengukuran Akurasi Prediksi

Dalam penelitian ini untuk mengukur akurasi prediksi menggunakan MAE dan MAPE sebagai hasil akhir dari pengukuran metode prediksi.

3.5.1 MAE (Mean Absolute Error)

Untuk mengukur nilai keakuratan metode yang digunakan dalam penelitian ini menggunakan MAE agar diketahui rata-rata selisih mutlak nilai sebenarnya (aktual) dengan nilai prediksi (peramalan). Semakin kecil nilai MAE, Chai. T., Draxler (2014), Semakin baik model tersebut dalam melakukan peramalan, dituangkan dalam rumus MAE sebagai berikut :

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} | A_i - F_i | \qquad (3.1)$$

di mana :
n = ukuran sampel
Ai = nilai data aktual ke-*i*
Fi = nilai data peramalan ke-*i*

3.5.2 MAPE (Mean Absolute Perecentage Error)

Sedangkan untuk menghitung prosentase keakuratan metode prediksi digunakan MAPE yang mempunyai rumus hampir sama dengan MAE dengan

menambahkan pembagian dan prosentase. Semakin kecil prosentase nilai MAPE semakin baik hasil prediksi, Krisma, A., & Azhari (2019). Berikut rumus MAPE tanpa diberikan penjelasan karena variabel sudah dijelaskan pada MAE:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{A_i - F_i}{A_i} \right| \text{ x } 100\% \qquad (3.2)$$

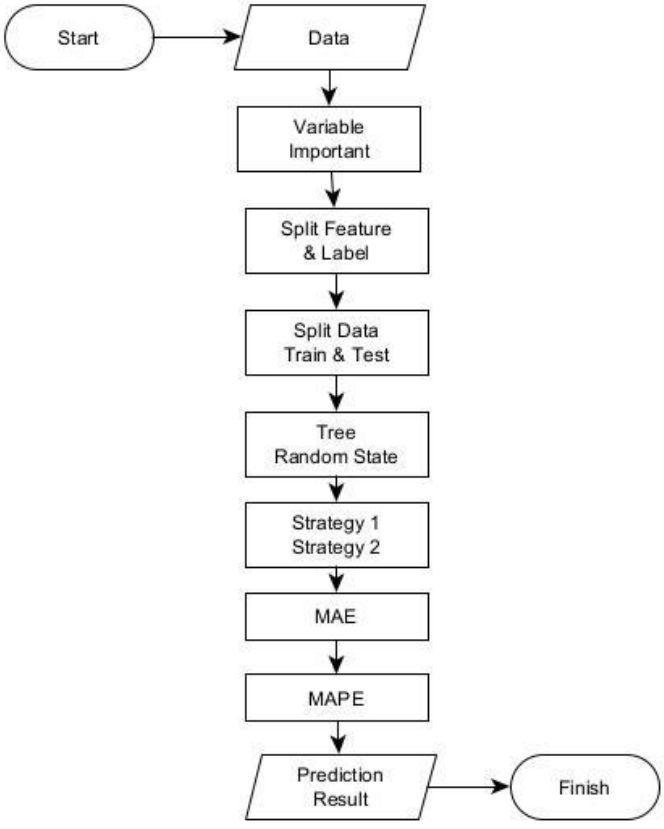
Dalam penelitian ini nilai prosentase yang dihasilkan akan ditampilkan 100 dikurangi hasil MAPE.

BAB IV

METODE RANDOM FOREST

4.1 Desain

Pada algoritma RF pertama kali yang dilakukan adalah penentuan variabel yang mempengaruhi perhitungan prediksi yang ada pada dataset. Setelah itu data diolah di *optimized* untuk menghasilkan prediksi yang terbaik. Desain diagram alur prediksi dengan metode RF tergambar pada *Flowchart* proses prediksi dalam penelitian ini pada Gambar 4.1 berikut:



Gambar 4.1 Flowchart proses prediksi FR
(Sumber : diolah dari Breiman, L. 2001)

Pada Gambar 4.1 terdapat pemilihan *VI (Variable Important)* dimana penentuan ini akan memfokuskan variabel yang paling berpengaruh melalui perhitungan algortima RF. Menurut Breiman terdapat beberapa tree dan membentuk hutan atau forest, dan pengembangan dari metode klasifikasi dan regresi tree. (Breiman 2001)

Tahapan berikutnya dengan melakukan penarikan terhadap sample secara acak sebanyak 2/3 dari data, dari data tersebut kemudian dijadikan tree. Kemudian dari beberapa tree tersebut membangun tree dari masing-masing sample dan mebentuk sebuah forest. Dari data pada random forest tersebut dilakukan perhitingan pada setiap variabel untuk mendapai nilai variabel important (VI) atau variabel yang paling berpengaruh. VI ini dipengaruhi dari nilai error yang dihasilkan OOB (out of bag) yang dihasilkan dari data testing yang diuji. Penjelasan rumus Error OOB sebagai berikut :

$$errOOB = \frac{1}{n-z}\sum_{i=1}^{n-z}(y_i - \hat{y}_i)^2 \quad (4.1)$$

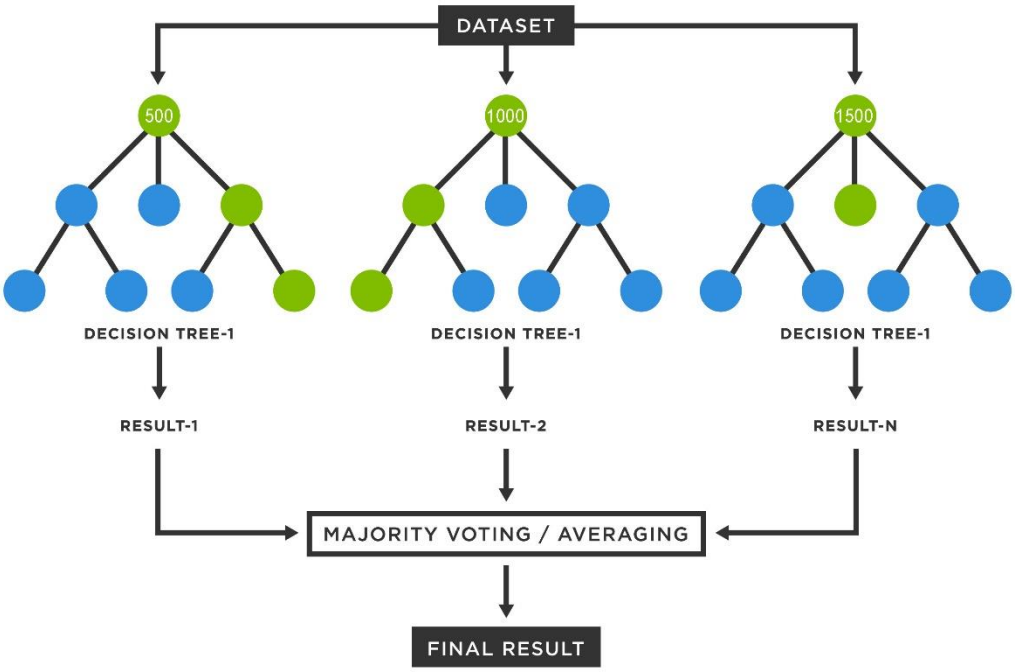
di mana,
n = data observasi
z = *sample* (data testing) untuk membentuk tree
y_i = data kei-*i* (data testing)
$\hat{y}_i$ = prediksi ke-*i.*

Untuk persamaan *variable important* (VI) diperoleh dari rumus berikut:

$$VI\left(X^j\right) = \frac{1}{s}\sum_{t=1}^{s}(errOOB_t^j - errOOB_i) \quad (4.2)$$

di mana,
$VI(X^j)$ = variable important variabel ke-j
s = jumlah tree yang dibangun
$errOOB_t^j$ = error pada tree ke-t dengan variabel j dan t
$errOOB_t$ = error pada tree ke-t variabel ke-t saja

Dalam penerapannya RF menghasilkan variabel mana yang paling dominan yang kemudian dapat dijadikan sebagai perhitungan prediksi selanjutnya. Adapun diagram metode RF dapat dilihat pada Gambar 4.2 berikut :



Gambar 4.2 Random Forest Diagram

4.2 Ujicoba

Pada tahap ujicoba ini akan diterapkan 2 strategi pada metode RF, yang pertama dengan menggunakan variabel yang berpengaruh, terdapat 5 variabel yang berpengaruh yaitu *distance, weight, actual_time, deviation* dan *total_hub*. Srategi yang kedua dengan memilih VI *(variable important)* dari kedua srategi ini dipilih strategi dengan nilai prediksi yang terbaik.

Untuk melakukan ujicoba ini menggunakan *google collab* untuk kedua strategi didahului dengan melakuan *import dataset* yang diambil dari data penelitian yang disimpan pada *google drive* seperti pada *coding* berikut :

```
#loading data
url =
'https://docs.google.com/spreadsheets/d/1hOjX5Q0IfiqhGzK_jtWTkUBaOnYO69nj/edit#gid
=868380143'
sheet_url_trf = url.replace('/edit#gid=', '/export?format=csv&gid=')
data = pd.read_csv(sheet_url_trf)
#dataset untuk random forest
dataRF = data.copy()
datasetRF1 = dataRF.filter(['distance','weight','actual_time','deviation','total_hub',
'current_sla','destination_postcode'], axis = 1)
datasetRF2 = datasetRF1.drop(['current_sla','destination_postcode'], axis = 1)
datasetRF2 = datasetRF2[datasetRF2.actual_time!=0]
```

Coding diatas terdapat variabel *destination_postcode* yang akan digunakan sebagai data hasil dari prediksi berdasar kodepos tujuan dengan dibandingkan dengan *actual_time* dan *prediction result*. Data traning dan data testing berbanding 75:25, dengan data training yang besar diharapkan mendapatkan akurasi prediksi yang lebih tepat.

4.2.1 Strategi Pertama

Strategi pertama menggunakan 5 variabel yang berpengaruh yaitu *distance, weight, actual_time, deviation* dan *total_hub*. Dalam setiap skenario prediksi, diperlukan beberapa percobaan dengan melakukan perubahan pada pohon *(tree) estimator* pada Random Forest untuk mengukur tingkat akurasi. Metode ini dijalankan dengan mengubah nilai nominator pada *tree* 500, 1000 dan 1.500, pada percobaan *tree estimator* 500 mendapatkan hasil seperti pada tabel 4.1

Tabel 4. 1 Hasil Prediksi RF 500 strategi pertama

index	Destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	90243	3	3.460385	0.460385
1	91251	3	3.159933	0.159933
2	74325	6	6.771483	0.771483
3	91511	16	17.3	1.3
4	75765	8	8.056	0.056
...	...	...	...	...
4687	79588	14	10.281	-3.719
4688	90222	2	2.583763	0.583763
4689	90222	6	6.744419	0.744419
4690	95114	3	4.493167	1.493167
4691	97511	6	5.334177	-0.665823

Pada ujicoba pertama dengan *tree* nominator 500 didapat hasil prediksi yang mendekati actual time, terdapat juga nilai dibawah 0 yang menggambarkan tingkat prediksi lebih cepat daripada actual time. Deviasi antara hasil prediksi dan *actual time* menunjukkan rentang perbedaan nilai prediksi dan *actual time*, rentang nilai tersebut akan dijadikan indikator seberapa ketepatan yang didapat. Dari 4.692 data akan diujicoba lagi hingga menemukan hasil yang optimal, setelah melakukan percobaan dengan *tree* estimator 500 selanjutnya dilanjutnya *tree estimator* 1000 dan 1500 dengan hasil pada table 4.2 dan 4.3.

Tabel 4.2 Hasil Prediksi RF 1000 strategi pertama

index	Destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	90243	3	3.452565	0.452565
1	91251	3	3.161611	0.161611
2	74325	6	6.760113	0.760113
3	91511	16	17.284	1.284
4	75765	8	8.073	0.073
...	...	...	...	...
4687	79588	14	10.2477	-3.7523
4688	90222	2	2.584018	0.584018
4689	90222	6	6.771456	0.771456
4690	95114	3	4.46925	1.46925
4691	97511	6	5.333133	-0.666867

Tabel 4.3 Hasil Prediksi RF 1500 strategi pertama

index	Destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	90243	3	3.461449	0.461449
1	91251	3	3.161704	0.161704
2	74325	6	6.757541	0.757541
3	91511	16	17.292	1.292
4	75765	8	8.068667	0.068667
...	...	...	...	...
4687	79588	14	10.240689	-3.759311
4688	90222	2	2.584367	0.584367
4689	90222	6	6.771305	0.771305
4690	95114	3	4.4535	1.4535
4691	97511	6	5.33564	-0.66436

Pada daftar hasil diatas menunjukkan terjadinya beberapa perbedaan nilai dari *tree estimator* 500, 1000 dan 1500 walaupun rentang perbedaan tidak signifikan, namun hal ini sangat berpengaruh terhadap ketepan prediksi yang didapat.

Langkah selanjutnya yaitu mengukur akurasi prediksi, dalam hal ini penulis menggunakan MAE dan MAPE dalam bentuk prosentase. Bagian

ujicoba yang mana yang mendapatkan nilai MAE paling kecil dan jumlah prosentase yang besar pada MAPE. Pada percobaan strategi pertama ini menghasilkan performa metode prediksi RF pada tabel 4.4

Tabel 4.4 Performa metode RF strategi pertama

Tree Estimator	MAE	MAPE
500	0.7180	16.2988
1000	0.7181	16.2966
1500	0.7179	16.2925

Dari hasil pengukuran akurasi, maka *tree estimator* 1500 mempunyai nilai terbaik dengan MAE 0,7179, MAPE 16,2925 nilai ini tidak jauh beda dengan pada *tree estimator* 500 dan 1000. Karena *tree estimator* 1500 mendapatkan nilai terbaik, penulis menjelaskan *coding* dan visualisasinya. Untuk melakukan perhitungan prediksi RF pada *tree estimator* 1500 seperti *coding* berikut:

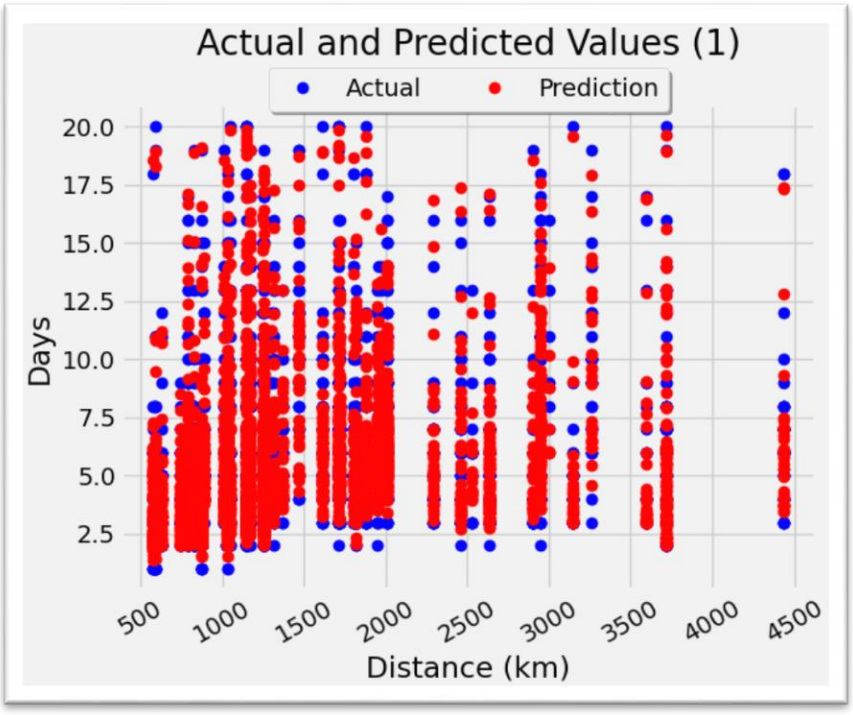
```
# menggunakan random forest regresor
from sklearn.ensemble import RandomForestRegressor
rfores = RandomForestRegressor(n_estimators = 1500, random_state = 42)
# menggunakan daa training
rfores.fit(train_features, datatraining_labels);

#MAE
predictions_rfores = rfores.predict(datatesting_features)
MAEerrors = abs(predictions_rforest - datatest_labels)
print('Mean Absolute Error:', round(np.mean(MAEerrors), 4))
#MAPE
mapeerrors = 100 * (errors / datatesting_labels)
print('MAPE:', round(mapeerrors, 4))
```

Sebelum mengelola data pada *google collab* dibutuhkan *library* dalam hali ini *library* yang digunakan adalah *random forest regressor*. Pada *coding* diatas terdapat juga *random_state*=42, dapat dijelaskan bahwa *Scikit-Learn*

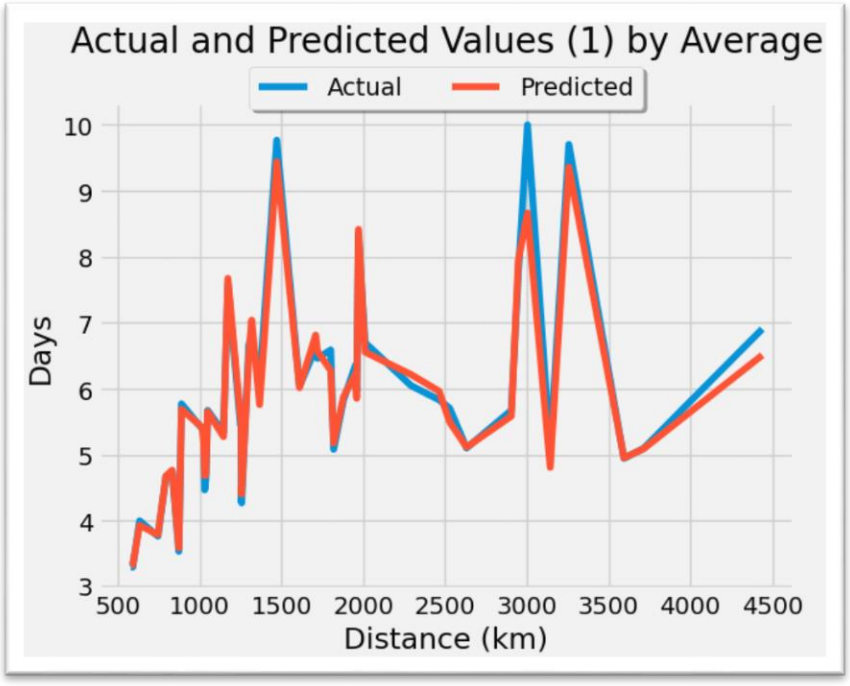
mempunyai fungsi pembagian dataset menjadi beberapa subset dengan cara yang berbeda. Parameter *random_state* yang memungkinkan mengatur *seed* dari generator acak. Jika tidak menetapkan *random_state*=42, setiap kali menjalankan kode akan menghasilkan kumpulan pengujian yang berbeda. Kharwal (2020). Perhitungan MAE dihitung berdasar dari nilai absolut prediksi dikurangi data testing, kemudian nilai MAPE ada prosentase dari nilai absolut dari MAE dibagi dengan data testing dikali 100.

Visulisasi dari hasil yang terbaik pada *tree estimator* 1500 dapat dilihat pada Gambar 4.3 dan Gambar 4.4



Gambar 4.3 RF Plot Actual & Prediction strategi pertama

Pada Gambar 4.3 memperlihatkan untuk jarak dibawah 2000 km mempunyai kerapatan yang cukup tinggi sedangkan untuk jarak 3000 km keatas tingkat kerapatannya lebih rendah.



Gambar 4.4 RF Average Actual & Prediction strategi pertama

Pada jarak 3000 km terdapat selisih yang jauh antaran aktual dan prediksi sedangkan jarak diatas 4000 km kecenderungan selisih mulai melebar.

Hasil dari prediksi strategi pertama dengan jumlah data testing sebanyak 4692 data dapat dilihat pada tabel 4.2 dan tabel 4.3 untuk nilai rata-rata, nilai maksimum dan minimum.

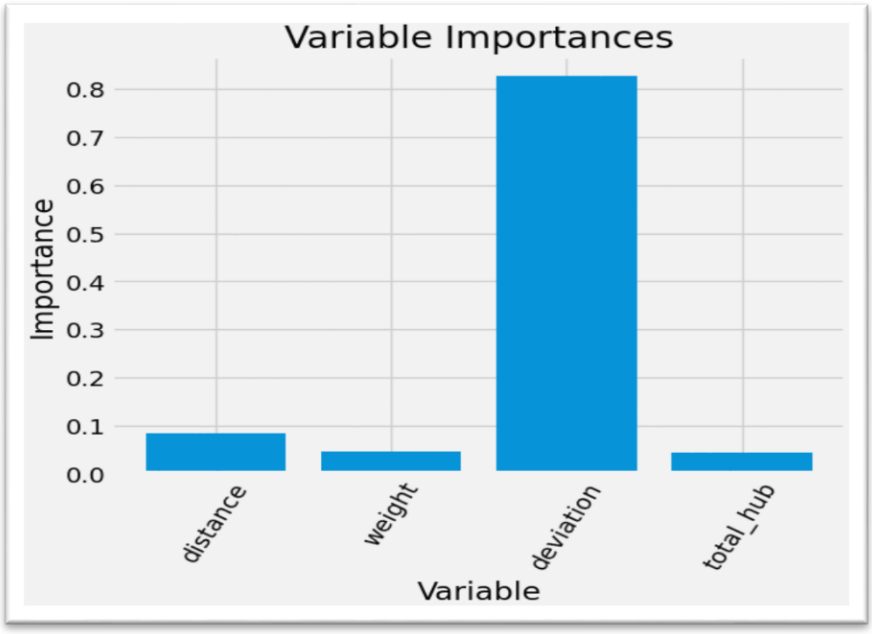
4.2.2 Strategi Kedua

Pada strategi kedua dilakukan perhitungan untuk menentukan variabel yang terpenting atau VI (*Very Important Variable)*. Dengan menentukan VI nilai prediksi menjadi lebih baik, seperti terlihat pada *coding* berikut:

```
#Mencari VI
Important_vr = list(rforest.feature_important)
fitur_important = [(feature, round(important_vr, 2)) for feature, important in zip(feature_list,
importan_vr)]
fitur_important = sorted(fitur_importanant, key = lambda x: x[1], reverse = True)
 [print('Var: {:20} Important_vr: {}'.format(*pair)) for pair in fitur_important];

Var: deviation       0.83
Var: distance        0.08
Var: weight          0.05
Var: total_hub       0.04
```

Dari c*oding* diatas dapat dilihat variabel *deviation* mempuyai nilai paling bepengaruh yaitu 0,83 diikuti variabel *distance* 0,08 variabel *weight* 0,05 dan variabel *total_hub* 0,04. Dari keempat variabel ini dipilih 2 variabel dengan nilai paling tinggi yaitu variabel *deviation* dan *distance*.



Gambar 4.5 Variable Importances

Sama halnya dengan srategi pertama untuk menghasilkan prediksi yang terbaik dilakukan beberapa *tree estimator* mulai dari 500, 1000 dan 1500 untuk srategi kedua dengan hasil pada tabel-tabel dibawah ini:

Tabel 4.5 Hasil Prediksi RF 500 strategi kedua

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	90243	3	2.649868	-0.350132
1	91251	3	2.999536	-0.000464
2	74325	6	6.786012	0.786012
3	91511	16	17.528442	1.528442
4	75765	8	8.438662	0.438662
...	...	...	...	...
4687	79588	14	11.32326	-2.67674
4688	90222	2	2.649868	0.649868
4689	90222	6	6.644822	0.644822
4690	95114	3	3.537141	0.537141
4691	97511	6	5.633046	-0.366954

Tabel 4.6 Hasil Prediksi RF 1000 strategi kedua

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	90243	3	2.65089	-0.34911
1	91251	3	2.999672	-0.000328
2	74325	6	6.792988	0.792988
3	91511	16	17.51956	1.51956
4	75765	8	8.430693	0.430693
...	...	...	...	...
4687	79588	14	11.335215	-2.664785
4688	90222	2	2.65089	0.65089
4689	90222	6	6.647617	0.647617
4690	95114	3	3.535018	0.535018
4691	97511	6	5.630614	-0.369386

Tabel 4.7 Hasil Prediksi RF 1500 strategi kedua

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	90243	3	2.650002	-0.349998
1	91251	3	2.999411	-0.000589
2	74325	6	6.788645	0.788645
3	91511	16	17.516713	1.516713
4	75765	8	8.431835	0.431835
...	...	...	...	...
4687	79588	14	11.342687	-2.657313
4688	90222	2	2.650002	0.650002
4689	90222	6	6.65478	0.65478
4690	95114	3	3.535874	0.535874
4691	97511	6	5.632832	-0.367168

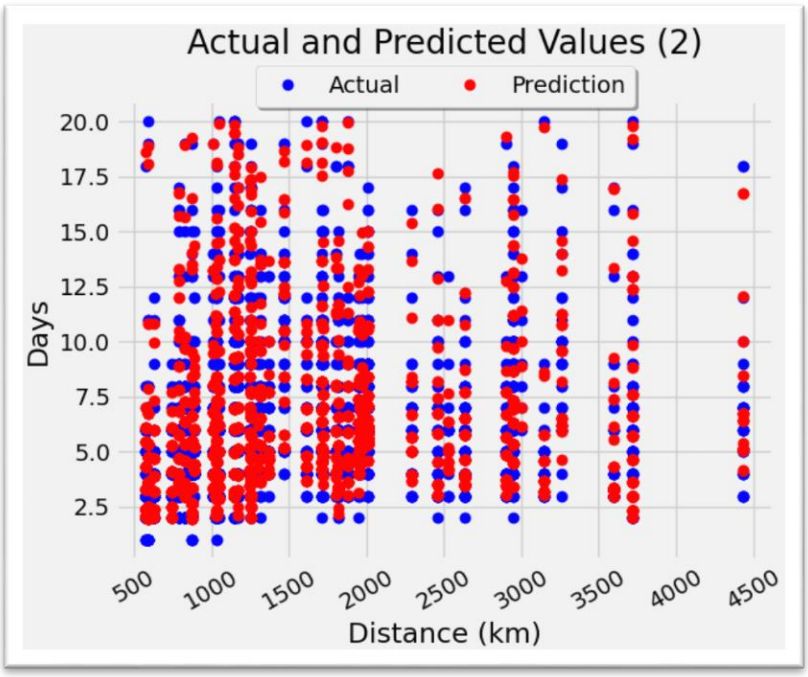
Pada strategi kedua ini dilhat dari tabel hasil prediksi *tree estimator* 500, 1000, 1500 menunjukkan nilai performa prediksi yang lebih baik dari strategi pertama, untuk membuktikannya diperlukan pengukuran nilai performa prediksi suatu model dengan menggunakan MAE dan MAPE seperti pada tabel 4.8.

Tabel 4.8 Performa metode RF strategi kedua

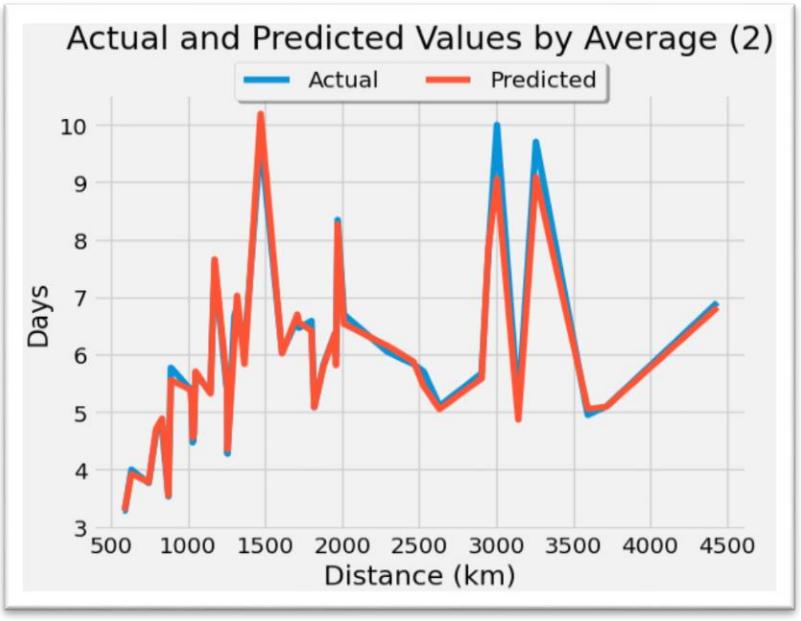
Tree Estimator	MAE	MAPE
500	0.6968	15.7982
1000	0.6968	15.7976
1500	0.6969	15.7992

Pada tabel 4.8 mempunyai performa hampir sama dari ketiga percobaan yang dilakukan. Namun *tree estimator* 1000 mempunyai nilai MAE 0.6968 dan MAPE 15.7976 lebih baik dari yang lainnya, sedangkan rata-rata, maksimum dan minimum nilai hampir sama dengan *tree estimator* 1500.

Visualisasi dari hasil yang terbaik untuk strategi kedua pada *tree estimator* 1000 dapat dilihat pada Gambar 4.6 dan Gambar 4.7



Gambar 4.6 RF Plot Actual & Prediction strategi kedua



Gambar 4.7 RF Average Actual & Prediction strategi kedua

Pada srategi kedua tingkat kerapatan plot semakin merata dari setiap jarak, hanya pada jarak 4500 km keatas terdapat tingkat kerapatan yang renggang. Untuk nilai rata-rata prediksi jarak 3000 km dan 3500 km terdapat deviasi yang tinggi dibandingkan yang lain, namun untuk jarak 4500 km keatas deviasi semakin berkurang.

Pada strategi kedua nilai MAE, prosentasi MAPE, menujukkan hasil yang lebih baik dari strategi pertama.

4.3 Komparasi performa ujicoba strategi 1 dan 2

Dari srategi pertama dan kedua dapat dilihat bahwa strategi kedua memiliki nilai yang lebih baik dengan melakuan penentuan *Variabel Important* seperti yang terlihat pada table 4.8. Dari kedua strategi dicari nilai terbaik untuk masing-masing strategi, strategi pertama hasil yang paling optimal pada

tree estimator 1500 sedangkan strategi kedua pada *tree estimator* 1000, kemudian dari kedua hasil yang optimal tersebut di komparasi sepeti terilihat pada tabel 4.9

Tabel 4.9 Komparasi Performa Srategi Pertama dan Kedua

Strategi	MAE	MAPE
Strategi Pertama	0.7179	16.2925
Strategi Kedua	0.6968	15.7976

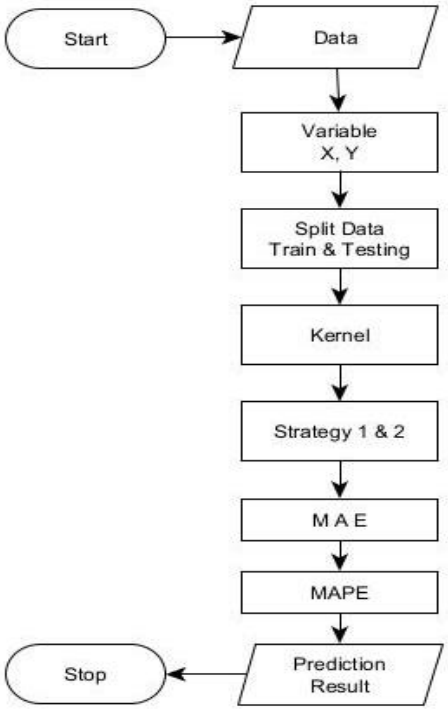
Pada algoritma RF strategi kedua dapat dijadikan strategi yang optimal dengan mendapatkan nilai yang tinggi, strategi kedua ini yang akan dijadikan acuan untuk perhitungan berikutnya secara *hybrid* dengan membandingkan dengan metode algoritma lainnya, dalam hal ini penulis mebandingkan dengan algoritma SVR.

BAB V

METODE SUPPORT VECTOR REGRESSION

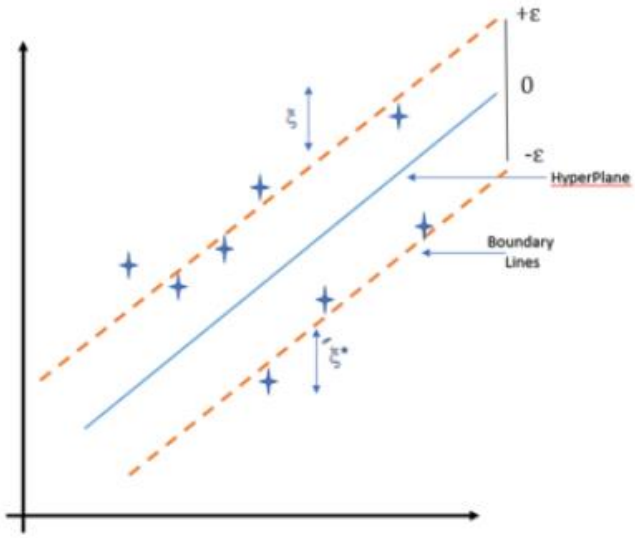
5.1 Desain

Dalam penelitian ini Support Vector Regression (SVR) digunakan untuk memprediksi nilai variabel *kontinue*, yang dapat memberikan nilai akurasi prediksi optimal nilai aktuak dan prediksi yan dapat diminimalkan tingkat kesalahannya dalam prediksi. Desain metode SVR pada penelitian ini tergambar pada *Flowchart* proses prediksi seperti pada Gambar 5.1 berikut :



Gambar 5.1 Flowchart proses prediksi SVR
(sumber : diolah dari Chun-Hsin Wu, 2004)

Pada gambar 5.1 terdapat beberapa pengujian dengan beberapa kernel yang sudah ada sehingga didapat hasil prediksi yang lebih baik. Sedangkan gambaran umum algoritma SVR adalah melakukan percobaan untuk mencari nilai terbaik dalam bentuk garis antara garis berupa *hyperplane* dan *boundary line* dapat dilihat pada Gambar 5.2



Gambar 5.2 Cara Kerja SVR (Sethi 2020)

Dari gambar 5.2 terdapat garis pemisah antar dua garis kelas data yang atas dan yang bawah dinamakan *Hyperplane*. Garis ini akan membantu nilai prediksi pada SVR, sedangkan pada hyperplane rumus persamaannya sebagai berikut:

$$y = W\,x + b \tag{5.1}$$

di mana:
y = nilai prdiksi
W = bobot atau weight
x = penjelas variabel
b = bias

Dalam SVR terdapat model liner dan non linear yang mampunyai persamaan yang berbeda pada model linear rumusnya sebagai berikut:

$$y = \sum_{i=1}^{N}(a_i - a_i^*) \cdot \langle x_i , x \rangle + b \quad (5.2)$$

Sedangkan mode non linear membutuhkan suatu kernel untuk membantu data agar dapat dihitung secara linear pada SVR dengan rumus sebagai berikut :

$$y = \sum_{i=1}^{N}(a_i - a_i^*) \cdot \langle \varphi(x_i) , \varphi(x) \rangle + b \quad (5.3)$$

$$y = \sum_{i=1}^{N}(a_i - a_i^*) \cdot K(x_i, x) + b \quad (5.4)$$

Terdapat beberapa seperti radial basis function, sigmoid, polynomial, gaussian, dan lain-lain yang berfungsi untuk menemukan hyperplane. Kernel yang sering digunakan adalah *Radial Basis Function* (RBF) dengan rumus :

$$k(x_i, x_j) = exp\left(-\frac{||x_i - x_j||^2}{2\sigma^2}\right) \quad (5.5)$$

5.2 Ujicoba

Untuk algortima SVR akan dilakukan 2 strategi sama halnya dengan algortima RF sebelumnya. Pada algoritma SVR ini ditentukan variabel terbaik yang sudah dihasilkan pada algoritma RF sebelumnya yaitu variabel *deviation* dan variabel *distance*.

Variabel dimaksud akan dijadikan sebagai sumbu X sedangkan sumbu Y adalah *actual time*. Strategi pertama yang digunakan adalah variabel *deviation* sedangkan strategi kedua variabel *distance*. Sedangkan kernel yang digunakan untuk ujicoba adalah kernel RBF *(Radial Basis Function* dan kernel *Polynomial*, alasan menggunakan 2 metode ini adalah data yang tidak dapat

dipisahkan secara linear.(Nova 2023). Kedua kernel dibandingkan untuk menentukan kernel mana yang mempunyai nilai terbaik dengan variabel yang paling optimal.

5.2.1 Strategi Pertama

Pada strategi pertama variabel yang ditetapkan sebagasi sumbu X adalah *deviation*. Dengan menggunakan *google collabs* dengan data pada *google drive* dilakukan prediksi dengan menggunakan SVR seperti *coding* berikut:

```
#dataset untuk SVR strategi pertama
col_x = 'deviation' #dipilih karena memiliki score tertinggi pertama
col_y = 'actual_time' #kolom yang akan diprediksi nilainya

dataSVR = data.copy()
datasetSVR = dataSVR.filter([col_x,col_y], axis=1)
datasetSVR = datasetSVR[datasetSVR.actual_time!=0]

x = datasetSVR.iloc[:, 0] #kolom selisih
y = datasetSVR.iloc[:, 1] #kolom actual time
```

Dengan sumbu X variabel *deviation* dan sumbu Y variabel *actual time* dengan perbandingan data training dan data testing 75:25, dengan banyaknya data training diharapkan tingkat prediksi didapatkan lebih baik. Untuk membandingkan nilai yang optimal dengan menggunkan RBF dan *Polynomial*. Hasil prediksi dari strategi pertama 2 kernel tersebut dapat dilihat pada tabel 5.1 untuk kernel RBF dan tabel 5.2 untuk kernel *Polynomial*.

Tabel 5.1 Hasil Prediksi SVR RBF strategi pertama

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	93117	6	7	1
1	98813	6	3	-3
2	76111	2	3	1
3	70714	2	3	1
4	71513	5	4	-1
...	...	...	...	...
4687	90243	3	5	2
4688	76114	5	5	0
4689	97727	13	11	-2
4690	90133	1	3	2
4691	92517	14	15	1

Hasil prediksi kernel RBF pada strategi pertama dengan menggunakan variabel *deviation* terdapat beberapa nilai plus dan minus dari selisih antara prediksi dan *actual* time, dimana nilai minus menunjukkan prediksi lebih cepat dari *actual time*.

Tabel 5.2 Hasil Prediksi SVR Polynomial strategi pertama

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	93117	6	4	-2
1	98813	6	4	-2
2	76111	2	4	2
3	70714	2	4	2
4	71513	5	4	-1
...	...	...	...	...
4687	90243	3	4	1
4688	76114	5	4	-1
4689	97727	13	6	-7
4690	90133	1	4	3
4691	92517	14	14	0

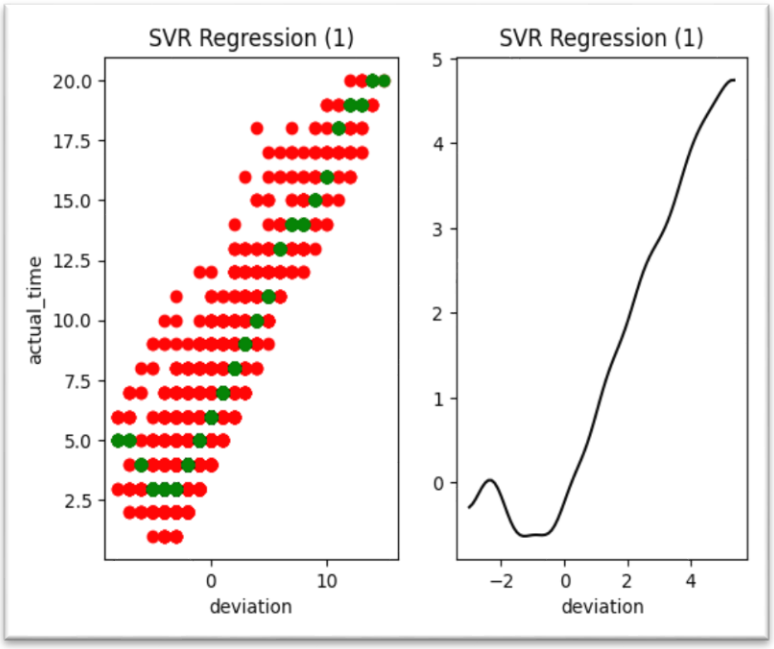
Pada ujicoba kedua dengan menggunakan kernel *Polynomial* nampak beberapa selisih antara prediksi dan *actual time* terdapat nilai minus yang artinya nilai prediksi lebih cepat dari *actual time.*

Langkah selanjutnya yaitu mengukur akurasi prediksi metode SVR, dalam hal ini penulis menggunakan MAE dan MAPE dalam bentuk prosentase. Bagian ujicoba yang mana yang mendapatkan nilai MAE paling kecil dan jumlah prosentase yang besar pada MAPE. Pada percobaan strategi pertama ini menghasilkan nilai keakuratan model prediksi SVR pada tabel 5.3.

Tabel 5.3 Performa metode SVR strategi pertama

Kernel	MAE	MAPE
RBF	1.0407	24.3927
Polynomial	1.7136	34.4070

Kernel RBF mempunyai nilai yang relatif lebih baik dibandingkan dengan kernel *Polynomial* dengan MAE 1,0407 MAPE 24,3927. Karena tingkat akurasi kernel RBF pada strategi pertama mempunyai nilai akurasi yang tinggi berikut disampaikan hasil visualisasi dari perdiksi SVR dengan kernel RBF pada Gambar 5.3



Gambar 5.3 Visualisasi SVR Strategi Pertama

Pada Gambar 5.3 dapat dijelaskan bahwa semakin tinggi nilai *actual time* nya tingkat prediksinya cenderung kurang optimal.

5.2.2 Strategi Kedua

Pada strategi kedua digunakan variabel berpengaruh urutan berikutnya yaitu variabel *distance* yang dijadikan sebagai sumbu X dan variabel *actual_time* sumbu Y. Dengan menggunakan teknik yang sama pada strategi pertama pada saat pertama kali dengan melakukkan pemanggilan data set terlebih dahulu seperti pada coding berikut :

```
#dataset untuk SVR strategi kedua
col_x = 'distance' #dipilih karena memiliki score tertinggi kedua
col_y = 'actual_time' #kolom yang akan diprediksi nilainya

dataSVR = data.copy()
datasetSVR = dataSVR.filter([col_x,col_y], axis=1)
datasetSVR = datasetSVR[datasetSVR.actual_time!=0]

x = datasetSVR.iloc[:, 0] #kolom selisih
y = datasetSVR.iloc[:, 1] #kolom actual time
```

Adapun data training dan data testing komposisinya sama dengan strategi pertama yaitu 75:25. Dengan membandingkan 2 kernel yaitu kernel RBF dan *Polymonial* didapat hasil prediksi pada strategi kedua ini pada tabel 5.4 dan tabel 5.5

Tabel 5.4 Hasil Prediksi SVR RBF strategi kedua

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	74261	14	4	-10
1	73593	5	3	-2
2	75313	5	5	0
3	93122	3	6	3
4	77212	4	5	1
...	...	...	...	...
4687	78571	6	6	0
4688	90552	3	4	1
4689	94765	11	6	-5
4690	74871	6	3	-3
4691	99368	2	4	2

Tabel 5.5 Hasil Prediksi SVR Polynomial strategi kedua

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	74261	14	4	-10
1	73593	5	4	-1
2	75313	5	4	-1
3	93122	3	4	1
4	77212	4	4	0
...	...	...	...	...
4687	78571	6	4	-2
4688	90552	3	4	1
4689	94765	11	4	-7
4690	74871	6	4	-2
4691	99368	2	5	3

Pada daftar hasil diatas menunjukkan terjadinya beberapa perbedaan nilai yang cukup signifikan hasil dari kernel RBF dan kernel *polynomial*, hasil yang cukup jauh ini dapat disimpulkan bahwa strategi kedua sama dengan strategi pertama yaitu kernel RBF mempunyai nilai yang lebih optimal. Namun demikian perlu dilakukan pengukuran keakuratan metode SVR, hasilpengukuran tersebut terlihat pada tabel 5.6

Tabel 5.6 Performa metode SVR strategi kedua

Kernel	MAE	MAPE
RBF	1.9885	36.8587
Polynomial	2.1298	38.2541

Kernel RBF tetap mempunyai nilai yang tinggi akurasinya dibandingkan dengan kernel *Polynomial*. Pada strategi kedua ini rata rata MAE mendapatkan nilai akurasi 1,9885 dan MAPE mendapatkan 36,8587. Dibandingkan dengan strategi pertama yang menggunakan variabel *deviation* sebagai variabel paling

berpengaruh maka variabel *distance* sebagai variabel berpengaruh urutan kedua mempunyai prediksi yang cukup jauh dengan strategi pertama.

Visualisai dan hasil prediksi menggunakan hasil perhitungan dari kernel RBF karena mempunyai tingkat akurasi yang lebih tinggi dibandingkan dengan kernel *polynomial* pada strategi kedua ini pada Gambar 5.4

Gambar 5.4 Visualisasi SVR Strategi Kedua

Pada Gambar 5.4 dapat dilihat perbedaan yang cukup tinggi tingkat akurasinya pada semua jarak. Dengan melihat hasil diatas Srategi kedua dengan menggunakan variabel *distance* mendapatkan nilai prediksi yang kurang baik dengan rentang antara *actual time* dan prediksi yang cukup jauh.

5.3 Komparasi performa ujicoba strategi 1 dan 2

Dari hasil strategi pertama dan kedua pada SVR dapat disimpulkan bahwa kernel RBF merupakan kernel dengan tingkat akurasi yang lebih baik

sedangkan variabel *deviation* mempunyai tingkat akurasi yang lebih baik dibanding dengan variabel *distance*. Hasil dari strategi pertama dan kedua dapat disimpulkan pada tabel 5.7.

Tabel 5.7 Komparasi Performa Srategi Pertama dan Kedua SVR

Strategi	MAE	MAPE
Strategi Pertama	1.0407	24.3927
Strategi Kedua	1.9314	36.2916

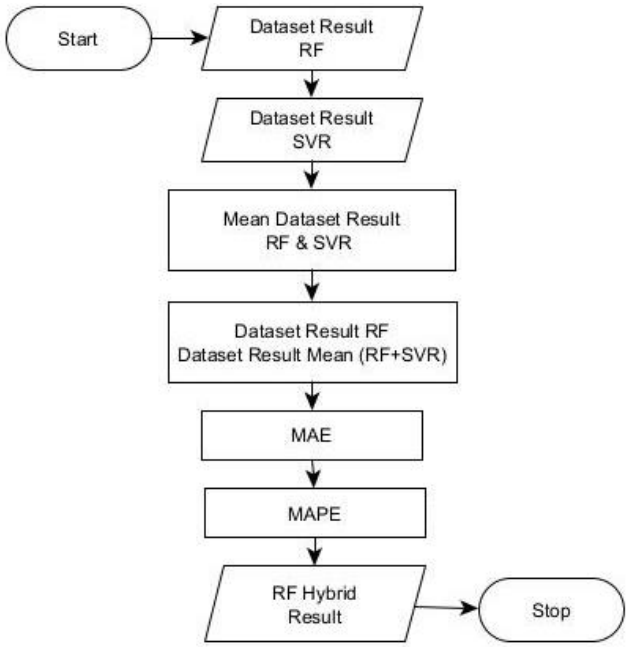
Dari hasil ini maka strategi pertama yang akan digunakan sebagai pembanding dengan metode lainnya.

BAB VI

METODE RANDOM FOREST HYBRID

6.1 Desain

Untuk melakukan analisa hasil prediksi yang lebih optimal diperlukan tambahan proses metode algoritma lagi dengan data sumber dari hasil prediksi beberapa metode sebelumnya dengan metode yang berbeda dalam hal ini metode RF dan SVR. Metode ini disebut metode *hybrid* dengan metode *Random Forest* sebagai metode acuan untuk melakukan komparasi dari hasil prediksi metode sebelumnya. RF dipilih dikarenakan mempunyai nilai akurasi yang lebih tinggi daripada SVR. Pada penelitian ini desain diagram alur metode *Random Forest Hybrid* dengan flowchart pada gambar 6.1



Gambar 6.1 Flowchart FR Hybrid (Sumber: data diolah peneliti)

Flowchart FR Hybrid ini diolah peneliti berdasar hasil prediksi dari metode FR dan SVR. Data diolah terlebih dahulu dari hasil prediksi rata rata 2 metode (FR dan SVR), setelah itu data diolah lagi dengan melakukan perhitungan rata rata metode FR dan hasil dari perhitungan rata rata prediksi metode FR dan SVR, hasil tersebut diuji lagi dengan tingkat akurasi dalam hal ini menggunakan MAE dan MAPE dilanjutkan hasil dari *RF Hybrid* berupa dataset, nilai rata-rata, maximum dan minimum.

6.2 Ujicoba

Pada ujicoba RF *Hybrid* akan dilakukan dengan membandingkan 2 pendekatan yaitu dengan perhitungan prediksi *non Hybrid* dan *Hybrid*. Perhitungan prediksi *Non Hybrid* dilakukan dengan membandingkan secara langsung hasil terbaik dari hasil prediksi RF dan SVR, hasil ini nanti akan dibandingkan pada metode *Hybrid* dengan RF, hasil ini untuk membuktikan apakah dengan RF *hybrid* hasil prediksi dan pengukuran prediksinya apakah lebih baik atau tidak.

6.2.1 Non Hybrid

Ujicoba *non hybrid* dengan melakukan perhitungan prediksi dari hasil metode terbaik RF dan SVR. Kemudian data di *compare* dicari nilai rata rata dari kedua metode tersebut, sedangkan pengukuran prediksi dengan menggunakan MAE dan MAPE serta dicari nilai rata-rata, maksimum dan minimum. Hasil dari prediksi *non hybrid* dan selisih antara prediksi dan *actual time* pada tabel 6.1

Tabel 6.1 Hasil Prediksi Non Hybrid

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	90243	3	3.325445	0.325445
1	91251	3	2.999836	-0.000164
2	74325	6	6.396494	0.396494
3	91511	16	17.75978	1.75978
4	75765	8	7.715347	-0.284653
...	...	...	...	...
4687	79588	14	11.167608	-2.832392
4688	90222	2	3.325445	1.325445
4689	90222	6	7.323808	1.323808
4690	95114	3	3.767509	0.767509
4691	97511	6	4.315307	-1.684693

Hasil prediksi *non hybrid* pada tabel 6.1 adalah komparasi dari hasil terbaik metode RF dan SVR yang menghasilkan nilai MAE dan MAPE serta nilai rata-rata, maximum dan minimum pada tabel 6.2

Tabel 6.2 Performa non hybrid

Metode	MAE	MAPE
Non Hybrid	0.8035	18.7871

Berdasarkan perhitungan pengukuran prediksi didapatkan nilai MAE 0,8035 MAPE 18,7871, hasil ini menunjukkan hasil yang lebih tinggi dari metode SVR namun lebih rendah dari RF. Dari komparasi metode RF dan SVR atau peneliti namakan *non hybrid* mendapatkan hasil prediksi yang baik ditandai dengan nilai MAE di angka 0 dan prosentase MAPE dibawah 20.

6.2.2 Random Forest Hybrid

Ujicoba *hybrid* dilakukan dengan perhitungan terlebih dahulu hasil dari metode RF dan SVR, kemudian hasil tersebut dihitung lagi dengan hasil metode RF yang terbaik. Metode ini penulis namakan metode Random Forest Hybrid

(RF *Hybrid*) karena metode RF sebagai acuan pembanding dari hasil komparasi hasil prediksi 2 metode yang berbeda (RF dan SVR). Hasil perhitungan prediksi RF *hybrid* dan selisih antara hasil prediksi *hybrid* dengan *actual time* pada tabel 6.3.

Tabel 6.3 Hasil Prediksi RF Hybrid

index	destination postcode	Actual Time	Predicted Values	Dev (Predicted-Actual)
0	90243	3	3.325445	-0.011833
1	91251	3	2.999836	-0.000246
2	74325	6	6.396494	0.594741
3	91511	16	17.75978	1.63967
4	75765	8	7.715347	0.07302
...	...	...	...	...
4687	79588	14	11.167608	-2.748589
4688	90222	2	3.325445	0.988167
4689	90222	6	7.323808	0.985713
4690	95114	3	3.767509	0.651263
4691	97511	6	4.315307	-1.02704

Setelah mendapatkan hasil prediksi kemudian dilakukan pengukuran prediksi dengan MAE dan MAPE seperti pada tabel 6.4

Tabel 6.4 Peforma RF Hybrid

Metode	MAE	MAPE
Hybrid	0.7254	16.77

Hasil dari *hybrid* didapatkan MAE 0,7254 MAPE 16,77, berdasarkan hasil tersebut metode *hybrid* yang dibandingkan dengan metode RF mendapatkan hasil yang baik dengan capaian MAPE dibaah 20 dan MAE dikisaran 0, sedangkan nilai rata rata yang relatif rendah dibawah 0 yang artinya perhitungan waktu kiriman lebih cepat dari *actual time*.

6.3 Komparasi performa ujicoba non hybrid dan RF hybrid

Metode RF *Hybrid* perlu dilakukan ujicoba dengan non *hybrid* agar perhitungan dengan metode RF *Hybrid* benar-benar bisa memberikan nilai prediksi yang lebih baik dengan metode komparasi biasa atau *non hybrid* yaitu mebandingkan secara langsung nilai 2 metode.

Dari hasil ujicoba diatas bahwa RF *Hybrid* dapat menunjukkan nilai akurasi prediksi dibandingkan dengan komparaasi biasa atau penulis sebut *non hybrid* 2 metode. Perbandingan hasil dari metode dimaksud dapat dilihat pada tabel 6.5

Tabel 6.5 Performa Non Hybrid dan RF Hybrid

Metode	MAE	MAPE
Non Hybrid	0.8035	18.7871
RF Hybrid	0.7254	16.77

BAB VII

PEMBAHASAN

7.1 Komparasi dan Performa Metode

Sebelum menjelaskan komparasi performa metode, perlu disampaikan bahwa mengapa perhitungan prediksi ini dilakukan. Hal ini dikarenakan bahwa hasil dari *test* perhitungan akurasi SLA dengan membandingkan hasil selisih dari *actual time* dan SLA yang ditetapkan nilainya dibawah 50% sebelum menggunakan metode prediksi dan metode prediksi secara *hybrid*. Data hasil *test* perhitungan akurasi dengan menggunakan MAE mendapatkan nilai 2,2586 dan MAPE 54,9464. Dengan hasil tersebut sangat mempengaruhi performa SLA yang dijanjikan ke pelanggan, oleh karena itu perlu dilakukan upaya untuk melakukan analisa prediksi yang lebih tepat dan optimal untuk menentukan SLA.

Komparasi performa metode dibagi dalam 2 kategori, yang pertama komparasi dan performa metode *Random Forest* (RF) dengan *Support Vector Regression* (SVR), kedua komparasi *Random Forest* (RF) dengan rata-rata hasil prediksi dari RF dan SVR yang disebut juga oleh penulis sebagai performa *Random Forest Hybrid*.

7.1.1 Performa dan Komparasi metode RF dan SVR

Dari hasil eksperimen strategi pertama dan strategi kedua pada metode *Random Forest* (RF) dengan menggunakan 3 *tree estimator* untuk masing-

masing strategi yaitu 500, 1000, 1500. Pada pengujian strategi pertama dengan melibatkan semua variabel atau attribut yaitu *distance, weight, actual_time, deviation* dan *total_hub* dalam penelitian ini mendapatkan hasil performa akurasi MAE 0,7179, MAPE 16,2924 dengan *tree estimator* 1500. Hasil ini mendapatkan nilai lebih rendah dari strategi kedua. Pada strategi kedua dilakukan optimalisasi variabel atau attribut yang paling penting atau *Very Important* (VI), dengan variabel *deviation* dan *actual time*, menghasilkan performa akurasi MAE 0,6968, MAPE 15,7976 dengan *tree estimator* 1000. Dari hasil kedua eksperimen tersebut bahwa untuk menentukan prediksi SLA yang lebih tepat adalah strategi yang kedua dimana nilai MAE, MAPE, nilai performa prediksi yang lebih baik pada metode RF dibandingkan dengan strategi pertama.

Hasil performa dari metode *Support Vector Regression* (SVR) dilakukan melaui experimen 2 strategi dengan masing-masing menggunakan kernel *Radial Basis Function* (RBF) dan kernel *Polynomial*. Kedua kernel ini digunakan karena data yang tidak dapat dipisahkan secara linear. Pada strategi pertama menggunakan variabel atau atribut yang paling optimal yaitu *deviation* dan *actual time* mendapatkan performa akurasi MAE 1,0407, MAPE 24,3927 dengan kernel RBF. Hasil ini mempunyai nilai performa akurasi yang lebih baik dibandingkan strategi kedua untuk metode SVR. Pada strategi kedua dilakukan eksperimen dengan menggunakan atrribut atau variabel *distance* dan *actual time*. Variabel *distance* sebagai attribut yang berpengaruh kedua setelah *deviation*. Strategi kedua ini menghasilkan

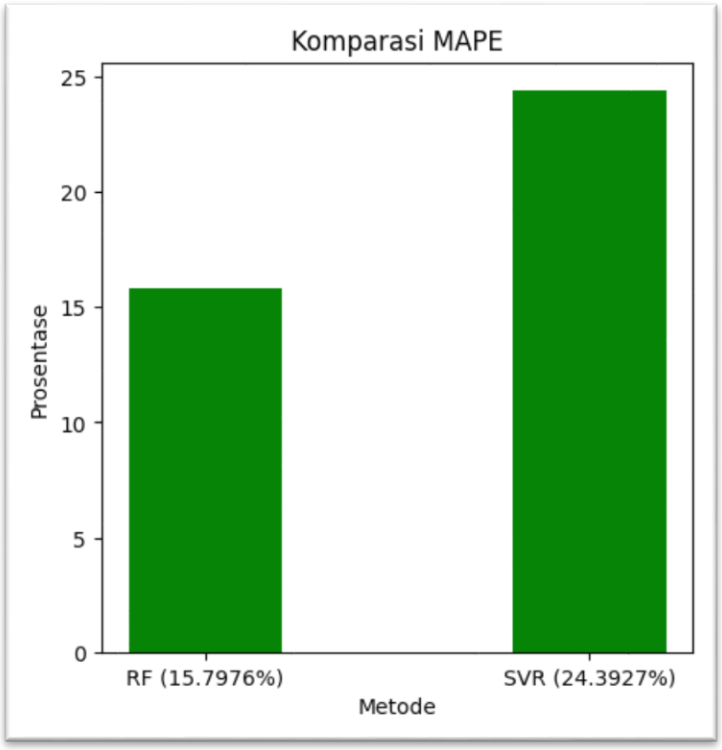
performa akurasi MAE 1,9885, MAPE 36,8587, dengan kernel RBF. Dari kedua strategi ini didapatkan performa yang lebih baik adalah strategi pertama karena tingkat akurasi MAE, MAPE lebih baik dari strategi kedua pada metode SVR.

Dari metode RF dan SVR yang digunakan kemudian dilakukan komparasi setelah ditentukan nilai performansi yang paling baik, seperti pada tabel 7.1 tentang hasil perbandingan metode RF dan SVR.

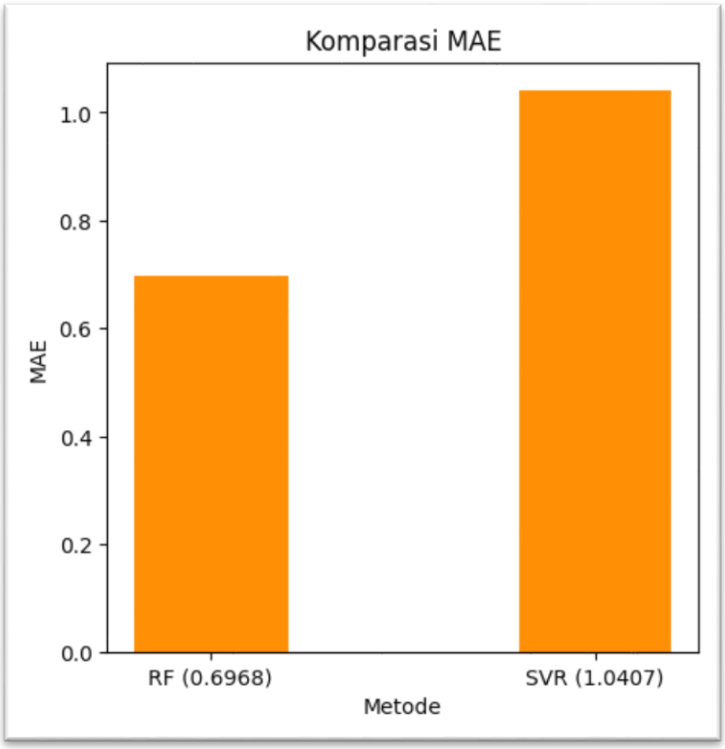
Tabel 7. 1 Komparasi performa metode RF dan SVR

Metode	MAE	MAPE
RF	0.6968	15.7976
SVR	1.0407	24.3927

Berdasarkan tabel 7.1 bahwa metode RF adalah metode dengan nilai akurasi prediksi yang lebih baik dibanding SVR untuk melakukan prediksi SLA untuk tujuan wilayah Kalimantan, Sulawesi, Maluku dan Papua dengan nilai akurasi prediksi MAE 0,6968 yang mendekati angka 0 merupakan nilai MAE yang baik. Untuk nilai MAPE 15,7976 merupakan nilai MAPE prakiraan baik dengan range diatas 10 dan dibawah 20. Berikutnya adalah nilai rata-rata 0,00158750, nilai maksimum 2,96208210 dan nilai minimum -6,9133. Visualisasi performa komparasi metode RF dan SVR terlihat pada Gambar 7.1 dan Gambar 7.2



Gambar 7.1 Komparasi Performa MAPE RF dan SVR



Gambar 7.2 Komparasi Performa MAE RF dan SVR

7.1.2 Performa dan Komparasi metode RF Hybrid

Pada experimen metode RF Hybrid terlebih dahulu prediksi dihitung dari nilai prediksi rata rata RF dan SVR dari hasil yang paling optimal pada masing-masing metode. Hasil dari prediksi rata-rata nilai RF dan SVR atau penulis beri nama non hybrid ini mendapatkan nilai MAE 0,8035 MAPE 18,7871, nilai rata-rata -0,067, nilai maksimum 2,9810 dan nilai minimum -7,4567. Hasil komparasi nilai rata-rata 2 metode ini akan dibandingkan dengan FR Hybrid sebagai experimen kedua, untuk melakukan pembuktian bahwa dengan FR Hybrid akan didapat hasil yang lebih baik.

Nilai rata-rata prediksi RF dan SVR kemudian dilakukan perbandingan dengan nilai prediksi RF sebagai metode RF Hybrid, karena metode RF dijadikan acuan sebagai metode hasil yang lebih baik dibandingkan SVR. Dari hasil nilai rata-rata tersebut kemudian dilakukan perhitungan yang menghasilkan performa akurasi MAE 0,7254 MAPE 16,77, nilai rata-rata -0,0296, nilai maksimum 2,9716 nilai minimum -7,1850. Hasil dari RF Hybrid ini menunjukkan bahwa tingkat akurasi untuk MAE tidak jauh dari 0 yang artinya hasil prediksinya baik, nilai MAPE 16,77 juga termasuk hasil yang baik karena berkisar dari interval nilai 10 sampai 20.

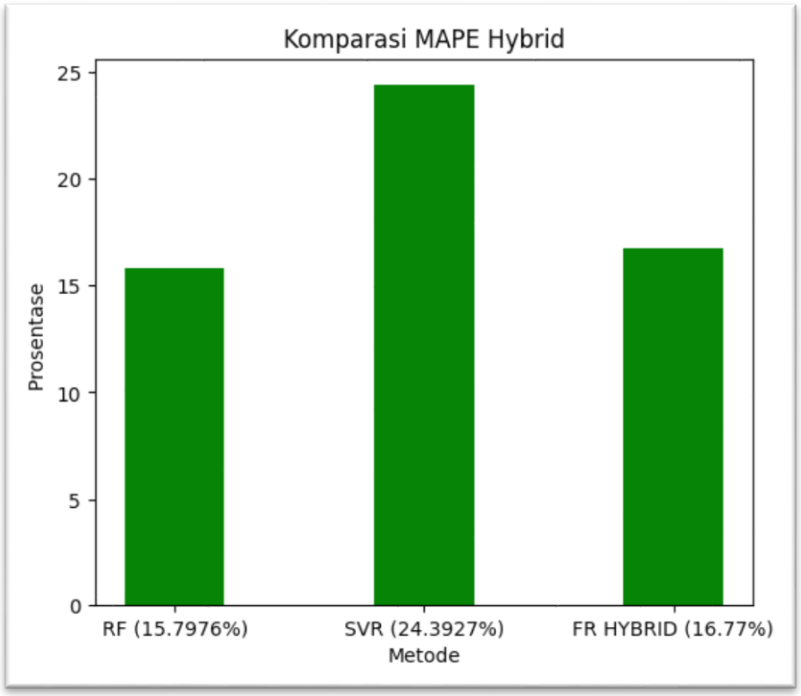
Untuk melihat komparasi hasil prediksi dari RF, SVR dan RF Hybrid dapat dilihat pada tabel 7.2

Tabel 7.2 Hasil perbandingan performa metode RF Hybrid

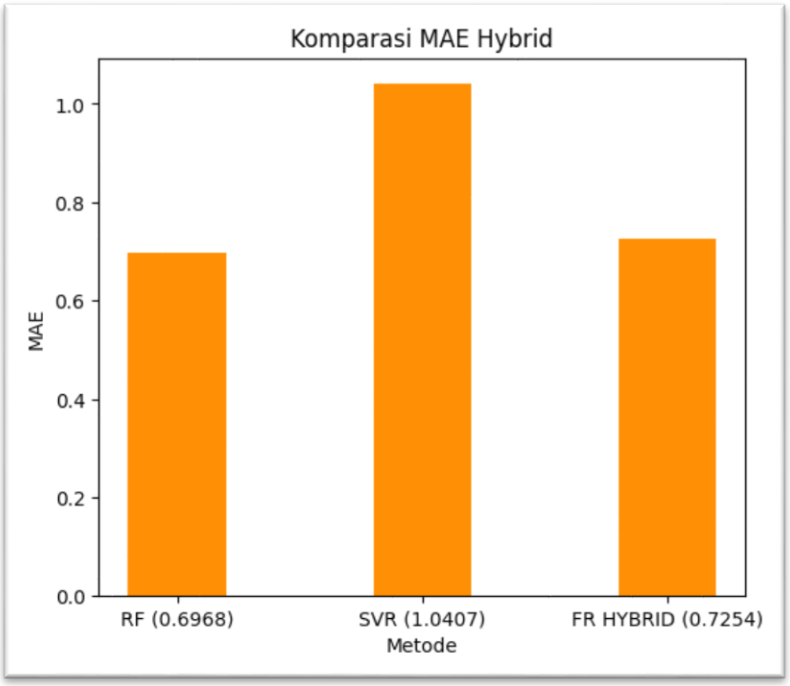
Metode	MAE	MAPE
RF	0.6968	15.7976
RF Hybrid	0.7254	16.7700

Pada tabel 7.2 hasil komparasi perhitungan prediksi bahwa RF lebih baik dari RF Hybrid dengan selisih yang tidak terlalu jauh. Nilai MAE pada RF sebesar 0,6968 sedangkan RF Hybrid 0,7254 mempunyai selisih -0,0286, untuk nilai MAPE jika dibandingkan RF 15,7976 dan RF Hybrid 16,77 mempunyai selisih 0,9724.

Beberapa pengaruh hasil RF Hybrid lebih kecil dari RF disebabkan bahwa hasil dari nilai prediksi metode SVR yang relatif rendah yaitu MAE 1,0407 dan MAPE 24,3927 dimana hasil MAE diatas 0 dan MAPE diatas 20. Hasil yang relatif kecil ini kemudian dihitung rata rata dengan RF yang mempunyai nilai prediksi lebih baik. Untuk melihat performasi komparasi dari metode RF, SVR dan RF Hybrid dapat dilihat pada Gambar 7.3 dan 7.4



Gambar 7.3 Komparasi Performa MAPE RF, SVR, RF Hybrid



Gambar 7.4 Komparasi Performa MAE RF, SVR, FR Hybrid

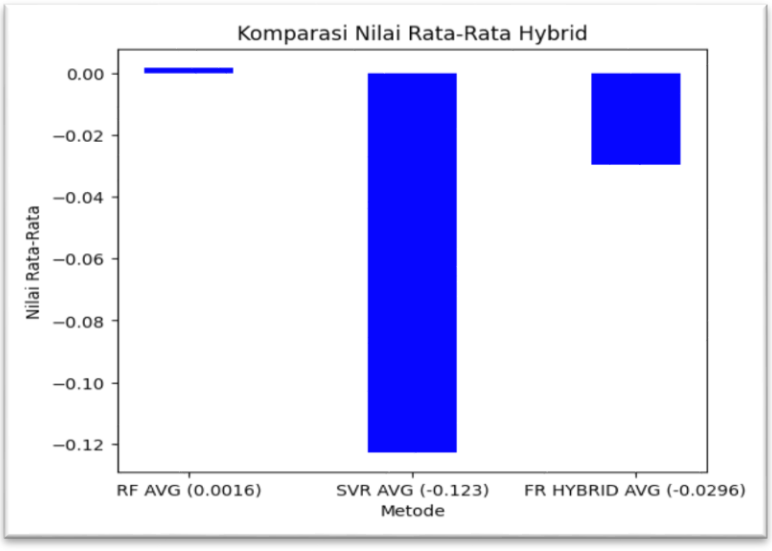
Pembuktian selanjutnya adalah nilai rata-rata, nilai maksimum dan nilai minimum. Nilai ini diambil dari dataset dari hasil prediksi dari metode

RF dan RF Hybrid dengan menghitung selisih antara hasil prediksi dan *actual time*. Selisih tersebut kemudian dicari nilai rata-rata, minimum dan maksimum yang dicapati seperti pada Tabel 7.3

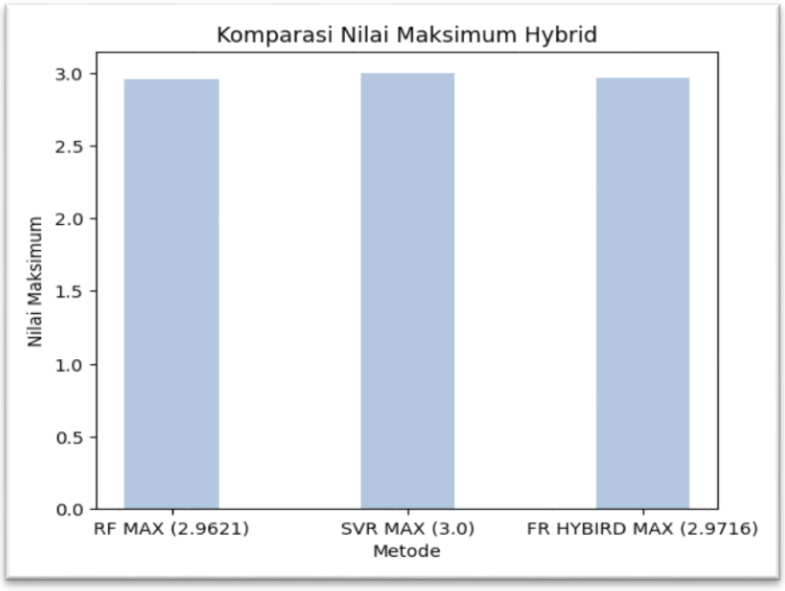
Tabel 7.3 Komparsi Performa Selisih Rata-rata, Max, Min RF dan RF Hybrid

Metode	AVG	MAX	MIN
RF	0.00158750	2.96208120	-6.9133
RF Hybrid	-0.0296	2.9716	-7.1850

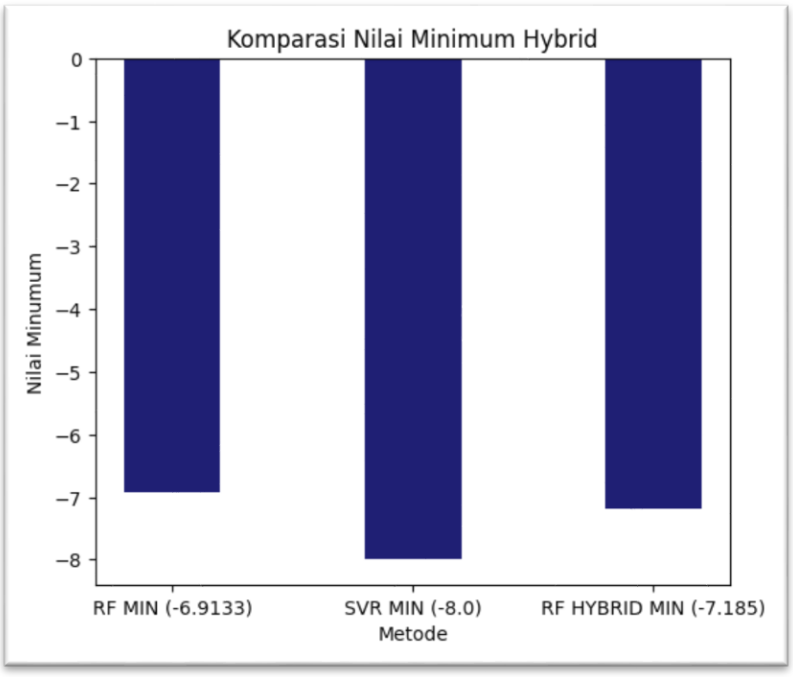
Pada tabel 7.3 metode RF mendapatkan nilai rata-rata 0,00158750 RF Hybrid -0,0296 selisih 0,0311407, nilai minus pada nilai rata rata metode RF Hybrid menggambarkan bahwa rata rata prediksi waktu nya lebih cepat dari *actual time*. Sedangkan nilai maksimum RF 2,96208120 RF Hybrid 2,9716 selisih -0,0094797 dimana nilai maksimum RF Hybrid lebih besar, untuk nilai minimum RF -6,9133 RF Hybrid -7,1850 selisih 0,27166718. Jika dibandingkan dengan SVR tentunya lebih jauh lagi, pada Gambar 7.5, Gambar 7.6 dan Gambar 7.7 dapat dilihat komparasi dari ketiga metode.



Gambar 7.5 Komparasi nilai rata-rata RF, SVR, FR Hybrid



Gambar 7.6 Komparasi nilai MAX RF, SVR, FR Hybrid



Gambar 7.7 Komparasi nilai MIN RF, SVR, FR Hybrid

Dari hasil dan visualisasi data maka hasil prediksi yang paling optimal menggunakan metode RF daripada menggunakan RF Hybrid

walaupun dengan selisih yang tidak terlalu banyak. Harapan dengan menggunakan RF secara hybrid dapat lebih dioptimalkan, namun karena hasil dari SVR yang berbanding jauh dengan RF, hal ini sangat berpengaruh terhadap nilai rata-rata prediksi yang dihitung. Dengan analisa ini dapat membuktikan bahwa untuk melakukan perhitungan hybrid dapat dilakukan dengan metode yang mempunyai cara kerja relatif sama misalkan metode *Random Forest* dengan metode *Decission Tree* sebagai bahan penelitian berikutnya.

7.2 Implementasi Hasil

Dari hasil komparasi dan analisa beberapa metode dan telah ditentukan metode dengan performansi terbaik, maka selanjutnya dilakukan implementasi dari hasil prediksi dalam penelitian ini metode yang terbaik adalah metode prediksi dengan menggunakan *Random Forest.*

Data yang disiapkan sejumlah 18.769 terbagi dalam empat tujuan pulau Kalimantan, Sulawesi, Maluku dan Papua. Data ini kemudian diolah untuk dilakukan prediksi dengan metode yang tepat untuk menentukan SLA. Dalam setiap metode prediksi data terbagi menjadi *data testing* dan *data training* dengan komposisi 75:25 seperti pada tabel 7.4

Tabel 7.4 Komposisi Data Kiriman berdasar Tujuan

Pulau	Transaksi	Data Training	Data Testing
Kalimantan	11,407	8,580	2,827
Sulawesi	4,677	3,468	1,209
Papua	1,409	1,059	350
Maluku	1,276	970	306
Total	18,769	14,077	4,692

Metode RF sebagai metode terpilih kemudian hasil dari dataset prediksi RF dilakukan perhitungan rata-rata, nilai maksimum dan nilai minimum dari selisih antara *actual time* dibandingkan dengan hasil prediksi RF untuk diketahui tujuan mana yang mempunyai potensi perbaikan SLA yang lebih besar, seperti pada tabel 7.5

Tabel 7.5 Performa Nilai Rata-rata, Max, Min berdasar tujuan

Pulau	AVG	MAX	MIN
Kalimantan	0.01328	2.96208	-5.37000
Sulawesi	0.01809	2.32132	-4.21924
Papua	-0.06739	2.91760	-3.55271
Maluku	-0.09273	2.74984	-6.91333

Dari hasil perhitungan selisih pada tabel 7.5 menjelaskan bahwa tujuan pulau Kalimantan dan Sulawesi mempunyai perhitungan akurasi dengan model RF lebih tepat dengan nilai rata-rata paling kecil 0,01328 dan 0,01809 , namun terdapat nilai minimum yang cukup besar -5,37 dan -4,21934, terhadap tujuan dengan nilai minimun yang besar ini akan dilakukan penyesuaian di bidang operasi meliputi pola distribusi dan penjadwalan transportasi, begitu juga sama halnya untuk pulau Papua dan Maluku.

Metode prediksi ini sangat membantu untuk memberikan nilai awal SLA yang lebih baik dibandingkan SLA sebelumnya, terhadap hasil prediksi ini bisa digunakan untuk dilakkan evaluasi ketepatan waktu pengiriman dengan mengatur ulang pola operasi dengan disesuaikan

kapasitas produksi untuk masing-masing tujuan.

Harapan dengan penelitian ini dapat digunakan perusahaan untuk memperbaiki pola operasi untuk menghasilkan SLA yang lebih tepat agar kepercayaan pelanggan kembali tumbuh dan dapat bersaing lebih ketat dengan para kompetitor pada perusahaan sejenis.

7.3 Prediksi Service Level Agreement Kiriman Barang dan Dokumen dalam pandangan Islam

Pengiriman barang atau dokumen melalui PT Pos Indonesia dikategorikan sebagai akad *wakalah* (perwakilan) ditambah lagi dengan adanya SLA yang sudah ditetapkan kepada pelanggan yang berisi informasi waktu penyampaian barang atau dokumen. Dalam hal ini PT Pos Indonesia sebagai *mu'jir* yang memberikan jasa tenaganya untuk mengirimkan barang dan dokumen, sedangkan pelanggan sebagai *musta'jir* yang membutuhkan tenaga profesi Pos Indonesia untuk mengirimkan barangn dan dokumennya. M. Yazid Afandi (2009).

Dasar hukum *wakalah* terdapat dalam ayat al-Qur'an QS. An-Nisa (4) : 35 yang berbunyi :

وَاِنۡ خِفۡتُمۡ شِقَاقَ بَيۡنِهِمَا فَابۡعَثُوۡا حَكَمًا مِّنۡ اَهۡلِهٖ وَحَكَمًا مِّنۡ اَهۡلِهَا ۚ اِنۡ يُّرِيۡدَاۤ اِصۡلَاحًا يُّوَفِّقِ اللّٰهُ بَيۡنَهُمَا ؕ اِنَّ اللّٰهَ كَانَ عَلِيۡمًا خَبِيۡرًا

"Dan jika kamu khawatirkan ada persengketaan antara keduanya, maka kirimlah seorang hakam dari keluarga laki-laki dan seorang hakam dari keluarga perempuan, jika kedua orang hakam itu bermaksud mengadakan perbaikan, niscaya Allah memberi taufik kepada suami istri itu. Sesungguhnya Allah maha mengetahui lagi maha mengenal". (QS. An-Nisa: 35)

Dalam ayat ini memiliki makna bahwa ketika ada persegketaan antara dua orang maka diperbolehkan mengutus seseorang yang ahli (*Hakam*) sebagai wakil dalam menyelesaikan persengketaan yang terjadi. Dalam isi ayat diatas dijelas bahwa ada dua orang (suami istri) yang terlibat persengketaan dan sampai ingin melakukan perceraian, sehingga kedua pihak tersebut mengutus orang (*Hakam*) untuk mewakili mereka dalam menyelesaikan persengketaan yang terjadi.

Dalam Firman Allah dalam QS. Yusuf (12): 55, tentang ucapan yusuf kepada raja :

قَالَ اجْعَلْنِيْ عَلٰى خَزَاۤىِٕنِ الْاَرْضِ اِنِّيْ حَفِيْظٌ عَلِيْمٌ

Dia (Yusuf) berkata, *"jadikanlah aku bendaharawan negara (Mesir). Sesungguhnya aku adalah orang yang pandai menjaga lagi berpengetahuan."* (QS Yusuf:55)

Dalam ayat ini bercerita bahwa nabi Yusuf as. Mengusulkan untuk menjadikan dirinya sebagai bendaharawan di negara yang mewakili pemerintahan negara dalam mengelola keuangan negara, karena Nabi Yusuf adalah orang yang dapat dipercaya lagi pandai dan ahli dalam mengatur keuangan negara.

Dapat disimpulkan pula bahwa ayat ini mengandung pesan bahwa boleh melakukan perjanjian kerja atau mewakilkan suatu pekerjaan kepada orang lain yang memang dapat dipercaya dan telah ahli dalam bidang pekerjaan tersebut.

Adapun dalam hadits diriwayatkan Abu Dawud no 1438 berbunyi :

حَدَّثَنَا عُبَيْدُ اللَّهِ بْنُ سَعْدِ بْنِ إِبْرَاهِيمَ حَدَّثَنَا عَمِّي حَدَّثَنَا أَبِي عَنْ ابْنِ إِسْحَقَ عَنْ أَبِي نُعَيْمٍ وَهْبِ بْنِ كَيْسَانَ عَنْ جَابِرِ بْنِ عَبْدِ اللَّهِ أَنَّهُ سَمِعَهُ يُحَدِّثُ قَالَ أَرَدْتُ الْخُرُوجَ إِلَى خَيْبَرَ فَأَتَيْتُ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فَسَلَّمْتُ عَلَيْهِ وَقُلْتُ لَهُ إِنِّي أَرَدْتُ الْخُرُوجَ إِلَى خَيْبَرَ فَقَالَ إِذَا أَتَيْتَ وَكِيلِي فَخُذْ مِنْهُ خَمْسَةَ عَشَرَ وَسْقًا فَإِنْ ابْتَغَى مِنْكَ آيَةً فَضَعْ يَدَكَ عَلَى تَرْقُوَتِهِ

“Telah menceritakan kepada kami 'Ubaidullah bin Sa'dan bin Ibrahim telah menceritakan kepada kami Pamanku telah menceritakan kepada kami Ayahku dari Ibnu Ishaq dari Abu Nu'aim Wahb bin Kaisan dari Jabir bin Abdullah bahwa ia mendengarnya menceritakan, ia berkata, "Aku ingin pergi ke Khaibar, lalu aku datang menemui Rasulullah shallallahu 'alaihi wasallam, aku ucapkan salam kemudian berkata, "Sesungguhnya aku ingin pergi ke Khaibar." Kemudian beliau bersabda: "Apabila engkau datang kepada wakilku, maka ambillah darinya lima belas wasaq, dan apabila ia menginginkan tanda darimu maka letakkan tanganmu pada tulang bahunya!"

Dalam hadits Rasulallah diatas jelas tersirat bahwa pada masa Rasulullah beliau pun memiliki wakil yang dipercayainya dalam melakukan suatu pekerjaan.

Dari prepektif Islam bahwa pengiriman barang dan dokumen oleh PT Pos Indonesia dengan disertai penetapan SLA termasuk dalam akad *wakalah* yang disertai adanya upah *(ujrah)* atau disebut akad *wakalah bil ujrah* dengan menjalankan amanah ketentuan yang telah dietapkan yang salah satunya dengan adanya penetapan SLA waktu tempuh kiriman.

BAB VIII

KESIMPULAN

8.1 Kesimpulan

Dari hasil ujicoba dan pembahasan metode prediksi dapat diambil kesimpulan sebagai berikut :

a. Bahwa dari data kiriman barang dan dokumen PT Pos Indonesia Cabang Utama Malang untuk tujuan Wilayah 6 yang meliputi pulau Kalimantan, Sulawesi, Maluku dan Papua mulai Januari 2022 hingga Desember 2022 terdapat 18.769 data kiriman. Dari data tersebut variabel atau atribut yang digunakan adalah *distance, current_sla, weight, actual_time, deviation* dan *total_hub* dengan data variabel adalah *actual time* sebagai atribut untuk mengukur pencapaian prediksi dari masing masing metode.

b. Data dibagi menjadi 2 sebagai *data training* dengan *data testing* dengan komposisi data 75:25 atau 14.077 *data training* dan 4.692 *data testing*. Sebelum menggunakan prediksi dengan metode RF, SVR dan RF Hybrid pada data testing dilakukan pengukuran prediksi dengan MAE dan MAPE yang mendapatkan nilai 2,2589 MAE dan 54,9464 MAPE nilai ini tentunya jauh dari nilai akurasi prediksi yang diharapkan, oleh karena itu perlu dilakukan prediksi terhadap SLA yang sudah ditetapkan dengan beberapa metode.

c. Pada metode RF dilakukan eksperimen dengan 2 strategi, yang pertama dengan menggunakan seluruh variabel yang berpengaruh, yang kedua

dengan menggunakan variabel yang paling berpengaruh. Metode RF yang terbaik pada strategi kedua dengan MAE 0,6968 MAPE 15,7976. Eksperimen berikutnya pada metode SVR juga dilakukan 2 strategi, yang pertama menggunakan variabel yang paling berpengaruh yaitu *deviation* sedangkan yang kedua menggunakan variabel *distance*. Dengan menggunakan 2 kernel RBF dan Polynomial untuk masing-masing strategi menghasilkan nilai pengukuran terbaik dengan akurasi MAE 1,0407 MAPE 24,3927 pada kernel RBF. Dari 2 metode ini hasil pengukuran akurasi prediksi yang terbaik pada metode RF.

d. Metode RF Hybrid sebagai metode terakhir yang dilakukan dalam penelitian ini dengan menghitung hasil rata-rata nilai prediksi yang terbaik antara RF dan SVR terlebih dahulu, lalu dikomparasi dengan metode RF, metode RF sebagai metode yang mempunyai nilai akurasi yang terbaik. Untuk RF hybrid ini menghasilkan nilai pengukuran akurasi MAE 0,7254 dan MAPE 16,77. Hasil ini lebih rendah dari RF walaupun perbedaannya tidak signifikan, dengan demikian pada penelitian ini metode RF mempunyai nilai akurasi yang terbaik dibandingkan dengan SVR dan RF Hybrid.

e. Tujuan Kalimantan dan Sulawesi mempunyai nilai prediksi lebih baik dibandingkan dengan tujuan Maluku dan Papua, walaupun selisih tidak *significant.* Hal ini membuktikan metode RF yang dipilih memberikan nilai yang cukup baik dengan selisih rata-rata antaran *actual time* dan hasil prediksi pada sekitaran angka 0.

f. Hasil prediksi ini dapat dijadikan alat ukur untuk menentukan SLA yang sudah ditetapkan sebelumnya, beberapa nilai akurasi untuk beberapa tujuan perlu dilakukan penyesuaian dengan melakukan perbaikan pada jadwal transportasi dan simpul-simpul hub pada pola operasi di PT Pos Indonesia.

8.2 Saran

Dalam penelitian ini tentunya dapat dijadikan pengembangan untuk penelitian berikutnya, oleh karena itu terdapat beberapa saran dari penulis.

a. Penelitian dapat dilakukan pada dataset yang lebih besar dengan kantor asal dari produksi kiriman yang paling besar dengan wilayah kiriman ke seluruh Indonesia.
b. Pada metode RF Hybrid dilakukan pada metode dengan model perhitungan yang sejenis seperti *Decission Tree* dibandingkan dengan RF, sehingga hasil dari RF hybrid lebih optimal. Begitu juga jika menggunakan metode hybrid yang lain misal metode ARIMA maka metode pembanding rata-rata dilakukan dengan model yang sama misalkan dengan ARIMA-ANN dengan ARIMA-Kalman.

DAFTAR PUSTAKA

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45(1):5–32. doi: 10.1023/A:1010933404324.

Bristy, Anika Hossain, Alimozzaman Durjoy, and Md Hasibul Hasan. 2021. "E-Commerce Shipping: Prediction of on Time Delivery of Products Using Data Mining." (January).

Chai. T., Draxler, R. R. 2014. "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE) – Arguments against Avoiding RMSE in the Literature. Geoscientific Model Development Geosci Model Dev." Vol.7:1247.

Florio, Alexandre. 2021. "2021 Last-Mile Routing Research Challenge Article XII A Machine Learning Framework for Last-Mile Delivery Optimization." (September).

Heidy, Pratiwi. 2020. "Konsep Dasar Dalam Machine Learning." *Medium.Com* ISBN 9786230210839 Hal 1-7.

Hiles, Andrew. 2016. "Service Level Agreements : Winning a Competitive Edge for Support & Supply Services." 1005.

Hudzaifah, Muhammad, and Ali Akbar Rismayadi. 2021. "Peramalan Arus Lalu Lintas Berdasarkan Waktu Tempuh Dan Cuaca Menggunakan Metode Time Series Decomposition." *Jurnal Responsif : Riset Sains Dan Informatika* 3(2):207–15. doi: 10.51977/jti.v3i2.559.

Kharwal, Aman. 2020. "Why Random_state=42 in Machine Learning?" Retrieved May 2, 2023 (https://thecleverprogrammer.com/2020/12/17/why-random_state42-in-machine-learning/).

Kisgyörgy, Lajos, and Laurence R. Rilett. 2002. "Travel Time Prediction by Advanced Neural Network." *Periodica Polytechnica Civil Engineering* 46(1):15–32.

Krisma, A., & Azhari, M. 2019. "Perbandingan Metode Double Exponential Smoothing Dan Triple Exponential Smoothing Dalam Parameter Tingkat Error Mean Absolute Percentage Error (MAPE) Dan Means Absolute Deviation (MAD)." 4(2):81–87.

Liu, Sheng, Long He, and Zuo Jun Max Shen. 2021. "On-Time Last-Mile Delivery: Order Assignment with Travel-Time Predictors." *Management Science* 67(7):4095–4119. doi: 10.1287/mnsc.2020.3741.

M. Yazid Afandi. 2014. *Fiqih Muamalah*. Surabaya: UIN Sunan Ampel Pres.

Megantara, Enrico. 2022. “Prediksi Waktu Tempuh Bus Trans Metro Bandung Dengan Internet Of Things Dan Metode Machine Learning Arrival Time Prediction Bus Trans Metro Bandung with Internet Of Things and Machine Learning Method.” 8(6):3801–6.

Nova. 2023. “The Power of the Radial Basis Function (RBF) Kernel - AITechTrend.” Retrieved May 3, 2023 (https://aitechtrend.com/the-power-of-the-radial-basis-function-rbf-kernel/).

Pambudi, Hardian Kokoh, Putu Giri Artha Kusuma, Femi Yulianti, and Kevin Ahessa Julian. 2020. “Prediksi Status Pengiriman Barang Menggunakan Metode Machine Learning.” *Jurnal Ilmiah Teknologi Infomasi Terapan* 6(2):100–109. doi: 10.33197/jitter.vol6.iss2.2020.396.

Quah, Jo Wei, Chin Hai Ang, Regupathi Divakar, Rosnah Idrus, Nasuha Lee Abdullah, and Xin Ying Chew. 2019. *Timing-of-Delivery Prediction Model to Visualize Delivery Trends for Pos Laju Malaysia by Machine Learning Techniques*. Vol. 937. Springer Singapore.

Rio, Afrioni Roma, Berton Maruli Siahaan, Jurusan Fisika, Fakultas Matematika, Pengetahuan Alam, Universitas Sam Ratulangi, and Kampus Unsrat. 2022. “Prediksi Jarak Tempuh Kapal Motor Sangiang Menggunakan Supervised Machine Learning.” 2(2):134–36.

Sethi, Alakh. 2020. “Support Vector Regression In Machine Learning.” *Analytics Vidhya*. Retrieved May 3, 2023 (https://www.analyticsvidhya.com/blog/2020/03/support-vector-regression-tutorial-for-machine-learning/).

Sugiyono. 2014. *Metode Penelitian Kuantitatif, Kualitatif Dan R & D*. Bandung: Alfabeta.

Vanajakshi, Lelitha, and Laurence R. Rilett. 2007. “Support Vector Machine Technique for the Short Term Prediction of Travel Time.” *IEEE Intelligent Vehicles Symposium, Proceedings* 600–605. doi: 10.1109/ivs.2007.4290181.

Wu, Chun Hsin, Jan Ming Ho, and D. T. Lee. 2004. “Travel-Time Prediction with Support Vector Regression.” *IEEE Transactions on Intelligent Transportation Systems* 5(4):276–81. doi: 10.1109/TITS.2004.837813.

Wustenhoff, E., S. BluePrints-Sun BluePrints Online, and undefined 2002. 2002. “Service Level Management in the Data Center.” *Filibeto.Org*.