DOUGLAS QUAN

Toronto, Ontario

📞 437-351-1092 ✉ douglas.quan@mail.utoronto.ca 💼 linkedin.com/in/douglas-quan-ca ⬛ github.com/douglasquan

EDUCATION

University of Toronto   **Sep 2021 – May 2025**
Honours Bachelor of Science, Computer Science Specialist & Mathematics Minor   *CGPA: 3.51/4.0*
- Dean's List Scholar
- **Relevant Coursework:** Data Structures, Software development, Web Development, Database Management, Systems Programming, Artificial Intelligence, Machine Learning, Neural Network and Deep Learning, Information Security, Computer Vision, Computer Architecture, Computer Organization, Operating Systems, Computer Networks

EXPERIENCE

Lillup   **Jan. 2024 – current**
Development Engineer Intern (NLP, Machine Learning and LLM)   *Remote*
- Assisted in the development and optimization of NLP, ML, and LLM algorithms and applications
- Participated in the collection, cleaning, and analysis of large datasets to train and test AI models
- Contributed to the implementation of machine learning pipelines and model deployment strategies
- Collaborated with the team to conduct experiments and evaluate the performance of AI models

University of Toronto, Department of Economics   **Sep. 2023 – current**
Research Assistant, Supervisor: Loren Brandt   *Toronto, Ontario*
- Developed a **web scraping** mechanism in **Python** to scrape industrial policies and patenting data from PKUlaw.com
- Leveraged **Cohere Classify**, a **large language model (LLM)**, to classify over **300k** policy documents effectively
- Applied **linear regression** to quantify the temporal and spatial influence of policy documents
- Analyzed datasets with **Excel Pivot Tables** to identify relationshps between different industry classification systems
- Wrote Python scripts using **Pandas** and **Seaborn** to clean, preprocess, visualize, and analyze datasets

University of Toronto, Department of Linguistics   **Sep. 2022 – Sep.2023**
HLVC Research Assistant, Supervisor: Naomi Nagy   *Toronto, Ontario*
- Contributed in a Jyutping-to-Chinese transliteration project, focusing on data analysis, sourcing, and cleaning for an **RNN model**, achieved a character error rate (CER) of **0.21**
- Developed **Python** scripts to automate repetitive tasks and data entry process with **PyCantonese** and **ELAN API** on **63** Cantonese interview transcriptions
- Utilized **FileZilla** for file management of interview data, and **ELAN** for **NLP** analysis
- Ideated and Developed a **corpus-based imputation method** to solve inconsistency in transcriptions
- Recruited Torontonian participants who speak Cantonese as a heritage language for an interview

PROJECTS

MeetHomie | *Django, React.js, Javascript, REST API, JWT Authentication, Tailwind CSS*   **Jan. 2024**
- Deployed a **full-stack** scheduling platform designed to streamline the organization of regular one-on-one meetings
- Utilized **React.js** for the frontend development, creating a **responsive and interactive user interface**
- Developed a comprehensive **database schema** and implemented a **RESTful API** with thorough documentation and **Postman testing**, utilizing the **Django** framework for robust back-end functionality

WeBite | *Flask, React.js, REST API, SQLite, Tailwind CSS*   **Jan. 2024**
- Contributed as a **backend developer** in a food ordering app project aimed at connecting home cooks with customers
- Employed a **microservices architecture** for the backend with **Flask** and designed a **SQLite**-based database schema
- Adopted **Agile** development practices within the team, participating in sprint planning, daily stand-ups in a group of 7
- Utilized **Jira** as project management tool, efficiently tracking and managing backend tasks, bugs, and features

J2C Translator | *Python, Deep learning, PyTorch, NLP*   **Nov. 2023**
- Developed the first seq-to-seq Jyutping-to-Chinese-Character translator with **Transformer Model Architecture**
- Performed **data cleaning** on over **18000** training examples from three Cantonese conversational corpora datasets
- Improved validation accuracy by **17%** compared with the baseline model using **RayTune** for hyperparameter tuning
- Designed a weighting mask to solve the homophone ambiguity problem in J2C translation, achieving **94%** test accuracy

TECHNICAL SKILLS

Programming Languages: Python, Java, C, HTML/CSS, JavaScript, R, SQL
Developer Tools: Unix/Linux, GitHub, Vim, RStudio, Figma, Google Colab, Jupyter, Excel, kali Linux, VSCode, PyCharm, IntelliJ, Postman, Docker, Jira
Frameworks and Libraries: React.js, Django, Node.js, Flask, FastAPI, Bootstrap, Tailwind CSS, SQLite, PostgreSQL, Pandas, Scikit-learn, TensorFlow, PyTorch, Numpy, Matplotlib, OpenCV

Languages

Languages: Cantonese, English, Mandarin