2595 Washington Street
Boston MA – 02119
Phone No: +1-857-260-9294

SAKSHI SUMAN

sakshi.math.163@gmail.com
GitHub | LinkedIn | Portfolio

A computer scientist, mathematician and a Software Engineer solving complex challenging problems for 2+ years. Looking for a long-term career in research-oriented roles in the fields of Computer Vision with applications of Machine Learning/Deep Learning.

EDUCATION

- M.S. in Applied Mathematics (Machine Learning and Data Science), Northeastern University, Sep 2021 – Present; GPA: **3.95/4.0**
- B.Tech. in Computer Science Engineering, REVA University, Jul 2015 – Jun 2019. GPA: 8.55/10.
- **Coursework:** Data Structures & Algorithms, Calculus, Applied Linear Algebra, Probability, Mathematical Modeling, Machine Learning, Applied Statistics, Graph Theory, Databases, Operating Systems, Computer Architecture, Discrete Mathematics.

TECHNICAL SKILLS

- Python, Java, R, C/C++, MATLAB, Mathematica, SQL, PHP, Perl, HTML, CSS, TypeScript, XML, JSON, Visual Basic.
- PyTorch, TensorFlow, OpenCV, NumPy, pandas, Matplotlib, scikit-learn, SymPy, Spark, Git, Jupyter, Linux, Docker, PyCharm.
- Regression, Classification, Ranking, Recommendation Systems, Clustering, Dimensionality Reduction, Bagging, Boosting, Feature Engineering, Neural Networks, Deep Learning, Computer Vision, Natural Language Processing, Optical Character Recognition.

EMPLOYMENT

AI/ML Engineer Intern — **LinkedIn Corporation** — **May 2022 – Aug 2022**

- Worked with the Anti-Abuse AI team to develop an end-to-end Machine Learning model to detect fraud jobs using content-based features. Improved the Precision from 28% to 49% at 80% Recall compared to an existing behavioral model.
- Reduced manual human intervention of downstream team by 43 %.
- Worked in technical teams in development, deployment with product managers to formulate ML problems.

Software Engineer – Machine Learning — **Pelatro Solutions Pvt. Ltd.** — **Jun 2019 – Jun 2021**

- Predicted Next Best Action for an Offer Generator using K-Means Clustering with a 61 % average chance of achieving the intents.
- Optimized the duplicate row detection algorithm using a probabilistic approach and reduced time complexity from $O(n^2)$ to $O(n)$.
- Adapted Tesseract OCR's code, to increase accuracy in text-recognition for screen fonts from 50 % to 95 %.

Machine Learning Intern — **Walkter Beacon Lab Pvt. Ltd.** — **Jan 2019 – May 2019**

- Automated resume matching process using a word count model and decreased the time spent by recruiting by ~ 80 %.
- Performed sentiment analysis on user ratings for organizations and developed a smart scoring algorithm for work happiness.
- Designed an efficient data structure for user visit logging and calculation of user retention rate. Automated email system for ATS.

Teaching Assistant/Mentor — **Northeastern University** — **Sep 2021 – Present**

- Courses: Calculus I (including applications of mean value theorem, integration, finding area under curve), Probability & Statistics.
- Mentor at **Girls' Angle**, a math club supported by Google that provides comprehensive approach to math education for girls.

PROJECTS

- Brain CT Hemorrhage Classification & Segmentation Performed binary classification using Xception Net and transfer learning to classify brain CT scan slices achieved an F-Score of 0.76. Used class weighting to account for imbalance and improve F-score to 0.82. Applied Bayesian Hyperparameter Optimization to reduce training time by 70 %. Performed semantic segmentation using U-Net and achieved an IoU of 0.66. Leveraged multiple shades of CT scans and 3D convolutions to improve IoU to 0.71.
- Movie rating prediction using Matrix Factorization Derived update rules for Weighted Alternating Least Squares and predicted missing user ratings for MovieLens data to achieve a 62 % better MSE performance than baseline model.
- Data Modeling using Markov Chain Performed Time Series Analysis of monthly Sunspots from 1749 – 1983 with a Markov Chain. Performed autocorrelation and GoF test at 5 % significance level to determine valid states of chain.
- Image classifier for the SVHN dataset Built a CNN classifier model with 3 convolutional layers and 2 fully connected layers for digit recognition on street view images. Applied MaxPooling, BatchNormalization, Dropout and Early Stopping callback techniques to increase the validation accuracy on baseline from 55 % to 89.55 %.
- Transfer learning for pet classification Used pretrained MobileNet V2 model (trained on ImageNet dataset) as a feature extractor and trained additional new layers to classify cats and dogs. Applied freezing on pretrained layers and replaced last layer to achieve a classification accuracy of 99 %.

EXTRA ACADEMIC ACTIVITIES

- Featured as a stellar grad student in the Applied Math program (out of 229 students) for academic and career excellence.
- Attendee of virtual Grace Hopper Conference (vGHC) – 2021, WomenHack – Boston and networked with several companies.
- Member of IEEE computer society, student branch and volunteered at the IEEE International Smart Technologies, Bangalore, 2017.