Pengyu Wang

Website: phillipw.info
Github: github.com/cswpy

Email: pw1287@nyu.edu
LinkedIn: pengyu-wang

EDUCATION

- **New York University Abu Dhabi** — Abu Dhabi, UAE
  Bachelor of Science - Computer Science; Major GPA: 3.91 — *Sep 2019 - May 2023*
  Courses: *Algorithms, Operating Systems, Machine Learning, Deep Learning, Database Design, Big Data Analytics, Computer Networks*

SKILLS SUMMARY

- **Languages:**   Python, GoLang, SQL, C++, HTML/CSS, Bash, Java
- **Frameworks:**   Scikit, Numpy, PyTorch, Keras, Django, Flask, Pandas, Raft, Spark
- **Tools:**   Docker, Jenkins, dbt, Snowflake, BigQuery, Git, Regex, Linux, HPC
- **Databases:**   PostgresSQL, MySQL, SQLite, MongoDB, Hadoop, Hive, HBase, Flink, Kafka, HDFS

EXPERIENCE

- **TikTok Inc.** — San Francisco, USA
  Software Engineering Intern — *Jun. 2022 - Aug. 2022*
  - **Realtime Data Dashboard**: Implemented a streaming data pipeline with **Flink** in Non-China Monetization Data that monitors real-time Ads auditing queues, supporting 450k daily throughput at peak times at a global scale
  - Devised a **Lambda architecture** that integrates offline batch data with online streaming data to serve realtime dataflow
  - Built a future-proof pipeline with MQs like **Kafka** and **RocketMQ** and OLAP databases like **ClickHouse** and **Hive**

- **Nokia Corp.** — Hangzhou, China
  Software Engineering Intern — *Jan. 2021 - Mar. 2021*
  - **Project - Automatic Fault-Finding System for Documents**: Developed a website for automatically locating errors in technical documents and displaying on websites, serving 500+ engineers worldwide
  - Implemented a full CI/CD pipeline and release scheme on development and production servers with **Jenkins**
  - Devised a systematic testing framework using **PyTest** and accompanying guidelines for unit testing

- **combyne** — Munich, Germany
  Junior Data Scientist — *Jun. 2020 - Dec. 2020*
  - Delivered 5 monthly user retention reports with **BigQuery** and Google Analytics
  - Increased onboarding user retention rates by 2% and improved UX by working with cross-functional teams
  - Integrated **Jupyter** with data warehouses and compiled templates for stakeholders to harness data insights

PROJECTS

- **Grocery on Rails - online grocery platform with mobile app, admin panel, and RESTful API service**: Implemented a universal backend with authentication and CRUD service for clients and admin, and coordinated three sprints with three teammates as SCRUM master. Tech: Python, Flask, Flutter, MongoDB (March '21)
- **SimpleDB - class project for Database Systems Design**: Implemented core database components like Tuple, BufferPool, HeapFile, and Transactions management. Tech: Java, SQL, PostgreSQL (April '22)
- **Screen Helper - a webapp that helps users to foster healthy eye habits**: Developed the webcam utilization, video processing, pupil identification with DL model, and backend. Tech: Python, OpenCV, Flask. (November '20)
- **MindDoc**: Implemented a random forest that classifies electroencephalograph into healthy and MDD (Major Depressive Disorder) with 0.92 accuracy Tech: Python, Scikit-learn, Matlab, Numpy (July '21)

RESEARCH AND PUBLICATIONS

- **AI in Reducing the Use of Antibiotics in Urinary Tract Infections** — Abu Dhabi, UAE
  Cranmore Fellowship recipient — *Dec. 2020 - Mar. 2022*
  - Abstract accepted in Faculty of Clinical Information 22', paper under review
  - Trained a neural network that embeds EHR and feeds them into a Graph Neural Network for classification

- **Enabling Efficient & Accurate Evaluation on Complex Queries** — Abu Dhabi, UAE
  Capstone project — *Feb. 2022 - Present*
  - Extending relational databases to evaluate Package Queries that is a **NP-hard** problem
  - Utilizing synoptic techniques like sampling to further speed-up package query evaluation process

HONORS AND AWARDS

- Ranked 447/10724 globally in Google Hash Code Competition - February, 2020
- 2nd Place, NYUAD International Hackathon in Quantum Computing - April, 2022
- 1st Prize, Artificial Intelligence Competition of China - August, 2018

VOLUNTEER EXPERIENCE

- **Regional Manager & Head of IT at Education Without Barriers** — China
  Managed 15+ volunteers and liaised with a remote school that has 60+ kids — *Nov. 2017 - Jun. 2019*
- **Volunteer developer at wuhan.support** — China
  Involved in efforts to provide transparent & verified info during the COVID outbreak — *Feb. - Jul. 2020*