

BAB I PENDAHULUAN

1.1 Latar Belakang

Berlian adalah sebuah mineral yang mengendap selama bertahun-tahun yang membuatnya mempunyai nilai harga yang tinggi. Berlian dianggap sebagai batu permata yang merupakan salah satu batu permata yang lebih mahal dibandingkan batu permata lainnya[1]. Berlian memiliki sifat optik luar biasa yang menjadikannya semakin berguna untuk optik *multispectral*. Selain itu, ketahanan berlian, keindahan, warna, tekstur, dan bentuk, membuat kombinasi dari semua karakteristik tersebut berlian semakin populer karena menunjukkan kilau dan kemewahannya.

Semakin banyak orang yang tertarik pada berlian, para penggemar rela membelinya dengan harga jutaan. Namun, di sektor perdagangan berlian, pembeli dan investor menghadapi beberapa kesulitan dalam meramalkan harga batu berlian. Kesulitan ini disebabkan oleh variasi bentuk, ukuran, dan kemurnian batu. Hal ini menyebabkan banyak penggemar maupun investor berlian membeli batu dengan harga yang terlalu tinggi, dan saat ingin menjualnya kembali, harganya terlalu rendah, sehingga mengakibatkan kerugian [2].

Berdasarkan masalah diatas, salah satu cara yang dapat digunakan untuk memudahkan penggemar dan investor adalah dengan mengimplementasikan teknologi machine learning dengan algoritma Linear Regression, Decision Tree Regression, dan Random Forest Regression. Model prediksi dari algoritma tersebut mampu menganalisis data historis transaksi berlian lebih akurat. Hal ini memberikan panduan harga yang wajar bagi penggemar dan investor [3].

Berdasarkan uraian diatas, peneliti akan melakukan penelitian untuk mendapatkan prediksi harga berlian menggunakan tiga algoritma, yaitu Linear Regression, Decision Tree Regression, dan Random Forest Regression. Model yang akan dihasilkan akan dievaluasi menggunakan metrik MAE, RMSE, dan R2-score. Sehingga menghasilkan pemodelan prediksi harga berlian yang memiliki evaluasi

terbaik, dengan demikian memberikan informasi yang akurat dan relevan bagi para pelaku pasar dalam perdagangan berlian.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang sudah dijelaskan, terdapat beberapa rumusan masalah yang dapat diidentifikasi dalam konteks perbandingan algoritma machine learning untuk memprediksi harga berlian. Rumusan masalah tersebut adalah sebagai berikut:

1. Manakah hasil yang terbaik dari perbandingan algoritma Linear Regression, Decision Tree Regression, dan Random Forest Regression dalam memprediksi harga berlian?
2. Apakah ada perbedaan yang signifikan dalam akurasi prediksi harga berlian antara ketiga algoritma tersebut?
3. Berapa tingkat akurasi metode Linear Regression, Decision Tree Regression, dan Random Forest Regression dalam prediksi harga berlian?

1.3 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan masalah yang perlu diperhatikan. Batasan-batasan tersebut adalah sebagai berikut:

1. Penelitian akan fokus pada algoritma Linear Regression, Decision Tree Regression, dan Random Forest Regression dalam membandingkan algoritma tersebut pada prediksi harga berlian.
2. Data yang digunakan dalam penelitian ini dapat dari Kaggle <https://www.kaggle.com/datasets/sibelius5/diamond-prices>. Data yang relevan pada dataset diatas untuk prediksi harga berlian terdiri dari atribut seperti diamond price, inflation rate, interest rate, fed rate, dan gold price.
3. Pengujian evaluasi menggunakan metrik MAE, RMSE, dan R2-score.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah menghasilkan sebagai berikut:

1. Tujuan utama dari peneliti ini adalah membandingkan algoritma Linear Regression, Decision Tree Regression, dan Random Forest Regression mana yang terbaik untuk memodelkan dan memprediksi harga berlian pada semua data.
2. Penelitian ini bertujuan untuk mengidentifikasi kelemahan dan keunggulan masing-masing algoritma. Sehingga membantu dalam pemahaman lebih mendalam tentang di mana setiap algoritma berkinerja baik dan di mana mungkin kurang efektif.
3. Melalui penelitian ini, peneliti dapat memahami faktor-faktor yang paling berpengaruh dalam prediksi harga berlian. Termasuk fitur-fitur berlian yang paling penting dalam menentukan harganya.
4. Penelitian ini bertujuan untuk memberikan rekomendasi praktis kepada orang awam, peniliti berlian, penggemar, maupun investor, tentang algoritma mana yang sebaiknya digunakan untuk prediksi harga berlian.

1.5 Manfaat Penelitian

Hasil penelitian diharapkan dapat memberikan manfaat baik segi teoritis maupun praktis atau bagi peneliti selanjutnya, seperti sebagai berikut:

1. Hasil penelitian ini dapat membantu harga pasar saham dalam mengembangkan strategi untuk meningkatkan tingkat harga berlian di masa yang akan datang.
2. Penelitian ini memberikan kesempatan untuk mendalami konsep-konsep dalam bidang pembelajaran mesin dan kecerdasan buatan. Sehingga dapat menghasilkan peningkatan dalam pemahaman tentang algoritma yang berbeda.
3. Hasil dari penelitian ini dapat membantu dalam mengoptimalkan prediksi harga berlian. Dengan mengetahui algoritma mana yang paling efektif, penentuan harga berlian dapat menjadi lebih akurat, yang dapat menguntungkan dalam penjualan berlian dan penilaian aset.
4. Hasil dari penelitian ini dapat menjadi landasan untuk penelitian lebih lanjut dalam domain yang sama atau terkait, misalnya, mempertimbangkan

algoritma pembelajaran mesin yang lebih canggih atau faktor-faktor tambahan yang mempengaruhi harga berlian.

1.6 Sistematika Penulisan

Dalam proses penulisan skripsi, telah dibagi menjadi 5 bab, masing masing dengan deskripsi berikut :

- a) **BAB I Pendahuluan**
Bab ini menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.
- b) **BAB II Tinjauan Pustaka**
Bab ini menjelaskan penelitian terdahulu, kerangka teori serta landasan teori yang digunakan pada penelitian ini.
- c) **BAB III Metode Penelitian**
Bab ini menjelaskan jenis penelitian, metode, serta membahas alur penelitian, serta memberikan penjelasan alat dan bahan yang digunakan dalam penelitian.
- d) **BAB IV Hasil Dan Pembahasan**
Bab ini menjelaskan hasil analisis, pembahasan code, pemilihan fitur, implementasi algoritma, dan hasil evaluasi yang didapatkan.
- e) **BAB V Penutup**
Bab ini menjelaskan saran dan kesimpulan serta rekomendasi penelitian untuk dikembangkan jauh lebih baik lagi.