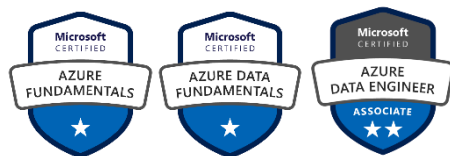


Harsha Buddana

Denton, TX, 76205 | 469.325.2204 | harshabuddana@gmail.com | LinkedIn:
<https://www.linkedin.com/in/harsha-buddana/>



Professional Profile

- 3+ years' experience in Python, SQL, ETL tools, Data Science and published a research paper on Text Generation.
- Dynamic and results-driven Data Engineer with a master's degree in data engineering
- Solid background in handling big data, developing ETL processes, and implementing scalable data architectures.
- Skilled in utilizing cutting-edge technologies, including Microsoft Azure, MySQL, PySpark, and Tableau
- Seeking a challenging Data Engineer role to leverage technical expertise and passion for data engineering.

EDUCATION

University of North Texas | Denton, Texas

May 2023

Master of Science | Data Engineering

GPA: 3.6/4.0

Relevant Course Work: Big Data, Distributed and parallel Database Systems, Advanced Data Analytics, Fundamentals of Database Management System, Directed Study on Natural Language Processing (NLP) Artificial Intelligence, Machine Learning.

Marri Laxman Reddy Institute of Technology and Management | Telangana, India

Bachelor of Technology, Electronics & Communication, ABET accredited

Aug 2021

GPA: 7.4/10.0

TECHNICAL SKILLS

Data Engineering Tech Stack: MySQL, Tableau, Power BI, PySpark, ETL, MapReduce/Hadoop, Microsoft Suite, Snowflake.

Cloud: Microsoft Azure: Data Lake, Data Factory, SQL Data Warehouse, Synapse, Cosmos DB, Data Bricks

Languages: Python, C.

Data Analytics Algorithms: Linear Regression, Logistic Regression, K-Mean, KNN, TensorFlow.

PRACTICUM/RESEARCH/INTERNSHIP EXPERIENCE

Clove Technologies, Data Science Intern | Telangana, India

May-Jul 2019

- In my association with Clove technologies, I acquired knowledge of various data handling, Machine learning and IoT applications.
- I worked on Data logging with DHT22 sensor interfacing with a Raspberry Pi and AC automation.

EMPLOYMENT EXPERIENCE

Peers Technology | Hyderabad, India

Jun – December 2021

Data Engineer:

- Designed and implemented ETL processes using Azure Data Factory, resulting in a 40% reduction in processing time.
- Collaborated with cross-functional teams to design and implement scalable data architectures and reported major metrics to the authorities.

CSCE Department, UNT | Denton, TX

Jan – May 2023

Instructional Assistant:

- Assisted group of 45 graduate students in their In-class assignments and graded their assignments for course Introduction to Bigdata and Data Science.

Sycamore Library, UNT | Denton, TX

Jun – December 2022

Student Library Assistant:

- Assisted patrons in accessing the library resources and made use of our database to record patron interactions and data important to the libraries.

RESEARCH AND PUBLICATIONS

Word Level LSTM and Recurrent Neural Network for Automatic Text Generation

Organized by: IEEE | Date: Jan 2021

This paper delineates the design and working of text generation using word-level LSTM-RNN.

[Word Level LSTM and Recurrent Neural Network for Automatic Text Generation | IEEE Conference Publication | IEEE Xplore](#)

PROJECTS

COVID-19 reporting using Azure Data Factory

Tech Stack: Azure: Data Storage, Data Lake Gen2, SQL Data warehouse, Data Factory, Databricks, Power BI.

Data Engineering Project on Covid19 Reporting – Using Azure Data Factory, Databricks, HDInsight, Azure Data Factory – An End-to-End ETL pipeline in addition to a Power BI report dashboard.

GitHub Link https://github.com/hbuddana/Azure_Data_Factory_COVID-19_Reporting

Facebook Ad Campaign Dashboard

Tech Stack: Python, Google Collab, Stream lit Cloud, NumPy.

This project aims to visualize social media sales prediction and ad-campaign analysis by optimizing Facebook ad-campaigns, predicting future sales from campaign data, and deploying a web application, involving tasks like data preprocessing, exploratory analysis, and machine learning model training.

Git Hub Link - <https://github.com/hbuddana/sdvproj>

Text-Summarization of COVID-19 Medical Articles using Distil BERT

Tech Stack – Python, NLP, NLTK, SpaCy, BERT

Text summarization of COVID-19 dataset using Distil BERT integrated with a Question- Answering model.

GitHub Link - <https://github.com/hbuddana/Text-Summarization-of-COVID-19-Medical-Articles-using-DistilBERT>