

WSN Project

Classification of ADL data collected from WSN

Anusha Prakash	(12IT10)
Bhuvan M S	(12IT16)
Siddharth Jain	(12IT78)

Introduction

- ☐ ADL activity data collected from IoT applications
- ☐ Processing of these data to build application
- ☐ WSN to collect data
- ☐ Data Science and ML to analyse the data collected

Problem Statement

□ The aim of our project is to classify the ADL of the user using the data sensed by sensor nodes in the wireless sensor networks by modeling separate user dependent machine learning model and analyse the performance of various classifiers.

Objectives

1. To develop ADL prediction model on the data collected from sensor nodes in a Wireless Sensor Network.
2. Evaluate each prediction model using Receiver Operating Characteristics (ROC), Precision-Recall Curves, F1 Measure.
3. To implement simple version of HEED protocol to illustrate how the ADL data could be collected from the Wireless Sensor Network.

Methodology

—

Workflow

- 1.Data cleaning and preprocessing
- 2.Applying different ML techniques and evaluating performance
- 3.Plotting ROC and PR curves



Data Cleaning and Preprocessing

- Merging Data Sensors and Activity files
- Deriving the Time Difference as a Duration feature.
- Converting it to .arff file format.

Data Scale and Feature Set

Data Scale: Number of Instance: 408 and 2335 for User A and User B respectively
Feature Set:

Feature	Type
Start Timestamp	Numeric
End Timestamp	Numeric
Duration	Numeric
Location	Nominal
Type	Nominal
Place	Nominal
Class: Activity	10 Nominal Classes

Learning

- *Naive Bayes Classifier* - Simple probabilistic classifier based on Bayes probability
- *MLPC* – Neural Network with one hidden layer with 2 hidden nodes.
- *Bayesian Network* - Directed Acyclic Graphs whose nodes represent random variables
- *Random Forest* – Ensemble model with tree – subsamples, using feature subspace
- *Adaboost* -
- *Logistic Regression* -

Evaluation metrics

- Accuracy = $(TP + TN) / (TP + TN + FP + FN)$
- Precision = $TP / (TP + FP)$
- Recall = $TP / (TP + FN)$
- F1 = $(2 * Precision * Recall) / (Precision + Recall)$
- ROC Curves - TP vs FP
- PR Curves - recall vs precision

HEED Protocol

Hybrid, Energy-Efficient Distributed Clustering

- *Hybrid, Energy-Efficient Distributed Clustering – Hierarchical Routing Protocol*
- Residual node energy is used as a metric for cluster selection to achieve power balancing
- Selection of CHs - Residual energy of each sensor node and Intra-cluster communication cost as a function of cluster density or node degree

Goals of HEED Protocol

- *Prolonging network lifetime by distributing energy consumption*
- *Terminating the clustering process within a constant number of iterations*
- *Minimizing control overhead*
- *Producing well-distributed CHs and compact clusters*

$$CH_{prob} = C_{prob} \times \frac{E_{residual}}{E_{max}},$$

Results and Analysis

—

User A

ROC and PR Curves

User A

Summary:

Algorithm	F1_Score	Accuracy
MLP_CLASSIFIER_2	0.328498863313	0.441176470588
LOGISTIC_REGRESSION	0.752633628582	0.764705882353
BAYES_NETWORK	0.720978373888	0.735294117647
NAIVE_BAYES	0.764874797931	0.764705882353
RANDOM_FOREST	0.64279135996	0.656862745098
ADABOOST	0.110876854162	0.254901960784

User B

ROC and PR Curves

User B

Summary:

Algorithm	F1_Score	Accuracy
MLP_CLASSIFIER_2	0.278021577872	0.349914236707
LOGISTIC_REGRESSION	0.287792488475	0.360205831904
BAYES_NETWORK	0.317570449807	0.391080617496
NAIVE_BAYES	0.289539715817	0.36192109777
RANDOM_FOREST	0.309216828748	0.368782161235
ADABOOST	0.0966266437965	0.245283018868

HEED Protocol

1,	0.0000,	0.0000,	4,	2,	3,	4,	5					
2,	20.0000,	20.0000,	5,	1,	3,	4,	5,	6				
3,	40.0000,	40.0000,	6,	1,	2,	4,	5,	6,	7			
4,	60.0000,	60.0000,	7,	1,	2,	3,	5,	6,	7,	8		
5,	80.0000,	80.0000,	8,	1,	2,	3,	4,	6,	7,	8,	9	
6,	100.0000,	100.0000,	8,	2,	3,	4,	5,	7,	8,	9,	10	
7,	120.0000,	120.0000,	8,	3,	4,	5,	6,	8,	9,	10,	11	
8,	140.0000,	140.0000,	8,	4,	5,	6,	7,	9,	10,	11,	12	
9,	160.0000,	160.0000,	8,	5,	6,	7,	8,	10,	11,	12,	13	
10,	180.0000,	180.0000,	8,	6,	7,	8,	9,	11,	12,	13,	14	
11,	200.0000,	200.0000,	8,	7,	8,	9,	10,	12,	13,	14,	15	
12,	220.0000,	220.0000,	8,	8,	9,	10,	11,	13,	14,	15,	16	

HEED Protocol

1,	-1,	4.0000
2,	-1,	4.0000
3,	-1,	4.0000
4,	-1,	4.0000
5,	-1,	4.0000
6,	-1,	4.0000
7,	-1,	4.0000
8,	-1,	4.0000
9,	-1,	4.0000
10,	-1,	4.0000
11,	-1,	4.0000
12,	-1,	4.0000
13,	-1,	4.0000
14,	-1,	4.0000
15,	-1,	4.0000
16,	-1,	4.0000
17,	-1,	4.0000
18,	-1,	4.0000
19,	-1,	4.0000
20,	-1,	4.0000

1	1	3.98
2	2	3.534
3	5	3.1
4	2	3.3005
5	5	3
6	16	3.505
7	19	3.5
8	1	3.06
9	5	3.064
10	2	3.01
11	5	3.0070
12	19	3.04
13	5	3.0305
14	1	3.0004
15	3	3.083
16	16	3.09
17	17	3.059
18	17	3.009
19	19	3.09
20	7	3.0707

Thank You

