GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

During the first week of the internship, I was introduced to my colleagues, I got to know about the best practices of the office and had to attend an orientation session about the role of a cloud developer, what are growth opportunities, what are typical daily tasks in the life of an associate software engineer and what is expected of me in the role. After the orientation session, I was introduced to the structure of my training and the possible tasks or the areas that I might be assigned which included a web development project for a German Startup and An Service development project for a hospital in the United States of America

After my orientation, I spent a day setting up system, installing Ubuntu, installing python and all the necessary things needed for the internship which included a virtual enviournment, VScode, Jupter Notebook, Conda, Docker and Github Desktop

I attended a visual studio code training where i was taught the best practices while using vs code ide to organize your code, breakpoints and ide specific functionalities and hacks of vscode such as easy debugging and running of your code using 'F5'

After my vscode training,I was also introduced to the concept of operating systems and kernels, I learnt about microsoft, Linux and MacOs.
We then dived deeper into linux to understand the linux architecture, open source and bash. Here are some of things i learnt while learning about linux:

a) working with files, directories and linux file structures, open ssh and numerical permissions

b) shell training to learn more about installing and updating the packages, debugging dependency errors

c) using regex in shell to automate things, package manager

d) IManaging system units and logs, and managing users, scp and rsync

Thus, by the end of the week I had a basic understanding of the architecture of the linux, how the file system of the linux works and automating some basic commands using linux

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

This week was all about wrapping up linux and brushing up our knowledge gaps of linux by looking at frequently used commands of linux and debugging some common linux issues such as :
the path not being specified in the proper direction
navigating through computer hardware and pc issues
particularly, the ubuntu gui taking a lot of time to load which I narrowed it down to the fan speed frequency issues and debugged it

After finishing the basic training in linux, I was advised to take a look at the courses from the linux foundation and generate detailed bug reports if i ever encounter a bug while using linux.
The week also consisted of introduction to programming

During the introductory segment, I was taught about the difference between a compiler and an interpreter and an introduction to high level, medium level and low level programming languages with a particular emphasis on python
Then i learnt about the basic structure and syntax of a python program which includes data types and operators in python, reading and printing the user data and logical statements and converting your simple code into a logical block to organise it better

I was also introduced to strings, arryas , lists and dictonaries and the common methods to manipulate them such as slicing, indexing, reversing, capitalizing etc

After the end of this week, I was comfortable writing basic python programs such as finding out if the year is leap or not, finding out if the string is a palindrome or not, finding out if the given numbers were odd or even etc

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

This week it was my turn to dive deeper in the python programming and we moved onto more complex topics
such as logical if-else statements, switch statements, pass by reference, pass by value, understanding functions
and some per-defined functions in python

After that, i was given a ste of problems which could only be solved by making python functions and I solved the same. After functions, I was introduced to the terminology 'clean code' and what exactly is clean code, why is it important and best practices to convert my code into clean code

After the session on principles of clean code, I was introduced to the idea of Object Oriented Programming by using classes and object. I was introduced to some common modules in python and I was introduced to more OOP concepts such as inheritance, polymorphism

I then spent a day building a dummy class called 'Employee' and the objects inside it. After that, i was introduced to the popular libraries pandas and Numpy for data visualisation and data manipulation and had to solve some simple problems surrounding it

The end of this week gave me a sound understanding regarding basic and intermediate python and I am now able to create programs of higher complexity and interestingly, I converted all of my previous messy python code into a more modularized format following the principles of clean code

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

The following week was a lot more hands-on with focus on applying the learnt concepts instead of learning the new concepts

I spent a day learning different types of data and how to work with all of them in python. I learnt about the pandas pd.read_csv function and all the intricacies around it as i was told that i would be using that a lot. After that, i used the sample csv data and performed some computations around it to generate insights from the same

After that, I learnt about what is web scraping using the popular python module called selenium. Selenium needs to be connected to a webdriver and since I was using chrome, I installed webdriver for the same. Selenium lets you enter your prompts and the webpage that you want to scrape the data frome and it gives you the data in a csv format fairly easily. I scraped that data from the popular social media app called reddit using the praw library and then analyzed the data using python and numpy.

Then I was introduced to the principles of cloud computing, what is cloud computing and why do we need cloud computing when there are already traditonal data storage facilites and advantages of on prem as comapred to cloud computing. Various models of cloud computing such as SaaS, Paas, Iaas etc

I learnt about major cloud providers in the markets such as IBM, GCP and AWS, understanding the advantages of one over the other. I then learnt more about AWS such as AWS free tier, evolution of aws, integrating the AWS with other machine learning solutions.

Then i learnt about major storage solutions such aws EFS, S3,EC2, Object based storage, file based storage and instance based storage. Then I wrapped the week up with a revision of all the services learnt in aws and a quiz about the basics of aws

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

This week I learned more about the object based storage system called s3 and functonalities of s3 such as hosting a static website via s3, tiers of s3 such as glacier, Infrequent access etc. I also learnt about whitelisting the IP addresses.

I learnt about database services using aws such as aws dynamodb, aurora, rds and graphql. I learnt more about the fully managed services such as aws lambda, aws fargate and did a case study about when to use fully managed services like aws lambda, the tradeoff between managed services and price.

I then learnt about the tiers of AWS and AWS free tier and how to use aws services so that we can stay in the aws free tier and cost and resource management services of aws such as aws cost explorer, aws budgets, aws cloudwatch and how to make the most use of them.

Diving deep into AWS, i then learnt about AWS security and privacy and threats looming over your data in the aws. The shared responsbility model of the aws where consumers are responsible for the security of their data on the cloud and the AWS is responsbile for the security of the cloud.

Then i did a case study about low latency delivery of the AWS services such as CDN and how they build a network closer to the user so that they can enjoy their shows with high throughput (the netflix example)

To further elaborate on the services that I learnt of the AWS, I learnt about the AWS Macie which is an intelligent threat detection tool which uses machine learning to find sensitive data among your s3 buckets, I also learnt about AWS SNS and AWS SQS which are notification services which run on a pub-sub model to send and recieve notifications
Then we started understanding what is containers and why do we need one

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

This week was all about docker and containers and I learnt about what is docker, what is the difference between docker, container and kubernetes and why do we really need docker.

After understanding the need and the importance of docker, i then started installing docker on my system using virtualization but unfortunately, my system did not support virtualization and hence had to install that first. After succesfully installing docker on my machine, I went to install docker engine for linux for the easy usage of docker.

Then I learnt the architecture of a docker, what is a docker application and a docker image and how to run a docker image on your local machine

After that, I installed the hello world docker image on my machine and ran it succesfully. I also learnt about frequently used docker commands such as docker run, docker -ps, docker -s etc.

After learning about common features of docker, I was taught how to create my own docker image and make my code and application reusable to learn

I also learnt another key feature of the software engineering workflow i.e. version control, what is version control the history of version control, types of version control and then we dived deeper into git , the architecture of git and how is it different from subversion, setting up on own github account, how to clone and fork a repo, git ssh, and i also learnt about various common github features such as git pull, git push, git merge, git commit, git rebase, git revert and then I uploaded all of my work done till date on github repo and made to understand the importance of keeping your repository clean and organized. I also started to participate in open source contests by opening issues and pull requests on remote repository of github

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

The following week I was introduced to the project that I will be working on. The project I will be working on revolves around securing your data even before it is released. My project involves anonymizing the user data using static and dynamic anonymization techniques so that the privacy of the user is not sacrificied and the data still remains statistically significant so that it can be used for research and analysis

My task is to use the raw, sensitive data and to build an end to end anonymization pipeline so that the user doesn't have to worry about inputting and anonymizing the data manually. This week was spent on brainstorming the solutions for building the pipeline and using the right tools to anonymize the data

My first hunch was to look at the solutions around AWS. I looked up services around compute and storage such as lambda, ec2, s3 and dynamodb. However, it so happened that I was told not to use s3 as s3 is used only for static data and the data in question would be updated fairly frequently and adding a lambda function on top of an ec2 instance for storing data would not be fruitful hence i had to look for other services

The next service that i came up with was AWS glue databrew which basically has all the data analysis and visualization functonality in the form of jobs and it had functonality to handle PII data. However, this was turned down because the data had quasi identifiers which could be linked with other data to get the original information and hence we needed a more sophasticated algorithm. Upon researching a little more, i found a static anonymizat ion tool called ARX which uses privacy models such as K-anonymity, L-diversity and t-closeness to anonymized your dara and hence it can not be traced back to the users. I downloaded the docker image for a local arx instance and got that running on my local machine and then I downloaded a module called pyarxaas which was a wrapper module providing python functions to work with the local arx machine and then i decided to use the data warehousing solution called snowflake to host my two seprate databases, conncet to snowflake using python and anonymize the data and generate the visualizations and this solution was approved

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

I spent majority of this week trying to understand how arx works and how to anonymize my data using ARX. Firstly, I installed the arx software and tried anonymizing a sample dataset following the instructions found in the arx documentation and then installed the local arx server using the github repository called 'Arx as A service' after installing the local docker image of the arx server, i ran it on the port 8080 and tried using the API to access the arx service.
However, the arx instance in question was a bit to messy to handle with the arx api and hence i decided to install the python module called pyarxaas. After several failed tries and debugging i realized that my system's current python version is not supported by the module as the module was outdated and hence i had no other option but to install the virtual manager module in python called pyenv as I could not downgrade my entire system's python version due to depenedencies. After succesfully installing pyenv, I installed python 3.8 and set it as the global version and installed pyarxaas in the python 3.8 and downgraded my numpy and pandas to match the version of pyarxaas

After installing pyarxaas, I used the very popular titanic dataset from kaggle and tried anonymizing it. Since the titanic dataset contains the names of the customer and the ticket number and the class of the passengers, it is an ideal candidate for anonymizing the data. I set some of the columns of the titanic dataset such as sex and embarked as insensitive. The column name has sensitive as the name data can easily be used to track back to the user. All the other columns such as Pclass, ticket and passenger id as quasidentifying as although they can not be used on their own to identify the data or the person but they can be used in conjuction with other columns or data to identify. I then built the hierarchies using pyarxaas and set the hierarchies using various privacy models such as kanonymity, ldiversity and tcloseness. I then identified the risk of reidentifcation and compared the risk of identification before the anonymization and after the anonymization

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

This week was more project focused as I was focused on creating the dummy dataset which has the same structure of the actual dataset. Fake dummy data can be created using a python library called Faker. I installed the faker library by using pip and then got started creating the dummy data. The dummy data here is called 'customer_data' which is the data collected from a hypothetical e-commerce website. The dataframe contains the columns Name Which has the name of the customers which will be a sensitive column, Gender column has genders of the customers which is insensitive which is not important as it cannot be traced back to the users, Age column contains the ages of the consumers which is a quasidentifying column as it can be used in conjuction with other columns or datasets to trace the data back to the user, Email column is a sensitive column as it can easily be used to reverse search and find not only the name of the user but a lot of other sensitive information of the user, Amount Spent is again a quasiidentifying column and zipcode is a quasidentifying column as well.

I wanted to set the hierarchies manually as arx has limited support for hierarchies. I then spent a day or two setting up and defining the hierarchies for Amount spent , zipcode and Age.
I divided all of the values of amount spent in four columns such as high, low, medium and low-hig for the first level, i generalized the columns even further.
For zipcode, i scraped the zipcodes for a city from the web and got a list of the areas. I then set up the hierarchy to generalize the zipcode values by zipcode-> area->city->state->country->*
I used the interval based hierarchy in pyarxaas to set up a interval based hierarchy for age. Then i used those csv files to set the generalize the privacy levels for the database.

I experimented with and used privacy models kanonymity, tcloseness and ldiversity to anonymize the customer data and came to the conclusion that identifying columns are better anonymized with tcloseness and ldiversity the rest of the quasiidentifying columns. Then performed the risk analysis for the same dataset

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)
ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

Annexure 1

Enrollment no: 190130111081

STUDENT'S WEEKLY RECORD OF INTERNSHIP

NAME OF STUDENT: Modi Shreshtha Pragnesh

DIARY OF THE WEEK: Dt: _____ **TO** _____

DEPARTMENT: Electronics And Communications Engineering **SEM:** 08

NAME OF THE ORGANISATION: Eternal Soft Solutions

NAME OF THE PLANT/SECTION/DEPARTMENT: Software and Cloud Engineering

NAME OF OFFICER INCHARGE OF THE PLANT/SECTION/DEPARTMENT: Mrs Poonam Patel

DESCRIPTION OF THE WORK DONE IN BRIEF

This week, i focused on making my code more usable and rewrote my code using the principles of Object Oriented programming.
After doing that, I worked making my snowflake account. After the account was created, I read the documentation about the history of snowflake, the architecture of the snowflake, pricing solutions of the snowflake and the capabilites of snowflake.
After that, installed python modules called snowflake connector
and sql alchemy. These modules allow us to access the snowflake warehouse right from the python script by running a connector instance and adding the credentials to the instance such as your name, the password, the warehouse that you want to use and the database and table that you want to use. After running the connector instance, i then uploaded my dataframe which was stored in a local csv file to the snowflake database

then i modified the code to:
add the functions to fetch the data from the snowflake table
set the hierarchy for the fetched data
Anonymize the fetched data
Upload the fetched data into a snowflake table so that it loops through all the tables in the snowflake table wrt the dataframe and only change the modified values in the snowflake table and insert the new rows

I developed two functions to do the same and added a more general function which takes in the data from any data frame and anonymize the same and return the anonymized dataframe

GUJARAT TECHNOLOGICAL UNIVERSITY
(Established under Gujarat Act No. 20 of 2007)

ગુજરાતટેકનોલોજીકલ યુનિવર્સિટી
(ગુજરાત અધિનિયમ ક્રમાંકઃ ૨૦/૨૦૦૭ દ્વારા સ્થાપિત)

TOTAL HOURS: 35

SIGNATURE OF STUDENT

✪ **The above entries are correct and the grading of work done by Trainee is**
EXCELLENT / VERY GOOD / GOOD / FAIR / BELOW AVERAGE / POOR

Signature of Faculty Mentor

Signature of officer-in-charge of Dept. / Section / Plant

Date:

Date:

✪ **Grading of Work, for trainee may be given depending upon your judgement about his Punctuality, Regularity, Sincerity, Interest taken, Work done etc.**

SUPPLEMENTRY NOTES
(add additional sheets if required)