ABHISHEK KUMAR

EMAIL PHONE LINKEDIN GITHUB WEBSITE
akumar58@buffalo.edu | +1 716-604-4649 | linkedin.com/in/akumar58 | https://github.com/abhinine4 | https://abhinine4.github.io/

SKILLS

Languages : (**Proficient**) Python, C++, SQL, (**Familiar**) Java, Go, ReactJs
ML Libraries. : PyTorch, TensorFlow, Keras, OpenCV, NumPy, Pandas, Matplotlib, Scikit-Learn, NLTK, SpaCy
ML/ DL : Regression, Classification, PCA, SVM, Random Forest, XGBoost, CNN, Transformers, GANs.
Database/Tools : MySQL, MongoDB, Solr, Elasticsearch, AWS, Git, Docker.

WORK EXPERIENCE

Machine Learning Research Assistant, IAD (SemaFor SRI-UB team) Sept 2022 – Aug 2023

- Implemented JPEG compression analysis and trained a model to **classify** manipulated images with 94% accuracy.
- Created a module to **localize** and predict manipulated regions in images using noise and edge level features.
- Leveraged PyTorch parallel training to train the 1.3 million parameter model, cutting down training time by 37%.
- Trained YOLO-v8 model with 0.69 F1 score on custom dataset to **detect** small object in manipulated regions.
- Containerized the application using Docker and setup end-to-end CI/CD pipeline on Gitlab.
- Designed a linear model over CLIP features that analyzes generated images and predicts aesthetic sores.
- Developed a tool that can perform controlled entity replacements in long texts using Python, SpaCy and NLTK.
- Performed data cleaning, preprocessing, feature engineering, and exploratory analysis on news datasets.
- Detected inconsistencies in multimodal news articles with entity mismatch and contradiction detection algorithms.
- Built a retrieval based chatbot on self-curated reddit dataset. Leveraged Solr for text indexing and searching.
- Implemented embedding based query matching for ranking and co-reference resolution to maintain short contexts

Teaching Assistant, (CSE 701/702) , University at Buffalo June 2022 – Aug 2022

- Guided students to develop deep learning projects with ML design patterns for video analysis in sports domain.
- Conducted classes, graded assignments, and reviewed students' technical presentations.

Software Engineer - Data, Infosys Feb 2018 – Aug 2020

- Developed employee task delegation system to manage production workflow data using Python and React.
- Built REST API to parse delegation data from JSON files received in hourly batches with Flask and Python.
- Designed scripts to extract, parse and transfer millions of parts data across 6 pricing interfaces using PLSQL.
- Reduced batch transfer failure rate by 20% by automating and monitoring batch jobs using RPA and Appworx.
- Built pricing page that shows available dependent and independent part prices from our database in Java.
- Managed deployment and maintenance pipelines for system availability and reliability using Jenkins and Git.

EDUCATION

Masters in Computer Science and Engineering, (University at Buffalo, GPA - 3.72/4.0) Sept 2021 - Aug 2023

- Courses : Machine Learning, Deep learning, Computer Vision, Video Analytics, Distributed Systems, Algorithms
- Thesis: Forensic Methods to Detect Manipulated News Media. Link: https://github.com/abhinine4/news_forensics

PROJECTS

Ear Hair-Cell Detection and Counting, (Python, Template Matching, Non-max Suppression, Dbscan Clustering)

- Successfully detected and extracted damaged inner and outer ear hair cells to measure deafness in animals using supervised and unsupervised machine learning algorithms.

Employee Attrition Prediction, (Python, Sklearn, numpy, random forest, sampling. correlation.)

- Developed machine learning model to predict employee attrition in companies with 88% accuracy. Applied advanced ML algorithms such as feature selection, dimensionality reduction and hyper parameter tuning.

Temporal Action Spotting, (PyTorch, Transformers, NumPy, Matplotlib, Boundary Regression, Action Classification)

- Achieved 5th rank in SoccerNet competition among 50+ participants for classifying 17 action classes and identifying temporal boundaries with 52% mAP using a Transformer + ResNet model with multiscale RGB features.

Player Re-Identification, (Pytorch, C++, transfer learning, ResNet, OpenPose, Bilinear Pooling, Batch Norm, Triplet Loss)

- Increased player re-identification score in broadcast soccer videos by 6%, using a dual branch deep learning model with appearance and body part features extracted by ResNets and an OpenPose subnetwork respectively.

CERTIFICATIONS

- Distributed Deep Learning with Horovod, NVIDIA, June 2022
- Machine learning, Data Science and Deep Learning with Python, Udemy, Jan 2019