

# Mingni Luo

(857)654-5079 | [luo.min@northeastern.edu](mailto:luo.min@northeastern.edu) | <https://www.linkedin.com/in/mingni-luo/>

## EDUCATION

### Northeastern University

Aug. 2021- Jun. 2023

*Master of Science in Data Science*

*Boston, Massachusetts*

*Coursework: Database Management, Machine Learning, Algorithms, Data Mining, Applications in Data Science.*

### Nanjing Agricultural University

Aug. 2015- Jun. 2019

*Bachelor of Management in Information Management and Information System.*

*Nanjing, China*

*Coursework: Data Structure, Probability Theory and Mathematics, Information System Analysis and Design*

## SKILLS

- Programming Language: Python(Matplotlib, Numpy, Pandas, Sklearn, Statsmodels), R, Java, C
- Tools & Software: Git, RStudio, SPSS, MATLAB, Jupyter NoteBook, Tableau
- Database: MySQL, Hive SQL, PLSQL
- Industry Knowledge: Data Mining, Statistics, Data Analysis, Data Visualization

## EXPERIENCE

### SUANHUA Intelligent Technology Company | Fintech

Oct. 2020 - Jun. 2021

*Data Analyst (Full-time)*

- Used Python to implement some rule-based models and classification models, including **logistic regression**, **Decision Tree** and **XGBoost** models and **deployed models** on a commercial bank platform, reducing the **rate of bad accounts from 2.5% to 1.2%** and the **overdue rate from 22% down to 15%**.
- Analyzed accuracy of data models and the stability of variables by **Hive SQL**, aiming at different types of customers, such as banks and loan companies, to **find suitable models** for the customer user group.
- Used **Hive SQL** and **R Studio** to clean the raw data provided by customers, including removing outliers and adjustments to models based on the customer samples.

### HAND Enterprise Solutions Company | ERP Consulting

May.2019 - Oct. 2020

*IT Consultant & Project Manager (Full-time)*

- **Implemented an application** for airlines that optimizes the process of purchasing costumes for flight crews and used **A/B testing** to continuously optimize the product during the development phase.
- Wrote project reports for customers after product was launched, analyzed customer data using PL SQL and visualized the product benefits, which **reduces the company's annual expense of clothing by 5%**.
- Designed the Customer relationship management software workflows for clients, communicated and cooperated with 10+ engineers from different teams to push the project forward.

## RESEARCH

### Covid-19 Vaccination Prediction and Model Evaluation

Nov. 2021-Dec. 2021

- Converted the original data to statistics on the vaccination progress by different countries, sum up "*the daily vaccination*", replace **NA value** with the numbers of the previous day, and then divide by the total number of people to get the estimated vaccination percentage
- Implemented feature engineering, generated sliding window of a week and the results of the previous day.
- Built-up **Prophet**, **ARIMA** and **RNN** models to predict the progress of vaccine used **cross-validation** to **choose hyper-parameters** and used **RMSE** and **R2** metrics to complete **model evaluation**: The RMSE of RNN lowest and The R2 of ARIMA is the best among the three. The prediction results of the three models are visualized at last.

### Automatic Reference Classification Based on Bayes Classifier

Oct. 2018-May. 2019

- Utilized Python to read training set of about **100,000 documents** and processed contents based on the specific format of references, combined with **regular expressions** to transform them to tidy data.
- Constructed a stop-word list for references through the Jieba library of Python and performed feature selection based on **word frequency**, Conducted TF-IDF calculation on the data through the **scikit-learn library** of Python.
- Built **Bayes classifier** using training set and evaluated the effect of the classifier by test set, raising accuracy rate and recall rate more than 90%, improving the efficiency of classifying references.