COREY YANG-SMITH

https://coreyyangsmith.github.io/ 🌐

coreyyangsmith@gmail.com ✉

Objective

Resourceful and hands-on AI engineer with production experience deploying LLM-powered tools across full-stack systems. Currently completing a PhD focused on automating software workflows with large language models. Eager to join a high-agency high-growth team shaping the future of AI-native workflows.

Education

Doctor of Philosophy, Software Engineering   **University of Calgary, Present**

Master of Engineering, Software Engineering   **University of Calgary, Apr 2024, GPA 3.97**

Bachelor of Science, Civil Engineering *(Minor in Entrepreneurship)*   **University of Calgary, Dec 2020, GPA 3.48**

Work Experience

Software & Machine Learning Developer, Urban Systems Ltd.   **May 2023 - Current**

- Spearheaded new AI practice area within company, scoping 100+ internal projects across 15+ practice areas and service offerings, working with colleagues to identify over $10,000,000 of potential cumulative annual savings through the development and deployment of applications & machine learning tools. Scaled department from 1 to 4 people.

- Built and deployed a Neo4j-based graph system to enhance ERP data with visual insights for client, branch, and career planning. Leveraged GraphSAGE to develop an employee recommendation engine supporting team building.

- Deployed a local large language model to over 700+ employees utilizing Open Web-UI and Ollama for model handling. Experimented with many LLM technologies and inference engines (vLLM, Qdrant, DSPy, langchain, etc).

- Developed a custom-trained computer vision model, fine-tuned on a pre-trained Yolov8, to identify defects in underground utility CCTV videos amongst 11 classes, reducing cost of video processing by 95%.

- Developed and deployed an asset management dashboard that streamlined forecasting and reduced risk, generating $20K in revenue within the first month.

- Advised on specialized hardware needed to run in-house AI machine, enabling the company to deploy LLMs locally to conserve data privacy and confidentiality with clients, and perform other deep learning tasks and training.

Teaching Assistant (Full Stack Web Development), University of Calgary   **January 2024 - April 2024**

- Facilitated lab sessions on a weekly basis to over 40 students, explaining new concepts thoroughly with examples and assisted in answering students' questions, focusing on the fundamental concepts of full stack web development.

Engineering Consultant, Urban Systems Ltd.   **June 2021 - April 2023**

- Developed automated photo organizing and reporting tool in Excel VBA, integrated with Microsoft Suite, reducing the time of daily report generation for site inspectors by 30mins/day equating to client cost savings of over $50,000/yr.

Estimating/Engineering Student, PCL Construction   **January 2018 - August 2018**

- Created historical costing database in Excel with data from over 300 completed jobs, leveraging advanced excel functions and VBA to assist in the estimating process. Observed pricing trends based on inflation, size of project, location, type of building, materials used, and more.

- Leveraged data to provide company insights on a variety of projects and metrics through the use of a business intelligence tool, Power BI.

Student Project Coordinator, Cyntech Construction Ltd   **May 2018 - December 2018**

- Developed pricing automation tool leveraging Excel VBA, reducing design-to-pricing time on projects by 20x while eliminating errors, enabling the company to bid on billion dollar jobs otherwise too resource-intensive to consider.

- Automated standard CAD drawings for helical piles using Excel VBA and AutoCAD Command Line, eliminating the time it takes to produce quality fabrication drawings and reducing errors.

- Created pile spacing tool leveraging Excel VBA, reducing time to iterate on engineering design of group piles by 50%.

Software Projects

Eye Tracking Research Tool - Data Collection Web App   **Express (TS), React (TS)**

- Developed eye-tracking data collection pipeline for video content using tensorflow.js, running face detection model and training custom CNN on user data to track eyes and collect predicted gaze, achieving ~80% model accuracy.
- Fine-tuned EfficientNetV2B0 with public eye dataset to improve eye tracking model, performing transfer learning and further calibrating on user-collected eye data to produce remarkable eye-tracking algorithm.
- Generated heat map visualizations over source videos from collected eye gaze data, processing user data using matplotlib and ffmpeg, allowing researchers easily analyze experiment results.
- Created Docker images for both front-end and back-end codebases, writing dockercompose file to easy deployment.

Calgary Housing Price Predictor - Big Data ML Predictor   **PySpark**

- Developed scalable scraping pipeline based on 9 publicly available Canadian datasets and extracted 16 unique features, inspecting data for validity and filtering results to produce 1,000,000+ data points to train our ML models.
- Explored 3 regression models, tuning various hyper-parameters to predict housing prices based on novel input, achieving 54% testing accuracy using a simple linear regression model.

Exploring Model Architectures for Brain Tumor Classification

- Performed extensive literature review on existing techniques for the identification and classification of brain tumors in MRI scans, evaluating the performance of three deep learning methods: VGG2D, ResNet50V2, and Xception.

3D Pile Spacing - Engineering Tool Web App   **React, Three.JS (TS)**

- Developed a 3D helical pile-spacing engineering tool leveraging three.js, allowing users to tweak design settings, automatically providing calculations and checking for design conflicts.

Goldnest - Personal Finance Web App   **Django (Python), React (JS)**

- Generated all documentation for front-end design including wireframes and design documents using Figma. Fully implemented front-end using React, utilizing charting libraries D3, e-charts, and Chart.js for dynamic visualizations.
- Handled user registration, login, and authentication with JWT, encrypting user data with Fernet and a private key to safely store user data in PostgreSQL, managed by a Django back-end.

Publications

Corey Yang-Smith, A. Abdellatif. "Tracing Vulnerabilities in Maven: A Study of CVE lifecycles and Dependency Networks"
To appear in Proceedings of the 2025 International Conference on Mining Software Repositories (MSR 2025).

Technical Skills

Programming & Development Tools

- **Python (FastAPI, Django)**: Experienced in data pipelines, machine learning, AI and backend development.
- **JavaScript (Node.js, React)**: Proficient in developing full-stack web applications and chatbot integrations.
- **Docker**: Created and managed persistent storage docker containers for service deployment and management.
- **Github Actions/Workflow**: Integrated CI/CD pipeline for deployed projects to enhance developer efficiency.
- **Redis:** Implemented Redis for in-memory caching to significantly reduce database load and improve performance.
- **Elasticsearch:** Deployed and integrated on-prem semantic search, enabling fast and relevant document retrieval.

Artificial Intelligence Tools & Libraries

- **Libraries/Frameworks:** LangChain, , LangGraph, llama-index, transformers, TensorFlow, PyTorch, Keras, DSPy
- **Inference Engines:** Ollama, vLLM, Triton
- **Vector Databases:** Qdrant, Chroma, Weaviate

Data Processing & Analytics

- **Tools & Technologies:** pandas, NumPy SQL, Spark, D3.js, matplotlib, Plotly, Tableau, Power BI, FME
- **Data Scraping**: GitHub API, Selenium, Requests, Beautiful Soup, lxml