Akhil Songa Github: <https://github.com/akhilsonga>

# Philadelphia, PA 19104 • [as5735@drexel.edu](mailto:as5735@drexel.edu) • +1 (267)-467-6401 • LinkedIn: <https://www.linkedin.com/in/akhil-songa-agnos>

# EDUCATION

**DREXEL UNIVERSITY |** Philadelphia, PA (Sept 2022 – 2024)

Master of Science in Data Science, Awarded a $8,000 Dean's Fellowship, Coursework: Data Acquisition and Pre-Processing, Data Structures and Algorithms, Data Analytics

**GITAM UNIVERSITY |**Visakhapatnam, India. 2018 – 2022, GPA: 3.4 (8.17/10)

Bachelor of Computer Science Engineering, Coursework: UNIX, Machine Learning, Probability and Statistics, Data Mining and Data Warehousing, Artificial Intelligence.

# EXPERIENCE

**GITAM UNIVERSITY |** Visakhapatnam, India

**Research Assistant |** Jan 2021 – Aug 2021

* Worked under Dr. Bhavani Madireddy on Data Visualization and published a Review article, “A Comparative Study of Data Visualization Techniques Using Covid-19 Data.” During the study, analyzed the feasibility and use-cases of various Data Visualization methods with respect to diverse categories of data.

**TECHNOCOLABS SOFTWARE |** Remote

**Data Science Intern |** Oct 2021 – Nov 2021

* Worked on Predicting a Startup’s Acquisition Status using CrunchBase datasets on a Startup’s financial information.
* Aided the decision-making ability of an investor or a jobseeker to determine if a given company would be acquired or goes for an IPO or shut down.
* Transformed the original dataset through Exploratory Data Analysis (EDA) and Data-Preprocessing methods, thereby making it efficient to train a regression-based machine learning model. Performed HyperParameter Tuning to further increase its efficiency

**INEURON |** Remote

**Data Science Intern |** Feb 2021 – Mar 2022

* Worked on various machine learning project pipeline roles, such as data loading, data cleaning, data validation, model selection, model tuning, & model monitoring pipelines, for various projects and implemented a machine learning pipeline in the Azure machine learning studio.

# PROJECTS

**QUORA QUESTIONS TOPIC CLASSIFIER USING TERM FREQUENCY INVERSE DOCUMENT FREQUENCY AND NON-NEGATIVE MATRIX FACTORIZER**

* Questions from the Quora website have been classified into multiple topics as politics, economy, technology, education, corporate jobs, books, and movies, by using TFIDF vectorization and the Non-Matrix Factorization model, model fitting was completed in just 1.077 seconds, which is 20% faster than my previous model.

**BANK: FOR THERA BANK, PREDICTING THE LIKELYHOOD OF LIABILITY CUSTOMER BUYING PERONAL LOAN USING VARIOUS DATA VISUALIZATION TECHNIQUES**

* By using several data visualization techniques to the dataset, I concluded that logistic regression, nave bias, and KNN are effective. After constructing a confusion matrix, I selected logistic regression with an accuracy of 95.46 %, which is higher than other regressions.

**AMAZON PRODUCT REVIEW SENTIMENT ANALYSIS USING SENTIMENT INTENSITY ANALYZER – NLTK**

* By using machine learning techniques like NLTK Vader, Sentiment Intensity Analyzer and vectorization a given review will be classified as negative, neutral, positive (Polarity Scores).

# ACHIEVEMENTS & PUBLICATIONS

* “A Comparative Study of Various Data Visualization Techniques using COVID-19 Data” (August 2021 Issue - IRJET).
* “The Societal and Transformational Impacts of Data Science” (September 2021 Issue - IJERT).
* “Vehicle Number Plate Recognition System Using TESSERACT-OCR” (April 2022 Issue - IJRASET).
* “Credit Card Fraud Detection using Various Machine Learning Algorithms” (September 2022 Issue - IJRASET).
* My Research-Interest Score is greater than 31% of all ResearchGate members and 84% of 2021 first-time authors.

# CERTIFICATIONS

• Natural Language Processing with Python, Udemy • Complete TensorFlow, Keras, Deep Learning Bootcamp, Udemy

• Programming for Everybody Python-University of Michigan. • Machine Learning Algorithms-Alberta Machine Intelligence-Institute.

# TECHNICAL KNOWLEDGE

Python, SQL, Java, C, C++, MySQL, Oracle SQL, MongoDB, Sci-kit, Matplotlib, Git, Tableau, Seaborn, HTML, CSS

, Microsoft Excel, Django, Flask, Regression models, Classification models, Spacy, YOLO v4, OpenCV, NLP, keras, Tensorflow, Pytorch, Sentiment analysis, PowerBI, Plotly, Artificial Neural Network, Convolutional Neural Network, Recurrent Neural Network, Long Short-Term Memory.