

# Kanika Chopra

[kdchopra@uwaterloo.ca](mailto:kdchopra@uwaterloo.ca) | [github.com/kanikadchopra](https://github.com/kanikadchopra) | [linkedin.com/in/kanikadchopra](https://www.linkedin.com/in/kanikadchopra)

## // Skills

**Languages:** Python, R, SQL, MATLAB

**Tools:** pandas, NumPy, Scikit-Learn, SciKit-Image, spaCy, nltk, matplotlib, Plot.ly, Tableau, OpenCV

## // Experience

**Wish – Incoming Data Analyst**

Jan 2021

**University of Waterloo – Undergraduate STAT231 Marker**

Sept 2020 – Present

**Intact DataLab – Data Scientist**

Toronto, Canada

Python, Pandas, NumPy, Scikit-Learn, OpenCV, Matplotlib

May – Aug 2020

- 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; decreased time-to-create by 95%
- Researched and implemented image augmentation techniques using *OpenCV* to improve classification by 3%

**Goldspot Discoveries – Data Scientist**

Toronto, Canada

Python, Pandas, NumPy, Scikit-Learn, Scikit-Image, rasterio, fast.ai, Plot.ly, SQL

Sept – Dec 2019

- Trained an *fast.ai* NLP classifier to conduct a sentiment analysis on 18,000+ gold-related tweets with 85% accuracy
- Visualized the sentiment distribution, U.S. geographic segmentation and top Twitter users between a range of dates
- Performed PCA, K-Means Clustering, SLIC and multi-band ratios on satellite images for mineral composition analysis
- Visualized interactive time-series plots for a stock price technical analysis with 20-day, 50-day, 200-day moving averages and relative indexes using *Plot.ly*

**RBC – Strategic Initiatives Analyst**

Toronto, Canada

Tableau, SQL, Selenium, BeautifulSoup

Jan – Apr 2019

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

## // Projects

**News Category Classifier**

Python, Scikit-Learn, NLTK, Pandas, NumPy

- Trained a news category classifier using Multinomial Naïve Bayes, SVM (Linear, Polynomial, Gaussian), Multinomial Logistic Regression, and Random Forest based on 60,000 HuffPost headlines
- Fine-tuned models using various parameters with GridSearchCV to achieve an accuracy of 88% with Linear SVM

## // Extracurriculars

**UW Data Science Club (President):** Leading a team of students to facilitate events promoting data science; hosted an Introduction to Data Science workshop for first year students

**Data Science for All: Women's Summit (Attendee):** Participated in a five-week program aiding women in data science by providing training and mentorship

**Big Brother Big Sisters (Mentor):** Mentored students from the Peel and Waterloo regions to help develop their interpersonal skills for 5+ years

## // Education

**University of Waterloo**

Waterloo, Canada

B. Math, Double Major in Mathematical Finance and Statistics

Sept 2017 – May 2022

Dean's List [Fall 2019, Winter 2020]