Arnav Kumar

a8kumar@uwaterloo.ca arnavcs arnavcs.github.io Arnav Kumar

Experience

Data Science Intern @ Tektorch.ai Apr 2022 - Aug 2022
- Extracted features such as job title, location, classification, and pay from a job listing website using Python, Beautiful Soup 4, NumPy, and Pandas to create a dataset for analysis.
- Developed an end-to-end data visualization pipeline to ingest data and display trends using Matplotlib, and NumPy.

Machine Learning Research Intern @ University of Alberta Jul 2021 - Aug 2021
- Trained and employed natural language processing models to help determine the cause of a change in depressive language in Tweets as part of a psychology study.
- Created an evaluation framework with graphs and videos to analyse subject mouse position data during exams.
- Developed automation pipeline in Bash to reduce time spent by 20x on the application of the evaluation framework.

Coding Instructor @ Code Ninjas West Springs Mar 2022 - Apr 2022
- Taught over 45 students to complete programming challenges in Scratch, MakeCode, and JavaScript.
- Managed the progress of the students and directly relayed feedback to the manager.

Projects

InterPlanetary File Explorer (IPFE) | *Python, Scikit-learn, Estuary, Co:here, Three.js*
- Created vector embeddings for files with their headers using Co:here's NLP embeddings to facilitate classification of files.
- Performed principal component analysis of the vector embeddings to reduce the dimensionality from 4096 to 3 to be plotted and displayed interactively in 3D space using Three.js.

Prognosing Idiopathic Pulmonary Fibrosis (IPF) | *Python, Tensorflow2, Pandas, Scikit-learn*
- Implemented an auto-encoder, linear regression, dense neural network, and bayesian model in order to accurately predict future lung capacity and give a confidence value using initial lung capacity data, age, sex, smoking status, and more.
- Obtained a Laplace Log Likelihood score of -6.9, much better than the baseline score of -8.1, and a standard deviation of approximately 200mL.

CLI Chess Engine | *Haskell*
- Employed the use of applicative functors and monad operations to write concise, readable code for finding the cartesian product of two lists and creating safe functions.
- Created custom Haskell data types (such as a position data type) and ensured that no data could be created which has no meaning in the context of the program.

Braille ASCII Art Generator | *Python, OpenCV*
- Used thresholding in OpenCV to determine which pixels of the output art should be shaded.
- Transformed the image array into braille characters using 3x2 pixel filters which returned a binary number corresponding to a braille character.

Blackjack | *C++*
- Used object oriented programming to implement a local multiplayer blackjack game with a computer playing as the house.
- Probabilistically optimized the holding value that the computer house uses.

Achievements

- 335^{th} (top 10%), with score 26, **William Lowell Putnam Mathematical Competition** 2022
- Estuary Sponsor Prize Winner, **UofTHacks X** 2023
- 10^{th} in Canada, **Asian Pacific Mathematics Olympiad** 2022
- 11^{th} in Canada, **Canadian Association of Physicists High School Exam** 2022
- 2 time qualifier, 18^{th} in Canada, **Canadian Mathematical Olympiad** ..... 2021 - 2022
- Invited to write, **United States of America Mathematical Olympiad** 2022
- Bronze medalist for machine learning project on prognosing IPF, **Canada Wide Science Fair** 2021

Education

Candidate for BSc. Computer Science @ University of Waterloo 2022 - Present
- Received the Ronald G. Dunkley National Scholarship ($18,000), the President's Entrance Scholarship ($2,500), and the Cenovus Energy STEM Scholarship ($14,000).
- Core average of 97% with all advanced level courses and recieved term distinction.

Skills

Languages: Python, C++, JavaScript, Haskell, Java, Bash, LaTeX, HTML, CSS, Markdown, RegEx
Frameworks: Linux, Unix, Pandas, Matplotlib, NumPy, Beautiful Soup, Tensorflow, Scikit-learn, OpenCV