|  |  |  |
| --- | --- | --- |
| Website:rajath1995.app | **Rajath Nagaraj**  rnagara1@mtu.edu  +1-906-370-7098 | 1911 Apt B Woodmar Dr  Houghton, MI, 49931 |

**SUMMARY**

Data science graduate student with 1.7 years of full-time experience in field of data services, specializing in field of machine learning. I have also served as student coordinator at Akuva Infotech. I love writing blogs on Machine Learning and datamining(<https://medium.com/@rnagara1>). I am constantly getting certified for various technologies from Udacity and Udemy. I am Interested in solving business problems using data driven approach.

**EDUCATION**

Michigan Technological University | Houghton, MI | **MS Data Science** **GPA 3.8/4.0 | Expected Dec 2020**

Visvesvaraya Technological University | Belgaum, Karnataka, India | **Mechanical Engineering**. **GPA 3.5/4.0**

**WORK EXPERIENCE**

**Infosys Limited | Mysore, India | Designation: Systems Engineer | Period - Jan 18 – July19**

**Project: American Family Insurance** - Cloud Data Lake on Amazon web services (AWS).

* Development of Python based tools for data quality check from on-premise Hadoop servers to cloud AWS and thereby reducing the time for validation to 1/3rd of usual time. The tool had a record time of 4mins to valuate 1million data.
* Closely working with onsite team, solving the issues on data type changes and timestamp related issues. Creating spark data frames in order to analyze differences in data like performing data profiling.
* Working with AWS EMR service to access the large data sets from cloud, creating tables from S3 buckets in order to support data analytics team. Also, I worked on Dynamo DB extracting the complex JSON and parsing them to CSV, also as part of project creating table and loading data to hive.
* Developed a tool to review the data migrated to snowflake cloud, the tool resulted in time reduction for the investigation and it could compare 1 million data under 4mins.
* Creating project report on day to day basis using SAS, Microsoft Excel and present the same to the onsite lead.

**Technical and Computer Skills**

* **Machine Learning** – Linear Regression, Multiple Regression, Polynomial Regression, Support Vector Machine, Decision Trees, Random Forest, KNN, Naïve Bayes, K- Means, FCM, GKFCM, NLP Genetic Algorithm, Particle Swarm Optimization, Scikit Learn, NLTK, Neural Network, TensorFlow and Keras.
* **Programming Languages and Tools –** Python, R-Prog, SAS (Basics),HTML,CSS, C; Informatica.
* **Big Data Technologies** – Hadoop, Hive, Sqoop.
* **Cloud Platforms** – AWS (Dynamo DB, Athena, S3, EMR), Google cloud.
* **Data Warehousing** – Oracle DB, MYSQL, Postgre and DB2.
* **Data Visualization –** Matplotlib, Seaborn, Ggplot, Tableau. Microsoft Office.
* **Data Preprocessing** – Informatica, Pandas, NumPy, Beautiful Soup, Twitter API Packages.

**Academic projects on Data Science and Machine learning:**

* Recommender system for IBM Watson studio articles using User based collaborative approach.
* Analyzing insights of Airbnb Seattle data using crisp data mining approach.
* Web Development – data dashboard using Flask, HTML Bootstrap, Plotly and Pandas.
* Analyzing US economic data and building a dashboard - Utilizing the data from AWS s3 bucket, using the data visualization package matplotlib, data processing module pandas and NumPy.
* Categorizing of hotel reviews as positive and negative using Natural Language Processing with Naïve Bayes as the prediction algorithm.
* Machine Learning pipeline on disaster management messages using NLP and Multioutput Classifier.
* Application of artificial neural network on bank customer exit prediction using Keras package.
* Image classification using TensorFlow Keras and implementation of image augmentation.

**Certifications:**

* Udacity certified Data Scientist and Data Visualization Nanodegree from Udacity.
* Tableau Desktop Specialist.
* Amazon Web Services Cloud Practitioner Certification.