

# Tharun Kumar Reddy Karasani

Masters in Data Science | University of Washington - Seattle | Ex-Goldman Sachs

## Profile Summary

**Objective -** Data Science Graduate student seeking a summer internship in Analytics, ML and DL domains

**Summary of Qualifications -**

- 2+ years experience in Data domain with focus on Data Engineering, BI and Analytics
- Problem solver and Data Savvy with a strong understanding of data's impact on business
- Practical experience with various ML, DL Algos, ETL, Visualization tools and large data systems
- Ability to interact with peers and stakeholders to drive business impact
- Quick and enthusiastic learner with high adaptability to changing work environments
- Organized multi-tasker with great focus on the effective outcome of the business problem
- Proponent of diversity, inclusion at workplace and can work independently and on team projects

## Education

Sep 2021 - Aug 2023	<b>University of Washington - Seattle</b> Master of Science in Data Science <b>Relevant UW coursework -</b> Applied Statistics, Statistical Machine Learning, Scalable Data Systems & Algorithms <b>Additional coursework -</b> Coursera: Deep Learning Specialization, Deep Learning with CV and NLP Udemy: Machine Learning A to Z, Deep Learning and AI
Jun 2015 - Apr 2019	<b>Vellore Institute of Technology (VIT) - Vellore, Tamil Nadu</b> Bachelor of Technology in Electronics and Communication Engineering Program Representative of the department (PR). <b>CGPA: 9.63 / 10.00 (US 4.0)</b>

## Work and Research Experience

Jan 2019 - Jul 2021	<b>BI Specialist - Advanced Analytics</b> <i>Goldman Sachs - Bengaluru, India</i> <ul style="list-style-type: none"><li>• Performed advanced analytics and Data science techniques on top of the BI layer to generate business insights for operational efficiency</li><li>• Achieved <b>70% reduction</b> in Trade Fail rate by deploying ensemble and pattern mining <b>ML algorithms</b> that exposed the trade flow bottlenecks</li><li>• Productionized unsupervised models that <b>reduced manual intervention by 30%</b> after identifying the manual touch point patterns in trade settlement workflow</li><li>• <b>Improved processing speed by 50%</b> for scheduled ETL jobs by pipelining data through <b>PySPARK refiners</b> and leveraging OLAP models</li><li>• Optimized the data engineering framework for multiple trade workflow systems by developing firmwide data source <b>integration APIs</b></li><li>• Automated and fastened the <b>ETL jobs SDLC process</b> by integrating it with <b>GitLab</b></li><li>• Enhanced the code fail resolution mechanism for benchmark and threshold breaches by leveraging <b>JIRA API</b> in the control framework</li><li>• Enabled Data-driven decision making by publicizing powerful <b>Tableau</b> dashboards on firm servers saving vendor server charges</li><li>• Partnered and cross trained with 4 global teams to understand their business requirements and workflow process</li></ul>
Jul 2018 - Sep 2018	<b>Machine Learning and Big Data Researcher</b> <i>VIT - Vellore, India</i> <ul style="list-style-type: none"><li>• Analyzed in real-time the usage of underground water in a village using <b>Machine Learning</b> and basic <b>Hadoop</b> ecosystem components</li><li>• Deployed a system of sensors, and coded microcomputers to capture and transmit the real-time field data to the executor nodes</li><li>• Monitored an average excess usage reduction of <b>200 gallons</b> per family post deployment</li><li>• Leveraged <b>regression models</b> to predict the water resource availability based on the current consumption metrics</li><li>• Automated the extraction, processing pipeline and storage using PySPARK and HDFS</li><li>• Productionized the <b>Tableau</b> dashboard to monitor the availability and usage of water resources</li></ul>
Apr 2017 - May 2017	<b>Embedded System Design Intern</b> <i>Electronics Corporation of India Limited (ECIL) - Hyderabad, India</i> <ul style="list-style-type: none"><li>• Developed a vehicle to vehicle communication (V2V) system for prototype vehicles</li><li>• Improved model efficiency by deploying smart IOT devices and connectivity over cloud</li><li>• Worked with different micro controllers and their interfaces with respective sensor modules</li></ul>

## Personal Info

<b>Address</b> 5021 Brooklyn Ave NE, Apt 40, Seattle, Washington, 98105
<b>Phone</b> +1 (206) 532-9924
<b>E-mail</b> karasani.tarunreddy@gmail.com
<b>LinkedIn</b> https://www.linkedin.com/in/tharun-kumar-reddy-karasani-b94332114/
<b>My Portfolio</b> https://karasanitarunreddy.wixsite.com/my-site
<b>GitHub</b> https://github.com/TharunKumarReddy5

## Framework Expertise

Data Engineering	<div><div></div><div></div><div></div><div></div><div></div></div>
Prescriptive Analytics	<div><div></div><div></div><div></div><div></div><div></div></div>
Project Management	<div><div></div><div></div><div></div><div></div><div></div></div>
Machine and Deep Learning	<div><div></div><div></div><div></div><div></div><div></div></div>
Data Visualization	<div><div></div><div></div><div></div><div></div><div></div></div>
Big Data Eco-System	<div><div></div><div></div><div></div><div></div><div></div></div>
API	<div><div></div><div></div><div></div><div></div><div></div></div>

## Languages

Python	<div><div></div><div></div><div></div><div></div><div></div></div>
SQL	<div><div></div><div></div><div></div><div></div><div></div></div>
C , C++	<div><div></div><div></div><div></div><div></div><div></div></div>
R	<div><div></div><div></div><div></div><div></div><div></div></div>
Java, PySpark	<div><div></div><div></div><div></div><div></div><div></div></div>
Scala	<div><div></div><div></div><div></div><div></div><div></div></div>

## Certifications

- ★ Alteryx Designer Advanced and Core
- ★ Machine Learning
- ★ Tableau specialist - Data school
- ★ SQL and Data Warehousing

## Project Experience

Jun 2021	<b>COVID-19 Infection dynamics and Vaccine Distribution</b> Regression Analysis, Regularization, FIR, ANN - <b>Research</b> <ul style="list-style-type: none"><li>Deployed epidemiology ML model to forecast the Coronavirus spread in INDIA and USA with and without vaccination</li><li>Proposed a vaccine distribution algorithm to reach the safety point with less fatalities</li><li>Captured the time varying nature of the virus spread by improving Time-dependent SIR model</li><li>Added vaccination state to the SIR model to understand the urgency in the states of India</li><li>Provided a Dashboard layer to perform post facto and ex ante analysis</li><li>Presented the <u>research paper</u> and published <u>findings</u> at ITNG 2021 conference</li></ul>
Jan 2021	<b>Trade Fail Root cause Analysis</b> Rule Mining, Bagging, Boosting, K-Medoids, Alteryx, Tableau, SQL - <b>Goldman Sachs</b> <ul style="list-style-type: none"><li>Leveraged Ensemble and Pattern mining ML Algorithms to identify trade settlement fail root causes and predict the fail possibility</li><li>Deployed trending clustering algorithms and distance techniques to categorize and prioritize the client data for fail resolution</li><li>Actively interacted with business heads and SMEs to eliminate noise in the predictions by knowing more about the nature of the business</li><li>Developed visualization layer using Tableau to exhibit weekly data patterns to stakeholders to facilitate their operations</li><li>Integrated GitLab to track the status of settlement fail resolution process that helped in resource management</li></ul>
Mar 2020	<b>Trading Metrics - Data Model Creation</b> Python, Alteryx, Tableau, SQL - <b>Goldman Sachs</b> <ul style="list-style-type: none"><li>Built an ETL pipeline that stitches all the trade data top to bottom and created a universal data model for business consumption</li><li>Cross trained with 3 global teams to onboard and integrate multiple system data models</li><li>Interacted with traders and sales teams to understand the client trading behavior and finalize useful data metrics</li><li>Exposed KPIs and trading opportunities over the Tableau Dashboard for daily consumption</li><li>Worked with Slang (GS Native) scripts to integrate with Presto for trade data extraction</li></ul>
Jan 2019	<b>Water Usage Monitoring &amp; Regulation (Real World Solution)</b> Regression, SQL, Hadoop - <b>Academics</b> <ul style="list-style-type: none"><li>Developed a system for monitoring of groundwater availability, usage and forecast in the city of Vellore, INDIA</li><li>An automated system was created to measure and transmit the amount of water consumed</li><li>Used ML regression models to forecast the water availability and visualized it using tableau</li><li>Designed a website for user details and login for water bill payments</li></ul>
Mar 2018	<b>Smart Face recognition and status monitoring</b> OpenCV, Hadoop - <b>Academics</b> <ul style="list-style-type: none"><li>Developed a face recognition system using OpenCV for automatic student attendance</li><li>Launched an Android application for teachers to monitor the status of students in class</li><li>Leveraged BIG DATA applications for data pipelining and processing</li></ul>

## Achievements

Jan 2021	<ul style="list-style-type: none"><li><b>ITNG21 Springer</b> conference Research paper presentation</li></ul>
Aug 2020	<ul style="list-style-type: none"><li>Goldman Sachs <b>COVID-19</b> Datathon Winner</li></ul>
Mar 2018	<ul style="list-style-type: none"><li>The prestigious "<b>Merit Scholarship</b>" winner for three consecutive years.</li><li>Idea presentation at "<b>TARP SET</b>" conference.</li><li>JOHNSON CONTROLS <b>Make-A-Thon Finalist</b>.</li></ul>

## Publications

Feb 2021	<b>Springer</b> - A Comprehensive Analysis of SARS-CoV-2 in India (Co-Author)
Mar 2018	<b>IEEE</b> - Monitoring of Groundwater level and Development of Control Mechanism based on Machine Learning Algorithm (Research Assistant)

## Volunteer and Leadership Activities

- Volunteer at Community Teamwork and **Analytics for good**
- Volunteer and Public Relations Fellow for **Make A Difference**, an NGO
- Program Representative Head of ECE department (PR) **SENSE VIT**
- Core Committee member in **The Electronics Club (TEC)**
- Senior Coordinator at **IEEE MTTS** chapter

## Software

Alteryx, KNIME	<div><div></div><div></div><div></div><div></div><div></div></div>
Tableau	<div><div></div><div></div><div></div><div></div><div></div></div>
IBM SPSS	<div><div></div><div></div><div></div><div></div><div></div></div>
Tensor Flow, Keras	<div><div></div><div></div><div></div><div></div><div></div></div>
R - Studio	<div><div></div><div></div><div></div><div></div><div></div></div>
MATLAB	<div><div></div><div></div><div></div><div></div><div></div></div>
HADOOP	<div><div></div><div></div><div></div><div></div><div></div></div>

## Other Interests

Research paper novelty improvements

Building Quad and Multi rotor copters

Playing Badminton

Social Causes

