

ANUSHA RAVI PATIL

(917)-667-5600 ◇ arp624@nyu.edu ◇ anushapatil5.github.io ◇ [GitHub](#) ◇ [LinkedIn](#) ◇ GHC20 Scholar

OBJECTIVE

Masters level Data Scientist with 1+ years of experience in Python. Seeking to leverage my Data Driven Thinking, Statistical Knowledge and Python skills in Industry.

SKILLS

Languages	Python, SQL, Java
Relevant Tools	Excel, PowerBI, seaborn, t-SNE, Spark, Hadoop, PyTorch, scikit-learn, Git, \LaTeX

EDUCATION

New York University, New York *May 2021 (Expected)*
Masters in Data Science
Relevant Coursework: Natural Language Processing, Machine Learning, Big Data, Linear Algebra
Graduate Teaching Assistant for the graduate course Big Data (Spring 2021), Center for Data Science

Ramaiah Institute of Technology, Bengaluru, India *June 2017*
B.E. in Electronics & Communication Engineering, First-Class Honors
Relevant Coursework: Data Structures, Operating Systems, Linear Algebra

EXPERIENCE

Data Scientist Intern | Cowen, NY *Jan 2021 - Present*

- Duties include conducting analysis and modeling of large complex data to discover patterns
- This role also involves the processing, cleaning and verifying the integrity of incoming transaction-level data

Data Scientist (Capstone) | GeneCentrix, Inc., NY *Sept 2020 - Dec 2020*

- This position is for the Capstone as part of the Fall 2020 semester at New York University, NY
- Automated the identification of drug-target interactions using NLP (**scispaCy**) with an AUC of 0.87
- Derived business insights by detecting novel drug-target interactions across biomedical text sources using **BioBERT**

Graduate Research Assistant | New York University, NY *June 2020 - Aug 2020*

- Worked on audio signal processing (**OpenL3**) to achieve privacy in reconstructed audio from LibriSpeech data
- Developed an inversion model that allows reconstructed audio to be used for speaker classification tasks (**PyTorch**)

Data Science Intern | 42hertz Inc. (Acquired by Cisco), India *Aug 2017 - Feb 2018*

- Shipped a collaborative filtering based [recommendation engine app](#) on the Shopify App Store. (**SQL**, **Git**)
- Used Pareto/NBD to calculate Customer Lifetime Value for an eCommerce music store 'CDNow'. [See blog](#)
- Used **PowerBI** and Shopify's interactive dashboards to obtain data driven insights

PROJECTS

Flight Delays Prediction

- Developed a framework to predict flight delays based on user behavior, weather data and flight history
- Achieved 0.78 AUC on the holdout set using ensemble methods (**scikit-learn**, **SVM**, **Regression**)

COVID-19: War of Twitter Narratives

- Analyzed political discourse (in US) on Twitter regarding the ongoing COVID-19 pandemic using classification.
- Performed data exploration using **LDA** and data visualization using **matplotlib** and **seaborn**

ACHIEVEMENTS

- **First Position** at Understood.org Datathon, hosted by NYU Center for Data Science
- Recipient of the **Indian Academy of Sciences Research Fellowship**, 2016 (194 selected across India)