HARIKA KANAKAM

hkanakam@iu.edu +1(812)778-4329 harikakanakam harikakanakam

PROFILE SUMMARY

Skilled data scientist with over 3 years of proven expertise in developing predictive models, machine learning solutions, and automation scripts to drive operational efficiency and achieve strategic business objectives. Currently pursuing Master's degree to deepen theoretical knowledge and stay at the forefront of industry advancements in Data Science.

EDUCATION

Masters in Data Science | *Indiana University*, GPA: 3.83/4 | Bloomington, IN, USA **Aug 2022 - May 2024**
Relevant Courses: Deep Learning, Machine Learning (Math oriented), Statistics, Applied Algorithms (Python), Cloud Computing.

Bachelors in Computer Science | *Madanapalle Institute of Technology and Science*| Madanapalle, India **Jul 2015 – Apr 2019**

TECHNICAL EXPERTISE

Python | MySQL | PostgreSQL | TensorFlow | PyTorch | Keras | scikit-learn | Pandas | NumPy | OpenCV | matplotlib | ggplot | Plotly

Cloud - Amazon Web Services(AWS), Azure Databricks | Visualization - Excel, Tableau, PowerBI | Machine Learning(ML) - Regression, Clustering, Decision Trees, Random Forest, XGBoost, SVM, Neural Networks, NLP, CNN, RNN, Word Embeddings.

WORK EXPERIENCE

Knowledge Lens: A Rockwell Automation Company **Bangalore, India**

Senior Big Data Engineer **Apr 2021 – Jul 2022**

- Designed a scalable, optimized and efficient architecture using SQL and communicated data insights through visualizations.
- Performed quantitative research to drive continuous improvement of product, resulting in a 10% reduction in production costs.
- Collaborated with cross-functional teams to conduct workload assessments and devised unique strategies to optimize operations.
- Monitored user sessions using AWS Cognito, CloudWatch Logs, Redis cache datastore delivering 40% spike in productivity.
- Revamped REST APIs by mentoring the team through code reviews resulting in 20% faster delivery and 35% bug reduction.
- Deployed network restricted applications using AWS Kubernetes pods to enhance network security and technical solutions.

Junior Data Scientist **May 2019 - Mar 2021**

- Collaborated with DevOps team to orchestrate Linux services and GIT CI/CD deployments, reducing execution time by 10%.
- Performed hands-on data analysis and modeling with large data sets, including discovering data from sources(local, s3, azure), getting data access, cleaning up data, and making them "model-ready".
- Automated the process of job monitoring and executed data migrations seamlessly, transferring data from MySQL to AWS S3.
- Utilized the capability of ML and data mining for predictive analysis, to forecast deployment status with 95% accuracy.
- Designed migration flow for microservices platform hosted on AWS using FastAPI, conducted analytics on stakeholder feedback.

Teck stack: PyCharm, AWS(EKS, sagemaker, s3, Athena, RDS, EMR, Secrets Manger, Cloud watch, Cognito, Load Balancer), Tensorflow, Keras, MySQL, Docker, Kubernetes, Fast API, Flask, Swagger, Redis, GIT, Jenkins, Bit bucket, ngrok, SSO, JWT, PostgreSQL, Mongo DB, Spark, Hive.

ACADEMIC EXPERIENCE

Indiana University, *Research Assistant* **Summer 2023**

- Conducted data extraction and transformation by parsing XML and SAS files to convert them into datasets to train ML models.
- Engineered a specialized algorithm that analyzes publicly available financial reports, such as annual filings (10-K Forms).
- Evaluated and analyzed time horizons of companies, helping investors with valuable insights to identify potential risk factors.

ACADEMIC PROJECTS

The Vision of Machine Unlearning **Fall 2023**

- Implemented machine unlearning algorithms to remove the effect of unwanted samples from the trained ML model.
- Implemented teacher-student (knowledge distillation) and zero-shot architecture for class-based unlearning by defining a custom loss function as an extension for the existing KL-Divergence loss, thereby outperforming the base implementation by 7%.
- Identified similar gradients by Gradient Matching and re-initialized to introduce model discontinuity, improving 15% efficiency.

User Interactive Image Segmentation **Fall 2023**

- Developed a user interactive platform to extract the objects with/without the background from the user uploaded images.
- Utilized Mask R-CNN for image classification and segmentation, achieving a 95% extraction rate for the classes known by model.
- Leveraged the application by using FastAPI and ReactJS frameworks for backend and frontend and hosted on AWS cloud, thereby providing a solution that is reliable, flexible, and scalable.

Optical Music Recognition Using Template Matching **Spring 2023**

- Developed an automated system for detecting music note symbols and analyzing their pitch in digital music sheets.
- Implemented robust pre-processing techniques, including Gaussian blur and staff line removal, significantly enhancing detection accuracy and making it accessible to musicians and educators for error-free music notation handling.
- Utilized template matching algorithms effectively to enhance symbol detection accuracy while reducing false positives.