

Min Sup Lee

(217) 305-5231 • leeminsup1989@gmail.com • www.linkedin.com/in/minsup-marc-lee/

Summary

Data scientist with a track record of Predictive Analytics, Process Optimization, and NLP. Experienced in Statistics, Engineering, and Machine Learning. Skilled in Python, R, and SQL as well as in TensorFlow and Keras for Deep Learning frameworks. Demonstrated experiences in speaking technical findings in business language.

Work Experience

Soothsayer Analytics, Data Scientist

Livonia, MI, May 2019 – Current

- Forecasted monthly revenue using Time Series techniques (ARIMA, Holt-Winter, TBATS, Prophet), XGBoost, and Neural Network – \$5B revenue client in a manufacturing industry
- Constructed the classification model to predict the 8-comfortable-levels of drivers by different temperature settings
- Optimized the production process to maximize the production rate, to minimize the waste, and to identify the key process drivers for plant engineers – \$8.4B revenue client in a chemical industry
- R&D – NLP (Topic Modeling, Word Embeddings, Speech Recognition), Computer Vision (Image Classification)

Anheuser-Busch InBev, Data Science Intern

Champaign, IL, May 2018 – Feb 2019

- Proposed more efficient transportation route that saved 17% of the logistics cost using simple EDAs
- Identified key drivers of sales volume of cities in Brazil and South Africa individually using 2 years' worth sales data, Google Analytics data, and additional custom data for each region
- Forecasted volume with WMAPE of <1% using Random Forest and XGBoost, and Neural Network

Oh Pharmaceutical, Project Coordinator

Crown Point, IN, July 2016 – Dec 2016

- Evaluated technical terms to construct an automated manufacturing facility that meets FDA and safety requirements
- Optimized the budget and the schedule to reduce cost up to 20% in \$40M project

G2G Solution Co. Ltd, Data Analytics Team

Seoul, Korea, Sept 2013 - Aug 2015

- Assisted to implement a new ERP system to increase the overall resource planning efficiency up to 20%
- Analyzed market data to identify the potential Dash Cam market size in the South Korea
- Examined the potential international Dash Cam market demand from CeBIT 2014 participation

Education

University of Illinois, Urbana-Champaign

- M.S. in *Statistics – Analytics* *Dec 2018*
- B.S. in *Industrial Engineering and Statistics* *May 2016*

Academic Project

Predictive Analytics on Financial Data

Fall 2018

- Engineered Financial data and articles to run ML/DL algorithms using H2O, Keras, and TensorFlow in R
- Achieved 77.9% accuracy using the Stacked Ensemble method with 0.881 AUC (Bitcoin Data, placed 1st/12 teams)
- Predicted stock movements in 3 directions using CNN, RNN, HRNN with up to 51% accuracy (CME LOB) on AWS

Clustering Gene Expressions by Ontology Annotation

Spring 2018

- Applied the latest Network Analysis and Clustering methods to identify ontology annotations using R and Python
- Concluded that autism, schizophrenia, and bipolar disorder share global gene expression patterns, characterized by astrocyte activation and disrupted synaptic processes

Predictive Analytics using Monte Carlo Algorithm

Fall 2017

- Reproduced and predicted the exchange currency trend of Ruble to USD using statistical sampling methods and Monte Carlo Markov Chain (up to 82% training accuracy)

Production Schedule Optimization for Cost Reduction (Client: Anheuser-Busch InBev)

Fall 2015

- Fetched production data from the client's database and conducted EDA to identify features that can be optimized
- Built a prototype production control system that reduced the production cost up to 15% using VBA, R, and Excel

Skill

Python, R, SQL, TensorFlow, Keras, SSMS, Azure Data Factory, Databricks, Spark, AWS EC2, Tableau, Power BI, Altair Smart Sight, H2O, SAS, Octave, C++, Inventor, AutoCAD, SIGMA, HDFS, Pig, Hive