

Kunj Chetan Mehta

New Brunswick, NJ | +1 8484371589 | kunj.mehta@rutgers.edu | [kunjmehta.github.io](https://github.com/kunjmehta) | linkedin.com/in/kunjmehta

EDUCATION

Master of Science in Computer Science | Rutgers University, New Jersey **Expected May 2023**
Bachelor of Technology in Computer Engineering | Mumbai University, India | CGPA – 8.52/10 **Oct 2020**
Relevant Coursework – Artificial Intelligence (AI), Machine Learning (ML), Data Warehousing and Mining, Distributed Cloud Computing, Digital Image Processing, Advanced Database Management System, Analysis of Algorithms, Data Structures, Operating System, Blockchain Technology, Mathematical Foundations of Data Science

EXPERIENCE

Teaching Assistant | Rutgers University | New Brunswick, NJ **Sep 2021 - Present**

- Conducted hands-on lab sessions teaching R and reviewed and grading weekly assignments and exams for 78 students across two courses – “Data 101” and “Database Systems for Data Science”

Business Analyst | Quantiphi, Inc. | Mumbai, India **Oct 2020 – Aug 2021**

- Built a solution deck for 7 Public Sector and Education industry clients to show how data and ML can be incorporated into their existing processes
- Designed go-to-market strategies for education and public sector solutions targeting different regions, resulting in 4 campaigns and interest from 6 customers
- Created and managed an internal repository to keep track of research advancements in machine learning; this was leveraged by 230 people in the organization including founders

Freelance Android Developer | IPLit Solutions LLP | Thane, India **Feb 2020 – Mar 2020**

- Developed and deployed an Android application for handsfree token printing for booking appointments in hospitals
- Currently in use in two hospitals across the city

Project Intern | Fractal Analytics | Mumbai, India **Jun 2019 – Jul 2019**

- Built and incorporated a standalone model for product classification of 50 SKUs (80% accuracy) into the collective model training pipeline as part of the consumer behavior analysis project for a Fortune 500 company
- Coded a script for scrapping images of representative products from e-commerce websites using Selenium and annotated 3500 images from the scrapped data to create a dataset for model training

PUBLISHED RESEARCH

- “Simplification with the Transformer - Its Drawbacks”(International Journal of Computer Sciences and Engineering) [\[pdf\]](#)
- “Abalone Age Prediction Problem: A Review” (International Journal of Computer Applications) [\[pdf\]](#)
- “Kernel Regression from Scratch in Python” (Towards Data Science, Medium.com) [\[link\]](#)
- “What Mainstream AI is Not Doing” (Towards Data Science, Medium.com) [\[link\]](#)
- “10 Points to Make it Big in the Data Industry” (Towards Data Science, Medium.com) [\[link\]](#)
- “Introduction to PySpark via AWS EMR and Hands-on EDA” (Towards AI, Medium.com) [\[link\]](#)
- “Fantasy Premier League x Data Analysis: Being Among the Top 2%” (Towards Data Science, Medium.com) [\[link\]](#)

PROJECTS

Image Colorization using Autoencoders | Convolutional Neural Networks (Python, Pytorch, streamlit, matplotlib) **Nov 2021**

- Built a 11-layer deep autoencoder neural network using residual connections that colorizes black and white images
- Trained the network on 10,000 images from FloydHub and deployed online via Streamlit [\[Repo link\]](#)

New York Taxi Fare Prediction | Big Data (Python, pandas, matplotlib, PySpark, AWS EMR, AWS EC2) **Oct 2020**

- Analyzed a 55-million-rows dataset on the cloud to determine varying trends in taxi fares across both location and time
- Augmented the data with features that help analyze trips to and from airports and across different boroughs of NY City
- Predicted taxi fares to an RMSE score of 4.28 by training a Random Forest model on the augmented dataset [\[Repo link\]](#)

FPL Teammaker | Data Analysis (Python, NumPy, pandas, streamlit, matplotlib, PuLP) **Sep 2020**

- Developed and deployed a customizable application that uses pandas and Exploratory Data Analysis to suggest an optimal team to be entered into the Fantasy Premier League fantasy soccer game
- Ranked in the top 2% in worldwide ranking among 8.2 million players by leveraging this application [\[Repo link\]](#)

Final Year Project | Natural Language Processing (Python, Pytorch, Flask, Selenium) **Jun 2019**

- Built a text simplification system that can work on text and simplify it by removing difficult-to-understand words
- Modeled and trained Transformer models that internalized the semantics of and recognized complex words in input
- Improved the performance of the application by preceding the transformer architecture with a Complex Word Identification (90.23% accuracy) model that flagged the complex words beforehand [\[Repo link\]](#)

Abalone Age Prediction | Machine Learning – Regression (Python, pandas, NumPy, matplotlib, scikit-learn) **Sep 2019**

- Determined the ages of abalones (snails) using classification techniques and leveraging their physical characteristics
- Improved the accuracy of determining age using regression techniques and obtained a MAE of 0.936
- Concluded that the dataset is not large enough to get the desired MAE of 0.5 implying correct age prediction [\[Repo link\]](#)

Alien Shooter | Python Game Development (Python)**Feb 2019**

- Expanded the 'Space Invader' game to include three modes of play: Arcade, Timed and Survival

[\[Repo link\]](#)**Reminder – Todo List | Android Development (Android, Java)****Jun 2018**

- Developed an Android application that acts as a combination of a reminder app and a notes app
- Published the app on Google Play Store, and currently has 50+ installs with a rating of 4.6

[\[App link\]](#)**TECHNICAL SKILLS**

- Development:** Python, Java, SQL, PySpark, R, JavaScript, Android, Julia, Adobe Photoshop, Adobe Indesign
- Frameworks and Libraries:** NumPy, pandas, matplotlib, scikit-learn, MLib, Looker, streamlit, Pytorch, seaborn, Flask
- Cloud Services:** AWS Sagemaker, AWS EMR, AWS Lambda, Google Big Query
- Machine Learning Theory:** Regression, Decision Trees, Clustering, Ensemble Models, EDA, CNNs, NLP, Time Series

CERTIFICATIONS

- | | |
|---|-----------------|
| • GANs Specialization Coursera | Ongoing |
| • Deep Learning Specialization Coursera | Ongoing |
| • LookML Developer Looker [link] | Mar 2021 |
| • Cloud Practitioner AWS [link] | Jul 2021 |
| • Machine Learning Engineer Nanodegree Udacity [link] | Oct 2020 |
| • Applied Data Science with Python Specialization Coursera [link] | Jul 2020 |
| • Machine Learning Coursera [link] | Apr 2019 |
| • Certified Mutual Fund Distributor in India, via the NISM Series V-A Certification Exam [link] | Oct 2020 |
| • Certified to teach mathematics, block coding and, Python to K-12 students on Cuemath, an e-learning startup | Apr 2020 |

ADDITIONAL ACTIVITIES

- Rehabilitated 10 runaway children as a volunteer for an NGO in Thane, India
- Mentored a team of 12 photographers through covering three college festivals
- Written 4 articles for Kshitij, the KJ Somaiya College of Engineering college magazine
- Visited Vechtdal College, Netherlands as a cultural exchange student