PARTH GUPTA

(347) 712-9803 |linkedin.com/in/guptaparth25/|pg2677@columbia.edu

Education

Columbia University New York, USA
Master of Science in Data Science, **GPA: 3.92/4.00** Sep 2021 - Dec 2022

Indian Institute of Technology Roorkee Roorkee, India
Integrated Master of Science in Applied Mathematics, CGPA: **9.51/10.00** July 2015 - July 2020
Honors: Department Gold Medal (Rank-1), **Best Thesis Award** and **INSPIRE** scholarship

Skills/Technologies

Computer languages	Python, C++, MATLAB, SQL, R
Software Packages	Scikit-learn, NumPy, Pandas, Pytorch, TensorFlow, Keras, OpenCV, XGBoost
ML Techniques	Classification, Regression, Clustering, Time Series Analysis, Recommendation Systems, Statistical Modeling, Bagging, Boosting
Data Visualization	ggplot2, Seaborn, Matplotlib, Tableau, Power BI

Experience

Columbia University New York, USA
Graduate Teaching Assistant Sep 2021 - May 2022

- Appointed as Teaching Assistant for Time Series, Panel Data Forecasting course in fall 2021 and Advanced Analytic Techniques in spring 2022, responsible for guiding students, teaching concepts, and evaluating assignments.

Adobe Noida, India
Software Development Engineer-1 Aug 2020 - Aug 2021

- Worked on Image Super Resolution feature to upsample RGB, CMYK and Grayscale images in PDFs; collaborated with 5 research scientists across 3 global teams.
- Implemented CNN in TensorFlow to upsample images. Significantly outperformed Bicubic upsampling in terms of PSNR by 4.278 and SSIM by 7.9% for scale factor 4 on Set5 dataset.
- Deployed and integrated model in Value Added Service (VAS) framework using OpenCV and C++.
- Fixed multiple bugs for Adobe PDF Print Engine (APPE) 5.6, 5.7 and 6.0 and solved several Xchange cases of OEMs.

Adobe Research Bangalore, India
Research Intern May 2019 - July 2019

- Proposed a novel memory-based architecture, Continuous Dynamic Key-Value Memory Network (CKVMN) to model similarity across time series of different products; collaborated with 2 interns and 2 research scientists.
- Forecasted metrics and achieved 31% improvement in mean MAPE for cold start products over LSTM network.
- Scrapped Best-Buy website using Requests and Beautiful Soup to collect SKU level meta-data for various products.
- Worked with two interns to create a live website showcasing project employing HTML, CSS, and JavaScript.
- Published results in WWW 2020 (The Web Conference) and filed a patent in the U.S. Patent Office.

Purdue University West Lafayette, USA
Visiting Undergraduate Student May 2018 - July 2018

- Designed a second order optimization method for training deep network utilizing sub-sampled gradient and Hessian.
- Developed a sub-sampling technique to calculate sub-sampled Hessian matrix based on leverage scores.

Data Science Projects

Ranking Toxic Comments by Severity Dec 2021 - Jan 2022

- Programmed Ensemble Ridge Regression model to predict toxicity score and attained ranking score of 0.866.
- Converted comments into vectors by utilizing combination of TF-IDF and Fast Text embeddings.

Amazon Recommender System for Video Games Oct 2021 - Dec 2021

- Led 5-member team to devise Content-Based & Collaborative Filtering approaches to recommend items.
- Evaluated effect of sentiment scores of reviews, metadata and summary of reviews on item recommendations.
- Achieved 16.4% and 17% overlap between recommended items and also view and also buy items respectively.