**ROOPAK TJ**

Data Engineer

16 Murdo Road Clayton – VIC-3168

Phone : +61497 522 844 Email : [mailstorpk@gmail.com](mailto:mailstorpk@gmail.com)

LinkedIn : https://www.linkedin.com/in/roopak-tj/ Github : 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**

**Master of Data Science**  July 2018 - June 2020

Monash University, Caulfield Campus

CGPA: 2.8/4

**Bachelor of Technology (Computer Science and Engineering).**  July 2012 - June 2016

Amrita School of Engineering

CGPA: 8.6/10

**Higher Secondary Education** March 2011

Kendriya Vidyalaya Bhusawal, Maharashtra, India

Percentage 84.4

**Work Experience**

**Dental One** **Melbourne**

*Data Engineer/SQL Developer July 2020 - Present*

* Lead Data Engineer/SQL Developer in the IT department at Dental One
* Devised and executed SQL scripts for report generation in native application and website
* Optimized SQL Query for improving performance using Tuning and Execution plans; reducing execution time by 60%
* Oversaw periodic transfer of historic data to cloud deployed Data-warehouse
* Took ownership for generating Tableau dashboard visualisations to enhance visibility of Business performance in terms of KPIs

**Cerner Healthcare Solutions. Banglore** *Software Engineer July 2016 -July 2018*

* Worked as a Software Engineer in Revenue Cycle Analytics Team
* Performed data analysis, design fabrication and end to end Apache Crunch(Framework for Hadoop Mapreduce in Java) Pipeline implementation for Revenue Cycle Analytics Project
* Validated client production data using CCL(Cerner Command Line), SQL and R scripts
* Led 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
* Volunteered and Mentored as a peer facilitator for group of 10 newly joined associates.

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

* Accomplished ETL jobs to load data from a SQL database, normalise it and populate in a noSQL HBase database utilising MapReduce jobs
* Performed testing and debugging on MapReduce programs employing MRunit tests
* Upskilled knowledge on Apache Hadoop, MapReduce, Hbase, SQL, Git, Linux commands and Healthcare domain

**Technical and Personal skills**

* Programming Languages: R, Python, Java
* Big Data skills Map Reduce, Hadoop, Apache Crunch, Apache Spark (Pyspark) for both Streaming and Processing, 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.
* Collaborative 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 accomplishing 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 Crunch 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 LibreO**ﬃ**cep**
* JANUARY 2014 onwards: Contributed and fixed bugs in a couple of Open source Organisations like Mediawiki and LibreOﬃce
* **ACM- ICPC Programming competition**
* DECEMBER 2013 onwards: Team was selected for the regional level of ACM ICPC

**Projects**

* **Fabricated a web application with MEVN technology stack:** Developed a Full Stack web application to

reduce Injuries in Rural Australia. Technology stack included Vue (Front-end), Express-Node(backend) and MongoDB(database). Cloud deployed MongoDB using Atlas was assigned as Database. The application also included Data Wrangling, Storing and Visualisation using Open Datasets including Google Javascript API for Maps and Places

Github : <https://github.com/RoopakTJ/Eu4ia-vue-app>

* **Real Time data Visualisation utilising Apache Spark, Kafka and Matplotlib:** Developed an application to emit data stored in a csv file through a kafka producer, receives it with Spark Consumer, performs transformations in Spark RDD and does real time analysis on data deploying Visualisations with Matplotlib package and Data Storage using MongoDB

Github : <https://github.com/RoopakTJ/DistributedDB>

* **Sentiment Analysis Tool with Twitter Data using Python:** Created a Sentiment analysis tool with Twitter Data. The script uses packages including Tweepy to extract tweets, textBlob for identifying the polarity of tweets and Matplotlib for graphical representation of public sentiment through Pie Chart and WordCloud
* **Apache Crunch Pipeline Implementation:** Partnered on end to end implementation of Apache Crunch pipelines in order to generate reports and KPIs for end users. Project also involved validation of client production data availing 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 with MapReduce jobs. Also involved testing and debugging MapReduce programs employing MRunit tests Knowledge of Apache Hadoop, MapReduce, Hbase, SQL, Git, Linux commands and Healthcare terminologies