**Email** : rnagara1@mtu.edu **Rajath Nagaraj Contact no**: **(906)-370-7098**

**LinkedIn**: <https://www.linkedin.com/in/rajath-nagaraj/> **Portfolio**: <https://rajat1995.app>

Data Science graduate student with 2+yrs of experience in field of analytics, data engineering and python development. My focus area is on Machine Learning, Cloud Services and Software development. I have been certified with AWS cloud and Tableau. I am interested in solving business problems using data driven approach.

**EDUCATION**

**MS in Data Science** 3.86/4.0 | Michigan Technological University | Houghton, MI **| Expected April 2021.**

**Bachelor of Engineering** 3.5/4.0 **|** Visvesvaraya Technological University | Bangalore, India **| 2013-2017.**

**Work Experience:**

**State Auto Insurance Companies : Digital Lab IT intern Sep 2020 – Dec 2020.**

* Developed a real time data ingestion framework using Python which can ingest data simultaneously into Apache SOLR, AWS Dynamo DB and S3 with the help of AWS step function.
* Application development using Python OOP for quoting home and auto insurance. Implemented the code on amazon workspace and utilized AWS services like Dynamo DB, S3, Lambda and state machine.
* To expedite the process of generating the insurance quote I researched about implementing a Machine Learning model on my own interest.

**Infosys Limited : Systems Engineer Jan 2018 – Jul 2019.**

* Development of Python based tools for data quality check from on-premise Hadoop servers to AWS and thereby significantly reducing time for validation to 1/3rd of usual time.
* Utilizing AWS EMR service to access the large data from s3 and create datasets to support analytics team. Additionally, worked with Dynamo DB extracting the JSON and parsing them to CSV.
* Analyzing the usability of data by performing Data Profiling using python and spark data frames.
* Developed a tool to review the data migrated to snowflake cloud and the tool achieved a speed to validate 100k records under 4mins.
* To showcase the day to day progress to the client by creating reports using Tableau and Microsoft excel. Also, took the leadership to present the demos and progress in absence of tech lead.

**Academic Projects -** <https://github.com/Rajath1995> | [https://medium.com/@rnagara1](https://medium.com/%40rnagara1) |

**Data Science and Machine Learning:**

* End to end implementation of sentimental analysis model on AWS SageMaker for IMDB dataset with the help of PyTorch, NLP and services like Lambda and API gateway as a result a publicly available end point was generated.
* Built a web application with help of a Machine Learning pipeline and NLP which can segregate the disaster response messages during natural calamities as a result a user-friendly UI page which handles the messages and Machine Learning model output.
* Developing a customer segmentation for Arvato Financial data using unsupervised learning (k-means) and identifying which population base would respond for the mail advertisements. The later part involves converting the model into supervised problem and create the leads for marketing campaign.
* Programmed a classifier using TensorFlow Keras to classify images and further increased the accuracy to 90%+ with help of image augmentation and transfer learning with MobileNet and VGG16.
* Implemented Generative adversarial network on MNIST to achieve fewer imperfections in the image.

**Data Mining:** <https://census5.herokuapp.com/>

* Analyzing insights, association analysis, and patterns from various datasets ex: Airbnb Seattle, Census, Music, and IRIS. Implemented core data mining algorithms like Random Forest, SVM, Decision Trees, clustering techniques, association rule and PCA.

**Web Development and Data Visualization Projects:** <https://public.tableau.com/profile/rajath.nag.nagaraj>

* Created visually attractive and simple to understand visualizations for Makeover Monday competitions.
* Developed my personal website using HTML&CSS and launched it over google (<https://rajat1995.app>).
* Created various web applications for ML projects using Stream Lite, Plotly and HTML web pages.

**Technical and Computer Skills**

* **Programming Languages and Tools –** Python, R-prog, HTML, C; Jupyter notebook and visual studio
* **Big Data Technologies** – Hadoop, Hive, Sqoop, Apache Solr – basic knowledge.
* **Cloud Platforms** – AWS(DynamoDB, Athena, S3, EMR, SageMaker), Google cloud (entry level skill).
* **Data Warehousing** – Oracle DB, Postgre and DB2 – intermediate. Workflow mgmt.– Apache Airflow.
* **Data Visualization –** Tableau, Flourish Studio, Matplotlib, Seaborn, Plotly and Tableau.
* **Data Preprocessing** – Informatica, Pandas, NumPy, Beautiful Soup, NLTK and Twitter API Packages.
* **Machine Learning –** Supervised, Unsupervised, SCIKIT Learn, Neural Network -TensorFlow and Keras.