**Key Accomplishments:**

**D-TAS DS Bootcamp Certification:**

1. Udemy - Data Science Bootcamp
2. Udemy - Feature Engineering for Machine Learning
3. Udemy - Machine Learning with Imbalance Data

**Spark NLP Certification:**

1. Certified Spark NLP Healthcare Data Scientist from John Snow LABS

**Internal Training Program:**

1. UniverCT internal SQL training completed
2. PySpark training (In-Progress)

**Areas of Focus:**

1. Pyspark
2. Natural language processing (Name entity recognition, Attention model, Transfer learning, LSTM, GRU)
3. Deep Learning (ANN and CNN)

**Kindly provide justification as your utilization from April to September is less than 80%:**

I have joined CitiusTech on 27th July 2022. During my initial days in CitiusTech, we were in the different induction programs like healthcare induction, information security etc. After that I’m part of data science bootcamp training programs from last 4 months. During this period, I have completed the certifications like Data Science Bootcamp, Feature Engineering for Machine Learning and Machine learning with Imbalance Data.

In my previous experience I had only experience in the field of Machine Learning but here most of the projects are related to Natural Language Processing and Deep Learning. So, currently I’m focusing on acquiring skills on different aspects of Natural Language Processing and Deep Learning. In that part I have recently certified Spark NLP Healthcare Data Scientist from John Snow LABS.

**PROUD:**

**P-PASSION:** I have been extremely passionate about the training I’m undergoing and I’m looking forward to exploring new technologies.

**R-RESPECT:** I highly appreciate all the CTzens and the valuable contributions they make.

**O-OPENESS:** I am always open to take the advice and also committed to helping those who require support.

**U-UNITY:** I strongly believe that teamwork can be incredibly powerful and help a team to achieve great results. Building a supportive environment encourages collaboration, boosts productivity, and improves job satisfaction.

**D-DEPTH:** I put in the effort to comprehend everything in a detailed and conceptual way.

**Process Adherence:**

I make sure to adhere to the process and complete it within the allotted timeline. Like filling the timesheet, applying leave requests, attending sessions, infosec assessments and other CT process. For the planned leave, I used to mention before.

**Cross-skilling / Learning and Development:**

**D-TAS DS Bootcamp Certification:**

1. Udemy - Data Science Bootcamp
2. Udemy - Feature Engineering for Machine Learning
3. Udemy - Machine Learning with Imbalance Data
4. Udemy - Pyspark Essential Data Scientist
5. Udemy - NLP - Natural Language Processing with Python

**Spark NLP Certification:**

1. Certified Spark NLP Healthcare Data Scientist from John Snow LABS

**Internal Training Program:**

1. UniverCT internal SQL training completed
2. PySpark training completed

**CT Contribution:**

Additional Certificate:

1. **Sponsored by CT:** Spark NLP Healthcare Data Scientist from John Snow LABS

Creation of Training Material:

1. **POC: Spend Analytics**

Currently we are working on that POC. So far, I worked on few areas such as, synthetic data creation, master table creation by using SQL joint operations, Data dictionary creation, WinScp and DBeaver configuration according to model requirement with our team.

**Kindly provide justification as your utilization from April to September is less than 80%:**

I have joined CitiusTech on 27th July 2022. During my initial days in CitiusTech, we were in the different induction programs like healthcare induction, information security etc. After that I’m part of data science bootcamp training programs from last 4-5 months. During this period, I have completed the certifications like Data Science Bootcamp, Feature Engineering for Machine Learning, Machine learning with Imbalance Data, Pyspark Essential Data Scientist and Natural Language Processing with Python

In my previous experience I had only experience in the field of Machine Learning but here most of the projects are related to Natural Language Processing and Deep Learning. So, currently I’m focusing on acquiring skills on different aspects of Natural Language Processing and Deep Learning. In that part I have recently certified Spark NLP Healthcare Data Scientist from John Snow LABS.

Currently I’m working on Spend Analytics POC. So far, I worked on few areas, such as,

1. Create a healthcare supply chain management Synthetic data by using Mockaroo tool.
2. Create a Master table from 5-6 different tables by using SQL joint operation.
3. By using pandas rearranged the Vendors name and Pricing for each item.
4. Completed the data dictionary part with our team.
5. Completed Spend Analytics model configuration in WinSCP with our team.
6. Configured DBeaver and set dependencies according to our model requirement.

**Performance on the project / assigned tasks 1:**

POC: Spend Analytics

Assigned Task: Synthetic data creation for healthcare supply chain management

1. Initially we tried to find the data for healthcare supply chain management into various sources like Kaggale, Data.word, Physionet, GitHub repository etc. But we are not able to find relevant data.
2. So, after that I found the list of vendors, manufacturers names along with medical items names on goggle. Based on those details we created a healthcare supply chain management Synthetic data by using Mockaroo tool.
3. This task requires to create a total of 5-6 tables with the provided table headers.

Table names: Item Master, Purchase order history, Vendor master, contract item lines, contract headers

**Performance on the project / assigned tasks 2:**

POC: Spend Analytics

Assigned Task: Create a master table of healthcare supply chain management data

1. In this activity, first I imported all the require csv files that were generated in the synthetic data creation task into a SQL server database.
2. Once all files are imported then I performed SQL joins on the tables to form a consolidated master table.
3. The outcome of the joint operation is stored in a csv file for future reference.
4. created a Master Table Connection Architecture that can help you better understand the link between tables and how they are interlinked. It facilitates an easier visualization of the relationship between tables.

**Performance on the project / assigned tasks 3:**

POC: Spend Analytics

Assigned Task: Data Pre-processing / Rearranged

1. In this activity, we used Panda’s python library to pre-process the data.
2. With the help of pandas, we restructured the Vendors’ names and prices for each item.
3. By considering the purchase order date, contact start and end date, we manually update the Vendors status to "No" for some of them.

**Performance on the project / assigned tasks 4:**

POC: Spend Analytics

Assigned Task: Data Dictionary Creation

1. Completed the data dictionary part for Spend analytics POC with our team.

**Performance on the project / assigned tasks 5:**

POC: Spend Analytics

Assigned Task: WinSCP Configuration for Spend Analytics POC

1. Completed Spend Analytics model configuration in WinSCP with our team.

**Performance on the project / assigned tasks 6:**

POC: Spend Analytics

Assigned Task: DBeaver Configuration for Spend Analytics POC

1. In the DBeaver configuration we updated data source, Entity configuration and entities table based on our model requirement.

synthetic data creation

master table creation by using SQL joint operations

pandas

Data dictionary creation

WinScp

DBeaver configuration according to model requirement with our team.

P-PASSION: I am passionate about the all the trainings and also passionate about learning new technologies.

I have been extremely passionate about the training I’m undergoing and I’m looking forward to exploring new technologies.

I am enthusiastic about the training I have received and also passionate about learning new technologies.

R-RESPECT: I highly appreciate all the CTzens and the valuable contributions they make.

O-OPENESS: I am always open to take the advice and also committed to helping those who require support.

U-UNITY: I strongly believed that a team could achieve brilliant things when they support each other. Also help to a boost in productivity and job satisfaction

I strongly believe that teamwork can be incredibly powerful and help a team to achieve great results. Building a supportive environment encourages collaboration, boosts productivity, and improves job satisfaction.

D-DEPTH: I put in the effort to comprehend everything in a detailed and conceptual way.

**Key skills:**

**Training Completed**

1. Machine learning projects (Regression and Classification)
2. Natural language processing projects specially on Document classification, NER, Sentiment Analysis.

I have joined CitiusTech on 27th July 2022. On the 1st month I was involved in the different-different induction program like healthcare induction. After that I’m part of data science boot camp training program from last 2 months. During this period, I have completed the Data science bootcamp certification, Feature Engineering for machine learning certification and Machine learning with Imbalance data certification.

In my previous experience I was done the projects in the Machine Learning fields but here my profile selected for Natural language processing so, currently I’m focusing to learn and upgrade my NLP skills. Recently I have certified Spark NLP Healthcare Data Scientist.