SOMANG (SO) HAN

sohan1@seas.upenn.edu| +1 (507) 321-3319 | Philadelphia, PA | https://somang1418.github.io

EDUCATION

University of Pennsylvania| Master of Science in Data Science   Philadelphia, PA | **Anticipated Graduation**: May 2022
St. Olaf College| Bachelor of Arts in Mathematics & Statistics, GPA: 3.79/4   Northfield, MN| May 2018

SKILLS

- **Programming Tools**: Python (scikit-learn, pandas, numpy, matplotlib, scipy, plotly, tensorflow, keras, pyspark) , R (ggplot2, r shiny, dplyr, leaflet, mlr) , SQL (MySQL, SQLite), Excel, SAS, Tableau, Power BI, Hadoop, AWS, HTML, CSS
- **Technical Skills**: Data Analytics, Data Visualization, Data Mining, Predictive Modeling, Machine Learning, Deep Learning, Big Data Engineering, Statistical Analysis, Customer Analytics, Fraud Analytics, E Commerce Analytics

PROFESSIONAL EXPERIENCE

University of Pennsylvania   Philadelphia, PA
Data Manager, Penn Online Learning Initiative (OLI)   Aug 2020 – Present
- Analyze Penn MOOCs data for 8 million enrollment on Coursera to uncover trends and provide recommendations
- Work collaboratively with OLI's Director of Operations to manage MOOCs database through Quick Base, Salesforces applications and retrieve data using SQL to create quarterly reports for tracking key metrics in OLI's growing portfolio

Bluestem Brands, Inc.   Eden Prairie, MN
Decision Scientist, Marketing   Jan 2019 – June 2020
- Extracted millions of customer data across multiple platforms (Hadoop, SQL Server, SAS) and transformed large complex data to develop machine learning models for new customer acquisition, customer marketing, and personalization in Python
- Led a team of 2 to enhance a customer experience tracking tool by K-Means Clustering in R Shiny and increased customer satisfaction of fingerhut.com by 10% in four months
- Built XGBoost ensemble models using Python on AWS to identify customer segments for direct mail marketing campaigns and improved the sales conversion rate by 15%
- Supported internal clients with ad-hoc analysis and statistical analysis to enable data-driven decisions by brand leaders

Amplifon, Inc.   Minneapolis, MN
Associate Data Analyst   Sep 2018 – Nov 2018
- Consolidated annual revenue reports for sales of Miracle-Ear products, validated metrics and KPIs by manipulation of complex Excel formulas and Pivot Tables, and improved the process efficiency by 5%

ID Insight, Inc.   Minneapolis, MN
Data Scientist Intern   June 2018 – Sep 2018
- Identified different types of bank fraud by K-Means Clustering in R; generated data visualization in Tableau; presented findings to C-level leadership
- Rebuilt a bank fraud detection model using generalized linear model in R and improved model performance by 20%

Center for Interdisciplinary Research, St. Olaf College   Northfield, MN
Statistical Consultant   Aug 2017 – May 2018
- Utilized statistical techniques in R to provide recommendations to 2 to 3 students per week for their projects
- Conducted NSF funded research in Topological Data Analysis with 3 students and published research findings on a peer-reviewed journal (SIAM Undergraduate Research Online - DOI: 10.1137/18S017302)

SELECTED PROJECTS

Data Scientist Job Description Generator   **(Github project)**
- Extracted 10,000 data scientist job descriptions from Indeed using web scripting (beautiful soup and requests), preprocessed the data, and generated a data scientist job description generator by Recurrent Neural Network (RNN)

Elo Customer Loyalty Score Prediction   **(Kaggle Competition: Top 11%)**
- Preprocessed 30,000 customer data and built ensemble models (XGBoost/LightGBM) to predict customers' loyalty score for Elo, one of the largest payment brands in Brazil

Topological Data Analysis on Network of Game of Thrones   **(Summer Research 2017 at St. Olaf College)**
- Examined Game of Thrones network map datasets by a Topological Data Analysis tool and implemented link analysis to closely look characters' degree centrality to find the most important community in the network map
- Presented results of the research at the Joint Mathematics Meetings (JMM) 2018

ADDITIONAL INFORMATION

- **Leadership**: Ambassador @ Women in Data Science (WiDS) 2020, Sergeant at Arms @ Toastmasters International, Guest Speaker @ MinneFRAMA (Business Analytics Conference), Co-chair @ St. Olaf Korean Cultural Association
- **Kaggle Expert**: Kaggle is a community for practicing data science problems (my profile: www.kaggle.com/somang1418)
- **Hobbies**: Alpine Skiing – I was in the South Korea Jr. National Alpine Skiing Team for 2 years, Baking – I like eating muffins