Identification of Violent Response using Feature Extraction Matrix Algorithm of a Time Series Data (FEM)

Princy Randhawa (princyrandhawa23@gmail.com)
Manipal University Jaipur
Vijay Shanthagiri
, Certisured
Hadeel Fahad Alharbi
University of Ha'il
Akshet Patel
Akshet Patel

Research Article

Keywords: Multivariate Regression Analysis, Physical Violence, Stretch Sensors, Smart jacket, Woman Safety

Posted Date: June 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1714291/v2 sensors

MDPI

Research Paper

Identification of Violent Response using Feature Extraction Matrix Algorithm of a Time Series Data (FEM)

Princy Randhawa [1], *, Vijay Shanthagiri[2]

[1] Department of Mechatronics Engineering, Manipal University Jaipur, Jaipur, India.; princy.randhawa@jaipur.manipal.edu;
[2] CEO, Analogica India Pvt. Ltd. Bangalore, India; vijay.shanthagiri@gmail.com
* Correspondence: princy.randhawa@jaipur.manipal.edu;

Abstract: There are several women safety devices in the market today. Women who suffer these atrocities are even denied basic human rights, as set out in the Criminal Code. Women who are not as fit (physically) as men need to be protected from the evils of society. The introduction of actions and procedures for healthier women is not adequate and needs to be well improved. However, these devices are not fool- proof. The summary of this paper is a partial result of a challenging problem faced during design and construction a fool-proof Smart Jacket for women's safety using fabric sensors. The jacket consists of fabric Sensors, Accelerometer, Gyroscope and Magnetometer, which are strategically placed to record maximum variations in signal for minimum movement in subject's body. The primary challenge is not in design or construction of a jacket, but in accurately classifying violent activity from animated activity. Both violent activity and animated activities have commonality in sensor excitation. There are subtle differences which needs to be extracted to train a Machine learning algorithm to learn these particular patterns. The process undertaken by other studies converges at a successful use of orientation sensors, fabric sensors and an effective system of integrating those sensors, in the form of a device, belt or a wearable gadget which can be worn by the subject and machine learning models have been applied on the data collected from these devices with varying degree of success. However, the major gap in the studies of the past is that none of the study led to a successful classification between normal rapid motion and violent assault. Thorough and granular study of previous research spanning all necessary spectrum of technology such as attack dynamics between an assaulter and the victim was undertaken to select the sensor position, leading to selection of motion parameters that would later become distinguishable features. The primary challenge in achieving distinct classification is the similarity in the data patterns of rapid motion of the body which are common in normal physical activities such as brisk action, playing games, running, or dancing. Using feature engineering and statistical analysis, new features have been created using a novel algorithm which uses the spatial-temporal parameters of the data and creates a sensor activation vector for a predefined length of time-stamped data. The training data has been collected from laboratory experiments in uncontrolled environment and the subjects were of different body types, thus ensuring that there would not be any bias in the data and the resulting model would perform well on the real data. The classical machine learning algorithms such as Decision Trees, KNN, SVM and Naïve Bayes have been used. The best result were obtained using Deep learning; the final result being a distinct classification between rapid movement from normal physical motion and rapid movement due to physical assault.

Keywords: Multivariate Regression Analysis; Physical Violence; Stretch Sensors; Smart jacket; Woman Safety;

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Sensors* **2022**, *22*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Received: date
Accepted: date
Published: date

1. Introduction

It is clearly marked from the facts and figures that incidents against woman have increased over the past decades [1]. Many gadgets have been devised which continuously track the events for the protection of the women such as sound grenade e-alarm, safety rod, pepper spray, smart pendant, safety torch with shock effect etc. [2]. These gadgets won't work when the subject is caught off guard so there is a need of a system/gadget that would be reactive against any physical movement without any human trigger and record maximum motion responses when it is worn by the subject in an uncontrolled environment. Based on earlier studies, it was clear that a significant amount of work had been done on body motion analysis and body centric calculations such as orientations, acceleration, and so on. Identifying the shortcomings and gaps in the previous research sheds light on the areas that were not addressed, one of which is body motion analysis with the goal of distinguishing signals produced by normal rapid motion from signals produced by the subject's resistance during an assault.

In our research work, a jacket was designed to record orientation, acceleration, alignment, stretching of the shoulders and elbows, as well as pressure applied to the body, in order to capture the entire body's motion. Sensors were strategically placed on the body to gather information. [16–18] The challenge was thus seen as a hardware problem with clearly defined goals, such as determining the optimal sensor data acquisition rate so that sufficient data can be acquired, resulting in sufficient variances in the data. The problem had been simplified to a machine learning problem when the proper hardware configuration had been established. Data was studied descriptively and inferentially to determine the solution to the challenge, which was approached as a data science problem. After doing a study of the data, it was discovered that using a single type of person to wear the jacket and generate the signals would result in biased results. As a result, three separate subjects were selected, and the data from each was integrated at the time of model development. The research was separated into two phases: the first phase took place in a controlled environment, while the second phase took place in a real-world, non-controlled environment. "Both standard machine learning models and deep learning models based on CNN were used in both phases," says the researcher. " In phase-one trials, the accuracy of the machine learning models was highly acceptable; nevertheless, the data obtained in a controlled environment was not a realistic depiction of real-world data. " Thus, Phase two was conducted in a far more robust and non-controlled manner than the previous phase. The machine learning models produced for non-controlled data were found to be unsatisfactory, which resulted in feature engineering and feature exploration, during which some of the sensors' data was removed and a novel technique was devised to preprocess the data before it could be used. Feature engineering is centered on creating or discovering new input feature vectors from your existing feature vectors. In general, it can be a part of data cleaning process or can be viewed as an advanced stage of machine learning when all means of extracting input data from the system/phenomenon is exhausted.

Feature engineering is based on transforming the feature vectors without changing the data itself. As one can recall, machine learning models are nothing but algebraic equations which predicts or classifies. Feature engineering techniques enables data scientists to view the data in a different dimension by changing the representational aspects of the data, but without changing the real numbered value and the ratios inherent between feature vectors [1], [2].

In this study, a temporal representation of data was constructed by rearranging the same data using the new approach, which resulted in considerably better machine learning models than the previous algorithm. The examination of the model was carried out in accordance with industry standards

This is the go-to method for data scientists to improve the model performance. Here is the reason why. Key information can be isolated and highlighted thus allowing the algorithm to focus on what are the most important features that affect the model.

2. Literature Review

Previous body motion analysis research has been conducted on a variety of platforms. Sensors and data gathering methods are frequently the difference. Some of them are standalone devices, such as pendants with a gyroscope, accelerometer, and GPS transmitter. Wearable gadgets that can monitor heart rate, temperature, pulse, and other physiological characteristics are another option. The issue statement in the earlier research on fabric sensors and wearable technology in general is different. It's worth noting that different study groups have handled the subject for various reasons. Because the technology is still in its early stages, the researchers are in an experimental mindset, focusing their efforts on seeing whether a viable hardware system can be constructed. The use of machine learning and AI algorithms is a common thread that runs across much of the prior research. Whatever the data and sensor, the numbers provided by the sensors must be processed using proven data centric processing methods, and machine learning models must be constructed to accomplish classification and prediction. When comparison research was required, the current study treated the prior study in the context of each previous study's aims, with no mention of which previous study had a superior outcome or strategy. The reader's attention has been drawn to the originality of the technique used in this research by highlighting the gaps observed in earlier studies. There was no prior research that used a data-centric strategy to create a fully functional smart jacket that could provide sensor inputs that corresponded to the activities. It was also discovered that although the previous study had the depth of research for a specific activity or body motion, the breadth of multisensory fusion from wearable sensors and fabric sensors had not been examined in the context of real-world deployment. For the protection of women's safety, many systems have been created, including FIGHT BACK, STREETSAFE, SENTINEL, ONWATCH, GUARDLY, SURAKSHA, BONITTA, and others. [7]–[13]. When a person is being resisted against an assault, methods or technologies that continually monitor the variables are required.

A smartphone that acts as a portable women's protection system for sending messages or calling the guardian has been proposed as one of the solutions to safeguard women. Pepper spray, a pistol, and a smart necklace are among the other items. [14] [15]. When an attack happens, the victim or a witness is overcome by the perpetrator after a brief struggle, and the victim is rendered powerless, unable to utilize her smartphone or other equipment as a safety mechanism to defend herself. "Peace is not the absence of conflict, but the existence of creative choices for reacting to likely conflict," Dorothy Thompson says.

It's a programme that can identify a violent occurrence without alerting the perpetrator and can automatically deliver warning signals when the cloth is strained. A solution for this kind of application is consistent and automatic categorization of physical activities. To collect such signals, you'll need a sensor or computer that records those factors. It is feasible to create a gadget that will continually capture the signal using wearable technology. Advanced and complicated data gathering, storage, and analysis methods were required due to the high amount, pace, and diversity of data.[16] [17] Wearable physiological metric monitoring systems based on microelectronic integration, wireless communication, and analytics are among these innovative technologies [41].

Advances in sensor technology and wireless networks have made it possible for sensors to be used for human activity detection in terms of capability, cost and reliability for various applications over the past decade. A large amount of data will soon be generated using these miniaturized sensors. In order to retrieve useful information, the need for processing techniques often increases in time as the data supply increases. Consistent and automatic classification of physical activities is needed for a solution for our form of application. “Human activity recognition (HAR) has become one of the most important and influential research topics in different fields”, “such as patient health care over the past few decades”. [43] [27] [38], fall detection of elderly people [44] [45][46][32], security and safety system [47] [48] etc.

The purpose of activity recognition is to recognize the acts performed by a person who has given a collection of self-inspections and the surroundings.[49] There are numerous of system available which automatically classify the human physical activity performed by the subject in many different applications in the field. Figure 1 summarizes the domain areas for human activity identification (HAR). Human physical activity refers to static (sitting, standing, lying) postures. [50] [51], transition activities (sit-to-stand, sit-to lie, stand to-walk) [52], [53], Dynamic motions (Walking, Running, stairs climbing, exercising, housed hold chores) [54] [55].

Figure 1: Applications areas of Human Activity Recognition (HAR)

In previous research, Human Activity Recognition (HAR) uses different types of approaches, such as: 1) Computer vision-based HAR- It uses cameras to record the various activities [56] [53] [57] 2) Environmental sensor-based HAR- In order to detect events, sensors and signals
such as sound sensors, light sensors or RFID tags are used, [58] [59] 3) Wearable sensor-based HAR-These are the sensors that are mounted on the various parts of the body, such as accelerometers, strain, stretch, and then analyze the information to identify the activities [21] [14], [60], [61] and 4) Time geography-based HAR- Using time and location data to classify human activities [62] [53], [63]–[65].

Many activity categorization algorithms based only on accelerometer data have been suggested for Android or other smart-phone platforms [15-18] .Differentiating between walking, running, cycling, and driving is done using accelerometer-based systems (Fang et al., 2012). Accelerometers have also been used to classify exercise activities (Muehlbauer et al., 2011). Various authors have proposed various classification schemes and feature extraction methods to identify different activities from a range of data; and have suggested various algorithms for the recognition of human postures such as decision trees, KNN, hidden Markov model, random forest, support vector machine, and artificial neural network to recognize different body activities [22-26] .Other approaches for classifying data from wearable sensors have also been employed. Model accuracy of over 95% has been attained using these approaches [21] . Other elements, such as the body's BMI, acceleration statistics, and physiological measures of the individual and during a particular activity, are recommended for further development. Singular value decomposition (SVD), multiscale entropy, fuzzy logic, and Naïve Bayes have all been used to automatically identify human body positions [31-36]

Limitations:
Different groups and industries have been researching sensor-based Human Activity Recognition (HAR) and Violent Activity Recognition (VAR), however, multiple questions have been addressed using different technologies. Based on the complexity of tasks they classify in the field of sensor-based activity detection, human activity recognition (HAR) techniques can be checked". [39], [40-42] The primary issues rely on the number of activities, the varieties of actions, the variety of sensors, comfortlessness, opacity & protocols for data collection which lead to a decrease in accuracy.[43], [44-46]
Figure 2 outlined the difficulties or shortcomings behind the detection by sensor systems of normal motion of violent action.

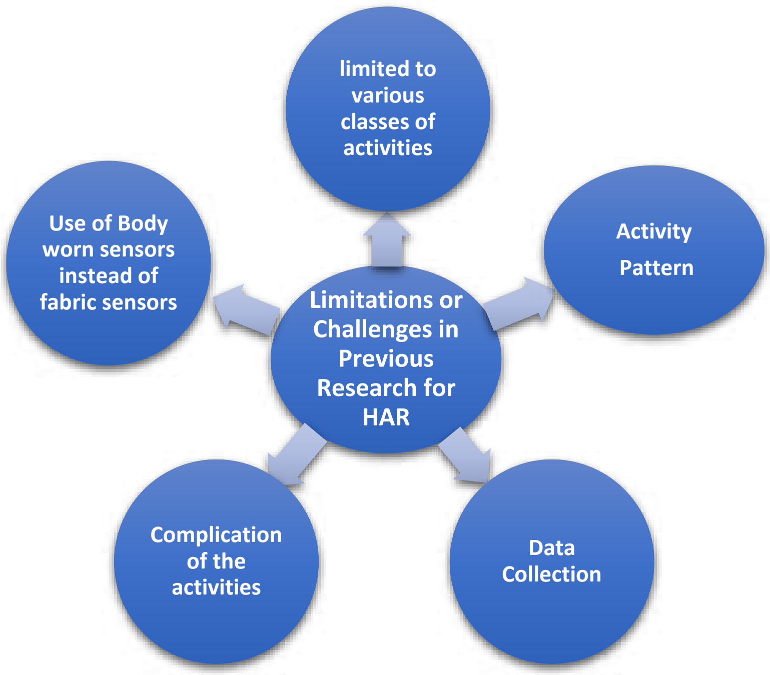

Figure 2: Limitations behind sensor-based recognition of human activity.

Hardware Architecture

The fabric sensors are strategically placed on the jacket in keeping the person 's reaction in mind when under violent attack as shown in Table 1.

Table 1: Sensor Positioning on the wearables

Fabric Sensor	**Sensor Positioning**
Pressure	Right and Left Wrist
Stretch	Right and Left Elbow and Shoulder
3 DOF A/M/G Accelerometer, Magnetometer, Gyroscope)	Wrist
Processing Unit (Adafruit Flora Microcontroller)	Chest

\

The rationale behind the sensor's position is that the first response to the shoulders and elbows immediately moves the subject / victim. The change of location causes the sensor values for pressure and elongation to change, which is further reported by the digitized flora microcontroller. Furthermore, on a computer, the wireless network transmitter sends the signal to the receptor. The jacket is made up of bus lines that bind the microcontroller to the sensors, acting around the jacket as a conductive circuit. Using a high conductivity cloth, the bus lines were made, cut into a fine line and woven into a fabric to prevent contact losses during violent movements [14,38,39,40,41]. The jacket was built using conductive bus lines intertwined only to be able to carry out such a violent motion study easily. The information is processed via conductive bus lines linking all sensors to the data collection flora system. The change in resistance of the elongation and pressure sensors is measured by means of a 10-ohm resistance mounted in parallel with each of the sensors and the alignment of the bus lines around the jacket is precisely arranged so that the Flora microcontroller independently detects the change in resistance of each sensor and there is a bit-enabled transition from the integrated analogue to the sensor[14] . The Figure 3 shows the Hardware architecture of the system.

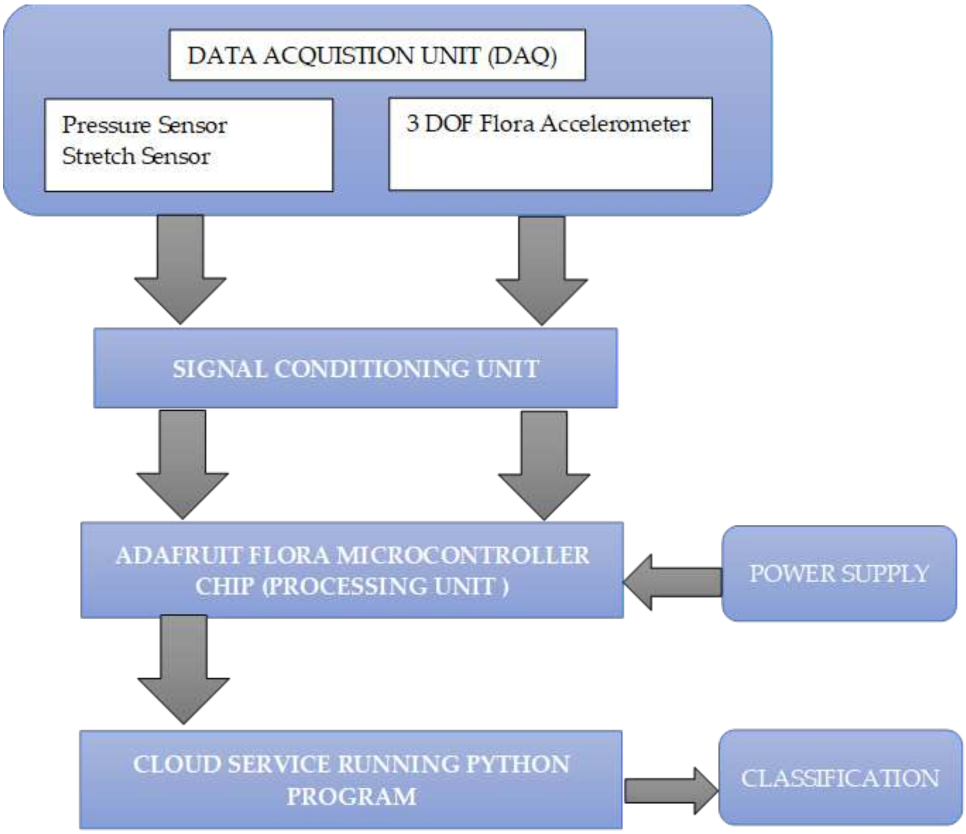

Figure 3: Hardware Architecture of the System

Software Architecture

The sensor data is collected from the fabric pressure and stretch sensor which is mounted on the jacket that is worn by the subject.

a) Data Collection

There is a change in sensor parameters when there is a sudden pressing or twisting of a subject's elbow or hand. The respondent was chosen for the collection of data by age (20-25) and weight (45-55) kg. The data is collected at a frequency of 50 milliseconds. Every 50 milliseconds, the device produces a single full frame represent a unique data set of all sensors, timestamps (epoch time) and expected to operate signs.

Feature vector containing 15 features + Time stamp is received every 50 ms.

20 feature vector containing 15 features + Time stamp is received every Second

A total of 50,000 data points are collected in Controlled environment and controlled envi -ronment.

Population Size = 50,000, Sample Size = 20,000

The subject performs the various activities in its own manner and was not limited to the manner in which the activities were to be carried out, but only in the order A1, A2, A3, A4, A5 and A6 as shown in Table 2.

b) Data Pre-processing

The Data which is collected from the sensors is a raw data which contains noisy information so it needs to be cleaned to get the relevant data before applying various machine learning algorithms.

Table 2: Description of Activities/Events

Activity	Description of the Activity
Stationary (A1)	The action is stationary and there are no data changes observed.
Walking (A2)	The person wears a jacket and walks at a regular speed around it.
Hurried Walking (A3)	The person runs around quickly to strengthen physical mobility.
Jump-Turn-Twist (A4)	The person Jumps and Twists in exercising or playing.
Tango (A5)	The person dances with another person who, as part of the tango style of dancing, often holds the subject's wrist.
Violent Motion (A6)	The person is being aggressively attacked.

Data collection rate for our study is 1 data frame per 50 milliseconds. Over 500 seconds per operation, almost 10,000 data points were obtained from the entire data set population [66]. For each task, each activity window size is about 5-7 seconds and about 15 percent of the frames have been randomly selected as the sample population. For identical operations, there is a certain amount of windows overlapping. For the signals at each interval, overlapping plays an important role in depending on the other interval signals. If we do not consider it, there is a risk of losing a significant data that is at the boundary of the window. This helps us to extract and pick important characteristics from the dataset during model growth. It also assists in the classification of transitional operations. For activity analysis, these frames were considered. For each frame type, a total of 8,000 to 10,000 data frames are reported and 1,500 sample frames are randomly chosen, consisting of one set of all acts [66-67] . Various pre-processing techniques are used to extract the meaningful data from the raw information such Principal component analysis (PCA), Multi-regression analysis (MRA) to ranked the highly correlated features as shown in Figure 4.

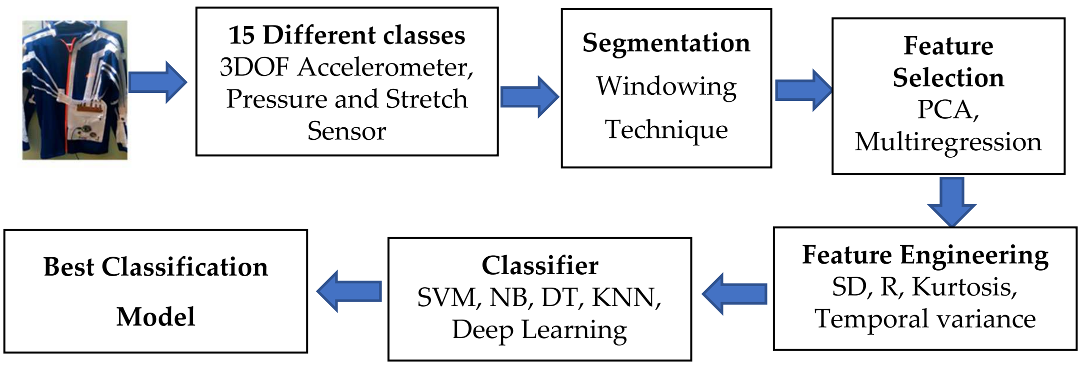

Figure 4: Software Architecture of the System

c) Implementation and Results

After pre-processing, different supervised training techniques have been used in the dataset such as Naïve Bayes (NB), Support Vector Machine (SVM) , Decision Tree (DT), K-Nearest Neighbor (K-NN) in the controlled and uncontrolled environment [66-67]. Figure 5 shows the classification accuracy for various algorithms. The support vector machine provides the best accuracy of 97.6 percent at a speed of 0.85 seconds. This figure summarises that the model used is perfectly capable of predicting most of the sample of violent data. It also summarizes the accuracy which we achieved on the data that is recorded in the controlled environment in which ranges and activity both are defined in the controlled environment. Because the data is controlled and too many outliners and too much variance was observed.

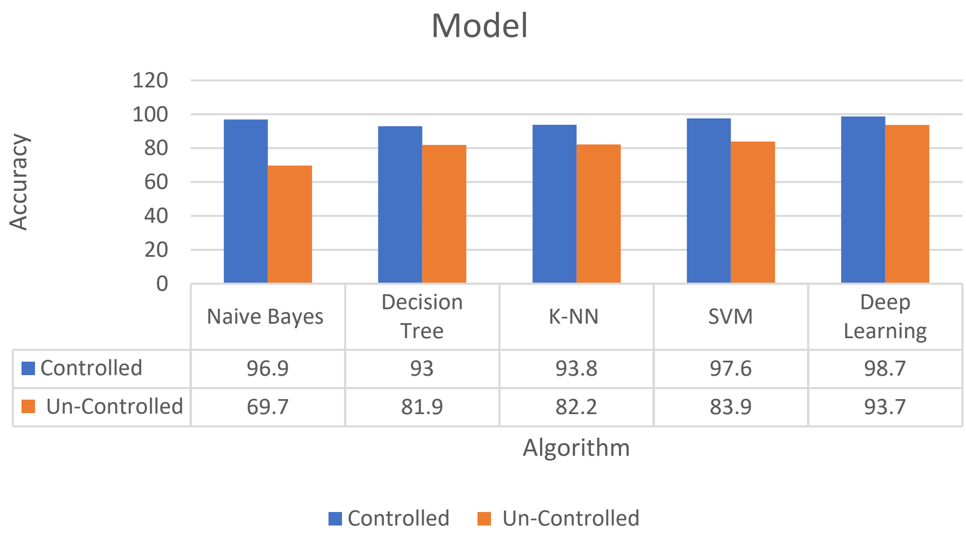

Figure 5: Accuracy model in controlled and uncontrolled environment

The accuracy is 98.7 percent in our dataset using Deep Learning Model. In controlled environment, the naïve Bayes algorithm was 96.9 but now in controlled environment using same algorithm is 69.7 which is remarkable decrease i.e. 27% decrease in accuracy that is a staggering number. Because the bias is less and variance is more and also noise is more so main challenge is not only the no. of subjects is more but also the overlapping of the features i.e. outliers and data pre-processing cannot remove. Similarly, other models show the decrease in accuracy percentage rate.

For the recognition of similar activities as shown in Figure 6 is difficult to recognize as hand gestures are same in both the activities i.e., normal motion and violent activity.

Figure 6: Similar Gestures of Hand in recognition of activity[14]

There are several limitations to the conventional methods to identify violent activity from normal motion from sensor data. Despite the development of traditional algorithms, there are major drawbacks or limitations which has been discussed below:
First, conventional predictions are based on hand-crafted attributes that rely on unique domain knowledge and human experience. "In the case of task-specific environmental setup, these models will perform well but in real-time, for more general scenarios, the performance is low, and it will take more running time to construct an effective activity recognition model".

Secondly, the characteristics learned are shallow in nature in the case of hand-designed methods and generally represent some statistical knowledge, such as mean, standard deviation, variance, amplitude, frequency, etc. Although low-level activities such as jogging, biking, walking, etc. can be easily detected using these shallow features, identifying complex activities that are identical would be difficult or almost impossible. It is necessary to select appropriate features to improve the accuracy in the uncontrolled- environment which leads to the feature engineering because we cannot change the data, data is the reality and the phenomenon and the experiment and what we can do is to look out the data in a different way. So, if you have dependent variable then you can label the variable and the label become the features and if you transform those features is actually go into the feature creation phase. Feature engineering is the process of the creation of those features or attributes that already exist in the dataset. So, using that dataset we are not adding more information to it by creating more features so we have many ways to do it but we have developed a new way of doing it using. temporal context. One of the key original work has been done as a part of study is the development of sensor activation table of matrix. Feature engineering is centered on creating or discovering new input feature vectors from your existing feature vectors. In general, it can be a part of data cleaning process or can be viewed as an advanced stage of machine learning when all means of extracting input data from the system/phenomenon is exhausted. Feature engineering is based on transforming the feature vectors without changing the data itself. As one can recall, machine learning models are nothing but algebraic equations which predicts or classifies. Feature engineering techniques enables data scientists to view the data in a different dimension by changing the representational aspects of the data, but without changing the real numbered value and the ratios inherent between feature vectors

Sensor Activation Table (SAT) and Matrix

Sensor activation table (SAT) is novel algorithm created as a part of the present study. The algorithm takes sensor data and transform the data into a new dimension representing the mean deviation, Range, Kurtosis, and the temporal variance in peaking time for each sensor. The SAT enhances the hidden Information in the physical activity which otherwise is easy to be misconstrued as noise and helps achieve a clear distinction between closely similar activities. For example, dance and violent attack have both similar patterns and it leads to false positives or false negatives. To overcome the situation, there was a need to see the data in a different way.
The principal logic in construction SAT is borrowed from Time domain analysis of sensor data. Data collected by sensors is segregated in batches or windows which has a fixed window size. It is assumed that each window captures the activity in a given instant of time. It is equivalent to saying that each window of size lets day 100 data points is a snapshot of activity in a single moment. If one were to collect a large amount of such moments, and then analyzed the data. There would be negligible loss of information.

In this study, each window is transformed into a single vector which holds a different kind of information derived out of the 100 points windows.
The data that first gets collected directly from the sensor amount to 15 features. Using features selection techniques, 4 of the features were found to add no significant value. Using the 11 remaining features, the data is then divided into windows of 100 data points.

Each window is subjected to the Senor Activation Algorithm which has been created in this study. The SAA will convert each window into a Sensor Activation vector (SAV). Stacking all the SAV's together will form the SAT. A 11 featured window with 100 data points is converted into vector containing 44 elements. This means that each of the features in the window is transformed into 4 fields which represents statistical measures describing the 100 points of that particular feature. A snapshot of the same is represented below in Table 3.

Table. 3: Representation of Feature Extraction using Sensor Activation Table

SR_SD	SR_Ran	SR_ttkVar	SR_Kurt	SL_SD	SL_Ran	SL_ttkVar	SL_Kurt	ER_SD	ER_Ran	ER_ttkVar	ER_Kurt	Activity
8.291198	26	522	1.74895	8.613533	29	202	1.872928	3.093206	8	36	1.674523	0
10.23935	36	783	1.982678	10.19338	36	606	1.988637	3.436979	12	68	2.034814	1
1.319892	6	1323	3.213505	1.329773	6	1344	3.13724	3.781471	14	2294	2.624003	2
3.601289	13	591	2.058669	3.594471	13	591	2.067014	7.483067	28	802	2.3483	3
9.650681	36	943	2.575868	9.52513	36	856	2.655095	8.939996	30	598	1.775615	4
12.6598	32	338	1.992752	8.697664	26	1152	2.349756	5.834928	21	331	3.131164	5

The 4 statistical measures extracted for every feature in the window.
Using concepts of inferential statistics, we have created newer methods like TT-K large variance etc., which is a part of the SAT as shown in Table 11 and Figure 7.

Standard Deviation: The standard deviation is a measure of the amount of variation or dispersion of a set of values. A low standard deviation indicates that the values tend to be close to the mean of the set, while a high standard deviation indicates that the values are spread out over a wider range.

$$\text{Standard Deviation} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \quad (1)$$

σ= population standard deviation, N = the size of the population, x_i =each value from the population, μ=the population mean

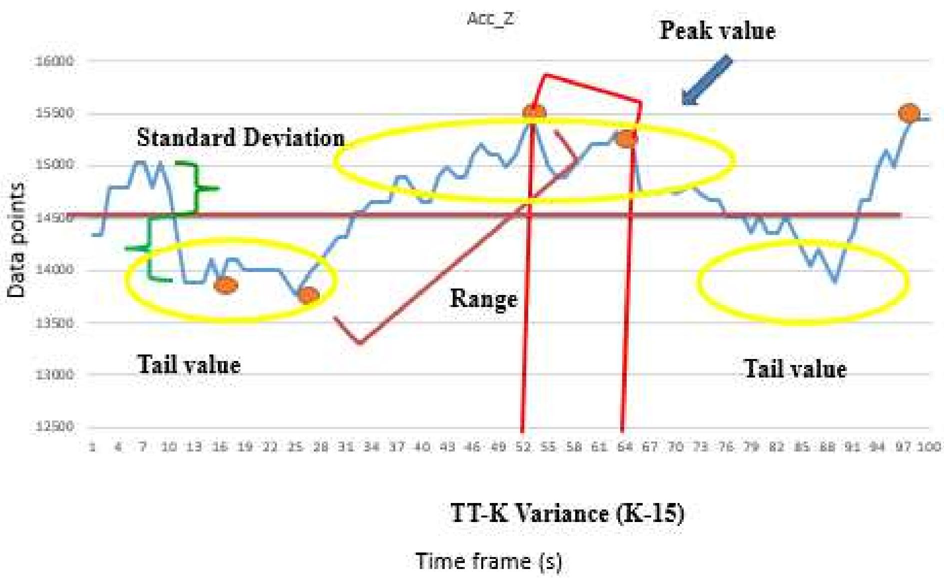

Figure 7: Features Extraction for Sensor Activation Table (SAT)

Range: Difference between the highest value and lowest value, i.e. means how far it is stretching.

$$\text{Range} = High\ value - Low\ value \tag{2}$$

Kurtosis: It is the representation how many % of points lying in the tail and how many points lying in the peak. Kurtosis is a measure of how differently shaped are the tails of a distribution

$$\text{Kurtosis} = \frac{weight\ of\ peak\ value}{weight\ of\ tail\ value} \tag{3}$$

The normal curve is called Mesokurtic curve. If the curve of a distribution is more outlier prone (or heavier-tailed) than a normal or mesokurtic curve, then it is referred to as a Leptokurtic curve". "If a curve is less outlier prone (or lighter-tailed) than a normal curve, it is called as a platykurtic curve". "If the kurtosis is less than zero, then the distribution is light tails and is called a platykurtic distribution". " If the kurtosis is greater than zero, then the distribution has heavier tails and is called a leptokurtic distribution". Kurtosis is measured by moments and is given by the following formula –

$$\text{Kurt} = \frac{\mu_4}{\sigma^4} \tag{4}$$

μ_4= $f$ourth central moment
σ_4 = standard deviation

Time to K large variance (TT-K Variance) - It is the variance of the data index where K small and K large values in sensor data and these large values and small values are mapped to the data index as shown in Figure 1.

The experiment conducted in an uncontrolled environment with a subject wearing the smart jacket yields us data points. After using principal component analysis, the more important features are selected for analysis. Out of 15 sensor values, 11 values are selected.

Table 4: Pseudo Code for Sensor Activation Table (SAT)

SR	SL	ER	EL	AccX	AccY	AccZ	GyX	GyY	GyZ	MgZ

Verify if the number of Columns in the Input Table is 11
{
<*Query the length of the Input Table*>

Length_of_Input_Table/100 = Total_Data_Frames
SAT_Table= Create a new Data Frame consisting of 44 columns
[For each data frame consisting of 11 columns * 100 rows]

(Loop through every column and calculate the)
{
Standard Deviation (),
Range (),
TTKL _Variance (),
Kurtosis ().
Add the 4 values to SAT_Table and repeat the same for remaining 10 columns.
}
<*Verify if the SAT_Table has 44 columns* >

< if the length of SAT_ Table= Total_Data_Frames>

Calculate TT-K variance.
}

TT-K = Time to K variance ultimately calculates the variance of time difference recorded when the data achieves K highest and K lowest points. K is determined based on the frame size. We take K= 10% of Length of Data Frame. For ex. If Length of Data Frame is 100, then K is chosen as 10.

Time stamping of each data point is the key here. To make things simple, every hundred points of data is time-stamped from 1 to 100. And every hundred points is selected as one Data Frame. Sensor activation table is created by calculating the SD, Range, Kurtosis and the TT-K Variance for each sensor data. For a list of X features, SAT will produce 4X features, thus increasing the features without losing information or adding new information which does not belong to the experiment.

A 'K' value of 10 means that the 5 highest and 5 lowest values are chosen from the 100 data points in the specific column. Their corresponding time stamp is recorded and the variance in the timestamp is calculated. This variance is called the TT-K_Var. What this effectively does is it creates a measure of pattern in the data with respect to time it takes for the data points to reach a certain highest or lowest point. For example, if a certain activity is robotic and is performed exactly at the same point and if it produces the same result, leading to a specific sensor value, then the TT-K_Var will be equal to the Range of the Data. If, however, the activity is random then there will be a large variance in the time stamps. TT-K variance is a measure of how disordered the sensor values are.

It can be debated if calculating the entropy of the values would be enough. However, entropy of the data values would not factor in the temporal relationship of the time at which the data reaches a certain level.

Pseudocode for calculating TT-K_var

```
        Query the length of the column whose TT-K_variance is to be calculated.
If (column length == 100)
{
Set K= 10% of Column Length.
Set K_Large = K/2
Set K_Small= K/2
}
            {
With the column as pivot sort the column in ascending order.
Select the first five values from Timestamp column as K_Small
Select the last five values from TimeStamp column as K_Large
Combine K_Large and K_Small into a list
Calculate the variance of the list and return the value as TTK_Variance for that par
-ticular column.
              }
```

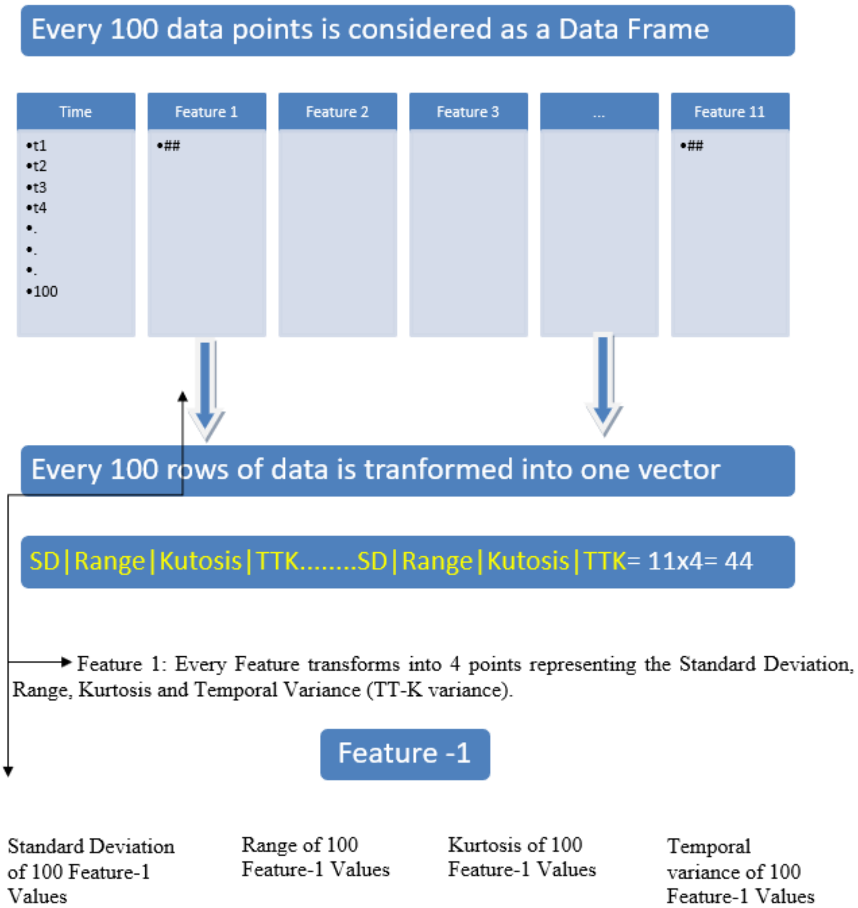

Figure 8: A granular insight logic of sensor Activation Table

Every 100 data points of a single feature are transformed into a vector of 4 numbers which signify the SD, Range, Kurtosis and Temporal Variance of those 100 points as shown in Figure 8 and 9.

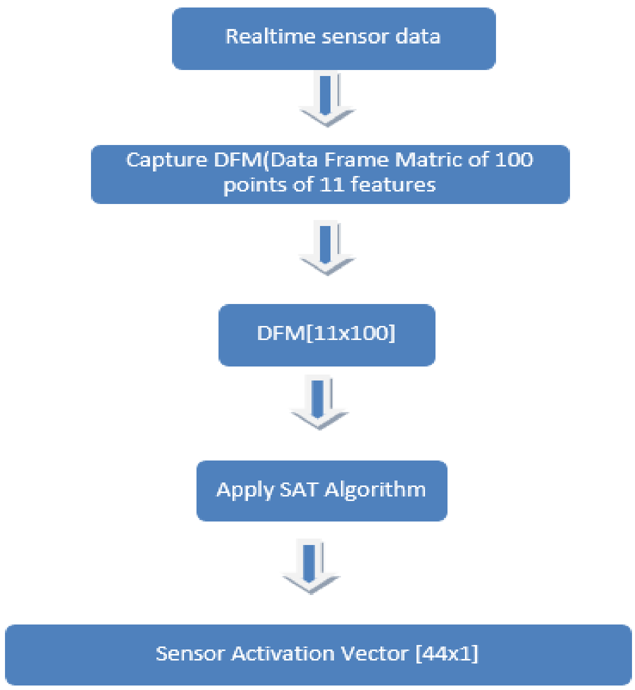

Figure 9: A Flow chart of Sensor Activation Table.

The creation of Sensor Activation matrix is the crucial step in the algorithm. Any machine learning algorithm which will be applied here on will be on the sensor activation matrix. In the present study, an average of 11 lakh data points have been collected for each activity which amounts to 6 lakhs data points. The Figure 10, a flow chart representing the transformation of the columns into Data frames and then into Sensor Activation Matrix is clearly described.

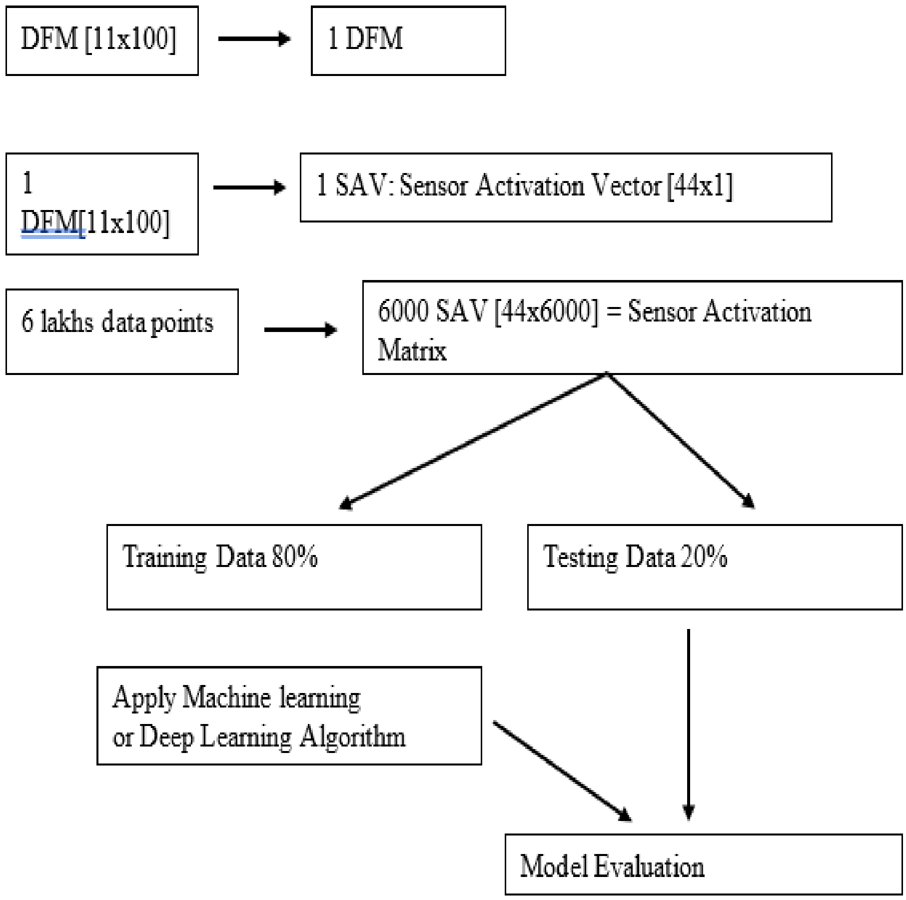

Figure 10: Creation of Sensor Activation Matrix

Model Evaluation

The large amount of dataset (200000-300000) is taken to classify the normal motion and violent attack in a controlled environment. For large amount of dataset, conventional machine learning algorithms does not perform well. In this case for classification used the deep neural network (DNN) or also known as Deep Learning by using framework Tensor Flow. The Model has achieved an accuracy of 99.8% which is good enough for the first run. Further we will use other algorithms to compare the results.

Results and Discussions

Various Algorithms have been applied using raw data and feature extracted data for various activities such as Walking (0), Brisk Walking (1), Sit-Stand (2) (Transitional Activity), Jump-Hop (3), Tango (4), Violent Attack (5). Figure 5.5 shows the pair plot of raw data which depicts that there is a confusion for the recognition of various activities which shows the overlapping of the scatter points. Figure 11 shows after the features are extracted and remove redundant features and attributes shows the correctly classified activities.

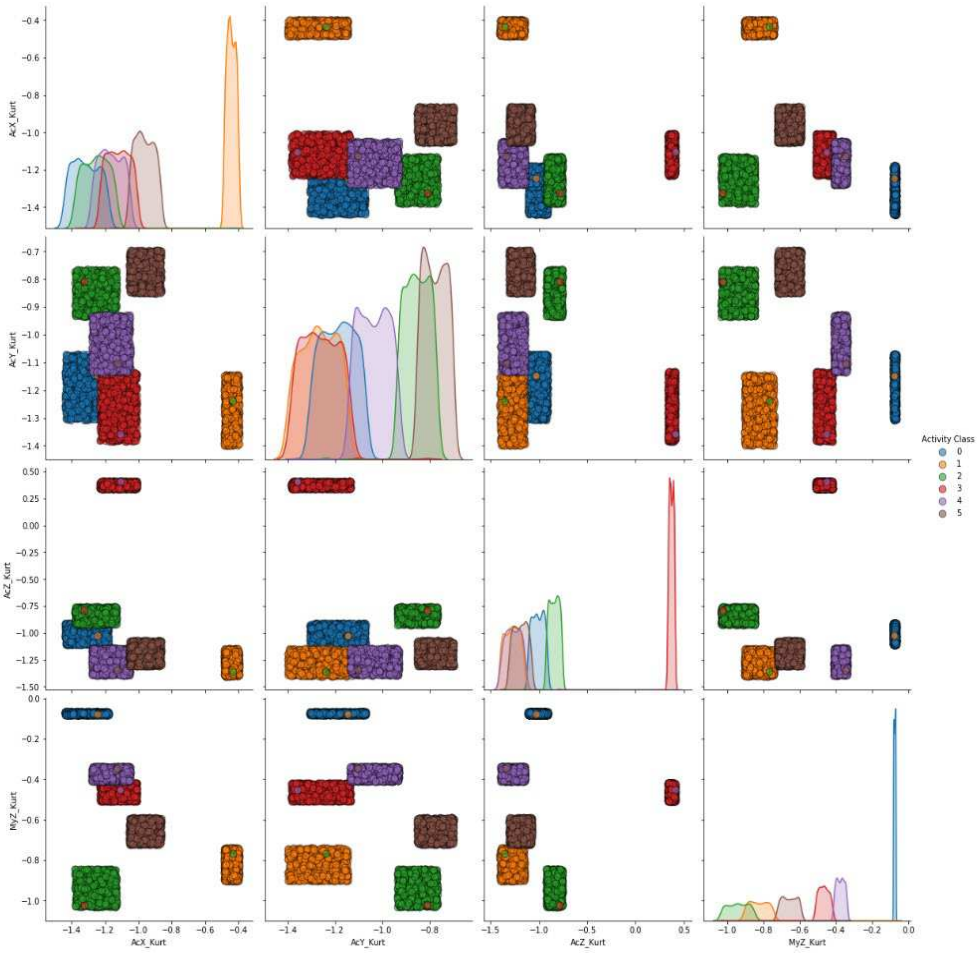

Figure 11: Pair plot correlation coefficient for various activities (Feature Extracted Data)

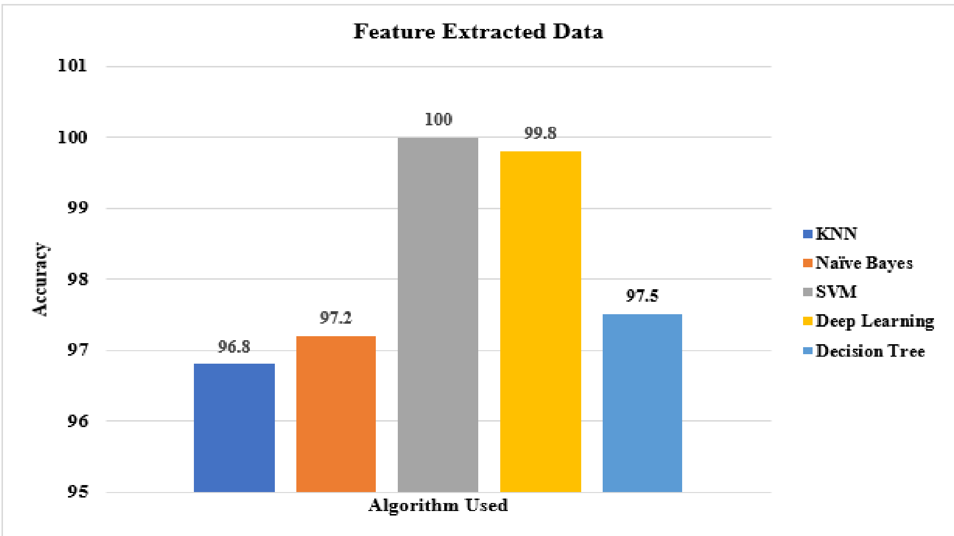

Figure 12: Classification Accuracy using various algorithms in un-controlled environment (Feature Extraction)

It must be cleared here that by raw data what the author means is non-engineered data. In both cases, i.e., raw data and feature engineered data; information has not been removed or added. It is the mere representation of the data which has been changed. The experimentation was conducted for laboratory conditions where there were restrictions on the extent of movements that the subject wearing the jacket would indulge in. A separate phase of experimentation was also conducted where in the previously imposed restrictions on movement were relaxed. Such data would now include extreme peaking and more closely resemble the real-world conditions. The subjects who were a part of the experimentation were of different body types as well and that ensured variation in the subject's movement. This reduced any chance of bias that would otherwise be found, if only one type of subject producing all the data.

The primary advantage that this study had was the fact that the smart jacket was built in-house and hence there was unlimited supply of data from experimentation. The author was not dependent on any external agencies to procure data. This led to individual study of each data and its sensor and how much it contributed to the overall information. Using multivariate analysis and principal component analysis, the features that were irrelevant were removed from the data set. Once the data was processed for missing data and outliers, the data was frozen for analysis. This non engineered data was further subjected to machine learning algorithms. It is evident from the plots in Figure 58 and Figure 67 that the algorithms used over the raw data before feature engineering produced lesser accuracy. The same algorithms when used over the feature-engineered data produced higher accuracy.

In this study, feature engineering was done on the basis of statistical measures and temporal nature of the data. Since the main objective was real-time analysis, it would have been to achieve real time classification on a single set of sensors parameters. The reason being that a single set would contain momentary information and the variation is quiet large between different activities. The feature engineering techniques used in this study considers chunks of data with a pre decided window size. What this means is that, instead of using a momentary set of data, we can now use a window of activity which would be about 100 data points. This is with the assumption that the movement or transition is captured in those 100 data points. These 100 data points are transformed using feature engineering and a new sensor activation matrix is created.

Even the real-time testing is done on the sensor activation vector which is created out of 100 data points. The conversion from raw data into sensor activation vector is computationally feasible owing to faster processing speeds.

Discussion and Conclusion

The entire study as mentioned earlier was data driven. It was clear earlier that model development would require a large amount of data and of great variation. The data was collected with the right strategy of noise cancellation. The following figure summarizes the entire process involved in this study. To counter the issue of lower accuracy in real-world conditions, a new approach was of feature engineering. Instead of just relying on feature reduction, this study viewed data. In terms of time domain responses; this study presents successful implementation of feature transformation. The data is seen both as a time varying and frequency varying phenomena. The real question is how different is sensor excitation for different activities. Adding the temporal aspect to the data was necessary to capture the noise related to random movements in highly animated activities such as dancing and violent attack. Variation brought in by different subjects in real world conditions added noise. However, this noise had valuable information related to the activity which could help clearly classify closely related activities.

Temporal transformation enabled for a different representation of the same data. A window size for recording data was decided and in our case, we choose it as 100 data points per window. Every window of 11 features with 100 points was transformed into a 44 feature vector. About 6000 such windows were transformed into a Sensor Activation Matrix of size 44x4000. This data set was further divided into testing and training data. An accuracy of 98% was observed as shown in the figure 13. This rise in accuracy can be rationally attributed to the data transformation. Further, real time testing of the model, which means, real time classification of the activity would not involve a new step. The real time data will have to be grouped as a window and the converted into sensor activation a matrix of size 44x1. This vector is used for prediction. The activity information and variations are captured in a window of 100 points. The Sensor activation vector is a single line of data embodying the variations in those 100 points.

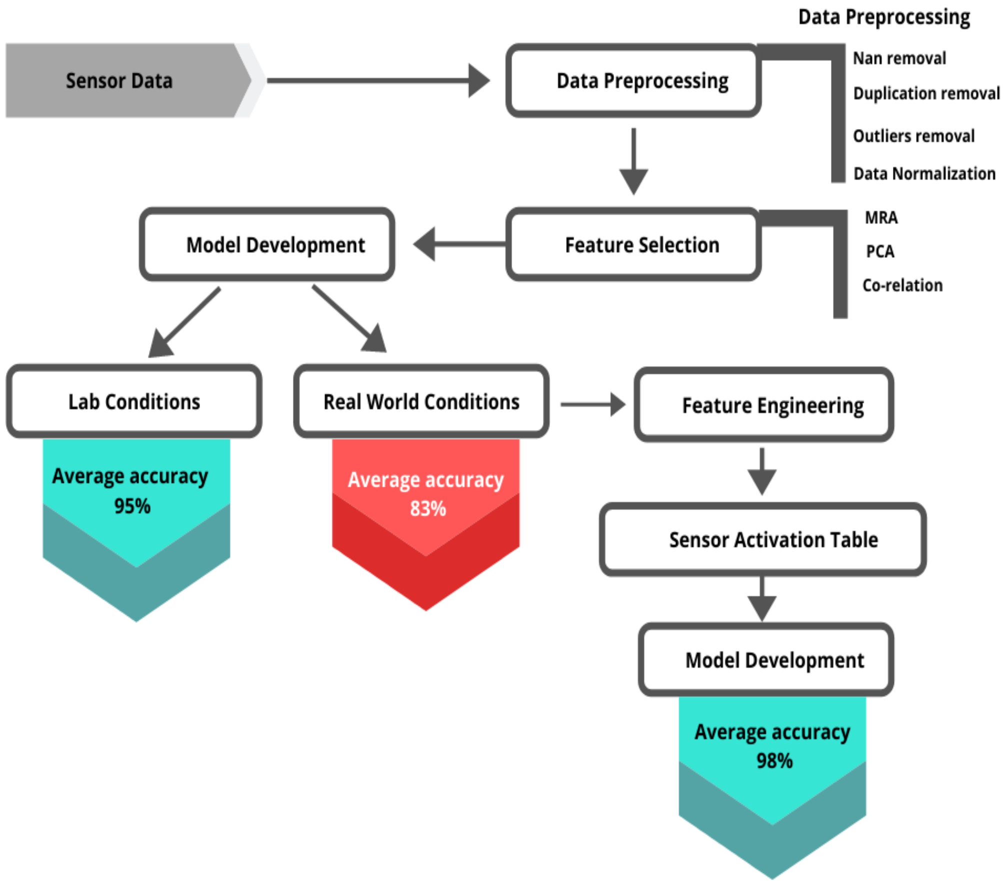

Figure 13: Summaries of the Study

Future Work

A deeper study on subject independent body motion was desired which would involve a separate thread of study which could analyze the effect of the resulting model with changing body type of the subject. As a part of future study, different kinds of subjects could be used to generate data and the model developed so far could be tested and verified. A large disparate body type which difference in age, height, weight and sex could reveal underlying shortcoming which can be explored and solved.

A limited number of body postures and activities were finalized for testing. However, as a future study, a granular approach including the sensor data glitches during activity transition could be undertaken. In future study, there could be a customized algorithm developed for prediction which can model even the glitches that occurs during a particular transition. For example, while transition from walking to brisk walking; there could be a sudden jump in the X axis of the accelerometer. However, if the subject were to transition from walking to dancing, then the changes could occur in many other feature vectors.

Activity analysis could have been better if the models were developed as a pair-wise classification. Meaning that the activities such as walking and brisk walking are easy to classify. A study of the transition between these activities will be studied in future research.

References

[1] Garcia E., Brena, R. F., Carrasco-jimenez, J. C. & Garrido, L, "Long-Term Activity Recognition from Wristwatch Accelerometer Data", 22500–22524 (2014). doi:10.3390/s141222500

[2] A. M. Khan, A. Tufail, A. M. Khattak, and T. H. Laine, "Activity recognition on smartphones via sensor-fusion and KDA-based SVMs," Int. J. Distrib. Sens. Networks, vol. 2014, 2014.

[3] H. Sagha et al., "Benchmarking classification techniques using the opportunity human activity dataset," Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern., pp. 36–40, 2011.

[4] C. Mattmann, F. Clemens, and G. Tröster, "Sensor for measuring strain in textile," Sensors, vol. 8, no. 6, pp. 3719–3732, 2008.

[5] R. Saini and V. Maan, "Human Activity and Gesture Recognition: A Review," 2020 International Conference on Emerging Trends in Communication, Control and Computing (ICONC3), 2020, pp. 1-2, doi: 10.1109/ICONC345789.2020.9117535.

[6] U. Mahbub and M. A. R. Ahad, "Advances in human action, activity and gesture recognition," Pattern Recognit. Lett., vol. 155, pp. 186–190, 2022, doi: https://doi.org/10.1016/j.patrec.2021.11.003.

[7] B. Sumathy, P. D. Shiva, P. Mugundhan, R. Rakesh, and S. S. Prasath, "Virtual friendly device for women security," J. Phys. Conf. Ser., vol. 1362, no. 1, 2019.

[8] R. M. Alisha, P. Vijayalakshmi, A. Jatti, M. Kannan, and S. Sinha, "Design and Development of an IOT based wearable device for the Safety and Security of women and girl children," 2016 IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol. RTEICT 2016 - Proc., pp. 1108–1112, 2017.

[9] R. Pawar, M. Kulabkar, K. Pawar, A. Tambe, and P. Smita Khairnar, "Smart Shield for Women Safety," Int. Res. J. Eng. Technol. e-ISSN, vol. 05, no. 4, pp. 56–2395, 2018.

[10] M. Pramod, C. V. Uday Bhaskar, and K. Shikha, "IOT wearable device for the safety and security of women and girl child," Int. J. Mech. Eng. Technol., vol. 9, no. 1, pp. 83–88, 2018.

[11] M. N. Islam et al., "SAFeBanD: A wearable device for the safety of women in Bangladesh," ACM Int. Conf. Proceeding Ser., no. November, pp. 76–83, 2018.

[12] M. Ahmed, M. Ramzan, H. U. Khan, S. Iqbal, M. A. Khan et al., "Real-time violent action recognition using key frames extraction and deep learning," Computers, Materials & Continua, vol. 69, no.2, pp. 2217–2230, 2021.

[13] B. R. Pragna, P. P. Mahabala, N. Punith, S. Pranav, and S. Ram, "Women Safety Devices and Applications," vol. 7, no. 07, pp. 175–178, 2018.

[14] Randhawa, P.;Shanthagiri, V.; and Kumar, A. (2017). "A Review on Applied Machine learning in Wearable Technology and its Applications.", International Conference on Intelligent Sustainable Systems (ICISS), (Iciss), 347–354..

[15] M. Cornacchia, K. Ozcan, Y. Zheng, and S. Velipasalar, "Using Wearable Sensors," vol. 17, no. 2, pp. 386–403, 2017.

[16] H. Khundaqji, W. Hing, J. Furness, and M. Climstein, "Smart Shirts for Monitoring Physiological Parameters: Scoping Review," JMIR mHealth uHealth, vol. 8, no. 5, p. e18092, 2020.

[17] J. Roski, G. W. Bo-Linn, and T. A. Andrews, "Creating value in health care through big data: Opportunities and policy implications," Health Aff., vol. 33, no. 7, pp. 1115–1122, 2014.

[18] L. M. Castano and A. B. Flatau, "Smart fabric sensors and e-textile technologies: a review," Smart Mater. Struct., vol. 23, no. 5, p. 53001, 2014.

[19] S. Waqar, S. John, and E. Textiles, "Piezoelectric energy harvesting from in- telligent textiles Comfort and durability in high-perfor- mance clothing," 2019. .

[20] A. Sadat, M. Sayem, S. H. Teay, H. Shahariar, P. L. Fink, and A. Albarbar, "Review on Smart Electro-Clothing Systems (SeCSs)," pp. 1–23, 2020.

[21] B. Michael and M. Howard, "Activity recognition with wearable sensors on loose clothing," pp. 1–13, 2017.

[22] D. J. Coming Lopez and C. -C. Lien, "Real-Time Human Violent Activity Recognition Using Complex Action Decomposition," 2020 International Computer Symposium (ICS), 2020, pp. 360-364, doi: 10.1109/ICS51289.2020.00078.
[23] L. M. Castano and A. B. Flatau, “Smart fabric sensors and e-textile technologies: a review,” Smart Mater. Struct., vol. 23, no. 5, p. 053001, 2014.
[24] Y. Wang, T. Hua, B. Zhu, Q. Li, W. Yi, and X. Tao, “Novel fabric pressure sensors: Design, fabrication, and characterization,” Smart Mater. Struct., vol. 20, no. 6, 2011.
[25] T. Dias and A. Ratnayake, “Integration of micro-electronics with yarns for smart textiles,” in Electronic Textiles: Smart Fabrics and Wearable Technology, 2015.
[26] A.-K. Witte and R. Zarnekow, “Transforming Personal Healthcare through Technology - A Systematic Literature Review of Wearable Sensors for Medical Application,” Proc. 52nd Hawaii Int. Conf. Syst. Sci., vol. 6, pp. 3848–3857, 2019.
[27] A. Mason, S. Wylie, O. Korostynska, L. E. Cordova-Lopez, and A. I. Al-Shamma’a, “Flexible e-textile sensors for real-time health monitoring at microwave frequencies,” Int. J. Smart Sens. Intell. Syst., vol. 7, no. 1, pp. 31–47, 2014.
[28] 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,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012, vol. 7657 LNCS, pp. 216–223.
[29] Y. Wu, S. Qi, F. Hu, S. Ma, and W. Yuchuan, “Recognizing activities of the elderly using wearable sensors: a comparison of ensemble algorithms based on boosting,” Sens. Rev., vol. 39, no. 6, pp. 743–751, Jan. 2019.
[30] N. Pannurat, S. Thiemjarus, E. Nantajeewarawat, and I. Anantavrasilp, “Analysis of optimal sensor positions for activity classification and application on a different data collection scenario,” Sensors (Switzerland), vol. 17, no. 4, 2017.
[31] A. Verma, R. A. Merchant, S. Seetharaman, and H. Yu, “An intelligent technique for posture and fall detection using multiscale entropy analysis and fuzzy logic,” IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON, pp. 2479–2482, 2017.
[32] P. Bet, P. C. Castro, and M. A. Ponti, “Fall detection and fall risk assessment in older person using wearable sensors: A systematic review,” Int. J. Med. Inform., vol. 130, no. August, p. 103946, 2019.
[33] Nassif, A.B., Talib, M.A., Nasir, Q., & Dakalbab, F.M. (2021). Machine Learning for Anomaly Detection: A Systematic Review. IEEE Access, 9, 78658-78700.
[34] Nunes, E.C. (2022). Machine Learning based Anomaly Detection for Smart Shirt: A Systematic Review. ArXiv, abs/2203.03300.
[35] K. K. Htike, O. O. Khalifa, H. A. M. Ramli, and M. A. M. Abushariah, “Human activity recognition for video surveillance using sequences of postures,” 2014 3rd Int. Conf. e-Technologies Networks Dev. ICeND 2014, pp. 79–82, 2014.
[36] M. Babiker, O. O. Khalifa, K. K. Htike, A. Hassan, and M. Zaharadeen, “Automated daily human activity recognition for video surveillance using neural network,” 2017 IEEE Int. Conf. Smart Instrumentation, Meas. Appl. ICSIMA 2017, vol. 2017-Novem, no. November, pp. 1–5, 2018.
[37] W. Lin, Y. Chen, J. Wu, H. Wang, B. Sheng, and H. Li, “A new network-based algorithm for human activity recognition in videos,” IEEE Trans. Circuits Syst. Video Technol., vol. 24, no. 5, pp. 826–840, 2014.
[38] Abid, M., Khabou, A., Ouakrim, Y., Watel, H., Chemcki, S., Mitiche, A., Benazza-Benyahia, A., & Mezghani, N. (2021). Physical Activity Recognition Based on a Parallel Approach for an Ensemble of Machine Learning and Deep Learning Classifiers. Sensors (Basel, Switzerland), 21.
[39] Petz, P., Eibensteiner, F., & Langer, J. (2021). Sensor Shirt as Universal Platform for Real-Time Monitoring of Posture and Movements for Occupational Health and Ergonomics. Procedia Computer Science, 180, 200-207.

[40] D. R. Seshadri et al., "Wearable Sensors for COVID-19: A Call to Action to Harness Our Digital Infrastructure for Remote Patient Monitoring and Virtual Assessments," Front. Digit. Heal., vol. 2, no. December 2019, pp. 1–11, 2020.
[41] Kan, C., & Lam, Y.L. (2021). Future Trend in Wearable Electronics in the Textile Industry. Applied Sciences, 11, 3914.
[42] L. Berglin, "Smart Textiles and Wearable Technology," BalticFashion, pp. 1–33, 2013.
[43] H. Lee, J. Cho, and J. Kim, "Printable skin adhesive stretch sensor for measuring multi-axis human joint angles," Proc. - IEEE Int. Conf. Robot. Autom., vol. 2016-June, pp. 4975–4980, 2016.
[44] J. Chen, K. Kwong, D. Chang, J. Luk, and R. Bajcsy, "Wearable sensors for reliable fall detection," Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc., vol. 7 VOLS, pp. 3551–3554, 2005.
[45] G. Chen, C. Huang, C. Chiang, and C. Hsieh, "A Reliable Fall Detection System Based on Wearable," Bioinformatics, pp. 267–270, 2010.
[46] D. Ajerla, S. Mahfuz, and F. Zulkernine, "A real-time patient monitoring framework for fall detection," Wirel. Commun. Mob. Comput., vol. 2019, 2019.
[47] J. Farringdon, A. J. Moore, N. Tilbury, J. Church, and P. D. Biemond, "Wearable sensor badge and sensor jacket for context awareness," Dig. Pap. Third Int. Symp. Wearable Comput., pp. 107–113, 1999.
[48] M. Jutila, H. Rivas, P. Karhula, and S. Pantsar-Syväniemi, "Implementation of a wearable sensor vest for the safety and well-being of children," Procedia Comput. Sci., vol. 32, pp. 888–893, 2014.
[49] Saguna, A. Zaslavsky, and D. Chakraborty, "Complex activity recognition using context driven activity theory in home environments," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 6869 LNCS, pp. 38–50, 2011.
[50] S. Rosati, G. Balestra, and M. Knaflitz, "Comparison of Different Sets of Features for Human Activity Recognition by Wearable Sensors," 2018.
[51] E. Dagan, E. M. Segura, F. A. Bertran, M. Flores, R. Mitchell, and K. Isbister, "Design framework for social wearables," DIS 2019 - Proc. 2019 ACM Des. Interact. Syst. Conf., no. June, pp. 1001–1015, 2019.
[52] M. Stoppa and A. Chiolerio, "Wearable electronics and smart textiles: A critical review," Sensors (Switzerland), vol. 14, no. 7, pp. 11957–11992, 2014.
[53] D. Roggen and A. Bulling, Signal processing technologies for activity- aware smart textiles, no. April. 2013.
[54] Kamal, S., & Jalal, A. (2016). A hybrid feature extraction approach for human detec -tion, tracking and activity recognition using depth sensors. Arabian Journal for science and engineering, 41(3), 1043-1051.
[55]E. Bermejo Nievas, O. Deniz Suarez, G. Bueno García, and R. Sukthankar, "Violence detection in video using computer vision techniques," in Lecture Notes in Computer Sci ence (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2011, vol. 6855 LNCS, no. PART 2, pp. 332–339.
[56] A. Bulling, "33 A Tutorial on Human Activity Recognition Using Body-Worn Inertial Sensors," dl.acm.org, vol. 46, no. 3, Jan. 2014.
[57] Sargano AB, Angelov P, Habib Z. A Comprehensive Review on Handcrafted and Learning-Based Action Representation Approaches for Human Activity Recognition. Applied Sciences. 2017; 7(1):110. https://doi.org/10.3390/app7010110
[58] R. San-Segundo, J. M. Montero, R. Barra-Chicote, F. Fernández, and J. M. Pardo, "Feature extraction from smartphone inertial signals for human activity segmentation," Signal Processing, vol. 120, pp. 359–372, 2016.
[59] D. Silvera-Tawil, D. Rye, and M. Velonaki, "Artificial skin and tactile sensing for socially interactive robots: A review," Rob. Auton. Syst., vol. 63, no. P3, pp. 230–243, 2015.

[60] O. Atalay, W. R. ichard Kennon, and M. D. awood Husain, "Textile-based weft knitted strain sensors: effect of fabric parameters on sensor properties," Sensors (Basel)., vol. 13, no. 8, pp. 11114–11127, 2013.
[61] A. M. Ngueleu et al., "Validity of instrumented insoles for step counting, posture and activity recognition: A systematic review," Sensors (Switzerland), vol. 19, no. 11, 2019.
[62] S. Corbett-Davies and S. Goel, "The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning," Jul. 2018.
[63] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," ACM Comput. Surv., vol. 46, no. 3, pp. 1–33, 2014.
[64] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, D. Howard, K. Meijer, and R. Crompton, "Activity identification using body-mounted sensors - A review of classification techniques," Physiol. Meas., vol. 30, no. 4, 2009.
[65] I. Poupyrev, N.-W. Gong, S. Fukuhara, M. E. Karagozler, C. Schwesig, and K. E. Robinson, "Project Jacquard: Interactive Digital Textiles at Scale," Proc. 2016 CHI Conf. Hum. Factors Comput. Syst., pp. 4216–4227, 2016.
[66] Randhawa, P., Shanthagiri, V., Kumar, A., & Yadav, V. (2020). Human activity detection using machine learning methods from wearable sensors. Sensor Review.
[67] Randhawa, P., Shanthagiri, V., & Kumar, A. (2020). Violent activity recognition by E-textile sensors based on machine learning methods. Journal of Intelligent & Fuzzy Systems, 39(6), 8115-8123.