

Inquisitive data scientist passionate about solving real-world problems using Machine Learning and Artificial Intelligence.
Proficient in understanding business requirements and translating data into useful insights by applying data science tools.

SKILLS

Tools and Languages	Python, R, SQL, TensorFlow, PyTorch, Keras, scikit-learn, SciPy, Pandas, NumPy, seaborn, ObsPy, Weka, Jupyter Notebooks, Excel, C/C++, HTML, CSS, MATLAB, Java, RStudio, Anaconda, Git, AWS
Courses	Machine Learning, Natural Language Processing, Data Structures & Algorithms, Data Science and Big Data Analytics, Probability and Statistics, Database Management, Deep Learning, Neural Networks
Technologies	Natural Language Understanding, Information Retrieval, Machine Translation, Predictive and Generative Algorithms, Classification, Sentiment Analysis, Data Mining, Feature Engineering

EDUCATION

MS, Computer Science, University of Colorado Boulder	<i>GPA: 3.9/4.0</i>	Dec 2022
B.Tech, Computer Science & Engineering, SRM Institute of Science and Technology	<i>GPA: 9.0/10.0</i>	Jun 2020

EXPERIENCE

Research Analyst Intern Data Analytics and Modelling Ruder Finn – Cision, BrandWatch, Excel, Python, Sentiment Analysis <ul style="list-style-type: none">Performed market research for over 10 clients using social and traditional media data, with focus on healthcareAnalyzed and investigated data to find insights that form the basis of improved market strategiesImplemented data science techniques like sentiment analysis to investigate public opinion of clients and their productsProcessed big data by creating and maintaining dashboards on enterprise analytics tools such as Cision and Brandwatch	May' 22 — Present <i>New York, NY</i>
Graduate Research Assistant Seismic Event Detection using Machine Learning University of Colorado Boulder – Python, PyTorch, NumPy, pandas, scikit-learn, matplotlib <ul style="list-style-type: none">Extracted, processed, and visualized one year of seismic time-series data from GeoNet using ObsPy and MatplotlibTrained a deep neural network using PyTorch with multiple convolutional and pooling layers, deployed using AWS EC2Improved precision from 86% to 93% by retraining on engineered features, created using Wavelet transformAnalyzed performance to identify the best model based on accuracy, precision, recall, and computational efficiencyPresented a poster at CIRES Rendezvous 2022	Jan' 22 — May' 22 <i>Boulder, CO</i>
Academic Intern Artificial Neural Networks National University of Singapore – Python, R, Weka, scikit-learn, ggplot, matplotlib <ul style="list-style-type: none">Designed a gradient boosting model (XGBoost) model to predict New York City taxi trip durationCreated a random sample of 10000 from 1.4 million initial data points and cleaned the data by removing noise, outliersVisualized the spatial-temporal data in R and Python using ggplot, matplotlib on attributes including speed, month, timePerformed feature engineering to extract features such as haversine distance, speed, total distance, maneuvers made by carCompared performance of models including logistic regression, support vector machine and naive bayes classifier	Jan' 18 <i>Singapore</i>
Academic Intern Big Data Analysis Hewlett Packard Enterprise Education Services – Hadoop, MapReduce <ul style="list-style-type: none">Built a multiple node Hadoop cluster from scratch and performed MapReduce operations	Dec' 17 <i>Singapore</i>

PROJECTS

Speech and Opinion Recognition using an Ensemble Classifier — Python, NLTK, scikit-learn, numpy, matplotlib, Flask <ul style="list-style-type: none">Achieved an accuracy of 94% detecting sentiment of speech by designing and training an ensemble natural language classifierEnsemble comprised of Naive Bayes, Support Vector Machine, Logistic Regression and Decision Trees with dynamic weightsPerformed dimensionality reduction, removed duplicates and stopwords, created vectors using TF-IDF on Amazon ReviewsDeployed the classifier using web app with flask backend and used CMU Sphinx API to convert speech-to-textPublication: T. Priyam, AMJ Muthukumaran and H. Vinayak, "Speech and Opinion Recognition from a Conversation," International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-6, April 2020
Visual Question Answering using Neural Networks — Python, TensorFlow Keras, Feature Extraction, VGG16, Binary-encoding <ul style="list-style-type: none">Cleaned and pre-processed a subset of the VizWiz - VQA dataset with 2000 training and 500 validation imagesTrained deep neural network using Tensorflow Keras to predict an object using an image and a question as inputExtracted features from images using VGG16 pre-trained on imagenet data and from questions using binary encodingAchieved a validation accuracy of 51% using the VQA accuracy metric which is more than the 47% reported by the VQA paper
Online Loan Prediction and Banking Application — Python, SVM, Flask, PostgreSQL, Object-Oriented Design, React <ul style="list-style-type: none">Implemented a Support Vector Machine model to predict loan amount for clients based on personal and financial dataDesigned a react web app to register and process loan applications using Flask and PostgreSQL as database server