

Sanjana Moodbagil Mallikarjuna

sanjanammoodbagil@gmail.com | linkedin.com/in/sanjana-m-moodbagil | sanjanamm98.github.io | New York City

EDUCATION

University of Southern California, Los Angeles, CA **Aug 2021 - May 2023**
Master of Science in Computer Science | GPA 3.72 / 4

BMS College of Engineering, Bangalore, India **Aug 2016 - Jun 2020**
Bachelor of Engineering in Information Science and Engineering | GPA 9.32 / 10

WORK EXPERIENCE

Software Engineer | **Karsun Solutions LLC** | Herndon, VA **Feb 2023 - Present**

- Built a dynamic dashboard to show agile decision and IT KPI insights by seamlessly integrating with multiple data sources using REST APIs
- Used Javascript (App Scripts) and MySQL to collect data from Coda, Couchbase, etc for backend data integration
- Optimized performance of multiple website scanning processes to collect metrics on ~15k websites reducing time taken by 50%
- Enhanced the fault tolerance and reliability of these scan processes with automated respawning
- Built custom visualizations for derived/aggregate metrics in Looker Studio using Javascript d3 library, NodeJS and Google Cloud Storage

Software Engineer Intern | **Karsun Solutions LLC** | Herndon, VA **Jun 2022 - Aug 2022**

- Devised a software to generate PII anonymized synthetic data from simple table, time-series and relational datasets using Synthetic Data Vault
- Automated the generation of metadata (dataset description, table hierarchy, column datatypes) for relational dataset using Python and MySQL
- Built an Exploratory Data Analysis (EDA) tool to interpret and compare the structure, format and statistical properties of the synthetic data
- Worked with AWS S3, Lambda and Dynamo DB to integrate backend with the client application

Risk Data Scientist | **Information Technology Services** | Part-time | **USC**, Los Angeles, CA **Jan 2022 - May 2022**

- Formulated IBM QRadar SIEM risk metrics to detect suspicious events by monitoring network flows and event logs
- Extracted data related to these key metrics into a data lake (Panaseer) and built visualization dashboards
- Designed an ML model pipeline to curate analytics data for training statistical models that forecasted security risks

Data Analyst Intern | **Caliper Business Solutions**, Bangalore, India **Jan 2021 - Jun 2021**

- Created generic transformation pipelines to build training data for ML models using BigQuery (on Holistics)
- Performed route analysis on unsupervised spatial data to understand driver behavior and freight route patterns using DBSCAN model
- Facilitated companies in planning for expected stoppages, potential stoppage hubs, expected destination time, etc.
- Built analytics dashboards on Time Budgeting and Transporter Performance datasets using SQL and Python

Research Intern | **Dalhousie University**, Halifax, Canada **Jun 2019 - Aug 2019**

- Devised a tool to analyze time series dataset of groundwater levels using signal processing and mathematical transformations
- Leveraged this tool to reconstruct the history of groundwater recharges to aquifers based on the nature of hydrographs like water levels, rain fall, temperature etc. Built a Tableau dashboard to show historical trends and forecasted recharge rates
- Programmed a data logging system using Arduino to log real time data using temperature and pressure sensors onto external storage

PROJECTS

Stock Search App | [Project video link](#) **Spring 2022**

- Built a NodeJS proxy server to integrate with Finnhub API for sourcing stock market data
- Hosted angular application backed by a proxy server in GCP. The app enabled users to see stock trends, top news articles and highcharts.
- HTML5 Local Storage is used to wishlist favorite stocks. Implemented similar features with Android Studio using Java as final project.

Analysis of Speech Patterns in Children to Detect Depression **Spring 2020**

- Extracted features from an audio dataset to identify children with depression
- Performed a comparative study between Logistic Regression, Random Forest, SVM-Gaussian and SVM-Linear to identify a supervised model that best analyzes the features. Random Forest performed the best with an accuracy of 94%.
- Experimented using transfer learning approach that increased input feature set but yielded an accuracy of 78% due to small dataset size. This was built using a pre-trained WaveNet feature extracting model with a CNN classifier.

Histopathological Cancer Detection | [Github link](#) **Fall 2019**

- Investigated the potential of ML algorithms to detect metastases in lymph nodes on body CT scan image data
- Trained supervised models such as MobileNet, Xception, and CNN to differentiate cancer cells from normal cells using TensorFlow
- Achieved an accuracy of **90%, 75% and 90%** respectively

SKILLS

- **Languages:** Python, JavaScript, Java, NodeJS, SQL, HTML/CSS
 - **Tools & DBMS:** MongoDB, Tensorflow, Keras, Sklearn, NLTK, Angular, Flask, Couchbase, SQL
 - **Platforms:** Git, AWS (EC2 DynamoDB, Lambda), GCP, Google AppScripts
-

ACHIEVEMENTS & LEADERSHIP

- Chosen for a research internship at Dalhousie University, Canada through Mitacs Scholarship - top 1% of Fall 2020 cohort.
- Organized coding and debugging competitions for Phase Shift, the college technical fest at BMS College of Engineering
- Taught visually impaired children as a volunteer at Samarthanam Trust, Bangalore, India