

# Divya Sree Velugula

[divya.velugula@gmail.com](mailto:divya.velugula@gmail.com) +1 (682)-2528817 [linkedin.com/in/divya-velugula/](https://www.linkedin.com/in/divya-velugula/)

[github.com/DivyaVelugula](https://github.com/DivyaVelugula)

## Big Data Developer, Data Engineer, Data Analyst

### Professional Experience:

Student Software Developer: UNT Data Science, Denton, USA

(07/2019-Present)

#### Responsibilities:

- Custom application programming and portal development.
- Migration and manipulation of data from legacy systems, HR and Student.
- Developed client specific reports using Report Builder in SSRS.

Associate Technical Analyst: Tech Mahindra Pvt. Ltd., Hyderabad, India

#### Responsibilities:

(3/2016-12/2018)

- Cloudera was used as a distribution platform for Hadoop.
- Created HIVE Tables on top of datasets which are in staging layer.
- Used Various Hive performance optimization techniques.
- Created Impala Tables as the target tables to load data into the integration layer.
- Extensively used Cloudera Stream-sets to transfer Raw data files into HDFS foundation layer.
- Used Parquet file format in all target tables in Impala for better performance.
- Used Spark core and Spark SQL for data transformations.
- Involved in Building data frames and RDD's using Spark SQL and Spark Core

### Education:

- Master's in Data Science, University of North Texas, Denton. (GPA-3.7 Graduation- Spring 2020)
- Bachelor of Technology, Jawaharlal Nehru Technological University.(GPA-3.7 Graduated- Spring 2016)

### Academic Projects:

- **Twitter-Data-Analysis:**  
Collected Tweets using twitter Streaming API and stored in HDFS. Analyzed the data by writing Analytics Queries in spark SQL and visualized data using AMcharts. (Python, Java, Apache Spark, Hadoop, Angular)
- **Walmart Sales Analysis:**  
Collected Walmart sales dataset from Kaggle and analyzed the sales over the past few years using a data visualization tool Tableau. (Sql, Apache Hadoop, Apache Hive)

### Skill Summary:

Big Data Technologies	: Hadoop, HDFS, SQOOP, HIVE, DWH, MapReduce, Apache Spark, PIG, FLUME, StreamSets
Skills	: SQL, PLSQL, Data Visualization(TABLEAU)
Scripting Languages	: UNIX Bash/Shell Scripting, Python.
Databases	: MS SQL Server 2016, Oracle 12C, MySQL
Version Control	: Git.
Others	: Office 365 Suite, JIRA, Confluence, RESTful
Developer Tools	: SQL Tuner, NiFi, HUE
High Level Lang.	: C, Python
Distribution N/w	: Cloudera, Hortonworks