Kunj Chetan Mehta

New Brunswick, NJ | +1 8484371589 | kunj.mehta@rutgers.edu | kunjmehta.github.io | linkedin.com/in/kunjmehta

EDUCATION

Master of Science in Computer Science | *Rutgers University, New Jersey* | **GPA** – 3.91/4 **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

TECHNICAL SKILLS

- **Development:** Python, Java, SQL, PySpark, R, C++, JavaScript, Android, MS Excel
- **Frameworks and Libraries:** NumPy, pandas, matplotlib, seaborn, scikit-learn, MLlib, streamlit, Pytorch, NLTK, Pillow, OpenCV, Flask
- **Cloud Services:** AWS Sagemaker, AWS EMR, AWS Lambda, Google Big Query
- **Databases and Visualization:** MySQL, AWS Redshift, Amazon RDS, Tableau, Power BI, Looker
- **Machine Learning:** Regression, Decision Tree, Clustering, EDA, CNN, NLP, Time Series, Ensemble & Multimodal Learning
- **Certifications:** Cloud Practitioner (*AWS*), LookML Developer (*Looker*), Machine Learning Engineer (*Udacity*), Deep Learning Specialization (*Coursera*), Applied Data Science with Python Specialization (*Coursera*) [link]

EXPERIENCE

Teaching Assistant | Rutgers University | *New Brunswick, NJ* **Sep 2021 – May 2022**

- Taught R, SQL and Amazon Redshift and graded 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**

- Researched and presented highlights of the US stimulus bills to internal stakeholders that informed Quantiphi's Public Sector business strategy
- Performed *market research on 200 organizations in the US Education industry* and came up with an effective go-to market strategy that *converted four cold leads*
- Presented solution decks showcasing how machine learning can be incorporated into their existing processes to four leads, *converting two of them*
- Analyzed and reported quarterly revenue figures to internal stakeholders using Looker dashboards
- Initiated and led the creation of 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 a model for classifying 50 products with a *80% accuracy* that was delivered 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]
- "How Gradient Descent Works?" (Towards Data Science, Medium.com) [link]

PROJECTS

Food AI| Multi-modal Representation Learning *(Python, PyTorch, Hugging Face, seaborn)* **May 2022**

- Beat the baseline retrieval performance in the original im2recipe paper by 80% for the Recipe1M cross-modal food recipe retrieval task by improving on the feature extraction pipeline
- Improved retrieval performance further by learning shared multi-modal representations using CCA and non-linear neural networks trained using Triplet Loss
- Enhanced the explainability of the system by incorporating Vision Transformers and cross-modal attention when learning shared representations [Repo link]

Movie Recommendation from Conversational Data | NLP *(Python, PyTorch, surprise, seaborn)* **May 2022**

- Built a movie recommendation system leveraging conversational user data, external critic data and domain adaptation techniques, which is a re-implementation of this paper
- Obtained a 3% improvement by performing hyperparam tuning on all three CF approaches: KNN, SVD and SVDpp
- Experimented with neural CF approaches employing Neural Matrix Factorization as an extension of the paper and obtained comparable results of RMSE=1.232 and MAE=0.9569 [Repo link]

Logo Detection and Classification | Computer Vision *(Python, PyTorch, seaborn)* **May 2022**

- Performed open set logo detection on "Logos in the Wild" dataset using YOLOv5 and obtained 12% improvement on the original results here
- Obtained a classification accuracy of 22.56% against 47 classes of the Flickr-47 dataset using a logo classification architecture consisting of YOLOv5 and template matching focused on both abstract and textual logos [Repo link]

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**

- Performed feature engineering to focus on trips to and from airports and across different boroughs of NYC
- Predicted taxi fares to a *RMSE score of 4.28* by training a Random Forest on the augmented data [Repo link]

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

- Developed and deployed an application that performs exploratory data analysis on the English soccer Fantasy Premier League (FPL) game data to suggest an optimal team to be entered into the game
- Ranked in the top 2% in worldwide ranking among 8.2 million players in the year 2020 [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]

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