My-Data-Science-Portfolio

View project on GitHub

/ Thip Rattanavilay

/ Welcome to My Data Science Portfolio

// About Me

I have over 17 years of IT experience in field of Network and System Infrastructure including Cloud computing, 6 years in Data Analytics focused on processing and analyzing large amount of data using Hadoop (Mahout, Hive, PIG), Python, R, MS Excel, MS Access, MS SQL, SAS, Kafka, Zookeeper, Matlab. Passion in Machine Learning, Data/Text Mining, Statistical Analysis & Predictive Modeling. Efficient in: data acquisition, storage, analysis, integration, predictive modeling, logistic regression, decision trees, data mining methods, forecasting, factor analysis, cluster analysis, ANOVA and other advanced statistical techniques.

This portfolio provides my journey into Data Science. Data science is the process of using algorithms, methods and systems to extract knowledge and insights from structured and unstructured data.

You can also find my latest work on my main GitHub repo.

// Project Intro/Objective

The purpose of this portfolio is a collection of all the data science projects throughout my journey into Data Science. In these Data Science journey, I have conduct several data science projects from origin to presentation. I have gather data, then prepare, clean, analyze, and present the analysis to an audience in a recorded video/presentation.

/// - My Data Science Projects:

1. Crimes in the City of Los Angeles Prediction
2. Covid 19 Data Visualisation
3. Hotel Cancellation Prediction
4. Data Analysis Modeling
5. Computer Gaming Addiction
6. Blood Donation Prediction
7. Cryptocurrency Price Speculation & Prediction
8. Air flight Safety
9. Disney Movie Gross Prediction
10. Real Estate Investment & California Price Prediction

/// Resources

- Kaggle
- Tableau
- Government Open Data
- Datahub
- University of California Irvine for ML
- NASA Earth Data
- Federal Aviation Data
- Justice and Crime datasets

/// Methods Used

- Inferential Statistics
- Machine Learning
- Data Visualization
- Predictive Modeling
- Random Forest
- Linear Regression
- KNN
- CRISP Model
- Keras
- Spark
- Python
- R Programming
- Zookeeper
- SciKit Learning

// Contact

- My Contacts/Email: Thip Rattanavilay
- My Github Repo: Github Repo

- Feel free to contact me with any questions or if you are interested in contributing!