DEVI SANDEEP ENDLURI

College Station, Texas | (979) 739-3429 | https://github.com/dsandeep0138
dsandeep97@tamu.edu | https://www.linkedin.com/in/dsandeep97

EDUCATION

Texas A&M University, College Station, Texas — Aug 2019 – (exp.) May 2021
Master of Science in Computer Science — **GPA**: **3.8 / 4.0**
Coursework: Deep Learning, Natural Language Processing, Analysis of Algorithms, Info Storage and Retrieval

Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, India — Aug 2010 – May 2014
Bachelor of Technology in Computer Science and Engineering — **GPA: 8.27 / 10**

EXPERIENCE

Pennsylvania State University, State College, Pennsylvania
Data Science Research Intern — May 2020 - Present

- Developed a fully automated end-to-end framework (ChartReader) to extract data from plots in scientific research papers
- Classified charts from research papers using VGG Neural Networks with an accuracy of 84.01 across 13 chart categories
- Applied Computer Vision techniques using OpenCV to detect axes with an accuracy of 80.22, plot labels, legends and to finally extract data from plots. Used Amazon Rekognition text detection to detect text from chart images.

Data Analytics at Texas A&M (DATA) Lab, Texas A&M University, College Station, Texas
Graduate Student Researcher (under Prof. Xia Ben Hu) — Jan 2020 – Present

- Working on a pipeline utilizing AutoML to automatically search for a best neural model for Natural Language Processing tasks such as Named Entity Recognition
- Constructed Knowledge graphs based on relations extracted from COVID-19 Open Research Dataset (CORD-19)

Qualcomm India Private Limited, Hyderabad, India
Senior Software Engineer — July 2014 – July 2019

- Facilitated design, development of innovative algorithms and maintenance of proprietary software CnE (Connectivity Engine) for intelligent switchover between 3G/4G and Wi-Fi without any user intervention
- Spearheaded various IMS critical value-add features (G2L Tuneaway, Dual VoLTE) for Qualcomm chipsets
- Interacted with 10+ internal and external teams to develop features end-to-end. Experience with partnership and collaboration with customers, ecosystem providers and support, during all stages of software product life cycle
- Awarded 5+ Qualstars, Orion Insta award in appreciation of outstanding contributions to Android Connectivity domain

PROJECTS & COMPETITIONS

Open Source Contributions: scrapy (#4634), tensorflow (#40610), scipy (#20), scikit-image (#4798, #4803), genism (#2869)

Real-time COVID-19 Twitter Data Analysis using Spark — April 2020

- Performed Real-time data analytics on COVID-19 over a Twitter Stream using Big Data Technologies of Hadoop Ecosystem such as Flume, Kafka and Spark Streaming. Built a Flask Web Application to display results and dashboards

Regression models to predict flight delays | *TAMIDS 2020 Data Science Competition* — April 2020

- Built Linear, Lasso, Ridge and Bagged Linear regression models to predict flight delays for 3rd and 4th Quarters of 2019. Presented 2018 flight delay data visually through dashboards using leaflet in R. Achieved test RMSE of 9.952

Deep Learning Image Colorization based on U-Net — Oct – Dec 2019

- Developed neural network regression and classification approaches to convert grayscale images to colorized RGB images with architecture inspired by U-Net, a convolutional method for image segmentation. Achieved accuracy of 70

Abstractive Text Summarization using pre-trained encoders (NLP project) — Oct – Dec 2019

- Modified existing text summarization model with pre-trained BERTSUM encoder model and decoder architecture by introducing recurrence in model to improve better copying of source text, achieved a ROGUE score of 19.03

Data Visualization model to analyze Tacos and Burritos data | *Goldman Sachs Challenge, TAMU Datathon* — 2019

- Derived insights from a list of 19,439 restaurants and businesses with menu items containing tacos and burritos from across the US. Delivered an interactive visualization tool using Tableau detailing the data analysis performed

AWARDS AND HONORS

- **Finalist** in **TAMIDS (Texas A&M Institute of Data Science) 2020 Data Science Competition**
- **17th** out of 70 teams in **ConocoPhillips Kaggle challenge** during TAMU Datathon, 2019
- **8th** out of 1000+ participants in **HackerEarth Machine Learning Challenge - Predict the DEFCON level**

TECHNICAL SKILLS

Languages: (Advanced) Python, SQL, R, C, C++; (Proficient) MATLAB, Java, Perl, Ajax, PHP, JavaScript
Frameworks: Python (NumPy, Pandas, Scikit-learn, matplotlib, TensorFlow, Keras), R (ggplot2), OpenCV, Spark, Kafka, Git, AWS
Certifications: Machine Learning, Deep Learning (Stanford University); R programming (Johns Hopkins University)