

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	University of Washington - Seattle Master of Science in Data Science Relevant UW coursework - Applied Statistics, Statistical Machine Learning, Scalable Data Systems & Algorithms Additional coursework - Coursera: Deep Learning Specialization, Deep Learning with CV and NLP Udemy: Machine Learning A to Z, Deep Learning and AI
Jun 2015 - Apr 2019	Vellore Institute of Technology (VIT) - Vellore, Tamil Nadu Bachelor of Technology in Electronics and Communication Engineering Program Representative of the department (PR). CGPA: 9.63 / 10.00 (US 3.9)

Work and Research Experience

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

Personal Info

Address 5021 Brooklyn Ave NE, Apt 40, Seattle, Washington, 98105
Phone (206) 532-9924
E-mail karasani.tarunreddy@gmail.com
LinkedIn https://www.linkedin.com/in/tharun-kumar-reddy-karasani-b94332114/
My Portfolio https://karasanitarunreddy.wixsite.com/my-site
GitHub 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 2020	<div>COVID-19 Infection dynamics and Vaccine Distribution<ul style="list-style-type: none">Deployed epidemiology ML model to forecast the Coronavirus spread in INDIA and USA with and without vaccinationProposed a vaccine distribution algorithm to reach the safety point with less fatalitiesCaptured the time varying nature of the virus spread by improving Time-dependent SIR modelAdded vaccination state to the SIR model to understand the urgency in the states of IndiaProvided a Dashboard layer to perform post facto and ex ante analysisPresented the <u>research paper</u> and published <u>findings</u> at ITNG 2021 conference</div>
Apr 2020	<div>Trading Metrics - Data Model Creation<ul style="list-style-type: none">Built an ETL pipeline that stitches all the trade data top to bottom and created a universal data model for business consumptionCross trained with 3 global teams to onboard and integrate multiple system data modelsInteracted with traders and sales teams to understand the client trading behavior and finalize useful data metricsExposed KPIs and trading opportunities over the Tableau Dashboard for daily consumptionWorked with Slang (GS Native) scripts to integrate with Presto for trade data extraction</div>
Mar 2019	<div>Trade Fail Root cause Analysis<ul style="list-style-type: none">Leveraged Ensemble and Pattern mining ML Algorithms to identify trade settlement fail root causesDeployed clustering algorithms to sample unbiased data for model to detect generalized patternsActively interacted with business heads and SMEs to eliminate noise in the predictions by knowing more about the nature of the businessDeveloped visualization layer using Tableau to exhibit weekly data patterns to stakeholders to facilitate their operationsIntegrated GitLab to track the status of settlement fail resolution process that helped in resource management</div>
Oct 2018	<div>Water Usage Monitoring & Regulation (Real World Solution)<ul style="list-style-type: none">Developed a system for monitoring of groundwater availability, usage and forecast in the city of Vellore, INDIAAn automated system was created to measure and transmit the amount of water consumedUsed ML models to forecast the water availability and visualized it using tableauDesigned a website for user details and login for water bill payments</div>
Sep 2018	<div>Smart Face recognition and status monitoring<ul style="list-style-type: none">Developed a face recognition system for automatic student classroom attendanceLaunched an Android application for teachers to monitor the status of students in classLeveraged BIG DATA applications for data pipelining and processing</div>
Jul 2018	<div>Target – Web Intelligence<ul style="list-style-type: none">Re-hosted existing company data from MySQL DB into Hadoop HDFSCrawled product data from top competitor company websites using pythonHandled massive amount of data by means of its cluster nodes and also satisfied the scaling needs of the Target business operation</div>

Achievements

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

Publications

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

Volunteer and Leadership Activities

	<ul style="list-style-type: none">Volunteer at Community Teamwork and Analytics for goodVolunteer and Public Relations Fellow for Make A Difference, an NGOProgram Representative Head of ECE department (PR) SENSE VITCore 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