

DEVI SANDEEP ENDLURI

(979) 739-3429 • dsandeep97@tamu.edu • [GitHub](#) • [Portfolio](#) • in/dsandeep97 • Bryan, Texas

EDUCATION

Texas A&M University, College Station, Texas **May 2021**
Master of Computer Science GPA: 3.857 / 4.0

Teaching Assistant: Programming Studio CSCE 315 (Spring 2020, Fall 2020)

Coursework: Deep Learning, NLP, Analysis of Algorithms, Software Engineering, Info Storage and Retrieval

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

EXPERIENCE

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

- Formulated an **AutoML** pipeline to automatically search for a best neural model for Named Entity Recognition NLP task
- Constructed Knowledge graph based on 0.9M co-occurrence relations extracted from COVID-19 Open Research Dataset

Penn State University, State College, Pennsylvania **May 2020 - Aug 2020**
Data Science Research Intern

- Productized a fully automated end-to-end framework in **Python** with **Amazon Rekognition** to detect text in image with 90+ accuracy and **OpenCV** to extract information from charts in scientific research papers
- Achieved an accuracy of 84.08 in chart classification using VGG Neural Networks in **Keras**, 98% and 68% in parsing x-axis and y-axis ticks, 83% in detecting legends and 42% in extracting data values

Qualcomm India Private Limited, Hyderabad, Telangana **July 2014 - July 2019**
Software Engineer

- Facilitated software design in **C** and **C++**, development of innovative algorithms, debug, and maintenance of proprietary software CnE (Connectivity Engine) for intelligent switchover between 3G/4G and Wi-Fi without any user intervention
- Accomplished various IMS critical value-add features (G2L Tuneaway, Dual VoLTE) for Qualcomm chipsets
- Awarded 5+ Qualstars, Orion Insta award in appreciation of outstanding contributions to Android Connectivity domain

PROJECTS

Animations in Cell Biology Learning Content (Skills: **HTML, CSS, JavaScript**) **Oct 2020 - Dec 2020**
• Developed 5+ interactive games, animations to improve biology learning experience of middle school students. [YouTube](#)

Real-time COVID-19 Twitter Data Analytics (Skills: **Python, Java, Flume, Kafka, Spark, Flask**) **Apr 2020**
• Created a production-ready end-to-end system for real-time data analytics on COVID-19 by pipelining Twitter Stream with Flume, Kafka using Spark Streaming. Deployed system on AWS with dashboards designed and displayed utilizing Flask

Deep Learning Image Colorization based on U-Net (Skills: **Python, Keras**) **Oct 2019 - Dec 2019**
• Implemented neural network regression and classification approaches using an architecture inspired by U-Net in Keras to convert grayscale images to colorized RGB images with an accuracy of 70%

Abstractive Text Summarization using pre-trained encoders (Skills: **Python, PyTorch**) **Oct 2019 - Dec 2019**
• Enriched existing text summarization model with pre-trained BERTSUM encoder model and decoder architecture written in PyTorch by introducing recurrence in model to improve copying of source text, yielded a ROGUE score of 19.03

- **[Best Model Award]** **[DSGO Virtual Hackathon 2020]** Built a predictive model to determine missing NO₃ values after analyzing data from different air quality monitoring sites in California. Outperformed other models with RMSE of 0.0017
- **[TAMU Datathon 2020]** Built a product search engine for various categories by crawling through walmart.com. [YouTube](#)
- **[Finalist]** **[TAMIDS 2020 Data Science Competition]** Developed Linear, Ridge regression models to predict flight delays for 3rd and 4th Quarters of 2019. Achieved test RMSE of 9.952. Presented 2018 flight delay data visually using leaflet in R
- **[Open Source Contributions]** scrapy ([#4634](#)), TensorFlow ([#40610](#)), scipy ([#20](#)), scikit-image ([#4803](#)), gensim ([#2869](#))

TECHNICAL SKILLS

Languages: Python (NumPy, Pandas, Scikit-learn, matplotlib, TensorFlow, Keras), SQL, R, C, C++, MATLAB, Java, JavaScript
Big Data & Machine Learning: OpenCV, Spark, Kafka, Excel, Git, AWS, Regression, Classification, Clustering, PCA, Data Mining
Data Science Technologies: A/B Testing, Data science pipeline (cleansing, wrangling, visualization, modeling, interpretation)
Certifications: Machine Learning, Deep Learning (Stanford University)