

SHARVARI KALGUTKAR

skalgutk@usc.edu | <https://github.com/Sharvari289> | <https://www.linkedin.com/in/sharvarikalgutkar/> | (323) 470-3032

EDUCATION

University of Southern California, Los Angeles, California

Aug 2022-May 2024

Candidate for Masters in Applied Data Science for fall 2022.

(CGPA 4/4)

Relevant Courses: DSCI 551: Foundations of Data Management, DSCI 552: Machine Learning for Data Science, CSCI 566: Deep Learning and Its Applications, DSCI 553: Foundations and Applications of Data Mining

Bharatiya Vidya Bhavan's Sardar Patel Institute of Technology, Mumbai, India

Aug 2018-Jun 2022

Candidate for Bachelor of Technology - Electronics and Telecommunication with *Distinction*

(CGPA 9.52/10)

PROFESSIONAL EXPERIENCE

AI Engineer (Intern), Scientist AI Technologies Pvt Ltd

Nov 2021-May 2022

- Designed and implemented road intersection analysis algorithms in Python to calculate five surrogate safety measures like Post Encroachment Time.
- Built an automated road inventory survey system to track road quality using object detection models namely Yolo V5 and Faster R-CNN with an average precision of 94%.
- Conceptualized and developed an algorithm to calculate angle of conflict between vehicles using bounding box annotations and video processing.

Research Intern – Deep Learning and Machine Learning, Skinzy Software Solutions Pvt Ltd

Oct 2020-Jan 2021

- Constructed a Mask-RCNN instance segmentation model in TensorFlow to detect skin abnormalities.
- Designed an algorithm to differentiate skin color from others using histogram segmentation.

PROJECTS

DISCI 551: HappinessQ | Firebase, SQL, Hadoop MapReduce, Flask.

- Built emulated distributed file storage using Firebase and SQL to analyze world happiness, unemployment, GDP.
- Implemented EDfs commands like mkdir, ls, getPartitionLocations, readPartition, cat, etc., and stored metadata about the file system, including file system structure, using Python.
- Created a website using Flask for searching and analyzing the data stored in EDfs using the functions implemented using partition-based map and reduce.

Medi Locker | Natural Language Processing, Named Entity Recognition, spaCy, Mongo DB.

- Created an application to digitize medical records and aid tracking of patient's medical history by providing three services: a Voice Prescription model, medical conversation summarizer and medical records database system.

EEG Brainwave Emotion Detection using Stacked Ensembling | Deep Neural Network, Principal Component Analysis

- Programmed a Stacked Ensemble model for emotion classification by combining outputs of eight base models such as deep neural networks and machine learning models such as Random Forest with a 97% accuracy.
- Authored and presented a paper at the international conference, the 12th International Conference on Computing, Communication and Networking Technologies, in July 2021. DOI: 10.1109/ICCCNT51525.2021.9579818.

Pneumonia Detection from Chest X-ray using Transfer Learning (Project Team Leader) | CNN, Transfer Learning, Image Processing, and Data augmentation

- Led a team of five to engineer three transfer learning models, namely ResNet50, VGG-16, and Inception V3 to aid pneumonia detection with a maximum recall of 98.97% and accuracy of 94.07%.
- Authored and presented a paper at the international conference, the IEEE 6th International Conference for Convergence in Technology in March-April 2021. DOI: 10.1109/12CT51068.2021.9417872.

Game Industry Sales Analysis (Project Team Leader) | Plotly, Seaborn, Matplotlib, NumPy, and Pandas.

- Analyzed key trends in gaming industry by performing multivariate, bivariate, and univariate analysis such as patterns in critic and user reviews w.r.t consoles, sales across financial periods, and performance analysis using Python.

TECHNICAL SKILLS

- Programming Language and Core Skills: Python, Machine Learning, Statistics, Deep Learning, Computer Vision, Data Analysis
- Databases and Frameworks: MySQL, XML, MongoDB, DynamoDB, FireBase, Apache Spark, Hadoop, MapReduce, HDFS, Resilient Distributed Databases
- Data Science libraries and tools: NumPy, Pandas, Scikit-Learn, TensorFlow, Keras, PyTorch, SciPy, Matplotlib, Plotly, Seaborn, XGBoost, Tableau, AWS ec2, Jupyter

ACTIVITIES

- Qualified as the National Finalist with a rank of 6 at the Business Data Analytics competition Anumaan, IIT Delhi in 2021 for analyzing key trends in sales data for six markets of an e-commerce company.
- Completed Certification in 2019-2020: DeepLearning.ai, Applied Machine Learning by the University of Michigan.
- Volunteered at the Abhyudaya SPJIMR for 2018-2020, teaching underprivileged students Science and Mathematics.
- Led the marketing team for the inter-college dance competition and negotiated with organizations to raise sponsorships.