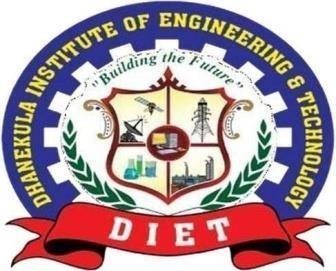
**HOTEL RECOMMENDATION SYSTEM**

# Submitted in partial fulfillment for the award of certificate of

**BACHELOR OF TECHNOLOGY**

**IN COMPUTER SCIENCE AND ENGINEERING – ARTIFICIAL INTELLIGENCE & MACHINE LEARNING**

# By

MANI DEEP MURALA (218T1A4235)

# DHANEKULA INSTITUTE OF ENGINEERING & TECHNOLOGY GANGURU, VIJAYAWADA - 521 139

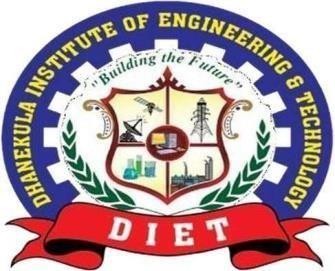
Affiliated to JNTUK, Kakinada &Approved By AICTE, New Delhi Certified by ISO 9001-2015, Accredited By NBA

# DHANEKULA INSTITUTE OF ENGINEERING&TECHNOLOGY

## GANGURU, VIJAYAWADA - 521 139

Affiliated to JNTUK, Kakinada &Approved By AICTE, New Delhi Certified by ISO 9001-2015, Accredited by NBA

# Department of Computer Science & Engineering– Artificial Intelligence & Machine Learning



This is to certify that the Summer Internship work entitled “**TECHNOLOGY VIRTUAL EXPERIENCE[HOTEL RECOMMENDATION SYSTEM]”** is a bonafide record of internship work done by MANI DEEP MURALA(218T1A4235) for the award of the Summer Internship in Computer Science and Engineering-Artificial Intelligence & Machine Learning by Jawaharlal Nehru Technological University, Kakinada during the academic year 2022- 2023.

## Head of Department:

Dr. K. SOWMYA

Professor, HOD CSE- AI& ML EXTERNAL EXAMINER

DHANEKULA INSTITUTE OF ENGINEERING & TECHNOLOGY

Department of Computer Science & Engineering Artificial Intelligence & Machine Learning VISION – MISSION - PEOs

|  |  |
| --- | --- |
| Institute Vision | Pioneering Professional Education through Quality |
| Institute Mission | Providing Quality Education through state-of-art infrastructure, laboratories and committed staff.  Moulding Students as proficient, competent, and socially responsible engineering personnel with ingenious intellect.  Involving faculty members and students in research and development works for betterment of society. |
| Department Vision | To empower the budding talents and ensure them with probable employability skills in addition to human values by optimizing the resources. |
| Department Mission | * To encourage students to become pioneers in the global competition with problem-solving skills * To make students become innovative with potential skills to explore the employment opportunities and/or to become entrepreneurs * To promote Research environment and inculcate corporate social responsibility |
| Program Educational Objectives(PEOs) | Graduates of Computer Science & Engineering will:  PEO1: Excel in problem solving and designing new products for a competitive and challenging business environment  PEO2: Contribute to technological innovation, research and society through the application of information technology in a diversified world. |

**PROGRAM OUTCOMES(PO’S)**

1. Engineering knowledge: apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. Problem Analysis: identify, formulate, review research literature, and analyze complex engineering problems reaching sustained conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. Design/Development Of Solutions: design solutions for complex engineering problems and design system components or process that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. Conduct Investigations Of Complex Problems: use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. Modern Tool Usage: create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modelling to complex engineering activities with an understanding of the limitations.
6. The Engineer And Society: apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. Environment And Sustainability: understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. Ethics: apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. Individual And Team Work: function effectively as an individual, and as a member or a leader in diverse teams, and in multidisciplinary settings.
10. Communication: communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. Project Management And Finance: demonstrate knowledge and understanding of the engineering and management principles and apply these to one’s own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. Life- Long Learning: recognize the need for, and have the preparation and ability to engage in independent and life- long learning in broadest context of technological change.

**PROGRAM SPECIFIC OUTCOMES(PSOs)**

PSO1: Designing and developing the Information Technology based systems with high professional skills.

PSO2: Qualify in national and international level competitive examinations for successful higher studies and get employment in IT enabled industries

Internship Mappings

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Project Title | P O 1 | P O 2 | P O 3 | P O 4 | P O 5 | P O 6 | P O 7 | P O 8 | P O 9 | P O 10 | P O 11 | P O 12 | P S O 1 | P S O 2 |
| HOTEL RECOMMENDATION SYSTEM |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |
| --- | --- |
| Mapping Level | Mapping Description |
| 1 | Low Level Mapping with PO & PSO |
| 2 | Moderate Mapping with PO & PSO |
| 3 | High Level Mapping with PO & PSO |

MANIDEEP MURALA

## 218T1A4235

3/1



## Contents

1. Internship carried out Company/Organization Details 2 Internship Log.

1. Domain area of the Internship.
2. Project report
3. **Internship carried out Company/Organization Details**



Codegnan is a IT solutions company which offers internships to the students We are hired for an internship for HOTEL RECOMMENDATION SYSTEM” project.

Our batch members

B.SATYA PRAKASH

D.SRI SAI

We will get a certification for doing the internship from codegnan technologies. From 22-MAY-2023 to 15-JULY-2023 we done this project

CODEGNAN ADDRESS:

#40-5-19/16, Prasad Naidu Complex, P.B.Siddhartha Busstop,Moghlrajpuram Vijayawada-520010

1. **InternshipLog**

**Day 1-5:**

Introduction to Data Science and Artificial Intelligence,Importance of Data in the 21st Century,Types of Data and its usage.

What is a Python Module? How is it different from a Python File? Creating Modules and Packages Importing Functions, Variables from different modules Python built-in modules.

Hands-On – Working on Python Built-in Modules and User-defined Custom Module Understanding Importance of Data Analysis Understanding Importance and Types of Data Analysis Understanding types of Data.

Understanding Importance of Visualization and Types of Graphs Understanding the Importance and Usage of Jupyter Notebook

**Day 6-10:**

Understanding and Working on NumPy Introduction to NumPy Advantages of NumPy over lists Creating NumPy arrays - 1-D, 2-D, N-D arrays Data types for ndarrays

Checking the attributes - shape, size, dimensions, dtype NumPy Arithmetic Operations NumPy universal functions Linear Algebra using NumPy Working on Appending and Concatenating.

Pandas for Data Analysis Getting Started with Pandas Introduction to Pandas Data Structures - Series, DataFrames Checking Attributes and Description.

Basic Essential functionality - Reindexing, Dropping entries from an axis Indexing, Selection, and Filtering Working on loc and iloc Functionalities Data Loading and Storage Reading and writing different file types (.txt,

.xlsx, .csv files).

Data Cleaning and Preparation Handling Missing Data - Filtering out Missing Data, Filling in Missing data Removing Duplicates Computing Indicator/Dummy Variables.

**Day 11-15:**

Data Aggregation and Group Operations Pivot Tables and Cross-Tabulation Working on TimeSeries data Hands- on: Case Study Working on Titanic Dataset for Cleansing Data, HR Analytics Data.

Data Visualization Using Matplotlib Introduction to Matplotlib Setting Labels, Titles, xticks, and yticks Multiple Line Plots, adding legend Bar charts - What are they, When to use it.

Scatter Plots and their Usage Pie charts and their usage Subplots and their Usage Hands-on: Case Study Working on Titanic Dataset for Visualization.

Python Interactive Visualization using Plotly for Dashboards Introduction to Plotly and Cufflinks Loading Plotly and Cufflinks Loading the Data.

Quick Visualization with custom bar charts Interactive Bubble charts Understanding and Working on Choropleth Maps Hands-on: Analyzing Gapminder dataset.

**Day 16-20:**

Data - Wealth of the 21st Century - Web Scraping using Python Why Web Scraping and Understanding its importance Installing BeautifulSoup Understanding web structures Scraping data from the web using Beautiful Soup - Static & Dynamic websites Performing Data Visualization over the scraped data.

Machine Learning Fundamentals Data Transformation and Preprocessing Handling Numeric Features Feature Scaling Standardization and Normalization.

More on different encoding techniques Train, Test and Validation Split Simple Train and Test Split Drawbacks of train and test split K-fold cross-validation Time-based splitting.

Overfitting And Underfitting What is overfitting? What causes overfitting? What is Underfitting? What causes underfitting? What are bias and Variance? How to overcome overfitting and underfitting problems?

**Day 21-25:**

Maths behind Linear Regression Ordinary Least Square Gradient Descent R - square Adjusted R-square Multiple Regression Performance Measures - MSE, RMSE, MAE Assumption of Linear Regression Ridge and Lasso regression Hands-on: Algorithm implementation with real use case datasets.

Building and Deployment of Machine learning model - Flask, Git, Github & PythonAnywhere Understanding steps in end-to-end ML projects Building a web service for Machine Learning Model Git Download and Github Usage Deploying the Final Trained Model on PythonAnywhere.

Understanding Classification Modelling Approach Introduction to the Classification problem Why the name Regression? and Implementation of the Sigmoid Function

Working on a dataset for Logistic Regression Performance Metrics for Classification Algorithms Accuracy Score Confusion Matrix, Precision-Recall F1-Score, ROC Curve and AUC, Log Loss

1. **Domain area of the Internship**

Machine Learning (ML) is a subset of artificial intelligence that focuses on developing algorithms and models that enable computers to learn from and make predictions or decisions based on data. By leveraging statistical techniques and pattern recognition, ML systems can improve their performance over time without explicit programming.

ML algorithms can be broadly categorized into supervised, unsupervised, and reinforcement learning. Supervised learning involves training a model on labeled data to make predictions or classifications, while unsupervised learning deals with discovering patterns and relationships in unlabeled data. Reinforcement learning, on the other hand, revolves around an agent learning to make decisions by interacting with an environment and receiving feedback in the form of rewards or penalties.

The applications of ML are diverse, ranging from image and speech recognition, natural language processing, recommendation systems, to autonomous vehicles and medical diagnoses. Its widespread use is transforming various industries, leading to more efficient processes and enhanced decision-making capabilities.

Machine learning algorithms are trained on data sets, which can be labeled (supervised learning) or unlabeled (unsupervised learning). During the training phase, the algorithms analyze the data and identify patterns, correlations, and features that are relevant to the task at hand. The algorithms then use this learned information to make predictions or take actions when presented with new, unseen data.

Machine learning has a wide range of applications, including image and speech recognition, natural language processing, recommendation systems, fraud detection, autonomous vehicles, medical diagnosis, and many others. It has become an essential tool in various industries, enabling computers to perform complex tasks and make accurate predictions based on large amounts of data.

1. **Project report**

Introduction:

The purpose of this project is to build a hotel recommendation system that suggests the best hotels to users based on their preferences and historical data. The system aims to enhance user experience and satisfaction by providing personalized hotel recommendations tailored to individual preferences, such as location, budget, amenities, and user reviews.

Data Collection:

The first step in the project involved collecting hotel data from various sources. This data includes hotel details, such as name, location, amenities, prices, and user ratings. Additionally, user behavior data, such as past bookings, searches, and reviews, were collected to understand user preferences and behavior patterns.

Data Preprocessing:

Before building the recommendation system, the collected data needed to be preprocessed and cleaned. This involved handling missing values, data normalization, and converting categorical features into numerical representations. Furthermore, user behavior data was aggregated and transformed into meaningful user preferences.

Exploratory Data Analysis (EDA):

EDA was performed to gain insights into the data and identify patterns and trends. Visualizations and statistical analysis helped understand user preferences, popular hotel locations, price distributions, and other relevant factors.

Recommendation System:

There are several approaches to building a recommendation system. For this project, we implemented a collaborative filtering-based recommendation system using the following two methods:

User-based Collaborative Filtering: This method identifies users with similar preferences and recommends hotels based on what similar users have liked or booked in the past.

Item-based Collaborative Filtering: This method identifies hotels that are similar in terms of user preferences and recommends hotels similar to those that a user has shown interest in.

Model Evaluation:

To evaluate the recommendation system's performance, we split the data into training and testing sets. We used metrics such as precision, recall, and Mean Average Precision (MAP) to measure the system's effectiveness in providing relevant hotel recommendations to users.

Model Deployment:

The recommendation system was deployed as a web application to provide real-time hotel recommendations to users. Users can input their preferences, and the system will suggest the top-rated hotels based on their selections and past behavior.

Results and Conclusion:

The evaluation results showed that the collaborative filtering-based recommendation system provided accurate and relevant hotel recommendations to users. The system significantly enhanced user satisfaction and engagement with the platform, leading to increased bookings and positive feedback.

Future Enhancements:

To further improve the hotel recommendation system, the following enhancements could be considered:

Incorporate additional data sources, such as social media activity, to gain a deeper understanding of user preferences.

Implement advanced recommendation algorithms, like Matrix Factorization or Neural Collaborative Filtering, to improve the accuracy and performance of the system.

Develop a mobile application to extend the recommendation system's reach and accessibility to a wider user base.

**SOURCE CODE :**

<https://drive.google.com/drive/folders/1_yGUlMcuDAulqhP7FwqDJ7M5k_ccf3gL>

**SAMPLE OUTPUT :**

