Santosh Kumaravel Sundaravadivelu

+61 431933607

santoshisjerry@gmail.com

https://www.linkedin.com/in/santoshkumaravel/

https://github.com/santoshrepo

OBJECTIVE

A Software engineer turned Data enthusiast. Aiming to use my knowledge and expertise in Data-driven Processing, Machine Learning, Artificial Intelligence for creating actionable solutions and strategies with a proven record of prioritizing and completing tasks in a timely manner when required in building the business and make it a place where I can learn more and showcase my skills.

TECHNICAL SKILLS

Programming Language: Python, R, Java, C/C++

Cloud Technologies: AWS S3, AWS EC2, AWS EMR, GCP

Big Data technologies: Hadoop

Database: MySQL, Microsoft SQL

Machine learning Tools/libraries: Numpy, Scikit-learn, Pandas, NLTK, Gensim, SciPy, Dask, Tweepy

Deep Learning Frameworks: Tensorflow, Word2Vec, Doc2Vec

Project Management Tools: JIRA, Confluence

Data Visualization Tools: PowerBI, Tableau, Google Charts

Data Visualization Libraries: Matplotlib, Seaborn, ggplot, Plotly, leaflets, R shiny, shiny dashboards

Application Server: JBoss, Websphere

Other: Maven, SSMS, Excel, SVN, Github, ETL Methodology, Agile Methodology

EMPLOYMENT HISTORY

1. Voop.Global, Data Scientist Intern February 2020 – June 2020

- Mined the data from different sources such as Twitter, Google trends, RSS Feeds, Wikipedia using various libraries and API like **Tweepy API, Wikipedia API**; built the knowledge base with the mined data using **Python**.
- Using the **RSS Feeds**, Covid19 Voop Dictionaries was created. For instance, Unemployment, Social connection are examples of Covid19 Dictionary.
- In the GCP **environment**, the **BERT** Deep learning model was trained on the mined data to determine the background.
- Extensive work to build the documents such as **SRS**(**software requirements specification**) and rollback plan.
- The dashboard was created in **PowerBI using DAX** for visualization(https://santoshrepo.github.io/Voop_project.html).

2. RMIT University, Research Assistant September 2019 – March 2020

- Used **pdfminer** library in **Python** to extract the contents in the PDF.
- Data extraction, manipulation, and analysis of unstructured data.
- **Topic modeling** was performed to understand the topics discussed by the local council in Victoria.
- **Sentimental analysis** was done over the topics to segregate the positive and negative sentiments.
- **Doc2Vec Algorithm** was built to understand the document from the local councils.

3. Newgen Software Technologies, Software Engineer June 2017 - June 2018

- Identified patterns in the data and translated them into insights on integration issues and vulnerabilities with improved solutions using Agile **Methodology**, Data extraction, manipulation, and analysis using **MySQL**.
- Setting up the **dev environment** and **deployment** in the production environment using **Java** and **Python**.
- Extensively worked on **extraction**, **loading data**, **transformation**, from various sources like **Microsoft SQL Server** and **MySQL.**
- Interacted with the **Stakeholders** for better implementation and weekly progress monitored using **JIRA** and **Confluence**.
- Versed in the complete Software life cycle from preliminary needs analysis to enterprise-wide deployment and support.

4. LTI - Larsen & Toubro Infotech, Software Engineering Intern December 2016 – February 2017

- Implemented graphical interface to manage VDIs(BOTS), which makes it much easier than the existing system
- Developed the interface using **HTML, CSS, Javascript**, and deployed in the **dev** environment.
- Analyzed insights and reported, and documented.

EDUCATION

Master of Data Science, RMIT University — Melbourne, Australia
Grade: 75% — July 2018 – July 2020

Bachelor's in Information Technology, Sri Sai Ram Engineering College — Chennai, India
Grade: 74% — June 2013 - June 2017

PROJECTS

Real-Time ROI for Sugarcane Land Investors(Melbourne Datathon 2019 Project)

- Extracted the **satellite images** from Sentinel 2A using API.
- Detected and masked the clouds with a different color using **pixel intensities**.
- Extracted the area by masking the potential sugarcane area from the TCI satellite images using the **CNN model** and calculated the ROI using the **RNN model**.

Types of accidents in Victoria Neighbourhood

- The application was built in **R shiny**.
- **HTML** was also integrated with the shiny app to build a better visualization.
- **Leaflets** library was used to visualize the maps and find the correlation between the accidents and house types.
- Different types of graphs were created to contribute to the analysis of the accidents. For instance, a Stacked bar chart was plotted to understand the number of accidents in the Suburbs.

Performance Comparison of Map-Reduce Algorithm

- Analyzed two big datasets Common Crawl and Amazon Review from **AWS S3** bucket.
- Established a connection with the jump host and accessed the **AWS EMR** Master.
- Analyzed the Performance by boosting the clusters, changing the data sizes among the datasets, changing the number of mappers and reducers.

Analyzing and Tracking the sentiment and topics on Social Media

- Connected with twitter was established by **Tweepy API** and fetched around 10k tweets over one week.
- The tweets were then cleaned and then normalized and tokenized to find the sentiment of the words/sentences.
- The trend of the Oneplus was found by the hashtags which were visualized to find the interesting findings.
- Implemented **Topic modeling to** find the relevant topics discussed in the tweets.
- Performed **sentiment analysis** on the tweets using the **Vader** sentiment.

AWARDS AND HONOUR

Pat on the Back Award, Newgen Software — Chennai, India
May 2018

Best Performer Award for the Security Bank, Newgen Software — Chennai, India
February 2018

EXTRA-CURRICULAR ACTIVITIES

Student Ambassador, RMIT University — Melbourne, Australia
June 2019 – July 2020

Event Volunteer, RMIT Student Union — Melbourne, Australia
October 2019

CERTIFICATES

Data Literacy, RMIT University — Melbourne, Australia
April 2020

Academic Integrity Awareness, RMIT University — Melbourne, Australia
April 2020

References

1. Dr. Mohammad Saiedur Rahaman, Research from RMIT University
   saiedur.rahaman@rmit.edu.au | +61 416800303
2. Kristin O'Brien, CEO and Founder Voop.global
   Kristin@voop.global |+61 467772881