Introduction:

Technology’s advancement has blessed mankind with smart world that consists of smart living appliances namely smart home devices, smartphones, wearables and other forms of applications, which has tremendously influenced human lifestyle and is continuing to shape the futuristic lifestyle as well. These technologies has empowered independent lifestyle of an individual, thus significantly reducing dependency on other people. With these smart technologies, the concept of Ambient Assisted Living (AAL) emerged. Ambient Assisted Living presents a system consisting of smart devices, home appliances, wireless networks primarily for healthcare monitoring and smart home living. This concept provides the solution to ensuring a safe and quality life for older citizens through preventing, curing and improving wellness and health conditions of older adults by assisting them in living comfortably in their preferred environment. Another sister concept in this regard is Ambient Intelligence (AML). It presents the ability of a computing system to sense its surrounding and interact with people around.

The concept of Ambient Assisted Living (AAL) and Ambient Intelligence (AML) originates at the first place from the advancement of Human Activity Recognition (HAR) through wireless sensor network and the Internet of Things (IoT). Data records from different sensor readings has paved the way to identify human activities separately and is leading to smart home systems consequently. Most HAR systems are based on camera or computer vision or wearable sensors.

One major feature of activity recognition is change detection via detecting sudden change in statistical metrics (e.g. Mean and Covariance), which represents a change in time series data within an indoor environment. Precise manipulation of the derived metrics using a robust algorithm would decide the class of activity performed within a timeframe. In general, activity recognition is a vital component of context-aware systems, which provides the understanding of the smart home applications to understand user requirement and adapt to the various circumstances. Nevertheless, a real-time indoor HAR system in a real environment is often limited by the constraints of indoor environments and makes it difficult to build a robust and scalable system.

Computer vision based HAR systems are useful for large coverage and pedestrian activity recognition. To eliminate the potential privacy issue related to camera based computer vision system in an indoor environment setup, HAR solutions at recent years are based on wearable sensors or devices including smartphones. Wearable approach is sometimes rendered restrained and potential limitations since the user need to always equip the sensing device while recording data, which doesn’t support seamless activity record process. In addition, the wearable approach requires transition between different positions of the user need to be perceived since the system depends on the target to determine the location of the wearable device with respect to the performed activity. [smart wall ref and medium aal article]

On the other hand, In the indoor environment, intelligent HAR system perceives the state of the physical environment and the interacting resident using sensors, reasons about the recorded data and applies Ambient Intelligence to take actions to achieve specified targets. During recording, embedded sensors in the home collects readings while residents independently perform their usual activities. Sensor-data is collected and stored in a database and later analyzed to generate target information such as patterns, predictions and transitions. The process of discerning relevant activity information from sensor streams is a non-trivial task and introduces many difficulties for traditional machine learning algorithms. These difficulties include spatio-temporal variations in activity patterns, sparse occurrences for some activities, and the prevalence of sensor data that does not fall into predefined activity classes.

“Smart Home in a Box” is an output of the Center of Advanced Studies in Adaptive System (CASAS) project at Washington State University, which is an example of a successful HAR system. The smart home kit is small and lightweight, extendable with minimum effort and can perform the key capabilities precisely. This box has been used in 30 volunteer resident houses to collect dataset and the dataset is published in UCI Machine Learning Repository, *Human Activity Recognition from Continuous Ambient Sensor Data Data*set. The dataset is fairly new, published on 20th September, 2019. This dataset represents ambient data collected in homes with volunteer residents with their usual daily activities at home. Ambient PIR motion sensors, door/temperature sensors, and Light Switch sensors are placed throughout the home of the volunteer. The sensors are placed in locations throughout the home that are related to specific target activity of daily living.

To this end, the present work is motivated to classify five distinct activities (Watch TV, Read, Phone, Cook, and Eat) from the dataset of 12 pre-defined activities including unlabeled activity namely “other activity”, on the basis of the CASAS dataset. The motivation is to precisely classify the activities while reducing the computational requirements through exhaustive data preparation. This originates from the idea to allow human activity recognition with less costs involved in computation so that we can incorporate the concept in the perspective of Bangladesh. The dataset is preprocessed, features with statistically significant values have been selected and finally we have applied three different classifier models to present a comparison output of the accuracy level.

The major contributions of the present paper include:

* Data preprocessing of the large CASAS dataset through Principal Component Analysis and Linear Discriminant Analysis
* Feature Selection based on statistical significance
* Classifier models comparison on the pre-processed dataset

[casas-cpd]

The paper follows the following structure: Section II presents the related works on the research objectives. Section III presents Methodology, where data preprocessing and feature selection approaches are discussed and classifier model approach follows the discussion. Section IV consists of the results from three consecutive steps of the research. In the following Section V, Discussion presents the observation and areas for further improvement.

Related Work:

[activity discovery and activity recognition] The research field of activity recognition is quite large considering the combination of embedded sensors, different environmental setups and algorithms to detect activity points. Hence, there are number of approaches explored in this field.

Naïve Bayes classifiers have produced satisfactory output for offline detection of activities […..]. Decision trees are used to learn logical transition of the activity […] while Gu et al […] utilizes KNN to detect mode sensor values associated with activities which helps in recognition.

Probabilistic graph based Markov models […..], conditional random fields[….], Bayesian network have been used successfully to recognize activities even in complex environments. Studies have found that probabilistic graphs along with neural network approaches […] are significant at mapping pre-segmented sensor sequence to activity labels.

Different types of sensor data are proven to be effective for classifying different types of activities. Ambulatory movements (e.g. Walking, Running, Standing, Sitting, Climbing Stairs and Falling) are classified in […….] using accelerometer placed on the body. Recently smartphones with accelerometer and gyroscope sensors are used as wearable device to recognize gesture and motion patterns […..].

More complex activities that requires more information than body movement, in that case the user’s interaction with key objects in the environment is recorded [….]. Shake sensors or RFID tags are tagged with the object and are selected based on the targeted activities. Environment sensors such as motion detector, light sensor, door contact sensors are used to recognize daily activities in other researches […].

At realistic activity recognition tasks, the recognizing activities are performed with interleaved activities […], embedded errors […], and concurrent activities performed by multiple individuals in the setup […]. Detecting activities in free movement setup, where the residents perform usual daily routines in a smart home environment was the next step of advancement […]. These recorded datasets have required on manual labelling to segment and analyze the data. Recent further advancements of activity recognition has brought automated segmentation […], spontaneous selection of objects to tag and monitor […], and for transfer of pre-learned activities to new environment setup [activity discovery and recognition].

[smart wall] An on-body approach is proposed by Kunze et al. […] that perceives if the target is walking and then apply pre-selected sensor reading pattern to predict the actual target’s position. This approach involves attachment of sensor onto the target and hence, the consequent dataset is small. On-body approach with device localization approach presented by Sztyler et al. […] predicts the target on-body position with F-measure calculation and cross-subject activity recognition.

Dedicated HAR architectures use various methods to perceive the complex concerns from recognizing sequential and concurrent human activities. Two key approaches are followed in HAR: data-driven and knowledge-driven technique. Naïve Bayes (NB) classifiers, Decision Trees, Hidden Markov Models, Bayesian Networks and Support Vector Machine (SVM) classifier are the machine learning techniques and probabilistic approaches in Data-driven method. The algorithms work on inductive reasoning to detect human activities in data-driven approach. Existing works including data-driven technique utilizes supervised approach using manually labeled data for training. The approach is restrained by complex method and additional computational cost.

The unsupervised approaches are often restricted by low performance in comparison with the supervised approach in indoor home environment. In the knowledge-based HAR, activities are modeled with their contextual information in the common ground as new activity record is detected via deductive reasoning. The construction of a common ground to present the set of concepts along with their relationships in a machine-interpretable approach is a restraint of knowledge-based HAR. Data-driven techniques are useful for detecting basic distinctive activities, on the other hand unsupervised approach is suitable for creating probabilistic models with expected accuracy score. [smart wall]

A knowledge-based approach utilizing the inter-frame algorithm convolutional neural network is applied in Chen et al. […], where distinguishing features are collected through cameras and learnt, filters non-target objects and estimate skeleton sequence from RGB images. [smart wall ref]

[Paragraph about casas dataset works]

Dataset:

Methodology:  
  
Performance Evaluation:  
  
Conclusion: