Swapnil Dubey

Portfolio | sdubey14@student.ubc.ca | linkedin.com/in/swapnildub/ | github.com/Swapnil-Dubey

Technical Skills

Technologies: C/C++, Python (TensorFlow, Scikit-learn, Matplotlib), R, SQL (in progress), Java, HTML/CSS, LaTeX, Scheme/Dr. Racket, Git/Github, Excel, Visual Studio
Testing: JUnit, GDB, unittest

Work Experience

Stanford Code in Place Section Leader Apr '24 - May '24
Stanford University *Remote*

- Led a cohort of 20 students through a comprehensive curriculum covering Python programming fundamentals, including various data structures, algorithms, achieving 90% student retention, demonstrating commitment and enthusiasm.
- Fostered an inclusive learning environment by developing interactive lesson plans and engaging in effective communication with students.
- Collaborated with 3 other instructors to refine teaching strategies based on comprehensive analysis of student performance data; this initiative has positively impacted students and increased retention rates by 15%.

Statistics Tutor Feb '24 - Present
Sealy Tutoring *Vancouver, Canada*

- Enhanced student performance by delivering tailored tutoring sessions for high school and undergraduate students in elementary and intermediate statistics.
- Expanded student knowledge by covering advanced topics such as Exploratory Data Analysis, Probability, Statistical Inference, Regression, A/B Testing, and Hypothesis Testing, ensuring students grasped complex statistical concepts.
- Innovated learning materials by creating customized lesson plans and problem sets, incorporating software tools like R and Python, which increased student proficiency in data analysis.

Orientation Leader Aug '23 - Sep '23
University of British Columbia *Vancouver, BC*

- Led a cohort of 30+ incoming first-year students at UBC, connecting them with campus resources and developing a welcoming and supportive community.
- Organized a series of social and educational events during orientation week, showcasing strong communication and interpersonal skills to address the need for community building among new students.
- Provided tailored support to individual students by identifying their specific needs and connecting them with appropriate campus resources

Technical Projects

Understanding & Predicting Heart Disease Presence | *(Academic)* | *R, Logistic Regression* Jul '24 - Aug '24

- Conducted an academic research project to predict heart disease using clinical data. The goal was to identify key predictors and develop an accurate model for early detection, aiding in effective clinical decision-making.
- Led data preprocessing and performed Exploratory Data Analysis (EDA). Identified critical predictors which guided feature selection and model development.
- Applied forward selection and LASSO regression to build predictive models, optimizing for accuracy while avoiding overfitting. Model achieved a 79.8% accuracy rate, validated through cross-validation, alongside a team of four.

Machine Learning Analysis of Internet Usage and Mental Health | *(Research Project)* May '24 - Present

- Analyzed the association between internet usage and perceived mental health among Canadians aged 15-44 using the Canadian Internet Use Survey (CIUS) dataset.
- Implemented decision tree algorithm using the R package `rpart` to model the relationship between internet usage and mental health and visualized the resulting classification tree using `rpart.plot`.
- Executed data pre-processing, including record filtering, feature engineering, outlier detection and hyperparameter tuning for model optimization.
- Achieved 78% accuracy in classifying unseen data and identified key predictors influencing mental health. Currently collaborating with a professor to publish findings in a peer-reviewed journal and contributing a book chapter on applying machine learning for social innovation.

Pet Pal: Pet Care Tracker in Java | *(Academic)* | *Java, Swing, JSON, Unit Testing* Jan '23 - Mar '23

- Developed a Java desktop application using Java Swing for the GUI, enabling users to efficiently record and manage their pet care activities, including feeding, watering, treating, and walking, for a Software development course.
- Using JSON for data storage and loading, developed a feature to record and update interactions, such as feeding and watering, with historical tracking for each pet, ensuring easy access to past pet-care activities.

- Integrated a recommendation system that provides personalized care suggestions based on recent pet activities, enhancing convenience and efficient management for users. The project was executed in iterative phases, demonstrating Agile delivery of software.

Water Potability Classification | *(Academic)* | *R, Supervised Machine Learning, Classification* Oct '22 - Dec '22

- Developed a K-nearest neighbors (KNN) model to classify water samples as potable or non-potable based on various chemical and physical properties, using a synthetic dataset of 3000+ records.
- Using R for data pre-processing, conducted exploratory data analysis using tidyverse and GGally to visualize relationships and prepare data for analysis. Performed k-fold cross-validation to evaluate model performance and fine-tuned the KNN algorithm to optimize accuracy.
- Achieved a classification accuracy of 70.6% on unseen data, as determined by k-fold cross-validation. Presented findings, representing UBC at Google's Solutions Challenge 2023.

Text Reminder | *(Hackathon)* | *HTML, CSS, JS, Python* Jan '23

- Developed an automated text-reminder application during nwHacks 2023. This applications allowed users to setup one-off or recurring reminders, enhancing time-management and productivity.
- Helped in developing the back-end of the website to ensure real-time scheduling of reminders.
- Collaborated in a team of four, to design a user-friendly GUI. The interface allowed users to easily input phone numbers, customize messages, and set schedules, resulting in positive feedback for usability and functionality from judges during demonstrations.

Amazon Review Emotion Analysis | *(Personal)* | *Python, Matplotlib, Tkinter, NLTK* Jul '22 - Aug '22

- Developed an Amazon product review analyzer using Rainforest API to fetch product reviews and analyze emotional tones with the text2emotion library, providing summarized emotion analysis of the reviews.
- Visualized the distribution of emotions from emotion analysis using Pie-charts from Matplotlib that provided user with a clear graphical representation of customer sentiments.
- Highlighted prevalent emotions, enabling visually-aided and strategic decision making.
- Designed the GUI using Tkinter to enable users to conveniently enter amazon product URL/info and view the sentiment analysis results.

PyGamez | *(Personal)* | *Python, SQL* May '22 - Jul '22

- Developed a console-based Python application consisting of three games: Rock paper scissor, Hangman and Quiz, connecting to a database using Python-MySQL connectivity to ensure real-time user score-keeping and progress.
- Integrated a MySQL database to store player scores and statistics providing the ability to track individual performance over time.

Technical Extracurricular Activities

Online courses

Remote

- Completed Harvard's CS50 series including CS50x, CS50AI, and CS50Python, gaining a strong foundation in computer science fundamentals, artificial intelligence, and Python programming.

Volunteer Experience

External Director Aug '23 - Feb '24

UBC Exchange and International Students Club *Vancouver, BC*

- Secured funding and sponsorship relations with $\sim$ 25 businesses for club initiatives.
- Strengthened sponsor partnerships by delivering on partnership expectations and fostering long-term relations.
- Improved financial management by working closely with the treasurer, ensuring efficient use of funds.

Activities/Interests

Cycling & Hiking Sep '22 - Present

- Rediscovered a passion for cycling and hiking; applying to UBC's Thunderbird Cycling Team and regularly exploring notable trails like Mt. Rainier, Grouse Mountain, and Lynn Peak.

Education

The University of British Columbia Vancouver, BC, Canada

3rd Year in BSc in Computer Science and Statistics (Dean's Honour List), GPA: 3.7/4.0 *Sept '22 - May '27*

Specialized Coursework: Relational Databases (in progress), Data Structures & Algorithms, Machine Learning, Statistical Inference, Regression

UBC Science Co-op Jul '24 - Present