Ankit Agrawal

Data Scientist | Kaggle Expert
801-800-1304 | kit.agrawal1@gmail.com | linkedin.com/in/ankitagrawal3

Technical Skills

Languages: Python, C/C++, SQL, HTML/CSS, R
Data Science: Data Analysis, Predictive Modeling, Big Data Queries, Data Mining, Visualization Tools, Machine Learning Algorithms, Forecasting, Natural Language Processing, Statistics, A/B testing
Developer Tools: Docker, VS Code, Visual Studio, PyCharm, Tableau, Github, Docker, GCP, AWS SageMaker, AutoML, Facebook Prophet, Hugging Face
Libraries: pandas, NumPy, Matplotlib, scikit-learn, Tensorflow, PyTorch, openCV, nltk, seaborn, Beautiful Soup

Experience

Data Scientist Nov 2019 - Oct 2021
Aakash 88, LLC *The Woodlands, TX*

- Implemented forecasting model pipeline to forecast electricity price & load for energy trading
- Performed data analysis using Tableau, unsupervised learning methods to gain insights in the data
- 13% increase in annual profits
- 35% reduction in analysis time through automation

Machine Learning Associate Researcher Feb. 2017 - July 2019
The University of utah *Salt Lake City, UT*

- Detecting optimal subspace configuration for SMACK parameters out of 5 Billion possible configurations using AutoML
- Reduced TIMEOUT errors by 23% across 14 benchmark categories
- Reduced Type II error by 3% across 14 categories by using AutoML and ETL data pipeline

Graduate Teaching Assistant Aug 2014 - Dec 2016
The University of Utah *Salt Lake City, UT*

- Assisted professors with courses like Discrete Mathematics, Advanced Algorithms, Data Mining
- Conducted office hours, help with grading assignments, projects and exams

Projects

Google Landmark Detection | *Python (PyTorch, pandas, seaborn), Google Colab* May 2018

- Given 6 million train images and 200,000 test images, implemented 2-layer image classification model
- Stage 1: Classify whether an image contains a landmark using VGG-19 transfer learning by injecting non-landmark images in the training dataset
- Stage 2: Predict appropriate landmarks in the images containing landmarks using ResNet transfer learning

SV-COMP classifier | *Python (keras, scikit-learn, matplotlib, pandas), LateX* Dec 2017

- Given 10,000 C-programs (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

Self Driving Car | *Python (keras, scikit-learn, matplotlib, pandas)* June 2017

- 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

Education

The University of Utah Salt Lake City, UT
PhD Candidate *Aug 2017 - July 2019*

The University of Utah Salt Lake City, UT
Masters in Computer Science *Dec 2016*