INTERNSHIP REPORT

ON

MACHINE LEARNING (PYTHON)

AT

EVOLVIER CENTRE

By

ALBIN SABU (MLM23MCA-2008)

Under the guidance of

MR. VISHNU PANICKER

INTERSHIP DURATION: 3 MONTHS

DEPARTMENT OF COMPUTER APPLICATIONS

MANGALAM COLLEGE OF ENGINEERING, ETTUMANOOR

(Affiliated to APJ Abdul Kalam Technological University)

APRIL 2025

ACKNOWLEDGMENT

I would like to express my sincere gratitude to Evolvier centre Learning Centre for providing me with the opportunity to intern and enhance my knowledge. I extend my heartfelt appreciation to my mentor, Vishnu Panicker, for their continuous guidance and support throughout the internship. Their valuable insights and encouragement played a crucial role in my learning process.

ABSTRACT

This report highlights my internship experience at Evolvier, where I worked extensively on Python programming, machine learning, natural language processing (NLP), and web development using Django. Throughout the internship, I gained practical experience in data preprocessing, implementing machine learning models, and developing a functional web application using Django. The internship helped me bridge the gap between theoretical knowledge and practical application, providing me with insights into real-world challenges and solutions in AI and web development.

INTRODUCTION

The objective of this internship was to gain practical exposure to core technologies such as Python, machine learning, NLP, and Django for web development. These technologies are widely used in the AI and software development industry for building intelligent applications that can process and analyze data efficiently. The internship aimed to provide hands-on experience in implementing machine learning models, understanding NLP techniques, and building a simple web application using Django.

TOPICS COVERED

1. Python

Python is a powerful programming language widely used in AI, data science, and web development. Python served as the foundational language for my internship. I began by learning the fundamental concepts, such as data types, loops, functions, and object-oriented programming. Additionally, I explored various libraries like NumPy and Pandas, which are essential for data manipulation and analysis. During my internship, I focused on the following key aspects of Python:

- Understanding basic syntax, loops, conditional statements, and functions.

- Exploring object-oriented programming (OOP) concepts such as classes and inheritance.

- Utilizing built-in libraries such as NumPy and Pandas for data manipulation.

- Writing efficient and optimized code for machine learning workflows.

2. Introduction to Machine Learning

Machine learning is a crucial aspect of artificial intelligence that allows systems to learn patterns from data and make predictions. Machine learning was one of the core topics I worked on during the internship. I started by understanding the fundamental principles of machine learning, including supervised and unsupervised learning. I studied various algorithms such as Linear Regression, Decision Trees, and Support Vector Machines (SVM).

I worked on real-world datasets, implementing machine learning models using Python's Scikit-learn library. Understanding model evaluation techniques such as accuracy score, confusion matrix, and cross-validation helped me assess model performance effectively. This experience provided me with insights into how machine learning can be leveraged to solve complex problems.

I covered the following key areas:

- Supervised and unsupervised learning concepts.

- Regression and classification techniques.

- Overfitting, underfitting, and model evaluation metrics.

- Feature engineering and data preprocessing for ML models.

3. Implementation of ML using Scikit-learn

Scikit-learn is one of the most widely used machine learning libraries in Python. After understanding the basics of machine learning, I applied my knowledge by implementing models using Scikit-learn. I loaded datasets, performed data preprocessing, and trained various models to predict outcomes. One of the key challenges was tuning the hyperparameters of machine learning models to achieve better accuracy.

To address this, I explored techniques such as GridSearchCV and RandomizedSearchCV, which helped me optimize model performance. Additionally, I learned the importance of feature scaling and feature engineering, which play a crucial role in improving model accuracy. I worked on implementing various ML algorithms using this library:

- **Linear Regression:** Used for predicting continuous values such as house prices.

- **Logistic Regression:** Implemented for classification problems such as spam detection.

- **Decision Trees and Random Forests:** Explored tree-based models for classification tasks.

- **Support Vector Machines (SVM):** Applied for both classification and regression problems.

- **K-Means Clustering:** Implemented for segmenting data into meaningful clusters.

- **Model evaluation:** Used techniques such as cross-validation, precision, recall, and F1-score for assessing model performance.

4. Introduction to NLP

Natural Language Processing (NLP) is a field of AI that focuses on the interaction between computers and human language. Natural Language Processing (NLP) is a crucial field in artificial intelligence, and my internship allowed me to explore its applications in depth. I started by understanding the basics of NLP, including text processing, tokenization, stemming, lemmatization, and Named Entity Recognition (NER).

One of my key projects involved analyzing textual data and extracting meaningful insights. I learned about different text representation techniques such as Bag of Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF), which are used for text classification tasks.

I gained insights into the following:

- Basics of text processing and linguistic structure.

- Tokenization, stemming, and lemmatization.

- Named Entity Recognition (NER) and Part of Speech (POS) tagging.

- Text classification and sentiment analysis.

5. Preprocessing using NLTK

The Natural Language Toolkit (NLTK) is a Python library that provides tools for text processing. For NLP tasks, I extensively worked with the Natural Language Toolkit (NLTK) library. Using NLTK, I performed text preprocessing steps such as stop-word removal, word tokenization, and stemming. These steps were crucial in preparing the text data for further analysis.

Handling noisy text data was a significant challenge, as it required careful preprocessing to ensure meaningful results. By experimenting with different text-cleaning techniques, I was able to improve the efficiency of NLP models. This experience helped me develop a strong understanding of how text preprocessing impacts model performance.

I used it for:

- **Tokenization:** Splitting text into words and sentences.

- **Stop word removal:** Eliminating unnecessary words such as "the," "is," and "in."

- **Stemming and Lemmatization:** Reducing words to their root forms for efficient processing.

- **POS tagging and Named Entity Recognition (NER):** Identifying key entities in text.

- **Text normalization:** Converting text into a structured format suitable for machine learning.

6. Introduction to Django

Django is a high-level Python framework used for web development. I explored Django's built-in features such as authentication, database management, and URL routing. This knowledge helped me develop a simple yet functional web application. I also learned how to integrate front-end templates with Django's back-end logic. My learning included:

- Understanding Django's Model-View-Controller (MVC) architecture.

- Setting up a Django project and managing applications.

- Creating models, views, and templates.

- Handling URL routing and request handling.

7. Loading Template and Building a Simple App using Django

One of my key projects involved building a simple web application using Django. The project involved developing a basic user interface, setting up URL routing, and handling user inputs through forms. I also connected the application to a database using Django ORM (Object-Relational Mapping).

Debugging and fixing errors in Django applications was a challenging yet rewarding experience. I encountered issues related to template rendering and database connectivity, which I resolved by referring to Django documentation and seeking guidance from my mentor. By the end of this module, I was able to build a fully functional Django application that demonstrated CRUD (Create, Read, Update, Delete) operations.

For practical implementation, I built a simple web application using Django, focusing on:

- **Setting up Django environment:** Installing Django and creating a project.

- **Building models:** Designing database schema using Django's ORM.

- **Creating views and templates:** Rendering data dynamically using HTML and Django templates.

- **Handling user authentication:** Implementing login and registration features.

- **Deploying the application:** Running the application on a local server.

CHALLENGES FACED

During my internship, I encountered several challenges that tested my problem-solving abilities. Understanding complex machine learning algorithms and tuning their hyperparameters required extensive experimentation and practice. Handling large textual datasets in NLP also posed difficulties, as it required efficient preprocessing techniques to improve model performance.

Working with Django was initially overwhelming, as it involved understanding multiple components such as models, views, and templates. Debugging issues related to database queries and template rendering required patience and a systematic approach. However, through continuous learning, mentor guidance, and practical implementation, I was able to overcome these challenges.

KEY LEARNINGS & SKILLS GAINED

Through this internship, I developed multiple technical and soft skills, including:

- **Technical Skills:**
  - Proficiency in Python programming
  - Hands-on experience with Machine Learning using Scikit-learn
  - NLP techniques and text preprocessing using NLTK
  - Web development using Django

- **Soft Skills:**
  - Problem-solving and debugging
  - Time management and meeting project deadlines
  - Collaboration and teamwork

CONCLUSION

My internship at Evolvier was a valuable learning experience that enhanced my technical and problem- solving skills. I gained in-depth knowledge of machine learning, NLP, and Django-based web development. The practical experience I obtained will be beneficial in my future projects and career endeavours.