In Sung Jang

Research Data Scientist at UChicago
hanlbomi@gmail.com, 312-709-6783, LinkedIn, Portfolio

Professional Experience

University of Chicago, Astronomy | Research Data Scientist **2020 - Present**

- **Big Data Analysis and Software Development**
  - Created end-to-end pipelines (Python and SQL), automatically processing over 1TB of unstructured dataset
  - Employed machine learning algorithms to uncover patterns and signals within extensive source catalogs (N > 1 million)
  - Led teams in enhancing flux measurement algorithms with predictive modeling, reducing errors by 95%
- **Quantitative Research and Engineering**
  - Securing funding of over $300k through NASA grant acquisition as a program lead
  - Collaborated with international teams, overseeing pull requests and conducting thorough code reviews
  - Independent research to develop data processing algorithms in astronomy; published over 3 papers.

Leibniz Institute for Astrophysics Potsdam (Germany) | Research Scientist **2016 - 2020**

- **Cloud Computing and Time Series Data Analysis** :
  - Leveraged supercomputers, reducing computation times tenfold in statistical modeling of unstructured data.
  - Developed algorithms for the time-series analysis of variable stars, enabling the identification of transient phenomena

Skills and Qualifications

- **Machine Learning**
  - Certificates : (1) Deep Learning Specialization (Univ. of North Texas), (2) Python for Data science, AI & Development (IBM), and (3) SQL Basics for Data Science Specialization (Univ. of Calibornia)
  - Applied Python sklearn and pandas packages to real-world, often uncleaned and semi-structured data
- **Statistics**: Probability, Distributions, ML methods, Hypothesis testing, A/B testing
- **Programming languages**: Python (+5 years, expertise in pandas, sklearn), SQL, R, Matlab, IDL
- **Software/Others**: Tableau, Linux/Terminal environment, Git

Education

Seoul National University, Ph.D in Astronomy&Astrophysics **2009 - 2016**
Inha University, BSc in Aerospace Engineering **2005 - 2009**

Projects and Honors

Business Intelligence via Machine Learning (Portfolio)

- Employed polynomial features and Linear Regression in Machine Learning to predict Chicago home prices using Zillow data
- Optimized bank marketing strategies through K-Nearest Neighbor classifiers, resulting in enhanced efficiency
- Leveraged Decision Tree algorithms to accurately forecast hotel booking demands

Kaggle Data Science projects

- Developed predictive models for various competitions, showcasing proficiency in machine learning and statistical analysis
- Continuously learned and adapted by studying Kaggle Kernels, exploring novel techniques and approaches

Data-Driven Research Publications

- 11 first-author journal articles with +300 citations in quantitative data analysis, ranking in the top 3% of the most cited work