mm9314@g.rit.edu | +1  
585  
202  
5217 |

**EDUCATION**

Education:  
- M.S in Data Science, Rochester Institute of Technology, Rochester, NY (Expected May 2025)  
- B.Tech in Computer Science, Jawaharlal Nehru Technological University Hyderabad, July 2018 - July 2022  
  
Skills:  
- Programming Languages: Java, Python, C++, R, JavaScript  
- Object Oriented Programming: Python, Java  
- Frameworks: PyTorch, Keras, Scikit, Tensorflow, Groovy, PySpark, Flask, React, NodeJS  
- Databases: SQL, MongoDB, SQLite, MySQL, NoSQL, PostgreSQL, DynamoDB  
- Technologies/Tools: JSON, Docker, Git, AWS, Kafka, GitLab, Numpy, Pandas, MLflow, Postman, Tableau, Power BI, MLOps  
- ML Algorithms and Techniques: Regression, Classification, Clustering, Recommender Systems, Deep Learning, NLP, A/B Testing, Time Series, Optimization, EDA, ETL  
  
Professional Experience and Internships:  
- R&D Data Governance Intern, Daiichi Sankyo Inc, Basking Ridge, NJ (03/2024 - Present)  
- Research Assistant, Rochester Institute of Technology, Rochester, NY (08/2024 - Present)  
- AI Infrastructure Engineer, SEO Content AI, Los Angeles, CA (Nov 2022 - July 2023)  
- Machine Learning Engineer, White Label Resell, Los Angeles, CA (June 2022 - March 2023)  
- Data Scientist Intern, Digital Clinics Research and Services, Hyderabad, India (Nov 2021 - Dec 2022)  
- Data Scientist Intern, Edgeforce Solutions, Hyderabad, India (Nov 2021 - Feb 2022)  
  
Projects:  
- Chronic Kidney Disease Predictor  
- Beverage Management System  
- Covid19 Bot  
- Flight Price Predictor  
- Generating Synthetic Anime Data  
- Number Plate Detection  
  
Research Papers:  
- Tubules Detection on Breast Carcinoma Whole Slide Images Using Artificial Intelligence (Sep 2022)  
- Web-Based Mitosis Detection on Breast Cancer Whole Slide Images Using FasterRCNN and YOLOv5 (Dec 2022)  
  
Achievements:  
- Kaggle Notebook and Dataset Expert  
- Open Source Contributor at OpenVino (Intel), Ivy  
- Best Overall Hack in CSH Hacks Hackathon 2023  
  
Suggested Skills:  
- Data Mining  
- Data Coding  
- Machine Learning

**SKILLS**

Skills:  
  
Programming Languages: Java, Python, C++, R, JavaScript  
Frameworks: PyTorch, Keras, Scikit, Tensorflow, Groovy, PySpark, Flask, React, NodeJS  
Databases: SQL, MongoDB, SQLite, MySQL, NoSQL, PostgreSQL, DynamoDB  
Technologies/Tools: JSON, Docker, Git, AWS, Kafka, GitLab, Numpy, Pandas, MLflow, Postman, Tableau, Power BI, MLOps  
ML Algorithms and Techniques: Regression, Classification, Clustering, Recommender Systems, Deep Learning, NLP, A/B Testing, Time Series, Optimization, EDA, ETL  
  
Professional Experience and Internships:  
  
Daiichi Sankyo Inc, Basking Ridge, NJ - R&D Data Governance Intern 03/2024 - Present  
- Developed an Informed Consent Forms (ICF) analysis tool using BERT and T5 models, also testing Amazon Bedrock for large-scale language processing.  
- Built a Flask-based frontend for the ICF tool, enabling secure modifications of document classifications based on user permissions.  
- Implemented advanced Large Language Models (LLMs) using Amazon Bedrock, achieving a 20% increase in classification accuracy for detecting explicit prohibitions against data sharing.  
- Utilized Amazon SageMaker for model training and fine-tuning, streamlining the handling of vast volumes of legal documents to ensure scalability and efficiency.  
- Optimized model inference and processing times with EC2 instances, improving overall performance and reliability of the ICF analysis tool.  
- Designed an on-premise chatbot delivering instant trial data, cutting access time by 70%.  
- Enabled visualizations in the chatbot, improving data comprehension and decision-making efficiency.  
- Automated milestone and site tracking, enhancing study performance and reducing missed deadlines by 30%.  
- Developed a secure chatbot architecture, ensuring 100% compliance with company data regulations.  
- Integrated RAG and AI, achieving 80% improvement in answer accuracy and user trust.  
- Led the migration of clinical data from Veeva Vault to Redshift, establishing an ETL pipeline that ensured data accuracy and improved analytics accessibility.  
- Built indexing and foreign key relationships in Redshift to optimize query performance, enabling high-speed analysis on clinical data.  
- Implemented data validation checks during migration to uphold data governance and regulatory compliance, ensuring reliable and accurate data in the new environment.  
- Collaborating with stakeholders to align the tool's capabilities with Daiichi Sankyo's data-sharing policies, compliance standards, and real-world data requirements.  
  
Rochester Institute Of Technology, Rochester, NY - Research Assistant 08/2024 - Present  
- Collaborating with Professor Haibo Yang to enhance federated learning models, utilizing gRPC and PyTorch to implement scalable decentralized training algorithms.  
- Analyzing and optimizing distributed environments with PyTorch RPC, reducing communication overhead by 15% through improved data transfer protocols.  
- Leading the integration of the FedDisco algorithm into existing frameworks, accelerating model convergence and reducing training time by 20%.  
- Transitioning between distributed communication algorithms, leveraging FedDisco's adaptive communication strategies for efficient, real-time model updates.  
- Developing advanced distributed training techniques for Large Language Models (LLMs), achieving a 30% increase in system performance and minimizing network latency.  
- Utilizing gRPC for dynamic, secure communication within federated learning models, enhancing data synchronization and increasing overall model accuracy by 10%.  
- Designing hands-on projects for the Intro to Machine Learning course, applying theory to real-world machine learning scenarios using Python and PyTorch.  
- Providing personalized support to over 25 students, fostering problem-solving skills.  
  
Job Description:  
AI Engineer  
Sunnyvale, CA  
  
Responsibilities:  
- Building and deploying relevance models that power recommender systems and personalized intent engines.  
- Keeping scalability and performance in mind through design and engineering choices.  
- Exploring novel approaches for machine learning ranking and modeling downstream impact.  
- Analyzing big data and tracking data to understand member preferences and identify opportunities for product improvement.  
- Working directly with partner data science, product, and infrastructure teams to take ideas from conception to production.  
  
Basic Qualifications:  
- Bachelor's degree in Computer Science or related technical field or equivalent technical experience.  
- 1+ years experience with machine learning, data mining, and information retrieval or natural language processing.  
- 1+ year of industry experience in software design, development, and algorithm-related solutions.  
- 1+ years experience in programming languages such as Java, Python, etc.  
  
Preferred Qualifications:  
- 1+ years of relevant work experience.  
- MS or PhD in Computer Science or related technical discipline.  
- Experience developing large-scale systems.  
- Knowledge of recommendation systems.  
- Ability to diagnose technical problems, debug code, and automate routine tasks.  
- Analytical approach coupled with solid communication skills and a sense of ownership.  
- Experience in Machine Learning and Deep Learning.  
- Experience in Big Data.  
- Strong technical background & Strategic thinking.  
  
Suggested Skills:  
- Data Mining.  
- Data Coding.  
- Machine Learning.

**PROFESSIONAL EXPERIENCE**

**PROJECTS**

Intro to Machine Learning Course:  
- Applied theory to real-world machine learning scenarios using Python and PyTorch.  
- Provided personalized support to over 25 students, fostering problem-solving skills and guiding successful project outcomes, significantly improving their understanding of distributed machine learning concepts.  
  
SEO Content AI, Los Angeles, CA - AI Infrastructure Engineer:  
- Enhanced AI-driven content generation efficiency by 25% through leading the integration of transformers within AWS microservices.  
- Developed a Chrome extension using Python, JavaScript, and NodeJS, boosting content quality and generation speed by 40%.  
- Utilized Docker for scalable and reliable deployment across cloud environments via AWS ECS and Fargate.  
- Improved project management and team collaboration through implementing effective development methodologies.  
- Optimized content creation for various use cases by conducting A/B testing with models like LLaMA-13B, Flan-T5, Vicuna-13B, and GPT-3.5.  
- Broadened audience reach by using GCP's Google Translate API for multi-language content translation.  
- Ensured high-quality, error-free content by implementing NLP algorithms with BERT, RoBERTa, and DistilBERT for grammar correction, significantly enhancing user trust.  
  
White Label Resell, Los Angeles, CA - Machine Learning Engineer:  
- Automated article generation for digital marketing using AWS Lambda, NodeJS, and API Gateway, achieving a 60x reduction in operational costs.  
- Integrated NLP and TensorFlow to analyze and generate over 130K high-quality articles within a week, improving content strategy.  
- Utilized DynamoDB for managing large datasets, ensuring efficient data access and manipulation for content generation processes.  
- Fine-tuned advanced NLP models like BERT, RoBERTa, ALBERT, and DistilBERT for contextual understanding, improving the relevance and quality of generated articles.  
- Explored generative models including GPT-3, GPT-Neo, GPT-J, and GPT-NeoX, to diversify content creation, ensuring a wide range of engaging and unique articles.  
- Leveraged Flan-T5 and T5 Codegen for specific content generation tasks, optimizing for both efficiency and creativity in article production.  
- Developed custom dataset models with Gensim, enhancing thematic accuracy and alignment with marketing goals.  
- Employed MLOps practices with Git and Docker to streamline the development, training, and deployment of machine learning models, facilitating continuous improvement and collaboration.  
  
Digital Clinics Research and Services, Hyderabad, India - Data Scientist Intern:  
- Developed a sophisticated image classification system using Faster R-CNN, tailored for the nuanced detection of cancerous cells in histopathological images, leveraging TensorFlow for model training and optimization.  
- Engineered a custom segmentation solution with Detectron2, enabling precise delineation of tumor boundaries in medical scans. This approach utilized QuPath for image management and Groovy for scripting complex analysis workflows, significantly enhancing the quality of medical image analysis.  
- Implemented YOLOv5 for the automated screening of pathological slides, achieving unprecedented speed and accuracy in identifying critical diagnostic markers, integrated with OpenCV for real-time image processing and enhancements.  
- Led the integration of MLOps practices, utilizing Docker for containerization and AWS for model deployment, ensuring scalable and robust AI solutions in clinical settings.  
- Mentored a diverse group of 10-20 junior colleagues (first and second-year students) on multiple ML and DL projects.