Prabowo Setiawan

prabowo.ast@gmail.com | Jakarta, Indonesia
https://www.linkedin.com/in/prabowo-setiawan | https://prabowst.github.io | https://github.com/prabowst

Skills

- Programming: Python (pyspark, scikit-learn, spacy, pandas, numpy, matplotlib, seaborn, tensorflow, pytorch), MATLAB, *familiar* with C++.
- Tools: SQL, AWS EMR, S3 and Redshift, GCP Dataproc and BigQuery, Airflow, Git, Jupyter, and Tableau.
- Techniques: Machine learning, statistical modeling and regression, clustering, neural networks, reinforcement learning, data visualization, ETL and ELT pipelines, dynamic and control system modeling.

Experience

Data Engineer & Analyst, FinAccel — Nov 2020 - Present

- Designs and implements data pipeline to process structured data in GCP and AWS environments; this includes S3 and SQL databases to Redshift and GCS to BigQuery.
- Deploys automated ETL and ELT processes with Airflow / GCP Composer through the help of EMR and Dataproc.
- Increases queries' performance by modeling data warehouse for the company's Data Science team use cases; **reduces runtime to under 15 seconds on average per query dashboards**.
- **Initiates the documentation project to eliminate 'middle-man' in knowledge transfer of tables' functions**.

Data Scientist/Analyst, Public Housing Savings Management — Aug 2020 - Nov 2020

- Developed liquidity management simulation (5 years or more) using Monte Carlo through application of Python and Numba for code optimization; **reduces the downtime of reporting from months to daily basis**.
- Applied machine learning on missing information of house ownership due to inaccessibility in confirming data quality; **provided multiple customer segmentation for specific business case** and **high success on predictive modelling of over 30% missing house ownership (over 90% TP/TN on confusion matrix)**.
- Dashboards creation and maintenance using Tableau for internal and external reporting.

Projects

Starbucks Customer Behavior - Rewards App — https://prabowst.github.io/starbucks.html

- Developed a thorough analysis on Starbucks rewards app program and built a predictive model to further improve its customers' classification; **eXtreme Gradient Boosting provided over 91% performance on TP/TN and provided feature importance insights for better marketing / rewards targets**.
- Studied the **overfitting and underfitting behaviors of models using learning-curve analysis**. Medium link for in-detail report: https://bit.ly/3gYCuC9

InfoGAN - Study on MNIST Dataset — https://prabowst.github.io/infogan.html

- Modeled an Information Maximizing Generative Adversarial Network with 1 categorical and 2 continuous codes as a form of application from the paper: https://arxiv.org/abs/1606.03657 implemented using tensorflow.
- **InfoGAN was able to distinguish digits using learned categorical code and control the representation of rotation and thickness of images using continuous codes;** hyperparameter-tuned the model to achieve clear image generation in just 50 epochs.

Education

- **MSc in Engineering**, Advanced Mechanical Engineering, **University of Leeds** *Graduated Distinction* **2019**
- **BE,** Mechanical Engineering, **Stony Brook University** *Graduated Summa Cum Laude* **2018**

Certifications

- Udacity Data Scientist Nanodegree
- Udacity Deep RL Nanodegree
- Udacity C++ Nanodegree

Relevant Coursework

- Data Science
- Data Engineering
- Recommendation Engine
- Deep Learning
- Software Engineering
- Deep Reinforcement Learning