

# INTERNSHIP DIARY

**Student Name** : VIGHNESH

**University Seat Number** : 4MW21AI058

**Academic Year** : 2024-25

**Department** : AIML

**Internal Guide** : Dr. Ganesha Prasad

**External Guide** : Kaveri

**Designation & Department** : Software developer

**Company/ Organization** : DLithe Consultancy Services Pvt Ltd

**Address** : No. 280, SLV Arcade, Dr Vishnuvardhan Rd, Phase 3,  
Banashankari, Bengaluru, Karnataka 560085, India

**Internship Dates** : From 18/02/2025 To 17/05/2025



**Department of Artificial Intelligence and Machine Learning Engineering**  
**Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal Udupi**

2024-25

### DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 1**

| Date  | Activity  | Signature of the External Guide |
|-------|---|---------------------------------|
| 18/02 | Introduction to the Internship and Organization. Gained insights into company goals, mission, and AI's role in industries. Understood industry expectations and internship structure. Explored how AI is applied in real-world scenarios and the significance of data-driven decision-making.   |                                 |
| 19/02 | Introduction to AI & Python Variables. Learned AI history, evolution and Python variable declaration rules. Explored naming conventions and common mistakes in variable handling. Understood how Python variables handle memory and how dynamic typing plays a role in data manipulation.   |                                 |
| 20/02 | Exploring Data Types & Operators. Practiced integer, float, string, boolean, and None types. Used arithmetic and logical operators. Learned operator precedence and debugging techniques. Experimented with type conversions and implicit/explicit type casting. Encountered issues with floating-point precision and researched solutions.             |                                 |
| 21/02 | Control Structures (if-else, loops). Implemented decision-making structures with real-world applications. Faced minor debugging challenges in nested conditions. Explored multiple use cases of loops including iteration over collections and implementing search algorithms. Understood the importance of proper indentation and logical structuring. |                                 |
| 22/02 | Basic Python Assignments: -Completed programs on eligibility checks and arithmetic operations. Focused on input validation and handling edge cases. Developed a voting eligibility program with enhanced exception handling and user-friendly messages.   |                                 |
|       |   |                                 |

**Signature of the Internal Guide**

**DAILY ACTIVITY REPORT****Student Name: VIGHNESH****USN:4MW21AI058****Name of the Organization: DLithe Consultancy Services Pvt Ltd****Week No: 2**

| <b>Date</b> | <b>Activity</b>   | <b>Signature of the External Guide</b> |
|-------------|---|--|
| 24/02       | Introduction to SDLC. Understood different phases of SDLC and its relevance in AI projects. Explored real-world industry applications of SDLC. Analyzed how different models impact AI development and how Agile methodologies benefit iterative improvements.              |  |
| 25/02       | SDLC Phases & Case Studies. Studied requirement analysis, design, testing, and deployment. Differentiated overlapping roles within SDLC phases. Researched case studies on how software failures were linked to poor SDLC implementation.                                   |  |
| 26/02       | SDLC Models (Waterfall, Agile, Spiral). Compared models and their advantages/disadvantages. Learned adaptability of models in AI/ML projects. Explored how Agile can be integrated into AI product development cycles and how iterative feedback helps in model improvement |  |
| 27/02       | Control Structures in Python. Implemented loops (for, while, break, continue, pass). Debugged common loop-related issues and optimized loop execution. Worked on an optimization problem where reducing loop iterations improved execution speed.                           |  |
| 28/02       | Data Structures (Lists, Tuples, Sets, Dictionaries). Understood different structures and their practical applications. Learned selection of optimal data structures for efficiency. Explored the time complexities of different operations and their impact on performance. |  |
| 01/03       | Assignments on SDLC & Python Concepts. Implemented Python programs using control structures and loops. Ensured modularization and efficiency in coding.   |  |

**Signature of the Internal Guide**

### DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 3**

| Date  | Activity  | Signature of the External Guide |
|-------|---|---------------------------------|
| 03/03 | Functions in Python. Learned about modular programming and importance of functions. Explored function reusability and efficiency. Worked on structuring a program into reusable modules.  |                                 |
| 04/03 | Function Declaration & Parameters. Explored function arguments, return values, and best practices. Handled multiple return values effectively. Implemented keyword arguments and default parameters to improve function usability.          |                                 |
| 05/03 | Writing Functions for Various Operations. Developed functions for mathematical, string, and logical operations. Debugged incorrect outputs in return handling. Explored lambda functions for concise and functional programming approaches. |                                 |
| 06/03 | Recursion in Python. Implemented factorial and Fibonacci series using recursion. Avoided infinite recursion and stack overflow errors. Analyzed the efficiency of recursive vs. iterative approaches.                                       |                                 |
| 07/03 | Object-Oriented Programming (OOP). Explored encapsulation, inheritance, polymorphism, and abstraction. Understood real-world applications of OOP. Created a class-based implementation for a vehicle management system.                     |                                 |
| 08/03 | Performed tasks on Functions & OOP. Implemented recursion, OOP concepts, and data type operations. Focused on optimizing recursion memory allocation. Applied design patterns to improve code structure and reusability.                    |                                 |

**Signature of the Internal Guide**

**DAILY ACTIVITY REPORT****Student Name: VIGHNESH****USN:4MW21AI058****Name of the Organization: DLithe Consultancy Services Pvt Ltd****Week No: 4**

| <b>Date</b> | <b>Activity</b>   | <b>Signature of the External Guide</b> |
|-------------|---|--|
| 10/03       | Standup Meeting – Review of last week’s topics.Discussed real-world use cases of Python programming in AI and ML.Explored how functions, recursion, and control structures are applied in practical scenarios.Analyzed different coding problems and their optimized solutions.   |  |
| 11/03       | Project Allocation: Assigned an AIML-based Product Recommendation System for E-commerce Websites. Understood the project scope, expected deliverables, and technical requirements.  |  |
| 12/03       | Introduction to NumPy – Basics of arrays, array operations, and vectorized computations. Implemented NumPy operations for handling large datasets efficiently. Hands-on practice with matrix operations, reshaping arrays, and broadcasting concepts.   |  |
| 13/03       | Introduction to Pandas – Understanding Series and DataFrames. Performed data cleaning, manipulation, and analysis using Pandas. Practical exercises on loading datasets, handling missing values, and filtering data.   |  |
| 14/03       | Advanced Pandas operations and data manipulation:Used .apply() function to modify data dynamically based on conditions (e.g., salary-based bonus calculations).Learned how to perform merging and joining of DataFrames using merge() with different join types: inner, left, and right.Discussed display options customization (e.g., setting max_rows) and reading JSON files for structured data handling. |  |
| 15/03       | Assignment on Fundamentals of Data Science using NumPy and JSON The assignment involved practical implementation of core concepts from the NumPy and JSON modules. Tasks included creating and manipulating multidimensional arrays, performing mathematical operations on matrices, reshaping arrays, filtering data based on conditions, and converting data formats.                                       |  |

**Signature of the Internal Guide**

### DAILY ACTIVITY REPORT

**Student Name: VIGNESH**

**USN:4MW21AI058**

**Name of the Organization: DLitthe Consultancy Services Pvt Ltd**

**Week No: 5**

| Date  | Activity  | Signature of the External Guide |
|-------|---|---------------------------------|
| 17/03 | Initiated the AI/ML-Based Product Recommendation System for E-commerce Websites project. Understood the project goals, scope, and the importance of recommendation engines in the e-commerce domain. Outlined key project modules including data collection, preprocessing, recommendation engine development, and integration.   |                                 |
| 18/03 | Focused on data collection from publicly available datasets on Kaggle. Initially began working with a large dataset containing approximately 2.6 million records covering user behavior, product details, and sessions. Faced challenges in managing such a large dataset on local machines due to performance limitations and processing overhead. As a result, shifted to a smaller yet meaningful dataset subset to ensure efficient handling, quicker iterations, and better development speed. |                                 |
| 19/03 | Attended a session on JSON (JavaScript Object Notation), a widely used data format in web applications. Learned about JSON structure, data types (string, number, boolean, null, array, object), and its role in data exchange between client-server systems. Practical session on loading and reading JSON data into Pandas DataFrames using read_json() method.   |                                 |
| 20/03 | Focused on data cleaning techniques using Pandas. Tasks included: Identifying and handling missing/null values using dropna() and fillna(), Detecting and removing duplicates, Addressing format inconsistencies and incorrect data entries and Understanding special placeholders like NaN (Not a Number) and NaT (Not a Time) in datasets.  |                                 |
| 21/03 | Concentrated on data preprocessing techniques required before implementing machine learning models. Applied operations like: Column normalization and feature engineering. Addressed practical project challenges: Due to the initial large dataset (2.6M records), processing and memory management became difficult on the local system. Decided to proceed with a lighter version of the dataset to maintain performance efficiency while retaining data quality for recommendation logic.       |                                 |

**Signature of the Internal Guide**

### DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 6**

| Date  | Activity  | Signature of the External Guide |
|-------|---|---------------------------------|
| 24/03 | Explored Matplotlib, its applications, and why it's widely used for data visualization. Learned about basic plotting functions, customizing graphs, and adding labels/titles. Practiced creating line plots and scatter plots to visualize trends in datasets.  |                                 |
| 25/03 | Continued working with Matplotlib and explored different plot customizations. Learned how to modify line styles, markers, colors, and grid lines. Hands-on practice on multiple plots in a single figure using subplots.  |                                 |
| 26/03 | Presentation on Matplotlib's advanced features: <ul style="list-style-type: none"> <li>• Subplots &amp; Grid Systems – Arranging multiple plots efficiently.</li> <li>• Labeling Axes &amp; Legends – Making plots more readable.</li> <li>• Bar Charts, Pie Charts, and Histograms – Visualizing categorical and distribution data. Explored real-world use cases of Matplotlib in data science.</li> </ul>  |                                 |
| 27/03 | Introduction to Seaborn – an advanced Python visualization library. Compared Matplotlib vs. Seaborn and understood why Seaborn is preferred for statistical visualization. Created heatmaps, box plots, and violin plots for better data insights. Explored Seaborn's built-in datasets and how to customize plots.   |                                 |
| 28/03 | Overview of Machine Learning, its types, and real-world applications. Covered Supervised, Unsupervised, and Reinforcement Learning. Discussed the Machine Learning Workflow: <ul style="list-style-type: none"> <li>• Data Collection &amp; Preprocessing</li> <li>• Feature Engineering</li> <li>• Model Selection &amp; Training</li> <li>• Evaluation &amp; Deployment</li> <li>• Brief introduction to popular ML libraries like Scikit-learn.</li> </ul> |                                 |

**Signature of the Internal Guide**

### DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLitthe Consultancy Services Pvt Ltd**

**Week No: 7**

| Date | Activity   | Signature of the External Guide |
|------|--|---------------------------------|
|      | I worked on handling missing values in a housing price prediction dataset. I used functions like <code>isna()</code> and <code>isnull()</code> to identify missing values and used <code>notna()</code> and <code>notnull()</code> to check for valid entries. I also calculated the total and column-wise count of null values. After identifying missing data, I performed label encoding on categorical columns to prepare the data for modeling. I then selected suitable machine learning models, evaluated them using metrics like accuracy and RMSE, and created visualizations to better understand the dataset.   |                                 |
|      | I focused on cleaning the dataset further by handling missing data. I began by identifying missing values using <code>isna()</code> and <code>isnull()</code> , which helped in understanding which columns required attention. I explored two main approaches: removing missing values using <code>dropna()</code> and filling them with appropriate replacements using <code>fillna()</code> . Depending on the nature of the data, I used statistical methods like replacing null values with the mean, median, or mode. In cases where temporal or sequential data was involved, I also considered forward-fill and backward-fill methods. This step was essential to ensure the dataset was complete and ready for further processing without introducing bias. |                                 |
|      | The session focused on encoding techniques used to convert categorical data into numerical form, which is required for most machine learning models. I learned about four major types of encoding: label encoding, which assigns each category a unique number; one-hot encoding, which creates binary columns for each category; ordinal encoding, used when categories have a meaningful order; and binary encoding, which is a combination of hashing and one-hot encoding. These techniques were implemented based on the type and importance of the categorical features, ensuring the model could interpret them correctly without misrepresenting relationships.  |                                 |
|      | The focus shifted to identifying and handling outliers. I explored methods such as the Z-score, IQR (Interquartile Range), and visual techniques like box plots to detect abnormal values that could skew the analysis or affect model performance. These outliers were either removed or treated to improve data quality.   |                                 |

**Signature of the Internal Guide**



## DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 8**

| Date | Activity   | Signature of the External Guide |
|------|--|---------------------------------|
|      | Worked on data preprocessing and basic plotting using the wine dataset (wine.csv). I cleaned the dataset by checking for missing values, understanding data types, and ensuring consistency in feature values. I then generated visualizations such as histograms, box plots, and correlation heatmaps to explore the distribution of features and identify any outliers or trends in the data. These steps helped lay a solid foundation for further analysis and model building. |                                 |
|      | I practiced making predictions using a trained model, applying both manual understanding and programming logic. I used the command <code>y_predicted = model.predict(x)</code> to generate predictions and then printed the intercept, slope (coefficient values), and the predicted values. This session helped me understand how linear regression works under the hood by interpreting the model's output and how each feature contributes to the target variable.              |                                 |
|      | We were given some tasks to complete. I worked on applying my previous learnings on data preprocessing and prediction models, revisiting both the wine dataset and earlier examples. I also reinforced my understanding of how predictions are evaluated and interpreted through coding exercises and mentor feedback.   |                                 |
|      | we gave a presentation on data preprocessing, using the Stock Price dataset that had been worked on previously. I presented the steps such as handling missing values, extracting date features (day, month, year), encoding categorical data, and organizing the dataset for visualization and model training. It was a good opportunity to communicate our approach clearly and apply the structured format taught by our mentor.  |                                 |
|      | Presentation focusing on data visualization. I showcased various plots we used to explore stock price trends over time, compare volumes, and detect outliers. I explained the use of line plots, bar charts, and box plots to visually interpret insights. This helped strengthen both my technical and presentation skills, especially in explaining the significance of data visuals to a non-technical audience.  |                                 |

**Signature of the Internal Guide**

# DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 9**

| Date | Activity   | Signature of the External Guide |
|------|--|---------------------------------|
|      | Practiced applying transformation techniques such as log and square root transformations to reduce skewness and normalize data for better results. |                                 |
|      | Worked on handling missing data in real-world datasets using various techniques like mean/median imputation and removal of null values.            |                                 |
|      | Applied encoding techniques such as Label Encoding and One-Hot Encoding to convert categorical data into numerical form for model compatibility.   |                                 |
|      | Implemented feature scaling methods like MinMaxScaler and StandardScaler to bring data into a uniform scale, ensuring consistency in training.     |                                 |
|      | Combined all preprocessing techniques (missing values, encoding, scaling, and transformations) into a single workflow on multiple CSV files        |                                 |

**Signature of the Internal Guide**

**DAILY ACTIVITY REPORT**

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 10**

| Date | Activity   | Signature of the External Guide |
|------|--|---------------------------------|
|      | Learned about skewed data distributions and how they affect model performance. Explored examples showing how skewness leads to inaccurate predictions in machine learning. |                                 |
|      | Practiced applying transformation techniques such as log and square root transformations to reduce skewness and normalize data for better results.                         |                                 |
|      | Applied encoding techniques such as Label Encoding and One-Hot Encoding to convert categorical data into numerical form for model compatibility.                           |                                 |
|      | Combined all preprocessing techniques (missing values, encoding, scaling, and transformations) into a single workflow on multiple CSV files.                               |                                 |
|      | Worked on handling missing data in real-world datasets using various techniques like mean/median imputation and removal of null values.                                    |                                 |

**Signature of the Internal Guide**

**DAILY ACTIVITY REPORT**

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 11**

| <b>Date</b> | <b>Activity</b>  | <b>Signature of the External Guide</b> |
|-------------|--|--|
|             | Built and compared Logistic Regression, Decision Tree, and Random Forest classifiers.            |  |
|             | Learned about cross-validation and used cross_val_score to evaluate model reliability.           |  |
|             | Practiced evaluation metrics—accuracy, precision, recall, and F1-score—on model outputs.         |  |
|             | Used Label Encoding to preprocess data and trained models using the transformed dataset.         |  |
|             | Measured and compared model performance using accuracy_score and interpreted prediction results. |  |

**Signature of the Internal Guide**

## DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLitthe Consultancy Services Pvt Ltd**

**Week No: 12**

| Date | Activity   | Signature of the External Guide |
|------|--|---------------------------------|
|      | Studied the concept and working of Histogram-Based Gradient Boosting (HGBT).                                   |                                 |
|      | Implemented HGBT using Scikit-learn and evaluated performance on classification tasks.                         |                                 |
|      | Learned about LightGBM, its advantages over traditional boosting, and installed the LightGBM library.          |                                 |
|      | Trained the LGBMClassifier and compared its performance with HGBT using precision and recall.                  |                                 |
|      | Summarized findings from both boosting techniques and discussed model selection strategies for large datasets. |                                 |

**Signature of the Internal Guide**

### DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 13**

| Date | Activity   | Signature of the External Guide |
|------|--|---------------------------------|
|      | <ul style="list-style-type: none"> <li>Got introduced to the House Price Prediction project.</li> <li>Reviewed the project's purpose: predicting house prices using ML based on location, BHK, and area.</li> <li>Studied the project structure, including files like app.py, train_model.py, and Cleaned_data.csv.</li> </ul> |                                 |
|      | <ul style="list-style-type: none"> <li>Installed Python environment, Flask, pandas, scikit-learn, and other required libraries.</li> <li>Set up and activated virtual environment (venv).</li> <li>Verified the availability of model file RidgeModel.pkl and data Cleaned_data.csv.</li> </ul>                                |                                 |
|      | <ul style="list-style-type: none"> <li>Analyzed the Cleaned_data.csv dataset — observed structure, features, and missing values.</li> <li>Understood why features like 'Location', 'BHK', and 'Area' were chosen for prediction.</li> </ul>  |                                 |
|      | <ul style="list-style-type: none"> <li>Learned how to run the Flask app via app.py and tested http://localhost:5000 in the browser.</li> <li>Explored the interface and used it to input house details and observe predictions.</li> </ul>   |                                 |
|      | <ul style="list-style-type: none"> <li>Understood AJAX-based form submission and how it improves user experience.</li> <li>Reviewed how Flask routes work and handled prediction requests on the backend.</li> </ul>   |                                 |

**Signature of the Internal Guide**

### DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 14**

| Date | Activity   | Signature of the External Guide |
|------|--|---------------------------------|
|      | <ul style="list-style-type: none"> <li>Deep-dived into train_model.py — understood data preprocessing steps.</li> <li>Learned how OneHotEncoder is used for categorical encoding (location).</li> </ul>                                      |                                 |
|      | <ul style="list-style-type: none"> <li>Studied StandardScaler for normalizing numerical features like BHK and Area.</li> <li>Understood how data is split into features and target variable (price).</li> </ul>                              |                                 |
|      | <ul style="list-style-type: none"> <li>Trained the Ridge Regression model using scikit-learn.</li> <li>Generated and saved the model as RidgeModel.pkl using joblib/pickle.</li> </ul>   |                                 |
|      | <ul style="list-style-type: none"> <li>Evaluated model performance and discussed potential limitations (e.g., minimal input validation).</li> <li>Understood how model is reused in app.py for live prediction.</li> </ul>                   |                                 |
|      | <ul style="list-style-type: none"> <li>Compared Ridge Regression with other regression techniques (Linear, Lasso).</li> <li>Documented the pros of Ridge Regression in this project (e.g., better performance with collinearity).</li> </ul> |                                 |

**Signature of the Internal Guide**

### DAILY ACTIVITY REPORT

**Student Name: VIGHNESH**

**USN:4MW21AI058**

**Name of the Organization: DLithe Consultancy Services Pvt Ltd**

**Week No: 15**

| Date | Activity   | Signature of the External Guide |
|------|--|---------------------------------|
|      | <ul style="list-style-type: none"> <li>Analyzed index.html to understand how Bootstrap and JavaScript are used for UI and AJAX.</li> <li>Investigated how the dropdown list for location is dynamically populated from the dataset</li> </ul>      |                                 |
|      | <ul style="list-style-type: none"> <li>Implemented simple frontend validation for numeric fields (BHK and Area).</li> <li>Explored how AJAX helps submit the form without refreshing the page</li> </ul>   |                                 |
|      | <ul style="list-style-type: none"> <li>Enhanced UI layout and styling using Bootstrap classes.</li> <li>Tested the application with various input scenarios, including edge cases</li> </ul>   |                                 |
|      | <ul style="list-style-type: none"> <li>Brainstormed and documented future improvements: bathroom count, age of property, error handling.</li> <li>Tested model robustness against unknown locations and invalid inputs.</li> </ul>                 |                                 |
|      | <ul style="list-style-type: none"> <li>Summarized project understanding and created a flow diagram for app structure and data flow.</li> <li>Prepared progress notes for review and suggested possible deployment options (Heroku/AWS).</li> </ul> |                                 |

**Signature of the Internal Guide**



ACTIVITY SUMMARY

**Student Name:**VIGHNESH

**USN:** 4MW21AI058

**Name of the Organization:** DLithe Consultancy Services Pvt Ltd

**Nature of the Industry:**

**Intern duties:** What were your specific duties and responsibilities during the internship?

**Learning:** What did you learn during the internship (please include specific skills and business/organizational concepts)?

**Guide Rating:** Was your guide allotted in the industry helpful and instructive? Please give reasons and examples for your answer.

**Internship Summary:**

**Signature of the Internal Guide**

**Signature of Head of the Department**