SAI PAVAN PADAMATA

Atlanta, GA 📞 (404)-528-3067 ✉ saipavanpadamata@gmail.com 🔗 linkedin.com/in/sai-pavan-padamata/ ⭘ variable-z

Education

Georgia State University   **Aug 2024 - Present**
MSA, Data Science and Analytics, Big Data and Machine Learning   *CGPA:4.08/4.3*

Prasad V Potluri Siddhartha Institute of Technology   **Jun 2018 - May 2022**
B. Tech, Computer Science   *CGPA:3.4/4.0*

Experience

Georgia State University

Graduate Research Assistantship - NLP, ML, Text Analytics   ***Aug 2024 - Present***

- Built a comprehensive data model for a hotel review system, leveraging 300,000+ entries to deliver actionable insights
- Designed and developed an OCR-based solution to extract, clean, and structure Bloomberg Terminal data for supply chain analysis.
- Streamlined data processing pipelines using Python, OCR (Tesseract), and regex, enabling seamless integration of structured datasets. Conducted comprehensive analysis of supplier, customer, and competitor data for 5+ years, enabling actionable insights into supply chain relationships.

Cognizant Technology Solutions

Programmer Analyst - AIA Modern Data Engineering   ***Nov 2022 - Jul 2024***

- Orchestrated the migration of 194+ reports from SAP BusinessObjects to Qlik Sense, meeting precise client requirements.
- Engineered scripts to optimize data loading and preprocessing from Oracle databases, halving file sizes and improving retrieval speeds.. This reduced data file sizes by 50%, enhancing retrieval speed and overall efficiency.
- Unified over 207 tables into key data models, aligning with Universe Design Tool specifications for seamless integration four key data models (Dispute, Retainer Master & Invoices, Vendors, Matter Summary) according to Universe Design Tool specifications.
- Enhanced user engagement by redesigning dashboards and reports, improving usability and access efficiency by 40%

Advanced Python Developer - Internship   ***Feb 2022 - Sep 2022***

- Optimized data ingestion processes using Hadoop tools, importing data into Hive tables with a 40% efficiency boost
- Executed advanced data transformations in Databricks, leveraging PySpark to accelerate processing time by 30%

Projects

Grocery Store Management System - Flask, MySQL, SQLAlchemy, HTML, CSS, JavaScript

- Developed a full-stack web application for grocery store management, supporting 500+ users for registration, login, cart management, order placement, and membership-based discounts.
- Normalized a MySQL database, reducing data redundancy by 35% and improving query performance by 20%
- Built employee functionalities for product management, shift scheduling, and customer memberships, enhancing operational efficiency.
- Integrated inventory stock validation and membership expiration checks, reducing manual errors by 30% and ensuring secure session-based operations

News Classification using NLP - *NLTK, Text Analytics, Sci-Kit Learn, Machine Learning, Deep Learning*

- Engineered a machine learning model for news article classification, distinguishing between genuine and fake content using NLP techniques with NLTK and Sci-Kit Learn, on a large-scale dataset of 700,000 records with 4 features.
- Delivered and evaluated various classification algorithms, achieving an accuracy rate of 94% and an F1-score of 0.8, which improved the model's reliability in detecting fake news.

Recognition of Objects with Convolutional Neural Network - *CNN, Tensorflow, Keras, Numpy*

- Constructed a Convolutional Neural Network (CNN) model for object recognition using the CIFAR-10 dataset, achieving an accuracy of 72%.
- Additionally, adapted the same CNN architecture for handwritten digit classification using the Keras MNIST dataset, which includes 60,000 training rows and 10,000 test records, showcasing versatility and effectiveness in various visual recognition tasks.

Technical Skills

- **Programming Languages & Web Technologies**: C, Python, Java, JavaScript, HTML, CSS
- **Databases & Integration Tools**: MySQL, Apache Sqoop, Apache Flume, Apache Spark, Kafka, Apache Hive, HBase, Pyspark
- **Business Intelligence & Visualization tools**: Power BI, Looker, Data Studio
- **Cloud Platforms** AWS (EC2, S3, RDS, Snow Family), Google Cloud (BigQuery, Dataproc, Pub/Sub)
- **Operating Systems** Unix, Linux, Windows)

Research Paper and Certificates

- **INFOSYS**: certified Software Programmer
- **AWS** : certified Cloud Practitioner, Solutions Architect Associate
- A research paper was published named "**News Classification using Natural Language Processing**" with IEEE index at 3rd International Conference of ICIEM 2022, Electronic ISBN: 978-1-6654-6756-8.