

ANKIT AGRAWAL

Data Scientist | Kaggle Expert

(801)-800-1304 | kit.agrawal1@gmail.com | <https://www.linkedin.com/in/ankitagrawal3/>

Business-minded data scientist with a demonstrated ability to deliver valuable insights via data analytics and provide data-driven solutions to business problems.

SKILLS

Languages: Python, R, C++, SQL, Matlab, HTML, CSS

Data Science: Data Analysis, Predictive Modeling, Data-Driven Personalization, Big Data Queries, Data Mining, Visualization Tools, Machine Learning Algorithms, Forecasting

Packages & Tools: pandas, scikit-learn, keras, tensorflow, pytorch, opencv, matplotlib, nltk, seaborn, beautifulsoup, Tableau, Github, Docker, Google Cloud Platform, AWS SageMaker, AutoML, Excel, LaTeX.

EXPERIENCE

Aakash 88 LLC

Data Scientist

The Woodlands, TX

2019 - Present

Performed data analysis using Tableau and unsupervised learning methods and feature engineering. Build and maintain ML pipelines to forecast hourly electricity price & load for wind farm generated energy trading for ERCOT market.

- 13% increase in annual profits, 35% reduction in analysis time through automation, feature selection and identifying new predictors.

The University of Utah

Machine Learning Research Associate

Salt lake City, UT

2017 - 2019

Run experiments on cluster (CHPC) to detect optimal subspace configuration for SMACK parameters out of 5 Billion possible configurations.

- Reduce TIMEOUT by 23% across 14 categories to increase the chances to score higher in SV-COMP.
- Reduce Type II errors by 3% for each of the 14 SV-COMP benchmark categories by using AutoML and ETL data pipelines.

The University of Utah

Teaching Assistant

Salt Lake City, UT

2014-2016

- TA for Discrete Mathematics, Advance Algorithms, Data Mining
-

PROJECTS

SV-COMP program classification: Given 10,000 C-programs (or 10 million lines of code), performed feature extraction, feature engineering, and implemented *classification trees* based models to classify them into appropriate SV-COMP benchmark categories.

- 98.3% *classification accuracy* compared to previous 67% accuracy.
- 3x *speedup* than previous implementation.

Google landmark detection: Given 6 million train images and 200,000 test images, implemented 2-layer image classification model firstly to classify whether an image contains a landmark using *VGG-19 transfer learning* by injecting non-landmark images in the training dataset and secondly, predict appropriate landmarks in the images containing landmarks using *ResNet transfer learning*.

Self-driving car: Given a train video with 25,000 frames and speed of car at each frame, implemented a *CNN - RNN regression model* to predict speed of the car for a test video with 8,000 frames.

- Achieved MSE < 10.
-

PAPERS

SMACK software verifier and verification toolchain, Z. Rakamaric et.al

EDUCATION

The University of Utah

PhD Candidate, Machine Learning

MS, Computer Science

Salt Lake City, UT

August 2019

December 2016