

INDEX

S.NO	TOPIC	PAGE NO.
1.	TITLE	
2.	INDEX	
3.	INTERNSHIP LETTER	
4.	DECLARATION	
5.	ACKNOWLEDGEMENT	
6.	ABOUT TRAINING	
7.	ABOUT UNICLOUD	
8.	OBJECTIVES	
9.	DATA SCIENCE	
10.	MY LEARNINGS	
11.	REASON FOR CHOOSING DATA SCIENCE	
12.	LEARNING OUTCOMES	
13.	SCOPE IN DATA SCIENCE	
14.	RESULTS	

INTERNSHIP LETTER

JioSign Document ID: c0b46181-936b-13f8-8193-776d0d0c0cbd



OFFER LETTER - INTERN

To,
Nishant Goel
House No – 32
Pocket – 13
Sector -20
Rohini
New Delhi

We are pleased to offer you an internship at UNICLOUD at our Noida office. Your internship shall commence on 2nd January 2025 for a period of 6 month. The terms and conditions of your internship with the Company are set forth below:

- Subject to your acceptance of the terms and conditions contained herein, your project and responsibilities during the Term will be determined by the supervisor assigned to you for the duration of the internship.

Code of Conduct and Policies

- As a part of your responsibilities, you will be provided with a company laptop for the duration of your Internship.
- You are responsible for the proper care and maintenance of the laptop. You should handle it with caution keeping it in secure location when not in use.
- You will be bound by the Company's Code of Conduct and all other rules, regulations, policies and orders issued by the Company from time to time in relation to your conduct, discipline and service conditions such as leave, work hours, dress policy, etc., as if these rules, regulations, policies et al, were part of this contract of Internship.
- Without prejudice to the generality of the foregoing, you shall at all times comply with the Company's policies and procedures (as may be intimated from time to time on the Company's intranet or through the Company's newsletters and webcasts or other written means), including but not limited to matters relating to independence, anti-bribery, prevention of insider trading

Regd. Address: Plot No.4, 1st Floor, Sukh Vihar, Delhi – 110051
Corporate Address: B- 8&9 Fifth Floor, Sector-1 Noida-201301 | Ph: +91 120 4157788, +91 9810237910



Notice period and termination

- During the Internship period described above, either party may terminate the employment by giving one month's notice or one month's salary in lieu of notice, to the other party, subject to the release date being approved by the reporting manager.
- The Company may terminate / suspend your services, at its discretion, at any time immediately upon written notice to you, if it has been alleged and prima facie established, through preliminary internal enquiry that you have committed (i) any heinous criminal act or any offence involving moral turpitude (the term "Moral Turpitude" includes crimes having an inherent quality of baseness, vileness, or depravity with respect to a person's duty to the society in general.), (ii) sexual harassment (adjudicated guilty as per the Company's policy and local laws) or (iii) other act that threatens or is likely to damage Company's reputation.

Confidentiality

- Remuneration: You shall at all times keep the details of your salary and employment benefits at the Company strictly confidential and shall not disclose such details to any other person within the Company.
- Use of Company's name: You shall use the Company's name, logo, trademark, or other identifiers strictly in the manner permitted by the Company's policies, or for the purposes of provision of services delegated to you to the extent required. Upon termination of your employment with the Company, you shall not use the Company's name, logo, trademark or other identifiers in any manner other than what is already a matter of public knowledge, provided however, you will not be in breach of this clause if you make reference to the Company's name solely to describe your former association with the Company subject to the confidentiality obligations which the Company might have undertaken in relation to any of its clients, vendors or other Company's personnel.
- Information: You shall always maintain the highest degree of confidentiality with respect to the records, documents and other Confidential Information relating to the business of the Company which may be known to you or confided in you by the Company, its representatives, authorized personnel, vendors, sub-contractors, clients, etc., by any means and you will use such records, documents, and information only in a duly authorized manner in the interest of the Company. For the purposes of this clause, 'Confidential Information' means information about the Company's business and that of its customers, sub-contractors, business partners or agents which is not available to the general public and which may be learnt by you in the course of your employment. This includes, but is not limited to, information relating to the Company, its customer lists, employment policies, personnel, and information about the Company's products, processes including ideas, concepts, projections, technology, manuals, drawings, designs, specifications, sales pitches, fees quotes, tender information, key personnel, customer contacts, thought leadership papers, resumes, records and other documents containing such Confidential Information, whether such information was disclosed to or accessed by you prior to or after the date hereof. You shall at all times, whether during or after the termination of your employment,



act with utmost fidelity and not disclose or divulge Confidential Information to any third party or make use of any such information for your own benefit.

- At no time will you remove any Confidential Information from the Company's offices without the permission of your reporting manager save and except for the purposes of performing the duties assigned to you in your capacity as an employee of the Company. Upon termination of your employment with the Company, or otherwise upon the Company's request, you will immediately surrender to the Company all files, books, magazines, reports, documents, manuals, audio and video tapes, discs, any knowledge databases entrusted to you, and any other data, information or material containing or reflecting Confidential Information. If the Company requests, you shall also confirm in writing to the Company that you have complied with this clause.
- You acknowledge and agree that disclosure of any portion of the Confidential Information or any breach of the provisions herein may result in irreparable injury and damage to the Company which will not be adequately compensable in monetary terms, that the Company will have no adequate remedy at law therefor, and that the Company may, in addition to all other remedies available to it at law or in equity, obtain such preliminary, temporary or permanent mandatory or restraining injunctions, orders or decrees as may be necessary to protect the Company against, or on account of, any breach by you of the provisions contained herein, and you agree to reimburse the reasonable legal fees and other costs incurred by the Company in enforcing the provisions of this Agreement. In addition, the Company will be within its rights to (i) advertise for public knowledge / notice (ii) notify to your prospective employer or regulatory body, any impropriety or breach of confidentiality obligations hereunder, at its absolute discretion.

Personal Data

- During the course of your employment with the Company, you may provide the Company with confidential data or information that can be linked to you personally, or otherwise personally identifies you, including but not limited to your financial information, email address, addresses, telephone numbers, shareholdings, physiological and/or mental health information, and medical records and history (your Personal Data). You acknowledge that the Company may collect, use, transfer, store or otherwise process (Process) such Personal Data as required per the Company's policies, to facilitate the conduct of the Company's business, to conduct background checks, check conflicts or maintain independence, finance and accounting purposes or for quality and risk management purposes. Company will Process your Personal Data in accordance with applicable law and professional obligations and shall ensure that any service provider who Processes Personal Data on our behalf adheres to such requirements. You hereby consent to the Processing of your Personal Data in the manner described above, whether by the Company or any service provider on the Company's behalf.

Please confirm your acceptance of the terms of this Internship offer by tomorrow failing which, we have the right to cancel the internship. We look forward to having you on our team!

IN WITNESS WHEREOF, the parties hereto have executed this Agreement as of the Effective Date.

Regd. Address: Plot No.4, 1st Floor, Sukh Vihar, Delhi – 110051
Corporate Address: B- 8&9 Fifth Floor, Sector-1 Noida-201301 | Ph: +91 120 4157788, +91 9810237910



UNICLOUD *Sandeep Goel*

By: _____
Signed on jiosign.com with eSignature

Print Name: Sandeep Goel
29.11.2024 15:55 IST
Sandeep Goel

Intern

Nishant Goel

Signature: _____
Signed on jiosign.com with eSignature

Print Name: Nishant Goel
29.11.2024 21:13 IST
Nishant Goel

DECLARATION

I, Nishant Goel, solemnly declare that the information provided in this internship report is true, accurate, and reflective of my internship experience at UNICLOUD.

I affirm that the projects, tasks, and responsibilities outlined in this report accurately represent my contributions and experiences during the internship period. Any references or citations included in this report are duly acknowledged and attributed to their respective sources.

Furthermore, I declare that this internship report is the result of my own work and efforts, with appropriate guidance and supervision from my internship supervisor at UNICLOUD.

I understand that any misrepresentation or falsification of information in this report may have serious consequences, including academic penalties and damage to my reputation.

By signing this declaration, I affirm my commitment to academic integrity and honesty in all aspects of my academic and professional endeavors.

Date: _____

Signature: _____

ACKNOWLEDGEMENT

I take this opportunity to express my sincere thanks and deep gratitude to all those people who extended their wholehearted cooperation and have helped me in completing this internship successfully.

First, I would like to thank our supervisor Mr. Prakash Chandra Saraswat, who mentored me, guided me and challenged me.

I also thank my family and friends who greatly supported me during the Internship.

Last but not the least, I would like to thank our founders for considering me a part of the organization and provide such a great Platform to learn and enhance my skills.

A very special thanks goes to all the faculties of Gautam Buddha University, Greater Noida under whom guidance I have been able to excel in my career and become a part of the Watchguard family.

Nishant Goel

235PMD001

Gautam Buddha University, Greater Noida

ABOUT COMPANY

Unicloud is one of the fastest growing Cloud & AI Consulting company, we strive to help our customers maximize returns on their cloud investments. We provide platform and expertise across the cloud lifecycle starting with Assessment, Migrations, Deployment, Optimization and SRE. In last few years we have helped over 100 enterprises realize \$ 44Mn+ in cloud savings.

Why choose UNICLOUD?

1. Cloud-Native Infrastructure – Scalable and flexible cloud environment for handling large datasets and machine learning workloads.
2. AI & Data Science Integration – Built-in AI tools, automated ML training, and real-time data analytics support.
3. Cost-Effective Solutions – Competitive pricing compared to AWS, Azure, and Google Cloud, making it ideal for startups and researchers.
4. Security & Compliance – Strong security framework with compliance support (e.g., GDPR, HIPAA) for safe data storage and processing.
5. Developer-Friendly Tools – Easy-to-use APIs, SDKs, and managed environments for deploying ML models without deep infrastructure expertise.
6. Industry-Specific Solutions – Tailored cloud services for sectors like finance, healthcare, and IoT, offering optimized performance.
7. Collaboration & Community Support – Active developer community, detailed documentation, and training programs for easy adoption.

Internship Objectives

The objective of this internship was to gain practical experience in data science within a cloud-based environment. My primary focus areas included:

- Understanding cloud-based data processing and storage.
- Working with large datasets hosted on cloud platforms.
- Applying machine learning algorithms to real-world problems.
- Collaborating with cross-functional teams to extract meaningful insights from data.

Roles and Responsibilities

During my internship, I was assigned the following key responsibilities:

1. **Data Preprocessing & Cleaning** – Extracting, cleaning, and preprocessing raw data stored on cloud databases.
2. **Exploratory Data Analysis (EDA)** – Performing statistical and visual analysis to understand data trends.

3. **Machine Learning Implementation** – Developing predictive models using Python and cloud-based ML services (e.g., AWS SageMaker, Google Vertex AI, or Azure ML).
4. **Data Visualization** – Creating dashboards and reports using tools like Power BI, Tableau, and Matplotlib.
5. **Cloud Integration** – Deploying data pipelines and models on cloud platforms such as AWS, Google Cloud, or Azure.
6. **Collaboration & Documentation** – Working with data engineers, software developers, and business analysts to ensure data-driven decision-making.

What is Data Science?

Data Science as a multi-disciplinary subject that uses mathematics, statistics, and computer science to study and evaluate data. The key objective of Data Science is to extract valuable information for use in strategic decision making, product development, trend analysis, and forecasting. Data Science concepts and processes are mostly derived from data engineering, statistics, programming, social engineering, data warehousing, machine learning, and natural language processing. The key techniques in use are data mining, big data analysis, data extraction and data retrieval.

Data science is the field of study that combines domain expertise, programming skills, and knowledge of mathematics and statistics to extract meaningful insights from data. Data science practitioners apply machine learning algorithms to numbers, text, images, video, audio, and more to produce artificial intelligence (AI) systems to perform tasks that ordinarily require human intelligence. In turn, these systems generate insights which analysts and business users can translate into tangible business value.

DATA SCIENCE PROCESS:

1. The first step of this process is setting a research goal. The main purpose here is making sure all the stakeholders understand the what, how, and why of the project.
2. The second phase is data retrieval. You want to have data available for analysis, so this step includes finding suitable data and getting access to the data from the data owner. The result is data in its raw form, which probably needs polishing and transformation before it becomes usable.
3. Now that you have the raw data, it's time to prepare it. This includes transforming the data from a raw form into data that's directly usable in your models. To achieve this, you'll detect and correct different kinds of errors in the data, combine data from different data sources, and transform it. If you have successfully completed this step, you can progress to data visualization and modelling.
4. The fourth step is data exploration. The goal of this step is to gain a deep understanding of the data. You'll look for patterns, correlations, and deviations based on visual and descriptive techniques. The insights you gain from this phase will enable you to start modelling.
5. Finally, we get to the sexiest part: model building (often referred to as "data modelling" throughout this book). It is now that you attempt to gain the insights or make the predictions stated in your project charter. Now is the time to bring out the heavy guns,

but remember research has taught us that often (but not always) a combination of simple models tends to outperform one complicated model. If you've done this phase right, you're almost done.

6. The last step of the data science model is presenting your results and automating the analysis, if needed. One goal of a project is to change a process and/or make better decisions. You may still need to convince the business that your findings will indeed change the business process as expected. This is where you can shine in your influencer role. The importance of this step is more apparent in projects on a strategic and tactical level. Certain projects require you to perform the business process over and over again, so automating the project will save time.

MY LEARNING

1) INTRODUCTION TO DATA SCIENCE

- Overview & Terminologies in Data Science
- Applications of Data Science
 - Unfamiliar detection (fraud, disease, etc.)
 - Automation and decision-making (credit worthiness, etc.)
 - Classifications (classifying emails as “important” or “junk”)
 - Forecasting (sales, revenue, etc.)
 - Pattern detection (weather patterns, financial market patterns, etc.)
 - Recognition (facial, voice, text, etc.)
 - Recommendations (based on learned preferences, recommendation engines can refer you to movies, restaurants and books you may like)

2) PYTHON FOR DATA SCIENCE

Introduction to Python, Understanding Operators, Variables and Data Types, Conditional Statements, Looping Constructs, Functions, Data Structure, Lists, Dictionaries, Understanding Standard Libraries in Python, reading a CSV File in Python, Data Frames and basic operations with Data Frames, Indexing Data Frame.

3) UNDERSTANDING THE STATISTICS FOR DATA SCIENCE

Introduction to Statistics, Measures of Central Tendency, Understanding the spread of data, Data Distribution, Introduction to Probability, Probabilities of Discrete and Continuous Variables, Normal Distribution, Introduction to Inferential Statistics, Understanding the Confidence Interval and margin of error, Hypothesis Testing, Various Tests, Correlation.

4) PREDICTIVE MODELING AND BASICS OF MACHINE LEARNING

Introduction to Predictive Modelling, Types and Stages of Predictive Models, Hypothesis Generation, Data Extraction and Exploration, Variable Identification, Univariate Analysis for Continuous Variables and Categorical Variables, Bivariate Analysis, Treating Missing Values and Outliers, Transforming the Variables, Basics of Model Building, Linear and Logistic Regression, Decision Trees, K-means Algorithms in Python.

Summary of Procedure of Analysing Data:

Data science generally has a five-stage life cycle that consists of:

- Capture: data entry, signal reception, data extraction
- Maintain: Data cleansing, data staging, data processing.
- Process: Data mining, clustering/classification, data modelling
- Communicate: Data reporting, data visualization
- Analyse: Predictive analysis, regression

Introduction to Data Science

Data Science

The field of bringing insights from data using scientific techniques is called data science.

Applications

Amazon Go – No checkout lines

Computer Vision - The advancement in recognizing an image by a computer involves processing large sets of image data from multiple objects of same category. For example, Face recognition.

Reporting / Management Information System

To track what is happening in organization.

Detective Analysis

Asking questions based on data we are seeing, like. Why something happened?

Dashboard / Business Intelligence

Utopia of reporting. Every action about business is reflected in front of screen.

Predictive Modelling

Using past data to predict what is happening at granular level. Big Data

Stage where complexity of handling data gets beyond the traditional system.

Can be caused because of volume, variety or velocity of data. Use specific tools to analyse such scale data.

Application of Data Science

• Recommendation System

Example-In Amazon recommendations are different for different users according to their past search.

• Social Media

1. Recommendation Engine
2. Ad placement
3. Sentiment Analysis

• Deciding the right credit limit for credit card customers.

• Suggesting right products from e-commerce companies

1. Recommendation System
2. Past Data Searched
3. Discount Price Optimization

• How google and other search engines know what are the more relevant results for our

search query?

1. Apply ML and Data Science
2. Fraud Detection
3. AD placement
4. Personalized search results

What is Python?

Python is an interpreted, high-level, general-purpose programming language. It has efficient high-level data structures and a simple but effective approach to object-oriented programming. Python's elegant syntax and dynamic typing, together with its interpreted nature, make it an ideal language for scripting and rapid application development in many areas on most platforms.

Python for Data science:

Why Python???

1. Python is an open source language.
2. Syntax as simple as English.
3. Very large and Collaborative developer community.
4. Extensive Packages.

• UNDERSTANDING OPERATORS:

Theory of operators: - Operators are symbolic representation of Mathematical tasks.

• VARIABLES AND DATATYPES:

Variables are named bounded

d to objects. Data types in python are int (Integer), Float, Boolean and strings.

• CONDITIONAL STATEMENTS:

If-else statements (Single condition)

If- Elif- else statements (Multiple Condition)

• LOOPING CONSTRUCTS:

For loop

• FUNCTIONS:

Functions are re-usable piece of code. Created for solving specific problem.

Two types: Built-in functions and User- defined functions.

Functions cannot be reused in python.

• **DATA STRUCTURES:**

Two types of Data structures:

LISTS: A list is an ordered data structure with elements separated by comma and enclosed within square brackets.

DICTIONARY: A dictionary is an unordered data structure with elements separated by comma and stored as key: value pair, enclosed with curly braces {}.

Statistics

Descriptive Statistic

Mode

It is a number which occurs most frequently in the data series.

It is robust and is not generally affected much by addition of couple of new values.

Code :

```
import pandas as pd
data=pd.read_csv( "Mode.csv")    //reads data from csv file
data.head()                      //print first five lines
mode_data=data['Subject'].mode() //to take mode of subject column
print(mode_data)
```

Mean

```
import pandas as pd
data=pd.read_csv( "mean.csv")    //reads data from csv file
data.head()                      //print first five lines
mean_data=data[Overallmarks].mean() //to take mode of subject column
print(mean_data)
```

Median

Absolute central value of data set.

```
import pandas as pd
```



```
data=pd.read_csv( "data.csv")    //reads data from csv file
data.head()                      //print first five lines
median_data=data[Overallmarks].median() //to take mode of subject column
print(median_data)
```

Types of variables

- Continous – Which takes continuous numeric values. Eg-marks
- Categorical-Which have discrete values. Eg- Gender
- Ordinal – Ordered categorical variables. Eg- Teacher feedback
- Nominal – Unorderd categorical variable. Eg- Gender

Algorithm for Machine Learning

Machine learning algorithms can be categorized into different types based on their learning approach. Here are some key algorithms:

Supervised Learning Algorithms

(Uses labeled data for training)

1. Linear Regression – Used for predicting continuous values (e.g., house price prediction).
2. Logistic Regression – Used for classification problems (e.g., spam detection).
3. Decision Trees – Uses tree-like structures for decision-making.
4. Random Forest – An ensemble of decision trees for better accuracy.
5. Support Vector Machine (SVM) – Finds the optimal boundary for classification tasks.
6. K-Nearest Neighbours (KNN) – Classifies data based on the closest neighbours.
7. Neural Networks (ANN, CNN, RNN) – Used in deep learning applications like image recognition and NLP.

Unsupervised Learning Algorithms

(Finds patterns in unlabelled data)

1. K-Means Clustering – Groups similar data points into clusters.
2. Hierarchical Clustering – Creates a tree-like cluster structure.
3. Principal Component Analysis (PCA) – Reduces dimensionality of data.

4. Autoencoders – Used in deep learning for data compression and feature extraction.

Reinforcement Learning Algorithms

(Learns by interacting with an environment)

1. Q-Learning – Uses a Q-table to learn optimal policies.
2. Deep Q-Network (DQN) – Combines Q-learning with deep learning.
3. Policy Gradient Methods – Learns policies directly using gradient ascent.

Conclusion

My internship at UNICLOUD has been a transformative learning experience. It provided me with an in-depth understanding of data science applications in cloud environments. The hands-on experience with cloud computing, data pipelines, and machine learning models has enhanced my technical and professional skills. I am grateful for this opportunity and look forward to applying my learnings in future endeavours.

Signature:

Nishant Goel

Date: 05-03-2025