Enhancing Activity Recognition Through Game Theory: Shapley Values and Stackelberg Games in Machine Learning Models

Elisa Borella
elisa.borella.4@studenti.unipd.it

Umut Berk Cakmakci
umutberk.cakmakci@studenti.unipd.it

Enrico Gottardis
enrico.gottardis@studenti.unipd.it

***Abstract*—This paper presents the findings of an empirical investigation introducing an innovative approach to investigate the decision-making process computed by machine learning algorithms regarding wearable sensor data in human activity recognition and how to enhance it using game theory principles. Traditional classification algorithms usually face major challenges when dealing with high dimensional datasets and an extremely large number of features, which can be highly inefficient and very costly. If we were able to calculate the importance of each feature and decide which ones to use, this would be very effective and beneficial for recognizing human activity. This study utilizes the Shapley value computation technique to determine the importance of features in machine learning models and employs the Stackelberg game framework to decide which ones to use immediately and which ones to use later (optimal sequencing of feature usage). The study presents experimental results and analysis, demonstrating the effectiveness of the proposed algorithm. Overall, using the publicly available datasets, the research provides valuable insights into the decision-making process in human activity recognition and opens up possibilities for further research in the field.**

***Index Terms*—Game Theory, Dynamic Game, Machine Learning, Stackelberg Game, SHAP Values, Shapley Values, Wearable Sensors, Human Activity Recognition (HAR)**

I. Introduction

Human Activity Recognition (HAR) using wearable sensors has gained significant attention due to its potential applications in various domains, including healthcare, sports, and lifestyle monitoring. The ability to accurately recognize and understand human activities based on sensor data has the potential to revolutionize fields such as personalized healthcare [1], activity tracking [2], and smart environments [3]. In the present investigation, a novel approach for analyzing wearable sensor data and determining the importance of different features in HAR using game theory principles is proposed.

One of the key challenges in HAR is understanding the relative importance of different features and metrics used for accurately recognizing human activities. While various machine learning algorithms have been successfully applied to HAR tasks, identifying the most important features that contribute to accurate predictions remains a critical research area. Additionally, incorporating the uncertainties associated with these features could enhance the decision-making process in HAR. To address these challenges, we employ game theory approaches, specifically the Stackelberg game [4] and Shapley values [5], to gain insights into the importance of different metrics.

The Stackelberg game, a well-known game theory framework, is employed to model the decision-making process in HAR. By characterizing the features as players in a Stackelberg game, we can investigate whether a diagnosis should be made solely based on the features immediately available (leader) or if it is necessary to include additional features that require further testing (follower) [6]. This approach introduces a dynamic setup that considers the payoffs involved in the decision-making process, such as the invasiveness and cost of additional testing.

Furthermore, we utilize Shapley values, a popular concept from cooperative game theory, to quantify the importance of each feature in the HAR model's predictions. The Shapley value analysis provides a comprehensive understanding of the contributions made by individual features and helps determine their relative importance [7]. By considering the interaction effects among features, we can avoid solely relying on individual feature importance metrics and gain a complete perspective on the feature contributions for more accurate decision-making.

The objective of this study is to check the importance values of all features using the Shapley method and divide metrics into 2 categories to follow the Stackelberg game: features A are immediately available and features B require further testing. Then test these features respectively and evaluate the outcomes in terms of duration and validation accuracy. Specifically, we aim to answer key research questions such as: (1) How can the Stackelberg game framework be utilized to make informed decisions regarding the recognition process in HAR when certain metrics require further testing? (2) How can Shapley values be computed and employed to identify the most important features for accurate HAR predictions?

To achieve these objectives, we follow a structured methodology that involves dataset selection, preprocessing, feature extraction, model training, Shapley value computation, and algorithm development. We perform experiments using real-world wearable sensor data and evaluate the proposed algorithm's performance and effectiveness in determining feature importance and aiding decision-making in HAR.

The outcomes of this research have the potential to advance the field of HAR by providing valuable insights into the importance of different metrics and facilitating more accurate decision-making processes. Additionally, the proposed approach opens up new avenues for future research, including the exploration of other game theory concepts and their applications in HAR. By leveraging advancements in game theory with wearable sensor data analysis, we can contribute to the development of more robust and accurate HAR systems.

The rest of this paper is organized as follows. Section II reviews the theoretical fundamentals of collaborative games, the idea of Shapley value, and Stackelberg games as well as the Dataset. In Section III, we describe our approach and expand the details of our data, their processing, and the application of the classification procedure. In Section IV, we discuss numerical evaluations, detailing the analysis performed with two approaches. Finally, Section V concludes the paper and discusses future work.

II. State of the Art

This section delves into the fundamental concepts of Shapley value and Stackelberg games, and it surveys the burgeoning body of literature exploring the utilization of these frameworks in the realm of machine learning methodologies for biomedical applications, especially as regards HAR, wearable devices, and physical activity.

A. Cooperative Games

To fully grasp the Shapley value concept, it is imperative to revisit the fundamental principles of cooperative games. Often called *coalitional games*, cooperative games in *characteristic form* [8] are described as a pair (N, v), where N is the finite set of players and v is the real-valued *characteristic function*. The main difference of coalitional games with respect to non-cooperative games is that in the former the finite set of N players, represented with $\{x_1, x_2, ..., x_\mathrm{N}\}$, acts with the goal of forming coalitions (hence the name *coalitional*), whereas in the latter type of games, players act individually and cooperation can only be achieved indirectly, through credible threats and bargaining for dynamic games. An example could be the *Tit-for-tat* strategy [9]. This contrast is also reflected in the object of study: non-cooperative game theory (NCGT) models the actions of the players, predicting the outcome of a certain interaction by maximizing the utility of each agent in a defined procedure [10](everyone plays his or her best response according to the beliefs they have regarding the actions of other players). In cooperative games the focus is stressed on which coalition will form among the set of N players, thus concentrating on defining the joint actions of the groups and the resulting collective payoffs.
Cooperative game theory (CGT) provides a higher-level approach than NCGT, as it fails to analyze the payoffs distribution within each coalition. Numerous cooperative games can be studied through NCGT, provided that sufficient assumptions hold. It would be wrong to extend this reasoning to every cooperative game, as there are instances in which the information available is insufficient to model the game according to NCGT or the resulting model would be too large and therefore impractical to be analyzed.
As mentioned before, in cooperative games the focus is placed on the behavior of a coalition of players: when this encompasses all agents in the game, it is referred to as a *grand coalition* and is characterized by enhanced efficiency. The main objective of a cooperative game is to find a reasonable distribution of the payoffs within the grand coalition, as it is assumed that it will be formed. Strictly linked to the notion of stability, the *core* [11] is the counterpart of the Nash Equilibrium for cooperative games. When considering the stability of a game, the incentive to unilaterally deviate from any proposed distribution of the total payoff has to be taken into account and this is captured by the core.
Having assumed that a grand coalition among N players will form, the *characteristic function* $v : 2^\mathrm{R} \rightarrow \mathbb{R}$ associates a set of payoffs to the set of all the possible coalitions, such that $v(S)$ corresponds to the collective payoff that a set of players $S \subseteq N$ creates by forming a specific coalition. Not only does the characteristic function satisfy $v(\emptyset) = 0$, but it also has multiple properties. Among these, there are:

- **Superadditivity**: The value of a union of disjoint coalitions is greater than or equal to the sum of the coalitions' separate values:
$$v(S \cup T) \geq v(S) + v(T)$$
where $S, T \subseteq \mathbb{N}$ represent two coalitions with a logical conjunction equal to the empty set.
- **Monotonicity**: If a coalition is larger with respect to another, its gain (the value of the characteristic function) will be higher:
$$S \subseteq T \implies v(S) \leq v(T)$$

In order to better understand cooperative games, it is compulsory to investigate also the meaning of *imputation* [12] and to describe the aforementioned *core*. An imputation, described as $x = (x_1, x_2, ..., x_\mathrm{n})$, is a distribution of the total payoff among the players such that each agent receives at least as much as if there was no cooperation (*individual rationality*) and the sum $x_1 + x_2 + ... + x_\mathrm{n}$ is equal to $v(N)$. The core instead is the set of imputations x such that $x(S) = \sum_{i \in S} x_i \geq v(S)$, for all $S \subseteq N$, with equality for $S = N$ (*efficiency*).
When investigating a cooperative game, assuming that a non-empty coalition T is formed and a player $j \in T$, it is interesting to check if player j has a right to request a share of the payoff $v(T)$, only for being in the coalition T. This is a sensible point if the payoff of the coalition T increases with the presence of the player j as will be delved into in the following of this work. If this is the case and the presence of player j increases the payoff of the coalition, then j can fairly claim a part of the payoff, which can be determined through the Shapley value. Before delving into the specifics of this approach, it's worth emphasizing the underlying concept: the abstract reasoning presented can be effectively translated into a novel approach for evaluating the significance of a given feature in the context of machine learning-based classification tasks. This methodology entails conceptualizing the input features of a classifier as players engaged in a coalitional game, where the payoff of each coalition reflects its collective influence on the classifier's

decision-making process. The ultimate objective is to establish a connection between this coalitional game framework and an individual quantification of feature importance, utilizing the Shapley value as the chosen measure.

B. The Shapley Value

The Shapley value is a concept used in cooperative game theory. It is a method used for determining a fair distribution of payoffs among the players in a coalitional game [13]. An alternative characterization of the Shapley value is as an instrument that assesses the relative importance of individual decision elements within a diverse set of criteria.
Formally, the Shapley value of the jth player in a cooperative game (N, v) is denoted as $\phi_j(v)$. We will interpret this value as the contribution of the jth feature to the overall performance of the classifier, where $v(S)$, also defined as the *worth* of the coalition S, represents the payoff achieved by the classifier, which can be interpreted as a measure of the goodness of the assigned class.
Following the reasoning above, the value of $\phi_j(v)$ can be seen as a fair distribution of v among the members of the coalition. As regards properties, the Shapley value satisfies various ones, including, among others, efficiency, symmetry, linearity and null player. Before describing these properties, it is vital to recall the idea of the Shapley value: it is a fair distribution of the payoff of the coalition S among its members, where each one of them will get a reward $v(S)$ because of the existence of the coalition and for being part of it. This remarks the similarities between coalitional and non-cooperative games: the selfishness that moves players in the latter, because everyone wants to maximize their payoff, can be found in the willingness to join a coalition, as it depends on the share the player gets from $v(S)$ [14].
The Shapley value $\phi_j(v)$, considering a coalitional game (N, v) is:

$$\begin{aligned}\phi_j(v) &= \sum_{S\subseteq N\setminus\{j\}} \frac{|S|!(N-|S|-1)!}{N!}(v(S\cup\{j\}) - v(S)) \\ &= \frac{1}{N}\sum_{S\subseteq N\setminus\{j\}} \binom{N-1}{N-|S|-1}^{-1}(v(S\cup\{j\}) - v(S))\end{aligned}$$

As mentioned before, the Shapley value satisfies many properties, which are shown in the following paragraph.

- **Efficiency**: The sum of the Shapley values of all players equals the value of the grand coalition so that the gain is distributed among the agents.

$$\sum_{j\in N}\phi_j(v) = v(N)$$

- **Symmetry**: The contributions of two players' values i and j should be the same if they contribute equally to all possible coalitions. That is, if $v(S \cup i) = v(S \cup j)$, $\forall S \subseteq (N \setminus \{i, j\})$, then $\phi_i(v) = \phi_j(v)$.
- **Linearity**: Given two coalition games with gain functions v and w, the distributed gain should correspond to the gains derived from v and the gains derived from w.

$$\phi_j(v + w) = \phi_j(v) + \phi_j(w) \quad \forall j \in N$$

  Also, $\forall\alpha \in \mathbb{R}$, the following holds:

$$\phi_j(\alpha v) = \alpha\phi_j(v) \quad \forall j \in N$$

- **Null player**: A player which does not contribute to $v(N)$ should not get a payoff. That is: if $v(S) = v(S \cup j)$, $\forall S \subseteq (N \setminus \{j\})$, then $\phi_j(v) = 0$.

In the context of cost allocation problems with concave costs (or concave utilities), the Shapley value-based cost-sharing rule achieves the minimum price of anarchy, which implies that it is an effective strategic decision rule for combining individual criteria into the optimal joint choice.

C. Stackelberg Games

In a Stackelberg game, also called the *Stackelberg competition*, the players are partitioned into two sets: the *leader* and the *follower*. The leader moves first, and the follower moves sequentially after observing the leader's action. The solution to this type of game is usually found by using the *backward induction*.
In certain Stackelberg competitions, it can be observed that possessing additional information before making a decision (that is the position of the *follower*) can lead to a lower payoff for a player, especially when the availability of this information is publicly known. This is referred to as the *First mover advantage*.
For our application, the features of the Datasets are considered as players and they will be divided into two subsets, one representing the leader and one representing the follower according to a Shapley value threshold.

III. Proposed Method

The proposed methodology involves two key components: utilizing Shapley values to quantify the importance of each feature in a machine learning model and applying the Stackelberg game model to strategically sequence the use of features, optimizing the decision-making process in human activity recognition algorithms. For our application, we chose two similar datasets to work on. These datasets contain real-world data, captured from the sensors that are placed on the subjects. The purpose of choosing two datasets is to show that different values of features can lead to different results even with same type of datasets.

A. Datasets Overview

*1) **HuGaDB Dataset**:* The HuGaDB (Human Gait Database) [3] serves as a comprehensive resource for human gait analysis, encompassing continuous recordings of various activities such as walking, running, taking stairs, sitting down, and more. The dataset comprises segmented and annotated data collected from a body sensor network consisting of six wearable inertial sensors (accelerometer and gyroscope)

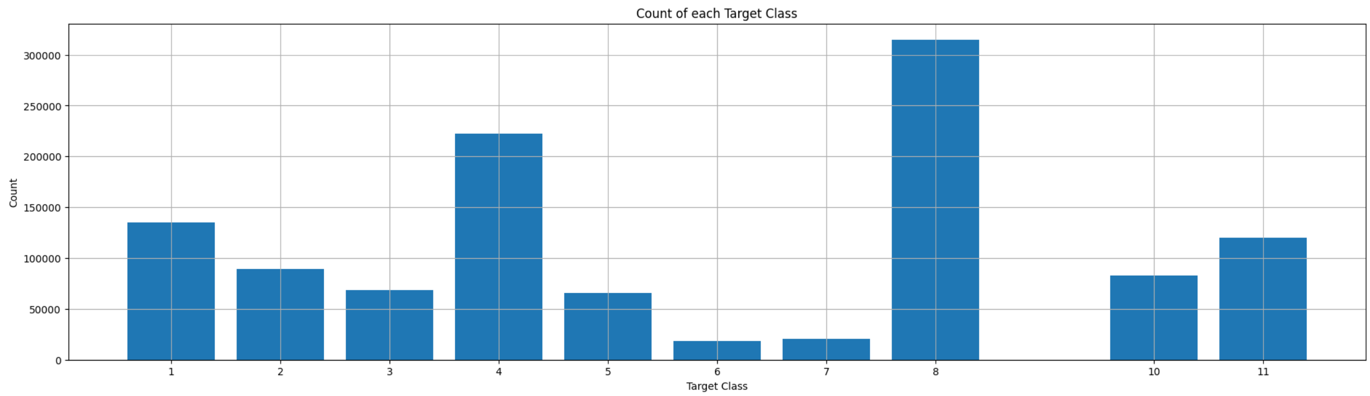

Fig. 1. Distribution of the target features for HuGaDB Dataset

placed on the right and left thighs, shins, and feet. Additionally, two EMG sensors were employed on the quadriceps (front thigh) to measure muscle activity.

Dataset Characteristics:

- Type: Multivariate, Time-Series
- Subject Area: Computer Science
- Tasks: Classification
- Feature Type: Real
- Instances: 2111962
- Number of Features: 36

Dataset Information: Participants performed a combination of activities, and data were recorded continuously. For example, a participant was instructed to perform a sequence starting from a sitting position, involving sitting, standing up, walking, going up the stairs, walking, and sitting down. The data, annotated with activities, provided a long, continuous sequence of segmented data. A custom data collector program was used to record a total of 2,111,962 samples from 18 participants, offering 10 hours of data.

Dataset Summary:

- Number of Activities: 12
- Number of Sensor Devices: 6
- Number of Subjects: 18

Experimental Setup: The dataset records body motions while volunteers perform 12 physical activities. Six wearable inertial sensors (accelerometer and gyroscope) and 2 EMG sensors, placed on the right and left thighs, shins, and feet, capture acceleration, rate of turn, and magnetic field orientation. Three pairs of inertial sensors and one pair of EMG sensors were strategically positioned on the right and left legs. Inertial sensors were located on the rectus femoris muscle, middle of the shinbone, and metatarsal bones, while EMG sensors were placed on the vastus lateralis. The data collection setup specified the gyroscopes' range as -2000 to 2000 deg/sec and the accelerometers' range as -2g to 2g, with 'g' representing gravity acceleration. All modalities are sampled at 50 Hz. The dataset is recorded in an out-of-lab environment, allowing diverse activities, body parts involvement, and execution dynamics.

Activity Set:

- Walking (label 1)
- Running (label 2)
- Going up (label 3)
- Going down (label 4)
- Sitting (label 5)
- Sitting down (label 6)
- Standing up (label 7)
- Standing (label 8)
- Bicycling (label 9, currently not present in the dataset)
- Up by elevator (label 10)
- Down by elevator (label 11)
- Sitting in the car (label 12, currently not present in the dataset)

*2) **Physical Activity Prediction Dataset**:* The Physical Activity Prediction dataset [1], transformed from the PAMAP2 dataset provided on 8/5/2012, serves as a comprehensive resource for techniques related to human behavior analysis through multimodal body sensing. The dataset can be used for activity recognition and intensity estimation while developing and applying algorithms of data processing, segmentation, feature extraction and classification.

Dataset Characteristics:

- Type: Multivariate, Time-Series
- Subject Area: Computer Science
- Tasks: Classification
- Feature Type: Real
- Instances: 2864056
- Number of Features: 32

Dataset Information: This dataset transformed from the original (PAMAP2) data, added a new column called "PeopleId," and merged all the datasets into one comprehensive CSV file. The dataset records data from 13 wide range of everyday, household and sports activities, including walking, cycling, vacuum cleaning, etc. The 8 subjects wear 3 Colibri wireless IMUs and a heart rate monitor, facilitating activity recognition and intensity estimation, offering over 10 hours of data. This dataset is instrumental for developing algorithms in

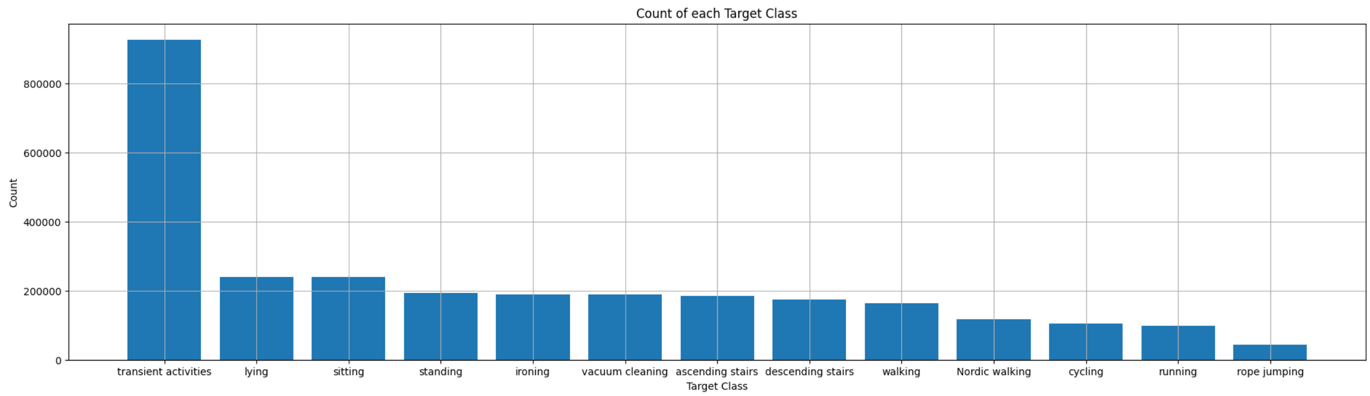

Fig. 2. Distribution of the target features for Physical Activity Dataset

data processing, segmentation, feature extraction, and classification.

Dataset Summary:

- Number of Activities: 13
- Number of Sensor Devices: 4
- Number of Subjects: 8

Experimental Setup: The dataset records body motion and vital signs while subjects follow a protocol comprising 13 activities. 3 Colibri wireless inertial measurement units (IMU) placed over the wrist on the dominant arm, on the chest, and over the dominant side's ankle. One additional chest sensor records the heart rate of the subject. IMU sensors record data at 100 Hz while additional chest sensors record heart rate at 9 Hz frequency. The dataset is recorded in an out-of-lab environment, allowing diverse activities, body parts involvement, and execution dynamics.

Activity Set:

- Transient Activities
- Walking
- Ironing
- Lying
- Standing
- Nordic Walking
- Sitting
- Vacuum Cleaning
- Cycling
- Ascending Stairs
- Descending Stairs
- Running
- Rope Jumping

B. Data Preprocessing

In order to exploit the information contained in the datasets effectively, we perform various processing operations, including handling the missing values, splitting the dataset, encoding (mapping) categorical variables and feature scaling.

In public datasets, there may be some missing values because the data is collected wirelessly most of the time. With high sampling frequency, there might be some connection losses and at that point, the algorithm writes NaN into the dataset. We need to convert them to 0 or, alternatively, we can drop these rows where any of the data in that row is NaN. Eventually, we have less data than before but that data is more accurate and more effective in the training phase.

Our datasets, usually composed of one large file, contain both all the features and the target values together. We need to split the features from the target value. Additionally, we need to split the data samples into training and test sets. Sometimes, we are required to split it further into training, validation and test sets. By dividing the data into 3 sets, we ensure that the model is tested on unseen data, which was neither used for training nor validation.

In these types of open-public datasets, some features can be categorical variables instead of numerical ones. Machine learning algorithms tend to get numerical data for the training phase. We need to convert these types of data to numerical values, usually floating point numbers. We can do it in many different ways using Python; including mapping, one-hot encoding (for multi-class classification tasks), etc.

The feature scaling phase is a necessary condition to standardize the features by scaling them to have a mean of 0 and a variance of 1. If the algorithm is sensitive to the scale of input features, this step becomes essential. We are also doing a feature scaling process to enhance our datasets' trainability.

In summary, we carefully prepare our datasets by fixing missing values, organizing data, and making sure everything is in a form that is easy for our computer programs to understand. This helps our models work better and learn more effectively. We also need to pay attention to details like the scale of our data, which can impact how well our models perform.

C. Models

In order to classify the data, the model chosen was the XGBoost classifier, with a *learning rate* equivalent to 0.1 and a *number of estimators* of 100. The *maximum depth* of each decision tree was set to 3, to prevent overfitting. This model is often used for classification problems due to its effectiveness, efficiency, and generalizability.

XGBoost predicts the probability of each class for a given data point. It does so by training an ensemble of decision trees, each of which classifies the data point into one of the available classes. The final prediction is determined by the class that receives the highest cumulative probability across all the trees. In order to achieve a correct classification, the dataset is randomly divided into 80% training and 20% validation data. First of all, we trained the XGBoost model with the full set of features, achieving:

- For the HuGaDB Dataset: an accuracy of 80.6% in approximately $19.5s$ of training
- for the Physical Activity Prediction Dataset: an accuracy of 89.8% in approximately $54.5s$ of training

Then, we extracted the Shapley values of each feature and created a subset of all the features composed only of those with the highest Shapley values. By exploiting only this subset of features, we trained a new XGBoost classifier.

We collected and evaluated the predictions done by this new model, and lastly, we trained a third model with the remaining features plus the predictions of the second model created. The main question we investigated was whether we should reach a diagnosis from the model with just the subset of features with the highest Shapley values or we ought to exploit also the remaining feature, thus combining the prediction of the first set of features with the remaining ones.

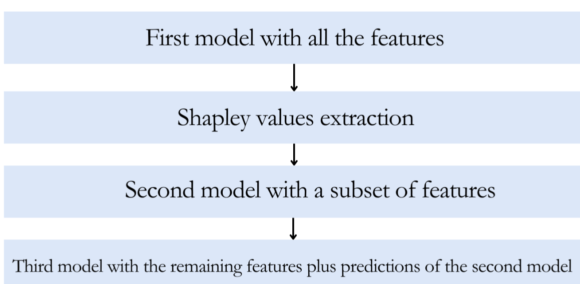

Fig. 3. Pipeline of the proposed approach

We developed an algorithm that automatically assigns each feature to the subsets. Shapley values are calculated separately for each feature and each instance and saved in an array. To do the assignment, we first take the averages of Shapley values feature-based. Then, we have the average Shapley values of each feature. We put these values into a decision mechanism that checks each value according to a threshold. For our research, we set the threshold value to 0 for simplicity. What we are doing here is simply assigning the positive average features to the first subset, which we call the *leader*'s feature set, and the negative average features to the second subset, which is the *follower*'s feature set.

IV. Results and Discussion

The results obtained with the Physical Activity Dataset are shown in Table I. The results obtained with the HuGaDB dataset are shown in Table II. The metric used to evaluate the performance of the model is accuracy, which is defined as follows: $Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$. Specifically TP (True Positives) and TN (True Negatives) indicate the number of samples correctly classified, while FP (False Positives) and FN (False Negatives) indicate the number of misclassified samples. In each Table are reported the results of the first model trained model with the whole set of features, the results of the second model trained only with a subset of features, and the results of the third model trained with the remaining features with the addition of the second model's predictions.

TABLE I
Results Physical Activity Dataset

Model	*Accuracy*
Model with all the features	0.898
Model with the first set of features	0.808
Model with the second set of features + predictions of the first	0.872

TABLE II
Results HuGaDB Dataset

Model	*Accuracy*
Model with all the features	0.806
Model with first set of features	0.551
Model with the second set of features + predictions of the first	0.789

Analyzing the results obtained, it can be seen that for both the model trained with the Physical Activity Dataset and the model trained with the HuGaDB the accuracy is very high when all features are present. After training the models only with features having a high Shapley value the accuracy decreases, noticeably in the model trained with the HuGaDB dataset and of a smaller amount in the model trained with the Physical Activity Dataset. The difference in accuracies so obtained is due to the characteristics of the datasets: the HuGaDB dataset has very low Shapley values on average, thus training the model with exclusively high Shapley values won't be enough to predict the correct target activity. This explains the need to train the model with the first subset of features and then with the rest of the features that have not been used before. On the contrary, the features of the Physical activity dataset are characterized by relatively high Shapley values on average, meaning that even if we exploit a smaller number of features (with respect to the total) we are able to obtain similar results on the classification task as if we had used all of the features and not just a subset of them. After these considerations, we can act according to our needs: if we need precision, so the objective is higher accuracy, then the network needs to be further trained. Instead, if computational time is the main priority and a smaller value of it is needed, then we can give up higher accuracy and stop training the model with the first subset of features.

In both cases, an improvement in accuracy is noticeable when a third model is trained with the remaining features (thus those with lower Shapley values) in addition to the predictions

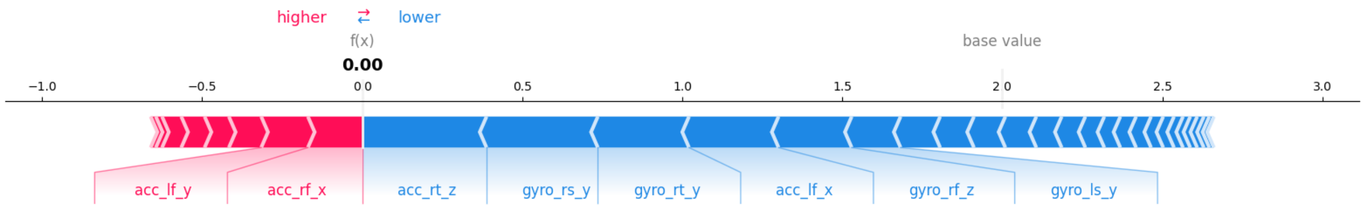

Fig. 4. Feature importance levels derived from Shapley values for an instance from the HuGaDB Dataset

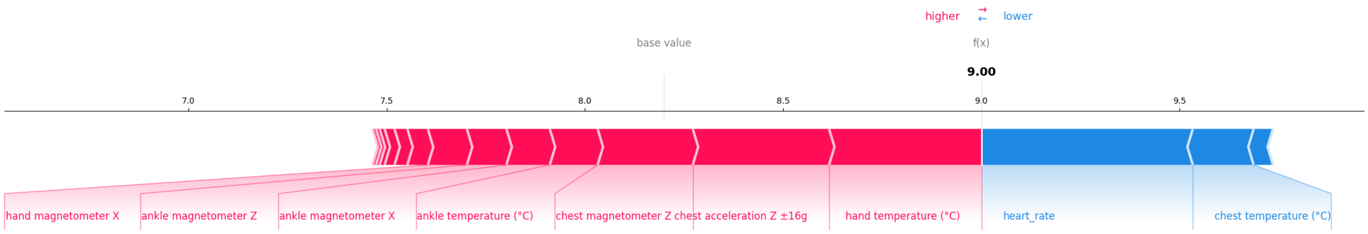

Fig. 5. Feature importance levels derived from Shapley values for an instance from the Physical Activity Dataset

from the second model. In the end, therefore, it is observable that the accuracy of the final model is very high despite not having been trained with all available features.

Furthermore, what was being investigated and is consequently shown by the results is that a delay in the transmission of data is tolerable. This can be derived from the fact that the accuracy of the third model is comparable to the model trained with all the features available from the beginning. This serves as a demonstration that it is not mandatory to wait to collect all the data, instead, it is possible to exploit data already collected to train a model with the features available and consequently feed the remaining data to the model just created.

The outcomes of this study provide valuable insights into the accurate classification of physical activities, even in scenarios where the features of the subject are not all immediately accessible. Specifically, in the context of sensor-based applications, our findings underscore the dispensability of having all sensor data readily available for achieving robust classification accuracy.

The observed result signifies that for the examined application relying on sensors, immediate access to the entirety of sensor data is not imperative for successful activity classification. This conclusion suggests the possibility of tolerating delays in the transmission of sensor data, particularly about the subset of information that is composed of the features with lower Shapley values.

Fundamentally, the study reveals the feasibility of optimizing the use of sensor data by prioritizing features with higher Shapley values, thereby allowing for more flexible data transmission timelines. This nuanced understanding of feature importance not only enhances the efficiency of the classification process but also offers practical implications for real-world applications, where delays in non-critical sensor data can be accommodated without compromising the overall accuracy of physical activity recognition algorithms.

CONCLUSIONS

An approach based on game theory has been proposed to support machine learning algorithms operating in the field of physical activity recognition. It was pointed out from the experimental results obtained that high accuracy of correct classification for physical activity can be achieved by performing accurate feature selection and implementing a game theory framework. In particular, feature selection based on their Shapley values and the implementation of a model inspired by the Stackelberg game proved to be effective in obtaining good classification results for physical activities, despite operating with only a subset of all available features.

Allowing for tolerance in data output delay enables cost savings by opting for more economical instruments, notwithstanding their inability to deliver real-time results comparable to high-end instruments. If the conclusion indicates that not all features are necessary in certain cases to achieve a result, cost reduction in instrumentation can be achieved by selectively purchasing and utilizing only those features that exert the most significant influence on determining the test results.

The integration of game theory into the feature selection process for physical activity recognition offers a promising avenue for both enhancing classification accuracy and optimizing sensor deployment costs. This approach is particularly useful in scenarios where complete datasets are not immediately available and where budget constraints are to be taken into consideration. Future research may explore the scalability of this approach to other domains of activity recognition and its applicability in real-world scenarios.

REFERENCES

[1] Alsareii, S.A.; Awais, M.; Alamri, A.M.; AlAsmari, M.Y.; Irfan, M.; Aslam, N.; Raza, M. Physical Activity Monitoring and Classification Using Machine Learning Techniques. Life 2022, 12, 1103. https://doi.org/10.3390/life12081103

[2] Logacjov, A.; Bach, K.; Kongsvold, A.; Bårdstu, H.B.; Mork, P.J. HARTH: A Human Activity Recognition Dataset for Machine Learning. Sensors 2021, 21, 7853. https://doi.org/10.3390/s21237853

[3] Chereshnev R., Kertész-Farkas A. (2018) HuGaDB: Human Gait Database for Activity Recognition from Wearable Inertial Sensor Networks. In: van der Aalst W. et al. (eds) Analysis of Images, Social Networks and Texts. AIST 2017. Lecture Notes in Computer Science, vol 10716. Springer, Cham.

[4] Tao Li, Suresh P. Sethi. 2017. A review of dynamic Stackelberg game models. Discrete and Continuous Dynamical Systems - B, 22(1): 125-159. doi: 10.3934/dcdsb.2017007 shu

[5] A. Ghorbani and J. Zou, "Data Shapley: Equitable valuation of data for machine learning," in Proc. Int. Conf. Mach. Learning, 2019, pp. 2242–2251.

[6] M. Bruckner and T. Scheffer, "Stackelberg games for adversarial prediction problems," pp. 547-555, 08 2011. https://doi.org/10.1145/2020408.2020495

[7] D. Fryer, I. Strümke and H. Nguyen, "Shapley Values for Feature Selection: The Good, the Bad, and the Axioms," in IEEE Access, vol. 9, pp. 144352-144360, 2021, doi: 10.1109/ACCESS.2021.3119110.

[8] G. Chalkiadakis, E. Elkind, M. Wooldridge, Computational aspects of cooperative game theory, Springer, 2012.

[9] R. Axelrod, The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration, Princeton, New Jersey: Princeton University Press, 1997.

[10] J. Nash, Non-Cooperative Games, Annals of Mathematics, 1951.

[11] B. Peleg , P. Sudhölter, Introduction to the Theory of Cooperative Games, Second Edition, Springer, 2007.

[12] S. Tijs, The imputation set and the core. In: Introduction to Game Theory. Texts and Readings in Mathematics. Hindustan Book Agency, Gurgaon, 2003.

[13] E. Balkanski, Y. Singer†, Mechanisms for Fair Attributions, https://scholar.harvard.edu/files/ericbalkanski/files/mechanisms-for-fair-attribution.pdf

[14] F. Guidolin, L. Badia, and M. Zorzi, "A distributed clustering algorithm for coordinated multipoint in LTE networks," vol. 3, no. 5, 2014.

[15] Z. Wang, D. Wu, J. Chen, A. Ghoneim and M. A. Hossain, "A Triaxial Accelerometer-Based Human Activity Recognition via EEMD-Based Features and Game-Theory-Based Feature Selection," in IEEE Sensors Journal, vol. 16, no. 9, pp. 3198-3207, May1, 2016, doi: 10.1109/JSEN.2016.2519679.

[16] D. Fryer, I. Strümke and H. Nguyen, "Shapley Values for Feature Selection: The Good, the Bad, and the Axioms," in IEEE Access, vol. 9, pp. 144352-144360, 2021, doi: 10.1109/ACCESS.2021.3119110

[17] N. Jothi, S. M. Syed-Mohamed and H. Rajagopal, "Hybrid Feature Selection using Shapley Value and ReliefF for Medical Datasets," 2022 International Conference on Inventive Computation Technologies (ICICT), Nepal, 2022, pp. 351-355, doi: 10.1109/ICICT54344.2022.9850833.

[18] H. Yu and L. Mo, "Research on Feature Selection of Human Physical Activity Recognition for IOT Wearable Devices," 2021 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD), Nanjing, China, 2021, pp. 1-5, doi: 10.1109/ICSMD53520.2021.9670841.

[19] C. V. San Buenaventura and N. M. C. Tiglao, "Basic Human Activity Recognition based on sensor fusion in smartphones," 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), Lisbon, Portugal, 2017, pp. 1182-1185, doi: 10.23919/INM.2017.7987459.