

# Mengying Wang

100 Exchange St., Apt224, Malden, MA 02148 857.210.2777

wang.mengyin@northeastern.edu

LinkedIn: [www.linkedin.com/in/mengying-mandy-wang](https://www.linkedin.com/in/mengying-mandy-wang)

## EDUCATION

---

**Northeastern University, Boston, MA**

**Sep 2018 - May 2020**

*Master of Science in Computer System Engineering*

Relevant Courses: Data Science & Engineering Methods, Parallel Machine Learning & AI, Data Management & Database Design, Network Structure & Cloud Computing

**Shanghai Institute of Technology, Shanghai, China**

**Sep 2013 - June 2017**

*Bachelor of Engineering in Software Engineering*

## TECHNICAL SKILLS

---

**Programming Languages:** Python, JAVA, Matlab, Shell, SQL, C#, C++

**Machine Learning Models:** Linear Regression, Logistic Regression, Random Forest, LightGBM, CatBoost

**Database:** MySQL, SQL server, MongoDB

**Machine Learning Tools:** TensorFlow, H2O, Keras, NumPy, Pandas, SciPy, Scikit-learn, Seaborn, Matplotlib, Google Colaboratory, Jupyter Notebook

**Development Tools:** AWS(EC2, AMI, Lambda, S3, CodeDeploy, Sagemaker, AWS CLI, CloudWatch, etc.), PySpark, Time Series(ARIMA), CUDA, CircleCI, Git, Docker, Linux, cluster

## TECHNICAL PROJECT

---

**Prediction Platform for Flights delays and Cancellations**

**Apr 2020 - Apr 2020**

- Applied **time series (ARIMA) analysis** to analyze the behavior of the monthly delayed figures based on 5819079 flight records in 2018, and forecast for the next year.
- Improved the process efficiency and time by models (**LR, RF, LightGBM, CatBoost** etc) combined with **Google Colaboratory** and **Discovery cluster**.
- Evaluated models by practical matrices (**AUC/ROC** etc) and optimized models by iterations, and **the accuracy of optimal model reached to 85.9%**
- For further improving the efficiency, used parallel computing with CPU and **GPU (CUDA)**, **the processing time has been reduced 4-20 times**.

**Database for Food Delivery System**

**Mar 2020 - Apr 2020**

- Developed a **database** with **15 tables** to effectively manage the data of food delivered **orders**, **the amount of tested orders is more than 5000**.
- Managed databases by **SQL** sequences included adding **table-level constraints**, created **Views** and **Stored Procedures** and **encrypted** passwords by end-to-end methods.

**Bacterial Phenotype Classification (with Broad institute)**

**Dec 2019 - Mar 2020**

- Refactored and optimized the existing scripts from **Python2 to Python3** and passed all tests.
- Effectively **managed Git repository** and kept closely working with other instructors and members.

## **AWS Based Online Note Taking System**

**Jan 2019 - Apr 2019**

- Used Amazon **AWS** as a cloud environment to deploy a note management project which is based on NodeJS in **SaaS** structure.
- Created **AMI (Amazon Machine Images)** to launch new **EC2** instances and triggered **AWS CodeDeploy** to automate software and **Lambda** function deployments to these instances, then set up continuous integration and deployment (**CI/CD**) with **Docker** on **CircleCI**.
- Stored notes using **S3 buckets** and configuring instance backups to S3 bucket.
- Created and managed stacks with **AWS CloudFormation** using the **AWS CLI**.

## **‘AI Add-in’ — Based on Interpretable Machine Learning**

**Mar 2019 - Aug 2019**

- Interpreted interpretable models (linear regression, decision tree, etc.) or appropriate surrogate models (**LIME**, **CART Decision tree**, etc) for non-interpretable models by several plots (**PDP**, **ICE**, **ALE**, etc.) based on **Interpretable Machine Learning** methods with the help of **H2O** platform.
- Processed multiple datasets successfully, such as Amazon/Yelp Reviews Prediction, Breast Cancer Prediction and Pregnancy Classification, with **more than 60,000 records** in a single dataset.

## **WORK EXPERIENCE**

---

### **Research Intern, Field Robotics Lab, Northeastern University**

**Dec 2020 - Present**

- Work on **depth estimation** and **SLAM VO** by **Unsurprised Deep Learning** methods.

### **Software Development Engineer Intern, Fedora Project (Google Summer of Code)**

**May 2017 - Aug 2017**

- Improve compatibility of Plinth, which is an open source web interface written by **Python** administering the functions of FreedomBox, migrate it from **Debian** to **Fedora** Server and make it compatible with other **Linux distributions**.
- In order to make systems smoother, **repacked** Plinth to make it adapt to both Fedora and **rpm** package format.

## **AWARDS**

---

### **Mathematical Contest in Modeling/Interdisciplinary Contest in Modeling**

Honorable Mention

**Apr 2016**

### **China Undergraduate Mathematical Contest in Modeling, Shanghai Division**

Second Prize

**Oct 2015**

## **EXTRA-CURRICULAR EXPERIENCE**

---

### **Member, GNOME Foundation**

**Jun 2017 - Jun 2019**

### **Member, Beijing Linux/GNU User Group(BLUG)**

**Sep 2016 - Aug 2018**

### **Annual meeting of Taiwanese Open Source (COSCUP), Taipei, Taiwan**

**Aug 2017**

Invited to present “Introduction to Open Source for Female College Students”.