Greater Vancouver Area
BC, Canada

Gautham Pughazhendhi

Portfolio | LinkedIn | GitHub | Kaggle

(+1) 672-514-9943
gautham.pughazhendhi@gmail.com

Education

Master of Data Science candidate, University of British Columbia, **Current Grade**: 98% — Sep 2021 — Jun 2022
Bachelor of Engineering (Computer Science and Engineering), Anna University, **CGPA**: 8.44/10 — Jun 2014 — Apr 2018

Experience

- **Machine Learning Consultant, Trusting Pixels Inc., Canada (MDS Capstone Project,** unpaid co-op) — May 2022 — Jun 2022
  - Building a **computer vision**-based (**CNN**) model using **PyTorch** to authenticate **thousands of photos** per day by detecting **cloning** (pimple/scar removal), **warping** (body shape enhancements), and **inconsistent noise patterns** in the edited images.
  - Contributing to the **design, development**, and **deployment** of an **image comparison product** by analyzing a large dataset of over **200 GB** to determine the amount of retouching done and validate its compliance with the retouching regulations of the company.
- **Teaching Assistant, The University of British Columbia, Canada** — May 2022 — June 2022
  - Teaching assistant for the course **DSCI 100 (Introduction to Data Science)**, responsible for educating the undergraduate students on the use of data science tools to summarize, visualize, and analyze data as part of the teaching team.
- **Software Engineer - Machine Learning, Sirius Computer Solutions, LLC, A CDW company, India** — Jun 2018 — Jul 2021
  - **Reduced the total budgeted manpower cost by 20%** for a US banking firm by building an **NLP-based contextual chatbot** with smart KB **article recommendations** to automate the firm's issue-creation for internal requests in Salesforce and ServiceNow. Fine-tuned the **intent classification and dialogue models** and contributed to the **design and architecture** of the **NLP pipeline**.
  - **Minimized the support team turnaround time by 20%** for a US global payments company by developing a **classification model using LSTM**, a recurrent neural network (RNN), and **eliminating the manual categorization** of emails. Built a **data pipeline** to clean, transform, and derive additional features; articulated and presented the **data insights** to the onsite team.
  - **Cut down the allocated workforce cost by 70%** by building multiple **conversational AI assistants** to automate the mission-critical IT operations of a US retail MNC; automated the CI/CD pipeline of bots using **Azure** DevOps and Kubernetes Service. Improved the **security** of bots and monitored the **performance of ML pipelines**. Created **Python API** for integration with external services.
  - Contributed to the design of **NoSource**, an **AI-based testing automation framework** to perform a website's QA testing without test suites by seamless identification of HTML elements in a page. Enabled cloud testing of web and mobile platforms for **OneSource**, a testing framework, and **reduced the procurement cost of test devices by 50%** by leveraging cloud-based testing simulations.

Major Academic Projects

- **StrapvizPy and StrapvizR (Graduate):** Created **Python and R packages** named StrapvizPy and StrapvizR to streamline bootstrapping samples, creating insightful plots and tables with statistics such as confidence interval, standard error for reports and papers. 2022
- **Forest Fire Area Prediction (Graduate):** Trained and tuned a **Support Vector Regression**(SVR) model to predict **forest fire areas** using weather and soil data; improved the model by removing outliers using the Cook's distance method. 2021
- **Earthosys (Undergraduate Capstone):** Developed an **ensemble model** using Random Forests to predict **tsunamis**. **Implemented an efficient search algorithm** on the **NASA** dataset to find the nearest coastal distance with coordinates from the **NOAA**'s tsunami dataset. Developed a **web application**'s backend, a **chatbot**, and an **IoT-based alert device** using **Raspberry Pi**. 2017 — 2018

Technical Skills

Languages: Python, R, SQL. **Frameworks:** Tensorflow, PyTorch, RASA, React, Flask, Django. **Databases:** MySQL, Postgres, Mongo.
Libraries: Scikit-Learn, Pandas, NumPy, Spacy, Gensim, Matplotlib, Altair, Keras, PySpark, SQLAlchemy.
Tools: Jupyter, PyCharm, VS Code, RStudio, git, Docker, Sourcetree, Android Studio. **Cloud Platforms:** Azure, AWS.

Hackathons

- **WiDS 2022 Datathon, Stanford University**: Ranked **1**st in Vancouver and **16**th overall in the 5th Annual WiDS Kaggle Datathon. Built an ML model to predict the energy efficiency of buildings to help policymakers target plans that maximize emission reductions.
- **HCL Commerce hackathon 2020**: Won the **Most Creative** award with my team for implementing a smart voice assistant to shop on an e-commerce site and integrating it with **Google Assistant**.
- **FLEX 2.0, Codes and Gears 2020, a 36-hour Hackathon, Sirius Computer Solutions (Winners):** Built a **computer vision**-based prototype to convert sign language into words on-screen to help differently-abled with speech and hearing loss. **Demo**
- **FLEX, Codes and Gears 2019 (Winners):** Developed a **computer vision**-based prototype to detect bad postures, control smart home devices, and convert gestures into action words by estimating body poses to help people with disabilities. **Demo**

Awards

Feather in the cap 2019; ACE performer of the year 2018 (Sirius Computer Solutions)
First place at **IEEE CS Project Expo 2016; Third** place at **IEEE CS Coding Contest 2016** (Velammal Engineering College)