

Rahul Sharma

(248)-525-5789 • rahulsharma0628@gmail.com • [rahulsharma0628.github.io](https://github.com/rahulsharma0628) • Malvern, PA

SUMMARY

3+ years as a skilled Programmer/Analyst with a demonstrated track record of achievements in examining large datasets, developing predictive analytics pipelines, and leveraging deep learning frameworks for data modeling to help businesses in critical data-driven decision making. Recognized for effective communication with clients.

EDUCATION

The Pennsylvania State University, USA

Aug 2019 – Dec 2020

M.S., Data Analytics

GPA: 4.0/4.0

Key Coursework: Applied Statistics, Predictive Analytics, Deep Learning, Databases, Data-driven Decisions

Research Assistant: Fairness in Machine Learning and Causal Inference (*Advisor: Prof. Youakim Badr*)

Indian Institute of Technology Bombay, India

Jul 2013 – Apr 2017

B.Tech, Chemical Engineering

Leadership: General Secretary (Chemical Engineering Association), Manager (Innovation Cell)

PROFESSIONAL EXPERIENCE

ML Engineer – PhenoMx Inc., New York (Capstone Project)

Aug 2020 – Dec 2020

- Led a team of four students to develop an MRI-based biomarker for early detection of Alzheimer's Disease (AD)
- Preprocessed brain MRI images and built deep learning models using correlation measure between 40 brain regions
- Engineered a production-ready predictive analytics pipeline to help medical practitioners classify A.D. patients

Analyst – Citigroup, India

Jul 2017 – Jul 2019

- Enhanced click-through rate by 35% measured using A/B testing by restructuring the alert dispatching framework
- Reduced false-positive alerts by 18% by analyzing traders' activities using ensemble machine learning methods
- Decreased legal entity onboarding time by 70% by developing an end-to-end application using RESTful APIs

RELEVANT PROJECTS

Video-to-Text Summarizer

Feb 2020 – Jul 2020

- Built a supervised, keyshots based video captioning and summarization model to flag inappropriate video content
- Attained 0.52 F1-score on the generated textual description by developing a *seq2seq* model with Attention

Time-to-Event Prediction for Alzheimer's Disease (A.D.) Progression

Feb 2020 - Apr 2020

- Leveraged 90+ clinical parameters to build deep learning-based survival models to predict A.D. stage-shift duration
- Achieved concordance-index of 79%; to be utilized by the medical experts to recommend personalized treatments

Model-based Book Recommendation System

Feb 2020 - Apr 2020

- Designed recommender systems based on Collaborative, Content-based Filtering and Matrix Factorization
- Applied Gradient Descent to minimize the RMSE by 10% as compared to KNN Means and Cosine Similarity

Sentiment Analysis on Amazon Cellphone Reviews

Oct 2019 - Dec 2019

- Accomplished 82% sentiment accuracy on cellphone reviews by building a Multinomial Naïve Bayes model
- Amplified sentiment model accuracy to 95% by engineering and implementing a negation identification algorithm

TECHNICAL SKILLS

Programming: Python, R, Java, SQL, HTML, CSS

Data Science: Machine Learning, Deep Learning, NLP, Data Mining, Statistical Modeling, Time Series Forecasting, A/B Testing, ETL, NumPy, Pandas, Scikit-Learn, spaCy, Keras, PyTorch, Minitab

Development: Git, Hadoop, Spark, Docker, NoSQL, MySQL, Hive, Test-driven development, Microservices

Others: Design of Experiments, Lean Six Sigma, Agile/Scrum Methodologies, JIRA, Project Planning

HONORS AND AWARDS

- **Winner - Wawa-HCL Hackathon;** Developed Sales Forecasting Models for Wawa using LSTM and Prophet
- **Recipient of Warren V. Musser Fellowship (2020) and Chancellor's Scholarship (2019)** at Penn State
- **Best Student Pitch - Lion Cage;** Annual competition for early-stage entrepreneurs to pitch products
- Secured All India Rank **1077 (top 1%)** in IIT-JEE Advanced 2013 among 1.5 million candidates

PUBLICATIONS

- **Sharma, R.,** Anand, H., Badr, Y., Qiu, R., "Time-to-Event Prediction using Survival Analysis Methods for Alzheimer's Disease Progression", *International Journal of Medical Informatics* (under-review)
- Badr, Y., **Sharma R.,** "Data Transparency and Fairness Analysis of the NYPD Stop-and-Frisk Program", *Journal of Data and Information Quality* (under-review)