

# Nadia Pasha

npasha6@gatech.edu | <https://github.com/nadiapasha> | <https://www.linkedin.com/in/nadia-pasha/>

---

## **Summary**

Student in the Data Analytics program at Georgia Institute of Technology, an expert in mathematics, proficient in python and its related Machine Learning ecosystem, comfortable applying new research and technologies to mine, analyze and utilize Data

## **Education**

***Georgia Institute of Technology*** (Expected December 2022)

Masters of Science in Analytics

Current GPA: 4

***National University of Science and Technology, Pakistan***

Masters of Philosophy in Mathematics

GPA: 3.9 - Presidential Gold Medalist

***The University of Texas at Austin***

Bachelor of Science in Mathematics

GPA: 3.9 - High Honors

## **Employment History**

***Data Analyst Intern, TVPHealth*** (summer 2021)

Completed an NLP project to analyze unstructured data and help HR department make better decisions in outreach program and improve employee retention. Main libraries used: SpaCy, PyTextRank.

***Math Expert Reviewer, study.com*** (2019-current)

Reviewed Calculus, Linear Algebra, Statistics and Probability problem solutions for accuracy

Instructed writers on how to improve their communication of mathematical ideas

## **Class and Personal Projects**

- *Natural Language Processing*: Used NLP to summarize news articles of interest from landing pages of news outlets. Main libraries used were newspaper3k, SpaCy, NLTK
- *Collaborative Filtering*: Built recommendation system to provide users personalized food recipes. Main libraries used were Pandas, Numpy, SciPy, scikit-surprise
- *Data Visualization*: Created visualization using Choropleth map of Olympic medals by country. Main library used was D3.js
- *Image classification*: Classified face images using PCA and ISOMAP. Main libraries used were Numpy, scikit-learn, Matplotlib

## **Machine Learning Skills**

- *Supervised learning*: Naive Bayes, Regression, Regression trees, Logistic Regression, SVM, KNN, CNNs
- *Unsupervised learning*: clustering via K-MEANS, Spectral Clustering
- Exponential Smoothing and CUSUM
- Outlier detection, data imputation, PCA, Validation, Cross-validation

## **Other Technical Skills**

- *Data Analysis*: SQL, R (dplyr), Python
- *Python's Machine Learning Ecosystem*: Pandas, Numpy, SciPy, PySpark, scikit-learn
- *Natural Language Processing*: SpaCy, NLTK
- *Web Development/Data Visualization*: Matplotlib, HTML, CSS, JavaScript, Vue.js, D3.js, Tableau
- *Simulation*: ARENA, Python
- *Distributed computing*: Databricks, AWS (S3, EC2, EMR), GCP, Azure
- *DevOps*: Git, Docker

## **Volunteer:**

***Local Community Women Writers Association, AMC Austin Chapter***

- Wrote letters and op-ed for local newspapers