

SHANGWEN YAN

343 Gold St., Brooklyn, NY 11201 | C: 347-320-7075 | sy2160@nyu.edu | github.com/Shangwen-Yan

EDUCATION

New York University Master of Science, Computer Engineering, GPA 3.8/4	05/2019 Brooklyn, NY, U.S.
Beihang University Bachelor of Engineering, Electronic Information Engineering, Outstanding Graduate	07/2016 Beijing, China

TECHNICAL SKILLS

Programming Languages:	Java, Python, PHP, R, Shell, Scala, Golang, JavaScript, HTML, CSS
Database Development:	SQL (MySQL, Oracle), NoSQL (HBase, Redis, MongoDB)
Back-End Related:	Tornado, Beego, Spring, SpringMVC, MyBatis, Spring Boot, Struts2, Hibernate
Front-End Related:	React, jQuery, EasyUI, ECharts, RESTful, JSONP, Postman
Machine Learning Related:	PyTorch, TensorFlow, GPU, CUDA, JupyterNotebook, HPC
Big-data Related:	Hadoop, ZooKeeper, Hbase, Hive, Flume, Kafka, Storm, SparkML, SparkStreaming
Others Skills:	Linux, Kubernetes, Argo, Nginx, Maven, SBT, SVN, GIT, VMware, IBMCloud

RELATED EXPERIENCES

DiDi Chuxing – Machine Learning Engineer Intern	05/2018-08/2018
<ul style="list-style-type: none">Developed machine learning platform DFlow which could be used by dragging components in graphic interface to create workflow, easily running tasks and managing member authority through projects.Integrated DFlow into DLab platform together with other tools such as DNotebook, DSQL, DCube.Investigated and contoured functions of platform, designed API in JSON format and provided contents for UI.Built database structures in MySQL and submitted to DBA servers which were not used in developing stage.Implemented backend API using Tornado framework in Python, tested requests using Postman.Stored configuration files on S3(Simple Storage Service). Created tasks, managed DAG workflows and allocated dockers using Kubernetes and Argo, implemented API using Beego framework in Golang.Uploaded logs and metrics to Kafka, consumed with SparkStreaming and saved on HDFS in parquet format.	
ECE Department, New York University – Graduate Assistant	09/2017- 05/2018
<ul style="list-style-type: none">Worked with ECE Faculty on designated research needs and gathered students' feedbacks & suggestions.Helped professors handle with PhD applicants' materials and answer students' questions.	
Tarena International Inc – Software Engineer Intern	09/2016-08/2017
<ul style="list-style-type: none">Participated in the development of E-Commerce platform "JingTao" and logistics backstage management system "HuiTong". Developed backend using Spring, SpringMVC, MyBatis in Java.Implemented cross-domain access through HTTP Client and JSONP. Improved Single Sign On function with tickets stored on Redis. Implemented load balance and high availability through Nginx in roll polling pattern.	

ACADEMIC PROJECTS

Movie Recommendation System (Python3, PyTorch, HPC, JupyterNotebook)	03/2018-05/2018
<ul style="list-style-type: none">Recommended movies to users based on their rating history, including rating scores and genres of those movies.Trained encoder and decoder to make weights orthogonal and further found hidden features.Clustered users in nine groups using K-Means, recommended those movies with highest scores rated by users in the same cluster to a specific user.Other labs such as prediction of taxi peak period, using CNN, SVMs, PCA, Back-Propagation, etc. GitHub link.Packages: sklearn, pandas, scipy, numpy, matplotlib, sys, io, flickrapi, torch, keras, etc.	
"Jobster" Social Networking Site (PHP, MySQL, JQuery, AJAX, JavaScript)	04/2018-05/2018
<ul style="list-style-type: none">Designed and implemented a website like LinkedIn, where students could upload resumes, follow companies, apply for jobs and connect friends. Companies could post information and job announcements. GitHub link.	
Analysis of Browsing Data from Telecom Users (Java, ZooKeeper, Hadoop, Spark, ECharts)	06/2017
<ul style="list-style-type: none">Analyzed user's favorite app, website, time online and places that they had surfed at most frequently.Distributed design: one job tracker, two first-stage engines to calculated data, one second stage engine to merge and store data. Visualized data and formed concise and elegant statistical graph using ECharts.	
Collection & Analysis of Web Traffic Data (Hadoop, Flume, Kafka, Hive, MySQL)	04/2017-05/2017
<ul style="list-style-type: none">Transmitted logs using Log4jAppender in Java project to Flume cluster (Avro serialization). Off-line data saved on HDFS, processed using MR, Hive and transferred to MySQL storage using Sqoop.Transmitted real-time data to Kafka and calculated through Storm to realize monitor and alert function.	
Analysis of recruitment needs in China (Java, Jsoup, HTML, JSON)	01/2017
<ul style="list-style-type: none">Design a recruitment needs analysis website using data crawled from various recruiting websites through Jsoup.	