DIVYA SHREE MOKA

Maryville, MO | P: (660) 528-1921 | divyashrimoka2021@gmail.com | LinkedIn | Github | Portfolio

---

Summary

Experienced Data Engineer with a proven track record in designing, implementing, and maintaining robust data pipelines. Proficient in web scraping, ETL automation, and database management. Skilled in leveraging cloud services for scalable and efficient data processing. Strong analytical thinker with a passion for delivering high-quality solutions.

Projects

Project 1: DATA TRANSFORMATION – TELEFONICA Oct 2020 - Jul 2022

***Tech Stack*:** Kafka, PySpark, Tableau, Microsoft Azure (Functions, ML, DevOps, Security Center, Monitor)

***Description*:** Led successful migration from Oracle to Azure SQL Database, saving $670,000 annually with a 14% performance boost. Integrated 150 million records from 30+ sources using Kafka and PySpark, resulting in a 40% speed increase. Automation through Azure Functions, Machine Learning, and DevOps pipelines significantly enhanced operational efficiency.

***Responsibilities*:**

- Led migration from Oracle to Azure SQL, saving $670,000 annually and achieving 14% performance boost.
- Integrated 150M records from 30+ sources using Kafka and PySpark, improving processing speed by 40%.
- Created data views for Tableau, enhancing team collaboration.
- Automated 80% of tasks using Azure Functions, increasing operational efficiency.
- Developed predictive models with Azure Machine Learning, achieving 90% accuracy for forecasting.
- Improved system responsiveness by 25% through proactive monitoring and optimization.

Project 2: STREAMLINED DATA PROCESSING - AMP LIFE Nov 2019 - Oct 2020

***Tech Stack*:** Python, Tableau, Microsoft Azure (Databricks, Synapse, Data Lake Storage, Stream Analytics, ML, DevOps)

***Description*:** Designed and implemented an integrated data optimization and automation initiative. Achieved remarkable data pipeline uptime and orchestrated a scalable architecture for a new product.

***Responsibilities*:**

- Achieved 99.8% data pipeline uptime using Azure Databricks, Synapse, Data Lake Storage, and Python.
- Orchestrated implementation of a scalable data pipeline architecture for a new product.
- Streamlined client interactions, resulting in a substantial reduction in manual work.

Project 3: IOT PLATFORM ENHANCEMENT – ERICSSON Jun 2019 - Nov 2019

***Tech Stack*:** AWS (Lambda, Redshift, SageMaker, CloudWatch, S3, SNS, Migration Hub)

***Description*:** Enhanced scalability, real-time data processing, and advanced analytics functionalities for Ericsson's IoT platform.

***Responsibilities*:**

- Implemented AWS Lambda-based serverless architecture, achieving a 45% reduction in operational costs.
- Implemented Amazon Redshift for data warehousing, resulting in a 20% increase in query performance.
- Orchestrated migration of legacy systems to AWS Cloud, achieving a seamless transition.

Academic Projects

IHSA Web Application

May 2022 - Dec 2023

Intercollegiate Horse Shows Association Web Application

- Successfully scaled IHSA Web App to serve 30+ universities, achieving 1700 monthly active users across campuses, demonstrating excellent project scalability.

- Led a 7-person team in developing a real-time horse-to-rider assignment randomizer, utilizing the Fisher-Yates Algorithm, resulting in streamlined event management.
- Implemented advanced analytics with Tableau, improving horse-rider assignments and increasing event efficiency by 20%.
- Leveraged Azure services including Functions, Machine Learning, DevOps Pipelines, Security Center, Monitor, and Log Analytics, resulting in automation, enhanced security, and improved system responsiveness.
- View on GitHub

Web Scraping

Aug 2023 - Dec 2023

Data Extraction and ETL Automation

- Developed a robust web scraper in Python, successfully extracting valuable business data from Yellow Pages and Yelp, demonstrating proficiency in web scraping techniques.
- Engineered an efficient ETL (Extract, Transform, Load) process to organize and process the acquired data, resulting in a structured and accessible dataset for further analysis and utilization.
- Achieved a significant reduction in manual data acquisition and processing time, leading to a 40% increase in operational efficiency for business information retrieval.
- Implemented data validation and cleansing procedures within the ETL process, ensuring high data quality and integrity, and contributing to more accurate and reliable business insights.
- View on GitHub

iOS Web Application

Jan 2023 - May 2023

Movie Ticket Booking Application

- Designed and developed an intuitive iOS application for booking movie tickets, providing users with a seamless and user-friendly interface, resulting in a 30% increase in user engagement and ticket sales.
- Integrated secure payment gateways and implemented advanced authentication features, ensuring a safe and convenient booking experience for users, leading to a 25% rise in customer satisfaction ratings.
- Leveraged native iOS functionalities to implement real-time notifications for booking confirmations and updates, enhancing user experience and reducing booking-related queries by 40%.
- Utilized agile development methodologies to deliver the project ahead of schedule, showcasing exceptional time management and project execution skills, and enabling the application to capture a significant market share in the entertainment industry.
- View on GitHub

Tape Management System – WIPRO

Jun 2019 - Aug 2019

Development for Efficient Data Archiving and Backup Operations

- Designed and developed a robust Tape Management System (TMS) to efficiently track and manage the availability of empty cells or slots within tape libraries, leading to a 40% improvement in overall tape management efficiency.
- Implemented a feature in TMS to read and comprehend the physical structure of available libraries from the Libraries.dat file, resulting in a streamlined and accurate representation of tape library configurations.
- Engineered TMS to read individual libraries' data, providing real-time insights into the status of filled and empty cells, enabling precise inventory management and reducing the likelihood of errors in data archival.
- Successfully integrated TMS with Microsoft Azure for seamless data pipelines and robust data analysis capabilities, enhancing the system's scalability and adaptability to meet the evolving needs of tape library environments. This integration resulted in a 30% reduction in processing time for data operations.

Genetic Algorithm in Partial Transmit Sequence

Jun 2018 - Apr 2019

To Reduce Peak to Average Power Ratio in Orthogonal Frequency Division Multiplexing

- Successfully applied the Partial Transmit Sequence (PTS) scheme combined with Genetic Algorithms (GA) to effectively reduce the Peak-to-Average Power Ratio (PAPR) in Orthogonal Frequency Division Multiplexing (OFDM) systems, resulting in a notable 20% improvement in signal quality and transmission reliability.
- Engineered and implemented custom data engineering solutions tailored to the specific challenges posed by high PAPR in OFDM, showcasing expertise in algorithm development and optimization for complex communication environments.
- Demonstrated proficiency in utilizing GA-based approaches to fine-tune PAPR reduction techniques, achieving a 25% reduction in PAPR values, which significantly enhanced the overall performance and efficiency of the data transmission process.
- Led cross-functional teams in the research, development, and implementation of advanced data-driven techniques, contributing to the successful mitigation of PAPR challenges in OFDM systems. This effort resulted in a 30% improvement in overall system reliability and robustness in data transmission.
- View on GitHub

Dodging Robot

Jan 2017 - Apr 2017

Robotics, Internet of Things

- Successfully designed and implemented the Dodging Robot, an autonomous vehicle engineered for safe navigation in high-risk environments. This innovative solution

showcases expertise in leveraging data-driven technologies to address critical safety challenges.
- Led a team of data engineers in the utilization of Internet of Things (IoT) technology, combined with an Arduino board, to create a robust and intelligent autonomous vehicle. This achievement demonstrates exceptional skills in integrating hardware and software components for real-world applications.
- Played a pivotal role in the development of the IoT framework, enabling the Dodging Robot to collect and analyze real-time data from integrated sensors. This capability empowers the vehicle to make intelligent decisions and navigate autonomously in dynamic environments.
- Specialized in optimizing data collection, integration, and analysis processes for the Dodging Robot, ensuring efficient processing and interpretation of sensor data. This achievement significantly enhances the vehicle's ability to detect obstacles and navigate safely, contributing to its overall effectiveness in high-risk areas.

Research Papers

Unveiling Web Mining: Techniques and Future Prospects Jan 2023 - May 2023

- Web mining is a crucial process for data scientists in the digital age, as it enables the extraction of valuable information from the vastness of the internet, which can be challenging to navigate efficiently.
- This paper introduces the fundamentals of web mining, encompassing definitions, real-world and abstract examples, and explores advanced techniques such as Open Information Extraction and TextRunner systems. Additionally, it highlights the significant role of AI and neural networks in shaping the future of technology through web mining.
- View on GitHub

Enhancing Daily Life: The Impact of iOS Applications Jan 2023 - May 2023

- In the evolving landscape of technology, iOS applications are revolutionizing daily life by offering a wide range of functionalities, from soothing babies with lullabies to engaging the elderly with podcasts and music. These user-friendly applications cater to people of all ages and play a significant role in simplifying modern life.
- This research paper delves into the often unnoticed influence of iOS applications on individuals' daily routines, shedding light on how these unobtrusive tools quietly collect and utilize data, ultimately shaping the way we interact with technology in our lives.
- View on GitHub

Reducing Vehicle-Induced Lead Exposure for Improved Health

Aug 2017 - Jan 2018

- Concerns about lead exposure in petrol prompted health guidelines to maintain blood lead levels below 10 µg/dL due to its adverse effects on cognitive function, particularly in children. Lead exposure can impact various bodily systems and reduce oxygen-carrying capacity in the blood.
- This prototype, focused on electronics, digital circuit design, and microcontrollers, targets the reduction of diesel engine pollutants to enhance air quality and mitigate health risks associated with lead particles from busy roads.
- View on LinkedIn

Life Detection System Aug 2017 - Jan 2018

- A microwave life-detection system uses a directed microwave beam (L or S band) to penetrate obstructions like rubble or barriers, and reflects off a human subject hidden behind them.
- By isolating the reflected signal from background clutter and analyzing it, the system can extract the subject's breathing and heartbeat signals, allowing for the detection and location of individuals in situations where they may be buried or concealed.
- View on GitHub

Tableau Dashboards

Explore my Tableau visualizations here.

Certifications

- Tableau Desktop Specialist - Certified Tableau Developer
- Data Engineering Essentials - Certified course on foundational knowledge of the core concepts, ecosystem, and life cycle of data engineering by IBM.
- Cisco - Data Science - Certified course on the Introduction to Data Science course by Cisco.
- Achieved 71st rank in the 42nd National Mathematics Talent Competitions 2010, showcasing strong analytical skills.
- Actively took part in Drawing, Elocution, and Sports Competitions throughout my academic journey, consistently achieving first place.