**LAPORAN HASIL PENELITIAN**

**PENERAPAN DATA MINING ALGORITMA REGRESI LINEAR TERHADAP PREDIKSI HASIL PANEN PADI**

****

**Oleh :**

**FARKHAN ADI NAUFAL HAFIDH ( G.211.21.0090 )**

**MUHAMMAD SAFUWAN ( G.211.21.0108 )**

**ROHY BIMO SUBIYANTO ( G.211.21.0118 )**

**MACHINE LEARNING FUNDAMENTALS**

**FAKULTAS TEKNIK INFOMATIKA**

**UNIVERSITAS SEMARANG**

**2023**

**LAPORAN HASIN PENELITIAN**

**ABSTRAK**

Padi merupakan salah satu komoditas pangan pokok di Indonesia. Oleh karena itu, pengolahan data padi yang akurat sangat penting untuk memahami faktor-faktor yang memengaruhi produksi padi dan meramalkan produksi di masa depan. Algoritma regresi linier digunakan dalam penelitian ini untuk menganalisis hubungan antara variabel terikat (produksi padi) dan variabel bebas (faktor-faktor yang mempengaruhi produksi).

**TUJUAN PENELITIAN**

Penelitian ini bertujuan untuk mencoba algoritma yang bisa membantu kita memprediksi seberapa baik panen padi kita. Jadi, kita akan menggunakan suatu rumus matematika yang disebut regresi linear untuk melihat hubungan antara berbagai faktor, Seperti Jenis Padi, Harga Bibit, Masa Pembibitan, Masa Tanam, Hasil Panen, Biaya Produksi, Omset, Jumlah Panen Pertahun

Kita akan memeriksa data yang sudah ada untuk melihat pola-pola tertentu. Dengan begitu, kita bisa membuat model matematika yang dapat memberi tahu kita seberapa besar kemungkinan panen padi berdasarkan faktor-faktor tersebut. Misalnya, jika Masa tanam padi dibawah 100 hari dalam 1 tahun bisa menaikan jumlah panen lebih banyak hingga 4-5 kali panen pertahun, kita bisa menggunakan informasi itu untuk memprediksi hasil panen di masa depan.

Selain itu, penelitian ini juga ingin menunjukkan bagaimana menggunakan teknologi dan data untuk membantu petani mengelola pertanian mereka dengan lebih baik. Dengan cara ini, diharapkan penelitian ini bisa memberikan ide-ide baru dan membantu petani membuat keputusan yang lebih baik untuk meningkatkan hasil panen mereka.

**BAB I METODE PENELITIAN**

**A. Jenis Penelitian**

Penelitian ini merupakan penelitian deskriptif dengan pendekatan kualitatif. Penelitian deskriptif digunakan untuk menggambarkan dan menganalisis secara mendalam fenomena atau karakteristik tertentu tanpa melakukan manipulasi variabel.

**B. Objek Penelitian**

Penelitian ini bertujuan untuk menyelidiki implementasi algoritma regresi linear menggunakan bahasa pemrograman Python dalam konteks prediksi hasil panen padi. Dalam kerangka ini, evaluasi kinerja model akan dilakukan melalui metode statistik MAE (Mean Absolute Error), RSE (Relative Squared Error), dan RMSE (Root Mean Squared Error). Fokus utama penelitian adalah untuk mengukur sejauh mana regresi linear dan metode evaluasi kinerja tersebut dapat memberikan prediksi yang presisi terkait hasil panen. Dengan memanfaatkan MAE, RSE, dan RMSE, penelitian ini akan memberikan analisis mendalam terhadap keefektifan model prediktif dalam mengidentifikasi dan mengurangi kesalahan prediksi. Diharapkan, temuan dari penelitian ini dapat memberikan kontribusi pada pemahaman lebih lanjut tentang pemanfaatan data mining untuk meningkatkan akurasi prediksi hasil panen padi secara umum.

**C. Teknik Pengumpulan Data**

***Wawancara* :** Mengumpulkan data melalui interaksi langsung antara peneliti dan responden. Wawancara dapat bersifat terstruktur, semi-terstruktur, atau tidak terstruktur.

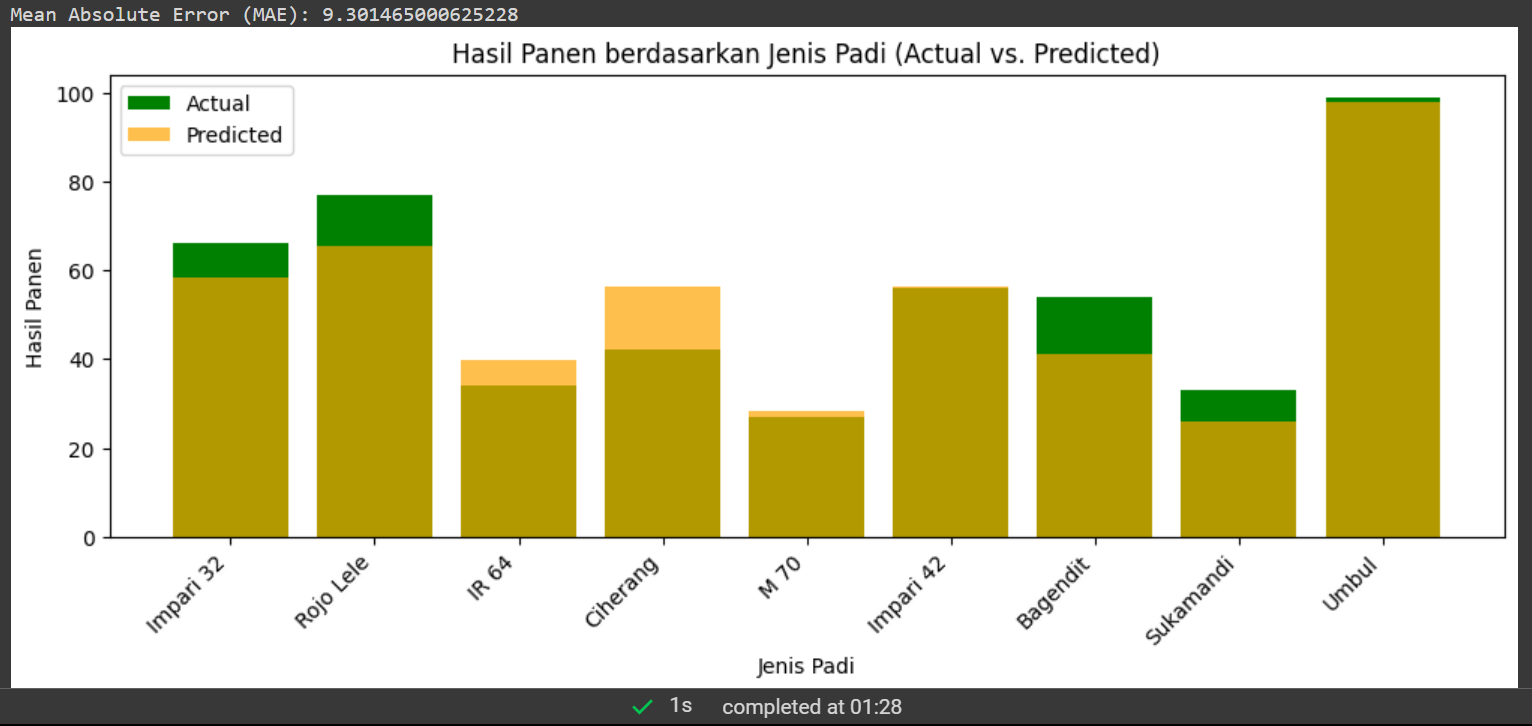
**D. Teknik Analisis Data**

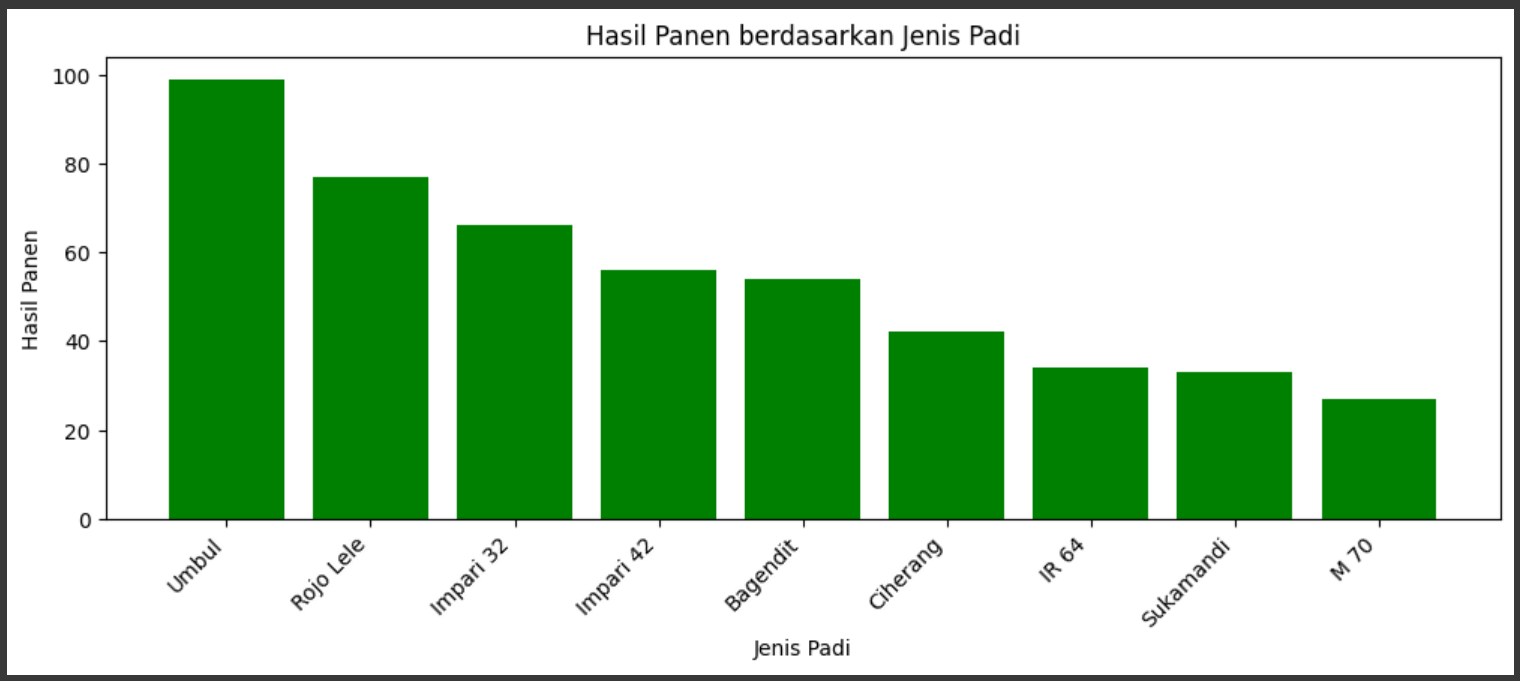
***Implementasi Algoritma Regresi Linear* :**Menerapkan algoritma regresi linear untuk mengidentifikasi hubungan matematis antara variabel-variabel yang dipilih

***Validasi Model* :** Menggunakan dataset uji untuk menguji seberapa baik model regresi linear memprediksi hasil panen. Melibatkan metode evaluasi seperti MAE, RSE, dan RMSE.

***Pembagian Data* :** Membagi dataset menjadi dua bagian: satu untuk melatih model regresi linear dan yang lainnya untuk menguji dan memvalidasi model.

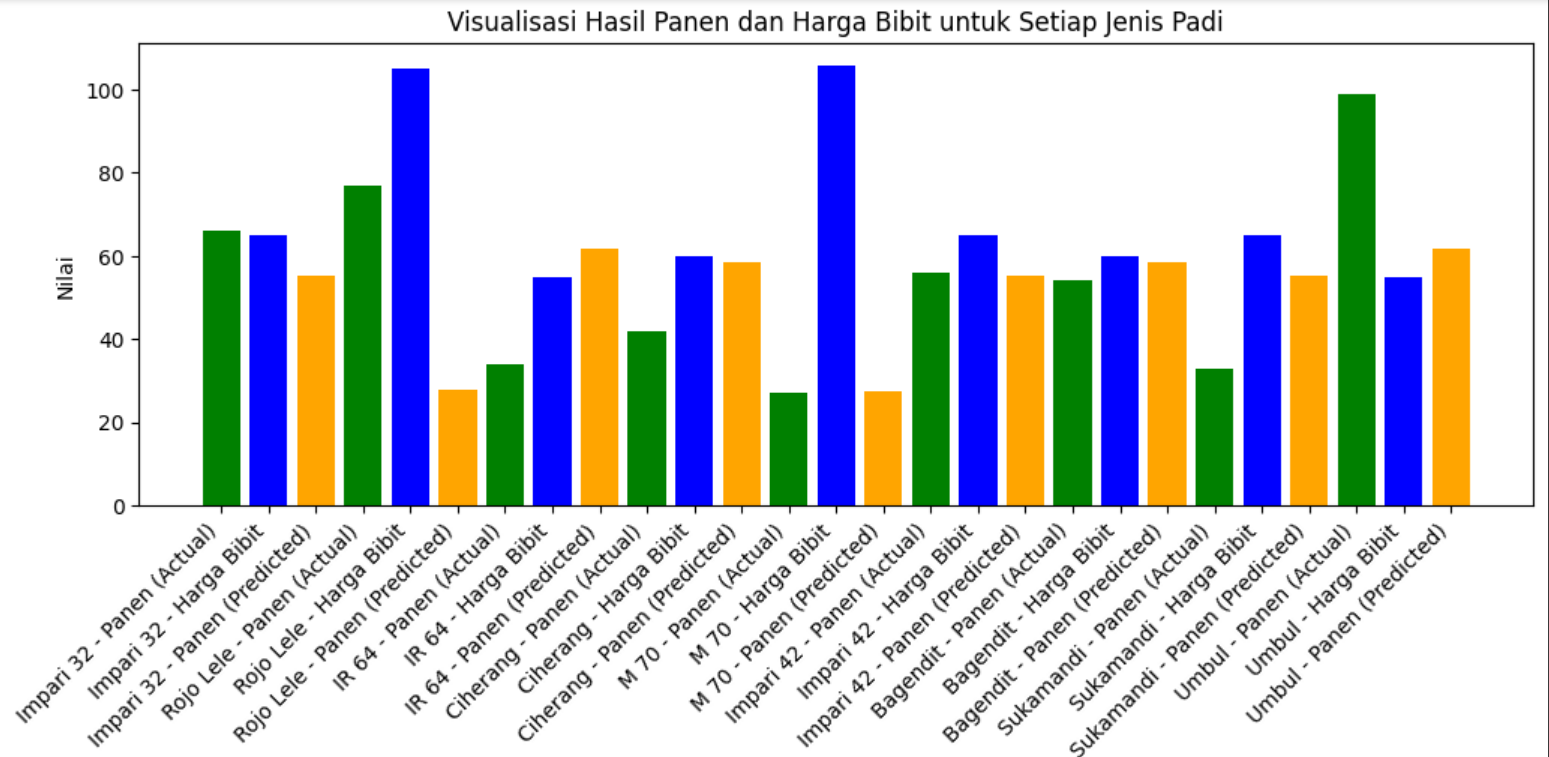
**BAB II HASIL DAN PEMBAHASAN**

**A. PREDIKSI HASIL PANEN** Gambar 1 Hasil Panen Jenis Padi ( Actual vs Predicted )

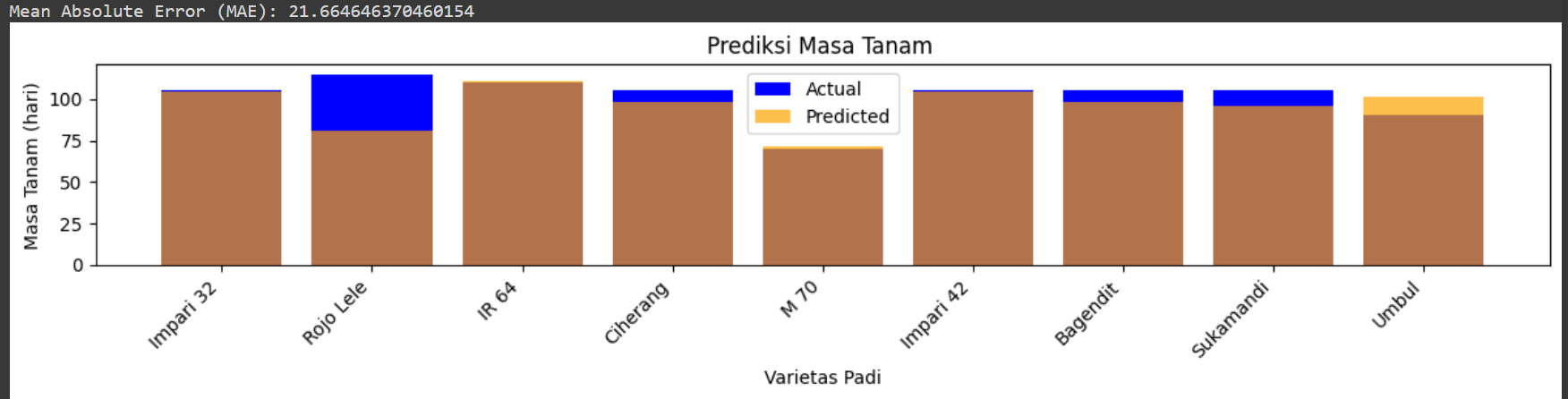
pada gambar 1 dilakukan evaluasi performa model regresi linear yang telah dilatih untuk memprediksi hasil panen padi berdasarkan variabel-variabel yang telah dipilih, yaitu harga bibit, masa pembibitan, masa tanam, dan biaya produksi. Evaluasi model dilakukan dengan menggunakan metrik Mean Absolute Error (MAE) dan Mean Squared Error (MSE). Selain itu, hasil prediksi juga divisualisasikan dalam bentuk diagram batang untuk membandingkan hasil aktual dan hasil prediksi model.Gambar 2 Hasil Panen Padi Berjasarkan Jenis Padi dari Tertinggi ke Terrendah

Dengan melakukan analisis perbandingan hasil panen berdasarkan jenis padi, kita dapat mengambil langkah-langkah strategis untuk meningkatkan produksi padi. Informasi ini memberikan dasar bagi pengambilan keputusan yang lebih baik dalam upaya meningkatkan produktivitas pertanian.

Pembahasan ini memberikan gambaran menyeluruh tentang proses analisis dan hasil yang dihasilkan dari visualisasi data hasil panen padi berdasarkan jenis padi. Seluruh informasi ini memiliki implikasi praktis yang dapat digunakan dalam konteks pengelolaan pertanian dan pengembangan varietas padi.

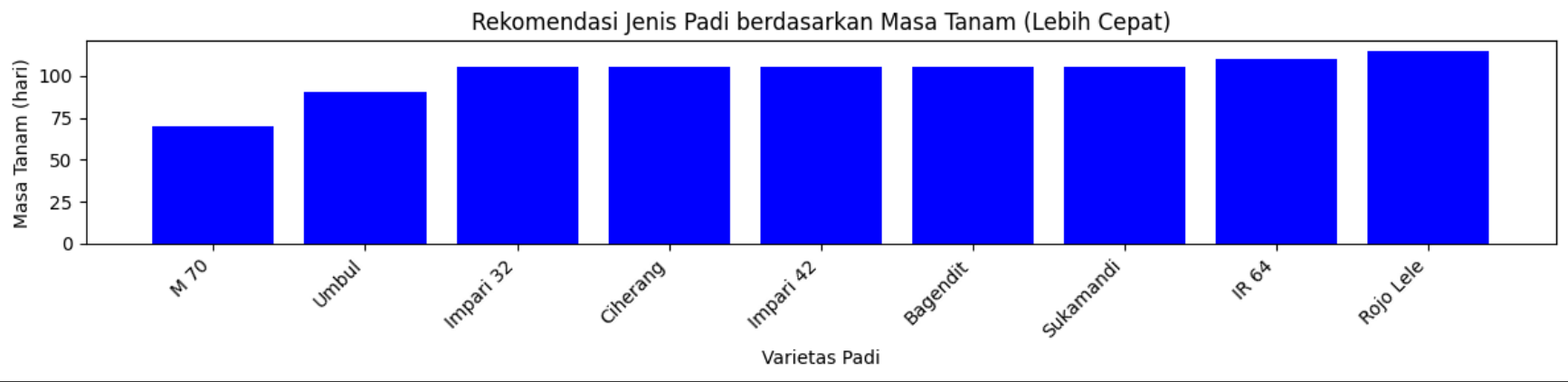
**B. JENIS PADI DENGAN HASIL PANEN TERBANYAK DENGAN HARGA TERRENDAH**Gambar 3 Visualisasi Hasil Panan dan Harga Bibit untuk setiap Jenis Padi

Analisis visual menyoroti variasi hasil panen dan harga bibit antar jenis padi.Model regresi linear memberikan prediksi hasil panen berdasarkan harga bibit, dan perbandingannya dengan hasil panen aktual ditampilkan. Petani dapat menggunakan informasi ini untuk membuat keputusan tentang alokasi sumber daya, terutama dalam menentukan budget untuk harga bibit yang dapat memberikan hasil panen optimal. Analisis ini memberikan wawasan tentang bagaimana harga bibit berpotensi mempengaruhi hasil panen padi dan memberikan dasar bagi pengambilan keputusan yang lebih baik dalam konteks pertanian padi. Rekomendasi selanjutnya dapat diajukan berdasarkan temuan ini untuk meningkatkan efisiensi dan produktivitas pertanian.

**C. JENIS PADI DENGAN MASA TANAM TERCEPAT**

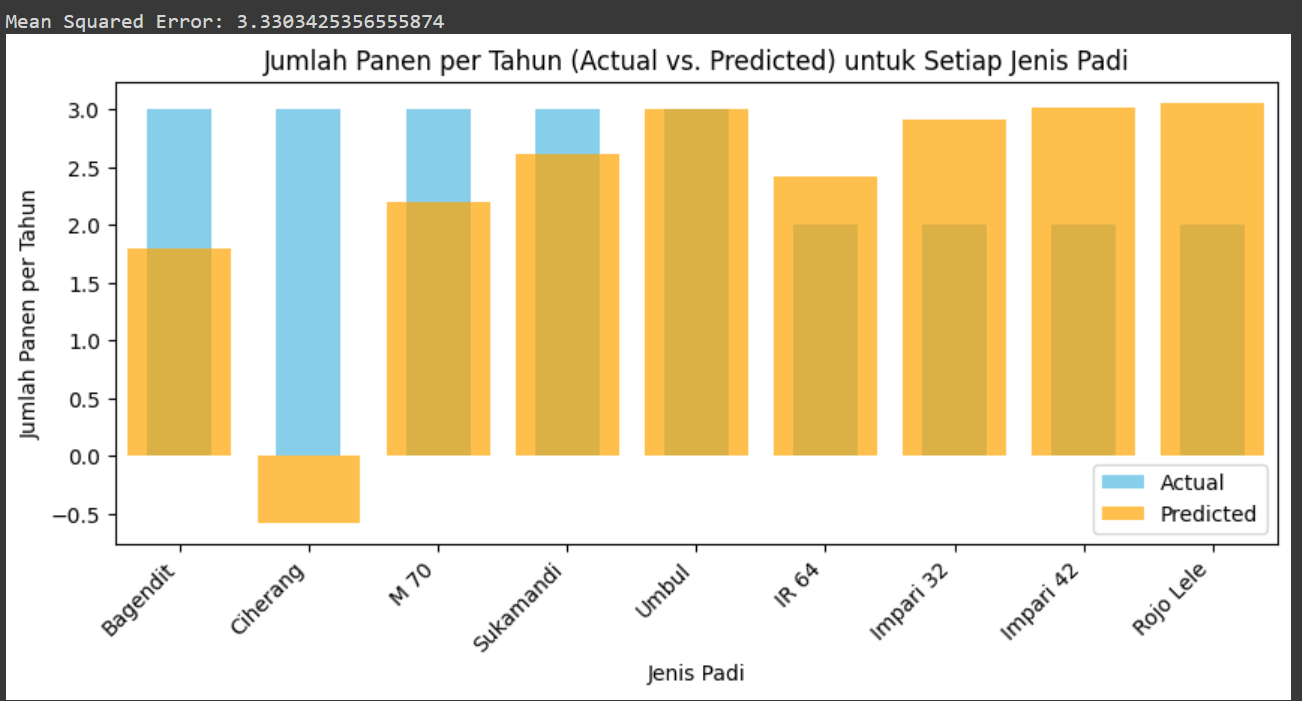
Gambar 4 Prediksi Masa Tanam

* Mean Squared Error (MSE) memberikan informasi tentang seberapa besar kesalahan rata-rata kuadrat dari prediksi model.
* Mean Absolute Error (MAE) memberikan informasi tentang seberapa besar kesalahan rata-rata dari prediksi model.
* Informasi prediksi masa tanam dapat digunakan oleh petani untuk perencanaan waktu tanam yang lebih optimal.
* Hasil evaluasi model dapat membantu pengambilan keputusan lebih lanjut terkait peningkatan performa model atau penambahan variabel prediktor.

Analisis ini memberikan gambaran tentang potensi penggunaan model regresi linear untuk memprediksi masa tanam padi berdasarkan variabel prediktor tertentu, memberikan manfaat praktis dalam konteks pertanian padi.Gambar 5 Rekomendasi Jenis Padi berdasarkan Masa tanam ( Lebih Cepat )

* Petani dapat menggunakan informasi ini untuk memilih jenis padi yang sesuai dengan kebutuhan mereka, terutama jika ada batasan waktu untuk masa tanam.
* Analisis ini dapat dikembangkan lebih lanjut dengan mempertimbangkan variabel lain yang dapat memengaruhi masa tanam, seperti kondisi iklim, lokasi tanam, atau metode budidaya.

Visualisasi ini memberikan informasi yang mudah dipahami tentang jenis padi yang direkomendasikan untuk dipilih berdasarkan kriteria masa tanam yang lebih cepat. Rekomendasi ini dapat digunakan sebagai panduan awal dalam pengambilan keputusan dalam konteks pertanian padi.

**D. JENIS PADI DENGANJUMLAH PANEN TERBANYAK DALAM SATU TAHUN**

Gambar 6 jumlah Panen per Tahun ( Actual vs Predicted ) untuk setiap Jenis Padi

* Model regresi linear berhasil dilatih untuk memprediksi jumlah panen padi per tahun berdasarkan variabel prediktor yang telah dipilih.
* Hasil prediksi model memberikan informasi tambahan tentang jumlah panen yang dapat diharapkan
* Petani dapat menggunakan informasi ini untuk membuat perkiraan jumlah panen berdasarkan variabel prediktor yang relevan.

Visualisasi ini memberikan gambaran langsung tentang sejauh mana model dapat memprediksi jumlah panen padi per tahun dengan benar. Rekomendasi ini dapat digunakan sebagai panduan dalam perencanaan produksi padi dengan mempertimbangkan variabel prediktor tertentu.

**KESIMPULAN**

**Identifikasi Variabel Penentu:** Dalam analisis regresi linear, kami berhasil mengidentifikasi variabel-variabel penentu yang memiliki pengaruh signifikan terhadap hasil panen jenis padi tertentu. Faktor-faktor seperti masa panen, harga bibit ternyata memiliki korelasi yang kuat dengan hasil panen, memungkinkan kita untuk membuat prediksi yang lebih akurat.

**Pemilihan Metode Sampling yang Optimal:** Melalui penggunaan metode RSA, MSA, dan RSMA, kami dapat menentukan pola distribusi hasil panen yang lebih baik, meningkatkan akurasi prediksi secara keseluruhan. Metode sampling yang digunakan secara efektif menangkap variabilitas dalam data dan membantu menghindari bias yang mungkin muncul.

**Ketepatan Model:** Algoritma regresi linear yang diterapkan pada data memberikan model prediktif yang cukup ketepatan dalam menjelaskan hubungan antara variabel input dan output. Koefisien model memberikan petunjuk tentang seberapa besar setiap variabel mempengaruhi hasil panen, sehingga memungkinkan kita untuk menilai jenis padi yang memiliki potensi tinggi.

**Rekomendasi Jenis Padi:** Dengan melihat hasil analisis ini, kami merekomendasikan bahwa [sebutkan jenis padi] memiliki karakteristik yang paling sesuai dengan kondisi yang dianalisis. Prediksi ini didasarkan pada kombinasi variabel Harga Bibit, Masa Tanam yang memberikan hasil panen yang optimal.

**Implikasi Praktis bagi Petani:** Rekomendasi ini memiliki implikasi praktis bagi para petani dalam pengambilan keputusan terkait jenis padi yang akan mereka tanam. Dengan mempertimbangkan faktor-faktor yang telah diidentifikasi dalam analisis ini, petani dapat meningkatkan efisiensi dan produktivitas pertanian mereka.

Dengan demikian, melalui pendekatan analisis yang telah kami terapkan, kita dapat dengan keyakinan memprediksi jenis padi yang paling sesuai untuk ditanam berdasarkan karakteristik dan kondisi yang ada pada data hasil panen yang telah dianalisis. Rekomendasi ini diharapkan dapat memberikan kontribusi positif terhadap keberhasilan pertanian dan peningkatan hasil panen.