Computer Vision-Based Human Activity Classification and Prediction

Naveed Younus Khattak[*], Ghassan Husnain[†], Waqas Ahmad[‡] , Shahab Ul Islam[§], Ihtisham Ul Haq[**]

Abstract

The detection and classification of human activities in computer vision systems are one of the most confounding tasks having Human-Computer Interaction (HCI) features, tracking, security purpose, and health monitoring nonetheless IoT assists in healthcare. Human activities are indwelling for the objective to recognize recurrently occurring actions. Sitting, standing, sleeping, and eating are a few indications of such human actions. It is tenacious to propose a new advanced model keeping in mind recently published work for the classification and prediction of human activities for better comprehension of the activities associated with humans. The model should perform fast and gives highly accurate results in comparison with existing models. Human activity classification also requires sufficient and easy step-by-step solutions for the day-to-day activities of humans. In this respect, this paper attempts to apply an advanced supervised machine-learning model of human activity classification and prediction. In the classification phase, the report demonstrates a precision rate of 97% and a recall rate of 97% accuracy. The overall accuracy of the classification model is 97%, which is reflected in the F1 score. In comparison with the existing research work score of 90%, this proposed model significantly improved the defect determination accuracy.

Keywords: Human Activity, Machine Learning, XG Boost, Random Forest.

Introduction

Human activities are functions, tasks, or tasks that humans work overtime to achieve a specific purpose. What people do or reason to happen. Human actions such as an act like sitting, standing, sleeping, etc. it is also human activity action that can do human in our regular life. Usually take some action, not a story of murder and other unnatural behaviour. What people do or cause. Human action as action. Event-what happened at a given time and place (Sang, V. N,2015).

Types of Human Activity

[*]Department of Computer Science, IQRA National University, Peshawar 25100, Pakistan, younaskhattak97@gmail.com

[†]Correspondence Author, Department of Computer Science, IQRA National University, Peshawar 25100, Pakistan, ghassan.husnain@gmail.com

[‡]Department of Computer Science, IQRA National University, Peshawar 25100, Pakistan, waqas.ahmad777@gmail.com

[§] Department of Computer Science, IQRA National University, Peshawar 25100, Pakistan, shahab@inu.edu.pk

[**]Department of Mechatronics Engineering, University of Engineering and Technology, Peshawar, 25100, Pakistan, ihtisham1022@gmail.com

The primary activity is human activities, involving the direct extraction of raw materials from the Earth. Examples of major activities are agriculture, fisheries, logging, and livestock rearing. So, primary activities consist of those activities which we are doing for human resource management. The main activity is directly dependent on the environment because these activities refer to the use of land, water, vegetation, and other construction materials and mineral resources of the earth. Thus, it includes hunting and gathering activities of livestock, fisheries, forestry, agriculture, and mining and quarrying (Wang,2018).

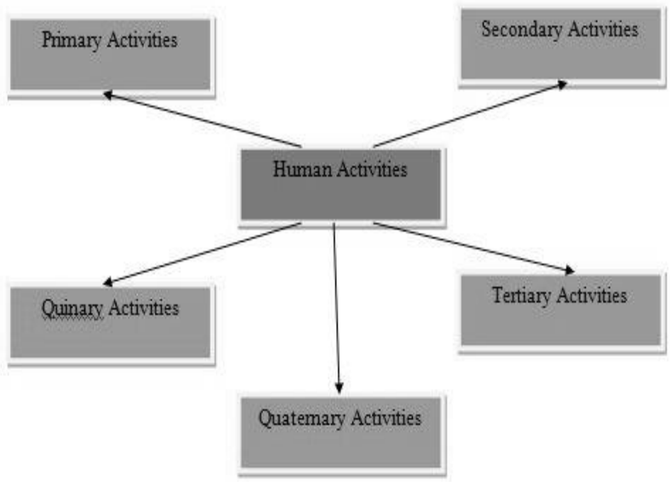

Figure 1: Types of Human Activity (Wang,2018**)**

The assembling and mechanical divisions are known as the subsequent segment, at times alluded to as the creation area, and incorporate all pieces of human activity that convert rough materials into things or items. Assembly and industrial sectors are called auxiliary sections, sometimes called the creative area, which includes all parts of human activities, these activities will become a crude feed item or merchandise (Osayamwen,2019). Comprehensive financial activities are divided into primary, secondary, and tertiary activities. Optional components include auxiliary processing of raw materials, nutrition manufacturing, materials, and assembly industry. This is part of a large part of the mechanical design (Ehatisham-Ul-Haq,2019). The higher administrative department of the tertiary activities has been ordered for fourth and fifth-grade activities. The fourth level of activity is a specialized level 3 in the knowledge area and needs to be categorized separately (Roobini,2019). From shared store administrators to assess experts, programming designers, and analysts, the interest in and utilization of data-based administrations have developed enormously (Roobini,2018). Faculty working in places of business, rudimentary and college study halls, clinics and specialist workplaces, theatres, bookkeeping, and financier firms fall into this class (Chen,2018).

The five exercises help that spotlight the creation, improvement, and translation of new thoughts and existing thoughts; information understanding, and the utilization and assessment of new advancements. The highest level of leaders or decision-makers in five events (Chen,2017).

Machine Learning

Artificial Intelligence (AI) is an accommodation of man-made reasoning that encourages the framework to mechanically learn and propel the training without the requirement for unequivocal programming. The primary intention is to permit the computer to naturally learn without manual intercession or help and alter the activity in like manner. Some of the important real-world machine-learning applications are shown in figure 2.

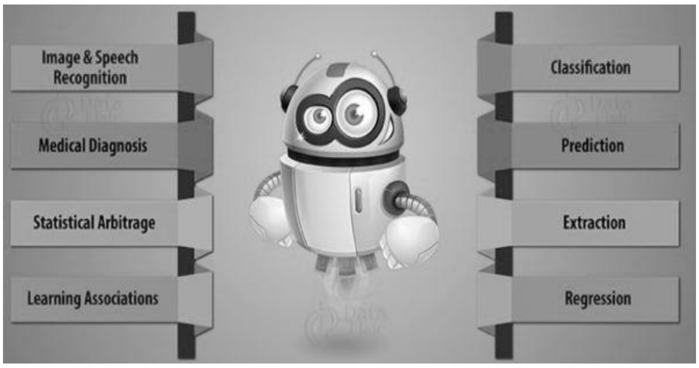

Figure 2: Machine Learning Areas

The figure real-world areas of machine learning where they using widely. Machine learning classifications are used for categorization, and regression using for time series data analysis. Extraction is used for the utilization of data, while prediction defines future trends from existing data. So, these all-use machine learning.

Classification

A supervised approach is used for classification problems and different models such as Support Vector Machine (SVM), K-nearest neighbours (KNN), Logistic regression, decision tree, and random forest. Various approaches were proposed for semantic human activities. Semantic human activities are those activities which a human act in our regular life. But here they take human standing walking, sleeping activities, etc., and proposed a supervised model based on the SVM classification model. They achieved 90% accuracy for the proposed model.

Contribution of Study

The objective of the study is the contemplation of an advanced technique and model for human activities classification which will perform step-by-step solutions for multiple-label classification. Nonetheless also perform a new supervised machine learning model for human activities classification and prediction. Finally, performance evaluation and comparison of the proposed model results with existing work.

Background Study

A deep learning-based technique was put forward to determine human motion with reasonable precision using sensed data. Employing datasets from the UCI datasets repository. Raw data were used for convolutional neural networks with long and short-term memory and recurrent neural networks were applied to analyse the performance of the framework. Activity recognition models can be recherché in healthcare to predict any disease through monitoring human behaviour. Semantic human activities were those activities which a human acts in our regular life.

Input weight is initialized using Gaussian random projection of the basic extreme learning machine. By doing so, more diversity can be generated to improve overall learning performance. Genuine exploratory information has been utilized to assess the exhibition of our proposed strategy (Chen,2018- Chen,2017).

A compression sequence technique based on IRC perceives the enthusiasm of physical activity for local life-building help. To completely compress and capture the IRC, a multi-viewpoint infrared motion detection framework was created, which includes three IRC detection modules, one module on the roof, and two modules on the inverted tripod facing each other. A grouping study of six normal sports activities was performed by merging the grouping program including the Hidden Markov Model and Bolster Vector Machine, which shows the feasibility of the framework. Another ineffective human motion confirmation model that relies on repetitive consideration learning is proposed, in which the operator is prepared to extract data from the anonymous sensor information by adaptively selecting a series of regions (Lin, 2016 – Guan,2016- He,2018).

Non-line-of-sight human crawling wave properties are used to arrange human movement activities. The proposed techniques depend on the timing information of the connection survey team created by the sensor. A human action confirmation framework that relies on motion design on mobile phones is proposed to arrange activities such as falling, walking, running, rising, and sinking stairs (Kim, 2015 - Li,2016- Malhotra,2018).

As the overview points out, most recent research has used deep learning to perceive HAR, although other types of deep learning have achieved accuracy, they have focused on CNNs. Various structural perspectives are analyzed for the ratification of the human motion framework (Mandal,2014- Markopoulos,2019). It further strives to reduce computational costs and achieve significant results in accuracy by highlighting selection strategies (Nazábal,2015- Slim,2019). In addition, it strives to demonstrate the use of repetitive nervous systems to draw highlights from long-term timing information, which can help improve accuracy and reduce dependence on regional information, including extraction and design (Badshah, 2019 - Zhang,2019).

Various delegates are utilized to gauge the multifaceted idea of exercises, discovering unpredictability uncovers 40% to 80% of the change in urban occupations, organizations, coherent zones, and dynamic groupings (Sousa Lima, 2019- Xie,2020- Gibert, 2020). Utilizing recorded patent data, they have demonstrated that the spatial union of front-line mechanical advances has extended since 1850, proposing that a reinforcing cycle be set up between the multifaceted extension of exercises and urbanization. These discoveries propose that the improvement of spatial contrasts might be identified with the extending multifaceted nature of the economy (Mliki,- 2020- Balland,2020).

Proposed Solution

In our proposed solution we design a simple and easily understandable methodology for human activity classification and prediction. For our work, we used Kaggle data sets. Initially, we take data to import it from specific libraries like pandas, NumPy, MatPlotLib, and Seaborn. These libraries are used for the data frame, matrix, and data visualization where we are understanding data features and their relationship with each other.

The data pre-processing techniques for missing values identification and imputation are done. Data cleaning by applying feature scaling for train and test data set. After that select model for our train and test data set for generating confusion matrix, classification reports, and prediction using the XGBoost machine learning model. Finally, we present results graphically and the average accuracy for our model is calculated and evaluate the verified results.

Results and Discussion

Data Set

The dataset comprises of comprehensive details regarding human actions such as standing, sitting, walking, and so on.

Table 1
Human activities data set

Total Instances	Total Feature
3608	563

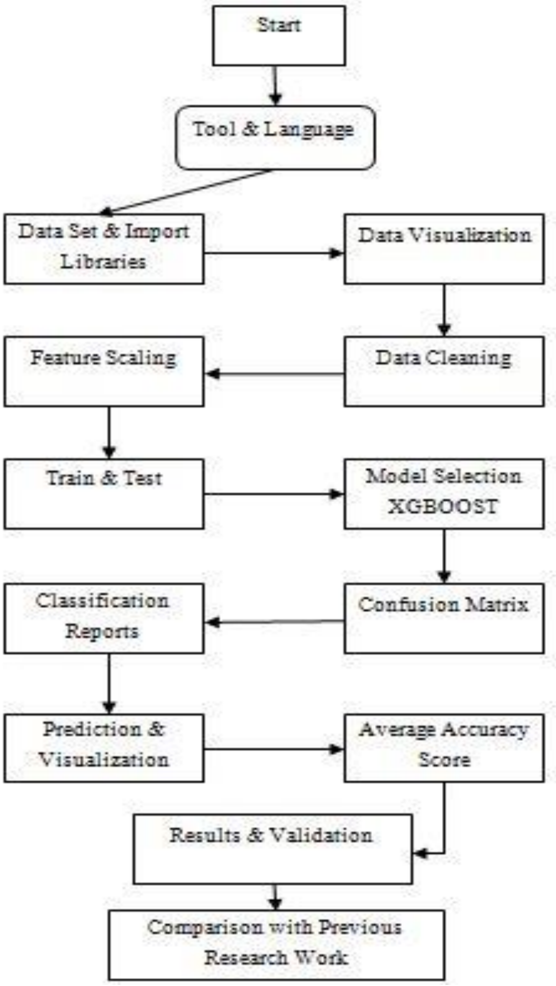

Figure 3: Proposed Solution

Tool & Language

In our proposed research work we select python language for experimental analysis. It is a high-level programming language. It contains limited code which is easier for human understanding.

Important Libraries

The figure below is an illustration of the initial read/write of the proposed work based on the Python libraries datasets.

```
In [1]: import pandas as pd
        import numpy
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        import math
```

```
In [2]: Data = pd.read_csv('train.csv')
        Data.head(5)
```

Out[2]:

	rn	activity	tBodyAcc.mean.X	tBodyAcc.mean.Y	tBodyAcc.mean.Z	tBodyAcc.std.X	tBodyAcc.std.Y	tBodyAcc.std.Z	tBodyAcc.mad.X	tBodyAcc.mad.Y
0	7	STANDING	0.279	-0.0196	-0.1100	-0.997	-0.967	-0.983	-0.997	-0.966
1	11	STANDING	0.277	-0.0127	-0.1030	-0.995	-0.973	-0.985	-0.996	-0.974
2	14	STANDING	0.277	-0.0147	-0.1070	-0.999	-0.991	-0.993	-0.999	-0.991
3	15	STANDING	0.298	0.0271	-0.0617	-0.989	-0.817	-0.902	-0.989	-0.794
4	20	STANDING	0.276	-0.0170	-0.1110	-0.998	-0.991	-0.998	-0.998	-0.989

Figure 4: Import Libraries

Data Visualization

Data visualization is an imperative phase for understanding the data. It is used for deciding data that what analysis is important for data also it is easy for choosing an algorithm or model based on the data set. Figure 5 shows the total types of human activity in a data set.

Out[5]: 3609

Figure 5: Human Activity

Data Cleaning

Data cleaning is also termed data wrangling in machine learning. It is a very long process and is time-consuming. Therefore, during this step, the dataset is analyzed to identify missing values and duplicate the features in a dataset.

As depicted in the Heatmap in Figure 6 above, the data set does not contain any missing values and is deemed clean. Typically, missing values are represented by white cells in the Heatmap. Since the Heatmap is not black or dark, it can be inferred that the data set has been successfully pre-processed for the subsequent analysis.

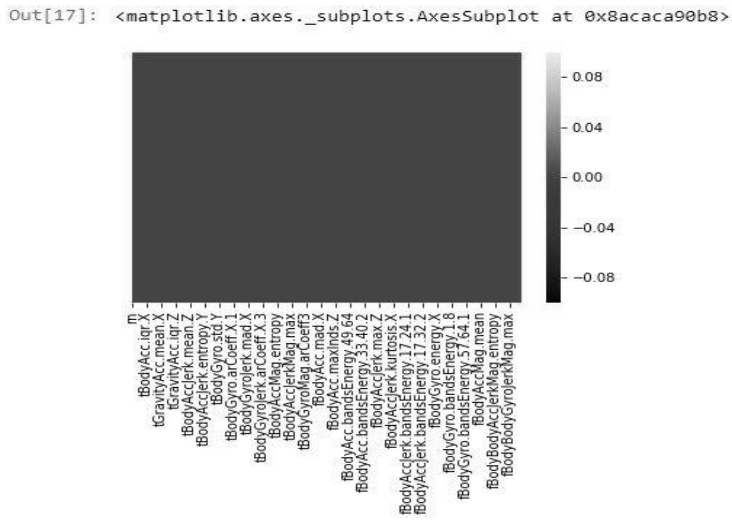

Figure 6: Heatmap for Data Cleaning

Feature Scaling

Machine learning algorithms utilize input data to generate corresponding output results. As a preparatory step, it is necessary to convert all categorical features into numerical labels. This conversion facilitates the subsequent enhancement of machine learning model performance.

Data Normalization

Feature scaling is a technique employed for standardizing the presence of independent features within a dataset across a defined range. During data preparation, feature scaling is performed to tackle differences in the magnitude, value, or unit of height change. Neglecting to perform feature scaling results in machine learning algorithms treating smaller values as inferior, irrespective of the unit of value, and weighing larger values more heavily. Below are some of the superior techniques that can be employed for feature scaling:

1. Standardization: Standardization comes first, followed by normalization. For the standardization approach in machine learning, apply the formula below. Rescaling the eigenvalues using this method results in a distribution with a mean of zero and a variance of one, as illustrated in (1).

$$Xnew = \frac{Xi - Xmean}{\text{Standard Deviation}} \quad (1)$$

2. Normalization: This technique involves rescaling the distribution value of features or observations to be between zero and on.

$$Xnew = \frac{\mathrm{Xi - min(X)}}{\mathrm{max(x) - min(x)}} \quad (2)$$

In our proposed study, we scale features using the common scalar feature scaling method.

Train and Test Class

Following the scaling of the entire data set, the data is divided into two classes: the dependent class, also known as the target class, and the independent class, which is not reliant on other classes. Subsequently, the data is split into training and testing sets for the proposed model. The scikit-learn model selection library can be utilized to split the data for model training and testing. The following are techniques.

X_i = Independent Classes
Y_j = Dependent Class

Model Selection

Here proposed XGBoost classifier for our classification problem. It is also known as the queen of machine learning. XGBoost is the usage of speed and slopes support selection tree. Overall, XGBoost is fast. It is very fast compared to the different implementations of the slope increase. The XGBoost rule organizes or prohibits permutations and recurring datasets, which is a precedent demonstration problem. The below codes are used for XGBoost in this proposed algorithm.

1. import ‘xgboost’ ()
2. from ‘xgboost’ import ‘XGBClassifier’ ()
3. Classifier = XGBClassifier(objective ='Classifier', colsample_bytree’ = 0.3, learning_rate’ = 0.05, max_depth’ = 10, alpha = 1, n_estimators = 1000)
4. Classifier. Fit’ (train_ X, train_ Y)

Confusion Matrix

This approach is employed to summarize the results of machine learning classification. The confusion matrix for our model is depicted in the following illustration. Figure 7 represents the dataset's confusion matrix which presents the total number of actual and predicted labels for classification. The true-positive, true-negative, false-positive, and false-negative labels are derived from a combination of the actual and predicted values. The accuracy of the model's classification and prediction can be evaluated using these parameters:

- True-Negative (TN) represents the sum of all correctly predicted negative instances.
- False-Positive (FP) indicates the number of instances that were predicted as positive but are actually negative.

- False-Negative (FN) indicates the number of instances that were predicted as negative but are actually positive.
- TP stands for True-Positive and refers to the number of precise forecasts that an instance is positive.

Therefore, the labels i.e., True positive, True negative, false positive, and false negative are inside this confusion matrix. We utilize the confusion matrix to compute the accuracy of our classification reports and prediction outcomes, which provides insight into the performance of our model.

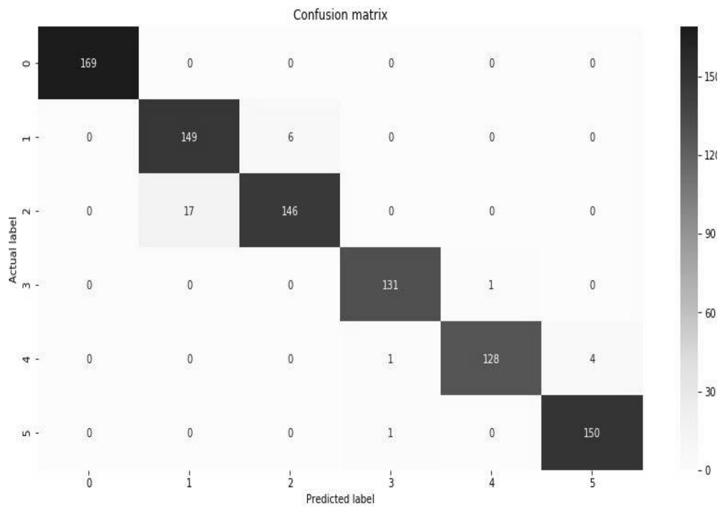

Figure 7: Confusion Matrix

Classification Results

The confusion matrix is utilized to evaluate the performance of our model. The table below provides a summary of the accuracy-based performance metrics of our proposed approach and model. Based on the aforementioned confusion matrix, the following performance metrics are computed: accuracy (AC), precision (PR), recall (RE), and F-measure (F1). Figure 8 depicts the complete classification.

	precision	recall	f1-score	support
LAYING	1.00	1.00	1.00	169
SITTNG	0.90	0.96	0.93	155
STANDNG	0.96	0.90	0.93	163
WALKING	0.98	0.99	0.99	132
WALKING_DOWNSTAIRS	0.99	0.96	0.98	133
WALKING_UPSTAIRS	0.97	0.99	0.98	151
micro avg	0.97	0.97	0.97	903
macro avg	0.97	0.97	0.97	903
weighted avg	0.97	0.97	0.97	903

Figure 8: Classification Report

These reports are generated from the confusion matrix by the following equation of precision, recall, and F1 score.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

$$\text{F1 Score} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{5}$$

Table 2
Performance Measure

AC (%)	PR (%)	RE (%)	F1 (%)
97	97	97	97

Following the report, recall (RE) and precision (PR) for classification are both 97% accurate. Our model has a 97% average accuracy (AC), which is also very remarkable. The F1 97% score is indicated by the average accuracy.

Prediction and Visualization

Prediction in data mining is to identify data points purely based on describing another relevant data value. Usually, regression analysis is used for prediction. Figure 9 shows our prediction of a new data set and visualization of it. Figure 11 shows our prediction based on the proposed data and shows all human activity predictions. As compared to our data the prediction result is 97% accurate. This is our new data.

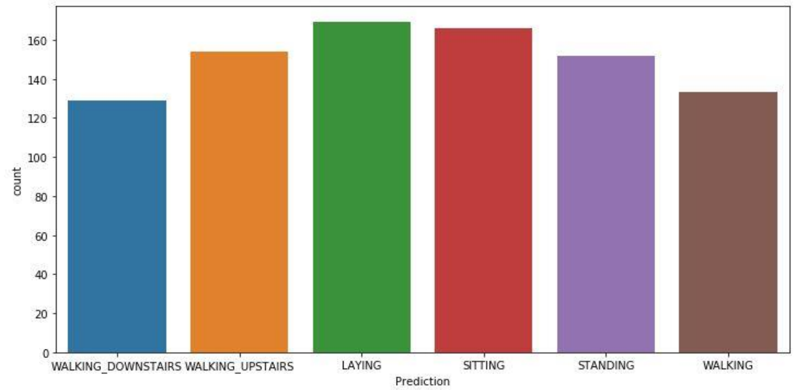

Figure 9: Prediction and Visualization

Results Validation

We used the Scik-learn package validation to identify the results' accuracy for results validation and verification, training, and testing. Figure 10 shows the proposed work through the Scikit-learn validation library.

```
Train set score: 1.00
Test set score: 0.97
```

Figure 10: Results Validation

Conclusions

We proposed a methodology for human activity classification and prediction model design. Valid data sets from Kaggle libraries are used for data frames, matrices, and data visualization. data set in Python Jupyter Notebook. Data visualization for understanding data and data cleaning for missing values are performed and data are trained and tested. The proposed classification model uses XGBoost algorithm and the machine learning model generates a confusion matrix for model performance identification.

***Table* 3**
Results Conclusion

	Comparison		
Research Work	**Data Set**	**Algorithm**	**Average Accuracy Score**
Mingqi Lv (2017)	Human Acceleration	SVM	90%
Our Proposed Research work	Human Activity	XGBoost	97%

The precision (PR) of our classification report is 97%, and the recall (RE) is 97%, resulting in an average accuracy (AC) of 97%, which is exceptionally high. The F1 score is also 97%. Compared to existing research [7], which had an accuracy of 90%, our proposed approach has significantly improved the defect determination accuracy. For practical applications, it would be beneficial to explore different deep learning techniques that are faster, more accurate, and require less time. Deep learning methods are more effective than current machine learning techniques for recognizing human activities. Additionally, it is critical to work on image data for detecting human activity and classifying it correctly.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

"The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper."

References

Balland, P. A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P., Rigby, D. L., & Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. *Nature human behaviour*, *4*(3), 248-254.

Badshah, M. (2019). Sensor-based human activity recognition using smartphones

Chen, Z., Jiang, C., & Xie, L. (2018). A novel ensemble ELM for human activity recognition using smartphone sensors. *IEEE Transactions on Industrial Informatics*, *15*(5), 2691-2699.

Chen, Z., Zhu, Q., Soh, Y. C., & Zhang, L. (2017). Robust human activity recognition using smartphone sensors via CT-PCA and online SVM. *IEEE transactions on industrial informatics*, *13*(6), 3070-3080.

Ehatisham-Ul-Haq, M., Javed, A., Azam, M. A., Malik, H. M., Irtaza, A., Lee, I. H., & Mahmood, M. T. (2019). Robust human activity recognition using multimodal feature-level fusion. *IEEE Access*, *7*, 60736-60751.

Guan, Q., Yin, X., Guo, X., & Wang, G. (2016). A novel infrared motion sensing system for compressive classification of physical activity. *IEEE Sensors Journal*, *16*(8), 2251-2259.

Ghahramani, Z. (2015). Human activity recognition by combining a small number of classifiers. *IEEE journal of biomedical and health informatics*, *20*(5), 1342-1351.

Gibert, D., Mateu, C., & Planes, J. (2020). The rise of machine learning for detection and classification of malware: Research developments, trends and challenges. *Journal of Network and Computer Applications*, *153*, 102526.

He, J., Zhang, Q., Wang, L., & Pei, L. (2018). Weakly supervised human activity recognition from wearable sensors by recurrent attention learning. *IEEE Sensors Journal*, *19*(6), 2287-2297.

Kim, Y., & Moon, T. (2015). Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks. *IEEE geoscience and remote sensing letters*, *13*(1), 8-12

Li, Y., Xue, D., Forrister, E., Lee, G., Garner, B., & Kim, Y. (2016). Human activity classification based on dynamic time warping of an on-body creeping wave signal. *IEEE Transactions on Antennas and Propagation*, *64*(11), 4901-4905.

Lv, M., Chen, L., Chen, T., & Chen, G. (2018). Bi-view semi-supervised learning based semantic human activity recognition using accelerometers. *IEEE Transactions on Mobile Computing*, *17*(9), 1991-2001.

Li, Y., Xue, D., Forrister, E., Lee, G., Garner, B., & Kim, Y. (2016). Human activity classification based on dynamic time warping of an on-body creeping wave signal. *IEEE Transactions on Antennas and Propagation*, *64*(11), 4901-4905.

Lin, Y., Wang, C., Wang, J., & Dou, Z. (2016). A novel dynamic spectrum access framework based on reinforcement learning for cognitive radio sensor networks. *Sensors*, *16*(10), 1675.

Malhotra, A., Schizas, I. D., & Metsis, V. (2018). Correlation analysis-based classification of human activity time series. *IEEE Sensors Journal*, *18*(19), 8085-8095.

Mandal, I., Happy, S. L., Behera, D. P., & Routray, A. (2014, September). A framework for human activity recognition based on accelerometer data. In *2014 5th International Conference-Confluence The Next Generation Information Technology Summit (Confluence)* (pp. 600-603). IEEE.

Markopoulos, P. P., Zlotnikov, S., & Ahmad, F. (2019). Adaptive radar-based human activity recognition with L1-norm linear discriminant analysis. *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, *3*(2), 120-126.

Mliki, H., Bouhlel, F., & Hammami, M. (2020). Human activity recognition from UAV-captured video sequences. *Pattern Recognition*, *100*, 107140.

Nazábal, A., Garcia-Moreno, P., Artes-Rodriguez, A., &

Slim, S. O., Atia, A., Elfattah, M. M., & Mostafa, M. S. M. (2019). Survey on human activity recognition based on acceleration data. *International Journal of Advanced Computer Science and Applications*, *10*(3).

Osayamwen, F., & Tapamo, J. R. (2019). Deep learning class discrimination based on prior probability for human activity recognition. *Ieee Access*, *7*, 14747-14756.

Roobini, M. S., & Naomi, M. J. F. (2019). Smartphone sensor based human activity recognition using deep learning models. *Int. J. Recent Technol. Eng*, *8*(1), 2740-2748.

Sang, V. N. T., Thang, N. D., Van Toi, V., Hoang, N. D., & Khoa, T. Q. D. (2015). Human activity recognition and monitoring using smartphones. In *5th International Conference on Biomedical Engineering in Vietnam* (pp. 481-485). Springer International Publishing.

Sousa Lima, W., Souto, E., El-Khatib, K., Jalali, R., & Gama, J. (2019). Human activity recognition using inertial sensors in a smartphone: An overview. *Sensors*, *19*(14), 3213.

Wang, C., Xu, Y., Liang, H., Huang, W., & Zhang, L. (2018). WOODY: a post-process method for smartphone-based activity recognition. *IEEE Access*, *6*, 49611-49625.

Wang, H. B., Xue, Y., Zhen, X., & Tu, X. (2018). Domain specific learning for sentiment classification and activity recognition. *IEEE Access*, *6*, 53611-53619.

Xie, P., Li, T., Liu, J., Du, S., Yang, X., & Zhang, J. (2020). Urban flow prediction from spatiotemporal data using machine learning: A survey. *Information Fusion*, *59*, 1-12.

Zhang, H. B., Zhang, Y. X., Zhong, B., Lei, Q., Yang, L., Du, J. X., & Chen, D. S. (2019). A comprehensive survey of vision-based human action recognition methods. *Sensors*, *19*(5), 1005.