Puneet Mathur

+1 (240)-615-7067

✉ puneetm@cs.umd.edu github.com/pmathur5k10
Google Scholar: www.tinyurl.com/puneet-mathur www.linkedin.com/in/themadaiguy

EDUCATION

University of Maryland College Park, *Maryland, MD USA*
Master of Computer Science *Expected:* ***May 2021***
Key Courses: Machine Learning, Advanced Computer Vision

Netaji Subhas Institute of Technology (NSIT), *New Delhi, India*
Bachelors of Engineering in Computer Engineering *Aug. 2014 – May. 2018*
Key Courses: Data Structures, Algorithms, Databases, Object Oriented Programming,

SKILLS

Programming Languages: Python, C++, C, SQL, Flask,
Machine Learning: Computer Vision, Medical Imaging, Natural Language Processing, Neural Networks, Deep Learning
Data Science Frameworks : Numpy, Pandas, Keras, PyTorch, Tensorflow, Scikit-Learn, Matplotlib, ETL Pipelines
Teaching Assistant: Algorithms, Evolutionary Computing

WORK EXPERIENCE

Estee, *Gurgaon, India*
Software Developer, Quantitative Strategy *Jun. 2018 – Jul. 2019*

- **Algorithmic Trading**: Automated quantitative backtesting, trading execution and equity-sector optimization for multi-factor trading logic in Python and SQL using Bloomberg for market data. Effectively managed to generate 6.7% profit for assets worth INR 20 million.
- **Data Analytics**: Analyzed long-term market trends using Matplotlib and Scikit-Learn for forward testing of trading hypothesis. Extracted statistical signals to forecast probable loses in 78% of divergent cases
- **Mutual Fund Trading App**: Designed and developed backend REST API's on AWS Lambda for mutual fund recommendation. Used k-Means clustering and time series analysis to design an optimal basket selection algorithm. Reduced the time utilization in data processing steps by 67% by parallelizing time intensive computations.

MIDAS Labs, IIIT, *Delhi, India*
Machine Learning Research Intern *Jan. 2018 – Mar. 2019*

- Designed a novel implementation of multi-headed self-attention networks for detecting disease in renal tissues, achieving 87% accuracy, at par with experienced nephrologists.

MakeMyTrip, *Gurgaon, India*
Software Development Intern *Apr. 2017 – Jul. 2017*

- Developed discount voucher and hotel catalog modules for Android mobile apps using the React Native framework. The app was downloaded 10K times in 2 months post-launch.

DATA SCIENCE PROJECTS AND PUBLICATIONS

Detecting Offensive Tweets in Hindi-English Code-Switched Language *AAAI 2020*

- Explored cross-lingual transfer learning to detect hate speech on multi-lingual social media posts.
- Developed a hybrid CNN-LSTM model to outperform supervised machine learning models like SVM, Decision Trees, Naive Bayes trained on TF-IDF, Bag of Words and N-gram features by 17%.
- Employed deep graph embeddings, community profiling and a semi-supervised expert-in-the-loop debiasing algorithm for hate speech detection modeling.

Exploring Classification of Histological Disease Biomarkers from Renal Biopsy Images *WACV 2019*

- Studied computer vision techniques to identify diseased kidney tissues using transfer learning, supervised feature extraction and self-attention based CNN architectures.
- Developed a novel Multi-Gaze Attention Network (MGANet) to effectively establish a state of the art model that gives an accuracy of 87.25% and 81.47% for glomeruli and fibrosis classification.

Author Profiling and Social Network Graphs for Suicide Ideation Detection *NAACL SRW 2019*

- Designed a deep learning based feature stacking approach to extract textual features, historical author profiling and graph embeddings to detect suicidal intent in tweets with a highly desirable recall of 0.96.

Detection of Social Media Disclosures of Sexual Harassment *NAACL SRW 2019*

- Experimented with several deep learning based models such as CNN, RNN, GRU, Bi-LSTM for detecting social media disclosures of sexual harassment.
- Used ULMFiT based Disclosure Language Model to show that augmenting the training data with additional domain-specific data achieves superior tweet classification.

Identification of Emergency Blood Donation Request on Twitter *EMNLP 2018, SMM4H*

- Classified tweets referring to the necessity of urgent blood donation requirement using linguistic feature-based SVM classification technique with an accuracy of 97.89