Zhongyu Pan

(778) 318 1663 | panzy0524@gmail.com
https://www.linkedin.com/in/zhongyu-pan-104a7b14b/
Montréal, QC

Education

Master of Data Science in Computational Linguistics
University of British Columbia Graduated July 2020
Curriculum: Machine Translation, Sentiment Analysis, Computational Semantics, Advanced Corpus Linguistics, Data Manipulation, Algorithms and Data Structures, Data Visualization, Machine Learning, Databases and Data Retrieval.

Bachelor of Science in Financial Mathematics
Hefei University of Technology Graduated May 2019

Skills

- SQL (SQL Server, MySQL, PostgreSQL)
- Python (PyTorch, Tensorflow, scikit-learn, Pandas, NumPy, SciPy, MatPlotLib)
- Tableau (Data Analysis & Data Visualization)
- Natural Language Processing
- Web Development (HTML, CSS, JavaScript; React.js)
- Others: Git, Jira, Agile, ETL, Data Pipeline, Project Management
- Languages: Native proficiency in Mandarin & English, conversational proficiency in French

Work Experience

Data Science Instructor / Content Creator (Contract) – LinkedIn – Remote May 2021 - Present
- Released Natural Language Processing with PyTorch course on LinkedIn Learning Platform, reached 3000 learners in 2 months.
- Drafted Python NLP project code and presentations, held weekly meetings with content manager.

Agile Developer *(Data Wrangling Team)* – SAP – Vancouver, BC Sep 2020 - April 2021
- Employed **SQL** for data retrieval and in-depth analysis of user data, enhancing workflow efficiency by integrating re-editing capabilities to SAC (SAP Analytics Cloud) using **React.js** in an MVC pattern.
- Achieved a notable 20% reduction in data retrieval time, optimizing system responsiveness and user experience.
- Developed backend logging functionality using **Node.js**, facilitating usage tracking specifically focused on user click behaviors within the application.
- Integrated data collection services with data wrangling processes, consolidating user data flow for seamless analysis and insights extraction.
- Implemented usage tracking mechanisms employing a **pub/sub (publish-subscribe)** approach to capture and analyze user click behavior within SAC workflows, enhancing analytical capabilities and user-centric decision-making processes.

Data Scientist (Capstone Project) – Minerva Intelligence – Vancouver, BC April 2020 - June 2020
- Conducted comprehensive **data collection and cleansing** from BC geology database using an **ETL** process, ensuring data integrity and consistency.
- Built an information extraction system for BC mining files corpus, implemented **Artificial Intelligence** models for mapping text to desired information fields.
- Actively participated in weekly company meetings, presenting findings and insights derived from the analysis, fostering collective decision-making.

Financial Data Analyst – Beijing Slow Snowball Investment – Beijing, China June 2018 - June 2019
- Researched and developed quantitative trading algorithms using **Excel and Python**, ensuring robustness through meticulous back-testing procedures.
- Analyzed extensive financial datasets, synthesizing findings into research reports with graphical representations, delivering valuable business insights.
- Published reports in Quantitative community, showcasing methodologies and results from rigorous back-testing, contributing to industry knowledge.

Projects

Information Extraction System for Geology Reports – Client's Project (Minerva Intelligence)
- Constructed an **information extraction system** deployed within the company's infrastructure based on BC geology reports **database**.
- Implemented Convolutional Neural Network (**CNN**) in **PyTorch** with enhanced features and **machine learning** algorithms with **sklearn** to automatically extract targeted information from unstructured text in geology reports.
- Significantly improved **text data classification** accuracy from 88% to 96% by refining the baseline CNN.

Lyrics Transliterator – University Project
- **Web scraped** over 1000 Chinese and English lyrics data from Lyrics.com using **Python's BeautifulSoup**.
- Developed a **Seq2Seq** Model for the system to output transliterations from original song lyrics.
- Applied **data transformation and filtering** techniques, including aggregating and filtering functions, to enhance the efficiency of the reporting process.
- Collaborated with a team of Master of Data Science students at UBC to create a transliteration app aiming to enable singers to perform songs in foreign languages by leveraging **artificial intelligence techniques**.
- Initial focus: Building a transliteration system for song lyrics from Chinese to Pinyin (phonetic English form), for the convenience of music lovers..

Matcha Order Chatbot – Personal Project
- Developed an order chatbot for a matcha business, focusing on **prompt engineering** and user experience optimization.
- Applied **prompt engineering** techniques to **fine-tune** prompt generation algorithms for improving user interactions.
- Used advanced **Natural Language Processing models**, including **transformer architecture**, to ensure accurate understanding of user inquiries.