

ANUSHA RAVI PATIL

+1 (917)-667-5600 ◊ arp624@nyu.edu ◊ [anushapatil5.github.io](https://github.com/arp624) ◊ [GitHub](#) ◊ [LinkedIn](#) ◊ GHC20 Scholar

OBJECTIVE

Masters level Data Scientist with 1+ years of experience, currently interning at **Cowen Inc., NY**. Seeking to leverage my data driven thinking, statistical knowledge and Python expertise in industry.

EDUCATION

New York University, New York, NY May 2021 (Expected)
Masters in Data Science (GPA: 3.6)
Relevant Coursework: Big Data, Advanced Python, Natural Language Processing, Machine Learning, Deep Learning
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 (GPA: 3.9)
Relevant Coursework: Data Structures, Operating Systems, Linear Algebra

SKILLS

Programming Languages: Java, Python, SQL, Git
Database & Distributed Frameworks: Spark, Hadoop, MapReduce, Snowflake
Tools: Git, MS Excel, Power BI, Google Analytics, \LaTeX

EXPERIENCE

Data Scientist Intern | Cowen Inc., New York, NY Jan 2021 - Present

- Maintain and improve data pipelines utilizing healthcare claims data using **Visible Alpha API**
- Develop and enhance medical, pharmacy and EHR claims data pipeline using (**Python, SQL**) to make strategic decisions about handling majority and edge cases in Claims data
- Derive insights using financial history, projections, current earnings, and improve quality of data.

Data Scientist (Capstone) | GeneCentrix Inc., New York, 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, New York, 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), Bengaluru, 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
- Performed feature transformations on input variables and achieved 0.78 AUC on the holdout set using (**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, 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)