Ding Luo

4231 12th Ave NE, Seattle WA | dingluo@uw.edu | 206-483-5633 | Portfolio

SUMMARY OF QUALIFICATIONS

- Highly analytical and process-oriented with 2-year experience in **Business Intelligence**, **Data Science** and **Data Visualization**
- Exceptional communication skills and the ability to interact with both technical and non-technical audiences
- **Technical Skills:** SQL, R, Tableau, Python (TensorFlow), Excel, JavaScript (D3.js), HTML, CSS, SSIS, SSAS, C++, MATLAB
- **Research Areas:** Causal Analysis (Regression & Survival Analysis), Data Visualization, Deep Learning (RNN, LSTM, CNN), Social Network Analysis (Network Inference, Structure), NLP, Supply Chain Finance (Optimization), Predictive Modeling

EDUCATION

University of Washington, Information School *Seattle, WA*

M.S. in Information Mgmt., Specialized in Data Science & Business Intelligence. GPA: 3.76/4.00 *2016-2018 (expected)*

- **Relevant Courses:** Business Intelligence Systems, Customer Analytics, Interactive Information Visualization, Social Media Data Mining, Social Network Analysis, Machine Learning & Econometrics, Probability & Statistics, Database Mgmt.

Shanghai Jiao Tong University (SJTU) *Shanghai, China*

Antai College of Economics and Management, B.S. in Operations Research. GPA: 3.60/4.00 *2014-2016*

University of Michigan - SJTU Joint Institute, Minor in Electrical and Computer Engineering *2012-2014*

PROFESSIONAL EXPERIENCE

PATH *Seattle, WA*

Monitoring, Evaluation, and Data Analysis Intern - BID Initiative *Aug. 2017-Present*

- Conduct statistical test on immunization data using **Generalized Linear Mixed Model** in R, with the results showing significant improvements in data quality and data use by New Electronic Immunization Registries in over 1000 health facilities in Tanzania
- Develop automated data integration and data wrangling methods by Power Query to reshape large data sets (100k+)
- Analyze the timeliness of vaccine data via **Survival Analysis** to evaluate quality of immunization services (Paper in Progress)

Institute for Health Metrics and Evaluation (IHME) *Seattle, WA*

Research Assistant - Digital Data Representation *June 2017-Sept. 2017*

- Built Ensemble Methods using **Random Forest** for inferring demographics (age & race groups) from Twitter users' profile data
- Created six age models by different feature sets, and estimated accuracy, precision, recall, and F1 metrics for model selection
- Performed effective feature engineering by leveraging data processing techniques (visualization) to domain knowledge (linguistic analysis on text data), designed creative features (N-Gram features from text data) and improved age models' accuracy by 20%

Starbucks *Seattle, WA*

Data Analyst and Report Developer Intern - Global Store Development *June 2017-Aug. 2017*

- Identified business requirements, and defined metrics and KPIs with stakeholders for vendor performance analysis
- Reviewed data model, and retrieved data from multiple database systems by writing sophisticated queries in Oracle SQL
- Used LOD calculations to deal with multiple levels of data granularity in **Tableau**, created professional dashboards with Agile project management and delivered a refined reporting architecture with exhaustive documentation

IBM *Shanghai, China*

Business Analyst Intern - Global Business Services *July 2015-Aug. 2015*

- Took an in-depth study of client company's end-to-end time-based logistics management mode, launched a new index model based on business processes, and utilized MATLAB, C++ to improve the model accuracy to over 85% which reduced cost by 5%

RELEVANT PROJECTS

Customer Online Grocery Shopping Behavior Analysis - AmazonFresh, *Seattle WA* *Jan. 2018-Present*

- Analyze the relationship between density of grocery store and customers' adoption of online grocery shopping using Regression
- Create grocery store density models using location data pulled from Google Maps API, size, popularity and review data from Yelp Fusion API in Python, and visualize interactive store penetration map for 22 cities in Tableau

Language Identification of Tweets using RNN in TensorFlow, *Seattle WA* *Feb. 2018*

- Trained a single character-based RNN that jointly models all languages (9 Western European languages) by minimizing the cross-entropy, embedded the first layer for both language and character, and selected the best model with highest F1 score (0.87)

Business Intelligence Solution Design on Sales Data, *Seattle WA* *Apr. 2017-May 2017*

- Built a Dimensional Data Model ERD in SSMS based on star schema and product analysis requirements
- Designed efficient ETL processes to move data through data processing pipeline for need of the business use cases in SSIS

Twitter Sentiment Analysis on Movie Box Prediction, *Seattle WA* *Apr. 2017-May 2017*

- Fetched real-time Twitter data about one movie via Twitter Streaming API, and used Python packages (scikit-learn) to analyze the association between sentiment composition from Twitter users and box office records of the movie by **Time Series Regression**

Text Authorship Identification Based on Writing Stylistics, *Shanghai China* *Mar. 2014-Sep. 2014*

- Built classification models using **SVM**, extracted features from POS tagging, and creatively implemented **Genetic Algorithm** for feature selection (top 20 features) which improved the model accuracy by over 35% and reduced training time by 10%