Methods for Recognition and Classification of Human Behaviour Through Sensor-Driven Insights

Mark Turos
Faculty of Engineering, Environment and Computing
Coventry University
Coventry, U.K.
turosm@uni.coventry.ac.uk

Abstract – This paper presents a compelling case for using machine learning to tackle human behaviour classification. Utilising a comprehensive dataset of sensory data (accelerometers, gyroscopes, magnetometers) from subjects performing 19 different activities, the spatial data then underwent aggregation, pre-processing, standardisation, and PCA application. In order to ensure optimal use with various machine learning models, including Linear Regression, Decision Trees, Gaussian Naive Bayes, and SVMs. The results indicated a high accuracy in activity classification by the majority of the models, demonstrating the potential for adaptation or expansion of the work explored in this paper, optimistically leading to applications in the key industries.

Ⅰ. Introduction - Sensor-Driven Insights of Human Behaviour

Human Activity Recognition or HAR [1] has become an increasingly important area of research, empowered by widespread availability of cost-effective sensory devices. This trend then is able to fuel advancements in a multitude of industries such as healthcare, sports, VR, enabling important applications such as rehabilitation, health monitoring, VR haptic feedback suits [2], athletic performance and many others. Traditional methods, relying on hand-crafted features, proved economically impractical, paving the way for machine learning algorithms to efficiently recognize and adapt to diverse movement patterns [3].

The current state of HAR research, while innovative, has yet to fully embrace cutting-edge techniques such as Convolutional Neural Networks (CNNs) [4], which show promise in learning complex data patterns. However, challenges persist, including data privacy and the need for real-time feedback, underscoring the importance of advancements in edge computing [5][6][7].

The field's dependency on visual methods also faces privacy and environmental limitations, suggesting a pivot towards sensor-based approaches to mitigate bias and security risks.

This paper explores sensor-based HAR by analysing a dataset from 8 participants performing 19 activities, captured through sensors, such as accelerometers, gyroscopes and magnetometers, on five body parts at a 25 Hz frequency, resulting in a vast collection of spatial data[8]. Employing machine learning models like Logistic Regression, Decision Trees, Naive Bayes, and SVMs, the study achieves precise activity classification, underlining the potential for accurate activity prediction while addressing the limitations of visual-based systems.

Ⅱ. Dataset Description

In this paper we have not collected the data ourselves, but rather provided by Billur Barshan and Kerem Altun at UC Irvine [8].

The dataset itself has been generated via inertial electrical sensors attached to the subject’s body. The main point of interest being the 19 activities performed by the subjects that have the demographics of an even split study of 1:1 male:female ratio for a sample size of 8 people and an age range of 20 - 30. The sensors collected information over a time span of 5 minutes per activity, each split by 5-second segments i.e. 60 segments per activity per person. This in turn resulted in the generation of 480 signal segments for each of the 19 activities. The 19 activities dictionary is shown in ***Table*** Ⅰ.

TABLE Ⅰ
Activities Performed - Dictionary

Activity ID	Activity
A1	Sitting
A2	Standing
A3	Lying on back
A4	Lying on right side
A5	Ascending stairs
A6	Descending stairs
A7	Standing in elevator
A8	Standing in moving elevator
A9	Walking in parking lot
A10	Walking on treadmill (4 km/h)
A11	Previous act. with a 15° incline

A12	Running on treadmill (8 km/h)
A13	Exercising on stepper
A14	Exercising on cross trainer
A15	Cycling on exercise bike (horizontal)
A16	Cycling on exercise bike (vertical)
A17	Rowing
A18	Jumping
A19	Playing basketball

When looking at the sensors themselves, in order to capture a broad spectrum of datapoints, 3 different types of sensory devices were being used, accelerometers, gyroscopes and magnetometers. The sensors themselves are generating each datapoint at a sampling frequency rate of 25 Hz/s which over the 5-second segments results in 125 data entries, 25 Hz/s x 5s = 125 entries.

Each being placed strategically on the subject’s anatomy, more specifically Torso(T), Right Arm(RA), Left Arm(LA), Right Leg(RL), Left Leg(LL) which results in 5 body parts and 3 types of sensors each therefore 5bp x 3st = 15 inputs. This is then further expanded via the 3 degrees of freedom x,y,z such that 3dg x 15inp = 45 total distinct characteristics which is then further multiplied by the 125 entries resulting in an impressive 5625 features.

One of the main challenges of utilising sensory data is that it is very difficult to spot trends via analysing the actual data points themselves. Meaning that since the data provided is just an amalgamation of x,y,z coordinates in a 3D plane the values themselves don't inform the researcher of any outliers, negative/positive trends or do not provide any justification. This is built into the nature of this and many other similar datasets. However, this is a rather favourable position for us as this paper will then be able to highlight the strength of using machine learning techniques in generating outlooks, predictions, classifications or showcasing trends that would otherwise be omitted.

The dataset is extensive, structured in a hierarchy of activity, person, and segments across 19 activities, 8 participants, and 60 segments, creating a sizable directory of approximately 415 MB. After initial aggregation, the dataset encompasses over 1.1 million rows, each detailing the 45 distinct features across the various body parts and sensors, showcasing the depth and scale of data available for analysis.

Ⅲ. Methodology

I would like to briefly discuss the machine learning techniques and algorithms used in this paper and the reasoning behind their selection.

1. Logistic Regression

One of the most popular statistical methods used in classification tasks is Logistic Regression. It models the probability of a binary outcome as a function of one or more predictor variables [9]. A logistic function is then used to estimate the likelihood of an event occuring, thus making it a primary tool in our arsenal in tackling the problem highlighted by this paper.

2. Decision Trees

Decision Trees are algorithms that take data and partition it into various subsets. This is done via the construction of a tree-like model that uses the laid out features of the data set [10]. It is an effective method as it iteratively splits the dataset, optimises the separation based on features' values thus leading to an effective structured prediction mechanism. Another key addition, as the multidimensional input of spatial data is handled well through the segmentation showcased in this model, thus facilitating an intuitive approach to prediction of complex datasets.

3. (Gaussian) Naive Bayes

The Gaussian Naive Bayes classifier utilises Bayes' theorem under the assumption that features can independently predict the outcome. The Gaussian attribute is used as a method of verification of whether or not the spatial data can follow a normal distribution, quite an impactful consideration as it will allow for future models to be parameterized accordingly.

The main hypothesis for using this methodology is that there's a >0 chance that certain activities' features can be independently used to predict the outcome [11], therefore if hypothesis is proved that will result in a sizable decrease in the data needed. This will result in efficiency gains. Looking at activity groups such as A3/4, A5/6, A8/9 Gaussian Naive Bayes can provide a robust framework for uncovering subtle patterns and similarities thus being able to classify accordingly.

4. SVMs (Support Vector Machines)

SVMs are a set of supervised learning methods used for classification [12], regression problems where the core of the mechanism is searching for a hyperplane that best separates the different classes in the feature space. This is then achieved by maximising the margin between the support vectors and said hyperplane. They are an incredible tool at handling high complex data with a sizable amount of features.

Which in turn makes it an excellent addition to our arsenal of tackling the problem set out in this paper. The ability of SVMs to distinguish between the high-dimensional sensor data representing different activities and their ability to work in multi-dimensional spaces allows us to exploit the patterns within the coordinate data.

Ⅳ. Experimental Setup

The dataset, structured in a hierarchy with 19 activity folders, each containing 8 participant folders and 60 segment files per participant, was first aggregated into a single text file, with data from each sensor on corresponding body parts labelled (e.g., T_xacc, T_yacc).

To prepare for Principal Component Analysis (PCA) [13], the data underwent pivoting to organise the 125 features per segment into a single row (e.g., T_xacc_1 to T_zmag_125), followed by standardisation to normalise the distribution. This step is of imperative importance to address imbalances and enhance algorithmic performance.

Given the data's high dimensionality, PCA was applied to reduce it to a manageable size while retaining essential variance, demonstrated by a logarithmic graph indicating a few principal components effectively capture the dataset's variability (***Figure1***). This preparatory work enables the use of machine learning techniques suited for the refined dataset structure.

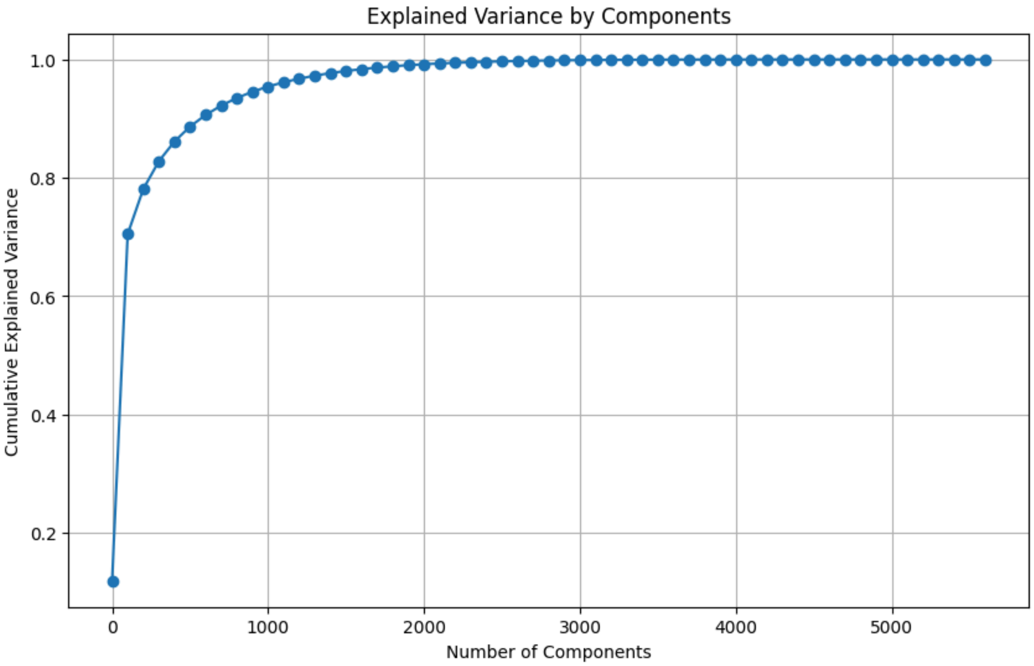


Figure 1 - Explained Variance by Components

Therefore, the PCA dataset is generated holding a significant number of components, selected to capture 90% of the variance, this allows for the expansion of the classification algorithms' capabilities whilst also maintaining a manageable and relatively efficient dataset. ***Figure 2*** represents the initial 30 principal components as a way of showcasing the efficacy of the PCA algorithm in selecting powerful values to abstractly represent the data.

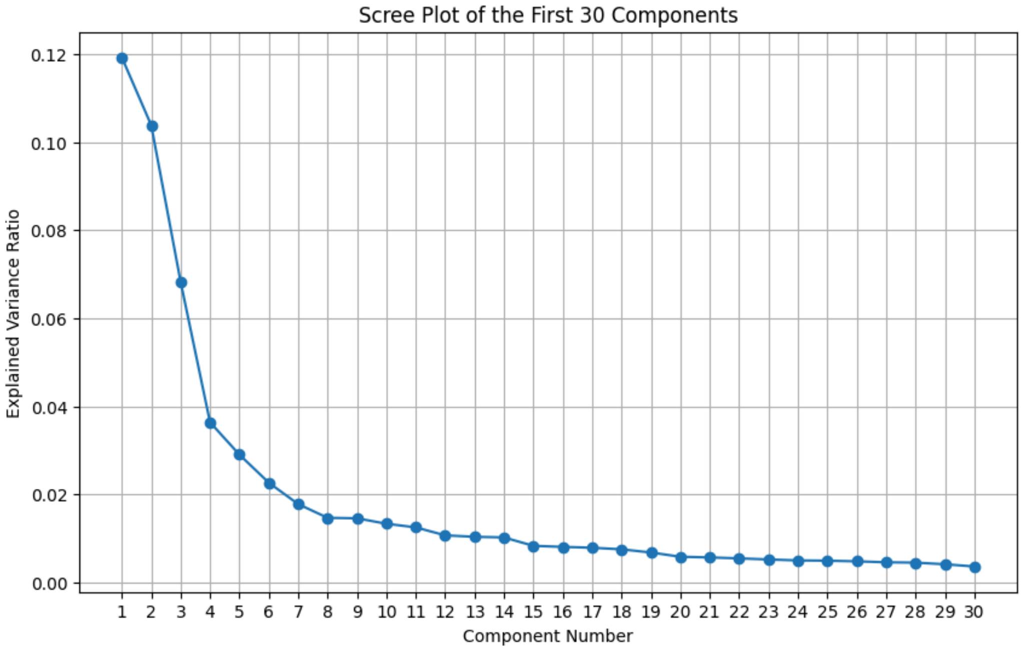


Figure 2 - Scree Plot of 30 Initial Components

Ⅴ. Results

In the results and evaluation section of this study I present an analysis of the performance of the Logistic Regression, Decision Tree, Gaussian Naive Bayes, SVMs, algorithms. As well as their application to the PCA transformed dataset.The comparison focuses on F1 scores (***Table*** Ⅱ), harmonic means of precision and recall, offering a balanced metric for assessing model accuracy by equally considering both precision and recall to provide a comprehensive evaluation of the classifiers' effectiveness.

TABLE Ⅱ
F1 score comparison

Model	F1 Score
Logistic Regression	0.93
Decision Tree	0.91

Gaussian Naive Bayes	0.85
SVM	0.95

As can be seen in the table, LR, DT and SVM (**0.93** / **0.91** / **0.95**) performed similarly due to the nature of our dataset that is well suited for linear models. Another factor could be the pre-processing done on the dataset before modelling, such as normalisation, PCA, this then allows for highly representative features to be used. Conversely, the Gaussian Naive Bayes model underperformed with an F1 score of 0.85, likely due to its independence assumption clashing with the interdependent nature of human activity data and the similar functionalities of accelerometers and gyroscopes. This discrepancy results in the conclusion that GNB is not optimally effective for the context of this paper, compared to its counterparts.

The study further evaluates model performance via the use of confusion matrices (***Figure 3-6 in Appendix***), visualisations that effectively illustrate the prediction accuracy across the different activities. LR, DT and SVMs show similar performance profiles, strong diagonal elements and weak off-diagonal values indicate high accuracy. With the mentionable exceptions of activities A7 and A8, standing in elevator vs moving elevator, where differentiation seems to be challenging, possibly due to the minimal sensor variation, as accelerometer and magnetometer would show closely linked values due to the nature of the activity. As a final point I would like to emphasise the Decision Tree model, with characteristics of a depth of tree of: 22 and number of leaf nodes being 276. This can be both a positive and a negative, as the tree is able to be flexible and handle the large dataset, on the other hand there’s a high chance of overfitting which will result in poorer results in unseen data. This can be optimised via pruning and regularising.

The Gaussian Naive Bayes model's lower efficacy, as previously indicated by its F1 score, is corroborated by its confusion matrix. Here, weaker diagonal values and higher off-diagonal counts reflect the model's weakness when applied to our data, further underscoring the importance of considering model-specific strengths and weaknesses in relation to the dataset's characteristics.

Overall the models showcase evidence of well suitability for this kind of classification problem, with relatively high and appropriate F1 score, above average to good confusion matrix scores and quite efficient in terms of runtime as most models took only around 120 seconds to train and run on a quite large dataset, of course with the exception of GNB where the performance is not as adequate. On the other hand, these models are relatively rudimentary and simplistic and cannot be fully trusted in a real world scenario where the prediction holds weight in terms of affecting people’s lives or products. In that case I recommend a stronger approach, using a more modern deep learning framework to combat these inadequacies and provide a strong prediction, a more detailed and accurate one that can be used further.

Ⅵ. Conclusion

I would like to initially summarise this paper and its findings. Overall the main purpose of this study was to use traditional machine learning algorithms, such as Linear Regression, Decision Trees, Gaussian Naive Bayes and SVMs and apply them to the problem of human activity classification via sensory gathered data.

This proved to be relatively more challenging than expected simply due to the difficulty of the data, unlabelled, abstract, unstandardized and unorganised. Upon data aggregation, standardisation and feature extraction the modelling proved to be quite comfortable, bearing in mind that I mainly used the default models provided by the scikit-learn library and did not venture into changing the parameters too much. As I feel that would be outside the scope of this paper specifically.

As a main point of future improvement, I recommend taking the prepared data and applying either more adjusted models or even higher complexity algorithms such as deep learning's CNNs or RNNs and using the available popular PyTorch or Tensorflow libraries for ease-of-use. This will then hopefully provide more interesting outcomes with higher discrepancies between the models, which will result in a more selective alternative to what has been currently laid out here.

Finally I would like to underscore the need for careful navigation of the legal, social, and ethical effects of this and similar studies. Privacy concerns necessitate robust data protection measures in line with legal standards like GDPR. Socially, the technology promises significant benefits in healthcare and lifestyle but must be accessible to all to prevent a digital divide. Ethically, accuracy and fairness are paramount to avoid unintended consequences from misclassifications or biases. As we harness these technologies, balancing innovation with ethical responsibility is essential for their positive integration into society.

Ⅶ. Appendix

Figures

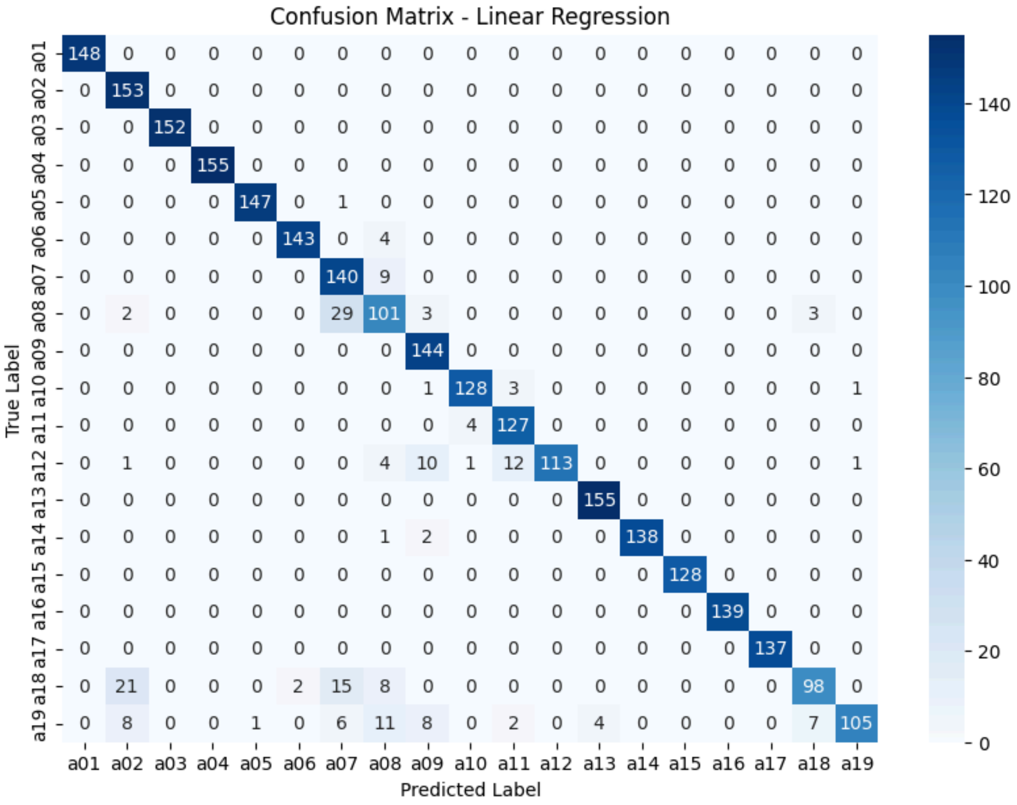
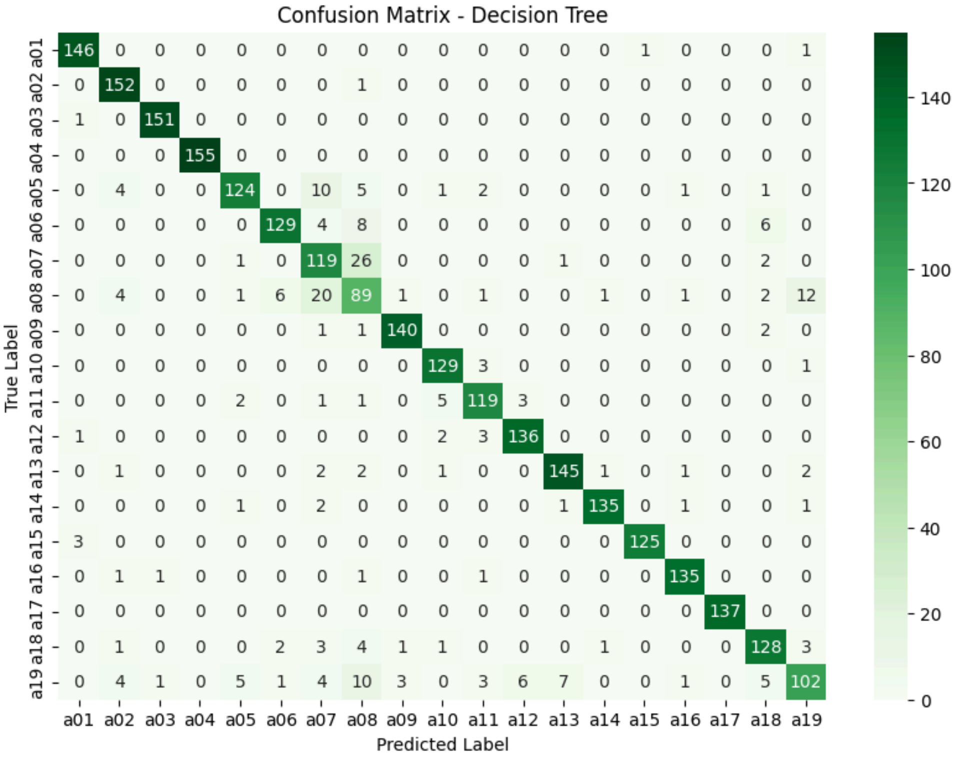
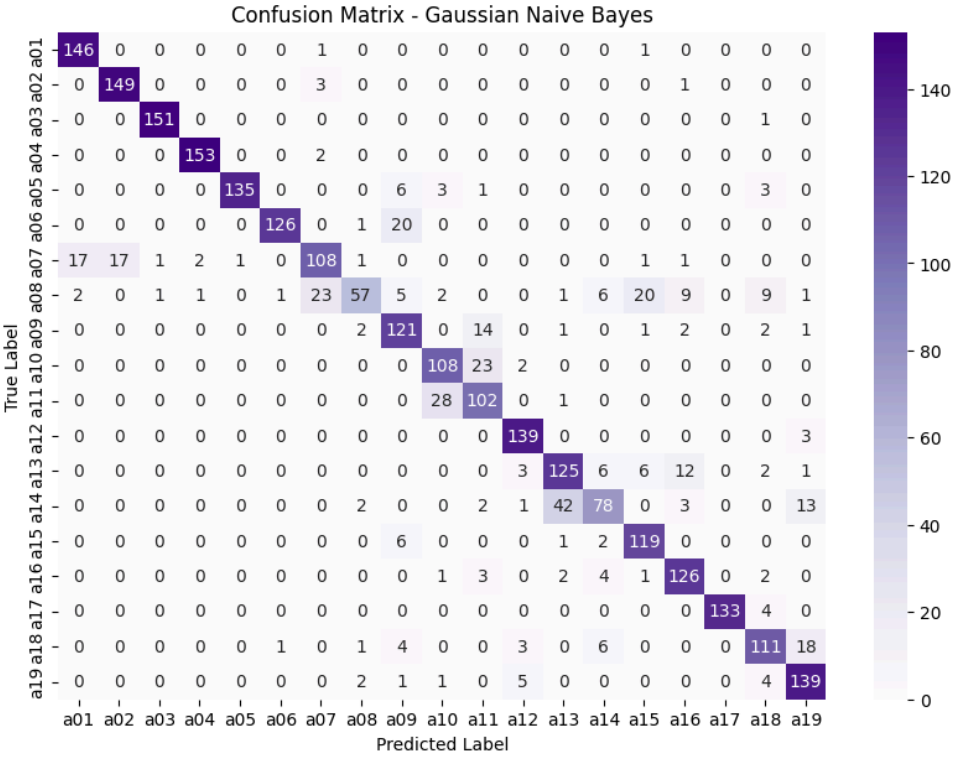


Figure 3 - Confusion Matrix - Logistic Regression



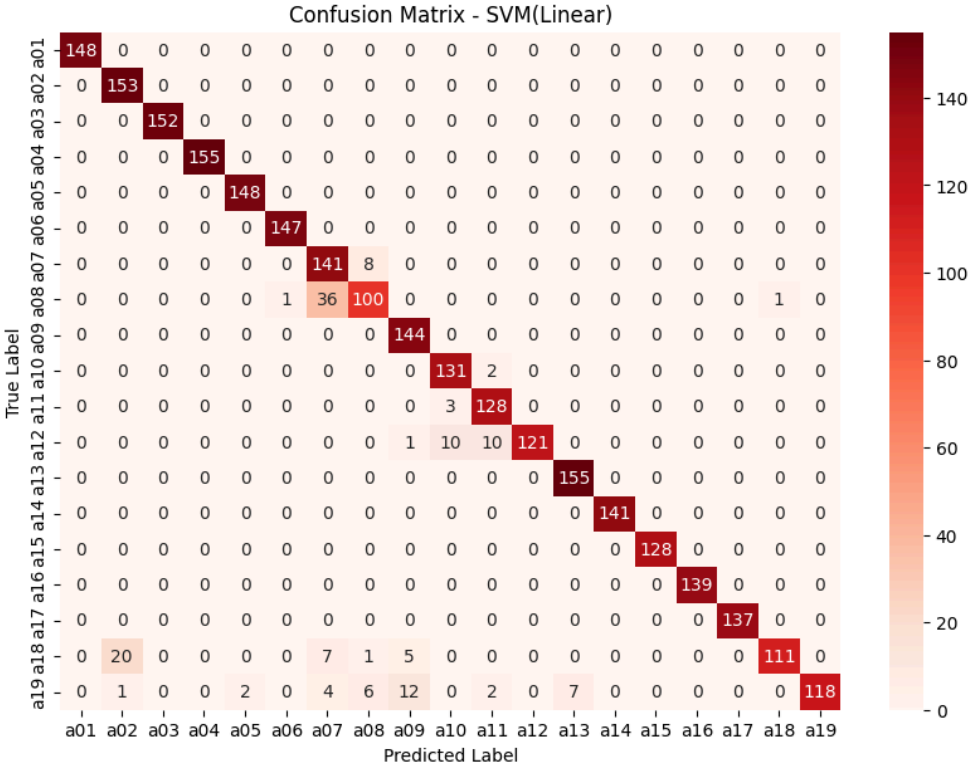
	a01	a02	a03	a04	a05	a06	a07	a08	a09	a10	a11	a12	a13	a14	a15	a16	a17	a18	a19
a01	146	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
a02	0	152	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
a03	1	0	151	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a04	0	0	0	155	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a05	0	4	0	0	124	0	10	5	0	1	2	0	0	0	0	1	0	1	0
a06	0	0	0	0	0	129	4	8	0	0	0	0	0	0	0	0	0	6	0
a07	0	0	0	0	1	0	119	26	0	0	0	0	1	0	0	0	0	2	0
a08	0	4	0	0	1	6	20	89	1	0	1	0	0	1	0	1	0	2	12
a09	0	0	0	0	0	0	1	1	140	0	0	0	0	0	0	0	0	2	0
a10	0	0	0	0	0	0	0	0	0	129	3	0	0	0	0	0	0	0	1
a11	0	0	0	0	2	0	1	1	0	5	119	3	0	0	0	0	0	0	0
a12	1	0	0	0	0	0	0	0	0	2	3	136	0	0	0	0	0	0	0
a13	0	1	0	0	0	0	2	2	0	1	0	0	145	1	0	1	0	0	2
a14	0	0	0	0	1	0	2	0	0	0	0	0	1	135	0	1	0	0	1
a15	3	0	0	0	0	0	0	0	0	0	0	0	0	0	125	0	0	0	0
a16	0	1	1	0	0	0	0	1	0	0	1	0	0	0	0	135	0	0	0
a17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	137	0	0
a18	0	1	0	0	0	2	3	4	1	1	0	0	0	1	0	0	0	128	3
a19	0	4	1	0	5	1	4	10	3	0	3	6	7	0	0	1	0	5	102

Figure 4 - Confusion Matrix - Decision Tree



	a01	a02	a03	a04	a05	a06	a07	a08	a09	a10	a11	a12	a13	a14	a15	a16	a17	a18	a19
a01	146	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
a02	0	149	0	0	0	0	3	0	0	0	0	0	0	0	0	1	0	0	0
a03	0	0	151	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
a04	0	0	0	153	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
a05	0	0	0	0	135	0	0	0	6	3	1	0	0	0	0	0	0	3	0
a06	0	0	0	0	0	126	0	1	20	0	0	0	0	0	0	0	0	0	0
a07	17	17	1	2	1	0	108	1	0	0	0	0	0	0	1	1	0	0	0
a08	2	0	1	1	0	1	23	57	5	2	0	0	1	6	20	9	0	9	1
a09	0	0	0	0	0	0	0	2	121	0	14	0	1	0	1	2	0	2	1
a10	0	0	0	0	0	0	0	0	0	108	23	2	0	0	0	0	0	0	0
a11	0	0	0	0	0	0	0	0	0	28	102	0	1	0	0	0	0	0	0
a12	0	0	0	0	0	0	0	0	0	0	0	139	0	0	0	0	0	0	3
a13	0	0	0	0	0	0	0	0	0	0	0	3	125	6	6	12	0	2	1
a14	0	0	0	0	0	0	0	2	0	0	2	1	42	78	0	3	0	0	13
a15	0	0	0	0	0	0	0	0	6	0	0	0	1	2	119	0	0	0	0
a16	0	0	0	0	0	0	0	0	0	1	3	0	2	4	1	126	0	2	0
a17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	133	4	0
a18	0	0	0	0	0	1	0	1	4	0	0	3	0	6	0	0	0	111	18
a19	0	0	0	0	0	0	0	2	1	1	0	5	0	0	0	0	0	4	139

Figure 5 - Confusion Matrix - Gaussian Naive Bayes



	a01	a02	a03	a04	a05	a06	a07	a08	a09	a10	a11	a12	a13	a14	a15	a16	a17	a18	a19
a01	148	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a02	0	153	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a03	0	0	152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a04	0	0	0	155	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a05	0	0	0	0	148	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a06	0	0	0	0	0	147	0	0	0	0	0	0	0	0	0	0	0	0	0
a07	0	0	0	0	0	0	141	8	0	0	0	0	0	0	0	0	0	0	0
a08	0	0	0	0	0	1	36	100	0	0	0	0	0	0	0	0	0	1	0
a09	0	0	0	0	0	0	0	0	144	0	0	0	0	0	0	0	0	0	0
a10	0	0	0	0	0	0	0	0	0	131	2	0	0	0	0	0	0	0	0
a11	0	0	0	0	0	0	0	0	0	3	128	0	0	0	0	0	0	0	0
a12	0	0	0	0	0	0	0	0	1	10	10	121	0	0	0	0	0	0	0
a13	0	0	0	0	0	0	0	0	0	0	0	0	155	0	0	0	0	0	0
a14	0	0	0	0	0	0	0	0	0	0	0	0	0	141	0	0	0	0	0
a15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	128	0	0	0	0
a16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	139	0	0	0
a17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	137	0	0
a18	0	20	0	0	0	0	7	1	5	0	0	0	0	0	0	0	0	111	0
a19	0	1	0	0	2	0	4	6	12	0	2	0	7	0	0	0	0	0	118

Figure 6 - Confusion Matrix - SVM(Linear kernel)

Code

Original data: https://archive.ics.uci.edu/dataset/256/daily+and+sports+activities

Google drive (all files excluding original data): https://drive.google.com/drive/folders/1NicF0__uU1sBcoMd2z9YLoiT313E_wKV?usp=sharing

Github link (source code): https://github.com/mturos19/7072CEM

References

[1] Beddiar, D.R., Nini, B., Sabokrou, M. et al. Vision-based human activity recognition: a survey. Multimed Tools Appl 79, 30509–30555 (2020) https://doi.org/10.1007/s11042-020-09004-3

[2] Kang, D., Lee, CG. & Kwon, O. Pneumatic and acoustic suit: multimodal haptic suit for enhanced virtual reality simulation. Virtual Reality 27, 1647–1669 (2023). https://doi.org/10.1007/s10055-023-00756-5

[3] A. Pentland, "Looking at people: sensing for ubiquitous and wearable computing," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 107-119, Jan. 2000, doi: 10.1109/34.824823.

[4] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998. doi: 10.1109/5.726791.

[5] Al-Faris M, Chiverton J, Ndzi D, Ahmed AI. A Review on Computer Vision-Based Methods for Human Action Recognition. Journal of Imaging. 2020; 6(6):46. https://doi.org/10.3390/jimaging6060046

[6] L. Mo, F. Li, Y. Zhu and A. Huang, "Human physical activity recognition based on computer vision with deep learning model," 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Taipei, Taiwan, 2016, pp. 1-6, doi: 10.1109/I2MTC.2016.7520541.

[7] H. Meng, N. Pears and C. Bailey, "A Human Action Recognition System for Embedded Computer Vision Application," 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 2007, pp. 1-6, doi: 10.1109/CVPR.2007.383420.

[8] Barshan,Billur and Altun,Kerem. (2013). Daily and Sports Activities. UCI Machine Learning Repository. https://doi.org/10.24432/C5C59F.

[9] Agus Eko Minarno, Wahyu Andhyka Kusuma, Rizalwan Ardi Ramandita; Classification of activity on the human activity recognition dataset using logistic regression. AIP Conf. Proc. 25 July 2022; 2453 (1): 030003. https://doi.org/10.1063/5.0094789

[10] L. Fan, Z. Wang and H. Wang, "Human Activity Recognition Model Based on Decision Tree," 2013 International Conference on Advanced Cloud and Big Data, Nanjing, China, 2013, pp. 64-68, doi: 10.1109/CBD.2013.19.

[11] Maswadi, K., Ghani, N.A., Hamid, S. et al. Human activity classification using Decision Tree and Naïve Bayes classifiers. Multimed Tools Appl ***80****, 21709–21726 (2021). https://doi.org/10.1007/s11042-020-10447-x*

[12] K. G. Manosha Chathuramali and R. Rodrigo, "Faster human activity recognition with SVM," International Conference on Advances in ICT for Emerging Regions (ICTer2012), Colombo, Sri Lanka, 2012, pp. 197-203, doi: 10.1109/ICTer.2012.6421415.

[13] Abdi, H. and Williams, L.J. (2010), Principal component analysis. WIREs Comp Stat, 2: 433-459. https://doi.org/10.1002/wics.101