Activity Recognition in Smart Home Data and Perspective of Adversarial Network on such Classification Network

Mahbuba Tasmin  
ID : 1610064042

Sharif Uddin Ruman  
ID: 1611557642

Md. Shahnawaz Zulminan

ID: XXXXXXXXX

Taoseef Ishtiak  
ID: 1612142042

Taoseef.ishtiak@northsouth.edu

Sajid Ahmed  
ID: XXXXXXXXX

N.M. Shihab Islam  
ID: 1420339042

Arif ur Rahman Chowdhury Shuhan  
ID: 1612142042

Sifat JahanID: 1611702642

Abdur Raufus SaleheenID: 1610472642

.

# Methodology

The primary goal of this research work is to activity detection through recognizing the pattern of data mined in **DATSET NAME** [dataset paper citation]. Primarily 5 distinct activities have been selected to train for pattern recognition purpose. The end goal of the research work is to inject adversarial attack on the model to confuse the network and identify the actual activity after the injection. To suggest more amicable work based on such data, the research team is exploring variety of fields in health, administration and security issues where such dataset generation and model implementation will be useful for activity recognition. Before fitting the dataset into the explored classifiers for activity recognition several preprocessing techniques have been applied for statistical analysis of the attributes the dataset to reduce those number of features that do not contribute to training. The research team believes the feature selection approach not only reduce the number of training time and computational cost but also will reduce the variance of the model, thus avoiding overfitting. The following section describes the feature selection techniques that the research team has applied for feature selection and the theoretical background of the techniques.

## A. Low Variance Feature Removal:

The low variance feature selection technique removes the features which is found to be constant mostly. The constant value of a feature is not very interesting to find pattern and can be removed from the dataset. For dataset with large attributes the scikit-learn library automatically identifies the features which have the lowest variance. The heuristic approach before running the feature selection techniques is to use a threshold value to use as cut-off. The feature elimination is run when any features comes beneath this threshold value. On the given threshold the library computes the covariance against each tuple of the dataset and generates the result. The research team has kept a threshold of 80% as the threshold.

B. L1 Based Feature Selection:

In SVM the parameter C controls the sparsity of the vectors. The smaller C is the fewer features elected. In large number of samples, the L1 model perform at random where it depends on the number of non-zero coefficients, the logarithm number of features, the amount of noise, the smallest absolute value of non-zero coefficients and the structure of the design matrix. The design matrix must contain the property of not being too correlated.

C. Tree-based Feature Selection:

The tree-based estimators are used to compute the importance of features and to discard the irrelevant features.

D. Feature Selection with Random Forest:

Random forest classifier uses the tree-based strategies to rank the features for improving purity of the node.

Table **X**. describes the result acquired by running the above mentioned approaches and the selected features which have been found important across different approach.

Table 1 Feature Selection Results Across Different Approaches

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Original Set of Features** | **Selected Features with Low Variance Feature Removal** | **Selected Featured with L1 Based Feature Selection** | **Selected Features with Tree-based Feature Selection** | **Feature Selection with Random Forest** |
| lastSensorEventHours | lastSensorEventHours | lastSensorEventHours | lastSensorEventHours | lastSensorEventHours |
| lastSensorEventSeconds | lastSensorEventSeconds | lastSensorEventSeconds | lastSensorEventSeconds | lastSensorEventSeconds |
| lastSensorDayOfWeek | lastSensorDayOfWeek | lastSensorDayOfWeek | lastSensorDayOfWeek |  |
| windowDuration | windowDuration | windowDuration |  | windowDuration |
| timeSinceLastSensorEvent | timeSinceLastSensorEvent | timeSinceLastSensorEvent |  |  |
| prevDominantSensor1 | prevDominantSensor1 | prevDominantSensor1 | prevDominantSensor1 |  |
| prevDominantSensor2 | prevDominantSensor2 | prevDominantSensor2 |  |  |
| lastSensorID | lastSensorID | lastSensorID |  |  |
| lastSensorLocation | lastSensorLocation | lastSensorLocation | lastSensorLocation |  |
| lastMotionLocation | lastMotionLocation | lastMotionLocation | lastMotionLocation | lastMotionLocation |
| complexity | complexity | complexity |  |  |
| activityChange | activityChange | activityChange |  |  |
| areaTransitions | area transitions | area transitions |  |  |
| numDistinctSensors |  |  |  |  |
| sensorCount-Bathroom | sensorCount-Bathroom | sensorCount-Bathroom |  |  |
| sensorCount-Bedroom | sensorCount-Bedroom | sensorCount-Bedroom |  |  |
| sensorCount-Chair | sensorCount-Chair | sensorCount-Chair |  |  |
| sensorCount-DiningRoom | sensorCount-DiningRoom | sensorCount-DiningRoom |  |  |
| sensorCount-Hall | sensorCount-Hall | sensorCount-Hall |  |  |
| sensorCount-Ignore | sensorCount-Ignore | sensorCount-Ignore |  |  |
| sensorCount-Kitchen | sensorCount-Kitchen | sensorCount-Kitchen | sensorCount-Kitchen | sensorCount-Kitchen |
| sensorCount-LivingRoom | sensorCount-LivingRoom | sensorCount-LivingRoom | sensorCount-LivingRoom |  |
| sensorCount-Office | sensorCount-Office | sensorCount-Office |  | sensorElTime-Bedroom |
| sensorCount-OutsideDoor | sensorCount-OutsideDoor | sensorCount-OutsideDoor |  |  |
| sensorCount-WorkArea | sensorCount-WorkArea | sensorCount-WorkArea |  |  |
|  |  |  |  |  |
| sensorElTime-Bathroom | sensorElTime-Bathroom | sensorElTime-Bathroom | sensorElTime-Bathroom | sensorElTime-Bathroom |
| sensorElTime-Bedroom | sensorElTime-Bedroom | sensorElTime-Bedroom |  |  |
| sensorElTime-Chair | sensorElTime-Chair |  | sensorElTime-Chair | sensorElTime-Chair |
| sensorElTime-DiningRoom | sensorElTime-DiningRoom | sensorElTime-DiningRoom | sensorElTime-DiningRoom |  |
| sensorElTime-Hall | sensorElTime-Hall | sensorElTime-Hall |  |  |
| sensorElTime-Ignore | sensorElTime-Ignore | sensorElTime-Ignore |  |  |
| sensorElTime-Kitchen | sensorElTime-Kitchen | sensorElTime-Kitchen | sensorElTime-Kitchen | sensorElTime-Kitchen |
| sensorElTime-LivingRoom | sensorElTime-LivingRoom | sensorElTime-LivingRoom |  |  |
| sensorElTime-Office | sensorElTime-Office | sensorElTime-Office |  |  |
| sensorElTime-OutsideDoor | sensorElTime-OutsideDoor | sensorElTime-OutsideDoor | sensorElTime-OutsideDoor | sensorElTime-OutsideDoor |
| sensorElTime-WorkArea | sensorElTime-WorkArea | sensorElTime-WorkArea | sensorElTime-WorkArea |  |

From Table. 1, it can be inferred that more number of features has been eliminated in the tree based and random forest based approach. **(Shihan ke eishb feature er rank ber korar code disi ota kroe dile er pore add kore dis)** The next approach of the research team would be train a number of models against all the selected features generated by different approach and to find out the best combination of features.