Sravani Ganta

sravaniganta1999@gmail.com | +1 (226)-972-3043 | linkedin | Github

SUMMARY

Experienced Cloud Engineer with extensive experience in architecting and deploying machine learning solutions on Azure. Demonstrated ability to optimize compute resources, design scalable data pipelines, and enhance model accuracy and system efficiency. Proven expertise in database management, including high availability and disaster recovery solutions, performance optimization, and complex issue resolution. Proficient in RDBMS and various operating systems. Eager to leverage a diverse skill set and passion for data-driven solutions to address complex business challenges.

PROFESSIONAL EXPERIENCE

LTIMindtree, Canada **April 2024 – Present**

Role: Cloud Engineer, Client: Microsoft

- Architecting and deploying machine learning solutions on the Azure platform, leveraging Azure Machine Learning to streamline the end-to-end machine learning lifecycle from data preparation and model training to deployment and monitoring
- Optimizing Compute Instances and Clusters, developing sophisticated Designer pipelines, and streamlining Prompt Flow processes, enhanced model accuracy and overall system efficiency by 30%, reducing latency and improving scalability
- Designed scalable pipelines on Azure, integrating data stores, blobs, and container registries for seamless data ingestion, preprocessing, model training, and deployment. Ensured robust performance and reliability throughout the lifecycle
- Pioneered advancements in DNS lookup and Virtual Networks, leveraging Azure ML Studio and Azure ML SDK V2 to develop and deploy robust machine learning models, achieving a 30% increase in model accuracy and a 25% reduction in processing time
- Designed and implemented scalable pipelines for data ingestion, preprocessing, model training, and deployment
- Collaborating with cross-functional teams to integrate cognitive services such as natural language processing, computer vision, and speech recognition into diverse applications and applying cutting-edge algorithms and frameworks to optimize data-driven, decision-making, achieving 30% reduction in model inference latency and enhancing overall system responsiveness

AIFocal, Canada **November 2023 - December 2023**

Role: Full Stack Developer

- Developed a user-centric admin dashboard for a chatbot application, enabling efficient real-time monitoring and management of user interactions. The dashboard improved data processing efficiency by 40%.
- Executed a responsive UI/UX design, integrated with backend APIs for dynamic data handling. This integration resulted in a 30% reduction in data retrieval times and a 25% increase in overall dashboard usability.
- Enhanced Next.js, React, and performance optimization expertise, contributing significantly to data-driven decision-making and operational efficiency.

University of Waterloo **September 2023 – December 2023**

Role: Teaching Assistant, Supervisor: Prof. Abdalla Mohamed Hussein

Course: ECE 358 - Computer Networks

- Provided guidance and assistance to undergraduate students in problem-solving and completing assignments
- Conducted online office hours and provided Quercus support to enrolled students in the course

University of Waterloo **April 2023 – August 2023**

Role: Machine learning Researcher, Supervisor: Prof. En Hui Yang

- Worked on designing and implementing weight initialization strategies, by transferring knowledge from larger models to smaller models using pre-trained ResNet and ShuffleNet architectures, coupled with Kaiming normalization on CIFAR-100, ImageNet datasets and accomplished a significant enhancement with a 10% faster convergence and a 0.5% increase in accuracy for the child models.

Accenture Private Ltd, India **October 2020 – August 2022**

Role: Database Administrator, Client: Microsoft

- Implemented robust High Availability and Disaster Recovery solutions (Mirroring, Log Shipping, Replication) for critical production servers and resolved maintenance plans, SQL job failures ensuring the execution of scheduled tasks
- Enhanced server capabilities by creating communication and automation tools (Linked Servers, Endpoints, Operators, Database Email). Monitored server performance proactively using Dynamic Management Views (DMVs) and ensured data integrity with Database Consistency Checking (DBCC) commands. Proficiently resolved complex issues like locking, blocking, and deadlocks using SQL Server Profiler, leading to smoother transaction processes and improved efficiency
- Managed SQL Servers (2012, 2014, 2016) focusing on installation, configuration, performance optimization and also administered SQL Logins, roles, and authentication modes in alignment with security policies. Executed database migrations, and addressed database fragmentation, disk space concerns, optimizing storage and overall database performance

SKILLS

Programming Languages: Python, SQL, C/C++, Java, ASP.Net, HTML, KQL
Frameworks/Libraries: TensorFlow, NumPy, PyTorch, Pandas, Seaborn, Matplotlib, Keras, PySpark, NLP, NLTK, Airflow
Tools: Azure Machine Learning, Azure Cognitive Services, Power Bi, Tableau, Visual Studio, GitHub, Airflow, SFG Demo Software.
Skills: DBA, Database Management Systems, Operating Systems, Data Wrangling & Cleaning, Data Manipulation Libraries, Data Exploration & Visualization

EDUCATION

University of Waterloo — **September 2022 – December 2023**
Master of Engineering in Electrical and Computer Engineering — *GPA- 3.9/4.0*

Jawaharlal Nehru Technological University — **June 2016 – August 2020**
Bachelor of Technology in Electrical and Instrumentation Engineering — *CGPA- 9.2/10*

RESEARCH

Weight Initialization Techniques for Enhanced Deep Learning | *Resnet, Shufflenet, Kaiming Normalization, Weight Initialization.*

- Implemented an innovative weight initialization method that included down sampling the larger model weights to align with the smaller model layers, followed by a process of random sampling and Kaiming normalization. Additionally, developed Resnet and Shufflenet models with diverse configurations, involving weight transfers between fully connected and convolutional layers, both with and without Kaiming normalization, evaluated on CIFAR-100 and Imagenet datasets
- Achieved a 10% faster convergence rate and a 0.5% accuracy improvement over baseline models by initializing weights from larger model convolutional layers to smaller ones using Kaiming Normalization.

PROJECTS

SQL Migration, Pipeline Dynamics, and Power BI Insights | *Microsoft SQL Server, Microsoft Azure, Data Visualization, ETL, Power BI*

- Orchestrated the migration of an on-premises SQL database to Azure, ensuring seamless data integration from various sources including CSV files. This foundational work laid the groundwork for advanced data manipulation and analysis.
- Engineered a robust ETL pipeline leveraging Azure Data Factory, facilitating the efficient transfer of data from on-premises environments to Azure Storage. This included the creation of Azure resources and the configuration of a self-hosted integration runtime for cloud connectivity.
- Implemented Azure Databricks for enhanced data processing, employing Spark SQL for table aggregation. This step was crucial for handling large volumes of data and performing complex computations efficiently.
- Finalized the data processing framework with the integration of Power BI, enabling sophisticated data visualization and insights. This allowed for the dynamic representation of data analytics, aiding in decision-making processes.

Admin Login page for Chatbot Management| *React, Next.js, JavaScript, Tailwind CSS*

- Developed an admin page using Next.js and React to efficiently manage chatbot user information, including user activity, last seen status, payment dates, and renewal countdowns.
- Designed and implemented an intuitive user interface for admin operations.
- Successfully fetched and displayed user data, formatted for readability.
- Calculated and displayed renewal countdown based on payment dates.

Medical Condition Classification and Drug Review Analysis | *NLP, python, sentiment analysis*

- Developed an NLP and ML-based drug recommendation system using trigram TF-IDF vectorization and the Passive Aggressive Classifier, achieving a 98.2% accuracy rate in evaluations with a drug dataset. This involved preprocessing text data by removing HTML, special characters, and stopwords, and applying stemming and lemmatization.
- Implemented a comprehensive text analysis approach, utilizing both bag of words and TFIDF vectorizer techniques, and integrated Naive Bayes and Passive Aggressive Classifier models for sentiment analysis and drug recommendation.

Google Cloud Data Pipelining with Airflow | *Google API, Airflow, AWS, EC2, python*

- Developed and automated a data pipeline to extract, transform, and clean YouTube video reviews using Google API, Python, and Apache Airflow, storing data securely on Amazon S3.
- Designed and implemented ETL processes to extract data from multiple sources, transform it using custom Python scripts, and load it into a centralized data warehouse.
- Deployed the pipeline on AWS EC2 instances, ensuring high availability and scalability to handle large volumes of data.
- Implemented monitoring and alerting mechanisms using AWS CloudWatch and Airflow's built-in features to ensure the reliability and performance of the data pipeline.
- Automated the data pipeline with Apache Airflow and stored the processed data securely on Amazon S3.