RAMDAYAL NIRMALKUMAR REWARIA

Lead Data Science Engineer

📞 +91 9137092308   @ rr1220@georgetown.edu   🔗 https://ramdayalrewaria.quarto.pub/   📍 Mumbai, MH, India

EDUCATION

Master of Science, Data Science and Analytics

Georgetown University
08/2022 - 05/2024   Washington DC

- Top 5% in Cohort; Merit based Scholarship
- GPA (3.98/4)

Bachelor of Technology, Data Science

Narsee Monjee Institute of Management Studies
07/2017 - 05/2021   Mumbai

- Top 20% in Cohort
- GPA (3.25/4)

EXPERIENCE

Lead Data Science Engineer | Co-Founder

SwipeSwipe   06/2024 - Ongoing   Delaware, USA

- Built and maintained ETL pipelines to ensure smooth data flow into the unified model, optimized data storage solutions for performance and cost-efficiency, and ensured data quality, integrity, and security
- Created and implemented algorithms and machine learning models for advanced data analysis, deriving actionable insights while ensuring scalability and maintainability
- Developed a recommendation engine to extract data from SerpAPI, comparing results from Amazon product links to Google Shopping API, and provided recommendations based on similarity ranking resulting in a 20% increase in relevant product suggestions and enhancing user satisfaction by 15%

Data Analytics Intern

Magoosh Inc   01/2024 - 05/2024   San Francisco, CA, USA

- Assisted in creating and structuring data models for 2 different products offered to 25 school partners, enhancing the efficiency and usability of educational data by 30%
- Contributed to the development of customer-facing content platforms and proprietary data analytics, providing valuable insights and improving customer experience for 50 school partners, leading to a 20% increase in user satisfaction
- Automated the data workflow pipeline aiding in the formulation of data optimization strategies, regularly prepared detailed reports to track project progress and outcomes, and optimized the processing times, reducing it by 99.98%, saving approximately 500 hours annually

Data Management Analyst

Georgetown University   01/2023 - 12/2023   Washington, DC, USA

- Modernized the workflow pipeline by transitioning from Stata to a powerful combination of Python and R, resulting in a 40% reduction in data processing time and improved efficiency
- Analyzed Teacher Wellbeing over 7 years using models like XGBoost, Random Forest, and Neural Networks, predicting depression rates with an accuracy of 82%, providing insights that influenced policy changes affecting over 10,000 teachers

Qualitative Research Analyst

Center for Global Health Practice and Impact   08/2022 - 2023   Washington, DC

- Assisted with the LISTEN Process, analyzing 150+ interview transcripts from Kenya and Eswatini residents to derive actionable insights, contributing to reports that guided international health policy changes
- Played a pivotal role in advancing strategies for up-scaling HIV and AIDS treatments, impacting over 100,000 individuals by harnessing local innovations

Data Scientist

ICICI Bank   07/2021 - 08/2022   Mumbai, India

- Spearheaded automation of the Society Ecosystem, achieving a 200% boost in computing times using Python, reducing project completion times from 10 hours to 3 hours
- Refined a Customer Lifetime Value model, enhancing accuracy to 82%, facilitating targeted banking strategies by segment analysis, resulting in a 15% increase in customer retention
- Streamlined data pre-processing and cleaning for private banking projects, shifting from SQL to Python, reducing process time by 50% and saving approximately 300 hours annually

Machine Learning Engineer

Konsultera Solutions Pvt. Ltd   04/2020 - 12/2020   Mumbai, India

- Developed a custom TensorFlow object detection model for monitoring social distancing with 81% accuracy, improving compliance monitoring by 25% in public spaces
- Led optimization of a 2500-case law firm database utilizing advanced NLP techniques and Spacy, reducing data retrieval times by 60%

PROJECTS

Human Activity Recognition | TensorFlow, PyTorch, LSTM, CNN, Python

08/2023 - 12/2023   Location

Implemented a Human Activity Recognition system with LSTM and Multi-Channel CNN, processing accelerometer and gyroscope data from smartwatches to classify six activities, achieving a high accuracy of 94%.

Reddit Analysis | EC2, S3, Sagemaker, Azure, Spark, Python, NLP, ML

08/2023 - 12/2023   Location

Utilized PySpark, SparkNLP, and various cloud tools to perform extensive EDA, NLP, and ML on a dataset of over 3 million rows, focusing on analyzing trends and patterns in entertainment Subreddits.

Soccer Match Summarization | Spacy, NLTK, Generative AI, NLP, Streamlit, Python

01/2023 - 05/2023   Location

Developed a real-time Streamlit application with a comprehensive data pipeline that collected, processed, and visualized over 10,000 UEFA Champions League commentaries, utilizing Spacy, NLTK, and GPT-3 for extractive and abstractive summarization, and evaluated via ROUGE scores.

Asteroid Diameter Prediction | APIs, Regression, Python, Visualization

01/2023 - 05/2023   Location

Utilized Ridge, Lasso, Random Forest, XGBoost, and MLP Regressor to forecast asteroid diameters, targeting potential hazards with a precision of RMSE 0.164.

SKILLS

`Python` `R` `SQL`

`Adobe Premiere Pro` `Apache Airflow`

`Data Structure and Algorithms` `Apache` `API`

`AWS` `Azure` `Cloudera` `Computer Vision`

`Cloud Computing` `D3.js` `Data Quality`

`Data Science` `Database management` `Docker`

`Dplyr` `EC2` `EDA` `ETL pipelines`

`Ggplot2` `Git` `Jupyter` `Machine Learning`

`MatplotLib` `MongoDB` `MySQL`

`Natural Language Processing` `Neural Networks`

`NLP` `Numpy` `Oracle` `Pandas` `Plotly`

`Power BI` `PySpark` `MS Excel`