Hui (Henry) Chen

+1 347-223-1312 | hchen60@nyit.edu | github.com/hchen98 | hchen98.github.io

Education

New York Institute of Technology   **New York, NY**
Master of Science in Data Science   **Sep 2021 – Exp. May 2022**

Bachelor of Science in Computer Science   **Jan 2018 - May 2021**
- Concentration: Big Data Management and Analytics (CGPA: 3.51)
- Courses: Data Structures & Algorithms, Distributed Database, Information Retrieval, Intro to Big Data
- Involvement: Tech Lead, Google Developer Student Club (DSC)

Skills

- Python, Java, PHP, Flask, pandas, PySpark, selenium, Hadoop, PyTest, React Native, matplotlib, Jinja2, AWS EC2, JWT, Jenkins, Docker, Git, MySQL, MongoDB, HTML5, CSS3, JavaScript

Technical Projects

Personal Website: hchen98.github.io (for additional information and projects)

Scholar Recommendation App: Scholar Seek   **Feb 2021 – May 2021**
- Showcased a cross-platform mobile app that allows students to easily create their profile and get modular content-based filtering in scholarships, colleges, and majors to the NYIT engineering department.
- Implemented web-scraping for scraping 2.7m rows of semi-structured scholarship data and 3.5k+ rows of unstructured US-college data by using selenium with anti-captcha and real-time authentication.
- Designed and documented RESTful APIs for backend server to enable a secured and encrypted token integration of the client-side devices, recommendation models, and MongoDB through Flask and JWT.
- Configured backend for production and test environments, performed functional testing, and containerized the recommendation models, RESTful APIs, and web-scrapings to AWS EC2 through Docker.
- Utilized: Agile methodology, Waterfall Model, React Native, Python, Git, JavaScript, MongoDB, Flask, Jinja2, Selenium, AWS EC2, Docker, RESTful API, JWT, PyTest, Postman, Ngrok, and shell programming.

Big Data: MR Movie   **Sep 2020 – Dec 2020**
- Implemented a movie collaborative filtering using Cosine Similarity and Spark through the user ratings.
- Performed parallel computations for 1 million rows of MovieLens data in different nodes in the local cluster.
- Analyzed and interpreted model results by comparing different threshold and co-occurrence threshold values.
- Utilized: Git, Python, Apache Spark, Hortonworks Sandbox, shell script, and numpy.

Linear Regression: Airbnb Open Data (NYC)   **Sep 2020 – Dec 2020**
- Built linear regression models for Airbnb price prediction by computing all or without all other features.
- Created an interactive density map for 44k+ rows of semi-structured data for visualization and analytics through Folium.
- Established feature correlational matrix and importance graph for data preprocessing and feature selection.
- Evaluated and interpreted models' prediction results with an MAE of 39%, R2 of 41%, and RMSE of 55%.
- Utilized: Git, Python, scikit-learn, linear regression, Folium, seaborn, Jupyter Notebook, Scipy, pandas, numpy, matplotlib, and RapidMiner Studio.

Autonomous RC Car + Virtual Driving   **Feb 2019 – Apr 2019**
- Utilized PCA servo driver and Raspberry Pi to control RC car steering speed by integrating Donkey Car API.
- Trained a supervised classification CNN autopilot model with 11 layers by using collected image data with Keras for generating a hierarchical file that contains steering and throttle value.
- Showcased the project and result at NYIT Ventures' Pitch Contest to college faculties and students.
- Utilized: PCA 9685 Driver, Arduino, Python, Keras, TensorFlow, Donkey Car, Flask, 3D Printer.

Work Experience

SkyMobile Inc (startup)   **New York, NY**
Full-Stack Web Developer, Intern   Aug 2019 – Mar 2020
- Redesigned all existing websites for cross-platform and 1k+ records relational database.
- Improved website SQL injection vulnerabilities by implementing server-side script through PHP OOP.
- Implemented a secure payment system for online shopping for 200+ clients across NYC by using Stripe API.
- Reduced website latency by 10% by configuring DNS records and intergrading with Cloudflare DNS.
- Leveraged knowledge in Agile methodology, Waterfall Model, Git, Bootstrap UI, jQuery, Apache Server, MySQL, HTML5, CSS3, Ajax, jQuery, Apache JMeter, cPanel, Stripe API, and MVC design pattern.