Roopak TJ | Data Engineer

2/46 Belsize Avenue, Carnegie – VIC-3163

☐ +61-497522844   •   ✉ mailstorpk@gmail.com
https://www.linkedin.com/in/roopak-tj-0697179b/ | https://github.com/roopaktj

Data Engineer with 2+ years of experience in building data intensive pipelines, tackling challenging architectural and scalability problems in Healthcare industry. Data Science Graduate from Monash University, Australia and AWS Certified Cloud Practitioner

Education

Academic Qualifications

Master of Data Science, Caulfield campus   **Monash University**
Post Gradguate Degree , CGPA: 3.25/4   *2018-2020*

Amrita School of Engineering, Amritapuri campus   **Amrita University**
Bachelor of Technology, Computer Science and Engineering , CGPA: 8.6   *2012-2016*

Kendriya Vidyalaya Bhusawal, Maharashtra   **Higher Secondary Education**
, Cental Board of Secondary Education, 12th Grade   *2011*
84.6 percentage

Work Experience

Cerner HealthCare Solutions   **Banglore**
Software Engineer   *July 2016–July 2018*

- Worked as a Software Engineer in Revenue Cycle Analytics Team.
- Responsible for data analysis, design fabrication and end to end Apache Crunch Pipeline implementation for Revenue Cycle Analytics Project.
- Validation of client production data using CCL(Cerner Command Line), SQL and R scripts.
- Worked on end to end implementation and enhancements in Claim, Cost Reimbursement and Balance AR Apache crunch pipelines
- Conducted knowledge transfer sessions to peers on Java and healthcare functional topics

Cerner HealthCare Solutions   **Banglore**
Seasonal Intern   *January 2016- July 2016*

- Implementation of ETL jobs to load data from a SQL database, normalise it and populate in a noSQL HBase database using MapReduce jobs
- Testing and debugging MapReduce programs using MRunit tests
- Knowledge of Apache Hadoop, MapReduce, Hbase, SQL, Git, Linux commands and Healthcare terminologies.

Technical and Personal skills

- **Programming Languages:** R, Python, Java

- **Big Data skills** Map Reduce, Hadoop, Apache Crunch, Apache Spark, Vertica DB, Apache Kafka, Mongo

DB

- **Data Visualisation skills** R, Python, Tableau and D3
- **Office Suite skills** PowerPoint, Word, Data analysis and Visualization using Excel
- **Platforms :** GNU/Linux (Ubuntu), Windows, Mac OS.
- **Version Control System :** Git
- **Cloud Technology :** AWS Cloud Practitioner Certification
- **Platform/IDE :** Eclipse/Intellije IDE for Java projects, Sublime Text, Jupyter notebook/Pycharm for Python,PostgreSQL
- **Industry Software Skills :** Issue tracking products : Jira and Bugzilla.
  Collabrative code review software : Crucible
  Continuous Integration and development software: Jenkins
- **General Business Skills :** Good presentation skills, Agile team working experience.

Most Proud Of

- **Industrial Experience Expo Winner 2020 Semester 1**

  JUNE 2020: Winner of Monash Industrial Experience Expo implementing a Full stack web application to reduce injuries in Rural Australia
- **AWS Certified Cloud Practitioner**

  FEB 2020: Cleared AWS Cloud Practitioner Certification which effectively demonstrates an overall understanding of AWS Cloud Technology
- **Cerner NOTT Award** *'Award for Excellence in Cerner'*

  JULY 2016: Awarded with NOTT for the exceptional work on the end to end implementation of an Apache Cruch pipeline as per Client requirement.
- **Article published in Magazine** *Open Source For You*
  JULY 2014: Article on "Swapping Space for Linux" got published in the "Open Source for You" Magazine.
- **Open Source Contributions** *Mediawiki and LibreOffice*
  JANUARY 2014 onwards: Contributed and fixed bugs in a couple of Open source Organizations like Mediawiki and LibreOffice.
- **ACM- ICPC** *Programming competition*
  DECEMBER 2013 onwards: Team was selected for the regional level of ACM ICPC.

Projects

- **Implemented a web application with MEVN tech stack:** Implemented a Full Stack web application to reduce Injuries in Rural Australia. Technology stack used include Vue (Front-end), Express-Node(backend) and MongoDB(database). Cloud deployed MongoDB using Atlas was used as Database. The application also included Data Wrangling, Storing and Visualisation using Open Dataasets including Google Javascript API for Maps and Places.
  Github : https://github.com/RoopakTJ/Eu4ia-vue-app
- **Real Time data Visualisation using Apache Spark, Kafka and Matplotlib:** Implemented an application which emits the data stored in a csv file through a kafka producer, receives it using Spark Consumer and does real time analysis on the data using Visualisations with Matplotlib package and Data Storage using MongoDB.
  Github : https://github.com/RoopakTJ/DistributedDB
- **Sentiment Analysis Tool with Twitter Data using Python:** Implemented a Sentiment analysis tool

with Twitter Data using Python programming language. The script uses packages like tweepy to extract tweets, textBlob for identifying the polarity of tweets and matplotlib for graphical representation of the public sentiment using Pie-Chart and WordCloud.

- **Apache Crunch Pipeline Implementation:** Worked on end to end implementation of Apache Crunch pipelines which generated reports and KPI's for the end user. Project also involved validation of client production data using CCL(Cerner Command Line) and SQL scripts
- **Cerner Internship Project**: Implementation of ETL jobs to load data from a SQL database, normalise it and populate in a noSQL HBase database using MapReduce jobs. Also involved testing and debugging MapReduce programs using MRunit tests Knowledge of Apache Hadoop, MapReduce, Hbase, SQL, Git, Linux commands and Healthcare terminologies