Parameter Setting and Reliability test of a Sensor System for Person Detection in a car wearing winter wear.

Course: Information Technology

Modules: Autonomous Intelligent Systems
and
Machine Learning

By Dr. Peter Nauth and Dr. Andreas Pech

Milestone-4

Shiva Kumar Biru
Mat No: 1436929

Faiz Mohammad Khan
Mat No: 1427310

Report on Passenger Detection with different Threshold using RandomForestClassifier with ADC Data

Link to Dataset

https://drive.google.com/file/d/1JK0Lkapl47YT4KQe00_sFq6uZ5Q2Moay/view?usp=sharing

Introduction

This report aims to present a comprehensive analysis of passenger detection in car seats utilizing RandomForestClassifier with ADC (Analog-to-Digital Converter) data processed through Fast Fourier Transform (FFT). Unlike the previous Milestone (MS3), which employed SVM, this analysis explores the effectiveness of RandomForestClassifier in handling high-dimensional FFT data. Additionally, a different threshold of **-t 160 9000** is applied in the sensor setup to enhance data preprocessing.

Data Collection and Description

The dataset comprises ADC data collected from sensors installed in car seats, capturing analog signals indicative of passenger presence. These analog signals are converted to digital format using ADC. Subsequently, FFT is applied to extract features characterizing passenger occupancy. Each observation in the dataset represents a unique data point, with features obtained from FFT analysis. The target variable indicates whether the seat is occupied or empty.

Data Splitting and Preprocessing

Similar to the Milestone 3, the dataset is randomly divided into training and testing sets using an 80-20 ratio. The training set, encompassing 80% of the data, is utilized for training the

RandomForestClassifier, while the remaining 20% serves for evaluating the classifier's performance. All datasets are transformed into NumPy arrays to facilitate compatibility with scikit-learn's machine learning algorithms.

Classifier Model

RandomForestClassifier is chosen for its capability to handle high-dimensional data and its ensemble learning approach, which aggregates the predictions of multiple decision trees. This classifier constructs a multitude of decision trees during training and outputs the class that is the mode of the classes of individual trees. By averaging over multiple trees, RandomForestClassifier can mitigate overfitting and enhance generalization performance.

Data Preprocessing with Threshold Adjustment

In contrast to previous Milestone, a different threshold (**-t 160 9000**) is applied during data preprocessing in the sensor setup. This threshold modification aims to improve the signal-to-noise ratio and enhance the discriminative power of the extracted features. The adjusted threshold ensures that only relevant ADC data within the specified range are considered for further analysis, potentially enhancing the classifier's performance.

Model Evaluation and Performance Metrics

The performance of the RandomForestClassifier is assessed using various metrics, including accuracy, precision, recall, and F1 score. Additionally, confusion matrices are generated to visualize the classifier's predictions compared to actual labels, providing insights into its performance across different classes. The evaluation metrics help in determining the classifier's effectiveness in distinguishing between occupied and empty car seats.
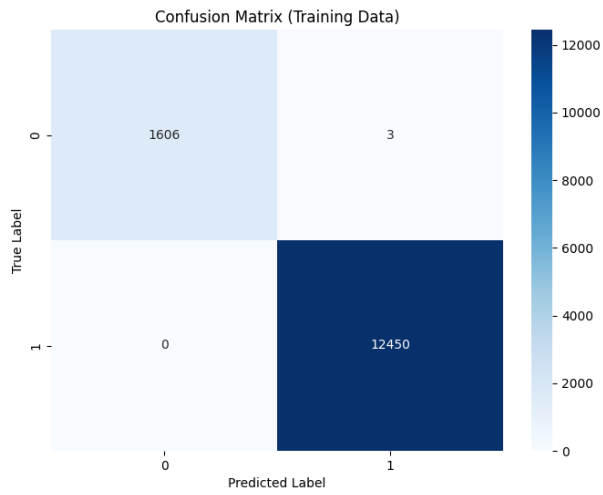
Results
The performance of the RandomForestClassifier classifier is evaluated based on the generated confusion matrices. Performance metrics such as accuracy, precision, recall, and F1 score are computed to assess the classifier's effectiveness in distinguishing between passenger present and seat empty scenarios.
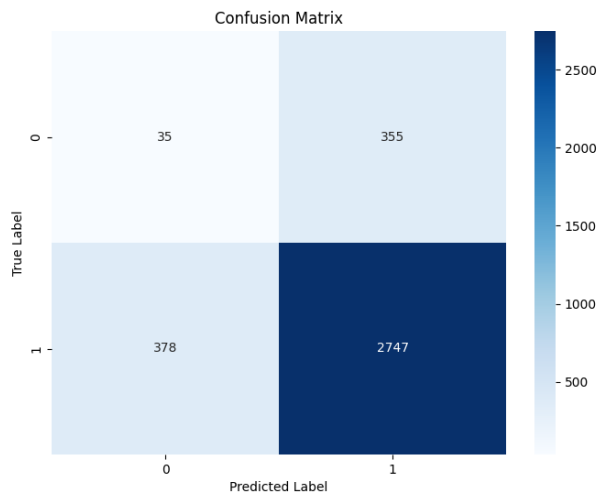
Label 1 shows Passenger Detection
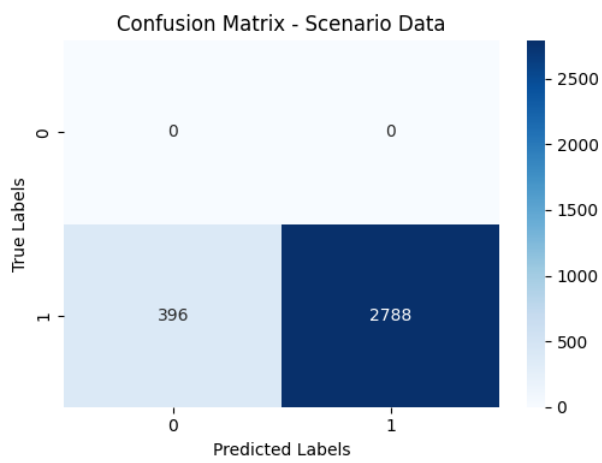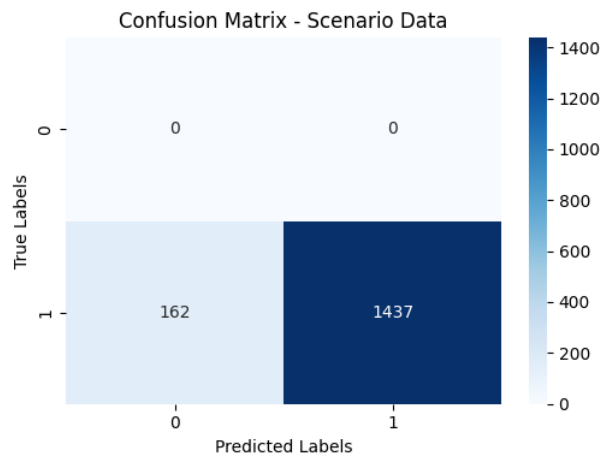**Label 0 shows No Passenger Detection**

Training Data

Confusion Matrix (Training Data)

Testing Data

Confusion Matrix

Scenarios
Person 1 – Jacket 1

Confusion Matrix - Scenario Data

Person 1 – Jacket 2 Seated Constant

Confusion Matrix - Scenario Data

Person 1 – Jacket 2 always Moving

Confusion Matrix - Scenario Data

Person 2 – Jacket 1

Confusion Matrix - Scenario Data
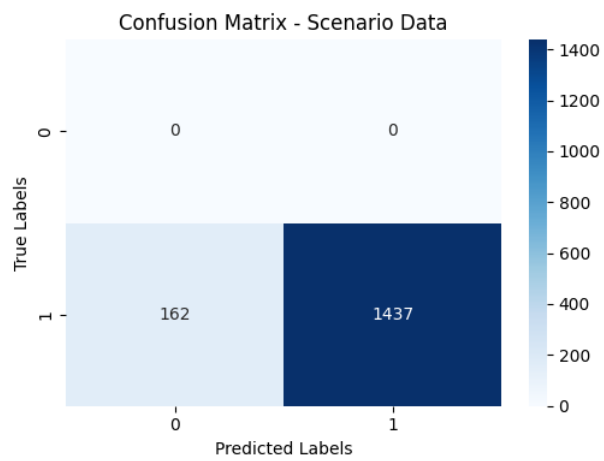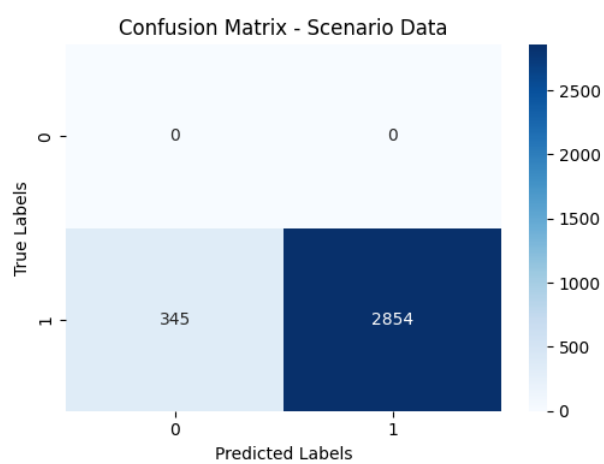
Person 2 – Jacket 2 Constant Seated

Confusion Matrix - Scenario Data
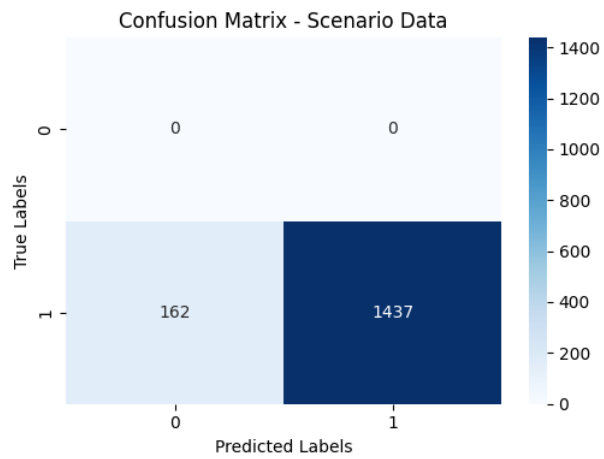
Person 2- Jacket 2 always Moving

Confusion Matrix - Scenario Data

Person 3 – Jacket 1

Confusion Matrix - Scenario Data

Person 3 – Jacket 2

Confusion Matrix - Scenario Data
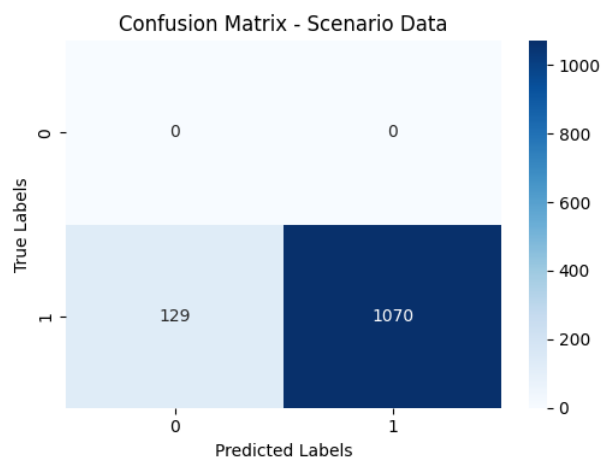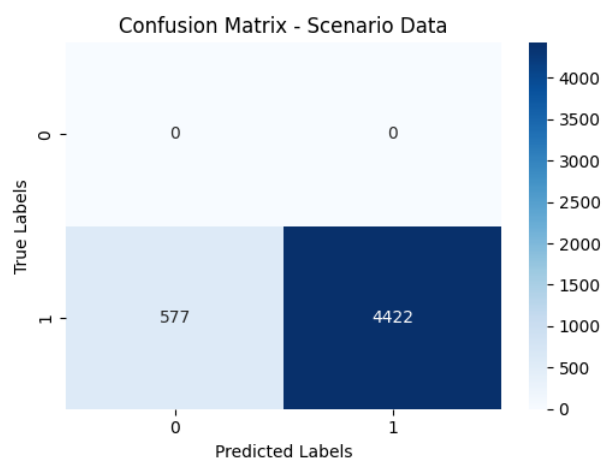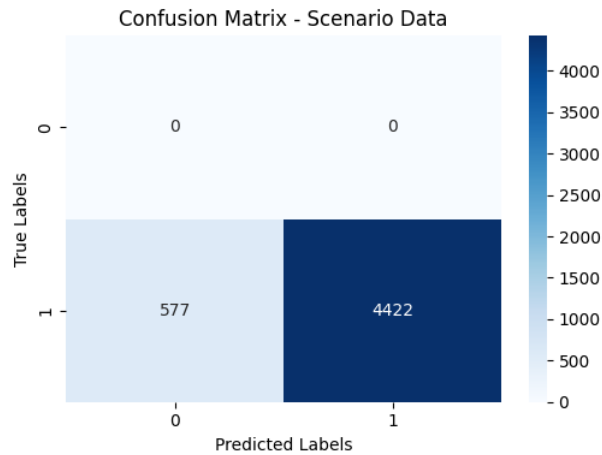
Conclusion

In conclusion, RandomForestClassifier demonstrates promising performance in passenger detection based on FFT features extracted from ADC data collected from car seats. The utilization of a different threshold in the sensor setup enhances data preprocessing, potentially improving the quality of features extracted for classification. Further experimentation and parameter tuning may lead to refinements in model performance. The findings of this study contribute valuable insights into the application of ensemble learning techniques for passenger detection in automotive systems.

By employing RandomForestClassifier with ADC data and incorporating threshold adjustments in data preprocessing, this study provides a novel approach to passenger detection in car seats, offering potential enhancements in model performance and reliability.