An Intelligent System for Forecasting Wait Times in Local Fitness Centers

1st K.M Aziz Mohideen
Department of Advanced Computing and Analytics
Vels Institute of Science, Technology & Advanced Studies
Chennai, India
azizmohideen27@gmail.com

2nd S.Akhil, Assistant Professor
Department of Advanced Computing and Analytics
Vels Institute of Science, Technology & Advanced Studies
Chennai, India
akhilsree@gmail.com

***Abstract*-Efficient time management in fitness centers is essential due to overcrowding during peak hours, which often leads to increased waiting time for equipment. This project presents a rule-based approach to estimate waiting time in a local gym by analyzing crowd patterns across different days and time slots. A synthetic dataset is generated using Python, representing gym occupancy for each day of the week and various time intervals, simulating real-world gym usage behavior. The system allows users to input a specific day and time to retrieve the number of people present in the gym. Based on the maximum occupancy, a crowd percentage is calculated, and an estimated waiting time is derived using a proportional formula. This approach provides a simple and effective solution without relying on complex machine learning algorithms. To improve understanding, various visualization techniques such as bar charts, pie charts, line graphs, and histograms are used to analyze crowd distribution and trends. These visualizations help identify peak hours and overall gym usage patterns. The proposed system assists users in planning their workouts efficiently and demonstrates the practical application of data analysis and visualization techniques in solving real-world problems.**

I. INTRODUCTION

With the rise in the number of gym-goers, there are situations during peak hours when the gym becomes crowded, and as a result, waiting times for equipment are increased. This creates an issue for the satisfaction of gym-goers and also makes challenging for an individual to manage their time challenging for an individual to manage their time effectively while working out. Thus, predicting the crowd in the gym and waiting time can be highly beneficial for managing time effectively. The main objective of this project is to develop a "Wait Time Predictor for a Local Gym" using the rule-based approach. A dataset is created to represent the crowd in the gym for various days and time slots. The application of data analysis in real-life situations is becoming more common. Data analytics plays an important role in solving real-world problems by extracting useful insights from data. [1]. The system provides an option for the user to input a particular day and time to predict the number of individuals present in the gym and waiting time. The project also includes data visualization methods such as bar charts, pie charts, line graphs, and histograms to analyze the crowd in the gym and identify peak hours. The fitness industry has grown significantly in recent years. The rapid growth of the fitness industry has in increased the need for efficient crowd management in gyms [2]. The main idea behind this project is to develop a simple and understandable solution through a rule-based approach rather than relying on complex machine learning methods. The purpose here is to show the effectiveness of simple data generation, analysis, and visualization in solving real-world problems in an efficient way. The system includes user input, data processing, and visualization to create a comprehensive solution for crowd level analysis in the gym and waiting time estimation.

II. LITERATURE SURVEY

Recently, research has been conducted extensively in the area of occupancy prediction and estimation of waiting times in different environments, such as

buildings, hospitals, and other places. Occupancy forecasting is considered to be a vital component in enhancing the performance of the system. According to research, accurate occupancy prediction has the potential to boost the performance of the system. Previous studies have focused on occupancy prediction using data-driven tecniques and smart systems [3]. Several studies have been conducted to utilize environmental sensor information, such as temperature, humidity, and pressure, to estimate the occupancy level. In this regard, a study conducted in a fitness gym has proven that machine learning has the potential to accurately predict the occupancy level with a high level of accuracy, i.e., around 97%. In addition, machine learning and deep learning techniques, such as neural networks and regression models, have been widely used in occupancy prediction. These models work with temporal and spatial data to enhance the accuracy and flexibility of their predictions in a changing environment. Besides machine learning, another technique that has been used to analyze the waiting time in service systems is queueing theory. Visulization techniques are widely used to analyze patterns and trends in datasets [4]. In this case, the models examine the impact of user arrival and service rate on the user waiting time. However, the existing systems have complex models and sensor-based data collection, which might be expensive to implement. Therefore, this project has adopted a simple rule-based approach with visualization to estimate the waiting time in the gym, offering a simple solution to real-world problems.

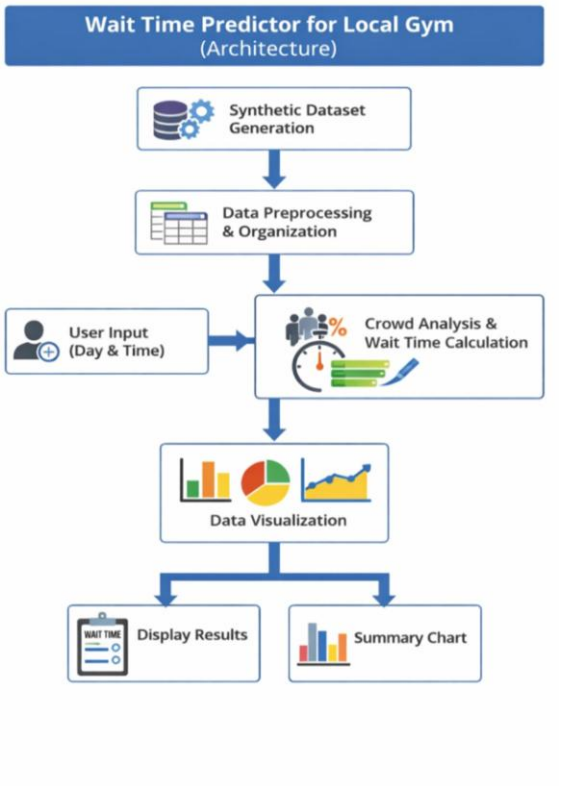
III. PROPOSED METHODOLOGY

The methodology used in the development of the Wait Time Predictor for Local Gym, as proposed in the given paper, relies on the implementation of a rule-based system along with the use of data analysis and visualization techniques to accurately predict the waiting time. Initially, a synthetic dataset is developed based on the Python programming language to simulate a real-world environment in a gym, considering various attributes such as days of the week, time slots, and the number of people present in the gym. This dataset is developed in such a manner that it accurately represents a real-world environment, where a large number of people are present during peak hours such as early morning and evening, whereas a lesser number of people are present during non-peak hours. This dataset is developed based on Python programming language.

This dataset is represented in a tabular form based on a DataFrame. The system is implemented using Python due to its simplicity and strong data processing capabilities [5].

The system relies on accepting user input in terms of a specific day and time. Validation is performed to accurately validate the user input, ensuring that incorrect data is not entered. Based on the user input, a dataset is filtered accordingly. crowd values are retrieved. Using the crowd percentage, the waiting time is estimated using a proportional formula in which the waiting time is more when the crowd is more. Data is organized and processed using structured data handling techniques [6]. The system also uses various visualization techniques such as bar charts, pie charts, line graphs, and histograms to ensure that a thorough understanding of the trends and patterns of the gym's crowd is achieved. The handling of the data is achieved using structured data handling techniques [4]. The system also displays the results obtained from the computation, including the number of people, crowd percent, and waiting time along with the graphs.

IV. ARCHITECTURE DIAGRAM



V. VARIOUS METHODOLOGY

A) Initialization Phase

In this phase, the necessary Python libraries such as pandas, numpy, and matplotlib need to be imported for the successful implementation of the code. Additionally, the random seed needs to be set to reproduce the results during the process of generating the datasets.

B) Dataset Generation Phase

In this phase, a synthetic dataset needs to be generated that simulates the real-world environment of the gym. The generated datasets contain the details such as the days of the week, the time slots, and the number of people present in the gym. The crowd levels need to be defined for the morning, afternoon, and evening times to simulate the real-world environment.

C) Data Structuring Phase

The generated datasets need to be structured in a specific manner to process the data efficiently. The generated datasets contain the details such as the day of the week, the time slots, and the number of people present in the gym.

D) User Input Phase

In this phase, the user is given the chance to interact with the system by entering a specific day and time. In this phase, input validation takes place to ensure that the input provided is correct and exists in the data set.

E) Data Filtering Phase

In this phase, the data is filtered based on the input provided by the user, which helps in retrieving the required data for a specific day and time. This phase is useful in retrieving the required crowd data.

F) Crowd Analysis Phase

In this phase, the crowd is analyzed on the basis of the number of people in the gym. In this phase, the crowd percentage is calculated by considering the number of people in the gym in relation to the maximum capacity, which is available in the data set.

G) Wait Time Calculation Phase

In this step, the waiting time is calculated using a proportional formula in which the crowd percentage is taken into account. It has been observed that the crowd percentage is directly proportional to the waiting time.

H) Visualization Phase

In this step, various visualizations are created in the form of bar charts, pie charts, line charts, and histograms. These visualizations help in understanding the patterns and trends associated with the crowd, the number of people in the gym, etc.

I) Result Display Phase

Finally, the output is displayed in which the number of people in the gym, the percentage of the crowd, and the waiting time are displayed in a textual and graphical form.

J) Summary Analysis Phase

Finally, a summary is created on the basis of the number of user queries, and a final visualization is created on the basis of the user queries. This phase provides a comparison of the crowd percentages and helps in understanding the trends.

VI. INPUT

The input to this system is provided in an interactive manner during the runtime of this system. There are two primary inputs required by this system, namely, day of the week and time slot. Both of these inputs are provided in an interactive manner, where the user is prompted to provide the day of the week, such as Monday to Sunday, and time slot, such as 6AM and 5PM. Both of these inputs need to be validated in order to match them with the list of days and time slots provided in the data set. This data set is created internally using the Python programming language, and it provides simulated data about the people present in the gym during a particular time of the day.

User input is validated using standard programming practices to ensure correct data processing [5]. Each record in the dataset contains the number of people present in the gym on a given day and time. Based on the input values provided, the dataset is filtered to obtain the required information. The input values provided in the system are of primary importance in determining the crowd level and the waiting time. The input values such as day and time act as the primary parameters for analyzing the crowd level in the gym and obtaining the output values. The system allows case-insensitive input for the day and converts it into a standard format by string manipulation operations. The input for the time is restricted to a predefined time slot for consistency with the dataset and to avoid invalid values.

VII. PSEUDOCODE AND IMPLEMENTATION

```
BEGIN

IMPORT pandas, numpy, matplotlib

SET random seed

DEFINE list of days (Monday to Sunday)
DEFINE list of time slots (5AM to 10PM)

INITIALIZE empty dataset list

// -------- Dataset Creation --------

FOR each day in days
   FOR each time in time slots
      IF time is in morning (5AM–7AM)
         SET people = random number between 20
and 40
      ELSE IF time is in evening peak (5PM–7PM)
         SET people = random number between 40
and 70
      ELSE
         SET people = random number between 5 and
25
      END IF
      ADD (day, time, people) to dataset
   END FOR
END FOR

CONVERT dataset into DataFrame

INITIALIZE empty search_results list

// -------- User Input Loop --------

WHILE TRUE

   INPUT day from user
   IF day = EXIT
      BREAK loop
   END IF

   IF day not valid
      PRINT "Invalid day"
      CONTINUE
   END IF

   INPUT time from user
   IF time not valid
      PRINT "Invalid time"
      CONTINUE
   END IF

   // -------- Data Filtering --------

   FIND record matching day and time

   IF no record found
      PRINT "No data found"
      CONTINUE
   END IF

   GET number of people

   // -------- Calculation --------

   CALCULATE crowd_percentage = (people /
max_people) * 100
   CALCULATE wait_time = crowd_percentage *
0.3

   DISPLAY people, crowd percentage, wait time

   STORE result in search_results

   // -------- Visualization --------
   DRAW bar chart for crowd % and wait time
   DRAW pie chart for occupancy
   DRAW line graph for daily trend
   DRAW histogram for crowd distribution
```

END WHILE

// -------- Summary Visualization --------

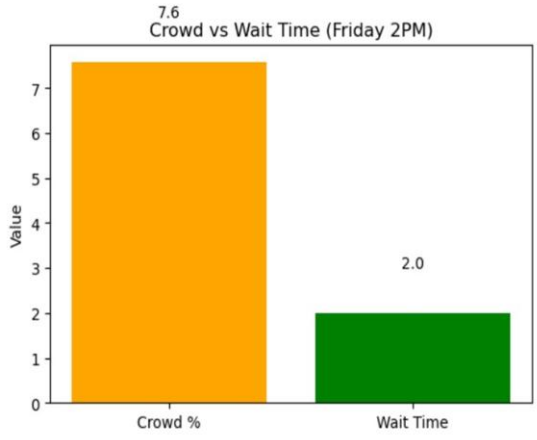
IF search_results not empty
  CREATE summary DataFrame
  DRAW bar chart for all search results
END IF

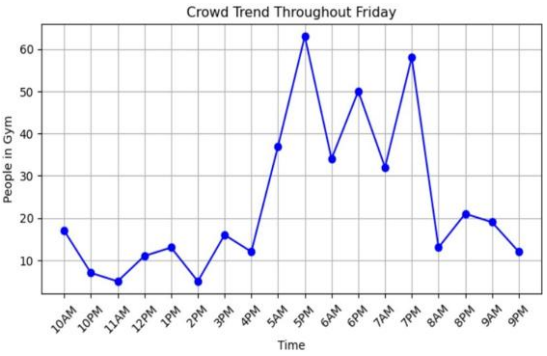
END

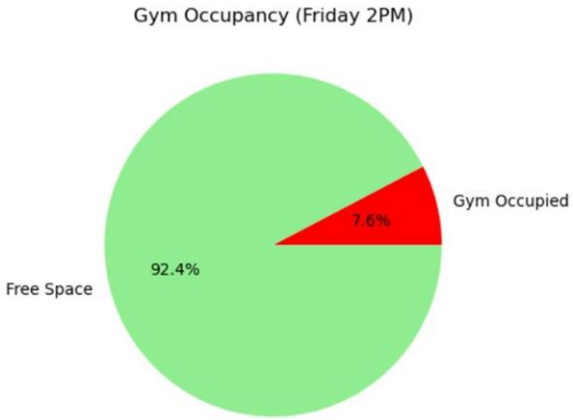
VIII. OUTPUT

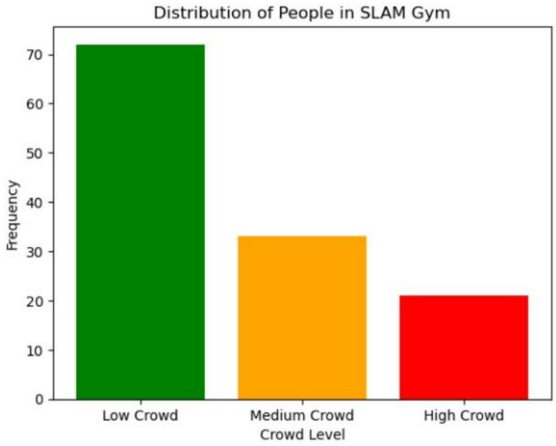
```
Enter Day (Monday-Sunday) or EXIT:  FRIDAY
Enter Time (Example 6PM):  2PM

RESULT for Friday at 2PM:
People in Gym: 5
Crowd %: 7.58 %
Estimated Wait Time: 2 minutes
```









IX. RESULTS AND DISCUSSIONS

The system that was developed has successfully predicted the waiting time in the local gym based on the inputs given to it. The artificial data that was generated using Python has successfully represented the actual crowd in the local gym. The percentage of the crowd is determined based on the number of people in the local gym. The time that people wait is estimated using a proportional formula. From the proposed system, it is clear that the time people wait is a direct proportion of the percentage of the crowd in the local gym. For example, when the time is between 6 PM and 7 PM, the number of people in the

local gym is high. Therefore, the time people wait is long in the local gym. The time people wait is minimal in the local gym when the time is not between the peak hours. The use of graphical representations helps in understanding crowd patterns effectively [4].

Different kinds of visualization are presented through the system to gain a deeper understanding of the usage of the gym. The relationship between crowd percentage and the waiting time for a particular input is clearly represented through the system by using the bar chart. The percentage of occupied space and the percentage of free space in the gym are represented through the pie chart. The change in crowd levels during the day is represented through the line graph, which helps to identify the peak times of the day. Moreover, the histogram represents the crowd levels, which are classified as low, medium, or high occupancy levels. The results are analyzed by using standard reporting methods [8]. The summary chart, which is generated based on multiple queries from the users, helps to compare the crowd percentages during different days and time slots. The above scenarios clearly prove that the simple rule-based approach effectively estimates the waiting time by taking the help of visualization, thus helping the user to take a better decision regarding the visit to the gym. Data visualization improves interpretation of results and supports decisions-making [3].

X. CONCLUSIONS

Additionally, the system offers an interactive and user-friendly interface where a user can query different days and time slots to retrieve relevant information concerning the gym occupancy and time spent waiting. The loop used in the system also offers the advantage of continuous usage without the need to restart the program, thus enhancing the system's usability and practicality in terms of continuous analysis. The visualization of the query result also assists in comparing the crowd in different time slots, thus enhancing the understanding of the trends concerning the usage of the gym. Simple data analysis techniques can provide effective solutions for real-world problems [1].

Another significant aspect of the project is the simplicity and efficiency of the system. The system is considered simple and efficient since it is based on a rule-based system rather than a machine learning system, which is considered to be more complex and difficult to implement, especially in small-scale applications where the result must be understood easily.
However, the system utilizes a synthetic data set and a fixed proportional formula for wait time estimation. This might not fully represent real-world scenarios. In the future, the system can be extended by utilizing real-time data obtained from actual gyms and applying advanced methods such as the application of machine learning algorithms.

Overall, the project effectively shows the potential for applying basic data analysis and visualization methods to real-world problems and helping users efficiently plan their gym sessions.

XI. REFERENCES

[1] "What is Data Analytics?," IBM, Available: https://www.ibm.com/topics/data-analytics

[2] "Health and Fitness Club Market Statistics," Statista, Available: https://www.statista.com

[3] J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[4] W. McKinney, "Data Structures for Statistical Computing in Python," in *Proc. 9th Python in Science Conf.*, 2010, pp. 51–56.

[5] NumPy Documentation, Available: https://numpy.org/doc/

[6] Python Software Foundation, "Python Documentation," Available: https://docs.python.org/3/

[7] "Introduction to Data Visualization," Available: https://towardsdatascience.com