Project 11: Identify abnormal driving behavior using spatio-temporal analysis

Bhavya Khakhar
B.Tech in CSE
Ahmedabad University
bhavya.k1@ahduni.edu.in

Harsh Loriya
B.Tech in CSE
Ahmedabad University
harsh.l2@ahduni.edu.in

Krishna Patel
B.Tech in CSE
Ahmedabad University
krishna.p5@ahduni.edu.in

Priyam Shah
B.Tech in CSE
Ahmedabad University
priyam.s@ahduni.edu.in

***Abstract*—This study investigates the use of machine learning (ML) with spatio temporal analysis for identifying normal and abnormal driving behaviors. By analyzing video data from GPS, we aim to develop models capable of distinguishing between typical driving patterns. We have used Ant-colony model and Logistic regression and their effectiveness is evaluated based on the accuracy and interpretability of the models. This is helpful contribution to enhance road safety and progress in the development of intelligent transportation systems.**

***Index Terms*—Abnormal Driving Behavior, Spatio-temporal Analysis, Trajectory Data, Feature Selection, Ant Colony Optimization, Logistic Regression**

I. Introduction

In recent years, in the era of sensor-equipped vehicles and advancements in data analytics, there has been a growing interest in understanding and identifying abnormal driving behaviours. Detecting such behaviours not only enhances road safety but also facilitates the development of intelligent transportation systems. Traditional approaches often assume uniform driver behaviour within specific road segments, overlooking variations that may indicate abnormal driving patterns based on the cretain parameters. This project aims to address this limitation by employing spatio-temporal analysis techniques to discern abnormal driving behaviours from trajectory data.

II. Methodology

Data Understanding: The trajectory dataset consists of spatial coordinates (left, top) and dimensions (width, height) of detected objects in each frame, along with corresponding labels that we have initialised from the code indicating normal or abnormal driving behaviour. We were given the normal data files and abnormal data files. We separated them and mad two different folder of both the files.

Data Preprocessing: Initially, the data undergoes preprocessing steps such as labelling each data instance with its respective folder name and removing certain irrelevant columns (for eg. in our dataset, it was latitude/longitude). Data preprocessing also include normalization or scaling of numerical attributes, encoding categorical variables, and partitioning the data into training and testing sets for model evaluation.

Feature Extraction: Relevant features, including spatial and temporal attributes such as frame number, object position, and dimensions, are extracted from the pre-processed trajectory data using Ant colony algorithm. These features serve as input for subsequent analysis and classification tasks.

Ant Colony Optimization for Feature Selection: Ant colony optimization is used to select the most discriminative features to easily identify the variation in the driving behaviour. Ant colony optimisation is a search algorithm technique that was inspired from the movements of ants. Here, this algorithm is used to explore the feature space efficiently and identify subsets of features that maximise classification performance. To optimise the accuracy of the model we used logistic regression for classification of Data

Logistic Regression Classification: Following feature selection, logistic regression is used as a binary classifier to distinguish between normal and abnormal driving behaviours to optimise the accuracy and improve the effectiveness of the model. Logistic regression is a widely used statistical method for binary classification tasks.

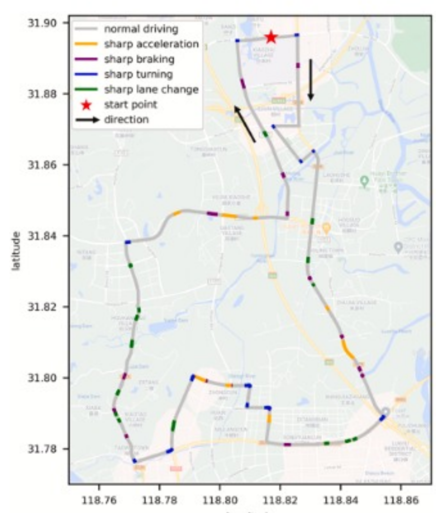


Fig. 1: Dataset

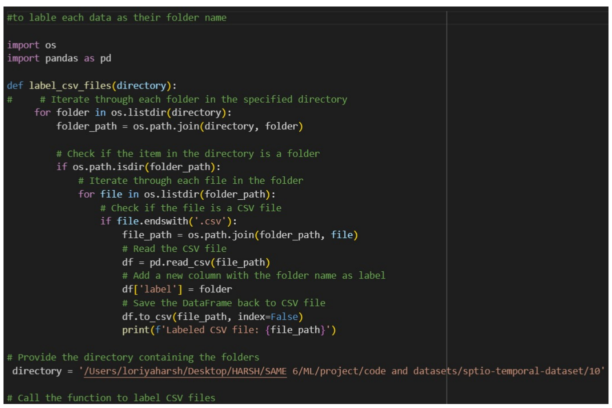
III. Results

The approach of whole project is evaluated on real world trajectory data which was obtained by breaking video data in seperate frames. Experimental results made by extracting features from ant colony algorithms demostrate the effectiveness by accurately differentiating between normal and abnormal driving beheviours. The logistic regression classifier is used to improve the classification performance. Find the section of figures and their working for more information.

TABLE I: Accuracy

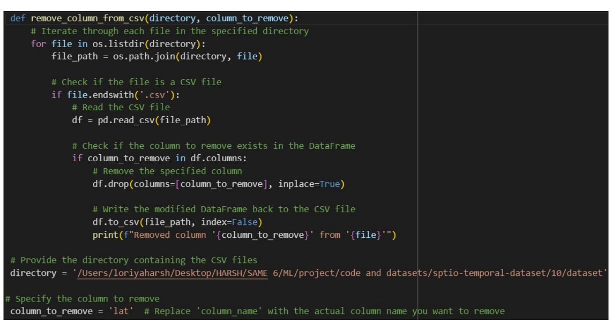
Algorithms	Accuracy
Ant Colony	68.82 percent
Logistic Regression	72.27 percent

A. Figures and their working



```python
#to lable each data as their folder name

import os
import pandas as pd

def label_csv_files(directory):
#     # Iterate through each folder in the specified directory
    for folder in os.listdir(directory):
        folder_path = os.path.join(directory, folder)

        # Check if the item in the directory is a folder
        if os.path.isdir(folder_path):
            # Iterate through each file in the folder
            for file in os.listdir(folder_path):
                # Check if the file is a CSV file
                if file.endswith('.csv'):
                    file_path = os.path.join(folder_path, file)
                    # Read the CSV file
                    df = pd.read_csv(file_path)
                    # Add a new column with the folder name as label
                    df['label'] = folder
                    # Save the DataFrame back to CSV file
                    df.to_csv(file_path, index=False)
                    print(f'Labeled CSV file: {file_path}')

# Provide the directory containing the folders
 directory = '/Users/loriyaharsh/Desktop/HARSH/SAME 6/ML/project/code and datasets/sptio-temporal-dataset/10'

# Call the function to label CSV files
```

Fig. 2: Data Labelling

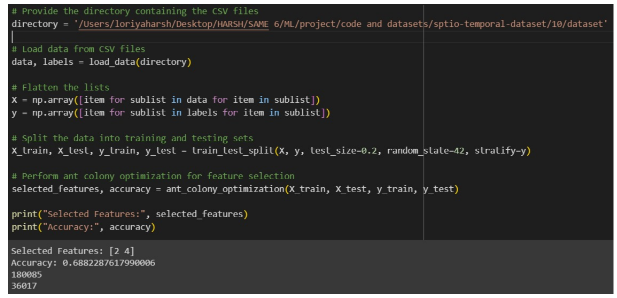


```python
def remove_column_from_csv(directory, column_to_remove):
    # Iterate through each file in the specified directory
    for file in os.listdir(directory):
        file_path = os.path.join(directory, file)

        # Check if the file is a CSV file
        if file.endswith('.csv'):
            # Read the CSV file
            df = pd.read_csv(file_path)

            # Check if the column to remove exists in the DataFrame
            if column_to_remove in df.columns:
                # Remove the specified column
                df.drop(columns=[column_to_remove], inplace=True)

                # Write the modified DataFrame back to the CSV file
                df.to_csv(file_path, index=False)
                print(f"Removed column '{column_to_remove}' from '{file}'")

# Provide the directory containing the CSV files
directory = '/Users/loriyaharsh/Desktop/HARSH/SAME 6/ML/project/code and datasets/sptio-temporal-dataset/10/dataset'

# Specify the column to remove
column_to_remove = 'lat'  # Replace 'column_name' with the actual column name you want to remove
```

Fig. 3: Data Cleaning

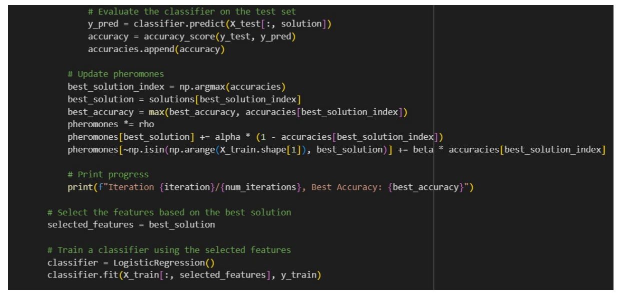
IV. Discussions

The successful identification of abnormal driving behaviours holds significant implications for various applications, including traffic management, driver assistance systems, and anomaly detection in surveillance scenarios. The use of spatio-temporal analysis techniques coupled with machine learning algorithms enables a nuanced understanding of driver behaviour and enhances the capabilities of intelligent transportation systems.



```python
# Provide the directory containing the CSV files
directory = '/Users/loriyaharsh/Desktop/HARSH/SAME 6/ML/project/code and datasets/sptio-temporal-dataset/10/dataset'

# Load data from CSV files
data, labels = load_data(directory)

# Flatten the lists
X = np.array([item for sublist in data for item in sublist])
y = np.array([item for sublist in labels for item in sublist])

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42, stratify=y)

# Perform ant colony optimization for feature selection
selected_features, accuracy = ant_colony_optimization(X_train, X_test, y_train, y_test)

print("Selected Features:", selected_features)
print("Accuracy:", accuracy)

Selected Features: [2 4]
Accuracy: 0.6882287617990006
180085
36017
```

Fig. 4: Ant Colony Optimisation



```python
            # Evaluate the classifier on the test set
            y_pred = classifier.predict(X_test[:, solution])
            accuracy = accuracy_score(y_test, y_pred)
            accuracies.append(accuracy)

        # Update pheromones
        best_solution_index = np.argmax(accuracies)
        best_solution = solutions[best_solution_index]
        best_accuracy = max(best_accuracy, accuracies[best_solution_index])
        pheromones *= rho
        pheromones[best_solution] += alpha * (1 - accuracies[best_solution_index])
        pheromones[~np.isin(np.arange(X_train.shape[1]), best_solution)] += beta * accuracies[best_solution_index]

        # Print progress
        print(f"Iteration {iteration}/{num_iterations}, Best Accuracy: {best_accuracy}")

    # Select the features based on the best solution
    selected_features = best_solution

    # Train a classifier using the selected features
    classifier = LogisticRegression()
    classifier.fit(X_train[:, selected_features], y_train)
```

Fig. 5: Logistic Regression Optimisation

V. Conclusion

In conclusion, the project presents a novel approach for identifying abnormal driving behaviours using spatio-temporal analysis and machine learning techniques. By leveraging trajectory data and employing feature selection with ant colony optimization, the proposed method achieves promising results in distinguishing abnormal driving behaviours from normal ones. Future work may involve further refinement of feature selection strategies and exploration of advanced machine learning algorithms to enhance classification performance.

References

[1] Yu Zhang, Yingying He, and Likai Zhang, "Recognition method of abnormal driving behavior using the bidirectional gated recurrent unit and convolutional neural network," *Physica A: Statistical Mechanics and its Applications*,. Available: https://doi.org/10.1016/j.physa.2022.128317

[2] Yongfeng Ma et al., "Real-time detection of abnormal driving behavior based on long short-term memory network and regression residuals," *Transportation Research Part C: Emerging Technologies*,. Available: https://doi.org/10.1016/j.trc.2022.103983

[3] Xiaodi Huang et al., "Abnormal driving behavior detection based on an improved ant colony algorithm," *Applied Artificial Intelligence*,. Available: https://doi.org/10.1080/08839514.2023.2216060