

Ravi Ranjan

• helloravi@live.com • [856-214-5123](tel:856-214-5123) • [Linkedin \(linkedin.com/in/rvranjan\)](https://www.linkedin.com/in/rvranjan)

Data Engineer and tech lead with over 9 years of experience dedicated to maximizing the performance of softwares and researching a highly-effective data solution. Proven track record of architecture & development of large scale software solutions and driving customer satisfaction. Skilled in providing technical mentorship resulting in improved code quality, system reliability, and team productivity.

SKILLS

Programming Languages : Python, Java, C++, C#, R, Scala, PHP, & UNIX shell.

Big Data & ML : Hadoop, SSIS, Airflow, Databricks Delta Lake, PySpark, PyTorch, TensorFlow, Scikit-learn, NumPy, PyTable, Terraform, Alation, Pandas, Hive, Spark & Kafka.

Cloud Technologies : Azure : Databricks, ETL, ML, AI, Data Lake, Google Cloud (GCP) : AutoML, Vertex AI, BigQuery, Snowflake, OpenStack, Amazon Web Service (AWS) : EC2, S3, ELB, EMR, Elastic Beanstalk, Lambda, CloudFront, Redshift, RDS & DMS.

Databases : SQL, NoSQL, Postgresql, SQLite, Oracle, DynamoDb, MongoDB & Neo4Js.

Framework / CMS : Spring Boot, Django, NodeJS, Laravel, Magento, Wordpress, Drupal, Slate CRM, Salesforce, Sugar CRM, Oracle CRM, extra.

Frontend Languages : Html, Css, ReactJs, Bootstrap, JQuery, d3s, Tableau & PowerBI.

Functional Skills / Tools : Analytics, Office, Excel, Graphs, Git, Jira, Docker, Podman, Puppet, Chef, Agile, Scrummaster and Charts.

PROJECTS

• Data pipeline developments • Mathematical modeling using R • ETL projects over cloud • Pyspark cluster optimization • Statistical modeling • Data governance projects • Big data processing using Databricks • Healthcare data analysis • Deep learning models • Stocks & cryptocurrency analysis • Smart devices and IOT for devices • IVR Software development • Finance API Developments • E-Commerce platform development • ML Algorithms on datasets • Web development • Data Analysis & Project management Projects • Contribution on Stackoverflow, git and other open source platforms.

SMART BILLBOARDS WITH VIEWER ANALYSIS

- Design, development & Implementation of python-3 and supported machine learning modules on Raspberry Pi and other hardware devices to build prototype models that can analyze objects.
- Perform object detection, Face & Eyes Detection using OpenCV, Computer Vision and machine learning algorithms. Trained model over 2 Millions images and achieve success rate up to 98%.
- Build complete end to end CICD pipeline over AWS & databricks for live data analysis and frontend reporting. Use AWS Kinesis Data Streams to capture the live data feeds.

APPLICANTS DATA ANALYSIS, FORECASTING & CLASSIFICATIONS

- Use tensorflow, SciPy, Keras and DL libraries to analyze the admission data for college applicants. Create an algorithm which classifies new applicants into their potential categories.
- Use time series forecasting for future applicants and achieve up to 20% of cost reduction in resource allocations while following agile practices and scrum framework.
- Trained Machine learning model over a variety of data sets, that's helped identify the top 30% potential candidate and give a chance to focus more on them.

EXPERIENCE

Rowan University (Planning & Management Dept.)

Oct-2021 - May-2024

Data specialist (AI & ML)

Glassboro, NJ

- Created a data pipeline using Hadoop ecosystem using HDFS to read files, MapReduce to process the data and HBase to store the data. Use Pig and Hive to query the data for analytics.
- Define and deliver the implementation of machine learning algorithms on applicants and create a Supervised Learning model results in to manage workload in 25% less admission counselors.
- AWS cloud architecture, frontend, backend and database queries. Extensive use of MySQL and spreadsheet along with tableau to create charts, included in main website stats.
- Develop a detailed project plan, including tasks, timelines, milestones, and resource allocation and make sure that project adheres to agile best practices, and quality guidelines.
- Provide investigational support for identifying process optimizations for cooper medical and organize the data into databases, data warehouses, and data lakes. End results of the data pipeline in Spark and Snowflake were 30% faster and cleaner than conventional methods.

Pemex- Crypto Exchange

Mar-2022 - Aug-2022

Research Scholar

NewYork, NY

- Research and publish a blog on "[Retailers Planning To Accept Crypto Payments](#)" and paper on "[Effectiveness Of Machine Learning Algorithms In Cryptocurrency Price Forecasting](#)".
- Worked on various machine learning algorithms like LSTM and KNN to check the effectiveness of stocks and crypto pricing. Presents findings to stakeholders and brokers.
- Worked on multiple big data technologies, developed and consumed stock data API.
- Provide technical support to process investigations, also work on data science libraries LSI distributed models, gensim's TF-IDF and gensim's similarity index.

Virtual Employee Pvt. Ltd.**Jan-2019 - May-2021***Lead developer (Data Engineering)**Amersfoort, Netherlands*

- Completed entire project lifecycle including Data Validation, Manipulation, Validation, Mining, Algorithm and Visualization using multiple approaches to choose the best.
- Worked with Snowflake utilities, SnowSQL, SnowPipe, Big Data model techniques using Python. Also implemented Spark using Scala and utilizing Data frames and Spark SQL API for faster processing of data. Perform SQL Server Integration Services (SSIS) is a Microsoft SQL.
- Develop expertise in migrating large-scale, complex datasets to AWS using services such as AWS Database Migration Service , AWS Glue, AWS Lambda and storing data to DynamoDB.
- Developed, tested and debugged several ETL pipelines that resulted in increased loading yields by 14%. Optimized data warehouse solutions for efficient data consolidation, reducing processing time by 78%.

Igniters Hub Pvt. Ltd.**Feb-2017- Jan-2019***Lead Engineer (Data Engineering)**Noida, India*

- Use Tensorflow and other deep learning models for object detection and eye sight counts. Use over millions images to classify and achieve model accuracy up to 92%.
- Involved in design, develop and support automated data pipelines using Sqoop, Pig, Spark and Hive and store the data on HDFS and S3. Also provided support for real-time multivariate monitoring models using statistical software like SAS and SPSS.
- Created Visual report, Graphical analysis and Dashboard reports using Tableau. Apply new research on visual analytics to complete the innovation challenge.

Tekege Solutions Pvt. Ltd.**Apr-2015 - Feb-2017***Software Developer**New Delhi, India*

- Develop and maintain interactive Voice Response (IVR) software written over Java and convert plain Java codes to Spring Boot Framework results in significant improvement in speed.
- Develop large scale enterprise applications on Java, Spark and similar technologies, particularly focused on data structure and algorithm, complete projects in 15% less time than estimated.
- Worked on SugarCRM module development and integration with hardware Ericsson cellular devices. Which ease calling directly from the app and increase in sales over 10%.
- Develop module for API integration with AWS & Azure Platform which helped clients like WHO and Averdi to get better insight of their data and broadcast it using Kafka.

EDUCATION**Masters of Science in Data Science**

Rowan University • Glassboro, NJ • May 2023

Bachelor of Engineering (EEE)

RGPV University • MP, India • 2014

Certifications (Online & Offline)

• Databricks • Azure • AWS • Python • Data science • SQL • AI • Tableau • Scrum • PMP

Projects & Research Papers

• 2 Publish Research papers • 12+ Software development projects • 9+ Data engineering projects