

NADHIF RIF'AT RASENDRIYA

AI/ML PROJECT

SALARY PREDICTION







MACHINE LEARNING (ML)



Machine learning (ML) adalah cabang dari kecerdasan buatan (artificial intelligence/AI) yang memungkinkan komputer untuk belajar dari data tanpa diprogram secara eksplisit. Dalam machine learning, algoritma digunakan untuk menganalisis data, mengenali pola, dan membuat prediksi atau keputusan berdasarkan data yang ada.

Berikut adalah beberapa konsep dasar dalam machine learning:

1. **Data:** Machine learning membutuhkan data untuk belajar. Data ini bisa berupa angka, teks, gambar, suara, atau informasi lainnya.
 2. **Model:** Model adalah representasi matematis dari masalah yang ingin diselesaikan. Model ini dilatih menggunakan data untuk mempelajari pola-pola tertentu.
 3. **Algoritma:** Algoritma adalah metode yang digunakan untuk melatih model agar dapat mengenali pola dalam data.
 4. **Training:** Proses melatih model menggunakan data yang telah diberi label (supervised learning) atau tanpa label (unsupervised learning).
 5. **Testing:** Setelah model dilatih, model diuji dengan data baru untuk mengukur seberapa baik model tersebut dalam membuat prediksi.
- 
- 

JENIS JENIS MACHINE LEARNING

Machine learning dapat dibagi menjadi beberapa jenis, antara lain:

- **Supervised Learning:** Metode ini memerlukan data latih sebagai panduan dalam pembelajaran mesin. Algoritma dilatih dengan data yang sudah memiliki label (jawaban yang benar). Contohnya adalah klasifikasi gambar atau prediksi harga rumah.
- **Unsupervised Learning:** Algoritma dilatih dengan data yang tidak memiliki label. Tujuan utamanya adalah untuk menemukan pola atau struktur dalam data. Contoh: pengelompokan (clustering).
- **Reinforcement Learning:** Algoritma belajar melalui trial and error, mendapatkan umpan balik dalam bentuk reward atau punishment berdasarkan tindakan yang diambil.

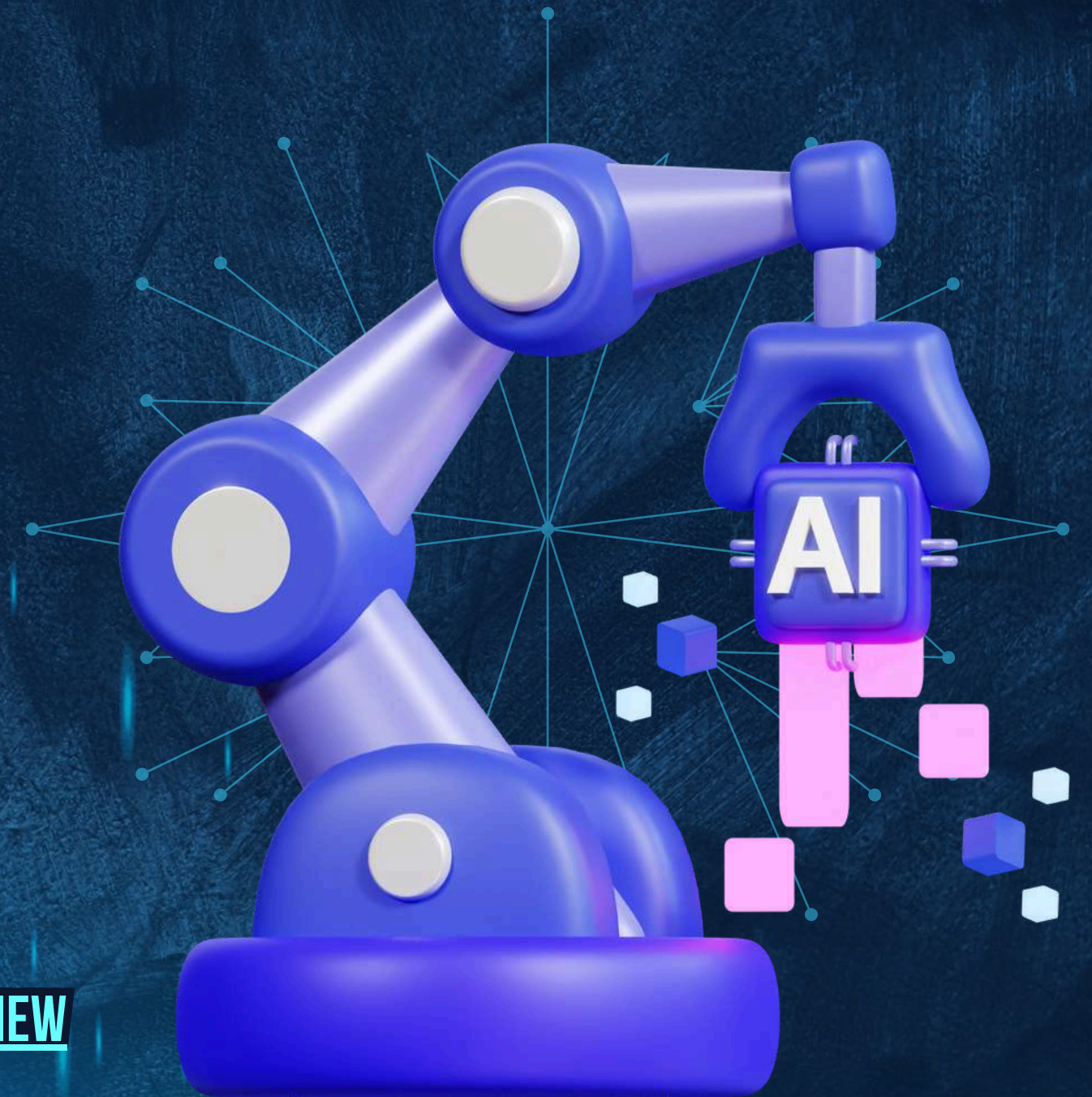


SUPERVISED LEARNING

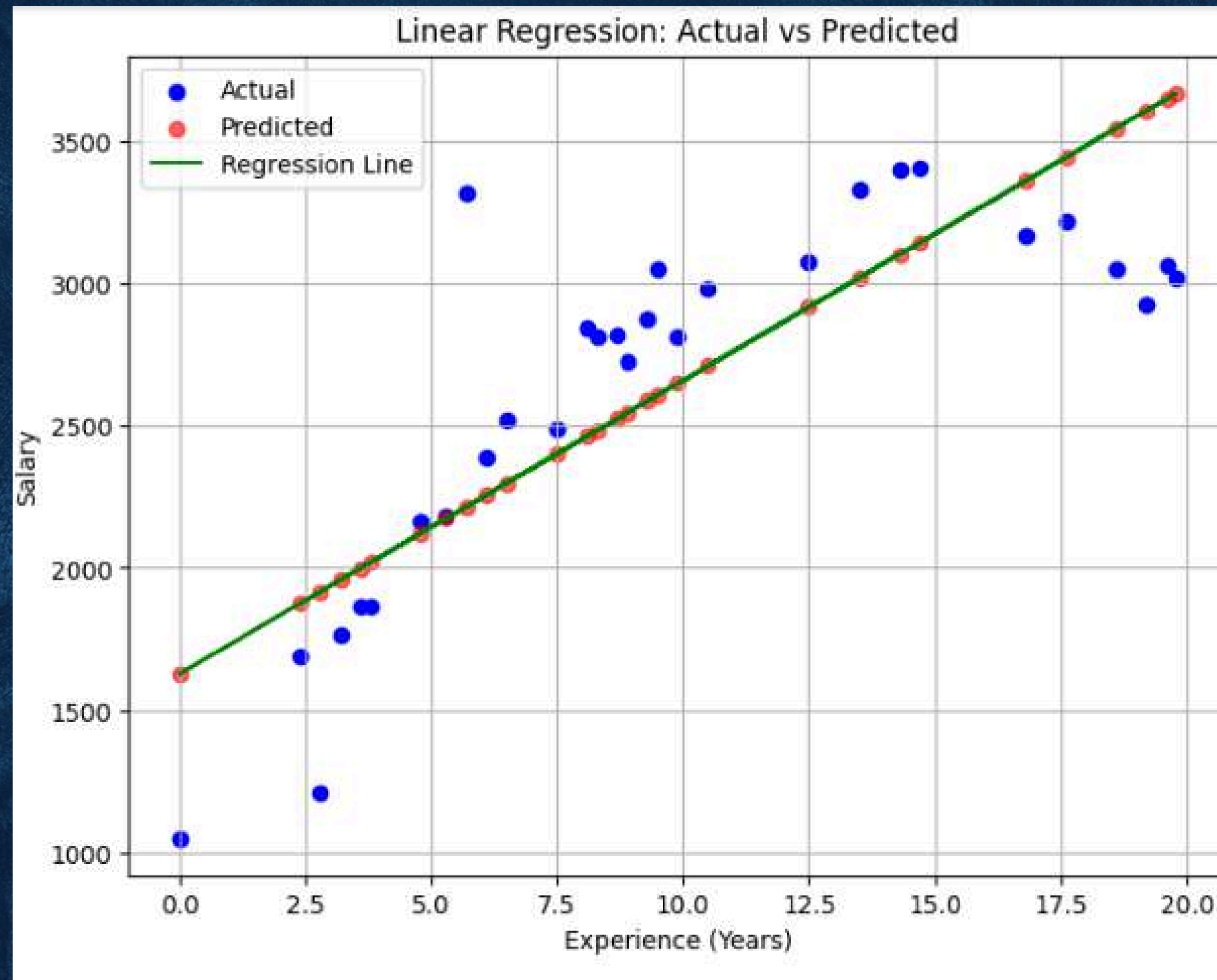
SALARY PREDICTION

DATASET:

[HTTPS://DRIVE.GOOGLE.COM/FILE/D/1FOESCJLEPAZKQJXBNCJHHU6P3PE7LHF0/VIEW](https://drive.google.com/file/d/1FOESCJLEPAZKQJXBNCJHHU6P3PE7LHF0/view)



LINEAR REGRESSION



LINEAR REGRESSION RESULTS:

Mean Squared Error

Train: 91704.29

Test: 160984.16

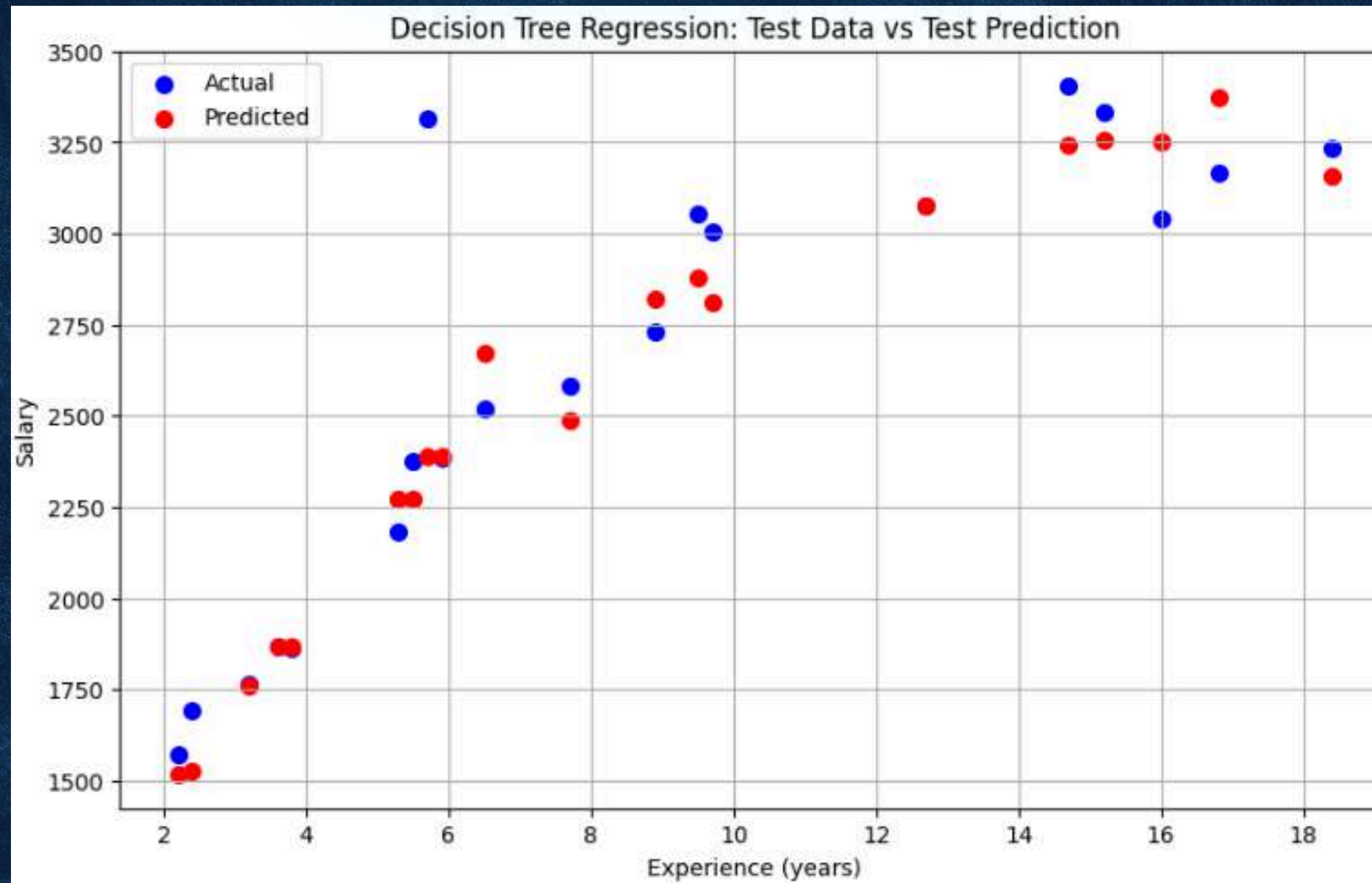
R^2 Score

Train: 0.796

Test: 0.595



DECISION TREE



DECISION TREE RESULTS:

Mean Squared Error

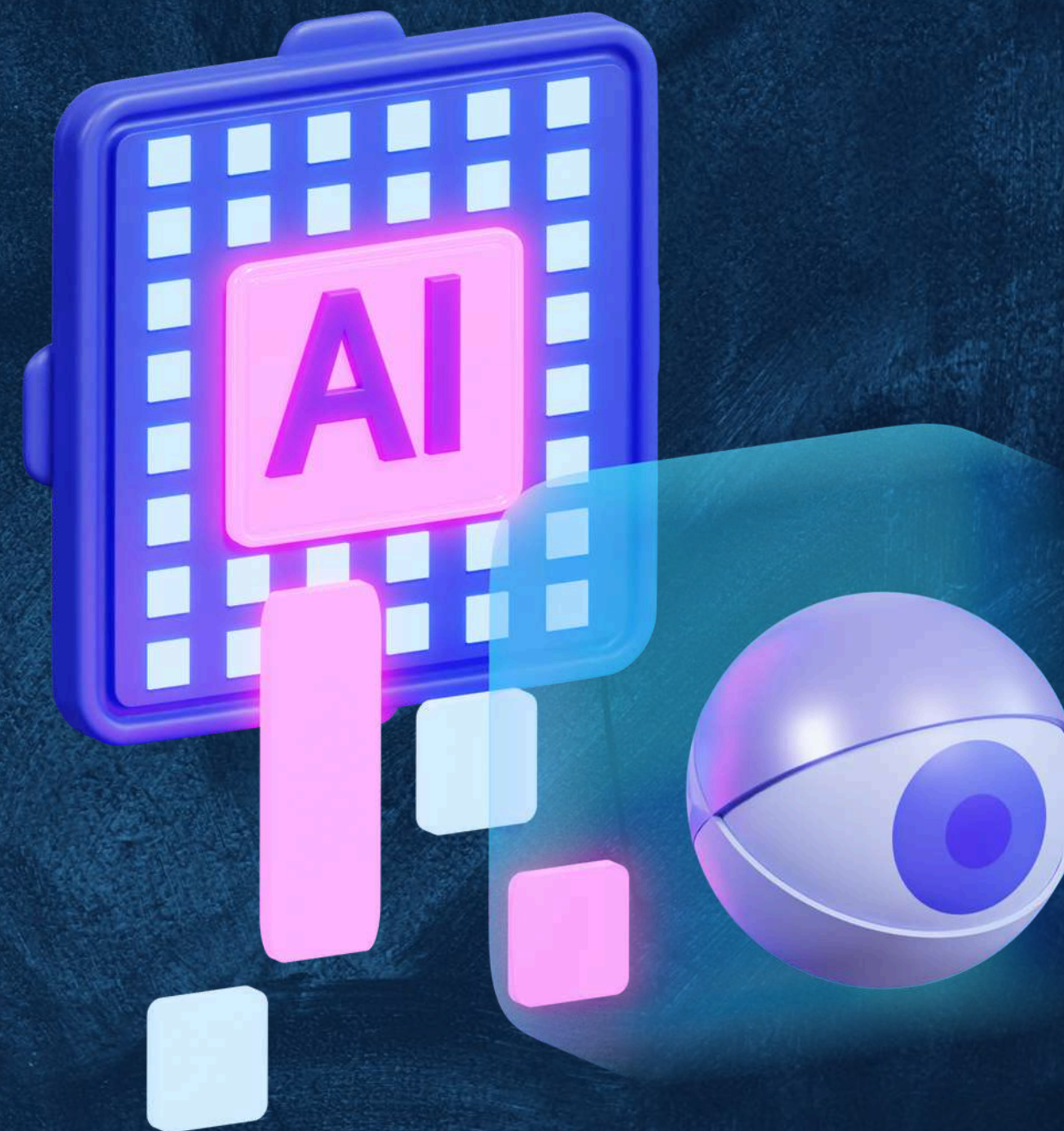
Train = 1000.71

Test = 56870.37

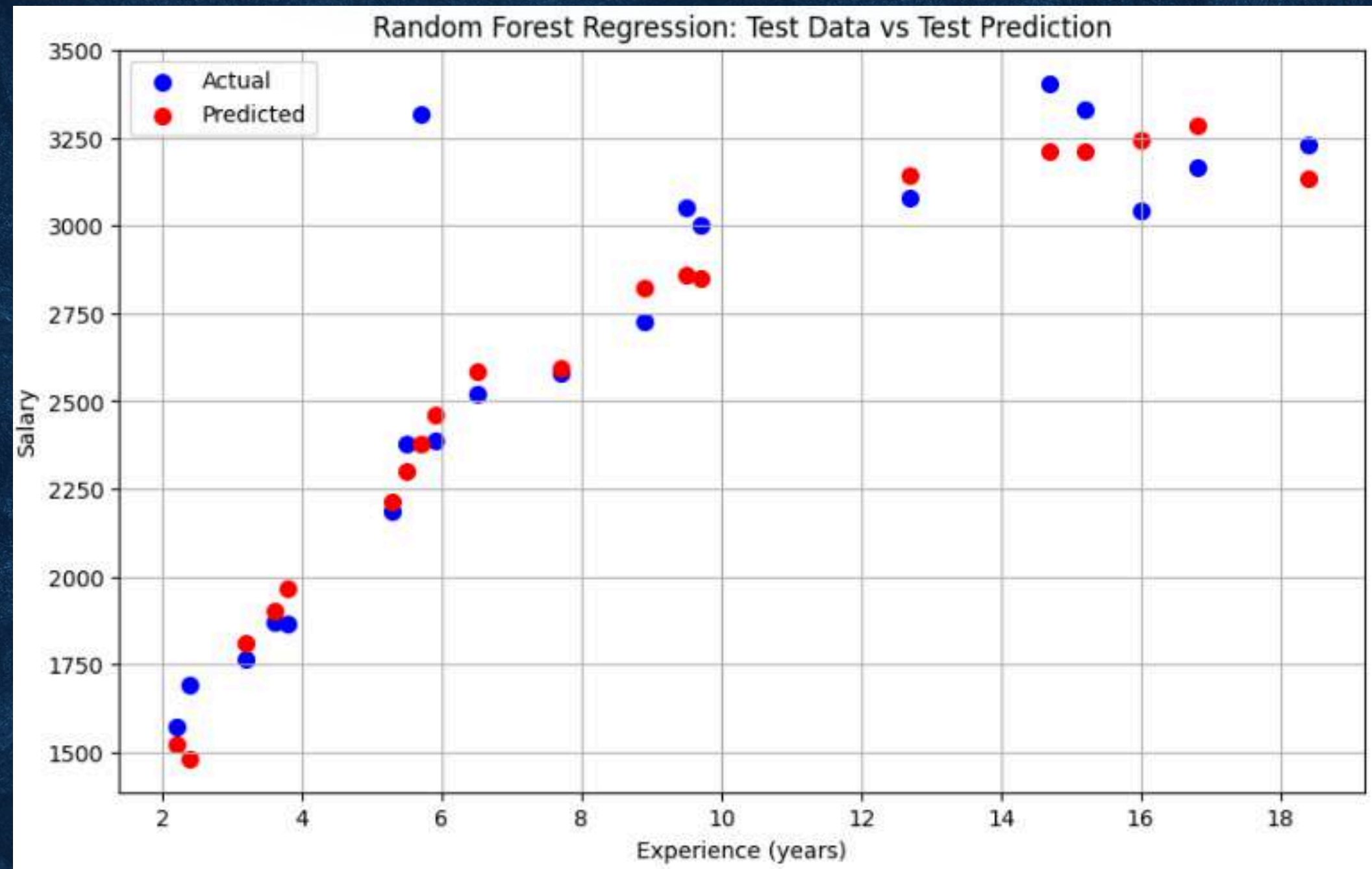
R^2 Score:

Train = 1.00

Test = 0.84



RANDOM FOREST



RANDOM FOREST RESULTS:

Mean Squared Error:

Train = 2714.12

Test = 57520.00

R^2 Score:

Train = 0.99

Test = 0.84



A decorative horizontal bar with a dotted pattern is positioned above the main title.

SOURCE CODE



[GITHUB.COM/NADHIF-ROYAL/SALARYPREDICTIONMACHINELEARNING](https://github.com/nadhif-royal/salarypredictionmachinelearning)



ANALISIS PERBANDINGAN



| Model | MSE (Train) | MSE (Test) | R^2 (Train) | R^2 (Test) |
|-------------------|------------------|------------|---------------|--------------|
| Linear Regression | Tinggi | Tinggi | Rendah | Rendah |
| Decision Tree | Rendah (Overfit) | Tinggi | Tinggi | Rendah |
| Random Forest | Rendah | Rendah | Tinggi | Tinggi |



KESIMPULAN

Dari ketiga metode, Random Forest memberikan hasil terbaik dengan performa yang stabil dan generalisasi yang baik pada data dengan pola non-linier. Linear Regression cocok untuk hubungan linier sederhana, tetapi kurang akurat pada data non-linier. Decision Tree mampu menangkap pola non-linier, namun cenderung overfitting dan kurang stabil dibanding Random Forest. Untuk dataset ini, Random Forest adalah pilihan paling tepat.



A decorative horizontal bar composed of a grid of small white dots.

TERIMAKASIH



[LINKEDIN.COM/IN/ROYALNADHIF50/](https://www.linkedin.com/in/royalnadhif50/)



[INSTAGRAM.COM/ROYAL_NADHIF/](https://www.instagram.com/ROYAL_NADHIF/)



[GITHUB.COM/NADHIF-ROYAL/](https://github.com/NADHIF-ROYAL/)

