Human Activity Recognition

Jake Sant[1] [117699M], Aiden Williams[2] [372001L], Ethan Zammit[3] [4802L]
Department of Artificial Intelligence
University of Malta
jake.sant.18@um.edu.mt[1], *aiden.williams.19@um.edu.mt*[2], *ethan.zammit.19@um.edu.mt*[3]

***Abstract*—Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.**

***Keywords*—Human activity, Classification, Deep learning, SVM**

I. Introduction

Human Activity

II. Aims and Objectives

The main objective our group aimed to accomplish was to understand and make use of deep learning and classical machine learning techniques to correctly classify a number of stationary and moving activities in the UCI dataset [1]. We also aimed to build our own dataset, consisting of many records of activities. We then wished to apply and evaluate the classification methods on our own database. The new activities in our own dataset are laying, sitting, standing, walking straight, walking upstairs, walking downstairs, driving, jumping, swimming and performing push ups.

III. Background Research

A. Smartphone Accelerometers

B. Support Vector Machines

Support Vector Machines (or SVMs) are a type of classifier used for classification and regression. SVM seeks to classify samples as different classes on a hyperplane in multidimensional space. The classification process generates multiple hyperplanes to find the one which minimises an error value. Generally, support vector machines are used for binary classification problems. However, they can be adapted to multiclass problems by converting or reducing them into a set of multiple binary classification problems. However some samples cannot be classified linearly, hence SVMs can also make use of kernels to transform non-linear space into linear space, which was proposed by Boser et al. [2].

C. Convolutional Neural Networks

IV. Literature Review

Anguit et al. [3] classified activities into two different categories: static and dynamic activities. Using a One-Vs-All approach and a Laplacian kernel, they obtained highly accurate rates of classification. Demrozi et al. [4] compiled a number of studies on human activity recognition (Human Activity Recognition using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey) and found that out of 149 papers published between January 2015 and September 2019, 53 examined deep learning models and the remaining 96 were classical machine learning models.

V. Design and Methodology

A. Feature Extraction

Excluding the label for each activity, each session in the dataset contains 589 features. These features were extracted by applying a number of different statistical measures to the different extracted signals in each sliding window.

Each signal has the statistical measures in Table 1 applied to them.

Statistical Measure	**Description**
mean	Mean value in the window
min	Smallest value in the window
max	Largest value in the window
std	Standard deviation
entropy	Entropy
mad	3
iqr	Inter-quartile Range
energy	3
sma	Signal magnitude area
arCoeff	3
correlation	3
angle	Angle between 2 vectors of signals
band energy	3

Table 1

The statistical measures in Table 2 are applied only to FFT signals.

Statistical Measure	Description
maxInds	1
skewness	2
kurtosis	3
meanFreq	3

Table 2

B. Signal Processing

C. Support Vector Classifier

1) Data Pre-processing: Firstly, the training and testing datasets were converted to Pandas dataframes and the labels for each activity were separated into their own variables. Using a LabelEncoder, each activity label is converted into a numerical value.

2) Comparing different classification models: In total, four different classic machine learning classifiers were used on the dataset developed by Anguita et al [3]. These classifiers are Gaussian Naïve Bayes, AdaBoost, Stochastic Gradient Descent and a Support Vector Classifier. These were trained and tested using the respective datasets, and their accuracy, F-beta, precision and recall scores were recorded.

Classifier	Accuracy	F-Beta	Precision	Recall
Gaussian Naïve Bayes	0.7134	0.7252	0.7555	0.7134
AdaBoost	0.4065	0.2520	0.4289	0.4065
Stochastic Gradient Descent	0.9600	0.9603	0.9605	0.9600
Support Vector Classifier	0.9668	0.9676	0.9682	0.9668

Table 3: Performance scores of each classifier

As can be seen in Table 3, the Support Vector Classifier had the highest scores, each result being over 0.96. It is for this reason that the SVC was chosen to classify the UCI dataset [1] and, later on, the dataset built by ourselves.

D. Classification

The RBF kernel is defined as the exponential function $exp(-\gamma|x - x'|)^2$ A primer on kernel methods, JP Vert et al, where x and x' are two feature vectors, and is the gamma parameter in the classifier. Gamma's value is scale, meaning that the parameter is the reciprocal of the number of features multiplied with the variance of the input data. For this implementation, we opted for a One-Vs-All approach. A One-Vs-All approach divides the data points into just two classes: a certain activity X and the other classes. Therefore records labelled as *SITTING* are a single class, and the other activities are treated as having a single label.

VI. Evaluation

A. Evaluation Metrics

In order to determine which classifier provided the most accurate predictions, four different performance metrics were recorded and acted as a score. These metrics are **accuracy**, **f-beta**, **precision** and **recall**.

Accuracy is the ratio of the true labels y on the set of predicted labels y', where $accuracy(y, y') = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} 1(y'_i = y_i)$.

Precision is the ratio of true positives during prediction, and it is calculated as $P = \frac{T_P}{T_P + F_P}$, where T_P is the number of true positives and F_P is the number of false positives.

Recall is the ratio of correctly identified positive labels, and it is calculated as $R = \frac{T_P}{T_P + F_n}$, where T_P is the number of true positives and F_n is the number of false negatives.

Beta (or **F-Beta**) is calculated based on the precision and recall scores of the classifier, wherein precision is multiplied by some parameter γ, thereby giving more importance to the precision value. This is calculated as $F_\gamma = (1+\gamma^2) * \frac{P*R}{(\gamma^2 * P) + R}$. For evaluation purposes, a γ value of 0.5 was used.

B. Evaluating SVM

Figure 1 above is the confusion matrix representing the accuracy of the classification of each activity as a percentage. The leading diagonal represents correctly labelled activities. Each activity group classified had an accuracy of 96% or higher, with half activities being classified at an accuracy of 99-100%. 1% of driving activities were misclassified as sitting. This was to be expected as during driving sessions there were times (i.e. being stuck in traffic) where the driver was stationary and in the same position as that of someone seated down. 2% of laying activities were misclassified as sitting, this is understandable as the two stationary activities have similar positions.

VII. Conclusion and Future Work

The purpose of this report is to discuss the problem, view the existing solutions and to present the work that has been done to date mainly the automatic user profiling prototype. Further development regarding both the improvement of such a prototype and the development of the automatic itinerary generation will take place in the upcoming weeks.

VIII. Distribution of Work

- Jake Sant
- Aiden Williams
- Ethan Zammit

References

[1] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A Public Domain Dataset for Human Activity Recognition Using Smartphones," in *21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, Bruges, 2013. [Online]. Available: https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2013-84.pdf

[2] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory - COLT '92*. New York, New York, USA: ACM Press, 1992, pp. 144–152. [Online]. Available: http://portal.acm.org/citation.cfm?doid=130385.130401

[3] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7657 LNCS, pp. 216–223, 2012. [Online]. Available: http://link.springer.com/10.1007/978-3-642-35395-6_30

[4] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human Activity Recognition Using Inertial, Physiological and Environmental Sensors: A Comprehensive Survey," *IEEE Access*, vol. 8, pp. 210 816–210 836, 2020.