Janish Rajesh Parikh

New Brunswick, NJ | +1 (609)-721-8313 | janishparikh5@gmail.com | GitHub | LinkedIn | RPubs

EDUCATION

Master of Science in Computer Science ***Sep 2021 – May 2023***
Rutgers University-New Brunswick | CGPA: 4.0 *New Brunswick, NJ*
Specialization: Machine Learning and Artificial Intelligence

Bachelor of Technology in Computer Science & Engineering ***Aug 2017- May 2021***
Indian Institute of Information Technology, Vadodara | CPI: 8.78 *Gandhinagar, India*

Relevant Coursework: Machine Learning, Natural Language Processing, Massive Data Mining, Introduction to Artificial Intelligence, Data Analytics, Data Structures & Algorithms, Software Engineering, Database Systems

EXPERIENCE

Data Science Intern | OnPoint - *Koch Industries* ***May 2022–Present***

- Researching and developing SOTA Machine Learning Models using waveform data to predict electrical network equipment failure to be able to alert an electrical grid operator about an incoming failure
- Conducted research to test multiple hypothesis and effectively communicated the findings to key stakeholders
- For predicting vegetation failures, improved the Balance Accuracy Score by 11% and the Precision Score for the critical class by 4% by performing feature engineering, outlier treatment, complex model evaluation and hyperparameter tuning
- Collaborated with multivariate teams to insert trained models and gauge performance improvement
- Streamlined and optimized data analysis/visualization and data preparation pipelines using Dask, Vaex, Plotly and SHAP

Undergraduate Teaching Assistant | *Rutgers University* ***Sep 2021 – May 2022***

- I taught students about Data Analytics, Statistical Modeling, and Machine Learning Algorithms using Python/R and guided them through their Capstone Project in Data Science
- Conducted regular lectures, interactive sessions for clarifying doubts and graded exams and assignments

Data Science Intern | *COSGrid Networks* ***Feb 2021–Jun 2021***

- Ideated, designed, implemented, and pipelined an end-to-end product for cyber-situational awareness using big data technologies Spark, Kafka, Druid, Elasticsearch and AWS
- Developed Machine Learning applications for IoT Device Classification, and Real-Time Attack Detection
- Ingested real-time NetFlow data through Spark Structured Streaming, processed more than 10,000 records per minute during peak hours and alerted the users of any malicious activity under a minute
- Project was selected amongst the top 6 finalists in 'Cyber Security Grand Challenge!' organized by DSCI

PROJECTS

Image to Image Translation *[Link]*

- Explored Supervised I2I using Pix2Pix GAN to translate Street View Images to Aerial View Images and vice-versa
- Implemented CycleGAN framework for the task of translating Real Pizza to Synthetic and Live Pizza Image Domains
- Researched the drawbacks of CycleGAN framework and proposed an enhanced CycleGAN by incorporating VGGPerceptual Loss in CycleConsistency that attained a 10% improvement by reducing unrealistic artifacts

Maze Solver *[Link]*

- Designed multiple AI agents using Repeated A*, Inference, and Bayesian Networks for optimally searching a hidden target within a maze using Python and NumPy
- Optimized these agents to find the targets in (101) *(101) dimension mazes under 20 milliseconds
- Built a CNN with Dense layers using PyTorch to imitate these agents obtaining accuracy of 92% in solving the mazes.

Conversational Movie Recommendation System *[Link]*

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

TECHNICAL SKILLS

- **Languages:** Python, SQL, R
- **Tools & Framework:** PyTorch, scikit-learn, NumPy, Pandas**,** Seaborn, PySpark, Kafka, AWS Sagemaker
- **Developer Tools**: Git, Docker, AWS, Elasticsearch, PostgreSQL, Parquet, Linux, Databricks, AWS Lambda
- **Theoretical Knowledge:** A/B Testing, Ensemble Learning, Gradient Boosted Trees, Big Data, GANs, Time Series Analysis, MLOps