Dharani Doppalapudi

Bloomington, IN | +1 (317) 696-9036 | dharanidoppalapudi04@gmail.com | LinkedIn | GitHub

EDUCATION

Indiana University Bloomington, IN, USA May 2023
Master of Science in Data Science **GPA: 3.7/4.0**
Coursework: Applied Algorithms, Data Mining, Computer Vision, Deep Learning Systems, Applied Machine Learning, Big Data, Data Visualization

VNR Vignana Jyothi Institute of Engineering and Technology, TS, India April 2019
Bachelor of Technology in Information Technology **GPA: 3.6/4.0**
Coursework: Advanced Data Structures, Database Management System, Web Programming, Cloud Computing, Data Analysis, Statistics

Indian School of Business, TS, India November 2018
Completed a two-year course in Technology Entrepreneurship Program focusing on convergence of business and technology, principles of product management and road mapping. Collaborated with cross functional teams and designed dashboards using **Tableau** to present complex data in concise manner for business analysis and drive business value.
Coursework: **Agile** Methodologies, Software Quality Assurance, Product Development, **Business Analysis**, Customer Relation Management, Tableau

PROJECTS

Argentina Real Estate Listings August 2022
Tech: PySpark, Google Cloud Platform, Tableau, Matplotlib, Google Big Query, SQL, Quality analysis, Data Analytics
- Automated **ETL** pipelines on **GCP**. Conducted Exploratory Data Analysis on 1 million rows dataset using **PySpark** and **Big Query**.
- Performed data cleaning and **quality assurance** and executed complex **SQL** queries and visualized the insights.

Depression detection using Computer Vision February 2022
Tech: Computer vision, Deep learning, Convolutional neural network, Image Classification, Python, PyTorch, Keras, PIL, BERT, Word2vec
- Created depression detection model using **3D CNNs**, spectrogram images from audio, and Google's **SpecAugment** strategy.
- Achieved 84% recall, 81% accuracy with VGG16 and Google's **BERT**. Performed speech denoising using 1D CNN.

Sentiment Analysis on self-driving car tweets December 2021
Tech: NLTK, Scikit-learn, NLP, Python, Applied Machine Learning, Data Mining
- Created sentiment analysis models using bag of words, MPQA subjectivity lexicon, dependency triples approaches, using **NLTK** and **Scikit-learn** libraries. Analyzed crowd Flower's self-driving car sentiment dataset using **Natural Language Processing**.

Home Credit Default Risk September 2021
Tech: Logistic Regression, XGBoost, Multi-layer Perceptron, Python, Matplotlib, Seaborn, Exploratory Data Analysis
- Creating Designed a classifier to predict the capability of a person to repay the loan for a Kaggle Competition.
- Used Logistic Regression, SVM, KNN, **XGBoost**, Random Forest and **MLP**. Performed EDA, data visualization using Matplotlib.

WORK EXPERIENCE

Indiana University Bloomington - Research Intern April 2022 – Present
IN, USA| Computer Vision, Python, OpenFace, Object Detection, Corner Detection, Gaze Detection, Deep Learning
- Explored and implemented **Gaze detection** models. Created models to **detect valence and arousal levels** from videos. Implemented hand and toe detection models. Used SIFT algorithm features for object detection and image matching.
- Performed **engagement analysis** by combining and correlating gaze direction and emotion levels under Professor David Crandall.

ValueLabs – Data Analyst June 2019 – July 2021
TS, India | Python, Machine Learning, Exploratory Data Analysis, Data Visualization, Predictive Analysis, DevOps, Microsoft Excel
- Created a python application to project and identify high-risk and low-risk zones for covid-19 cases in Charleston, SC, USA, and estimated infection rates. Trained an **ensemble** machine learning model to anticipate covid-19 risk, resulted in a **40%** performance improvement. Deployed using **AWS SageMaker** ensuring scalability.
- Presented dynamic **visualizations** of travel recommendations, risk zone maps based on results in **Tableau**.
- Performed **ETL** operations from Excel files, optimized **SQL** queries to efficiently ingest large data into system. Automated infrastructure deployment using **CI/CD** pipelines reducing downtime from **154** hours to **68** hours. Designed **PowerBI** dashboards to display **data analysis** results to management for better business decisions.

ValueLabs – Software Intern January 2019 – March 2019
TS, India | Python, NLTK, Scikit-learn, SciPy, Artificial Intelligence, Matplotlib, NLP
- Performed text summarization on Deep Mind Q&A dataset by removing punctuation, stop words, lemmatization, and generated summaries of the stories. Trained **TF-IDF** vectorizer to get unique words. Used **cosine similarity** to check the similarity of generated text and the highlights in the dataset. Implemented this on our employee portal blog.

TECHNICAL SKILLS

Languages and tools: C, Python, R, Java, MYSQL, Git, MS Office, PyCharm, Tableau, PowerBI, Jenkins, A/B Testing, Docker, Agile, GCP
Libraries: NumPy, Pandas, Scikit-learn, OpenCV, OpenFace, Scikit-image, Keras, PIL, Tensorflow, PyTorch, Matplotlib, Seaborn

CERTIFICATIONS

- Oracle Certified Associate - Java SE8 Programmer
- VNR Vignana Jyothi Institute of Engineering and Technology – Python programmer
- Udemy – Product Manager