Prabowo Setiawan

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

Skills

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

Professional Experience

Data Engineer & Analytics, FinAccel — *Nov 2020 - Present*

- Design and implement data pipeline to process structured data in GCP and AWS environments; this includes S3 and SQL databases to Redshift and GCS to BigQuery.
- Deploy automated ETL and ELT processes with Airflow / GCP Composer through the help of EMR and Dataproc.
- **Increase queries' performance by modeling data warehouse** for the company's Data Science team use cases; **reduce runtime to under 15 seconds** on average per query dashboards.
- **Improve business process** of Operations team through **automation and intricate pipeline designs** while maintaining reusability and flexibility on the users' end.
- Research and be a part of proof of concept team for multiple migration projects, data governance, as well as documentation project for better tables and pipelines understanding.

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; **reduced downtime of reporting** from monthly to daily basis.
- Applied machine learning techniques on missing information of house ownership due to inaccessibility in confirming data quality; **provided multiple user 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 [link]

- 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: medium link

InfoGAN - Study on MNIST Dataset [link]

- Modeled an Information Maximizing Generative Adversarial Network with 1 categorical and 2 continuous codes as a form of application from the paper 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

University of Leeds — *2018 - 2019*

Master of Science in Engineering in Advanced Mechanical Engineering, *Graduated with Distinction*
Grade: 71

Stony Brook University — *2014 - 2018*

Bachelor of Engineering in Mechanical Engineering, *Graduated with Summa Cum Laude*
GPA: 3.86