

Satyaraja Dasara (Satya)

812-606-9483 | sdasar95@gmail.com | linkedin.com/in/sdasara95 | github.com/sdasara95

Education

M.S. in Data Science

Indiana University Bloomington

May 2020

GPA: 3.87

B.S. in Computer Science

R.V. College of Engineering, Bengaluru, India

June 2017

GPA: 3.33

Experience

Data Science Intern

Edgetensor Technologies, USA

May 2019 – December 2019

- Developed webcam eye-tracking software using regression analysis in Python to identify regions of interest on the screen achieving **20% less pixel error** than existing open source projects.
- Automated suspicious activity detection in airport and parking lot videos using object detection and tracking (YOLOv3, DeepSORT) improving security and **cutting security expenditure in half**.
- Analyzed and visualized results using Pandas and Matplotlib to evaluate and improve proprietary deep learning models leading to **cloud computing cost reduction by 70%**.
- Built and trained deep learning models using Tensorflow to predict eye gaze from image and video datasets to identify distracted drivers and reduce car accidents.
- Developed real-time object detection and image segmentation applications using Deepstream SDK on NVIDIA Jetson TX2.
- Communicated the results to other team members and business stakeholders.

Data Science Intern

Quadratic Insights, India

January 2018 – June 2018

- Deployed a backend Document Analysis service using OpenCV, Tesseract and Django to automate filling of data into MS Excel from scanned bank documents **reducing man-hours by 50%** for clients.
- Analyzed and predicted the impact of a trade union's promotional events on it's members' churn rate using ensemble learning **achieving 83% recall, 12% improvement** over their baseline model.
- Developed load forecasting models for electricity board using GRU and Time Series resulting in **30% less energy wastage**.
- Performed data cleaning in Python and R and visualized data and findings using Tableau and R Shiny to derive insights.

Data Science Projects

Google Analytics Customer Revenue Prediction (*Sklearn, Lifetimes*)

- Analyzed Google Analytics transaction level data to identify revenue generating customers, forecast revenue generated by them and predict customer lifetime value (CLV) for better targeted marketing.
- Implemented regression, Random Forest and gradient boosting models for identification and forecasting of customers and Betageofitter and Gammafitter models for predicting the lifetime value of a customer. Achieved **88.5% recall**.

Super Resolution using Generative Adversarial Network (*Keras, Tensorflow, OpenCV, GCP*)

- Built a GAN for 4x image super resolution from 100x100 pixel low resolution images in DIV2K dataset using VGG loss.
- Achieved good perceptual quality with an **average PSNR of 23** and **SSIM of 0.62** after training for 55 epochs on GCP.

Amazon Sports Review Classification (*Sklearn, NLTK*)

- Performed text analysis on Amazon reviews for sports products to predict good and bad reviews. Trained ensemble of Naives Bayes, Random Forest, logistic regression, XGboost and Neural Network models and achieved an F1 score of **0.914**.

Speech Denoising (*Librosa, Tensorflow*)

- Developed 1D CNN, 2D CNN, LSTM to filter noisy speech audio samples. Trained on 1200 audio signals and tested on 400 test signals achieving **mean SNR of 15 dB** for LSTM and **11dB** for 1D CNN.

Skills

Languages: Python, R, C++, JavaScript, SQL, HTML, CSS

Libraries: NumPy, Pandas, Scikit-Learn, OpenCV, NLTK, Gensim, Tensorflow, Keras, Matplotlib, Altair, caret, dplyr, h2o, ggplot

Frameworks/Tools: Git, Hadoop, Spark, Hive, Pig, HBase, MongoDB, Tableau, GCP, AWS, Docker

Machine Learning: Logistic Regression, Decision Tree, MCMC, SVM, Clustering, Random Forest, XGBoost, Ensemble Learning

Deep Learning: Gradient Descent, CNN, GRU, LSTM, VAE, GAN, KL Divergence, Batch Normalization, Mini-Batch Sampling, t-SNE

Statistics: A/B testing, Bayesian inference, Cross Validation, Time Series, Multicollinearity, AIC, VIF, PCA, Chi-squared test

Coursework: Design and Analysis of Algorithms, Machine Learning, Statistics, Deep Learning, Computer Vision,

Data Visualization, Elements of AI, Health Informatics, Big Data Management