

DARSHAN SENTHIL

New Brunswick, NJ, United States | +1(848)-313-9989 | www.linkedin.com/in/darshan-senthil
<https://github.com/Darshan-Senthil> | www.darshansenthil.com | <https://public.tableau.com/darshan.senthil>

Experienced Data Analyst and Machine Learning Engineer proficient in leveraging advanced analytical techniques to drive strategic decision-making and innovation. Passionate about uncovering hidden patterns in data to solve real-world problems, optimize business performance, and facilitate informed decision-making.

SKILLS

Programming Languages: Python, SQL, R, Scala, Bash, Java, Julia

Developer Tools: Tableau, PowerBI, Git, Salesforce, PySpark, Google Cloud Platform, Databricks, Alteryx, SAS, Looker, Snowflake, Dataiku, QlikView

Technologies/Frameworks: AWS (S3, EC2), Azure, Airflow, Linux, Redshift, Flask, Docker, Kubernetes, Kafka, MongoDB, Hadoop, FastAPI, Spark, TensorFlow, PyTorch

EXPERIENCE

Research Assistant, Rutgers School of Public Health (October 2023 - Present)

- Implemented custom CI/CD pipelines, reducing deployment time by 45% and ensuring efficient, reliable deployment of data pipelines and analytical solutions.
- Leveraged OCR technology to extract information from images, enabling precise text detection and processing with an accuracy surpassing 90%.
- Designed a model leveraging advanced clustering algorithms for categorizing content, achieving over 92% accuracy in data clustering.
- Incorporated NLP methods to enhance the model's capabilities in precise text classification, sentiment analysis, as well as title and summary generation.

Methodologies: LLM, Hugging Face, Selenium, Jenkins, Python, R, OCR, T5, GPT, SQL, Git, WebDriver

Data Analyst, Heldrich Center for Workforce Development (June 2023 - August 2023)

- Uncovered significant correlations between Career and Technical Education (CTE) Program participation metrics and employment rates, resulting in a remarkable 5-fold increase in CTE program enrollment and employment rates.
- Established data pipelines and databases for real-time visualization and reporting across teams within the New Jersey Education Department.

Methodologies: Python, AWS, Apache - Spark, Hadoop, Airflow, PySpark, Power BI, Microsoft Excel, R, PostgreSQL

Data Analyst, MAD STREET DEN (February 2021 - August 2022)

- Directed data-driven decision-making as **Head of Onboarding** - Customers in Analytics, ensuring integration and satisfaction of over 20 new clients.
- Implemented MapReduce on the Databricks platform, resulting in a reduction in runtime by over **58.6%**, enhancing overall performance.
- Drove significant revenue growth, achieving **5x to 12x** increases for over 36+ global enterprises through analytics, and strategic guidance.
- Developed an alerting system that monitors metrics, addresses declines and issues, and sends alerts via Slack, achieving a **75%** workload reduction.
- Supervised and mentored a team of 4 interns within the analytics department, providing guidance on advanced data analysis techniques.
- Deployed a verification system that decreased data verification time by **98.61%**, generating precise final reports on missing data.

Methodologies: NoSQL, Supply Chain, Bash, Scala, AWS, PySpark, Google Cloud Platform, Domo, Databricks, Tableau, JIRA, SAS, PyTorch, A/B Testing

Machine Learning Engineer, DappLogix Software Pvt Ltd

(January 2020 - November 2020)

- Created an LSTM-based sentiment analysis model for a resume parser, achieving a 94% accuracy rate.
- Enhanced NLP techniques and Image Processing for robust text classification and sentiment analysis by implementing advanced methodologies.
- Established APIs for data collection, integrated custom pipelines, and employed Selenium for automation, facilitating seamless data transfer.

Methodologies: *TensorFlow, TF-IDF, Word to Vectorization, Stemming, Label Encoder, Tokenization, MongoDB, Looker, JavaScript*

PROJECTS

Colourization of Gray Scale Images

- Implemented a customized GAN architecture tailored for video colorization, combining spatial and temporal information to create smooth color transitions.

Methodologies: *TensorFlow, PyTorch, Convolutional Neural Networks, GAN*

Personalized Movie Recommendation System

- Developed and evaluated Collaborative Filtering and Item-based Filtering recommendation system algorithms for personalized movie recommendations.

Methodologies: *Singular Value Decomposition (SVD), K-Nearest Neighbors, Slope-One*

Crime News Sentiment Analysis

- Created a sentiment analysis system for New York crime news, employing web scraping methods for data acquisition and implementing time-series visualization techniques for sentiment score monitoring.

Methodologies: *Transformers, VADER, TextBlob, BeautifulSoup, NLTK, Tableau, WebDriver*

Car Price Analysis Modeling

- Implemented a predictive system for estimating used automobile prices by analyzing diverse price-related variables and employing regression techniques to construct precise price prediction models.

Methodologies: *Python, R, Data Visualization - Matplotlib, Seaborn, Plotly, ML Algorithms*

COVID-19 Data Analysis

- Conducted an analysis of Covid-19 data, utilizing statistical methods and machine learning algorithms to identify trends, patterns, and insights, thereby contributing to research efforts in pandemic epidemiology.

Methodologies: *Python, Tableau, Hypothesis Testing, Clustering Algorithms, Time Series Analysis*

EDUCATION

Rutgers, The State University of New Jersey (GPA: 3.96)

Masters of Science in Computer Science

Course highlights: *Math for Data Science, Statistics for Machine Learning, Artificial Intelligence, Machine Learning II, Brain Inspired Computing, Advanced Database Management System*

Anna University

Bachelors in Computer Science

Course highlights: *Probability and Statistics, Data Analytics, DBMS, Software Engineering, Operating Systems*

AWARDS AND QUALIFICATIONS

- International Student Orientation Leader, Rutgers University
 - CURATOR, Head of TEDxSriSairamIT 2019
 - Best Student Project Award, Computer Society of India 2019
 - Title winner, BLOCKATHON 2018
-

CERTIFICATION

- Google Data Analytics Professional Certification
- IBM Data Engineering Certification
- AWS Cloud Certification
- Tableau Desktop Certification