

**PENERAPAN METODE *RANDOM FOREST* PADA PREDIKSI PENILAIAN
NILAI ASET KJPP SIG MALANG BERBASIS WEB**

SKRIPSI

**Oleh:
MUNIRUL HUDA
NIM. 19650069**



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

**PENERAPAN METODE *RANDOM FOREST* PADA PREDIKSI
PENILAIAN NILAI ASET KJPP SIG MALANG BERBASIS WEB**

SKRIPSI

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

Oleh :
MUNIRUL HUDA
NIM. 19650069

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

HALAMAN PERSETUJUAN

**PENERAPAN METODE *RANDOM FOREST* PADA PREDIKSI
PENILAIAN NILAI ASET KJPP SIG MALANG BERBASIS WEB**

SKRIPSI

Oleh :
MUNIRUL HUDA
NIM. 19650069

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 27 November 2023

Pembimbing I,




Dr. M. Faisal, M.T
NIP. 19740510 200501 1 007

Pembimbing II,



Dr. M. Imamudin Lc, MA
NIP. 19740602 200901 1 010

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



Dr. Fachrul Kurniawan, M.MT, IPM
NIP. 19771020 200912 1 001

HALAMAN PENGESAHAN

PENERAPAN METODE *RANDOM FOREST* PADA PREDIKSI PENILAIAN NILAI ASET KJPP SIG MALANG BERBASIS WEB

SKRIPSI

Oleh:
MUNIRUL HUDA
NIM. 19650069

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 11 Desember 2023

Susunan Dewan Penguji

Ketua Penguji : Dr. Fachrul Kurniawan, M.MT, IPM
NIP. 19771020 200912 1 001

Anggota Penguji I : Puspa Miladin N. S. A. B, M.Kom
NIP. 19930828 20180201 2 238

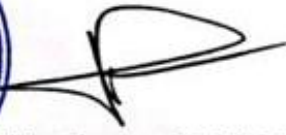
Anggota Penguji II : Dr. M. Faisal, M.T
NIP. 19740510 200501 1 007

Anggota Penguji III : Dr. M. Imamudin Lc, MA
NIP. 19740602 200901 1 010

()
()
()
()

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




Dr. Fachrul Kurniawan, M.MT, IPM
NIP. 19771020 200912 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Munirul Huda
NIM : 19650069
Fakultas / Prodi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : Penerapan Metode *Random Forest* Pada Prediksi Penilaian Nilai Asset KJPP SIG Malang Berbasis WEB

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

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

Malang, 29 November 2023
Yang membuat pernyataan,



Munirul Huda
NIM.19650069

HALAMAN MOTTO

...Fall seven times, Stand up eight...

HALAMAN PERSEMBAHAN

Skripsi ini saya persembahkan untuk Kedua Orang Tua saya,
Bapak Agus Sholah dan Ibu Sri Yuliati Ningsih, Keluarga,
Seluruh Dosen, Sahabat, Teman-Teman Seperjuangan,
Seluruh Pihak yang Terlibat, dan Diri Saya Sendiri

Terima kasih

KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Alhamdulillah rabbil 'alamin, puji syukur kehadiran Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Skripsi yang berjudul “Penerapan Metode *Random Forest* Pada Prediksi Penilaian Nilai Aset KJPP SIG Malang berbasis WEB“. Shalawat serta salam semoga selalu tetap tercurahkan kepada baginda Nabi Muhammad SAW yang telah menerangi dunia dengan cahaya iman dan islam. Penulis mengucapkan terima kasih kepada pihak-pihak yang membantu proses penulisan Skripsi ini. Ucapan terima kasih ini, penulis sampaikan kepada:

1. Prof. Dr. M. Zainuddin, M.A., selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Sri Harini, M.Si., selaku Dekan Fakultas Sains dan Teknologi Universitas Maulana Malik Ibrahim Malang.
3. Dr. Fachrul Kurniawan, M.MT, IPM, selaku Ketua Program Studi Universitas Maulana Malik Ibrahim Malang.
4. Dr. M. Faisal, M.Kom selaku dosen pembimbing I yang sabar untuk memberikan bimbingan, memberikan nasehat, saran dan memberikan pengarahan sehingga penulis dapat menyelesaikan skripsi ini.
5. Dr. M. Imamudin Lc, Ma selaku pembimbing II yang sabar memberikan bimbingan, pengarahan dan nasehat sehingga penulis dapat menyelesaikan skripsi ini.

6. Dr. Fachrul Kurniawan, M.MT, IPM selaku penguji I yang telah memberikan saran dan kritik yang telah berikan selama proses pengujian skripsi ini.
7. Puspa Miladin Nuraida Safitri A. basid, M.Kom selaku penguji II yang telah memberikan saran dan kritik yang telah berikan selama proses pengujian skripsi ini.
8. Hani Nurhayati, M.T, selaku Dosen Wali yang senantiasa membantu serta turut memberikan ilmu, saran dan motivasi belajar untuk penulis selama menjalani masa studi di Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
9. Seluruh dosen Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang yang telah memberikan ilmu, pengetahuan, pengalaman dan wawasan yang banyak bagi penulis.
10. Orang tua saya yaitu Ayahanda saya Agus Sholah, Ibunda saya Sri Yulianti Ningsih dan kakak saya yang telah banyak memberikan perhatian, nasihat, doa, motivasi, dukungan baik moral maupun materil bagi penulis sehingga dapat menyelesaikan skripsi ini dengan penuh perjuangan.
11. Teman-teman yang telah memberikan semangat, saran dan juga doa kepada penulis.
12. Semua pihak yang telah memberikan bantuan baik dukungan nyata ataupun semangat serta pihak-pihak yang tidak dapat penulis sebutkan.
13. Terimakasih untuk diri sendiri yang selalu sabar, selalu bertahan dan semangat dengan penuh perjuangan walaupun hampir berada pada posisi terendah, sehingga penulis dapat menyelesaikan skripsi ini dengan baik.

Dengan penuh kesadaran, penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan. Oleh sebab itu, penulis berharap kritik serta saran untuk penelitian kedepannya. Penulis berharap semoga skripsi ini dapat memberikan manfaat kepada pembaca khususnya bagi diri penulis. Semoga Allah SWT melimpahkan keberkahan atas skripsi penulis dan senantiasa selalu diberikan petunjuk dan rahmat-Nya serta segala bentuk kebaikan dalam menarungi kehidupan kita di dunia.

Wassalamu'alaikum Wr. Wb.

Malang, 29 November 2023

Penulis

DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
HALAMAN MOTTO	vi
HALAMAN PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xiv
ABSTRAK	xv
ABSTRACT	xvi
المخلص.....	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	6
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	7
BAB II TINJAUAN PUSTAKA.....	8
2.1 Penelitian Terkait.....	8
2.2 Kantor Jasa penilai Publik	14
2.3 Sistem Informasi Geografi (SIG)	16
2.4 <i>Decision Tree</i> (Pohon Keputusan).....	17
2.5 Metode <i>Random Forest</i>	19
2.6 Mean Absolute Percentage Error (MAPE).....	21
BAB III METODOLOGI PENELITIAN	23
3.1 Pengumpulan data	24
3.2 Perancangan Sistem.....	25
3.2.1 Preprocessing.....	25
3.2.2 Data Splitting	26
3.2.3 Pengujian <i>Random Forest</i>	27
3.2.4 Analisa Peforma.....	34
BAB IV HASIL DAN PEMBAHASAN.....	35
4.1 Hasil <i>Datasets</i>	35
4.2 Hasil Preprocessing.....	36
4.2.1 Hasil Data Cleaning	37
4.2.2 Hasil Data Exploration.....	37
4.3 Hasil Data Splitting	39
4.3 Hasil Pengujian <i>Random Forest</i>	40

4.3.1 Hasil Pengujian Rasio 90:10.....	40
4.3.2 Hasil Pengujian Rasio 80:20.....	41
4.3.3 Hasil Pengujian Rasio 70:30.....	42
4.3.4 Hasil Pengujian Rasio 60:40.....	44
4.3.5 Hasil Pengujian Rasio 50:50.....	47
4.4 Analisa Hasil Pengujian	50
4.5 Hasil Tampilan Aplikasi GIS	54
4.5.1 Halaman Main Map	54
4.5.2 Halaman Uji <i>Random Forest</i>	55
4.6 Integrasi Penelitian dengan Islam.....	56
BAB V KESIMPULAN DAN SARAN	60
5.1 Kesimpulan	60
5.2 Saran	61
DAFTAR PUSTAKA	

DAFTAR GAMBAR

Gambar 2. 1 Komponen GIS.....	16
Gambar 2. 2 <i>Random Forest</i>	21
Gambar 3. 1 Desain Sistem.....	23
Gambar 3. 2 <i>Decision Tree</i>	28
Gambar 3. 3 Kemungkinan Pohon Split 1	30
Gambar 3. 4 Random Pohon Split 2.....	32
Gambar 3. 5 Hasil <i>Random forest</i>	33
Gambar 4. 1 Correlation Matrix.....	38
Gambar 4. 2 Hasil MAE.....	51
Gambar 4. 3 Hasil MAPE	51
Gambar 4. 4 GIS Peta Asset.....	54
Gambar 4. 5 Tampilan table data asset	55
Gambar 4. 6 Halaman Hasil Uji <i>Random Forest</i>	56

DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	11
Tabel 3. 1 Data Penelitian	24
Tabel 3. 2 Contoh Data	25
Tabel 3. 3 Contoh Perhitungan Data	29
Tabel 3. 4 Random Subset 1	30
Tabel 3. 5 Random Subset 2	31
Tabel 4. 1 Atribut Dataset	35
Tabel 4. 2 Penggantian Null Value kolom Indikasi Nilai	37
Tabel 4. 3 Hasil Rasio dan Splitting Data	39
Tabel 4. 4 Data Pengujian 90:10	40
Tabel 4. 5 Data Pengujian 80:20	41
Tabel 4. 6 Data Pengujian 70:30	42
Tabel 4. 7 Data Pengujian 60:40	44
Tabel 4. 8 Data Pengujian 50:50	47
Tabel 4. 9 Tabel Hasil Pengujian <i>Random Forest</i>	50

ABSTRAK

Huda, Munirul. 2023. **Penerapan Metode *Random Forest* Pada Prediksi Penilaian Nilai Aset KJPP SIG Malang berbasis WEB.** Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. M. Faisal, M.Kom. (II) Dr. M. Imamudin Lc, Ma.

Kata kunci: Nilai Aset, *Random Forest*, *Machine Learning*, *Decision Tree*, *Regression*, *GIS*.

Menyoroti kebutuhan akan metode prediksi harga yang lebih akurat dalam konteks perdagangan aset. Penelitian ini bertujuan untuk mengembangkan aplikasi machine learning yang dapat memprediksi harga aset berdasarkan spesifikasi aset. Metodologi penelitian melibatkan pengambilan data aset dari perusahaan penilaian harga aset, dan penerapan algoritma machine learning *Random Forest* yang terintegrasi dengan Web GIS untuk memfasilitasi evaluasi *Random Forest*. Pada rasio data training 70:30, hasil eksperimen menunjukkan hasil maksimal dengan Mean Absolute Percentage Error (MAPE) sebesar 11,29. Hasil ini mengindikasikan bahwa model *Random Forest* mampu memberikan estimasi harga aset dengan tingkat akurasi yang signifikan. Implikasinya, penelitian ini memberikan kontribusi pada pengembangan metode prediksi harga aset yang lebih handal dan dapat diandalkan dalam konteks perdagangan. Kesimpulannya, *Random Forest* dapat dijadikan alat yang efektif untuk memprediksi harga aset, membuka potensi penerapan yang luas di berbagai sektor industri.

ABSTRACT

Huda, Munirul. 2023. **Penerapan Metode *Random Forest* Pada Prediksi Penilaian Nilai Asset KJPP SIG Malang berbasis WEB.** Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. M. Faisal, M.Kom. (II) Dr. M. Imamudin Lc, Ma.

Highlighting the need for more accurate price prediction methods in the context of asset trading, this research aims to develop a machine learning application capable of predicting asset prices based on asset specifications. The research methodology involves obtaining asset data from asset pricing assessment companies and implementing the *Random Forest* machine learning algorithm integrated with Web GIS to facilitate *Random Forest* evaluation. With a 70:30 training data ratio, experimental results show a maximum outcome with a Mean Absolute Percentage Error (MAPE) of 11.29. These findings indicate that the *Random Forest* model can provide asset price estimates with a significant level of accuracy. The implication is that this research contributes to the development of more reliable and dependable asset price prediction methods in the trading context. In conclusion, *Random Forest* can be an effective tool for predicting asset prices, opening up potential applications across various industrial sectors.

Key words: Asset Value, *Random Forest*, *Machine Learning*, *Decision Tree*, *Regression*, *GIS*.

المخلص

الهدى، منير. 2023. تطبيق طريقة الغابات العشوائية (*Random Forest*) في التنبؤ بتقييم قيمة الأصول في مكتب خدمات التقييم العام سوغينج-إروان-غوناوان على أساس الموقع الإلكتروني. البحث الجامعي. قسم الهندسة المعلوماتية، كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: د. محمد فيصل، الماجستير. المشرف الثاني: د. محمد إمام الدين، الماجستير.

الكلمات الرئيسية: قيمة الأصول، الغابة العشوائية، التعلم الآلي، شجرة القرار، الانحدار، نظم المعلومات الجغرافية.

يسلط الضوء على الحاجة إلى طرق أكثر دقة للتنبؤ بالأسعار في سياق تداول الأصول. يهدف هذا البحث إلى تطوير تطبيقات التعلم الآلي التي يمكنها التنبؤ بأسعار الأصول بناء على مواصفات الأصول. تتضمن منهجية البحث أخذ بيانات الأصول من شركات تقييم أسعار الأصول، وتطبيق خوارزمية التعلم الآلي "الغابات العشوائية" المدمجة مع الموقع الإلكتروني لنظام المعلومات الجغرافية لتسهيل تقييم الغابات العشوائية. عند نسبة بيانات التدريب 70:30، أظهرت النتائج التجريبية نتائج قصى مع متوسط النسبة المئوية المطلقة للخطأ (MAPE) من 11.29. أشارت هذه النتائج إلى أن نموذج الغابات العشوائية قادر على توفير تقديرات أسعار الأصول بدرجة كبيرة من الدقة. ضمناً، يساهم هذا البحث في تطوير طرق أكثر فعالية وموثوقة للتنبؤ بأسعار الأصول في سياق التجارة. في الختام، يمكن أن تكون الغابات العشوائية أداة فعالة للتنبؤ بأسعار الأصول، وإطلاق العنان لإمكانية التطبيق على نطاق واسع في مختلف القطاعات الصناعية.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada masa sekarang khususnya Indonesia usaha bidang pajak terus mengalami pertumbuhan. Perusahaan yang berjalan pada bidang jasa tersebut memiliki peran yang sangat penting dalam perekonomian negara. Salah satu perusahaan jasa yang mengalami perkembangan tersebut adalah jasa penilaian atau appraisal. Appraisal merupakan suatu profesi yang menyediakan jasa dalam menilai sebuah aset (Kartika & Oktarini, 2022). Appraisal sendiri merupakan sebuah proses pemberian sebuah taksir berupa angka pada benda nyata yang proses tersebut dilakukan oleh tenaga kerja profesional yang biasa disebut appraiser (penilai). Appraiser merupakan seseorang yang memiliki keahlian atau kualifikasi khusus dalam menilai atau menentukan nilai suatu properti atau aset, seperti tanah, bangunan, peralatan, barang, atau aset lainnya. Appraiser biasanya bekerja sebagai bagian dari tim profesional di KJPP (Kantor Jasa Penilai Publik) untuk menyediakan penilaian yang terpercaya dan independen dalam menentukan nilai aset properti (Reza et al., 2020).

KJPP adalah singkatan dari "Kantor Jasa Penilai Publik" dalam bahasa Indonesia. Secara harfiah, KJPP mengacu pada kantor yang memberikan layanan jasa penilaian publik. Penilaian publik dalam konteks ini berkaitan dengan penilaian nilai properti atau aset, seperti tanah, bangunan, peralatan, dan lain sebagainya, yang dapat digunakan untuk berbagai keperluan, termasuk transaksi jual beli, pembiayaan, akuntansi, perpajakan, serta tujuan hukum dan regulasi lainnya (Rivai,

2023). Perusahaan jasa penilaian adalah entitas yang bersifat independen dalam melakukan penilaian terhadap properti milik perusahaan atau individu. Dalam menjalankan aktivitasnya, setiap perusahaan memiliki fokus untuk mencapai tujuan perusahaan yang pada dasarnya adalah untuk menghasilkan pendapatan yang maksimal. Pendapatan ini memiliki peran yang sangat penting dalam mengembangkan bisnis dan mencerminkan tingkat kesuksesan dan prestasi suatu perusahaan. Salah satu faktor yang memengaruhi keberhasilan dan kesuksesan suatu bisnis adalah produktivitas tenaga kerja (Kartika & Oktarini, 2022).

KJPP memiliki peran penting dalam memberikan penilaian independen dan objektif terhadap nilai aset. Hal ini membantu memastikan bahwa nilai-nilai yang diakui dalam transaksi atau laporan keuangan tercermin secara akurat sesuai dengan kondisi pasar dan faktor-faktor lain yang relevan. Penilai di KJPP harus mempertimbangkan berbagai faktor seperti lokasi, ukuran, kondisi fisik, pasar, dan faktor-faktor lain yang dapat memengaruhi nilai aset. Perkembangan Kantor Jasa Penilai (KJPP) adalah hasil dari berbagai faktor ekonomi, hukum, dan teknologi yang memengaruhi industri penilaian properti. beberapa perkembangan penting yang memengaruhi KJPP diantaranya adalah teknologi dan data. Penggunaan teknologi dalam penilaian properti telah mengalami kemajuan signifikan, seperti Sistem informasi Geografis (SIG) (Kartika & Oktarini, 2022).

Seperti halnya KJPP Sugeng irawan dan rekan yang merupakan sebuah kantor jasa penilai publik yang bergerak dibidang jasa studi kelayakan, monitoring project, penilaian properti, penilaian bisnis, due diligence dan restrukturisasi, studi highest and best used yang merupakan wadah bagi para profesional dalam rangka

pengabdian diri terhadap masyarakat, dunia usaha dan lembaga keuangan. Khususnya dalam hal merancang, melindungi dan mengamankan investasi dunia usaha.

Kantor KJPP Sugeng Irawan dan Rekan berkomitmen untuk menyediakan layanan profesional penilaian dan konsultasi yang berkualitas tinggi serta dapat dipercaya dalam berbagai bidang seperti bisnis, manajemen, properti, industri, dan keuangan. Kami menjalankan misi ini secara independen, mengikuti standar penilaian Indonesia (SPI), Kode Etik Penilai Indonesia (KEPI), serta peraturan dan ketentuan lain yang berlaku. Selain itu, kami melaksanakan prosedur-prosedur jaminan kualitas guna memastikan pelayanan yang optimal. Terdapat beberapa permasalahan yang mengakibatkan visi dan misi Pada kantor kjpp SIG Malang belum tercapai, diantaranya pada kantor ini masih belum adanya sistem pengelolaan data yang cukup baik, dimana data hanya tersimpan dalam bentuk media excel. Hal tersebut bertentangan dengan perkembangan yang dialami oleh KJPP pada umumnya sehingga data akan terus bertambah dan pengelolaan data akan menjadi lebih rumit., sehingga dibutuhkan aplikasi seperti GIS agar data mudah dikelola dan mengikuti perkembangan KJPP pada bidang teknologi. Kegunaan GIS sendiri mempermudah penginputan data berupa titik lokasi agar proses penginputan data lebih mudah dan mengurangi kesalahan dalam menginputkan data. Selain pembuatan GIS penulis berinisiatif menambahkan fitur dalam memprediksi harga aset agar mempermudah pihak KJPP untuk melakukan penilaian.

Oleh sebab itu dengan adanya permasalahan tersebut maka dilakukan sebuah penelitian ini untuk perancangan pembuatan Sistem informasi Geografis

berbasis web dan menggunakan metode *Random Forest* sebagai prediksi harga nilai asset Sistem Informasi Geografi (SIG) adalah suatu sistem informasi yang berfungsi untuk mengintegrasikan, menyimpan, mengambil, memproses, menganalisis, dan menghasilkan data yang berkaitan dengan informasi geografis atau geospasial, dengan tujuan untuk mendukung proses pengambilan keputusan dalam proses perencanaan. Teknologi GIS menggabungkan berbagai proses pengolahan data dari basis data yang umumnya digunakan saat ini, termasuk pengambilan data sesuai dengan kebutuhan, analisis statistik, dan visualisasi yang khas, serta memberikan sejumlah keunggulan melalui analisis geografis yang ditampilkan dalam bentuk peta (Eka Dwi Cahyono, 2001). Dalam prediksi penilaian aset peneliti menggunakan metode *Random Forest*, dimana *Random Forest* dapat meningkatkan tingkat ketepatan karena prosesnya melibatkan pembuatan simpul anak secara acak pada setiap tingkat simpul atasnya, serta penggabungan hasil klasifikasi dari berbagai pohon (*tree*). Hasil klasifikasi yang paling sering muncul kemudian dipilih sebagai hasil akhir (Saadah & Salsabila, 2021a).

Pada Al – Quran terdapat ayat yang membahas tentang dengan adanya ilmu maka manusia dapat mengetahui sesuatu yang tidak diketahui. Dengan adanya ilmu menjadikan manusia untuk menemukan sesuatu yang bermanfaat bagi dirinya sendiri maupun orang lain, serta setiap orang lebih memperhatikan apa yang akan terjadi pada hari esok yakni pada surah Al – Hasyr ayat 18 yang berbunyi:

يَا أَيُّهَا الَّذِينَ آمَنُوا اتَّقُوا اللَّهَ وَلْتَنْظُرْ نَفْسٌ مَّا قَدَّمَتْ لِإِعْدٍ وَاتَّقُوا اللَّهَ إِنَّ اللَّهَ خَبِيرٌ بِمَا تَعْمَلُونَ

“Wahai orang-orang yang beriman! Bertakwalah kepada Allah dan hendaklah setiap orang memperhatikan apa yang telah diperbuatnya untuk hari esok (akhirat), dan bertakwalah kepada Allah. Sungguh, Allah Mahateliti terhadap apa yang kamu kerjakan.” (QS. Al-Hasyr: 18)

Allah menegaskan kepada hamba-hambaNya yang beriman agar memenuhi tanggung jawab iman dengan bertaqwa kepada-Nya, baik dalam kesunyian maupun keramaian, serta dalam segala aspek kehidupan. Mereka diminta untuk menjaga segala perintah, syariat, dan batasan yang telah ditetapkan-Nya, serta mempertimbangkan konsekuensi baik dan buruk dari amal perbuatan mereka, yang dapat membawa manfaat atau malapetaka di akhirat.

Jika mereka meletakkan akhirat sebagai prioritas utama dan sebagai fokus utama dalam hati mereka, maka mereka akan fokus untuk melaksanakan amalan-amalan akhirat. Mereka akan berusaha keras untuk meningkatkan amalan-amalan yang dapat membawa mereka ke surga, sambil membersihkannya dari segala halangan yang dapat menghalangi mereka. Kesadaran bahwa "sesungguhnya Allah Maha Mengetahui apa yang kamu kerjakan" mendorong mereka untuk bersungguh-sungguh dan konsisten dalam beramal (Ghoffar, 2000).

Dari tafsir di atas dapat dijelaskan pentingnya introspeksi diri dan *muhasabah*. Setiap individu diingatkan untuk selalu mengkaji dirinya sendiri. Jika terdapat kesalahan, mereka harus segera memperbaikinya dengan melepaskan diri dari dosa, bertaubat sungguh-sungguh, dan menjauh dari segala hal yang dapat mengantarkan pada kesalahan tersebut. Jika mereka menilai bahwa mereka kurang rajin dalam menjalankan perintah Allah, mereka harus berupaya dengan sungguh-sungguh, memohon pertolongan kepada Allah untuk memperbaiki diri, dan

membandingkan karunia dan kebaikan Allah dengan kemalasan mereka, sehingga mereka merasa malu.

Berdasarkan ayat diatas, peneliti mengimplementasikan ilmu metode *Random Forest* dalam memprediksi nilai asset serta pengembangan GIS pada KJPP Sugeng dan rekan Kota Malang.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat diidentifikasi sebuah permasalahan yaitu bagaimana memprediksi nilai aset pada KJPP Kota Malang menggunakan Metode *Random Forest*.

1.3 Batasan Masalah

Dalam memastikan pembahasan sesuai dengan rumusan dan tujuan masalah, dibutuhkan penentuan batasan masalah sebagai berikut:

1. Data yang digunakan merupakan data yang diperoleh dengan cara observasi pada KJPP Sugeng irawan dan rekan Kota Malang.
2. Data yang digunakan merupakan data survet pada tahun 2021 sampai tahun 2023.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini sesuai dengan identifikasi masalah yang telah dijelaskan yakni mengetahui tingkat akurasi penilaian asset menggunakan Metode *Random Forest*.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini ntuk membantu pengurus KJPP Sugeng Irawan dan rekan Kota Malang agar lebih mudah dalam pengelolaan data serta dapat memprediksi nilai asset yang akan dinilai.

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Menurut maharani Penentuan lokasi menggunakan Sistem Informasi Geografi (SIG) memungkinkan untuk menggambarkan informasi mengenai jarak dan tujuan sesuai dengan kebutuhan pengguna. SIG dapat digunakan secara akurat sebagai alat untuk mengelola data lokasi ATM Bank Kaltim di Kota Samarinda. Salah satu metode yang dapat digunakan adalah Haversine Formula, yang digunakan dalam aplikasi SIG untuk menentukan lokasi ATM terdekat dengan pengguna. Haversine Formula adalah persamaan yang digunakan dalam navigasi untuk menghitung jarak antara dua titik pada permukaan bola (Bumi) berdasarkan garis bujur dan lintang. SIG yang dibuat dapat memberikan informasi tentang lokasi ATM terdekat dan panduan untuk mencapainya. Hasil penelitian ini menunjukkan bahwa Haversine Formula dapat diterapkan untuk menghitung lokasi ATM terdekat dan memberikan informasi tentang ATM kepada pengguna atau nasabah Bank Kaltim (Maharani et al., 2017).

Pada penelitian dwiati dalam rangka mengatasi masalah banjir, penelitian tersebut merancang dan mengimplementasikan prototipe sistem drainase perkotaan yang berkelanjutan dengan optimalisasi menggunakan Sistem Informasi Geografi. Hal ini diharapkan akan mempercepat proses pengambilan keputusan dalam upaya penanggulangan banjir yang terprogram dan terencana, yang pada gilirannya dapat membantu meningkatkan ketahanan perkotaan terhadap bencana banjir (Wismarini & Ningsih, 2010).

Surya Afnarius, dkk (2017) melakukan penelitian mengenai permasalahan pada masyarakat pengunjung bukit tinggi yang mengalami kesulitan dalam mencari informasi mengenai lokasi tempat ibadah sesuai dengan kebutuhan pengunjung. Dalam mengatasi hal tersebut peneliti bertujuan untuk mengembangkan Aplikasi Sistem informasi Geografi (SIG). Hasil penelitian tersebut Aplikasi SIG memungkinkan pengguna untuk mencari, melihat, dan mendapatkan informasi tentang tempat ibadah di Kota Bukittinggi. Hasil pengujian dengan metode *blackbox* menunjukkan bahwa aplikasi ini telah memenuhi semua fungsi sistem yang diharapkan. Dengan demikian, Aplikasi SIG ini akan memberikan kontribusi positif dalam mendukung pariwisata halal di Kota Bukittinggi dengan memudahkan pengunjung dalam menjalankan ibadah sesuai kebutuhan mereka (Afnarius et al., 2017).

Barini Harahap, dkk (2016) melakukan sebuah penelitian mengenai Prediksi indeks harga saham menggunakan metode *Random Forest*, pada penelitian ini menghasilkan nilai MAPE sebesar 4.1173%, 8.6391%, dan 9.1708%. Dari hasil penelitian ini dapat disimpulkan bahwa metode SVR baik untuk prediksi jangka pendek sedangkan metode SVR-RF baik untuk prediksi jangka panjang (Harahap et al., 2015).

Pada penelitian Agi Pratomo, dkk (2019) melakukan prediksi pergerakan harga saham dengan metode *Random Forest* menggunakan trend Deterministic Data Preparation. Dari hasil percobaan didapat akurasi prediksi dari setiap perusahaan yaitu PT Astra International Tbk, PT Garuda Indonesia Tbk dan PT Indosat Tbk dengan menggunakan data Trend Deterministic yaitu 59,50%, 61,40% dan 59,44%.

Sedangkan akurasi menggunakan data Non Trend Deterministic yaitu 75,27%, 75,76% , dan 75,10% (Pratomo et al., 2019).

Rismita wahyu melakukan sebuah penelitian mengenai Prediksi indeks harga Televisi menggunakan metode *Random Forest* dan framework flask. Berdasarkan hasil analisis menggunakan algoritma *Random Forest* dengan data training dan data testing yang dibagi dalam perbandingan 75%:25%, ditemukan bahwa dalam pelatihan awal dengan model *Random Forest* regressor menggunakan parameter default menghasilkan tingkat akurasi sebesar 75,75%. Upaya untuk meningkatkan tingkat akurasi dilakukan melalui penyesuaian hyperparameter. Setelah melakukan tuning hyperparameter, parameter *n_estimators* (jumlah pohon) yang optimal adalah sebanyak 1.000 pohon dan parameter *max_features* (jumlah fitur maksimum) yang optimal adalah 4 variabel. Kemudian, dilakukan pelatihan kedua dengan mengimplementasikan model *Random Forest* regressor yang telah disesuaikan dengan parameter optimal hasil tuning hyperparameter tersebut, dan akhirnya mencapai tingkat akurasi sebesar 75,81% (Akhir et al., 2020).

Pada penelitian Vannisa Wanika dan Ika Elviana (2018) melakukan prediksi harga ponsel menggunakan metode *Random Forest*. Dalam penelitian ini, dilakukan sebuah proses klasifikasi menggunakan metode *Random Forest* untuk memprediksi harga ponsel berdasarkan spesifikasinya. Proses klasifikasi ini melibatkan tujuh variabel prediksi dan satu variabel respon, yang akhirnya menghasilkan tingkat akurasi sebesar 81%. Hasil dari penelitian ini juga menunjukkan bahwa tingkat akurasi tertinggi terdapat pada kategori harga ponsel yang tergolong murah (Vannisa Wanika & Ika Elviana 2018).

Siti Saadah dan Hanifa Salsabila melakukan penelitian Prediksi harga Bitcoin menggunakan metode *Random Forest*. Pada penelitian tersebut dengan menerapkan metode *Random Forest* hasilnya MAPE sebesar 1.50% dengan tingkat akurasi mencapai 98.50%. Hasil ini menunjukkan bahwa algoritma *Random Forest* adalah salah satu pendekatan yang mampu memberikan kinerja yang unggul dalam hal prediksi, terutama untuk data yang memiliki sifat acak (Saadah & Salsabila, 2021a).

Tabel 2. 1 Penelitian Terkait

No	Reference	Objek	Keterbaruhan Penelitian
1	Dian Pramesti dan Wiga Maulana Baihaqi(2023). Perbandingan Prediksi Jumlah Transaksi Ojek Online Menggunakan Regresi Linier dan <i>Random Forest</i>	Bagaimana mengukur tingkat akurasi prediksi jumlah transaksi ojek online menggunakan regresi Linier dan <i>Random Forest</i>	Berdasarkan perbandingan metrik, model regresi linier memberikan kinerja yang lebih baik dalam hal kesalahan prediksi (RMSE dan MSE) serta MAPE yang lebih rendah dibandingkan dengan model <i>Random Forest</i> yaitu nilai RMSE 1.6, MSE 2.6, dan MAPE 1.1. Sedangkan model <i>Random Forest</i> memperoleh nilai RMSE 921.19, MSE 948600.05, dan MAPE 0.91. Dalam melakukan prediksi tarif baru, tarif baru yang dihasilkan akan mengikuti data pada atribut yang kita masukkan pada model. Perbedaan dari penelitian yang dilakukan adalah peneliti menggunakan <i>Random Forest</i> regression dalam memprediksi nilai aset berbasis GIS
2	Syakirah Fachid dan Agung Triayudi (2022). Perbandingan Algoritma Regresi	Membandingkan tingkat akurasi algoritma Regresi Linier dan Regresi <i>Random Forest</i>	Hasil penelitian yang diperoleh dengan menggunakan model Regresi Linier menghasilkan tingkat akurasi sebesar 94% dengan

	Linier dan Regresi <i>Random Forest</i> Dalam Memprediksi Kasus Positif Covid-19	dalam memprediksi Kasus positif Covid- 19	RMSE sebesar 3031.127 dan MAPE sebesar 47.66. Sedangkan pada model <i>Random Forest</i> dengan menggunakan hyperparameter $\text{max_depth} = 10$ dan $\text{n_estimators} = 100$ dapat menghasilkan tingkat akurasi sebesar 97.7% dengan RMSE sebesar 1886.555 dan nilai MAPE sebesar 14.85. Perbedaan dari penelitian yang dilakukan adalah peneliti menggunakan <i>Random Forest</i> regression dalam memprediksi nilai aset berbasis GIS
3	Muhammad Reza Redo dan Artia Irianti (2021). Perbandingan performa Algoritma Neural Network, Regresi Linier, dan <i>Random Forest</i> dalam simulasi prediksi angka kematian pasien COVID-19 di Indonesia	Membandingkan tingkat akurasi algoritma Neural Network Regresi Linier dan Regresi <i>Random Forest</i> dalam memprediksi angka kematian pasien Covid-19	Hasil penelitian pada paper ini adalah algoritma yang paling mendekati nilai asli dalam melakukan prediksi dengan nilai RMS 94.991 +/- 0.000 selanjutnya diikuti oleh algoritma neural network dengan nilai RMS 285.956 +/- 0.000 simpangan terjauh terjadi pada algoritma linear regression yaitu positive 3285 menyimpang dari nilai sebenarnya. Perbedaan dari penelitian yang dilakukan adalah peneliti menggunakan <i>Random Forest</i> regression dalam memprediksi nilai aset berbasis GIS
4	Budi Prasajo, Emy Haryanti (2021). Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode <i>Random Forest</i>	Bagaimana mengukur tingkat akurasi prediksi kelayakan kredit pinjaman menggunakan metode <i>Random Forest</i>	Proses penelitian mengikuti langkah-langkah CRIPS-DM. Dalam tahap pelatihan, sebanyak 80% dari total 1000 data digunakan, sementara tahap pengujian melibatkan 20% sisanya yang dipilih secara acak. Hasil evaluasi dari algoritma <i>Random Forest</i> menunjukkan tingkat akurasi sebesar 0,83, atau

			83%, sehingga dapat diklasifikasikan sebagai model klasifikasi yang sangat baik. Perbedaan dari penelitian yang dilakukan adalah peneliti menggunakan <i>Random Forest</i> regression dalam memprediksi nilai aset berbasis GIS
5	Henri Tantyoko, Dian Kartika Sari, Andreas Rony Wijaya (2023). Prediksi Potensial Gempa Bumi Indonesia Menggunakan Metode <i>Random Forest</i> Dan Feature Selection	Bagaimana mengukur tingkat akurasi prediksi gempa bumi di Indonesia menggunakan metode <i>Random Forest</i> dan Feature Selection	Sebuah model yang baik dapat diidentifikasi dengan tingkat kesalahan yang sekecil mungkin. Oleh karena itu, penulis melakukan pemilihan fitur untuk memproses hanya fitur-fitur yang memiliki korelasi yang kuat. Ketika melakukan prediksi dengan menggunakan <i>Random Forest</i> (RF) dengan pemilihan fitur, ditemukan bahwa F1 score mencapai 92,23%, meningkat sebesar 5,02% dibandingkan dengan prediksi tanpa pemilihan fitur. Pendekatan RF dengan Seleksi Fitur juga unggul jauh jika dibandingkan dengan metode machine learning konvensional seperti SVM, Naïve Bayes, dan <i>Decision Tree</i> . Perbedaan dari penelitian yang dilakukan adalah peneliti menggunakan <i>Random Forest</i> regression dalam memprediksi nilai aset berbasis GIS
6	Bagiya Wahyudi, Ina Kuswandi (2022). Prediksi Peringkat Aplikasi di Google Play Menggunakan	Bagaimana mengukur tingkat akurasi prediksi peringkat aplikasi menggunakan metode <i>Random Forest</i>	Prediksi rating aplikasi akan dilakukan dengan menggunakan algoritma <i>Random Forest</i> sebagai metode prediksi yang dipilih. Dalam penelitian ini, dataset yang digunakan berasal dari

	Metode <i>Random Forest</i>		Google Play Store, yang terdiri dari 10,840 baris data dan 13 atribut. Hasil penelitian menunjukkan bahwa penggunaan algoritma <i>Random Forest</i> menghasilkan tingkat akurasi rata-rata sekitar 93,8%. Perbedaan dari penelitian yang dilakukan adalah peneliti menggunakan <i>Random Forest</i> regression dalam memprediksi nilai aset berbasis GIS
--	-----------------------------	--	--

2.2 Kantor Jasa penilai Publik

KJPP adalah singkatan dari "Kantor Jasa Penilai Publik." Ini adalah sebuah lembaga profesional yang berfungsi sebagai penilai independen untuk menentukan nilai properti atau aset yang akan digunakan dalam berbagai transaksi, seperti penjualan, hipotek, perhitungan pajak, atau tujuan lainnya (Dahnir et al., 2020). KJPP memiliki peran penting dalam memastikan bahwa penilaian properti atau aset dilakukan dengan objektif dan adil. Berikut Beberapa tugas umum KJPP diantaranya:

1. Penilaian Properti: KJPP melakukan penilaian properti secara profesional untuk menentukan nilai pasar yang obyektif. Hal ini dapat mencakup properti komersial, perumahan, atau jenis properti lainnya.
2. Penilaian Aset: Selain properti, KJPP juga dapat menilai aset lainnya seperti perusahaan, peralatan, atau barang berharga lainnya.
3. Laporan Penilaian: Setelah melakukan penilaian, KJPP menyusun laporan penilaian yang berisi informasi rinci tentang metode yang digunakan, data yang

digunakan, dan nilai yang ditetapkan. Laporan ini sering digunakan dalam berbagai transaksi hukum.

4. Konsultasi: KJPP juga dapat memberikan saran dan konsultasi kepada klien mereka mengenai nilai aset atau properti dan faktor-faktor yang memengaruhi nilai tersebut.

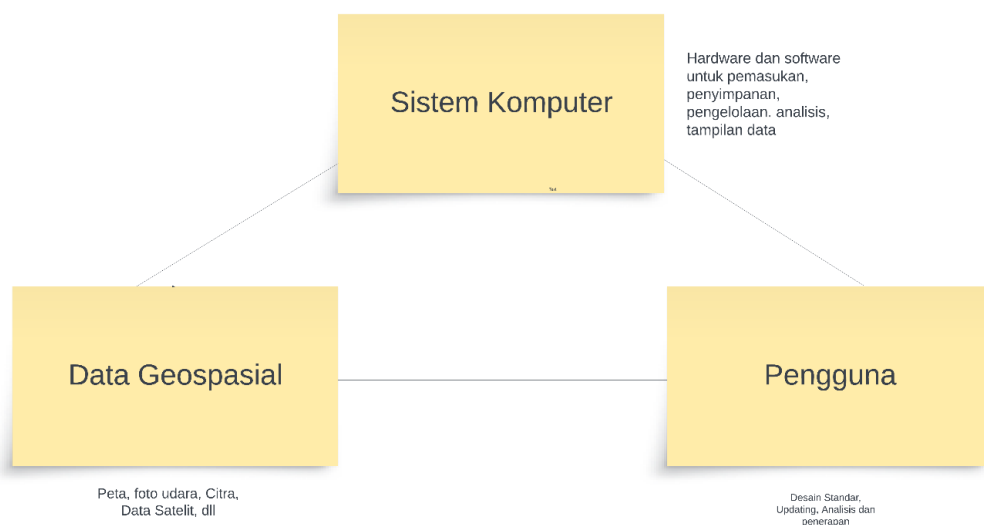
KJPP biasanya diatur oleh pemerintah atau badan pengawas keuangan di negara mereka untuk memastikan profesionalisme dan integritas dalam melakukan penilaian. Mereka harus mengikuti standar etika dan praktik penilaian yang ketat agar dapat menjaga kualitas dan keandalan penilaian yang mereka hasilkan.

KJPP Sugeng Irawan Kota Malang merupakan kantor jasa penilai publik yang ada di Kota Malang tepatnya di Komplek Ruko Taman Borobudur Indah Kav.23, Jl. Puncak Borobudur Malang. Selama ini Kantor KJPP telah mengikuti perkembangan seiring dengan perkembangan ekonomi dan industri dalam cakupan internasional. Beberapa tren perkembangan utama dalam dunia kjpp diantaranya adalah globalisasi, diversifikasi layanan, peningkatan regulasi, peningkatan teknologi dan automasi. Pada praktiknya KJPP irawan masih terdapat kendala dalam perkembangan di bidang teknologi, dikarenakan KJPP belum mengimplementasi GIS sebagai pengelolaan data. KJPP irawan masih mengalami kesulitan dalam pengolaan data penilaian asset dimana data hanya tersimpan dalam media excel. Seiring berjalan waktu data yang tersedia akan semakin bertambah sehingga data akan semakin sulit untuk dikelola dan dimanage pada saat adanya evaluasi data. Oleh karena itu, penulis berinisiatif merancang Sistem Informasi Geografi dalam pengelolaan data dan juga memberikan fitur prediksi penilaiaan

harga asset dengan menggunakan beberapa atribut yang dapat diolah pada data yang tersedia.

2.3 Sistem Informasi Geografi (SIG)

Sistem Informasi Geografi (GIS) adalah suatu alat bantu manajemen dalam bentuk informasi berbantuan komputer yang erat kaitannya dengan sistem pemetaan dan analisis terhadap berbagai hal dan peristiwa yang terjadi di permukaan bumi. Teknologi GIS menggabungkan operasi pengolahan data berbasis database yang umumnya digunakan saat ini, seperti pengumpulan data berdasarkan kebutuhan dan analisis statistik, dengan memanfaatkan visualisasi khas serta berbagai keunggulan yang ditawarkan melalui analisis geografis dengan gambar-gambar peta (Eka Dwi Cahyono, 2001). Komponen-komponen yang membentuk GIS diantaranya sistem komputer, pengguna, data geospasial seperti pada gambar 2.1



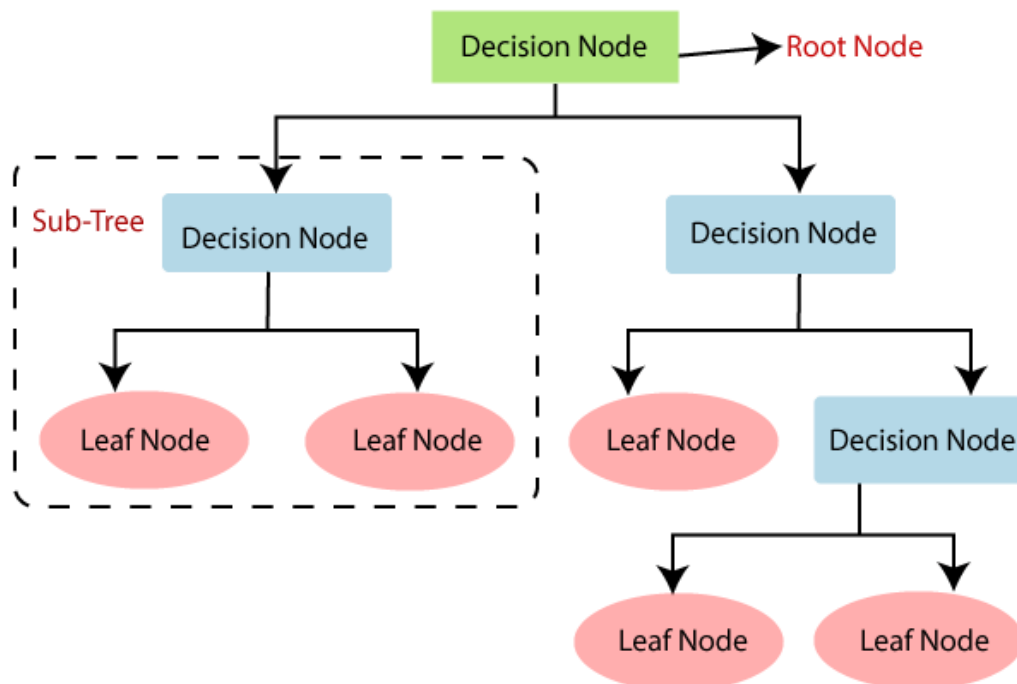
Gambar 2. 1 Komponen GIS

Data yang dikelola dalam GIS terbagi menjadi dua jenis, yaitu data geospasial, sering disebut sebagai data spasial, dan data non-spasial yang disebut atribut. Pada ilustrasi di atas, data atribut tidak diperlihatkan karena fokus utama dalam GIS adalah representasi visual dari data spasial. Namun, dalam beberapa situasi, GIS juga dapat mempergunakan data atribut, baik dengan cara langsung maupun tidak langsung.

2.4 Decision Tree (Pohon Keputusan)

Decision tree atau dalam bahasa Indonesia sering disebut pohon keputusan, adalah salah satu teknik dalam ilmu data mining dan machine learning yang digunakan untuk pengambilan keputusan. Ini adalah model prediksi yang digunakan dalam berbagai bidang, termasuk pengklasifikasi dan penghasil aturan (Song & Lu, 2015).

Decision tree bekerja dengan mengorganisir data menjadi struktur pohon, di mana setiap *node* dalam pohon mewakili sebuah keputusan atau pilihan, dan cabang-cabang pohon mewakili hasil dari keputusan tersebut. Pada setiap *node*, algoritma *decision tree* melakukan pemilihan fitur (atribut) dari data yang akan digunakan untuk memisahkan data menjadi dua atau lebih kelompok yang lebih homogen berdasarkan nilai atribut tersebut. Proses ini berlanjut hingga mencapai daun (*leaf*) pohon, yang mewakili prediksi atau keputusan akhir (Charbuty & Abdulazeez, 2021).

Gambar 2. 2 *Decision Tree*

Menurut gambar 3.2, pada awalnya ada simpul akar yang menggambarkan seluruh populasi atau sampel yang kemudian dibagi menjadi dua atau lebih kelompok. Proses pembagian ini disebut *splitting*. Ketika sebuah *sub-node* dibagi menjadi *sub-node* lebih lanjut, ini disebut simpul keputusan. Setiap sub-bagian dari pohon secara kolektif disebut cabang.

Decision Tree merupakan suatu diagram alir yang menyerupai struktur pohon, dimana setiap node internal mengindikasikan pengujian terhadap suatu atribut. Setiap cabang menggambarkan hasil dari pengujian tersebut, sementara node daun menyatakan kelas-kelas atau distribusi kelas. *Node* akar (*Root node*) memiliki beberapa cabang keluar tetapi tidak memiliki cabang masuk. *Node* internal memiliki satu cabang masuk dan beberapa cabang keluar, sedangkan *node*

daun hanya memiliki satu cabang masuk tanpa cabang keluar. *Node* daun mewakili hasil akhir yang mencerminkan label kelas dari kombinasi atribut yang membentuk suatu aturan. (Patel & Prajapati, 2018).

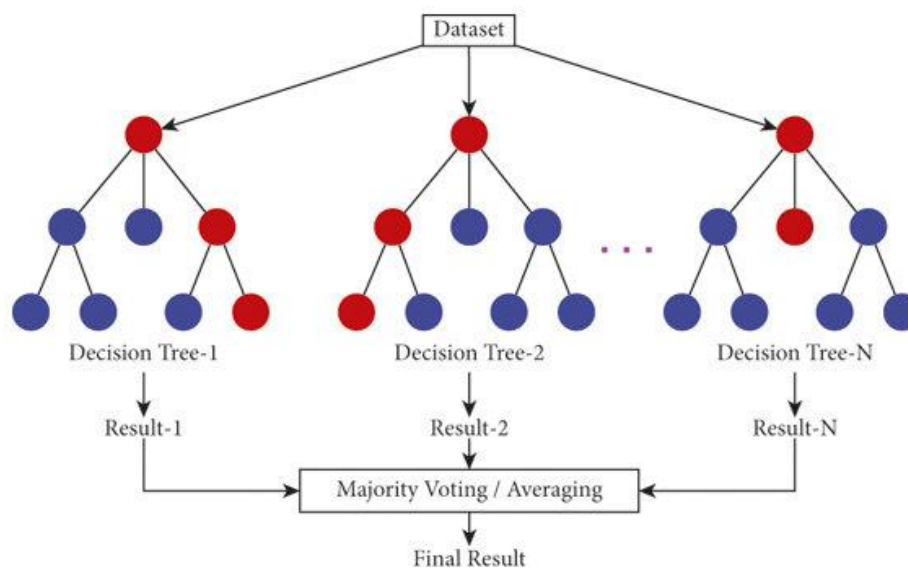
2.5 Metode *Random Forest*

Random Forest adalah inovasi yang dikenalkan pertama kali pada tahun 2001 oleh Leo Breiman melalui karyanya yang berjudul "*Random Forest*". Metode *Random Forest* merupakan evolusi dari *Decision Tree*, yang menggabungkan beberapa pohon prediksi. Setiap pohon dalam *ensemble* ini bergantung pada nilai-nilai vektor sampel yang acak dan independen, serta distribusi yang sama untuk semua pohon yang ada dalam kelompok hutan tersebut (Santra & Christy, 2012). Metode pengajaran *decision tree* merupakan salah satu teknik klasifikasi yang sangat diminati. Tingkat akurasi yang diberikannya sebanding dengan teknik klasifikasi lainnya, dan juga memiliki efisiensi yang tinggi. Model klasifikasi yang dihasilkan melalui teknik ini direpresentasikan sebagai pohon dan dikenal dengan istilah *Random Forest*. ID3 oleh Quinlan (1986), C4.5 oleh Quinlan (1993), dan CART oleh Breiman, Friedman, Stone, dan Olshen (1984) adalah beberapa algoritma pembelajaran pohon keputusan yang terkenal (Pratomo et al., 2019).

Random Forest termasuk dalam jenis algoritma yang mengadopsi pembelajaran ansemble. Dalam konteks *Random Forest*, ansemble ini terdiri dari sejumlah pohon yang beragam dalam nilai-nilainya, dan seluruh pohon tersebut digabungkan menjadi sebuah kesatuan. Algoritma ini merujuk pada *decision tree* sebagai fondasi utama dari ansemble tersebut. Model *Random Forest* sering digunakan untuk tugas klasifikasi, seperti prediksi, karena kemampuannya yang

lebih efisien terutama ketika menghadapi data yang besar. Model ini juga efektif baik pada data diskrit maupun kontinu, dan algoritma klasifikasinya dikenal mampu mencapai akurasi yang tinggi tanpa memerlukan penyesuaian parameter yang kompleks. *Random Forest* juga efektif dalam mengatasi data yang memiliki missing value, mampu menyederhanakan keputusan yang kompleks, dan membangun lebih dari satu pohon keputusan, sehingga memungkinkan untuk meningkatkan hasil prediksi (Akhir et al., 2020).

Teknik dasar dalam *Random Forest* adalah menggunakan *decision tree*. Pada *decision tree*, data dimasukkan ke *root node*, lalu diarahkan ke *leaf node* untuk mengidentifikasi prediksi kelas. *Random Forest* merupakan salah satu metode klasifikasi yang terdiri dari kumpulan pohon terstruktur. Setiap pohon memberikan suara (*vote*) untuk menentukan kelas dengan suara terbanyak. Dengan kata lain, *Random Forest* terdiri dari sejumlah *decision tree* yang tidak memiliki korelasi dan digunakan untuk mengklasifikasikan data ke dalam kelas tertentu. Ide dasar dari *Random Forest* adalah membentuk hutan dengan banyak pohon yang dihasilkan. Algoritma *Random Forest* berbeda dari algoritma *decision tree* karena tidak melibatkan pemangkasan variabel. Metode ini menggabungkan banyak pohon untuk membuat klasifikasi dan prediksi kelas. Dalam *Random Forest*, pembuatan pohon dilakukan melalui pelatihan dengan sampel data. Variabel yang digunakan untuk pemisahan terbaik ditentukan secara acak dengan mengumpulkan suara dari setiap pohon. Akhirnya, kelas dengan suara terbanyak menjadi kelas yang dipilih sebagai keputusan akhir (Saadah & Salsabila, 2021).



Gambar 2. 2 *Random Forest*
 Sumber: levelup.gitconnected.com

Berdasarkan skema gambar diatas, Berikut Langkah – Langkah *Random Forest* :

1. Tahap pertama pembentukan Subset:
 - a. Untuk setiap sampel yang dipilih secara acak, pohon keputusan dibentuk dengan membagi data ke dalam subset yang semakin kecil.
 - b. Ulangi Langkah a sebanyak k kali sehingga diperoleh k buah pohon acak;
2. Setelah membentuk subset, dilakukan pembagian (splitting) dengan menerapkan algoritma terbaik untuk memisahkan simpul. Proses prediksi dilakukan dengan menggabungkan semua hasil prediksi dari K pohon. Pada kasus klasifikasi, dilakukan pemilihan mayoritas suara, sedangkan pada kasus regresi, dilakukan perhitungan rata-rata.

2.6 Mean Absolute Percentage Error (MAPE)

MAPE (Mean Absolute Percentage Error) adalah salah satu metrik evaluasi yang digunakan untuk mengukur tingkat akurasi dari suatu model peramalan atau

prediksi. MAPE mengukur sejauh mana nilai-nilai prediksi mendekati nilai-nilai aktual dalam bentuk persentase. Semakin kecil nilai MAPE, semakin baik model prediksi tersebut (de Myttenaere et al., 2016). Berikut Rumus dari MAPE:

$$\frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - f_i}{A_i} \right| \times 100 \quad (2.1)$$

n adalah jumlah pengamatan atau observasi

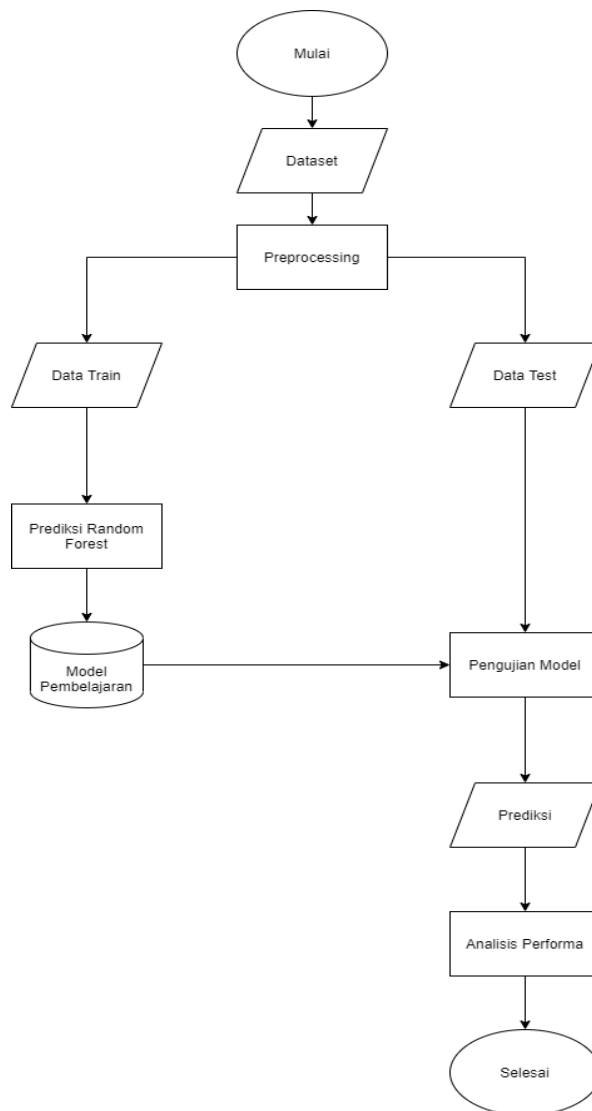
A_i adalah nilai aktual dari observasi ke- i

F_i adalah nilai Prediksi dari observasi ke- i

BAB III

METODOLOGI PENELITIAN

Pada bab ini akan dibahas tentang beberapa hal, yaitu alur atau tahapan yang akan dilakukan pada penelitian. Tahapan penelitian terdiri dari Identifikasi Masalah, Studi Literatur, pengumpulan data, perancangan sistem, pengujian dan kesimpulan atau menggunakan sistem. Seperti terdapat pada gambar 3.1.



Gambar 3. 1 Desain Sistem

3.1 Pengumpulan data

Informasi yang dibutuhkan dalam penelitian ini adalah informasi mengenai data asset yang sudah dinilai oleh Kantor KJPP. Dalam penelitian ini pengumpulan data dilakukan dengan cara wawancara dengan pengurus atau pihak terkait dan mengumpulkan data nilai asset dari tahun 2021 sampai dengan tahun 2023. Data yang terkumpul sebanyak 577 data dengan 5 atribut atau karakteristik yang digunakan. diantaranya sebagai berikut.

Tabel 3. 1 Data Penelitian

No	Atribut	Deskripsi
1	Luas tanah (lt)	Pada kriteria ini diisi dengan nominal luas tanah asset yang akan dinilai per meter Contoh 200
2	Luas Bangunan	pada kriteria ini diisi dengan nominal luas bangunan asset yang akan dinilai per meter Contoh 300
3	Indikasi nilai tanah	Pada kriteria ini diisi dengan nominal indikasi nilai tanah per meter Contoh 800.000
4	Koordinat asset	Pada kriteria ini diisi dengan nominal Koordinat asset Contoh -7.651581, 111.556163
5	Kesimpulan Nilai	Pada kriteria ini diisi dengan nominal Kesimpulan dari harga atau nilai asset Contoh 561.000.000

Tahapan pengumpulan data ini merupakan tahapan mempersiapkan data-data yang diperlukan sebelum diolah. Pada tahapan ini terdapat 6 atribut atau kriteria yang digunakan diantaranya 4 variabel input dan 1 variabel output. Semua atribut inputan tersebut disimbolkan dengan simbol K1 untuk kriteria pertama dan seterusnya, untuk atribut output disimbolkan dengan O. Pada tahap ini juga

mempersiapkan data yang akan digunakan dalam penelitian ini dengan format .xlsx.

berikut contoh data yang digunakan.

Tabel 3. 2 Contoh Data

NO	LT (Meter persegi)	LB (Meter persegi)	Indikikasi nilai tanah (Juta)	Koordinat Aset (Derajat)	Kesimpulan Nilai (juta)
1	210	259	3.700.000	'-7.472638, 112.724235	1967
2	178	180	3.700.000	'-7.471227, 112.725659	1304
3	78	62	2.300.000	'-7.894553, 112.607671	358
4	119	78	5.300.000	'-7.920552, 112.618714	953
5	192	168	3.000.000	'-7.918626, 112.644697	1002
6	137	45	1.350.000	'-8.189372, 113.678535	253
7	150	68	1.500.000	'8.190835,113.714027	475

3.2 Perancangan Sistem

Pada tahapan perancangan Sistem akan dibahas beberapa hal, yaitu alur dari program yang akan dilakukan pada penelitian. Sistem dimulai dari tahapan input data penerima, preprocessing, proses neural network backpropagation, dan diakhiri dengan hasil prediksi.

3.2.1 Preprocessing

Preprocessing adalah tahapan yang dilakukan untuk memeriksa data yang akan digunakan dengan tujuan untuk memaksimalkan hasil. Peneliti menggunakan 2 tahapan yakni data *cleaning* dan data *exploration*.

3.2.1.1 Data Cleaning

Data *cleaning* dilakukan untuk memperbaiki dan memodifikasi data untuk membuatnya siap digunakan dalam analisis, seperti menghapus duplikat, menghapus data tidak relevan, dan mengisi data yang hilang. Terdapat beberapa kolom yang tidak diperlukan seperti halnya kolom nomor laporan, tanggal penilaian, tanggal laporan, dll. Kolom-kolom tersebut akan dihapuskan dikarenakan kolom-kolom tersebut tidak berpengaruh dalam proses prediksi nilai.

3.2.1.2 Data Exploration

Memahami dan menjelajahi data melalui visualisasi dan analisis statistik, seperti memahami jenis data, variabel, dan sumber data, memeriksa data untuk menentukan apakah ada data yang hilang, duplikat, atau tidak sesuai, memvisualisasikan data, menghitung statistik seperti rata-rata dan standar deviasi, dan melakukan analisis univariat, bivariat, atau multivariat.

3.2.2 Data Splitting

Dalam artikel yang dipublikasikan oleh Liu, Q., Yan, S., & Zhou, Y. pada tahun 2012 berjudul "Ratio of training to testing samples and classifier learning" di jurnal *Pattern Recognition Letters*, dibahas mengenai dampak rasio sampel data training dan data testing terhadap kinerja klasifikasi dengan menggunakan berbagai algoritma, termasuk Neural Network. Dalam jurnal tersebut, disimpulkan bahwa terdapat pengaruh yang signifikan dari rasio data training terhadap akurasi hasil model.

Buku *Hands on Machine Learning with Tensorflow and Keras* (2021) Oleh Neurel Geron, dijelaskan bahwa tujuan utama dari penggunaan data testing adalah untuk mengoptimalkan kinerja model pada data yang belum pernah dikenal oleh model sebelumnya. Oleh karena itu, disarankan untuk membuat set data testing. Cara untuk membuat data testing, seperti dijelaskan dalam buku tersebut, adalah dengan memilih beberapa contoh secara acak, umumnya sekitar 20% dari seluruh kumpulan data (atau lebih sedikit jika ukuran kumpulan data sangat besar), dan mengisolasi contoh-contoh tersebut.

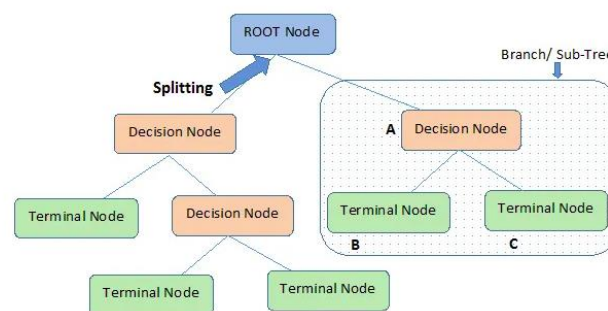
Pada penelitian ini, pemisahan antara data training dan data testing mengikuti pendekatan yang telah dijelaskan dalam tinjauan pustaka, yaitu menggunakan metode random state. Pada tahap ini, data yang telah melalui proses awal diacak secara acak dengan memanfaatkan nilai random state yang telah ditetapkan sebelumnya. Penggunaan nilai random state yang sama bertujuan untuk menjamin konsistensi dan reproduktibilitas pemisahan data secara acak pada setiap eksekusi. Setelah proses pengacakan, data yang telah diacak akan dibagi menjadi dua subset, yakni data training dan data testing, sesuai dengan perbandingan yang telah ditetapkan sebelumnya. Sebagai contoh, jika perbandingan yang diinginkan adalah 80% data training dan 20% data testing, pembagian data akan dilakukan dengan proporsi tersebut.

3.2.3 Pengujian *Random Forest*

Pohon Keputusan (Decision Tree) merupakan salah satu model prediksi yang menggunakan struktur pohon atau hierarki. Dalam struktur ini, setiap simpul internal mengindikasikan uji coba pada suatu atribut, setiap cabang menunjukkan

hasil dari uji coba tersebut, dan setiap simpul daun mewakili label kelas. Titik awal dari pohon keputusan ini disebut sebagai akar pohon (Wu et al., 2008).

Random Forest, di sisi lain, adalah suatu teknik ensemble learning yang menggunakan banyak *decision tree* untuk meningkatkan kinerja prediktif dan mengurangi overfitting. *Random Forest* secara acak memilih subset dari data pelatihan dan atribut untuk setiap pohon keputusan yang dibuat, sehingga setiap pohon memiliki variasi yang sedikit berbeda. Setelah itu, prediksi dari setiap pohon diambil, dan hasilnya diagregasi untuk menghasilkan prediksi akhir. Untuk memahami bagaimana *Random Forest* bekerja, kita dapat memahami bagaimana konsep perhitungan manual dari *decision tree*. Berikut adalah ilustrasi pohon keputusan pada *Random Forest*.



Gambar 3. 2 *Decision Tree*
Sumber: www.analyticsvidhya.com

Classification and regression tree (CART) adalah istilah yang digunakan untuk menggambarkan algoritma pohon keputusan (*decision tree*) yang digunakan untuk klasifikasi dan regresi dalam machine learning. CART akan menghasilkan pohon klasifikasi jika variabel respon mempunyai skala kategorik dan akan

menghasilkan pohon regresi jika variabel respon berupa data kontinu. Tujuan utama CART adalah untuk mendapatkan suatu kelompok data yang akurat sebagai penciri dari suatu pengklasifikasian (Sumartini & Purnami, 2015). Untuk memahami bagaimana pohon regresi pada CART tersebut dibuat, kita dapat memahaminya pada konsep *decision tree* dengan simulasi membangun *decision tree* sederhana dengan data rumah sederhana sebagai berikut:

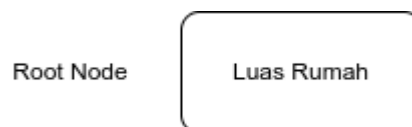
Tabel 3. 3 Contoh Perhitungan Data

Data	Data Feature		Data Target
	Luas Rumah	Jumlah Kamar	Harga (dalam juta)
Rumah 1	800	2	500
Rumah 2	1000	4	600
Rumah 3	2000	5	1000
Rumah 4	2000	7	1300

Tabel di atas adalah data rumah dengan dua variabel sederhana, yaitu luas rumah dan jumlah kamar sebagai data feature, dan harga rumah sebagai data target. Berdasarkan gambar *decision tree* dan tabel di atas, kita dapat menguji skenario dengan langkah langkah berikut.

1. Memilih atribut *ROOT NODE*

Atribut luas rumah akan digunakan untuk pemisahan awal sebagai inisiasi awal pemisahan data.



2. Menghitung kemungkinan Splitting menggunakan nilai tengah

$$\text{Nilai tengah 1} = (\text{Luas Rumah 1} + \text{Luas Rumah 2}) / 2 = 900$$

$$\text{Nilai tengah 2} = (\text{Luas Rumah 2} + \text{Luas Rumah 3}) / 2 = 1500$$

Pada luas rumah 3,4,5 tidak terdapat nilai tengah

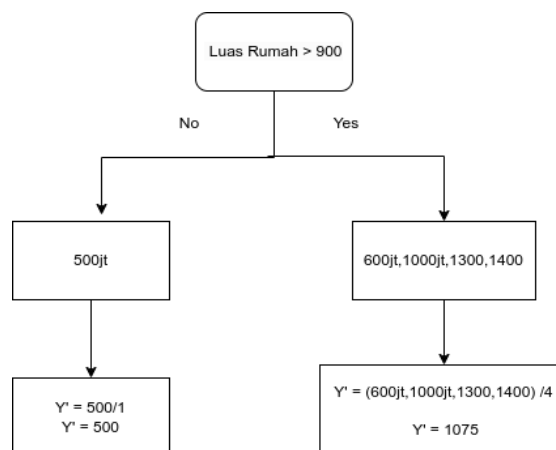
3. Menentukan Akar / Split

Cara menentukan akar adalah dengan mencari calon split dengan jumlah error terendah. Contoh perhitungan manual dibawah adalah bagaimana kita menghitung error calon cabang. Berikut adalah tabel kemungkinan nilai tengah dan errornya.

Tabel 3. 4 Random Subset 1

Split		Data Feature		Data Target	Data Prediksi	Error
Luas Rumah > 900	Data	Luas Rumah	Jumlah Kamar	Harga (dalam juta)	Harga Prediksi (dalam juta)	Absolute Error
FALSE	Rumah 1	800	2	500	800	300
TRUE	Rumah 2	1000	4	600	1075	475
	Rumah 3	2000	5	1000	1075	75
	Rumah 4	2000	7	1300	1075	225

Berikut adalah gambar pohon split 1 untuk menghitung nilai error pada akar pemisahan Luas Rumah > 900



Gambar 3. 3 Kemungkinan Pohon Split 1

Y adalah nilai asli dan Y' adalah hasil prediksi dari *node* masing-masing, dan nilai error akan dihitung menggunakan jumlah dari $(Y - Y')$

$$\text{Error} = \sqrt{(500-500)^2} + \sqrt{(1075-600)^2} + \sqrt{(1075-1000)^2} + \sqrt{(1075-1300)^2} + \sqrt{(1075-1400)^2}$$

$$0 + 475 + 75 + 225 + 325 = 1000$$

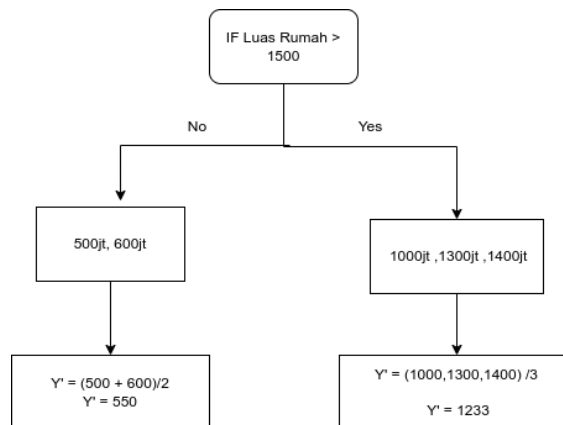
Maka, jumlah error dari split 900 adalah 1000.

Selanjutnya, akan dihitung jumlah error dari calon akar 2. Berikut adalah tabel split 2 nilai tengah dan errornya.

Tabel 3. 5 Random Subset 2

Split		Data Feature		Data Target	Data Prediksi	Error
Luas Rumah > 1500	Data	Luas Rumah	Jumlah Kamar	Harga (dalam juta)	Harga Prediksi (dalam juta)	Absolute Error
FALSE	Rumah 1	800	2	500	550	50
	Rumah 2	1000	4	600	550	50
TRUE	Rumah 3	2000	5	1000	1233	233
	Rumah 4	2000	7	1300	1233	67

Berikut adalah gambar pohon split 1 untuk menghitung nilai error pada akar pemisahan Luas Rumah > 900



Gambar 3. 4 Random Pohon Split 2

Y adalah nilai asli dan Y' adalah hasil prediksi dari *node* masing-masing, dan nilai error akan dihitung menggunakan jumlah dari $(Y - Y')$

$$\text{Error} = \sqrt{(550-500)^2} + \sqrt{(550-600)^2} + \sqrt{(1233-1000)^2} + \sqrt{(1233-1300)^2} + \sqrt{(1233-1400)^2}$$

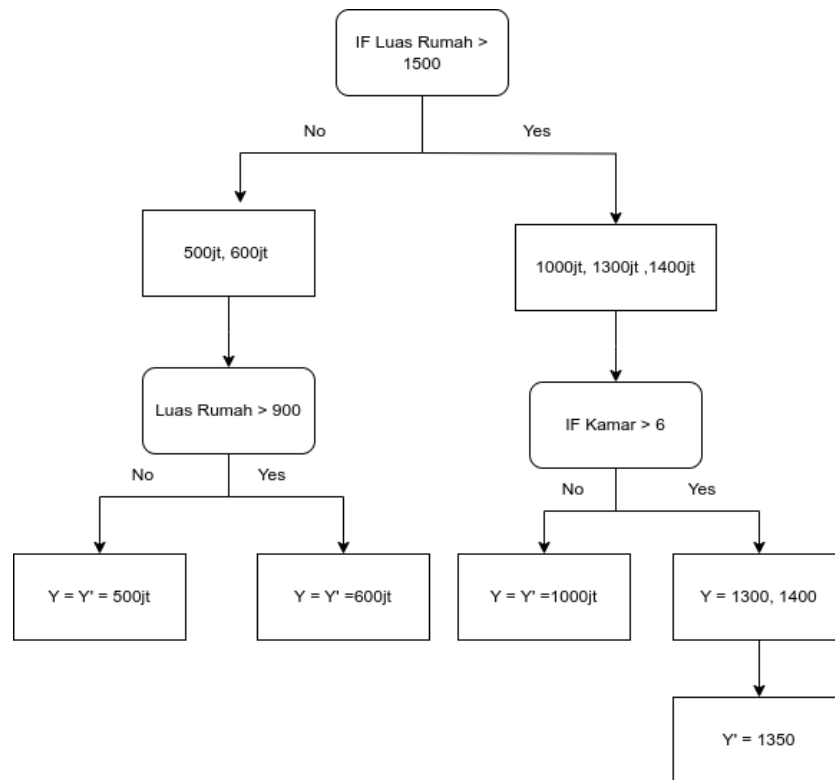
$$50 + 50 + 233 + 67 + 167 = 567$$

Maka, jumlah error dari split 1500 adalah 567

Karena error terendah ada pada calon akar 2, maka akar yang dipilih adalah calon akar dua.

4. Ulangi langkah 1 hingga langkah 4

Ulangi langkah diatas hingga tidak tree dapat mengklasifikasi harga rumah secara maksimal, sehingga didapatkan *Decision Tree Regressor* sebagai berikut:

Gambar 3. 5 Hasil *Random forest*

Hingga Mean Squared Error dari *Decision Tree* tersebut adalah:

$$\text{Error} = \sqrt{(500-500)^2} + \sqrt{(600-600)^2} + \sqrt{(1000-1000)^2} + \sqrt{(1350-1300)^2} + \sqrt{(1350-1400)^2}$$

$$0 + 0 + 0 + 50 + 50 = 100$$

Sehingga, jumlah error adalah 100,

Lalu didapatkan Rata rata error sebagai berikut:

$$MSE = \text{Error} / \text{Jumlah data yang di prediksi}$$

$$MSE = 100 / 5 = 20\text{jt}$$

Dapat disimpulkan bahwa setiap prediksi yang dilakukan *decision tree* di atas, terdapat error 20 jt. Dengan harga rumah yang diprediksi memiliki range harga 500 jt, 600jt, 1000jt, 1300jt dan 1400jt, error 20 jt termasuk error yang sangat kecil

Disimpulkan bahwa *Random Forest* regression dapat digunakan untuk memprediksi nilai harga rumah dengan data features atau atribut atribut yang dibutuhkan.

3.2.4 Analisa Peforma

Dalam Penelitian ini tahap analisa peforma menggunakan *Mean Absolute Percentage Error* (MAPE) sebagai metode pengujian yang bertujuan untuk menentukan akurasi dari sebuah model. Model yang diambil pada akhir merupakan model yang memiliki nilai MAPE tertinggi dimana hal tersebut dipengaruhi oleh beberapa parameter dalam model *Random Forest* diantaranya adalah max kedalaman, max sample split, max feature, dan jumlah pohon pada tree. MAPE dapat dirumuskan sebagai berikut.

$$\frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - f_i}{A_i} \right| \times 100 \quad (3.1)$$

n adalah jumlah pengamatan atau observasi

A_i adalah nilai aktual dari observasi ke-i

F_i adalah nilai Prediksi dari observasi ke-i

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil *Datasets*

Dataset ini mencakup informasi tentang penilaian harga aset, khususnya rumah, di beberapa lokasi di Kota Malang, Banyuwangi, Surabaya, Gresik, dan Blitar. Data ini digunakan untuk penelitian yang bertujuan memprediksi harga aset menggunakan metode *Random Forest Regression*. Atribut-atribut yang disertakan dalam dataset ini melibatkan informasi tentang aset, luas tanah (LT), luas bangunan (LB), jumlah lantai, jumlah kamar mandi, jumlah kamar, dan indikasi nilai per meter. Berikut hasil data yang telah berhasil didapatkan.

Tabel 4. 1 Atribut Dataset

Data Atribut	Deskripsi
id	Nomor identifikasi unik untuk setiap entri.
Bulan	Bulan ketika penilaian dilakukan.
No	Nomor urut dari penilaian.
No_laporan	Nomor laporan penilaian.
Tanggal_laporan	Tanggal laporan dibuat.
Tanggal_penilaian	Tanggal penilaian aset.
NO_SP	Nomor Surat Perintah.
Tanggal_SP	Tanggal Surat Perintah dikeluarkan.
Nama_debitur	Nama pemilik aset/debitur.
Nama_Pemberi_Tugas	Nama pemberi tugas untuk penilaian.
Tujuan	Tujuan dari penilaian (dalam hal ini, lelang).
Obyek	Jenis obyek penilaian (misalnya, Rumah Tinggal atau Tanah Kosong).
Lokasi	Lokasi lengkap aset.
Kota/Kabupaten	Kota atau kabupaten tempat aset berada.
LT	Luas tanah.
LB	Luas bangunan.
CP	Kontak person.
Supervisi	Informasi supervisi.
Penilai	Nama penilai.
Pelaksana	Nama pelaksana penilaian.
Fee_total	Total biaya penilaian.

dpp	Dasar pengenaan pajak.
ppn	Pajak pertambahan nilai.
Indikasi_nilai	Indikasi nilai aset.
Kesimpulan_nilai	Kesimpulan nilai dari penilaian.
Report	Jenis laporan penilaian (misalnya, Long Report).
Titik_koordinat.	Koordinat geografis lokasi aset.Koordinat geografis lokasi aset.Koordinat geografis lokasi aset.
Nama.	Nama Pemilik aset/debitur.
Jumlah_kamar	Jumlah kamar di rumah.
Jumlah_kamarmandi	Jumlah kamar mandi di rumah.
Jumlah_lantai	Jumlah lantai rumah.
Pusat_kota	Indikasi apakah properti berada di pinggir kota atau pusat kota.

Terdapat beberapa variabel yang digunakan dalam memprediksi harga nilai asset diantaranya adalah Luas Tanah, Luas Bangunan, Indikasi Nilai, Jumlah Lantai, Jumlah Kamar Mandi, Jumlah Kamar, dan Kesimpulan Nilai. Total data yang digunakan setelah proses preprocessing sebesar 100 data ini masih bisa ditambahkan bila hasil prediksi belum mencapai nilai yang baik.

4.2 Hasil Preprocessing

Pada tahap awal pengolahan data, hasil preprocessing menjadi kunci utama untuk memastikan keakuratan dan keterandalan analisis data. Sub bab Cleaning Data dan Data Exploration bertujuan untuk membersihkan dan menjelajahi dataset sebelum digunakan dalam model dan analisis lebih lanjut. Model Machine learning yang hanya dapat menerima angka, mengakibatkan preprocessing menjadi satu proses yang wajib dan dapat menentukan tingginya tingkat akurasi Cleaning data melibatkan identifikasi dan penanganan nilai yang hilang, outlier, atau kesalahan lainnya, sementara Data Exploration membawa kita pada perjalanan untuk memahami struktur dan karakteristik dataset dengan lebih mendalam.

4.2.1 Hasil Data Cleaning

Proses cleaning data yang dilakukan melibatkan penggantian nilai null pada kolom "Indikasi Nilai Rumah" dengan nilai rata-rata yang ada di kolom tersebut. Hal ini bertujuan untuk mengisi nilai yang hilang (null) dengan nilai tengah dari data yang telah ada dalam kolom "Indikasi Nilai Rumah", sehingga memastikan data tersebut tetap konsisten dan dapat digunakan secara optimal dalam analisis atau pemodelan lebih lanjut.

Tabel 4. 2 Penggantian Null Value kolom Indikasi Nilai

No.	SP.	Indikasi Nilai Awal	Indikasi Nilai Setelah Preprocessing
1	00195/2.0045-02/PI/07/PS.0205/1/VII/2021	-	2996214
2	00215/2.0045-02/PI/07/PS.0205/1/VIII/2021	-	2996214
3	00216/2.0045-02/PI/07/PS.0205/1/VIII/2021	-	2996214
4	00232/2.0045-02/PI/07/PS.0205/1/IX/2021	-	2996214
5	00297/2.0045-02/PI/07/PS.0205/1/X/2021	-	2996214
6	00429/2.0045-02/PI/07/PS.0205/1/XII/2021	-	2996214

Identifikasi dan jelaskan karakteristik utama dari data, seperti adanya nilai yang hilang, outlier, atau distribusi yang tidak biasa.

4.2.2 Hasil Data Exploration

Tabel korelasi adalah suatu tabel yang menampilkan nilai korelasi antara dua atau lebih variabel dalam bentuk matriks. Nilai korelasi mengukur sejauh mana dua variabel bergerak bersama-sama atau saling terkait dalam suatu dataset. Nilai korelasi berkisar antara -1 hingga 1, di mana: 1 menunjukkan korelasi positif

sempurna (kedua variabel bergerak searah dengan kekuatan penuh), -1 menunjukkan korelasi negatif sempurna (kedua variabel bergerak berlawanan dengan kekuatan penuh), 0 menunjukkan tidak ada korelasi linear.

Tabel korelasi biasanya digunakan dalam analisis statistik dan eksplorasi data untuk mengidentifikasi hubungan antara variabel-variabel. Berikut adalah tabel korelasi yang dihasilkan.

```

LT          0.030373
LB          0.814454
Indikasi_nilai  0.543462
Jumlah_kamar  0.290673
Jumlah_kamarmandi 0.308836
Jumlah_lantai  0.389262
Kesimpulan_nilai 1.000000
Name: Kesimpulan_nilai, dtype: float64

```

Gambar 4. 1 Correlation Matrix

Dari hasil angka correlation matrix diatas, didapatkan hasil:

1. Kesimpulan_nilai dan LB (Luas Bangunan): Korelasi positif yang tinggi (0.814454) menunjukkan bahwa ada hubungan yang kuat antara luas bangunan (LB) dengan kesimpulan nilai. Ini berarti semakin besar luas bangunan, semakin tinggi kemungkinan nilai kesimpulan.
2. Kesimpulan_nilai dan Indikasi_nilai: Korelasi positif (0.543462) antara indikasi nilai dan kesimpulan nilai menunjukkan adanya hubungan yang cukup kuat. Ini bisa berarti bahwa indikasi nilai yang tinggi cenderung berkorelasi dengan nilai kesimpulan yang tinggi.
3. Kesimpulan_nilai dan Jumlah_lantai: Korelasi positif (0.389262) antara jumlah lantai dan kesimpulan nilai menunjukkan bahwa properti dengan

jumlah lantai yang lebih tinggi mungkin cenderung memiliki nilai kesimpulan yang lebih tinggi.

4. Kesimpulan_nilai dan Jumlah_kamar / Jumlah_kamarmandi: Korelasi positif yang moderat (0.290673 untuk Jumlah_kamar dan 0.308836 untuk Jumlah_kamarmandi) menunjukkan bahwa ada hubungan positif antara jumlah kamar atau jumlah kamar mandi dengan kesimpulan nilai. Artinya, properti dengan jumlah kamar atau kamar mandi yang lebih banyak cenderung memiliki nilai kesimpulan yang lebih tinggi.
5. Kesimpulan_nilai dan Jumlah_kamar / Jumlah_kamarmandi: Korelasi positif yang moderat (0.290673 untuk Jumlah_kamar dan 0.308836 untuk Jumlah_kamarmandi) menunjukkan bahwa ada hubungan positif antara jumlah kamar atau jumlah kamar mandi dengan kesimpulan nilai. Artinya, properti dengan jumlah kamar atau kamar mandi yang lebih banyak cenderung memiliki nilai kesimpulan yang lebih tinggi.

4.3 Hasil Data Splitting

Proses Data Splitting menghasilkan dua datasets yaitu training data dan juga testing data. Berikut adalah hasil pemisahan data training dan data testing menggunakan rasio yang berbeda-beda. Pemisahan data menggunakan metode Random State.

Tabel 4. 3 Hasil Rasio dan Splitting Data

No	Rasio (Training/Testing)	Jumlah Data Training	Jumlah Data Set
1.	90:10	90	10
2.	80:20	80	20
3.	70:30	70	30
4.	60:40	60	40
5.	50:50	50	50

Dengan menggunakan random seed, kita dapat membuat pembagian data yang sama setiap kali pemisahan dilakukan. Ini penting untuk memastikan hasil eksperimen yang dapat direproduksi. Ketika kita membagi data untuk mengembangkan dan mengevaluasi model, hasil yang konsisten dapat membantu dalam perbandingan model dan penemuan kesalahan.

4.3 Hasil Pengujian *Random Forest*

4.3.1 Hasil Pengujian Rasio 90:10

Hasil Pengujian dengan split 90:10 Terdapat pada tabel berikut:

Tabel 4. 4 Data Pengujian 90:10

Nama Debitur	Obyek	Kota/Kabupaten	LT	LB	Asli	Predict	Error
Agung Prasetyo	Rumah Tinggal	Kab. Malang	90	108	406000000 .0	416102300 .0	10102300. 0
Hariyanto Erifin	Rumah Tinggal	Kab. Malang	78	36	165810000 0.0	172551200 0.0	67412000. 0
Vivin Nurlaila	Rumah Tinggal	Kab. Malang	249	98	618000000 .0	623040000 .0	5040000.0
Susy Pratiwi	Rumah Tinggal	Kab. Malang	71	44	574000000 .0	905310000 .0	331310000 .0
Gyindra Joevilka Arleo	Rumah Tinggal	Kota Malang	102	37	350000000 .0	373620000 .0	23620000. 0
Agung Priambodo	Rumah Tinggal	Kota Malang	129	68	255000000 .0	317200000 .0	62200000. 0
Sendy Mutiara Karlina	Rumah Tinggal	Kota Malang	139	67	747450000 .0	693208000 .0	54242000. 0
Amalina Samantha	Rumah Tinggal	Kab. Malang	60	25	583000000 .0	596220000 .0	13220000. 0

Model Pengujian diatas menggunakan 8 features berdasarkan 8 variabel diantaranya adalah LT, LB, Kota/kabupaten, indikasi nilai, jumlah kamar, jumlah kamar mandi, jumlah lantai, pusat kota, terhadap variabel output yaitu nilai asset Berikut merupakan hasil dari model yang telah dibuat:

MAE = 83681975.0

MAPE = 15.10%

4.3.2 Hasil Pengujian Rasio 80:20

Hasil Pengujian dengan split 80:20 Terdapat pada tabel berikut:

Tabel 4. 5 Data Pengujian 80:20

Nama Debitur	Obyek	Kota/Kabupaten	LT	LB	Asli	Predict	Error
Agung Prasetyo	Rumah Tinggal	Kab. Malang	90	108	406000000.0	416102300.0	10102300.0
Hariyanto Erifin	Rumah Tinggal	Kab. Malang	78	36	1658100000.0	1725512000.0	67412000.0
Vivin Nurlaila	Rumah Tinggal	Kab. Malang	249	98	618000000.0	623040000.0	5040000.0
Susy Pratiwi	Rumah Tinggal	Kab. Malang	71	44	574000000.0	905310000.0	331310000.0
Gyndra Joevilka Arleo	Rumah Tinggal	Kota Malang	102	37	350000000.0	373620000.0	23620000.0
Agung Priambodo	Rumah Tinggal	Kota Malang	129	68	255000000.0	317200000.0	62200000.0
Sendy Mutiara Karlina	Rumah Tinggal	Kota Malang	139	67	747450000.0	693208000.0	54242000.0
Amalina Samantha	Rumah Tinggal	Kab. Malang	60	25	583000000.0	596220000.0	13220000.0
Andika Veriastianto	Rumah Tinggal	Kota Malang	87	45	1211500000.0	1148195500.0	63304500.0
Rhoma Dona	Rumah Tinggal	Kota Malang	135	85	498000000.0	532002300.0	34002300.0

Rahmawan Abdi							
Hendra Kristiawan	Rumah Tinggal	Kota Malang	90	60	508800000.0	560887500.0	52087500.0
Bpk. Anton Wijaya	Rumah Tinggal	Kab. Malang	720	707	3894000000.0	3155138000.0	738862000.0
Ari Sulistiono	Ruko/Rukan	Kab. Pasuruan	58	74	2883000000.0	2795188000.0	87812000.0
CV Leading Digital Semesta	Ruko/Rukan	Kab. Jember	176	209	7100000000.0	4615360000.0	248464000.0
Edy Wijaya	Rumah Tinggal	Kab. Jember	254	150	10600000000.0	1045608000.0	14392000.0
Khusnul Khotimah	Rumah Tinggal	Kota Malang	105	158	4880000000.0	6069675000.0	118967500.0

Model Pengujian diatas menggunakan 8 features berdasarkan 8 variabel diantaranya adalah LT, LB, Kota/kabupaten, indikasi nilai, jumlah kamar, jumlah kamar mandi, jumlah lantai, pusat kota, terhadap variabel output yaitu nilai asset

Berikut merupakan hasil dari model yang telah dibuat :

MAE = 120314881.25

MAPE = 13.17%

4.3.3 Hasil Pengujian Rasio 70:30

Berikut adalah Hasil Pengujian dengan Rasio 70:30.

Tabel 4. 6 Data Pengujian 70:30

Nama Debitur	Obyek	Kota/Kabupaten	LT	LB	Asli	Predict	Error
Agung Prasetyo	Rumah Tinggal	Kab. Malang	90	108	406000000.0	416102300.0	10102300.0
Hariyanto	Rumah	Kab.	78	36	165810000	172551200	67412000.

Erifin	Tinggal	Malang			0.0	0.0	0
Vivin Nurlaila	Rumah Tinggal	Kab. Malang	249	98	618000000.0	623040000.0	5040000.0
Susy Pratiwi	Rumah Tinggal	Kab. Malang	71	44	574000000.0	905310000.0	331310000.0
Gyndra Joevilka Arleo	Rumah Tinggal	Kota Malang	102	37	350000000.0	373620000.0	23620000.0
Agung Priambodo	Rumah Tinggal	Kota Malang	129	68	255000000.0	317200000.0	62200000.0
Sendy Mutiara Karlina	Rumah Tinggal	Kota Malang	139	67	747450000.0	693208000.0	54242000.0
Amalina Samantha	Rumah Tinggal	Kab. Malang	60	25	583000000.0	596220000.0	13220000.0
Andika Veriastianto	Rumah Tinggal	Kota Malang	87	45	121150000.0	114819550.0	63304500.0
Rhoma Dona Rahmawan Abdi	Rumah Tinggal	Kota Malang	135	85	498000000.0	532002300.0	34002300.0
Hendra Kristiawan	Rumah Tinggal	Kota Malang	90	60	508800000.0	560887500.0	52087500.0
Bpk. Anton Wijaya	Rumah Tinggal	Kab. Malang	720	707	389400000.0	315513800.0	738862000.0
Ari Sulistiono	Ruko/Rukan	Kab. Pasuruan	58	74	288300000.0	279518800.0	87812000.0
CV Leading Digital Semesta	Ruko/Rukan	Kab. Jember	176	209	710000000.0	461536000.0	248464000.0
Edy Wijaya	Rumah Tinggal	Kab. Jember	254	150	106000000.0	104560800.0	14392000.0
Khusnul Khotimah	Rumah Tinggal	Kota Malang	105	158	488000000.0	606967500.0	118967500.0
PT. Sari Bahari	Tanah Kosong	Kab. Malang	1142	0	106100000.0	794785000.0	266215000.0
Margaretha Nurkatika	Rumah Tinggal	Kab. Malang	109	54	654000000.0	612434000.0	41566000.0

Wiedya							
Novita Primadhyni Putri	Rumah Tinggal	Kota Malang	300	324	241900000 .0	248105500 .0	62055000. 0
Yusuf Afandi	Rumah Tinggal	Kab. Jember	427	318	885000000 .0	837386500 .0	47613500. 0
Irfan Arfian	Rumah Tinggal	Kab. Jember	326	415	490000000 .0	539120000 .0	49120000. 0
Yudi Hartanto	Rumah Tinggal	Kab. Banyuwangi	29	52	655000000 .0	631810000 .0	23190000. 0
Yudi Hartanto	Tanah Kosong	Kab. Banyuwangi	885	0	484000000 .0	468099000 .0	15901000. 0
Yudi Hartanto	Rumah Tinggal	Kab. Banyuwangi	1208	234	908500000 .0	800643500 .0	107856500 .0

Model Pengujian diatas menggunakan 8 features berdasarkan 8 variabel diantaranya adalah LT, LB, Kota/kabupaten, indikasi nilai, jumlah kamar, jumlah kamar mandi, jumlah lantai, pusat kota, terhadap variabel output yaitu nilai asset Berikut merupakan hasil dari model yang telah dibuat :

MAE = 106286529

MAPE = 11.29%

4.3.4 Hasil Pengujian Rasio 60:40

Berikut adalah Hasil Pengujian dengan Rasio 60:40.

Tabel 4. 7 Data Pengujian 60:40

Nama Debitur	Obyek	Kota/Kabupaten	LT	LB	Asli	Predict	Error
Agung Prasetyo	Rumah Tinggal	Kab. Malang	90	108	406000000 .0	416102300 .0	10102300. 0
Hariyanto	Rumah	Kab.	78	36	165810000	172551200	67412000.

Erifin	Tinggal	Malang			0.0	0.0	0
Vivin Nurlaila	Rumah Tinggal	Kab. Malang	249	98	618000000.0	623040000.0	5040000.0
Susy Pratiwi	Rumah Tinggal	Kab. Malang	71	44	574000000.0	905310000.0	331310000.0
Gyndra Joevilka Arleo	Rumah Tinggal	Kota Malang	102	37	350000000.0	373620000.0	23620000.0
Agung Priambodo	Rumah Tinggal	Kota Malang	129	68	255000000.0	317200000.0	62200000.0
Sendy Mutiara Karlina	Rumah Tinggal	Kota Malang	139	67	747450000.0	693208000.0	54242000.0
Amalina Samantha	Rumah Tinggal	Kab. Malang	60	25	583000000.0	596220000.0	13220000.0
Andika Veriastianto	Rumah Tinggal	Kota Malang	87	45	121150000.0	114819550.0	63304500.0
Rhoma Dona Rahmawan Abdi	Rumah Tinggal	Kota Malang	135	85	498000000.0	532002300.0	34002300.0
Hendra Kristiawan	Rumah Tinggal	Kota Malang	90	60	508800000.0	560887500.0	52087500.0
Bpk. Anton Wijaya	Rumah Tinggal	Kab. Malang	720	707	389400000.0	315513800.0	738862000.0
Ari Sulistiono	Ruko/Rukan	Kab. Pasuruan	58	74	288300000.0	279518800.0	87812000.0
CV Leading Digital Semesta	Ruko/Rukan	Kab. Jember	176	209	710000000.0	461536000.0	248464000.0
Edy Wijaya	Rumah Tinggal	Kab. Jember	254	150	106000000.0	104560800.0	14392000.0
Khusnul Khotimah	Rumah Tinggal	Kota Malang	105	158	488000000.0	606967500.0	118967500.0
PT. Sari Bahari	Tanah Kosong	Kab. Malang	1142	0	106100000.0	794785000.0	266215000.0
Margaretha Nurkatika	Rumah Tinggal	Kab. Malang	109	54	654000000.0	612434000.0	41566000.0

Wiedya							
Novita Primadhyni Putri	Rumah Tinggal	Kota Malang	300	324	241900000 .0	248105500 .0	62055000. 0
Yusuf Afandi	Rumah Tinggal	Kab. Jember	427	318	885000000 .0	837386500 .0	47613500. 0
Irfan Arfian	Rumah Tinggal	Kab. Jember	326	415	490000000 .0	539120000 .0	49120000. 0
Yudi Hartanto	Rumah Tinggal	Kab. Banyuwangi	29	52	655000000 .0	631810000 .0	23190000. 0
Yudi Hartanto	Tanah Kosong	Kab. Banyuwangi	885	0	484000000 .0	468099000 .0	15901000. 0
Yudi Hartanto	Rumah Tinggal	Kab. Banyuwangi	1208	234	908500000 .0	800643500 .0	107856500 .0
Salimul Rakhmad	Rumah Tinggal	Kota Malang	67	38	591650000 .0	603358500 .0	11708500. 0
Dwi Hindaryo Citro Sumbono	Rumah Tinggal	Kota Malang	87	45	430000000 .0	434440000 .0	4440000.0
Yulis Indah Aristyani	Rumah Tinggal	Kab. Malang	72	50	111700000 0.0	112958950 0.0	12589500. 0
Novia Tifany Putri Dwi Y.	Rumah Tinggal	Kab. Malang	69	38	158000000 .0	197687000 .0	39687000. 0
Heri Wiyono	Rumah Tinggal	Kab. Malang	77	48	607000000 .0	820735000 .0	213735000 .0
Puspita Ayu Wijaya	Rumah Tinggal	Kota Malang	168	336	763000000 .0	864643500 .0	101643500 .0
Nurul Hidayah	Rumah Tinggal	Kota Malang	77	51	502000000 .0	474857000 .0	27143000. 0

Model Pengujian diatas menggunakan 8 features berdasarkan 8 variabel diantaranya adalah LT, LB, Kota/kabupaten, indikasi nilai, jumlah kamar, jumlah kamar mandi, jumlah lantai, pusat kota, terhadap variabel output yaitu nilai asset Berikut merupakan hasil dari model yang telah dibuat :

MAE = 156089548

MAPE = 14.64%

4.3.5 Hasil Pengujian Rasio 50:50

Berikut Adalah hasil pengujian Rasio 50:50

Tabel 4. 8 Data Pengujian 50:50

Nama Debitur	Obyek	Kota/Kabupaten	LT	LB	Asli	Predict	Error
Agung Prasetyo	Rumah Tinggal	Kab. Malang	90	108	406000000.0	416102300.0	10102300.0
Hariyanto Erifin	Rumah Tinggal	Kab. Malang	78	36	1658100000.0	1725512000.0	67412000.0
Vivin Nurlaila	Rumah Tinggal	Kab. Malang	249	98	618000000.0	623040000.0	5040000.0
Susy Pratiwi	Rumah Tinggal	Kab. Malang	71	44	574000000.0	905310000.0	331310000.0
Gyindra Joevilka Arleo	Rumah Tinggal	Kota Malang	102	37	350000000.0	373620000.0	23620000.0
Agung Priambodo	Rumah Tinggal	Kota Malang	129	68	255000000.0	317200000.0	62200000.0
Sendy Mutiara Karlina	Rumah Tinggal	Kota Malang	139	67	747450000.0	693208000.0	54242000.0
Amalina Samantha	Rumah Tinggal	Kab. Malang	60	25	583000000.0	596220000.0	13220000.0
Andika Veriastianto	Rumah Tinggal	Kota Malang	87	45	1211500000.0	1148195500.0	63304500.0
Rhoma Dona Rahmawan	Rumah Tinggal	Kota Malang	135	85	498000000.0	532002300.0	34002300.0

Abdi							
Hendra Kristiawan	Rumah Tinggal	Kota Malang	90	60	508800000.0	560887500.0	52087500.0
Bpk. Anton Wijaya	Rumah Tinggal	Kab. Malang	720	707	3894000000.0	3155138000.0	738862000.0
Ari Sulistiono	Ruko/Rukan	Kab. Pasuruan	58	74	2883000000.0	2795188000.0	87812000.0
CV Leading Digital Semesta	Ruko/Rukan	Kab. Jember	176	209	710000000.0	461536000.0	248464000.0
Edy Wijaya	Rumah Tinggal	Kab. Jember	254	150	1060000000.0	1045608000.0	14392000.0
Khusnul Khotimah	Rumah Tinggal	Kota Malang	105	158	488000000.0	606967500.0	118967500.0
PT. Sari Bahari	Tanah Kosong	Kab. Malang	1142	0	1061000000.0	794785000.0	266215000.0
Margaretha Nurkatika Wiedya	Rumah Tinggal	Kab. Malang	109	54	654000000.0	612434000.0	41566000.0
Novita Primadhyni Putri	Rumah Tinggal	Kota Malang	300	324	2419000000.0	2481055000.0	62055000.0
Yusuf Afandi	Rumah Tinggal	Kab. Jember	427	318	885000000.0	837386500.0	47613500.0
Irfan Arfian	Rumah Tinggal	Kab. Jember	326	415	490000000.0	539120000.0	49120000.0
Yudi Hartanto	Rumah Tinggal	Kab. Banyuwangi	29	52	655000000.0	631810000.0	23190000.0
Yudi Hartanto	Tanah Kosong	Kab. Banyuwangi	885	0	484000000.0	468099000.0	15901000.0
Yudi Hartanto	Rumah Tinggal	Kab. Banyuwangi	1208	234	908500000.0	800643500.0	107856500.0
Salimul Rakhmad	Rumah Tinggal	Kota Malang	67	38	591650000.0	603358500.0	11708500.0

Dwi Hindaryo Citro Sumbono	Rumah Tinggal	Kota Malang	87	45	430000000 .0	434440000 .0	4440000.0
Yulis Indah Aristyani	Rumah Tinggal	Kab. Malang	72	50	111700000 0.0	112958950 0.0	12589500. 0
Novia Tifany Putri Dwi Y.	Rumah Tinggal	Kab. Malang	69	38	158000000 .0	197687000 .0	39687000. 0
Heri Wiyono	Rumah Tinggal	Kab. Malang	77	48	607000000 .0	820735000 .0	213735000 .0
Puspita Ayu Wijaya	Rumah Tinggal	Kota Malang	168	336	763000000 .0	864643500 .0	101643500 .0
Nurul Hidayah	Rumah Tinggal	Kota Malang	77	51	502000000 .0	474857000 .0	27143000. 0
Irwan Prasetyo	Rumah Tinggal	Kota Malang	118	121	220720000 0.0	237211600 0.0	164916000 .0
Sugiyono Mursini	Rumah Tinggal	Kab. Jember	121	115	428400000 .0	443970000 .0	15570000. 0
Iqbal Arief	Rumah Tinggal	Kab. Jember	100	170	250730000 0.0	233649200 0.0	170808000 .0
Distia Witantinin grum	Rumah Tinggal	Kab. Jember	235	350	959700000 .0	105919050 0.0	99490500. 0
Iwan Wahyudi	Rumah Tinggal	Kab. Jember	291	259	459000000 .0	506980000 .0	47980000. 0
Mohamad Romli	Rumah Tinggal	Kab. Banyuwangi	629	596	180000000 0.0	132298150 0.0	477018500 .0
Mohamad Romli	Rumah Tinggal	Kab. Banyuwangi	407	136	138000000 0.0	131017700 0.0	69823000. 0
Eko Agus Lukiantoro SE	Rumah Tinggal	Kab. Banyuwangi	582	342	675000000 .0	649705000 .0	25295000. 0

Model Pengujian diatas menggunakan 8 features berdasarkan 8 variabel diantaranya adalah LT, LB, Kota/kabupaten, indikasi nilai, jumlah kamar, jumlah

kamar mandi, jumlah lantai, pusat kota, terhadap variabel output yaitu nilai asset

Berikut merupakan hasil dari model yang telah dibuat:

$$\text{MAE} = 168542525$$

$$\text{MAPE} = 16.09\%$$

4.4 Analisa Hasil Pengujian

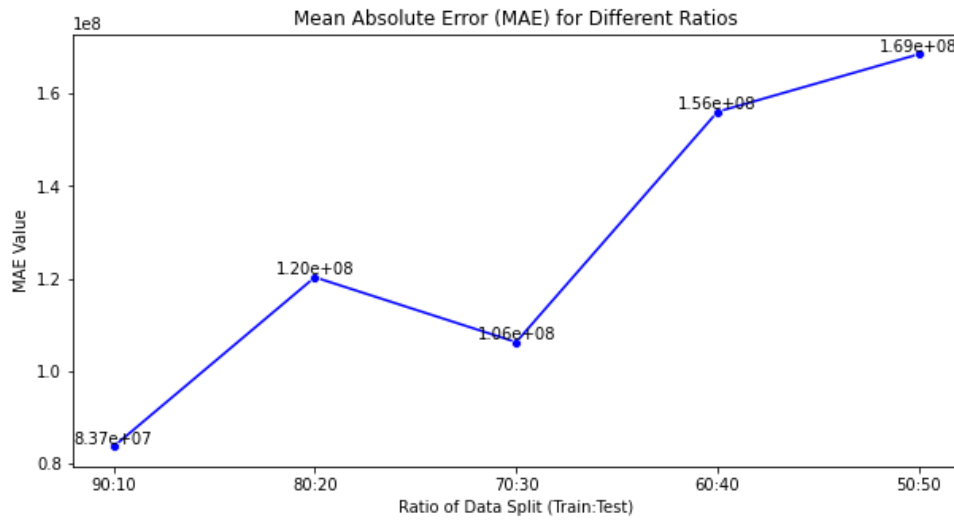
Hasil dari Pengujian *Random Forest* dengan dilakukan pembagian ratio split yang berbeda-beda dapat ditampilkan pada tabel berikut.

Tabel 4. 9 Tabel Hasil Pengujian *Random Forest*

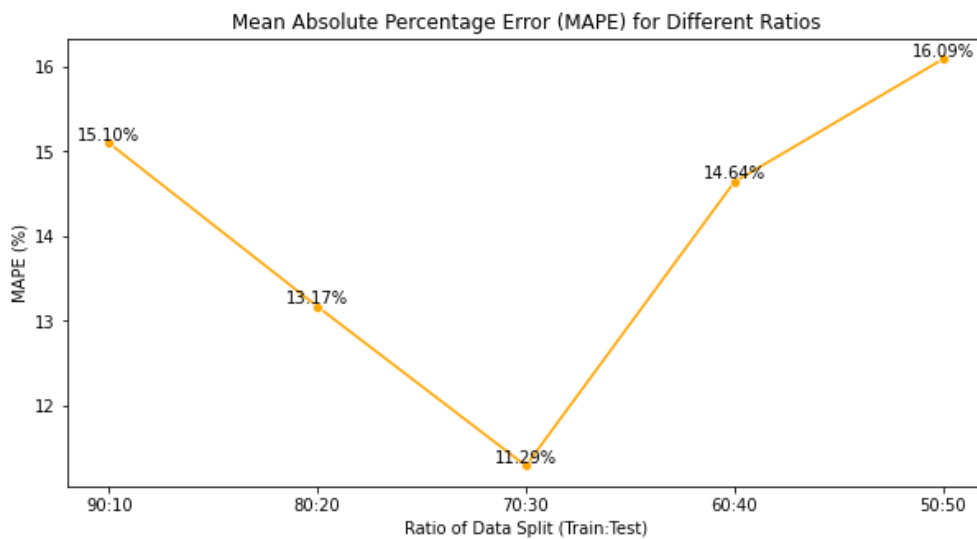
No	Rasio Split (Testing : Training)	Mean Absolute Error (MAE)	Mean Percentage Absolute Error (MAPE)
1.	90:10	83681975.0	15.10%
2.	80:20	120314881.25	13.17%
3.	70:30	106286529	11.29%
4.	60:40	156089548	14.64%
5.	50:50	168542525	16.09%

Berdasarkan nilai MAE, Rasio terbaik adalah rasio 90:10, sedangkan berdasarkan nilai MAPE, Rasio terbaik adalah rasio 70:30. Untuk menentukan performa model yang terbaik dan konsistensi model, analisis hasil pengujian *Random Forest* dilakukan. Visualisasi data (plot) dari tabel di atas digunakan untuk mempermudah dalam memahami data dan melakukan analisis, berikut adalah gambar dari Visualisasi data plot MAE dan MAPE.

Dari hasil pengujian diatas, didapat plot MAE grafik sebagai berikut.



Gambar 4. 2 Hasil MAE



Gambar 4. 3 Hasil MAPE

Dalam penelitian ini pengujian menguji performa model *Random Forest* pada rasio split data yang berbeda untuk prediksi harga rumah. Hipotesis awal adalah rasio split data yang lebih rendah seharusnya memiliki nilai MAE dan MAPE yang lebih rendah, sedangkan rasio yang lebih tinggi seharusnya memiliki nilai yang

lebih tinggi. Berikut adalah analisis perbandingan antara sistem pada setiap rasio split:

Sesuai hipotesis, nilai MAE cenderung meningkat seiring dengan peningkatan rasio data testing. Hal ini menunjukkan bahwa pada rasio split yang lebih tinggi, model memiliki kesalahan prediksi yang lebih besar dalam satuan mata uang. Pada rasio split 90:10, MAE mencapai nilai terendah, mengindikasikan bahwa model memiliki kinerja terbaik pada rasio ini dengan nilai MAE 83681975.0.

Secara keseluruhan, hipotesis tentang MAPE kurang terpenuhi. Meskipun terdapat peningkatan pada rasio yang lebih tinggi, namun pada rasio split 50:50, MAPE meningkat secara signifikan. Kesalahan yang lebih rendah pada data yang lebih sedikit dapat menyebabkan fluktuasi besar pada MAPE, terutama jika jumlah data testing sangat kecil.

Menganalisis hasil model machine learning melalui berbagai rasio split data merupakan langkah penting untuk memahami sejauh mana model mampu berkinerja pada data yang tidak pernah ditemui (data validasi) dan seberapa konsisten model dalam memberikan prediksi pada kondisi yang berbeda. Proses ini memberikan wawasan yang berharga terkait generalisasi model dan keandalannya dalam menghadapi variasi data. Berikut adalah analisis evaluasi hasil model dengan pendekatan tersebut:

Dengan menggunakan berbagai rasio split data, peneliti dapat mengamati perilaku model pada data validasi dengan melihat rasio data testing yang lebih kecil. Hasil yang konsisten dan baik pada data testing yang lebih kecil menunjukkan

bahwa model memiliki kemampuan generalisasi yang baik terhadap data yang tidak pernah ditemui sebelumnya.

Misalnya, pada rasio split 70:30, di mana data testing lebih kecil, jika model memberikan hasil yang baik, hal ini menandakan bahwa model mungkin dapat melakukan dengan baik pada data yang belum pernah ditemui sebelumnya.

Dengan membandingkan hasil pada setiap rasio split, peneliti dapat mengevaluasi konsistensi model dalam memberikan prediksi. Konsistensi yang baik menunjukkan bahwa model tidak terlalu sensitif terhadap perubahan komposisi data training dan testing.

Misalnya, jika terdapat variasi yang signifikan dalam hasil prediksi pada rasio split yang berbeda, hal ini dapat menandakan bahwa model mungkin kurang konsisten dan lebih rentan terhadap fluktuasi data.

Melalui analisis ini, peneliti dapat mencari rasio split yang optimal di mana model mencapai kinerja terbaiknya. Rasio ini dapat menjadi representasi dari seberapa baik model dapat menangani trade-off antara underfitting dan overfitting.

Misalnya, jika pada rasio split 70:30 model memberikan hasil terbaik, peneliti dapat menganggap rasio tersebut sebagai pilihan yang optimal untuk menjaga keseimbangan antara data training dan testing.

Melalui analisis hasil model pada berbagai rasio split data, peneliti dapat mengidentifikasi sejauh mana model mampu beradaptasi dengan data baru, sekaligus menilai konsistensinya dalam memberikan prediksi. Hasil ini dapat memberikan pandangan yang lebih holistik tentang kemampuan model dalam memahami pola pada dataset yang beragam. Pemilihan rasio split optimal dapat

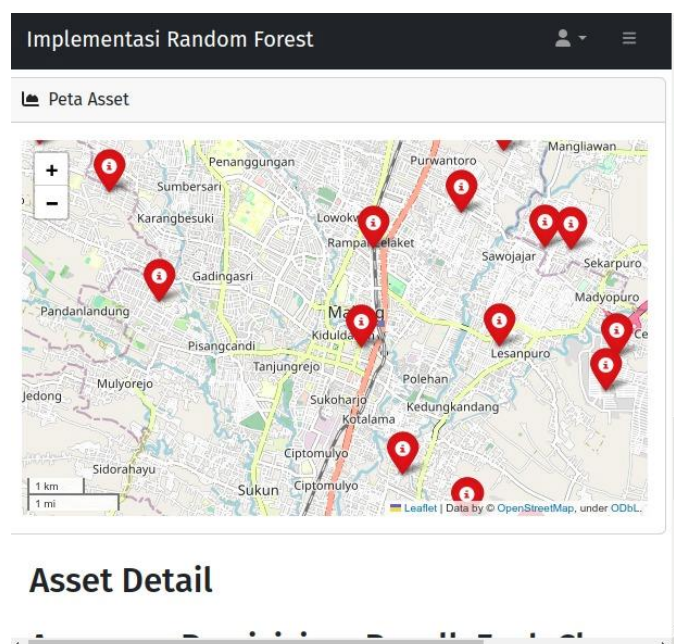
membantu peneliti menentukan konfigurasi model yang paling sesuai dengan karakteristik data yang dihadapi. Dengan demikian, analisis ini menjadi langkah penting dalam memvalidasi dan mengoptimalkan model machine learning untuk mencapai hasil yang konsisten dan dapat diandalkan.

4.5 Hasil Tampilan Aplikasi GIS

4.5.1 Halaman Main Map

Halaman main map terdapat dua bagian pada tampilan utama yang disusun vertikal. Pada bagian atas terdapat Dashboard Map yang memuat informasi Lokasi Asset dan juga Atribut Asset jika kita mengklik ikon asset pada dashboard tersebut. Diikuti oleh tampilan Table yang memuat informasi table data Asset yang rinci. Berikut adalah tampilan halaman “Main Map” yang telah dibangun.

1. Tampilan GIS Peta Asset



Gambar 4. 4 GIS Peta Asset

Peta Asset digunakan untuk mengetahui lokasi riil data asset, sehingga dapat memudahkan kita untuk memahami persebaran data dan lokasi data. Visualisasi juga di maksudkan untuk mempermudah dalam mengamati Pola Harga Asset per Kabupaten / Kota.

2. Tampilan Table Data Asset.

Titik_koordinat	Nama_debitur	Nama_Pemberi_Tugas	Kota_Kabupaten	LT	LB	Tujuan	Obyek	Fee_total	dpp	ppn	Indikasi_nilai
-7913450, 112.663797	Agung Prasetyo	PT. Bank Mandiri (Persero) Tbk.	Kab. Malang	90	108	Penjaminan Utang	Rumah Finggal	1320000	1200000	120000	2400000
-7919838, 112.578288	Hariyanto Erifin	PT. Bank CIMB Niaga Tbk.	Kab. Malang	78	36	Penjaminan Utang	Rumah Finggal	1320000	1200000	120000	4500000
-8.029701, 112.616644	Velin Nurhaila	PT. Bank Mandiri (Persero) Tbk.	Kab. Malang	249	98	Penjaminan Utang	Rumah Finggal	1320000	1200000	120000	900000
-790398, 112.671122	Sury Pratiwi	PT. Bank Mandiri (Persero) Tbk.	Kab. Malang	71	44	Penjaminan Utang	Rumah Finggal	1320000	1200000	120000	2900000
-7917422, 112.621929	Gyindra Jocvika Arico	PT. Bank Mandiri (Persero) Tbk.	Kota Malang	102	37	Penjaminan Utang	Rumah Finggal	1320000	1200000	120000	2400000
-8.021485, 112.632600	Agung Priambodo	PT. Bank Mandiri (Persero) Tbk.	Kota Malang	129	68	Penjaminan Utang	Rumah Finggal	1320000	1200000	120000	3100000
-7918224, 112.593710	Sandy Marlara Karlina	PT. Bank Mandiri (Persero) Tbk.	Kota Malang	139	67	Penjaminan Utang	Rumah Finggal	1320000	1200000	120000	3300000
-7989671	Amalina	PT. Bank Mandiri	Kab. Malang	60	25	Penjaminan	Rumah	1320000	1200000	120000	1300000

Gambar 4. 5 Tampilan table data asset

Dari tampilan GIS diatas, jika discroll dibawah maka terdapat tampilan data asset. Tampilan data asset digunakan untuk mengetahui detail data dan atribut data dari asset asset tersebut.

4.5.2 Halaman Uji *Random Forest*.

Halaman uji *Random Forest* terdapat table yang memuat informasi data features dan data target dan data prediksi, halaman ini berfungsi untuk mengetahui hasil uji *Random Forest*.

Nama Debitur	Obyek	Kota/Kabupaten	LT	LB	Asli	Predict	Error
Agung Prasetyo	Rumah Tinggal	Kab. Malang	90	108	406000000.0	416102300.0	10102300.0
Hariyanto Erifin	Rumah Tinggal	Kab. Malang	78	36	1658100000.0	1725512000.0	67412000.0
Vivin Nurlaila	Rumah Tinggal	Kab. Malang	249	98	618000000.0	623040000.0	5040000.0
Susy Pratiwi	Rumah Tinggal	Kab. Malang	71	44	574000000.0	905310000.0	331310000.0
Gyindra Joevilka Arleo	Rumah Tinggal	Kota Malang	102	37	350000000.0	373620000.0	23620000.0
Agung Priambodo	Rumah Tinggal	Kota Malang	129	68	255000000.0	317200000.0	62200000.0
Sendy Mutiara Karlina	Rumah Tinggal	Kota Malang	139	67	747450000.0	693208000.0	54242000.0
Amalina Samantha	Rumah Tinggal	Kab. Malang	60	25	583000000.0	596220000.0	13220000.0
Andika Veriastianto	Rumah Tinggal	Kota Malang	87	45	1211500000.0	1148195500.0	63304500.0
Rhoma Dona Rahmawan Abdi	Rumah Tinggal	Kota Malang	135	85	498000000.0	532002300.0	34002300.0
Hendra Kristiawan	Rumah Tinggal	Kota Malang	90	60	508800000.0	560887500.0	52087500.0
Bpk. Anton Wijaya	Rumah Tinggal	Kab. Malang	720	707	3894000000.0	3155138000.0	738862000.0

Gambar 4. 6 Halaman Hasil Uji *Random Forest*

Table diatas memuat informasi data features, Nilai Asset asli, dan data nilai Asset predict. Nilai Asset asli adalah nilai awal dari data sebagai data testing, dan nilai predict adalah nilai hasil prediksi *Random Forest*. Untuk mengetahui error dari masing masing prediksi, maka dilakukan $(y - y')$ atau data asli dikurangi dengan data prediksi.

4.6 Integrasi Penelitian dengan Islam

Dengan adanya sistem yang dibangun ini merupakan salah satu wujud dari *muamalah ma'a An-Nas* yaitu wujud tolong-menolong sesama manusia yakni memudahkan para pegawai kantor penilaian dalam menentukan pinjaman yang sesuai dengan kemampuan nasabah. Sebagaimana dijelaskan dalam Q.S Al Maidah ayat 2:

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا تَحْلُوا شَعَائِرَ اللَّهِ وَلَا الشَّهْرَ الْحَرَامَ وَلَا الْهَدْيَ وَلَا الْقَلَائِدَ وَلَا آمِينَ الْبَيْتِ الْحَرَامِ يَبْتَغُونَ فَضْلًا
مِّن رَّبِّهِمْ وَرِضْوَانًا يَوْمَآ حَلَلْتُمْ فَاصْطَادُوا وَلَا يَجْرِمَنَّكُمْ شَنَا نَقَوْمِ أَن صَدُّوكُمْ عَنِ الْمَسْجِدِ الْحَرَامِ أَن تَعْتَدُوا
وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ يَوْمَئِذٍ اللَّهُ لِلَّهِ شَدِيدُ الْعِقَابِ

“Wahai orang-orang yang beriman, janganlah kamu melanggar syair-syair (kesucian) Allah, jangan (melanggar kehormatan) bulan-bulan haram, jangan

(menggangu) hadyu (hewan-hewan kurban) dan qalā'id (hewan-hewan kurban yang diberi tanda), dan jangan (pula menggangu) para pengunjung Baitul Haram sedangkan mereka mencari karunia dan rida Tuhannya! Apabila kamu telah bertahalul (menyelesaikan ihram), berburulah (jika mau). Janganlah sekali-kali kebencian(-mu) kepada suatu kaum, karena mereka menghalang-halangi dari Masjidil Haram, mendorongmu berbuat melampaui batas (kepada mereka). Tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan permusuhan. Bertakwalah kepada Allah, sesungguhnya Allah sangat berat siksaanNya." (Q.S. Al-Maidah 2).

Tafsir Ibnu Katsir menjelaskan (Bertolong-tolonglah kamu dalam kebaikan) dalam mengerjakan yang dititahkan (dan ketakwaan) dengan meninggalkan apa-apa yang dilarang (dan janganlah kamu bertolong-tolongan) pada *ta`aawanu* dibuang salah satu diantara dua *ta* pada asalnya (dalam berbuat dosa) atau maksiat (dan pelanggaran) artinya melampaui batas-batas ajaran Allah. (Dan bertakwalah kamu kepada Allah) takutlah kamu kepada azab siksa-Nya dengan menaati-Nya (sesungguhnya Allah amat berat siksa-Nya) bagi orang yang menentang-Nya (Ghoffar, 2000).

Tafsir Ibnu Katsir menyimpulkan bahwa ayat ini mengajarkan untuk saling tolong-menolong dalam kebaikan dan ketakwaan. Masyarakat diminta untuk bersatu dalam melaksanakan perintah Allah dan menjauhi larangan-Nya. Larangan juga diberikan terhadap tolong-menolong dalam berbuat dosa, maksiat, dan pelanggaran terhadap ajaran Allah. Selain itu, ayat ini menekankan pentingnya takwa kepada Allah sebagai bentuk kewaspadaan terhadap azab-Nya yang berat bagi mereka yang menentang-Nya.

Dalam mengamalkan wujud dari *muamalah ma'a allah* yaitu dengan bertaqwa kepada allah dan memperhatikan apa yang telah diperbuat serta

memperhatikan hari esok merupakan wujud dari muamalah ma allah seperti tertera pada surat Al-Hasyr ayat 18 yang berbunyi.

يَا أَيُّهَا الَّذِينَ آمَنُوا اتَّقُوا اللَّهَ وَلْتَنْظُرْ نَفْسٌ مَّا قَدَّمَتْ لِغَدٍ وَاتَّقُوا اللَّهَ إِنَّ اللَّهَ خَبِيرٌ بِمَا تَعْمَلُونَ

“Wahai orang-orang yang beriman! Bertakwalah kepada Allah dan hendaklah setiap orang memperhatikan apa yang telah diperbuatnya untuk hari esok (akhirat), dan bertakwalah kepada Allah. Sungguh, Allah Mahateliti terhadap apa yang kamu kerjakan.” (Q.S. Al-Hasyr 18).

Allah Subhaanahu wa Ta'aala menegaskan kepada hamba-hamba-Nya yang beriman untuk mengikuti kehendak-Nya dan mengambil konsekuensi dari iman mereka, yaitu dengan tetap menjalankan takwa kepada *Allah Subhaanahu wa Ta'aala*, baik itu dalam keadaan tersembunyi maupun di hadapan publik, serta dalam setiap situasi. Mereka diarahkan untuk memperhatikan perintah dan batasan Allah, serta untuk mempertimbangkan dampak baik dan buruk dari tindakan mereka. Penting bagi mereka untuk fokus pada akhirat, baik dalam pemikiran maupun hati, sehingga mereka dengan sungguh-sungguh berusaha untuk meningkatkan amal perbuatan yang dapat membawa kebahagiaan di akhirat. Mereka menyadari bahwa *Allah Subhaanahu wa Ta'aala* melihat setiap tindakan mereka, tidak ada yang tersembunyi bagi-Nya, dan bahwa amal perbuatan mereka tidak akan sia-sia atau diabaikan oleh-Nya. Kesadaran ini memberikan motivasi ekstra bagi mereka untuk berusaha semakin keras melakukan amal perbuatan yang benar dan bermanfaat pada hari Kiamat (Ghoffar, 2000).

Dari tafsir diatas dapat dijelaskan pentingnya introspeksi diri dan *muhasabah*. Setiap individu diingatkan untuk selalu mengkaji dirinya sendiri. Jika terdapat kesalahan, mereka harus segera memperbaikinya dengan melepaskan diri

dari dosa, bertaubat sungguh-sungguh, dan menjauh dari segala hal yang dapat mengantarkan pada kesalahan tersebut. Jika mereka menilai bahwa mereka kurang rajin dalam menjalankan perintah Allah, mereka harus berupaya dengan sungguh-sungguh, memohon pertolongan kepada Allah untuk memperbaiki diri, dan membandingkan karunia dan kebaikan Allah dengan kemalasan mereka, sehingga mereka merasa malu.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini berfokus pada penggunaan metode *Random Forest* untuk memprediksi harga rumah dengan mempertimbangkan berbagai rasio split data, mulai dari 90:10 hingga 50:50. Hasil penelitian menunjukkan bahwa nilai Mean Absolute Error (MAE) dan Mean Absolute Percentage Error (MAPE) bervariasi tergantung pada proporsi data training dan testing. Pada Ratio 70:30 hasil pengujian mendapatkan nilai terbaik yaitu sebesar 106286529 pada nilai MAE dan 11.29% pada nilai MAPE. Sejalan dengan harapan, nilai MAE cenderung meningkat dengan peningkatan rasio data testing, menunjukkan bahwa model mungkin mengalami kesulitan dalam menggeneralisasi pola pada data yang lebih besar. Meskipun terdapat kecenderungan bahwa MAPE meningkat pada rasio yang lebih tinggi, fluktuasi signifikan pada rasio 50:50 menyoroti pentingnya jumlah data testing yang memadai untuk menghindari bias. Analisis ini memberikan wawasan yang berharga terkait dengan kinerja model *Random Forest* pada berbagai kondisi data, dan sementara *Random Forest* dapat memberikan prediksi harga rumah yang cukup baik, perlu diingat bahwa evaluasi yang lebih lanjut dengan dataset yang lebih besar dan beragam diperlukan untuk memvalidasi dan mengoptimalkan model.

Visualisasi GIS memiliki peran krusial dalam mempermudah pemahaman dan pengoperasian data geografis terkait harga rumah. Dengan memanfaatkan peta interaktif, kita dapat memvisualisasikan distribusi harga rumah di berbagai lokasi.

Visualisasi GIS tidak hanya meningkatkan pemahaman terhadap pola harga rumah, tetapi juga memfasilitasi pengambilan keputusan yang lebih tepat terkait investasi properti dan perencanaan urban.

5.2 Saran

Penelitian ini dapat ditingkatkan dengan ekspansi dataset untuk mencakup lebih banyak variabel dan data properti yang relevan, meningkatkan kualitas prediksi model. Jika variabel "indikasi nilai" tidak tersedia, teknik imputasi dapat diterapkan. Selain itu, analisis kualitatif, pemilihan fitur yang lebih relevan, dan penelitian lanjutan di area tanpa variabel "indikasi nilai" dapat memperkuat model. Ini membutuhkan pendekatan holistik untuk memastikan model machine learning dapat memberikan prediksi harga rumah yang konsisten dan akurat, bahkan dalam situasi yang kurangnya data.

DAFTAR PUSTAKA

- Afnarius, S., Yuliani, F., & Akbar, F. (2017). Aplikasi Sistem Informasi Geografi (Sig) Tempat Ibadah Dan Cara Aksesnya Berbasis Web Dan Mobile Di Kota Bukittinggi. *Research Report*, 382–387.
- Akhir, T., Widyastuti, R. W., Statistika, P. S., Matematika, F., Ilmu, D. A. N., Alam, P., & Indonesia, U. I. (2020). *Framework Flask*.
- Charbuty, B., & Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2(01), 20–28. <https://doi.org/10.38094/jastt20165>
- Dahnir, Yamin, M., Syahrin, A., & Sutiarnoto. (2020). Peranan Kantor Jasa Penilai Publik (KJPP) Dalam Kegiatan Pengadaan Tanah Untuk Kepentingan Umum (Antara Profesionalisme Dan Perbuatan Melawan Hukum/ Tindak Pidana Korupsi). *Iuris Studia: Jurnal Kajian Hukum*, 3(3), 255–263.
- de Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. *Neurocomputing*, 192, 38–48. <https://doi.org/10.1016/j.neucom.2015.12.114>
- Eka Dwi Cahyono. (2001). Sistem Informasi Geografis Angkutan Umum Di Surabaya Berbasis Web. *1Sistem Informasi Geografis Angkutan Umum Di Surabaya Berbasis Web*, 2–2.
- Ghoffar, A. (2000). *Tafsir Ibnu Katsir 1 a.pdf* (pp. 123–125).
- Harahap, B., Umbara, R. F., & Triantoro, D. (2015). Prediksi Indeks Harga Saham Menggunakan Metode Gabungan Support Vector Regression dan Random Forest. *E-Proceeding of Engineering*, 3(3), 5353–5363.
- Kartika, W., & Oktarini, I. D. (2022). Pengaruh Biaya Operasional Terhadap Produktivitas Tenaga Kerja Di Kjpp Herman Meirizki Surabaya. *Jurnal Revenue : Jurnal Ilmiah Akuntansi*, 3(1), 98–106. <https://doi.org/10.46306/rev.v3i1.59>
- Maharani, S., Harsa K, A., & Nalarwati, A. T. (2017). Sistem Informasi Geografi (SIG) Pencarian ATM Bank Kaltim Terdekat Dengan Geolocation Dan Haversine Formula Berbasis Web. *Jurnal Infotel*, 9(1), 1. <https://doi.org/10.20895/infotel.v9i1.139>
- Patel, H. H., & Prajapati, P. (2018). Study and Analysis of Decision Tree Based Classification Algorithms. *International Journal of Computer Sciences and Engineering*, 6(10), 74–78. <https://doi.org/10.26438/ijcse/v6i10.7478>
- Pratomo, A., Febrian Umbara, R., & Rohmawati, A. A. (2019). Prediksi Pergerakan Harga Saham dengan Metode Random Forest Menggunakan Trend Deterministic Data Preparation (Studi Kasus Saham Perusahaan PT Astra

- International Tbk, PT Garuda Indonesia Tbk, dan PT Indosat Tbk). *E-Proceeding of Engineering*, 6(1), 2545. <https://finance.yahoo.com/>
- Reza, V., Snapp, P., Dalam, E., Di, I. M. A., Socialization, A., Cadger, O. F., To, M., Cadger, S., Programpadang, R., Hukum, F., Hatta, U. B. U. B., Sipil, F. T., Hatta, U. B. U. B., Danilo Gomes de Arruda, Bustamam, N., Suryani, S., Nasution, M. S., Prayitno, B., Rois, I., ... Rezekiana, L. (2020). Tinjauan Yuridis Terhadap Perusahaan Jasa Penilai (Appraisal Company) [Studi. *Business Law Binus*, 7(2), 33–48. http://repository.radenintan.ac.id/11375/1/PERPUS_PUSAT.pdf<http://business-law.binus.ac.id/2015/10/08/pariwisata-syariah/><https://www.ptonline.com/articles/how-to-get-better-mfi-results/><https://journal.uir.ac.id/index.php/kiat/article/view/8839>
- Rivai, S. (2023). Computer Based Information System Journal Perancangan Aplikasi Perkantoran Elektronik Dengan Menggunakan Metode Object Oriented Analysis Design Berbasis WEB Pada KJPP DAR. *Cbis Journal*, 11(01), 42–54. <http://ejournal.upbatam.ac.id/index.php/cbis><http://ejournal.upbatam.ac.id/index.php/cbis>
- Saadah, S., & Salsabila, H. (2021a). Prediksi Harga Bitcoin Menggunakan Metode Random Forest. *Jurnal Politeknik Caltex Riau*, 7(1), 24–32.
- Saadah, S., & Salsabila, H. (2021b). Prediksi Harga Bitcoin Menggunakan Metode Random Forest (Studi Kasus: Data Acak Pada Awal Masa Pandemic Covid-19). *Jurnal Komputer Terapan*, 7(Vol. 7 No. 1 (2021)), 24–32. <https://doi.org/10.35143/jkt.v7i1.4618>
- Santra, a. K., & Christy, C. J. (2012). Genetic Algorithm and Confusion Matrix for Document Clustering. *International Journal of Computer Science*, 9(1), 322–328. <http://ijcsi.org/papers/IJCSI-9-1-2-322-328.pdf>
- Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, 27(2), 130–135. <https://doi.org/10.11919/j.issn.1002-0829.215044>
- Wismarini, T. D., & Ningsih, D. H. U. (2010). Analisis Sistem Drainase Kota Semarang Berbasis Sistem Informasi Geografi dalam Membantu Pengambilan Keputusan bagi Penanganan Banjir. *Jurnal Teknologi Informasi DINAMIK*, XV(1), 41–51.