SHARANYA AKKENAPALLY

Toronto, ON (open to relocate) | +1(514)-569-8534 | sharanyaakkenapally@gmail.com | LinkedIn | GitHub | Portfolio

SUMMARY

- AI Engineer with hands-on expertise in Generative AI, NLP, and production-ready machine learning applications, dedicated to implementing scalable AI solutions that optimize business operations and drive innovation.
- Experienced in data extraction, analysis, and collaborative cross-functional projects, with a focus on optimizing data processes and reporting efficiency to support business objectives.

WORK EXPERIENCE

Cognizant **Hyderabad, India**

Data Science Intern *Mar 2022 - June 2022*

- Collaborated cross-functionally with data engineers and product managers, ensuring scalable data pipelines for seamless integration and utilizing agile development methodologies, promoting iterative progress.
- Developed end-to-end fraud detection systems using LSTM-based RNNs with attention mechanisms, improving detection accuracy by 15%.
- Designed scalable ML pipelines for large-scale dataset, automated workflows using Python and SQL, reducing deployment time by 30%, identifying opportunities for improvement, and enhancing data architecture and reporting efficiency.
- Managed infrastructure with Docker, Kubernetes, and cloud-native tools, deploying and monitoring models at scale using CI/CD pipelines via Git.
- Implemented ETL processes and MLOps to streamline data flow, enhance scalability, and ensure real-time reporting for analytics.
- Created Power BI dashboards to communicate actionable insights to technical and non-technical stakeholders and to monitor KPIs in real-time, enhancing data-driven decisions aligned with business needs.

Technocolabs Softwares Inc **Hyderabad, India**

Data Science Intern *May 2021 - Aug 2021*

- Enhanced model accuracy by 15% through data augmentation techniques and temporal transformations.
- Developed feature engineering workflows using AWS and PySpark to streamline ML workflows, leveraging cloud infrastructure for scalable data storage and access, ensuring efficient handling of large datasets.
- Achieved 85% classification accuracy in time series forecasting for stock price prediction using Logistic Regression, SVM, and optimized models via Grid/Random Search.
- Applied advanced SQL for data querying and reporting to support business analysis and fulfill data requirements across departments.
- Utilized Python for data preprocessing, data wrangling, data analytics, predictive modeling, feature engineering, enhancing reporting capabilities and data transformation processes.
- Created and maintained comprehensive documentation for data architecture and processes, including UML sequence and data flow diagrams to support data transformation and analysis.
- Built interactive dashboards with Tableau to support strategic decision-making and effectively communicate insights, built automated reports enhancing data accessibility and efficiency for cross-functional teams.

The Sparks Foundation **Remote, India**

Data Science and Analytics Intern *Sep 2020 - Dec 2020*

- Processed text data using Natural Language Processing (NLP) techniques such as tokenization and word embeddings to generate high-quality feature vectors, enabling accurate classification.
- Collected, cleaned, and managed large datasets from multiple sources, ensuring data integrity and reliability for analysis.
- Designed and implemented a spam detection model with Support Vector Machine (SVM), achieving a 92% accuracy rate by fine-tuning hyperparameters through Grid Search for optimal performance.
- Visualized performance metrics and model outcomes using Matplotlib and Seaborn, presenting clear insights into model effectiveness and identifying key areas for further optimization in spam detection.

SmartBridge Pvt Ltd **Hyderabad, India**

AI Developer Intern *May 2020 - Aug 2020*

- Developed a CNN-based algorithm for real-time skin disease detection for classification, utilizing OpenCV for image preprocessing and feature extraction in healthcare data analysis.
- Engineered CNN architecture with Max-Pooling, Conv2D, Batch Normalization, and Dropout to enhance feature capture and prevent overfitting.
- Enhanced the model by integrating Transfer Learning, fine-tuning a ResNet-50 architecture, reducing training time by 22% and boosting model accuracy by 4%.
- Deployed the CNN model using Flask, implementing REST APIs supporting high-performance user experiences, and asynchronous processing to manage high user traffic, resulting in scalable application.

EDUCATION

Master of Applied Computer Science - *Concordia University, Canada* *Sep 2022 - May 2024*

B. Tech in Information Technology - *Sreenidhi Institute of Science and Technology, India* *Aug 2018 - Jul 2022*

TOP SKILLS

- **Programming languages & Databases:** C, Data Structures, C++, Python, R, Java, MySQL, NoSQL
- **Web Technologies**: HTML, CSS, JavaScript
- **Machine Learning:** PyTorch, Keras, TensorFlow, Hugging Face, Deep learning, Generative AI, NLP, Transformers, LLM Fine Tuning, MLOps, Sci-kit learn, EDA, OpenCV, CUDA
- **Data Visualization & Analysis:** Tableau, Power BI, MS Excel, R-Studio, Matplotlib, Seaborn, SpaCy, Pandas, Numpy
- **Cloud & DevOps:** Azure ML, AWS (Sagemaker), PySpark, HiveQL, HBase, Oracle cloud (OCI), Docker, Kubernetes, CI/CD pipelines, Git, Flask

PROJECTS

Taxi Demand Prediction **Oct, 2024**

- Conducted EDA and engineered time-based and interaction features (e.g., rush-hour precipitation) to capture temporal patterns and enhance prediction accuracy.
- Built and optimized an ensemble model (XGBoost, LightGBM, Random Forest) using GridSearchCV for hyperparameter tuning to improve hourly demand forecasting.
- Integrated time-series weather forecasts, including lagged features and seasonal adjustments, to account for real-time environmental factors in demand fluctuations.

Transformer based TTS System for Speech Synthesis **Feb - Apr, 2024**

- Developed a Transformer-based Text-to-Speech (TTS) model with the LJSpeech dataset, improving speech quality and performance using the SpeechBrain framework.
- Accelerated training using CUDA, cuDNN, and GPU optimizations, achieving 2.5x faster speeds than Tacotron2. Experience in distributed training and optimization of language models to solve resource allocation to enhance model efficiency.
- Applied scaled positional encodings and teacher forcing, reducing Mel Error to 8.27e-02 and lowering Stop Error by 10%.
- Enhanced efficiency with dynamic batching and Noam Scheduler for more natural and fluid speech synthesis.

Dialogue Summarization: Fine Tuning LLM using Prompt Engineering and PEFT **Dec 2023 - Jan 2024**

- Explored and Fine-tuned FLAN-T5 Language Model from Hugging Face for dialogue summarization using prompt engineering and zero/few-shot learning to improve adaptability for LLM-based solution.
- Applied PEFT (LoRA), updating just 2-12% of parameters, reducing model size for fine-tuning and optimizing the LLM.
- Improved ROUGE-1 by 17.47% and ROUGE-2 by 8.73% over human baseline summaries, ensuring efficient compression and accuracy.

Automated Retail Product Classification using CNN **Oct - Dec, 2022**

- Designed CNN architectures (ResNet-18, GoogleNet, AlexNet) to classify grocery products, addressing class imbalance with weighting and oversampling.
- Performed data augmentation, model scaling and utilized grid search for hyperparameter optimization, fine-tuning learning rates, batch sizes, and dropout rates and leveraged statistical techniques to support model evaluation and selection, ensuring robust performance metrics.
- Leveraged transfer learning by fine-tuning ResNet-18 with pre-trained ImageNet weights, resulting in an 8% increase in classification accuracy.

CERTIFICATIONS AND ACHIEVEMENTS

- Oracle cloud infrastructure 2024 generative AI certified professional
- Machine Learning certification - Stanford University
- The Data Scientist Toolbox certification, Johns Hopkins University
- AWS Academy- Cloud Foundations course - Amazon
- Silver Award at Ennovate - The International Innovation Show, Poland, 2021