Archer Zhang

Email: ez806@nyu.edu | Tel: 646-906-7909 | Website: https://archertakesitez.github.io/

SKILLS

Programming Languages: Python, SQL, R, Java, HTML, CSS, JavaScript, TypeScript, Shell
Expertise: Exploratory Data Analysis, Machine Learning, Deep Learning, Cloud Computing, AWS, TensorFlow, PyTorch, Scikit-learn, Keras, NumPy, Pandas, NLP, LLMs, RAG, Generative AI, Data Pipelines, Linux, Distributed Systems, Git

EDUCATION

New York University, New York, NY   Sep. 2023 – May. 2025
Master of Science in Data Science   GPA: 3.833/4.0
Coursework: Programming for Data Science | Cloud Computing | Big Data | Machine Learning | Database Systems
Brandeis University, Waltham, MA   Aug. 2019 – May. 2023
Bachelor of Science in Computer Science, Bachelor of Arts in Economics, Minor in Mathematics   GPA: 3.749/4.0
Coursework: Data Structures | Data Management | Deep Learning | Data Mining | Natural Language Processing | Algorithms

PUBLICATION

GraphBERT: Bridging Graph and Text for Malicious Behavior Detection on Social Media. Published by ICDM-2022.
https://archertakesitez.github.io/static/assets/papers/GraphBERT.pdf
Fair Graph Representation Learning via Diverse Mixture-of-Experts. Published by The ACM Web Conference 2023.
https://archertakesitez.github.io/static/assets/papers/Representation_learning.pdf

PROFESSIONAL EXPERIENCE

Kineviz Inc.   San Francisco, CA (remote)
Data Analytics Intern   Jun. 2024 – Dec. 2024
- Incorporated **LLMs** (**Large Language Models**) into a **chat GUI** via **Python** and **JavaScript**
- Developed a **Cypher** query parser with ANTLR in Python, incorporating Named Entity Recognition (**NER**) and vector search to validate and enhance Cypher queries, enabling advanced search capabilities in **Neo4j** graph databases
- Built a website leveraging **FastAPI** that employs **LLMs** to translate plain texts into user-defined **JSON** schemas

Liangyouyinli Technology Co., Ltd   Beijing, China
Data Scientist Intern   May. 2023 – Aug. 2023
- Utilized **PyTorch** to implement convolutional neural networks (**CNN**) for stock price prediction in China's stock market

YUSUR Technology Co., Ltd   Beijing, China
Software Development Engineer Intern   May. 2021 – Aug. 2021
- Queried datasets from **SQL** databases (**MySQL**, **PostgreSQL**) and **NoSQL** database (**MongoDB**)
- Utilized **Apache Spark** and implemented **Shell** and **Python** scripts to process the Iris dataset on **distributed and parallel systems**, benchmarking the performance of our self-developed Kernel Processing Unit (KPU)

ACADEMIC PROJECTS

PaperGist: A Cost-Efficient Cloud-Native Research Paper Summarization Platform   Spring 2025
- Built a serverless **cloud-native** platform on **AWS** using Lambda, API Gateway, EC2 (for inference), SQS, EventBridge, S3, and DynamoDB with content-based deduplication to enable scalable, cost-efficient research paper summarization

Deck-to-CPT: AI-Driven Reimbursement Code Discovery for HealthTech Start-Ups   Fall 2024
- Built an **AI**-powered web application to process PDF pitch decks, leveraging Named Entity Recognition (**NER**) to extract key information from PDFs and Retrieval-Augmented Generation (**RAG**) to recommend accurate Current Procedural Terminology (CPT) codes

Fixplainer: Failure Explainer for Multiple Object Tracking (MOT)   Spring 2024
- Built a **GUI** tool to extract features and generate **SHAP** plots explaining object tracking outcomes in **MOT** frames
- Applied **YOLOv8** and **BoT-SORT** as object detection and object tracking tools on video datasets to create training sets

GraphBERT: Bridging Graph and Text for Malicious Behavior Detection on Social Media   Nov.2021 – Jun. 2022
- Participated in designing a model that focuses on detecting malicious tweets and users using both semantic information encoded by **transformers** (i.e., **BERT**) and relational information encoded by **graph neural networks** (**GNNs**)
- **Preprocessed** datasets obtained from the Internet, including dealing with wrong data rows and labeling the data