Kanika Chopra

647-721-1223 kanikadatt@gmail.com linkedin.com/in/kanikadchopra github.com/kanikadchopra

Education

University of Waterloo **Sept 2017 – April 2022**
Bachelor's of Honours Mathematics in Statistics [GPA: 88%] *Waterloo, ON*

Experience

University of Waterloo **Feb 2021 – Dec 2021**
Teaching Assistant, CFM101 *Waterloo, ON (Remote)*

- Collaborated with Dr. James R. Thompson to design a new financial analytics course for first-year CFM students
- Designed four Python assignments along with solutions to test students' knowledge on financial data analytics tools
- Taught students how to use pandas, NumPy, matplotlib and yfinance for data analysis through tutorials

Uber **May 2021 – July 2021**
Data Analyst Intern, Safety and Insurance *San Francisco, CA (Remote)*

- Developed data pipelines to measure safety product metrics and designed a Tableau dashboard to monitor performance
- Collaborated with product managers, data scientists and engineers to understand the business use cases and present metric read-outs
- Conducted deep dive investigations into driver vehicle data to identify potential business opportunities

Wish **Jan 2021 – April 2021**
Data Analyst Intern, Data and Relevancy *San Francisco, CA (Remote)*

- Built four complex data-pipelines to productionize a classification model, decreasing time-to-run by 75% using Airflow
- Developed a deterministic shipping time model and Holt-Winters Additive forecasting model to predict future merchant cash flows
- Devised a deterministic model to predict empty packages shipped by merchants to decrease fraudulent activity with 95% accuracy

University of Waterloo **Sept 2020 – Dec 2020**
Undergraduate Marker, STAT231 *Waterloo, ON (Remote)*

- Graded assignments for over 400 students on topics covering statistical analysis, confidence intervals and hypotheses tests

Intact **May 2020 – Aug 2020**
Data Scientist, DataLab *Toronto, ON (Remote)*

- Conducted deep dive investigations on model predictions for 420,000 images to find improvement strategies
- Automated the development of a multi-page analysis spreadsheet using pandas to understand the accuracy and precision of model performance, decreasing time-to-create by 95%
- Researched and implemented image augmentation techniques using OpenCV to improve classification by 3%

Goldspot Discoveries **Sept 2019 – Dec 2019**
Data Scientist *Toronto, ON*

- Trained a fast.ai NLP classifier to conduct a sentiment analysis on 18,000+ gold-related tweets with 85% accuracy
- Performed PCA, K-Means Clustering, SLIC and multi-band ratios on satellite images for mineral composition analysis
- Visualized interactive time-series plots using Plot.ly for a stock price technical analysis with 20-, 50- and 200-day moving averages

Royal Bank of Canada **Jan 2019 – Apr 2019**
Strategic Initiatives Analyst Intern *Toronto, ON*

- Automated team reporting by determining metrics using GitHub verbiage stored in SQL server and then visualizing the data in a Tableau dashboard for management decisions
- Visualized event registration data with a Tableau storyboard to display a demographic and geographic analysis for the marketing team

Technical Skills

Languages: Python, R, SQL, Presto, MATLAB
Tools: Airflow, Tableau, SQLServer, Hive SQL, Treasure Data, Piper
Libraries: pandas, NumPy, SciKit-Learn, SciKit-Image, matplotlib, Plot.ly, spaCy, nltk, OpenCV, StatsModels

Research Projects

projplot (Dr. Martin Lysy) | *pandas, NumPy, seaborn* **In Progress**

- Developed a plotting tool a plotting tool, projplot, to assist users with testing optimization when building optimizers by providing additional visualizations

Classifying Tumour Types (STAT441) | *pandas, NumPy, SciKit-learn, matplotlib* **Dec 2021**

- Augmented image data using color masks, mosaics and random transformations to solve the class imbalance problem
- Experimented with pre-trained CNNs to classify tumour types based on the transformed histopathological data

Netflix Browsing Time Experiment (STAT430) | *R* **Aug 2021**

- Conducted an experiment to investigate three factors (preview length, match score, and tile size) to determine the optimal combination to minimize average browsing time spent on the homepage
- Performed factor screening by exploring the main and interaction effect plots to determine significant factors
- Implemented the method of steepest descent and response optimization to locate the optimum for significant factors

Leadership / Extracurricular

Tech+ **May 2021 – Jan 2022**

Co-Director

- Led a team of 50 students to discuss and advocate for diversity, equity, and inclusion in tech
- Built relationships with other clubs at the University of Waterloo to develop more accessible and inclusive practices

Tech+ Mentorship **May 2020 – Dec 2021**

Mentor

- Mentored students interested in data science by providing resume critiques and interview preparation tips

Waterloo Data Science Club **Jan 2020 – April 2021**

President

- Co-hosted a Clustering for Image Analysis workshop for high school students interested in STEM
- Organized a guided project to help 100+ students build a data science project and network with data scientists

Math Society Mentorship **Sept 2020 – Dec 2020**

Mentor

- Mentored incoming first year mathematics students at the University of Waterloo by providing advice relating to courses, co-op and transitioning to university

Big Brother Big Sisters **Sept 2014 – March 2020**

Mentor

- Mentored students experiencing adversity from the Peel and Waterloo regions to develop their interpersonal skills

Royal Bank of Canada **Jan 2019 – Apr 2019**

Student Partner

- Collaborated with a group of 10 students to improve the RBC co-op experience by delivering biweekly student scoop newsletters showcasing the voices and work of other co-op students

Awards

President's Scholarship of Distinction **Sept 2017**

Awarded to students who have achieved a 95% or higher admission average *Valued at $2000*

President's Research Award **Jan 2022**

Awarded to students who are under the supervision of a University of Waterloo researcher *Valued at $1500*

Talks/Panels

My Pride in Tech (Tech+) **June 2021**

Shared my experiences being a part of the LGBTQ+ community in the tech space

Career Stories of Women in Tech (Laurier Data Science Club) **March 2021**

Shared my journey entering data science and co-op experiences

Lightning Talk (Tech+) **Feb 2021**

Explained the position of a data scientist and data analyst to first- and second-year students

Working Through a Pandemic (UW Statistics Club) **Oct 2020**

Discussed how to navigate the challenges of remote internships

Resume 101 (UW Women in Computer Science) **June 2020**

Provided tips on how to cater your resume for data science positions