Shravani Hariprasad

+1 (619) 673-2889 shariprasad3296@sdsu.edu shravani-h-507280177 shravani-01

EDUCATION

San Diego State University **Aug 2022 – May 2024**
Master of Science in Big Data Analytics(GPA : 4/4) *San Diego, California*

Anna University **Aug 2017 – Jun 2021**
Bachelor of Engineering in Electronics and Instrumentation(GPA : 3.85/4) *Chennai, India*

EXPERIENCE

L&T Infotech **Jul 2021 – Jul 2022**
Senior Software Data Engineer *Mumbai, India*

- Developed and implemented automated solutions for natural language commentary generation and variance analysis in financial data, leveraging agile methodology throughout the software development lifecycle, leading to a 60% reduction in time spent on manual report creation and enhancing accuracy in financial reporting processes.
- Implemented the Isolation Forest machine learning model, resulting in a 20% reduction in false positives and an 85% improvement in the identification of impacted nodes when variance values exceeded predefined threshold levels. This enhanced financial accuracy and efficiency, empowering proactive data-driven decision-making.
- Python automation script to calculate and generate pivot tables for each financial account, streamlining data analysis processes, simplifying manual effort by 50% and enhancing data validation accuracy, resulting in a 20% decrease in error rates.
- Collaborated with the data engineering team to design ETL workflows using Apache Spark, enabling the extraction, transformation, and loading of data from diverse sources into the data warehouse, leading to a 30% reduction in data processing time.
- Utilized Power BI to visually represent financial data, enabling clients to identify trends and variances between report periods more efficiently, resulting in a 5% faster decision-making process for clients.

Cognizant Technology Solutions **Jan 2021 – Jun 2021**
Data Engineer Intern *Chennai, India*

- Proficient use of SQL for data extraction and structured querying, leading to a 20% increase in data retrieval and transformation efficiency.
- Developed data governance requirements, leveraged Python, Apache Spark and AWS services (S3, Amazon Glue - ETL processes) to automate tasks and achieve a 30% reduction in data processing time, significantly enhancing workflow efficiency.

RESEARCH AND PROJECTS

Deep Learning Image Classification with Google Street Views (GSV) - HDMA Lab **Jun 2023 – Present**

- Utilized deep learning models (Xception, YOLO) to address homelessness detection by analyzing Google Street View images in downtown San Diego.
- Curated a comprehensive dataset from Google Street View, focusing on key downtown streets. Extracted significant metadata for sidewalk-based analysis to identify homeless individuals and their habitats.
- Expected to improve model accuracy by 15%, showcasing the potential for real-world applications in addressing homelessness and urban planning.

Content Analysis of HIV Prevention Drug - PrEP(Co-author) **Mar 2023 – Present**

- Co-authoring an NIH grant-funded project analyzing public perceptions of the HIV prevention drug PrEP, through social media analysis, currently in the Exploratory Data Analysis phase.
- Utilizing Python for efficient data collection, preprocessing, and EDA to gain insights into post types, user demographics, sentiment analysis, and influential accounts.
- Preparing to employ Natural Language Processing (NLP) and machine learning techniques for sentiment analysis, classification of posts into positive, negative, or neutral sentiments, and providing data-driven recommendations to enhance PrEP awareness, reduce stigma, and guide targeted interventions for the NIH and healthcare stakeholders.

Image Caption Generation of the MSCOCO Dataset (Team Leader) **Feb 2023 – May 2023**

- Built a full-end ML project in PySpark on GCP, leveraging its distributed computing capabilities to efficiently handle large-scale COCO dataset and ETL workflows using AirFlow and improving data processing time by 50%.
- Pre-trained an Xception-LSTM model using TensorFlow within the GCP-based PySpark environment for generating image captions, cutting down the loss to 1.85%.
- Evaluated model performance using standard metrics, achieving a BLEU score of 0.35.

Interactive COVID-19 Data Visualization Project using Cyber-GIS Tool **Dec 2022 – Mar 2023**

- Collaborated on an advanced research project utilizing the Cyber-GIS tool to enhance COVID-19 data visualization.
- Leveraged the Adaptive Choropleth Mapper to create interactive map applications, improving data accessibility by 40%.
- Developed a spatial analysis platform for COVID-19 data, allowing in-depth correlation between confirmed cases, fatalities, and vaccination rates, contributing to more insightful data-driven decision-making.

Human Activity Recognition (Team Leader) **Aug 2022 – Dec 2022**

- Examined machine learning techniques for diverse human activity prediction in hospital environments with both Raw time-series and feature-engineered data.

- On feature-engineered data with classical machine learning classifiers, attained a maximum accuracy of 96.47% and the least error rate of 3.291%, surpassing the deep learning and time series models.

Analysis of public sentiments on climate change for formulating policies **Sep 2022 – Dec 2022**

- Addressed the research question, "To what extent can public sentiments on social media platforms such as Twitter influence national or global policies on climate change?"
- Employed Named Entity Extraction techniques to identify relevant organizations or locations within the discourse
- Forecasted that the Logistic Regression model achieved the highest accuracy of approximately 70% compared to other models, providing valuable insights into the research question's context and potential implications for policy formulation.

ACCOMPLISHMENTS

- One of the five recipients of the **2023-2024 SDSU Master's Research Scholarship ($10,000)** for outstanding academic performance and contributions to research.
- Completed training, worked on hands-on projects, and certification by L&T Infotech with proficiency in front-end(React.js) and back-end(Servlets, JSP, Java Spring Boot) frameworks, APIfication with microservices, database(SQL NoSQL), code quality, cloud deployment(Azure), and DevOps(Azure DevOps, Jenkins, Ansible).
- **Techgium Competition 2019-20** - Certificate of Recognition for innovative concept on Multilingual Text Extraction from Videos using YOLO by Larsen & Toubro Technology Solutions.

SKILLS

Languages	Python, R, C++, Java, JavaScript
Frameworks	Pandas, Matplotlib, scikit-learn, TensorFlow, Keras, PyTorch, NLTK, flask, React.js, Spring Boot, MERN
Tools	Tableau, PowerBI,ArcGIS Pro, Jira, Hadoop, Spark, Hive, Agile, Git, Airflow, DevOps, Docker, Kubernetes
Database	MySQL, MongoDB, Google BigTable, Snowflake
Cloud Technologies	AWS (EC2 Instance, S3 Bucket, Glue), Google Cloud (Big Query, Bucket, DataProc), Azure
Area of Interest	Data Science, Machine Learning, Software Engineering, Automation