Tech Salary Estimator PROJECT 🤑

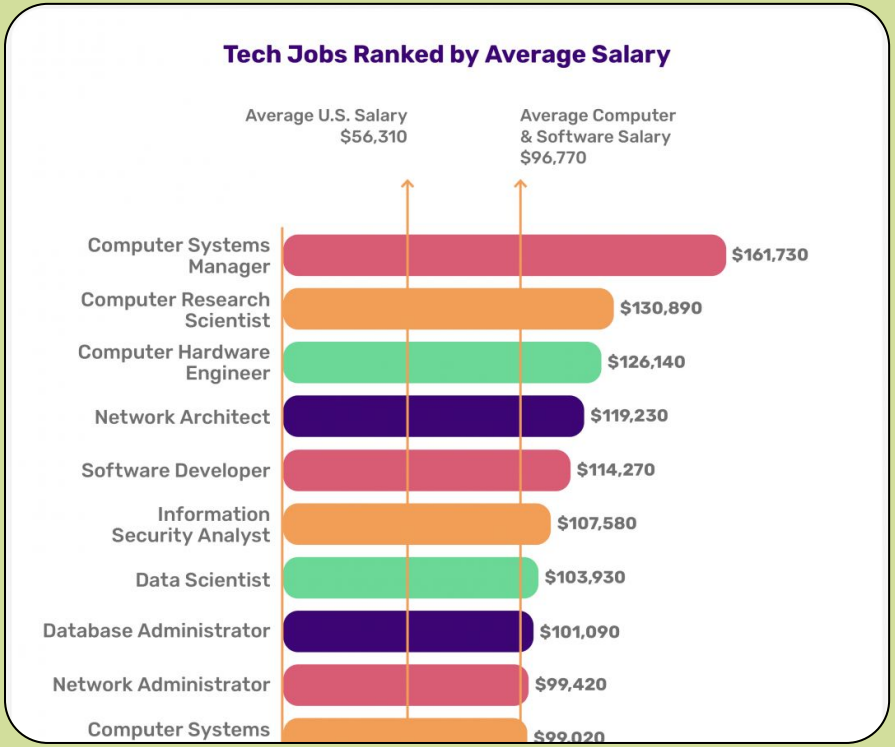
Presented By
Saowaluk Jirapornsirikul

CPSC481-01 | Eve Saowaluk Jirapornsirikul | CWID: 836243824

Why Tech Salary Estimator?

- As a computer science graduate student entering the tech industry, a key challenge is estimating a realistic salary based on personal background.

- Public sites give general ranges but lack personalization.
  Key factors like experience, location, job title/rank are often missing, yet heavily impact salary.



Section 1

Project Goal

- Develop AI Understanding🚀
- Develop a Personalized Estimation Tool 🧮
- Apply Machine Learning Concepts 🧠

How?

Tools & Technologies

- Python
- Pandas & Scikit-learn
- Random Forest
- Joblib
- Streamlit

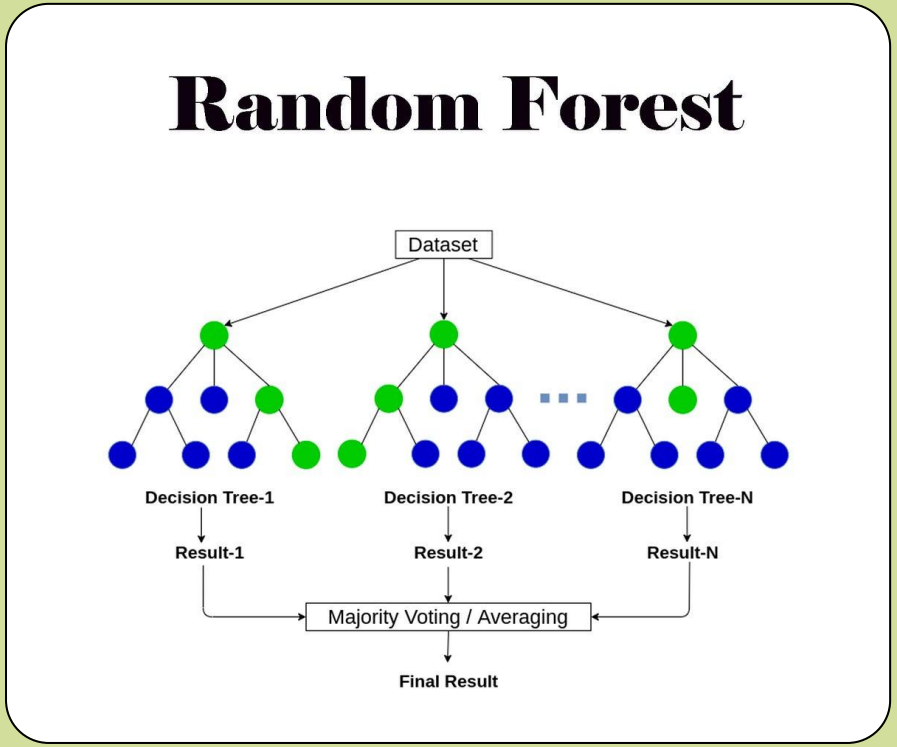
Datasets

Kaggle.com

- **Data from** Kaggle.com
  (https://www.kaggle.com/datasets/thedevastator/know-your-worth-tech-salaries-in-2016)
- **Selected Features**
  - 👩‍💻**Job Title** – Role or position in the tech industry
  - 🏅 **Job Title Rank** – Seniority or level (e.g., junior, senior)
  - 📈**Total Experience Years** – Total years of professional experience
- **Data Limitations**
  - The dataset is from 2016 and may not reflect current salary trends. ⚠️
  - However, the model and UI are designed to be reusable with updated datasets in the future.

Why Random Forest?

- Handles Missing Data
- Algorithm ranks features based on their importance in making predictions
- Scales Well with Large and Complex Data



https://medium.com/@abhishekjainindore24/everything-about-random-forest-90c106d63989

Implementation

1

Data Preprocessing
- Selected key features
- Cleaned missing values
- Encoded categorical variables

2

Model Training
- Used Random Forest Regressor
- Trained on historical data
- Saved model and input schema

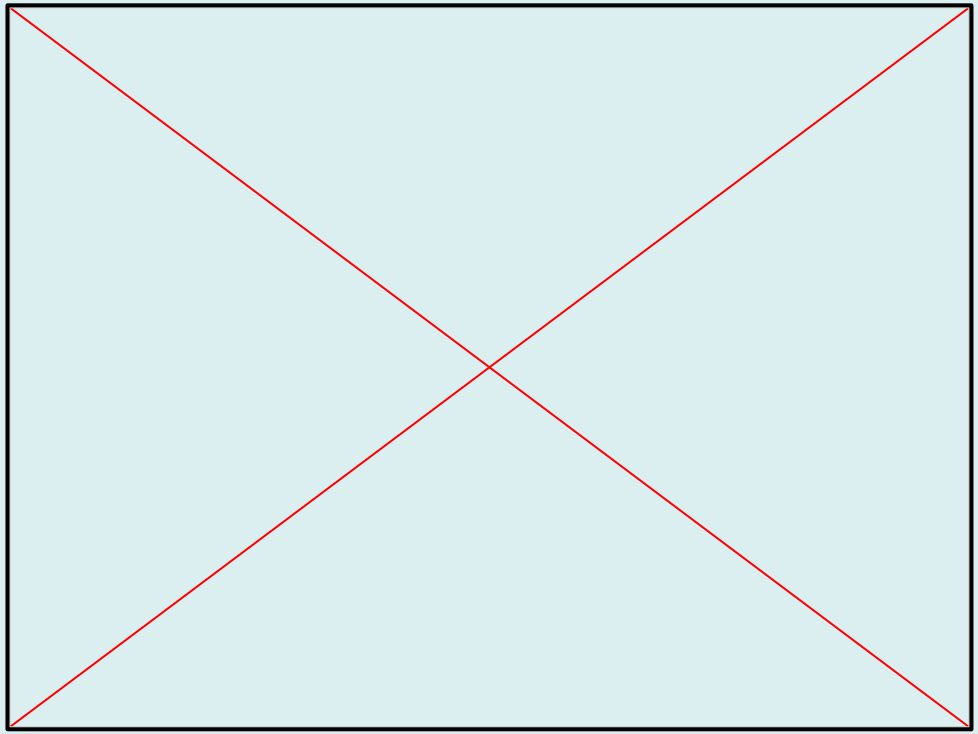
3

User Interface – Built with **Streamlit**

Final

DEMO

Streamlit



Final

🔮 Future Work

- While this project used 2016 salary data, the tool is designed to be easily updated with more recent datasets to provide more accurate and relevant predictions.
- Incorporate more features (e.g., skill sets, certifications, and education level)
- Improve model with additional algorithms
- Deploy on the web for public use or integrate with job platforms

THANK YOU