Aeint Thet Ngon

An692@georgetown.edu, (312) 375-3810
Aeintngon.github.io, https://www.linkedin.com/in/Aeintngon

Education

Georgetown University- Washington, D.C. May 2018
Masters of Science in Analytics Concentration in Data Science

Cornell College- Mount Vernon, IA May 2016
Bachelor of Arts in Economics and Business, and Mathematics, magna cum laude

Skills and Tools

- Software and Programming Languages: Python (numpy, scipy, pandas, NLTK), R (dplyr, ggplot2), Java, SQL, Stata
- Machine Learning: classification, regression, clustering, support vector machines, tensor flow, K-Nearest Neighbor, and random forest
- Statistical Methods: time series, regression models, hypothesis testing and confidence intervals, principal component analysis and dimensionality reduction
- Massive Data Tools: MySql, Spark, Amazon Web Services, MapReduce, Hadoop
- Data Visualization Tools: GGplot, Tableau, Plotly, Bokeh
- Development tools: Eclipse, JetBrains, Spyder, Jupyter Notebook, RStudio

Work Experience

Research and Technology Assistant, Georgetown University Jul 2017- Present

- Collaborate with the University Information Services' engineering team to grow the Expandable Open Source database, a Big Data archive of open source media used in universities researches
- Assist researchers from different universities to build new robots to crawl designated websites

Research Assistant, University of Iowa Jun 2016-Aug 2016

- Queried and analyzed 1.5 billion patient visits data from Healthcare Cost and Utilization Project
- Coded in R to manipulate data and connect patients across time and different hospital visits
- Cooperated with medical professionals to identify potential misdiagnosis
- Applied logistic regression and association rules to predict the likelihood of potential misdiagnosis
- Presented the analysis at University of Iowa Hospital and Cornell College's Annual Symposium

Peer Consultant, Quantitative Reasoning Studio, Cornell College Sep 2014-May 2016

- Aided students with different types of mathematical and statistical needs
- Collaborated with faculty to deliver statistics software training (R, Stata, Excel, Tableau)
- Led trainings to train students and other peer consultants on Tableau and Excel

Projects:

Forecasting Energy Emission- Tools: R Dec 2016

- Built a machine learning model using Support Vector Machine that predicted the energy production with root mean square error of 19%

Sentiment and Stock Analysis during 2008 Financial Crisis- Tools: Python Dec 2016

- Scraped Yahoo Finance, New York Times and Fox News for relevant data
- Analyzed trends of different industries using linear regression
- Performed sentiment analysis on how media reacted to the Financial Crisis

Amazon Review Data Analysis- Tools: Amazon Web Services, Python Apr 2017

- Generated a NLTK pipeline to process 7 million observations
- Developed a Naïve Bayes model to predict sentiment score
- Predicted the rating based on review text with 78% accuracy rate

Scholastic Awards

- Charles M. Cochran Economic and Business Award (awarded to top 1% based on major GPA)
- Edhill Mathematics Scholar Award (awarded for representing Cornell College at national math competitions)
- Omicron Delta Epsilon (Economic Honor Society)