Sozo Lab Task-2

Supri Bin Hj Amir
(supriamir@unhas.ac.id)

2021

<table>
<tr><td>01</td><td>Title</td><td>Capturing Sensor Data from Mobile Phones using Global Sensor Network Middleware</td></tr>
<tr><td></td><td>Summary</td><td>This research combines an open-source sensor data stream processing engine called 'Global Sensor Network (GSN)' with the Android platform to capture sensor data

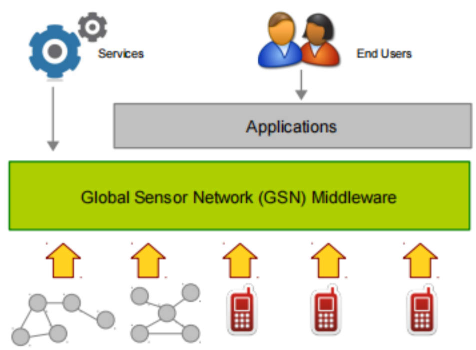
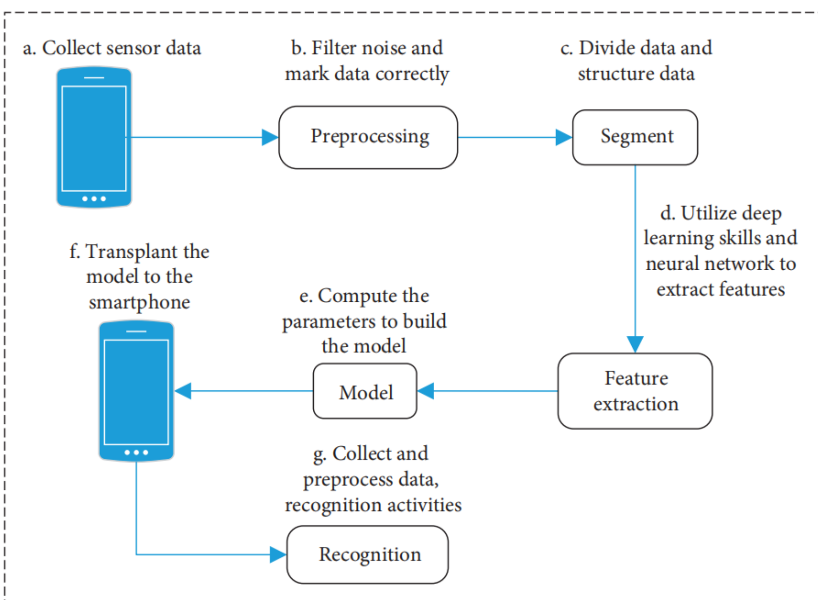


Fig. 2. Collect Data through Mobile Phones

How its work?
GSN gathers raw sensor data from mobile phones and organizes them according to the GSN standard data model and sends data to applications or services when requested.</td></tr>
<tr><td>02</td><td>Title</td><td>Real-time Smartphone Activity Classification using Inertial Sensor-Recognition of Scrolling, Typing, and Watching Video While Sitting or Walking</td></tr>
<tr><td></td><td>Summary</td><td>This research focus on
1. To develop the real-time Smartphone Activity Classification system.

• The system starts from to collect data using Mobile Phones sensor and Android Application within the smartphone, transferred to PC real-time using Bluetooth.
• Develop the ML model in training and testing processing (Labeled Dataset) then use the ML learning model to classify the new coming unlabeled datasets.

2. Investigate the best ML learning Algorithm to classify four different activities (scrolling, typing, watching videos, non-use) under two different conditions (sitting and walking) with an accuracy</td></tr>
</table>

		Evaluated seven different classification algorithms: Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Bootstrap Aggregating (Bagging), Adaptive Boosting (AdaBoosting), Random Forest (RF) and Extremely Randomized Trees (also called ExtraTree, ET). The best Algorithm with an accuracy of 78.6% using the Extremely Randomized Trees algorithm.
03	Title	Recognition of Transportation State by Smartphone Sensors Using Deep Bi-LSTM Neural Network
	Summary	• Proposed novel method that considers these factors comprehensively to enhance transportation state recognition. The deep Bi-LSTM (bidirectional long short-term memory) neural network structure, the crowd-sourcing model, and the TensorFlow deep learning system are used to classify the transportation states. • Data captured by the accelerometer and gyroscope sensors of smartphone is used to test and adjust the deep Bi-LSTM neural network model, making it easy to transfer the model into smartphones and conduct real-time recognition.  Figure 1: Flowchart of transportation recognition.

		• The experimental results show that this study achieves transportation activity classification with an accuracy of up to 92.8%. The model of the deep Bi-LSTM neural network can be used for other timeseries fields such as signal recognition and action analysis.
04	Title	Fall Detection using Recurrent Neural Networks
	Summary	The paper research Fall Detection using Recurrent Neural Network (RNN) model with underlying Long Short-Term Memory (LSTM) blocks. The method is tested on the publicly available SisFall dataset. This research used 3 class labeled dataset. • FALL: This class identifies the time interval when the person is experiencing a state transition that leads to a catastrophic change of state, i.e., a fall. • ALERT: the time interval in which the person is in a dangerous state transition; this state may lead to a fall, or the subject may be able to avoid the fall. • BKG: the default time interval when the person is in control of his/her own state.
05	Title	Mobile Sensor Data Classification for Human Activity Recognition using MapReduce on Cloud
	Summary	• The paper proposed the utilization of parallel computing using MapReduce on the cloud for training and recognizing human activities based on classifiers that can easily scale in performance and accuracy. • The sensor data is extracted from the mobile, offloaded to the cloud and processed using three different classification algorithms, Iterative Dichotomizer 3, Naive Bayes Classifier and K-Nearest-Neighbors. • The MapReduce based algorithms are mentioned in detail along with one of their performance on Amazon cloud.
06	Title	LabelSens: enabling real-time sensor data labelling at the point of collection using an artificial intelligence-based approach

	Summary	• In this paper introduce new techniques for labelling at the point of collection coupled with a pilot study and a systematic performance comparison of two popular types of deep neural networks running on five custom built devices and a comparative mobile app (68.5-89% accuracy within-device GRU model, 92.8% highest LSTM model accuracy).
07	Title	On-Device Deep Learning Inference for Efficient Activity Data Collection
	Summary	• The novel idea behind this is that estimated activities are used as feedback for motivating users to collect accurate activity labels. To enable us to perform evaluations, we conduct the experiments with two conditional methods. We compare the proposed method showing estimated activities using on-device deep learning inference with the traditional method showing sentences without estimated activities through smartphone notifications. • the preliminary results indicate that our proposed method has improvements in F1-score, precision, and recall for all machine learning classifiers compared to the traditional method
08	Title	Real-Time Monitoring System Using Smartphone-Based Sensors and NoSQL Database for Perishable Supply Chain
	Summary	• This study proposes a real-time monitoring system that utilizes smartphone-based sensors and a big data platform. Firstly, we develop a smartphone-based sensor to gather temperature, humidity, GPS, and image data. • The IoT-generated sensor on the smartphone has characteristics such as a large amount of storage, an unstructured format, and continuous data generation. • We propose an effective big data platform design to handle IoT-generated sensor data • The results showed that the proposed system is capable of processing a massive input/output of sensor data efficiently when the number of sensors and clients increases

<table>
<tr><td>09</td><td>Title</td><td>Development of a Wearable-Sensor-Based Fall Detection System</td></tr>
<tr><td></td><td>Summary</td><td>

- This paper develops a novel fall detection system based on a wearable device. The system monitors the movements of human body, recognizes a fall from normal daily activities by an effective quaternion algorithm, and automatically sends request for help to the caregivers with the patient's location.

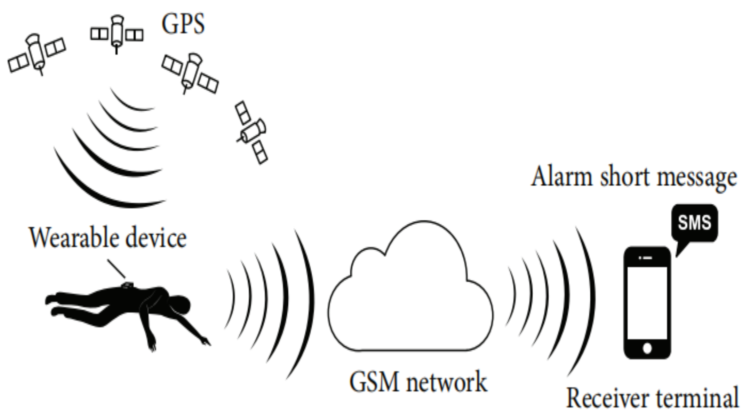


FIGURE 1: System architecture.

- Algorithm used in this fall alarm system is based on thresholds of sum acceleration and rotation angle information

</td></tr>
<tr><td>10</td><td>Title</td><td>Improving Fall Detection Using an On-Wrist Wearable Accelerometer</td></tr>
<tr><td></td><td>Summary</td><td>

- Falls are detected using a published threshold-based solution, although a study on threshold tuning has been carried out.
- The feature extraction is extended in order to balance the dataset for the minority class. Alternative models have been analyzed to reduce the computational constraints so the solution can be embedded in smart-phones or smart wristbands.
- Given the obtained results, the rule-based systems represent a promising research line as they perform similarly to neural networks, but with a reduced computational cost

</td></tr>
</table>

11	Title	Transfer learning approach for fall detection with the FARSEEING real-world dataset and simulated falls
	Summary	• The objective is to analyze if a combination of simulated and real falls could enrich the model. Falls are a sporadic event, which results in imbalanced datasets. • Several methods for imbalance learning were employed: SMOTE, Balance Cascade and Ranking models. • The Balance Cascade obtained less misclassifications in the validation set. There was an improvement when mixing the real falls and simulated non-falls compared to the case when only simulated falls were used for training. • When testing with a mixed set with real falls and simulated non-falls, it is even more important to train with a mixed set. to conclude that a model trained with simulated falls generalize better when tested with real falls, than the opposite. • The overall accuracy obtained for the combination of different datasets were above 95%.
12	Title	Vision-Based Fall Detection with Convolutional Neural Networks
	Summary	• The propose a vision-based solution using Convolutional Neural Networks to decide if a sequence of frames contains a person falling. To model the video motion and make the system scenario independent • we use optical flow images as input to the networks followed by a novel three-step training phase. • The proposed method is evaluated in three public datasets achieving the state-of-the-art results in all three of them • we presented a successful application of transfer learning from action recognition to fall detection to create a vision-based fall detector system which obtained the state-of the-art results in three public fall detection datasets, namely,URFD, Multicam, and FDD
13	Title	Validation of accuracy of SVM-based fall detection system using real-world fall and non-fall datasets

	Summary	• In this study, we examined the accuracy of a fall detection system based on real-world fall and non-fall datasets. • Five young adults and 19 older adults went about their daily activities while wearing tri-axial accelerometers. • Older adults experienced 10 unanticipated falls during the data collection. Approximately 400 hours of activities of daily living were recorded. • We employed a machine learning algorithm, Support Vector Machine (SVM) classifier, to identify falls and non-fall events. We found that our system was able to detect 8 out of the 10 falls in older adults using signals from a single accelerometer (waist or sternum). Furthermore, our system did not report any false alarm during approximately 28.5 hours of recorded data from young adults
14	Title	A Benchmark Database and Baseline Evaluation for Fall Detection Based on Wearable Sensors for the Internet of Medical Things Platform
	Summary	• A benchmark database, namely, a fall detection database, is presented to evaluate the performance of detection algorithms. This database collects sample data from 26 males and 24 females performing 15 kinds of activities, including falls and activities of daily life, such as walking, running, and walking upstairs. The subjects comprise 50 males and females ranging from 21 to 60 years of age, 1.55 to 1.90 m in height, and 40 to 85 kg in weight. • A full comparison between the existing databases and the proposed database is presented. Four baseline algorithms (the artificial neutral network, k nearest neighbor, support vector machine, and kernel Fisher discriminant) are used to evaluate the databases' reliabilities. • The algorithms have obtained different performance ratings using the different features and applying the same recognition methods. • The SVM-AdaBoost method has been used to compare and evaluate the performance of the benchmark algorithms based on our database, and the classification result is satisfactory • According to these results, the proposed database can be used to distinguish between a fall and an ADL
15	Title	Privacy Preserving Human Fall Detection using Video Data

<table>
<tr><td></td><td>Summary</td><td>
In this paper, we present a deep learning-based framework towards automatic fall detection from RGB images captured by a single camera.
Our framework learns human skeleton and segmentation based fall representations purely from synthetic data generated in a virtual environment.
This identifies personal information contained in the original images and preserves privacy which is highly desirable in health informatics
Our framework produces human proposals with body joint locations and segmentation information. These proposals are refined and transformed into multimodal visual representations for input to FallNet, a CNN model which uses modality-specific and multi-modal layers and learns highly discriminative feature embeddings for fall recognition.
We also present a human fall dataset which consists of human pose and segmentation data synthetically generated under different camera viewpoints.
Experiments on challenging public fall datasets show that our framework trained using only synthetically generated pose data successfully generalizes to unseen environments and achieves high precision and recall scores for fall recognition
</td></tr>
<tr><td>16</td><td>Title</td><td>Enhanced Human Activity Recognition Based on Smartphone Sensor Data Using Hybrid Feature Selection Model</td></tr>
<tr><td></td><td>Summary</td><td>
This research has proposed a hybrid method feature selection process, which includes a filter and wrapper method. The process uses a sequential floating forward search (SFFS) to extract desired features for better activity recognition.
Features are then fed to a multiclass support vector machine (SVM) to create nonlinear classifiers by adopting the kernel trick for training and testing purpose. We validated our model with a benchmark dataset.
</td></tr>
</table>

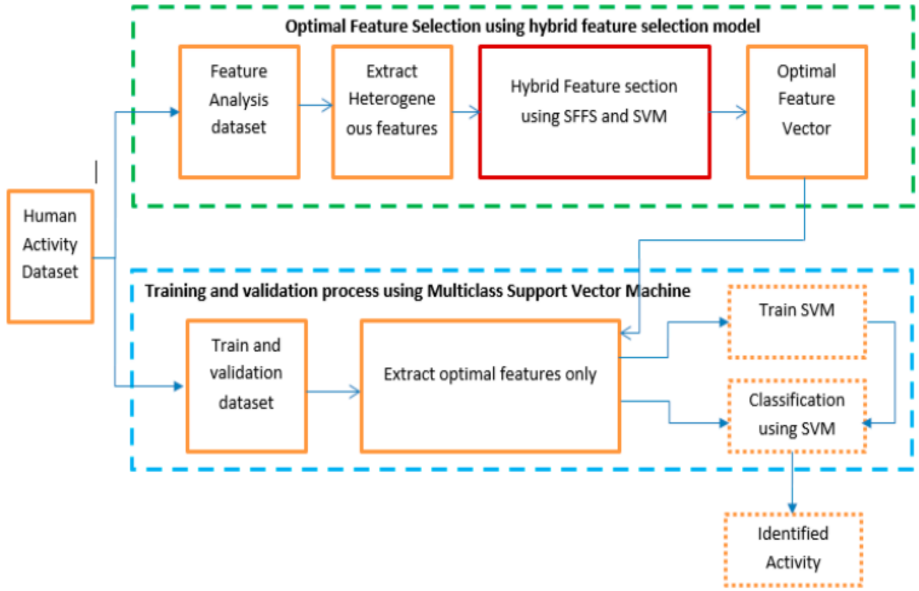
<table>
<tr><td></td><td></td><td>

Figure 1. Structure of the proposed model of activity identification.

the selected features are used for validation test using the SVM to identify the human activities.
The proposed system shows 96.81% average classification performance using optimal features, which is around 6% higher improved performance with no feature selection
</td></tr>
<tr><td>17</td><td>Title</td><td>Recognition of Daily Human Activity Using an Artificial Neural Network and Smartwatch</td></tr>
<tr><td></td><td>Summary</td><td>
In this study, we propose a human activity recognition system that collects data from an of the-shelf-smartwatch and uses an artificial neural network for classification.

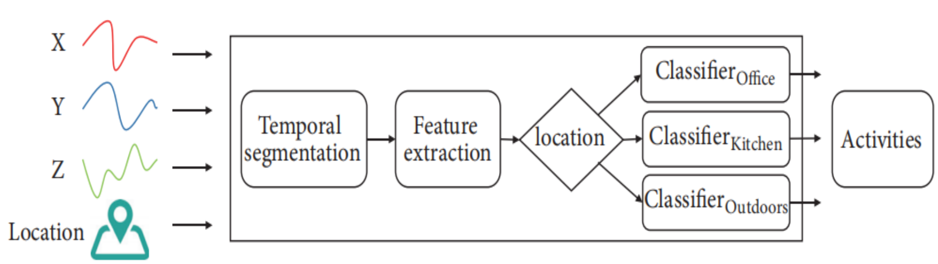


FIGURE 3: Proposed HAR system.</td></tr>
</table>

		• The proposed system is further enhanced using location information. We consider 11 activities, including both simple and daily activities. Experimental results show that various activities can be classified with an accuracy of 95%
18	Title	SparseSense: Human Activity Recognition from Highly Sparse Sensor Data-streams Using Set-based Neural Networks
	Summary	• In this paper, we rigorously explore the problem of learning activity recognition models from temporally sparse data • The process of operating a battery less sensor and transmitting the data captured is reliant on harvested power. Due to variable times to harvest adequate energy to operate sensors, the data-streams generated are highly sparse with variable inter-sample times. • Our work is built upon the insight that incorporating interpolation techniques to recover the missing measurements across large temporal gaps between received sensor observations in sparse data-streams leads to poor estimations and therefore, significant interpolation errors • We demonstrate significant classification performance improvements on real-world passive sensor datasets from older people over the state-of-the-art deep learning human activity recognition models • In contrast to previous studies that rely on interpolation pre-processing to synthesize sensory partitions with fixed temporal context, our proposed SparseSense network seamlessly operates on sparse segments with potentially varying number of sensor readings and delivers highly accurate predictions in the presence of missing sensor observations.
19	Title	Sensor Type, Axis, and Position-Based Fusion and Feature Selection for Multimodal Human Daily Activity Recognition in Wearable Body Sensor Networks
	Summary	• This research addresses the challenge of recognizing human daily activities using surface electromyography (sEMG) and wearable inertial sensors. • We propose a novel pipeline that can attain state-of-the-art recognition accuracies on a recent-and-standard dataset—the Human Gait Database (HuGaDB). Using wearable gyroscopes, accelerometers, and electromyography sensors placed on the thigh,

		shin, and foot, we developed an approach that jointly performs sensor fusion and feature selection. Being done jointly, the proposed pipeline empowers the learned model to benefit from the interaction of features that might have been dropped otherwise. • Using statistical and time-based features from heterogeneous signals of the aforementioned sensor types, our approach attains a mean accuracy of 99.8%, which is the highest accuracy on HuGaDB in the literature. • This research underlines the potential of incorporating EMG signals especially when fusion and selection are done simultaneously.
20	Title	Zero-Shot Human Activity Recognition Using Non-Visual Sensors
	Summary	• Activity recognition methods based on real-life settings should cover a growing number of activities in various domains, whereby a significant part of instances will not be present in the training data set. However, to cover all possible activities in advance is a complex and expensive task. • Concretely, we need a method that can extend the learning model to detect unseen activities without prior knowledge regarding sensor readings about those previously unseen activities. • we introduce an approach to leverage sensor data in discovering new unseen activities which were not present in the training set. • zero-shot learning is an extension of the supervised learning to overcome a well-known problem in machine learning when too few labeled examples are available for all classes. • We applied zero-shot learning to estimate occurrences of unseen activities. • Results show that our approach has achieved a promising accuracy for unseen new activities' recognition
21	Title	Semantic segmentation of real-time sensor data stream for complex activity recognition
	Summary	• Data segmentation plays a critical role in performing human activity recognition in the ambient assistant living systems. • It is particularly important for complex activity recognition when the events occur in short bursts with attributes of multiple sub-tasks

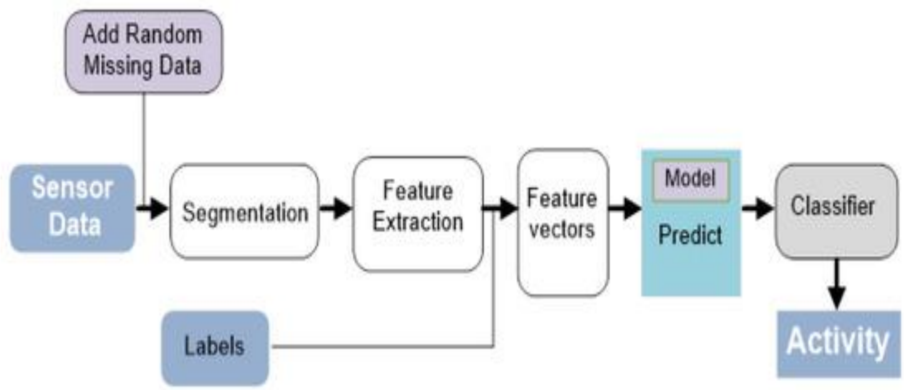
		• This paper proposes a semantic based approach for segmenting sensor data series using ontologies to perform terminology box and assertion box reasoning, along with logical rules to infer whether the incoming sensor event is related to a given sequences of the activity. • The proposed approach is illustrated using a use case scenario which conducts semantic segmentation of a real-time sensor data stream to recognize an elderly person's complex activities.
22	Title	A Smartphone Lightweight Method for Human Activity Recognition Based on Information Theory
	Summary	• Smartphones have emerged as a revolutionary technology for monitoring everyday life, and they have played an important role in Human Activity Recognition (HAR) due to its ubiquity. • The sensors embedded in these devices allows recognizing human behaviors using machine learning techniques. However, not all solutions are feasible for implementation in smartphones, mainly because of its high computational cost. • The proposed method, called HAR-SR, introduces information theory quantifiers as new features extracted from sensors data to create simple activity classification models, increasing in this way the efficiency in terms of computational cost. • Three public databases (SHOAIB, UCI, WISDM) are used in the evaluation process. The results have shown that HAR-SR can classify activities with 93% accuracy when using a leave-one-subject-out cross-validation procedure (LOSO).
23	Title	Real-time Activity Recognition in Wireless Body Sensor Networks: From Simple Gestures to Complex Activities
	Summary	• Real-time activity recognition using body sensor networks is an important and challenging task and it has many potential applications. In this paper, we propose a real-time, hierarchical model to recognize both simple gestures and complex activities using a wireless body sensor network. • We first use a fast, lightweight template matching algorithm to detect gestures at the sensor node level, and then use a discriminative pattern based real-time algorithm to recognize high-level activities at the portable device level.

		• We evaluate our algorithms over a real-world dataset. The results show that the proposed system not only achieves good performance (an average precision of 94.9%, an average recall of 82.5%, and an average real-time delay of 5.7 seconds), but also significantly reduces the network communication cost by 60.2%.
24	Title	HealthyLife: an Activity Recognition System with Smartphone using Logic-Based Stream Reasoning
	Summary	• This paper introduces a prototype we named HealthyLife which uses Answer set programming-based Stream Reasoning (ASR) in combination with Artificial Neural Network (ANN) to automatically recognize users' activities. • HealthyLife aims to provide statistics about user habits and provide suggestions and alerts to the user to help the user maintain a healthy lifestyle. • Besides detecting basic activities, HealthyLife is able to detect complex activities, which can be tracked for statistics for health-related purposes and rules can be used to map inferred activities and activity histories to suggestions for users, all within a logic-based rule framework
25	Title	Automatic Annotation for Human Activity Recognition in Free Living Using a Smartphone
	Summary	• Data annotation is a time-consuming process posing major limitations to the development of Human Activity Recognition (HAR) systems. The availability of a large amount of labeled data is required for supervised Machine Learning (ML) approaches, especially in the case of online and personalized approaches requiring user specific datasets to be labeled • we present (i) an automatic labeling method facilitating the collection of labeled datasets in free-living conditions using the smartphone, and (ii) we investigate the robustness of common supervised classification approaches under instances of noisy data. • We evaluated the results with a dataset consisting of 38 days of manually labeled data collected in free living. The comparison between the manually and the automatically labeled ground truth demonstrated that it was possible to obtain labels automatically with an 80-85% average precision rate.

		• Results obtained also show how a supervised approach trained using automatically generated labels achieved an 84% f-score (using Neural Networks and Random Forests); however, results also demonstrated how the presence of label noise could lower the f-score up to 64-74% depending on the classification approach (Nearest Centroid and Multi-Class Support Vector Machine).
26	Title	Modeling and discovering human behavior from smartphone sensing life-log data for identification purpose
	Summary	• In this research, we have collected user personal data from 37 students for 2 months which consist of 19 kinds of data sensors. • The goals of our research are to discover human behavior from the user smartphone life log data and based on those behavior data we want to build behavior model which can be used for user identification • We use and combine of many sensors instead only focus on one sensor because we realize that sometimes the users not have data from one or more sensors • Our system can handle the problem if one or more data sensors from users smartphone not available. Some of result from our system can achieve up to more than 80 % accuracy
27	Title	Validation Techniques for Sensor Data in Mobile Health Applications
	Summary	• Mobile applications have become a must in every user's smart device, and many of these applications make use of the device sensors' to achieve its goal. Nevertheless, it remains fairly unknown to the user to which extent the data the applications use can be relied upon and, therefore, to which extent the output of a given application is trustworthy or not. • To help developers and researchers and to provide a common ground of data validation algorithms and techniques, this paper presents a review of the most commonly used data validation algorithms, along with its usage scenarios, and proposes a classification for these algorithms. • The validation of the data collected by sensors in a mobile device is an important issue for two main reasons: the first one is the increasing number of devices and the applications that make use of the devices' sensors; the other is that also increasingly users

		rely on these devices and applications to collect information and make decisions that may be critical for the user's life and well-being. • This paper has presented a discussion on the different types of data validation methods such as faulty data detection, data correction, and assisting techniques or tools.
28	Title	Mobile Sensor Data Anonymization
		• Motion sensors such as accelerometers and gyroscopes measure the instant acceleration and rotation of a device, in three dimensions. • Raw data streams from motion sensors embedded in portable and wearable devices may reveal private information about users without their awareness. For example, motion data might disclose the weight or gender of a user, or enable their re-identification. • To address this problem, we propose an on-device transformation of sensor data to be shared for specific applications, such as monitoring selected daily activities, without revealing information that enables user identification. • We formulate the anonymization problem using an information-theoretic approach and propose a new multi-objective loss function for training deep autoencoders. • The trained autoencoder can be deployed on a mobile or wearable device to anonymize sensor data even for users who are not included in the training dataset. • The proposed anonymizing autoencoder lead to a promising trade-off between utility and privacy, with an accuracy for activity recognition above 92% and an accuracy for user identification below 7%.
29	Title	A Method for Sensor-Based Activity Recognition in Missing Data Scenario
		• There are numerous works in this field—to recognize various human activities from sensor data. However, those works are based on data patterns that are clean data and have almost no missing data, which is a genuine concern for real-life healthcare centers. Therefore, to

<table>
<tr>
<td></td>
<td></td>
<td>address this problem, we explored the sensor-based activity recognition when some partial data were lost in a random pattern

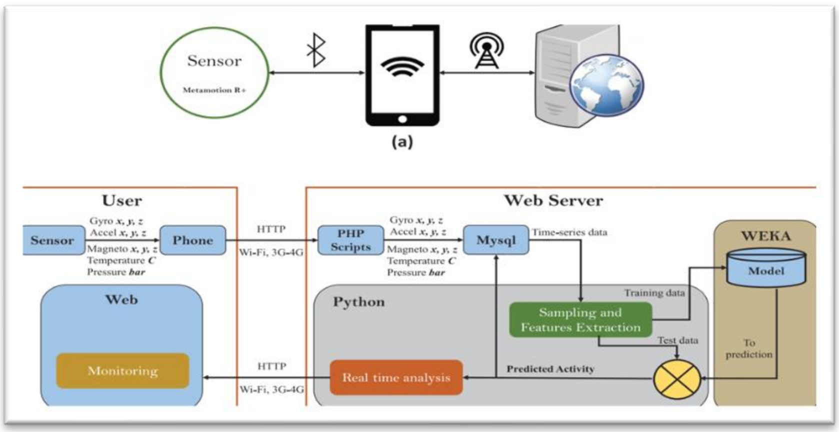


In our proposed approach, we explicitly induce different percentages of missing data randomly in the raw sensor data to train the model with missing data. Learning with missing data reinforces the model to regulate missing data during the classification of various activities that have missing data in the test module. This approach demonstrates the plausibility of the machine learning model, as it can learn and predict from an identical domain.
We developed a synthetic dataset to empirically evaluate the performance and show that the method can effectively improve the recognition accuracy from 80.8% to 97.5%. Afterward, we tested our approach with activities from two challenging benchmark datasets: the human activity sensing consortium (HASC) dataset and single chest-mounted accelerometer dataset. We examined the method for different missing percentages, varied window sizes, and diverse window sliding widths. Our explorations demonstrated improved recognition performances even in the presence of missing data

</td>
</tr>
<tr>
<td>30</td>
<td>Title</td>
<td>Wearable Internet-of-Things platform for human activity recognition and health care</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
<td>

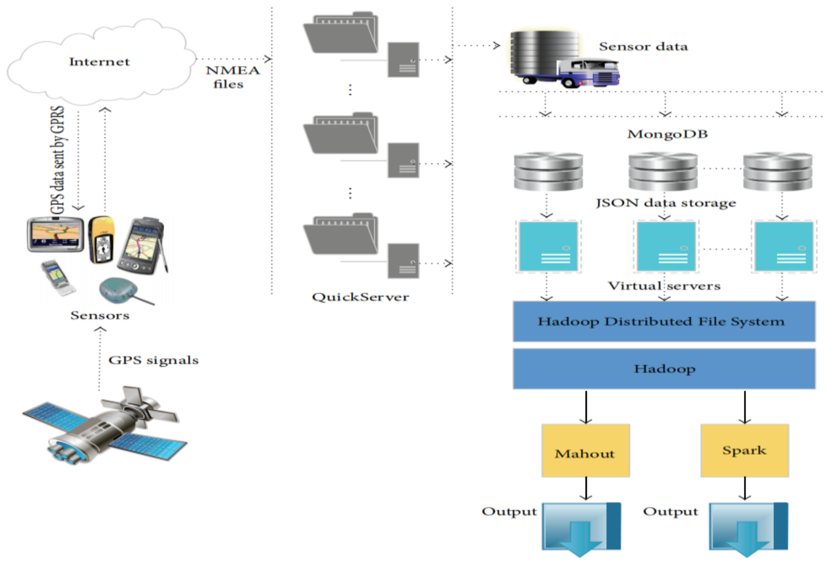
We propose to perform wearable sensors-based human physical activity recognition. This is further extended to an Internet-of-Things (IoT) platform which is based on a web-based application that integrates wearable sensors, smartphones, and activity recognition.
To this end, a smartphone collects the data from wearable sensors and sends it to the server for processing and recognition of the physical

</td>
</tr>
</table>

		activity. We collect a novel data set of 13 physical activities performed both indoor and outdoor. The participants are from both the genders where their number per activity varies. • During these activities, the wearable sensors measure various body parameters via accelerometers, gyroscope, magnetometers, pressure, and temperature. These measurements and their statistical are then represented in features vectors that used to train and test supervised machine learning algorithms (classifiers) for activity recognition. 
31	Title	Efficacy of Imbalanced Data Handling Methods on Deep Learning for Smart Homes Environments
	Summary	• The frequency and duration of human activities are intrinsically imbalanced. The huge difference in the number of observations for the classes to learn will make many machine learning algorithms to focus on the classification of the majority examples due to its increased prior probability while ignoring or misclassifying minority examples. • SMOTE and cost-sensitive learning are applied to temporal models and compared with ensemble learning to handle the class imbalance problem as well as to study the relation to two data pre-processing methods. • Experiments show that f-measures of the minority classes are increased when using SMOTE with both temporal models (LSTM and CNN) and based on both ways of extracting features (FTWs and ESTWs).

		• The experimental results indicate that handling imbalanced data is more important than selecting machine learning algorithms and improves classification performance. Moreover, handling imbalanced class problem from data level using SMOTE and ESTWs for these activity datasets outperforms the algorithm level.
32	Title	Dealing with Imbalanced data sets for Human Activity Recognition using Mobile Phone sensors
	Summary	• The wide spreading of smart-phones which are daily carried by humans and fit with tens of sensors triggered an intense research activity in human activity recognition (HAR). • Many statistical and logical based models for on-line or off-line HAR have been designed, however, the current trend is to use deep-learning with neural network. • These models need a high amount of data and, as most discriminative models, they are very sensitive to the imbalanced class problem. • We study different ways to deal with imbalanced data sets to improve accuracy of HAR with neural networks and introduce a new oversampling method, called Border Limited Link SMOTE (BLL SMOTE) which improves the classification accuracy of Multi-Layer Perceptron (MLP) performances. • These results show two advantages over classical approaches: the method makes it possible to improve overall and local performances and does not require extra external data.
33	Title	Modeling temporal aspects of sensor data for MongoDB NoSQL database
	Summary	• The next generation systems demand horizontal scaling by distributing data over autonomously addable nodes to a running system. • For schema flexibility, they also want to process and store different data formats along the sequence factor in the data. NoSQL approaches are solutions to these, hence big data solutions are vital nowadays. But in monitoring scenarios sensors transmit the data continuously over certain intervals of time and temporal factor is the main property of the data. • The key research aspect is to investigate schema flexibility and temporal data integration aspects together.

		• We need to know that: what data modelling should we adopt for a data driven real-time scenario; that we could store the data effectively and evolve the schema accordingly during data integration in NoSQL environments without losing big data advantages. In this paper we explain a middleware-based schema model to support the temporal oriented storage of real-time data of ANT+ sensors as hierarchical documents. We explain how to adopt a schema for the data integration by using an algorithm-based approach for flexible evolution of the model for a document-oriented database, i.e, MongoDB. • We define denormalized schema to have a document for each hour, which contains minutes as sub documents containing sensor data in an array of seconds' sub-documents. The normalization and denormalization of the document hierarchy decides the quality of a schema with respect to number of reads, updates and storage space utilization.
34	Title	Architecture and Implementation of a Scalable Sensor Data Storage and Analysis System Using Cloud Computing and Big Data Technologies
	Summary	• Sensors are becoming ubiquitous. Data produced by these sensors is much more dramatic since sensors usually continuously produce data. • It becomes crucial for these data to be stored for future reference and to be analyzed for finding valuable information, such as fault diagnosis information. • We describe a scalable and distributed architecture for sensor data collection, storage, and analysis. The system uses several open source technologies and runs on a cluster of virtual servers. We use GPS sensors as data source and run machine-learning algorithms for data analysis.

		 FIGURE 4: System architecture. • We demonstrated the architecture and test results for a distributed sensor data collection, storage, and analysis system. The architecture can be scaled to support a large number of sensors and big data sizes. It can be used to support geographically distributed sensors and collect sensor data via a high-performance server. • The test results show that the system can execute computationally complex data analysis algorithms and shows high performances with big sensor data. As a result, we show that, using open source technologies, modern cloud computing and big data frameworks can be utilized for large-scale sensor data analysis requirements.
35	Title	Data Management for the Internet of Things: Design Primitives and Solution
	Summary	• The solutions to manage and utilize the massive volume of data produced by these objects are yet to mature. Traditional database management solutions fall short in satisfying the sophisticated application needs of an IoT network that has a truly global-scale. Current solutions for IoT data management address partial aspects of the IoT environment with special focus on sensor networks. • In this paper, we survey the data management solutions that are proposed for IoT or subsystems of the IoT. We highlight the distinctive design primitives that we believe should be addressed in an IoT data management solution, and discuss how they are approached by the proposed solutions.

		• We finally propose a data management framework for IoT that takes into consideration the discussed design elements and acts as a seed to a comprehensive IoT data management solution. • The framework we propose adapts a federated, data- and sources-centric approach to link the diverse Things with their abundance of data to the potential applications and services that are envisioned for IoT. • The design primitives we propose cover the three main functions of handling data; how it is collected, how it is stored, and how it is processed. • The framework highlights the need for two-way, cross-layered design approach that can address both real-time and archival query, analysis, and service needs.
36	Title	Human Activity Recognition with Streaming Smartphone Data
	Summary	• With the widely used smartphones, dynamic data coming from built in sensors, such as human activity data, can be easily obtained. Many applications' developments, such as applications in healthcare, fitness monitoring, and elder monitoring, are based on this kind of dynamic data. Although there are many offline methods that have made a great progress in analyzing these kinds of data, it still has a big challenge to get good results from a streaming data perspective. • We use an online method called Very Fast Decision Tree (VFDT) to mimic the real scenario. There are two main improvements from the existing models: a. we train the model online and only use the examples data once for training instead of using them more than once; b. after building VFDT, the model can be adjusted to identify new activities by adding only small number of labeled observations. • Our experiment on the same existing activities shows that the proposed algorithm achieves an average accuracy of 85.9% for all subjects and single subject accuracy rates are between 60.5% and 99.3%. Moreover, the average accuracy of learning new activity from a different data is 84% and single subject accuracy rate goes to as high as 100%
37	Title	Developing an On-Demand Cloud-Based Sensing-as-a-Service System for Internet of Things

	Summary	• The increasing number of Internet of Things (IoT) devices with various sensors has resulted in a focus on Cloud-based sensing-as a-service (CSaaS) as a new value-added service, for example, providing temperature-sensing data via a cloud computing system. • However, the industry encounters various challenges in the dynamic provisioning of on-demand CSaaS on diverse sensor networks. • We require a system that will provide users with standardized access to various sensor networks and a level of abstraction that hides the underlying complexity. • We aim to develop a cloud-based solution to address the challenges mentioned earlier. • Our solution, SenseCloud, includes a sensor virtualization mechanism that interfaces with diverse sensor networks, a multitenancy mechanism that grants multiple users access to virtualized sensor networks while sharing the same underlying infrastructure, and a dynamic provisioning mechanism to allow the users to leverage the vast pool of resources on demand and on a pay-per-use basis. • We implement a prototype of SenseCloud by using real sensors and verify the feasibility of our system and its performance. SenseCloud bridges the gap between sensor providers and sensor data consumers who wish to utilize sensor data.
38	Title	A Human Activity Recognition System Using Skeleton Data from RGBD Sensors
	Summary	• The aim of Active and Assisted Living is to develop tools to promote the ageing in place of elderly people, and human activity recognition algorithms can help to monitor aged people in home environments. • Different types of sensors can be used to address this task and the RGBD sensors, especially the ones used for gaming, are cost-effective and provide much information about the environment. • This work aims to propose an activity recognition algorithm exploiting skeleton data extracted by RGBD sensors. The system is based on the extraction of key poses to compose a feature vector, and a multiclass Support Vector Machine to perform classification. Computation and association of key poses are carried out using a clustering algorithm, without the need of a learning algorithm. • The proposed approach is evaluated on five publicly available datasets for activity recognition, showing promising results especially when applied for the recognition of AAL related actions.

39	Title	Activity Feature Solving Based on TF-IDF for Activity Recognition in Smart Homes
	Summary	• The performance of daily activity recognition heavily depends on solving strategy of activity feature. However, the current common employed solving strategy based on statistical information of individual activity does not support well the activity recognition. • To improve the common employed solving strategy, an activity feature solving strategy based on TF-IDF is proposed in this paper. • The proposed strategy exploits statistical information related to both individual activity and the whole of activities. Two distinct datasets have been commissioned, to mitigate against any possible effect of coupling between dataset and sensor configuration. • The proposed strategies were evaluated using three classifiers on two distinct datasets, and results obtained in this study demonstrate the ability of strategy based on TF-IDF to dramatically improve the performance of activity recognition systems.
40	Title	Daily Human Physical Activity Recognition Based on Kernel Discriminant Analysis and Extreme Learning Machine
	Summary	• Wearable sensor based human physical activity recognition has extensive applications in many fields such as physical training and health care. • This paper will be focused on the development of highly efficient approach for daily human activity recognition by a triaxial accelerometer. • In the proposed approach, a number of features, including the tilt angle, the signal magnitude area (SMA), and the wavelet energy, are extracted from the raw measurement signal via the time domain, the frequency domain, and the time frequency domain analysis. • A nonlinear kernel discriminant analysis (KDA) scheme is introduced to enhance the discrimination between different activities. Extreme learning machine (ELM) is proposed as a novel activity recognition algorithm. • Experimental results show that the proposed KDA based ELM classifier can achieve superior recognition performance with higher accuracy and faster learning speed than the back-propagation (BP) and the support vector machine (SVM) algorithms.

41	Title	Energy-Efficient Real-Time Human Activity Recognition on Smart Mobile Devices
	Summary	• Nowadays, human activity recognition (HAR) plays an important role in wellness-care and context-aware systems. Human activities can be recognized in real-time by using sensory data collected from various sensors built in smart mobile devices. • Recent studies have focused on HAR that is solely based on triaxial accelerometers, which is the most energy-efficient approach. However, such HAR approaches are still energy-inefficient because the accelerometer is required to run without stopping so that the physical activity of a user can be recognized in real-time. • In this paper, we propose a novel approach for HAR process that controls the activity recognition duration for energy-efficient HAR. We investigated the impact of varying the acceleration-sampling frequency and window size for HAR by using the variable activity recognition duration (VARD) strategy. • We implemented our approach by using an Android platform and evaluated its performance in terms of energy efficiency and accuracy. The experimental results showed that our approach reduced energy consumption by a minimum of about 44.23% and maximum of about 78.85% compared to conventional HAR without sacrificing accuracy.
42	Title	Feature Selections Using Minimal Redundancy Maximal Relevance Algorithm for Human Activity Recognition in Smart Home Environments
	Summary	• In this paper, maximal relevance measure and minimal redundancy maximal relevance (mRMR) algorithm (under D-R and D/R criteria) have been applied to select features and to compose different features subsets based on observed motion sensor events for human activity recognition in smart home environments. • The selected features subsets have been evaluated and the activity recognition accuracy rates have been compared with two probabilistic algorithms: naıve Bayes (NB) classifier and hidden Markov model (HMM). The experimental results show that not all features are beneficial to human activity recognition and different features subsets yield different human activity recognition accuracy rates.

		• It is significant for researchers performing human activity recognition to consider both relevance between features and activities and redundancy among features. Generally, both maximal relevance measure and mRMR algorithm are feasible for feature selection and positive to activity recognition.
43	Title	Gender Recognition from Unconstrained and Articulated Human Body
		• Gender recognition has many useful applications, ranging from business intelligence to image search and social activity analysis. Traditional research on gender recognition focuses on face images in a constrained environment. • This paper proposes a method for gender recognition in articulated human body images acquired from an unconstrained environment in the real world. A systematic study of some critical issues in body-based gender recognition, such as which body parts are informative, how many body parts are needed to combine together, and what representations are good for articulated body-based gender recognition, is also presented. • This paper also pursues data fusion schemes and efficient feature dimensionality reduction based on the partial least squares estimation. Extensive experiments are performed on two unconstrained databases which have not been explored before for gender recognition.
44	Title	Position-Based Feature Selection for Body Sensors regarding Daily Living Activity Recognition
	Summary	• This paper proposes a novel approach to recognize activities based on sensor-placement feature selection. The method is designed to address a problem of multi sensor fusion information of wearable sensors which are located in different positions of a human body. • Precisely, the approach can extract the best feature set that characterizes each activity regarding a body-sensor location to recognize daily living activities. • We firstly preprocess the raw data by utilizing a low-pass filter. After extracting various features, feature selection algorithms are applied separately on feature sets of each sensor to obtain the best feature set for each body position.

		• We investigate the correlation of the features in each set to optimize the feature set. Finally, a classifier is applied to an optimized feature set, which contains features from four body positions to classify thirteen activities. • We obtain an overall accuracy of 99.13% by applying the proposed method to the benchmark dataset. The results show that we can reduce the computation time for the feature selection step and achieve a high accuracy rate by performing feature selection for the placement of each sensor. In addition, our proposed method can be used for a multiple-sensor configuration to classify activities of daily living. • The method is also expected to deploy to an activity classification system-based big data platform since each sensor node only sends essential information characterizing itself to a cloud server.
45	Title	Human Pose Recognition Based on Depth Image Multi feature Fusion
	Summary	• The recognition of human pose based on machine vision usually results in a low recognition rate, low robustness, and low operating efficiency. That is mainly caused by the complexity of the background, as well as the diversity of human pose, occlusion, and self-occlusion. • To solve this problem, a feature extraction method combining directional gradient of depth feature (DGoD) and local difference of depth feature (LDoD) is proposed in this paper, which uses a novel strategy that incorporates eight neighborhoods points around a pixel for mutual comparison to calculate the difference between the pixels. • A new data set is then established to train the random forest classifier, and a random forest two-way voting mechanism is adopted to classify the pixels on different parts of the human body depth image. • Finally, the gravity center of each part is calculated and a reasonable point is selected as the joint to extract human skeleton. • The experimental results show that the robustness and accuracy are significantly improved, associated with a competitive operating efficiency by evaluating our approach with the proposed data set.

My Interested Topics

Topics	References Paper
Missing Data	[29] A Method for Sensor-Based Activity Recognition in Missing Data Scenario
Imbalance Data	[3] Efficacy of Imbalanced Data Handling Methods on Deep Learning for Smart Homes Environments
Streaming Mobile Senser Data	[36] Human Activity Recognition with Streaming Smartphone Data
Real-time	[41] Energy-Efficient Real-Time Human Activity Recognition on Smart Mobile Devices
Feature Selection Extraction	[42] Feature Selections Using Minimal Redundancy Maximal Relevance Algorithm for Human Activity Recognition in Smart Home Environments