DEEPIKA VADLAMUDI

deepika.vadlamudi96@gmail.com | LinkedIn | (703)817-6839 | Github | Medium

SUMMARY

- Computer science professional with 4+ years of experience in Software Development, ML and Data Engineering
- Proficient in Python, utilizing it extensively for data analysis, machine learning, and statistical modeling
- Strong knowledge of command-line tools and utilities for effective server debugging
- Proficient in data processing technologies and tools, such as Apache Spark, Hadoop, and SQL, to ensure efficient and scalable data processing.
- Effective at collaborating and communicating with colleagues from technical and non-technical domains

EDUCATION

Stevens Institute of Technology, Hoboken, NJ
Master of Science in Computer Science. GPA: 3.83 Jan 2020-Dec 2021

Jawaharlal Nehru Technological University College of Engineering, Sultanpur
Bachelor of Technology, Computer Science. GPA: 3.70 Sep 2014 - Jun 2018

SKILLS

Programming Languages: Python, C++, MATLAB, SQL, C, Linux, CUDA, JavaScript
Libraries/Frameworks: PyTorch, Tensorflow**,** Hive, Spark, Pandas, NumPy, Scikit learn, BeautifulSoup
Technical Expertise: Algorithms, Machine Learning, Deep Learning, NLP, Data Mining, Adversarial Robustness, Recommender Systems, Analytics, Statistics, Optimization techniques, AWS, Azure

WORK EXPERIENCE

Field Data Science Engineer – ***Domino Data Lab***, New York City, USA Jan 2022– Dec 2022

- Conducted a comparative analysis of running a dense neural network for image recognition in a sequential and distributed manner using RayTune on Domino Platform, observed 70% speedup
- Implemented several machine learning algorithms, statistical techniques, boosting techniques and deep learning algorithms - and delivered proof of concept demos to existing and prospective users
- Identified user requirements that are not supported by product, implemented them as adhoc field solutions using python and linux script scripting
- Led effective communication and collaboration between cross-functional teams to ensure a shared understanding of project goals and customer requirements using the project management tool called Monday
- Effectively transformed raw data into meaningful visual representations using BigQuery enabling the customer success team to track and analyze usage patterns, identify growth opportunities, and optimize customer experiences
- Led training sessions about product, on-demand Spark, and advocated best practices for ML workflows for multiple users online and on-site

Environment: Python, SQL, Docker, Kubernetes, AWS, Linux, MLOps, Ray, TensorFlow, Keras, PySpark

Software Engineer, Risk Analysis – ***Amazon***, Hyderabad, India Oct 2018 – Dec 2019

- Developed an advanced model utilizing the FP growth algorithm for detecting fraudulent patterns, surpassing the effectiveness of previously employed apriori algorithm
- Implemented data preprocessing techniques, including data cleaning, transformation, and feature engineering, to optimize the dataset for association rule mining.
- Efficiently managed a large volume of over two thousand generated rules by employing advanced search techniques, ensuring the identification of the most relevant and actionable patterns on a daily basis
- Utilized ETL batch pipeline frameworks like Airflow to automate data ingestion, transformation, and integration processes, reducing manual effort and increasing productivity
- Developed data monitoring and alerting systems to detect anomalies and potential risks in real-time, enabling proactive risk mitigation measures using tools like Prometheus and Grafana.

Environment: Python, Java, Airflow, Nifi, Kafka, Spark, Hadoop, Hive

Machine Learning Engineer – *Ridhan Technologies*, Hyderabad, India May 2018 – Oct 2018

- Designed, implemented neural nets for anomaly detection in ECG signals
- Leveraged AWS technologies like S3, lambda, and Amazon Sagemaker

Environment: Python, SQL, JavaScript, CNNs, TensorFlow, PyTorch, AWS S3, Lambda, Amazon Sagemaker

Research Intern - ***Advanced Numerical Research and Analysis Group*** Hyderabad, India Sep 2017 - May 2018

- Examined performance of PCA over different data sets
- Created code using MPI to be computationally efficient by running it over multiple processors
- Drafted a detailed report about project using latex which could later be utilized for projects needing PCA as a module

Environment: Linux, MPI, C, Latex

Project Trainee - ***Institute of Development and Research in Banking Technology***, India May 2017 – Jul 2017

- Studied various machine learning and data mining algorithms on datasets over 100k instances
- Collaborated in research on particle swarm optimization
- Implemented PCA in MATLAB, experimented for efficiency using various mathematical techniques

Environment: MATLAB, Latex

ACADEMIC PROJECTS

Robust Learning of Halfspaces with Agnostic Noise

- Developed efficient algorithms to robustly learn halfspaces in presence of agnostic noise improving performance from exponential to polynomial time complexity
- Leveraged basic mathematical techniques such as SVD and gradient descent to achieve better computational complexity

Parallel Implementation of CNN for Image Classification using CUDA

- Conducted classification of given image on GPU, leveraged transfer learning, deployed prefetching and streams
- Gained 38x speed than sequential version; program is memory efficient, accomplished 100% occupancy

Plagiarism Detector

- Examined text files and performed binary classification. Labeled it plagiarized or not by comparing it to source text file
- Selected features for comparing similarity between two text files by analyzing correlations between different attributes
- Deployed plagiarism-classification model using Amazon SageMaker

Classification of Pima Indians Diabetes Dataset using AdaBoost with MLE as Base Classifier

- Implemented using MATLAB and Python; leveraged boosting technique to improve accuracy
- Analyzed results by varying different parameters; identified possible improvements

Parallel Implementation of PCA using MPI

- Modified PCA by leveraging different mathematical techniques to achieve higher performance using multiple processors
- Attained 9X higher performance than sequential PCA. Recognized as best project of 2018, JNTU-CS department

PUBLICATIONS

- Robust learning of halfspaces in the presence of agnostic noise – A study, implementation, design and analysis of algorithms - ProQuest Dissertations Publishing, 2021. 28867500 Dec 2021
- V. Ch. Venkaiah and V. S. P. Deepika: A novel self-healing key distribution scheme based on vector space access structure and MDS codes, International Journal of Communication Systems Sep 2019
- Co-authored “Mathematical Essentials”, a chapter in the book “Handbook of statistics Vol38 - Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications” edited by C.R. Rao and Venkat Gudivada, and published by Elsevier Sep 2018

GRADUATE WORK EXPERIENCE

Graduate Teaching Assistant – ***Stevens Institute of Technology*** Jan 2021 – Jan 2022

- Served as teaching assistant to courses: deep learning, probability and stochastic processes, mathematical foundations of machine learning
- Held office hours to clarify concepts, help students with assignments, project ideas and design
- Designed and graded assignments in python and MATLAB, drafted questions for term examinations