Bracelet-based Monitoring and Analysis Tool for Daily Life Behaviors

Binmei Hu\(^1\), Ting Li\(^2\)

\(^1\)Northeastern University, Shenyang City, Liaoning Province, China
\(^2\)Northeastern University, Shenyang City, Liaoning Province, China

Email of all the authors:
\(^*\)Corresponding Author: Binmei Hu; email: 1444240096@qq.com
\(^*\)Corresponding Author: Ting Li; email: 568068758@qq.com

Abstract: As China’s economy develops and people’s attitudes change, the birth rate of newborns in China is decreasing year by year, but the population over 65 is increasing year by year. China is moving towards an aging society. The problem of providing for the aged is gradually becoming a national problem. Faced with such a social problem, we want to solve such a problem through the development of science and technology. In response to the fall of an elderly person and the fact that he was left unattended, this paper presents a monitoring and analysis tool of daily life behavior based on bracelet. The tool collects data about the human body through a bracelet sensor, the machine learning algorithm is used to recognize the human motion behavior, and then the neighboring people or relatives are notified by alarms or text messages, so as to achieve the effect of guardianship, which not only avoids the elderly falling unnoticed, but also reduces the difficulty of guardianship without watching all the time.

1. Introduction

1.1. WEARABLE DEVICE

Wearable devices are usually composed of multiple sensors, which can be worn directly by the user, and generate corresponding physical signals, biological signals and location information when the user moves. Mobile phones and bracelets are typical wearable sensor devices, which can effectively collect the information of body position, motion, pulse and body temperature without affecting the activities of users. Researchers have found that different types of sensor information can effectively identify different types of activity. At present, there are mainly the following kinds of sensors used in sensor-based activity recognition at home and abroad:

1.1.1. ACCELERATION SENSOR

Accelerometer is the most important wearable sensor in motion recognition, it is responsible for measuring the linear acceleration of the vehicle, and it is very effective in the recognition of repeated body movements. Many researchers have used acceleration sensors to carry out related research. Lester[1] used a combination of accelerometers and other equipment, and used HMM to classify, with an accuracy of 90%; Yang [2] collected data from seven objects wearing watches with built-in accelerometers; Deng [3] collected data on six activities carried out by 30 subjects; Moataz used acceleration data to categorize activities that were performed on a 50 Hz smartphone accelerometer.
1.1.2. **GYROSCOPE**
Gyroscope is also one of the important sensors to collect human body information, which is used to collect and measure the angular velocity of the vehicle. It is often used in conjunction with accelerometers in activity recognition and performs well in collecting human data. The SBHAR data set published in the Anguita’s study\(^4\) was collected by two accelerometers and a gyroscope.

1.1.3. **BIOSENSOR**
Biosensor is a new technology, which can monitor human activities by analyzing life signals, including blood pressure, heart rate and so on. Sung detected the body temperature of the soldier by means of a body surface temperature sensor.

1.1.4. **IMU**
Inertial measurement unit (IMU) is a kind of measurement unit composed of many sensors. Accelerometers, gyroscopes and magnetometers are important parts of IMU. The PAMAP2 data set, published by Reiss\(^5\) in 2012, was captured by three IMUs and a heart rate detector.

1.2. **Related Algorithm**

1.2.1. **Naive Bayesian Model**
Naive Bayesian method is a classification method based on Bayesian theorem and independent hypothesis of characteristic condition. The core of naive Bayesian is Bayesian law, and the cornerstone of Bayesian law is conditional probability.

\[
P(B|A) = \frac{P(B|A)P(B)}{P(A)} \tag{1}
\]

Where \(P(A|B)\) denotes the probability of event \(A\) occurring when event \(B\) has occurred, which is called the conditional probability of event \(A\) when event \(B\) has occurred.

Naive Bayesian algorithm is based on the idea that for the given items to be classified, the probability of the occurrence of each category under the condition of the occurrence of this item is calculated, and the corresponding category of the maximum probability option is determined as the category of the items to be classified.

The formal definition of the Naive Bayes Classification is as follows:

Let \(x = \{a_1, a_2, ..., a_m\}\) be an item to be classified, where \(a\) is a feature attribute of \(x\);

Let \(C = \{y_1, y_2, ..., y_n\}\);

For each \(y\), calculate \(P(y_1|x), P(y_2|x), ..., P(y_n|x)\);

If \(P(y_k|x) = \max \{P(y_1|x), P(y_2|x), ..., P(y_n|x)\} \ (1 \leq k \leq n)\), then \(x \in y_k\).

If each feature attribute is conditionally independent, the following derivation can be made:

\[
P(y_i|x) = \frac{P(x|y_i)P(y_i)}{P(x)} \tag{2}
\]

In this case, the denominator is constant for all categories, and each characteristic attribute is conditionally independent, so there is

\[
P(x | y_i)P(y_i) = P(a_1 | y_i)P(a_2 | y_i)...P(a_m | y_i)P(y_i) = P(y_i)\prod_{j=1}^{m} P(a_j | y_i) \tag{3}
\]

From the above formula, it is concluded that calculating the conditional probability \(P(a|y)\) of each partition is the key step of Naive Bayesian Classification. When the feature attribute is discrete, only the frequency of each partition in each class in the training sample is needed to estimate \(P(a|y)\); and when the feature attribute is discrete, only the frequency of each partition in each class can be calculated. When the characteristic attribute is a continuous value, it is generally assumed that the attribute value obeys a normal distribution, which is

\[
g(x;\eta,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\eta)^2}{2\sigma^2}} \tag{4}
\]

For \(p(a|y)\), there is
At this time, the mean value and the standard deviation of the feature item in each category of the training sample are calculated, and the required estimation value can be obtained. In addition, for the case of \( P(a|y) = 0 \), that is to say, a certain feature item division does not appear in a certain category, Laplace calibration can be introduced, and the result will not be affected when the training sample set is large enough.

1.2.2. SVM

Support vector machine (SVM) is also a common classification algorithm, which improves the generalization ability of learning machine by seeking structural risk minimization, minimizes the empirical risk and confidence range, and achieves the goal of obtaining good statistical rules in the case of less statistical samples. Faced with a set of training instances, each of which is marked as belonging to one of the two categories, the SVM training algorithm creates a model that assigns a new instance to one of the two categories so that it becomes a non-probabilistic binary linear classifier. The SVM model represents instances as points in space, so the mapping separates instances of individual categories by as wide a distinct interval as possible. New instances are then mapped to the same space and the category is predicted based on which side of the interval they fall. Simply put, the solution of SVM is to find a hyper plane in an n-dimensional vector space with two kinds of points, which separates the two types, and the interval between the hyper plane and the point set is the largest.

Suppose we have given some data points in an instance that belong to two different classes. Now we need to find a linear classifier to divide the data into two classes. If the data points are represented by \( x \) and the categories by \( y \) (with a value of 1 or -1), the learning goal of a linear classifier is to find a hyper plane in the n-dimensional data space. The hyper plane equation can be expressed as

\[
P(ak | yi) = g(ak, \eta yi, \sigma yi)
\]

\[
w \cdot x - b = 0
\]

\[
1 = \frac{w \cdot x - b}{||w||}
\]

Where \( \cdot \) represents the dot product, \( w \) represents the normal vector of the hyper plane, and the parameter \( b \) determines the displacement of the hyper plane from the origin along the normal vector \( w \).

When the data is linearly separable, two hyper planes can divide the data and no points fall between the planes. We need to maximize the distance between these two hyper planes, whose equations can be described as follows:

\[
w \cdot x - b = 1
\]

\[
w \cdot x - b = -1
\]

The distance between two planes is \( 2/||w|| \) from the principle of geometry, so we minimize \( w \). For preventing data points from falling into the edge, we can make \( yi (w^* xi - b) > = 1 \) (1 <= i <= N), then maximizing the edge will be equivalent to minimizing the following function \( F(w) = (w^2)/2 \).

1.2.3. M5

Model tree is a kind of decision tree which adopts linear regression function in leaf nodes. It can realize classification by transforming classification problem into function optimization problem. A model tree represents a linear function that, like a typical regression equation, predicts the value of a variable (called a class) through a series of independent variables (called attributes). Training data in the form of tables can be directly used to construct decision trees. In the data sheet, each row (sample) is represented as \( (x1, x2, ..., xn, y) \), where \( xi \) represents the value of the ith attribute and \( y \) is the class value (target). For a given data set, the model tree divides the sample space into rectangular regions with parallel edges, and determines a corresponding regression model for each region.

The model tree tests the value of a particular attribute at each internal node and predicts the class value at each leaf node. Given a new sample to predict its class value, the tree interprets it from the root node; At each internal node, a left branch or a right branch is selected according to a certain at-
tribute value of that sample, and when the selected node is a leaf node, the output is predicted by the model of the leaf node.

The structure of the model tree is recursively generated, starting from the entire training sample set. In each layer of the model tree, the most recognizable attribute is selected as the root node of the sub-tree, and the sample arriving at this node is divided into several subsets according to the value of its node attribute. M5 adopts variance induction as a heuristic method and fills the leaf nodes with constant values as the model. For discrete attributes, each branch of the internal node represents a possible value of the parent node's attribute; For consecutive attributes, the algorithm determines the segmentation points to discretize the attribute values. This construction algorithm is called recursively for each subtree of the model tree. When the variance of the set of class attributes of a sample arriving at a node is small enough or the number of samples is small enough, the construction algorithm of the tree stops. This node is a leaf node.

The difference between model tree algorithm and simple linear regression is that the segmentation of input space is carried out automatically by the algorithm. It can effectively learn, can deal with input attributes up to hundreds of dimensions of the problem. The result of the model tree is simple and easy to understand, and it is a numerical prediction algorithm with wide application prospects.

2. Material and Methods

2.1. Data Acquisition

2.1.1. Data Acquisition Tool
The data acquisition tool uses WT901BLECL attitude and angle sensor module produced by Weite Intelligent Company. The module integrates high-precision gyroscope, accelerometer and geomagnetic field sensor, and adopts advanced digital filtering technology, which can effectively reduce the measurement noise and improve the measurement accuracy. Bluetooth BLE 4.0 module embedded in the internal support Android/IOS system can connect to mobile phones for wireless transmission is the core module inside the bracelet. As there is no open source interface for bracelet products on the Chinese market, sensor modules with bracelet functionality are selected to facilitate data collection and reduce the project cycle.

2.1.2. Data Acquisition Method
Android BLE is used to connect the sensor module and store it in the SQLite database of Android. Android 4.3 introduces the core functions of BLE and provides corresponding APIs, through which applications can scan Bluetooth devices, query services, read and write the characters of devices and so on.

This paper mainly creates five classes when building Android application: BluetoothService, DeviceControlActivity, DeviceScanActivity, MyDatabaseHelper, Database. The main functions include scanning Bluetooth device, reading Bluetooth data, data display and storage.

As you scan a Bluetooth device and read Bluetooth data, the API classes used are: Bluetooth adapt (features include turn on Bluetooth scanning, Using a known MAC address, instantiate a Bluetooth Device for connecting to a Bluetooth device, and so on), Bluetooth Device (functions include connecting Bluetooth devices and obtaining information about their names, addresses, and binding status), Bluetooth Gatt (functions include reconnecting Bluetooth devices, discovering Bluetooth device services, and so on), Bluetooth GattService (functions include further obtaining Bluetooth data for bi-directional transmission), and Bluetooth Gatt Characteristic (functions include reading data sent from peripheral devices). The specific steps include: declaring required permission, initializing work before connecting Bluetooth, scanning Bluetooth device, connecting Bluetooth device, discovering service, reading data, registering and listening to Bluetooth device to realize real-time reading data of Bluetooth device, and disconnecting.
Since the bracelet data obtained from Bluetooth is hexadecimal, it needs to be converted according to the formula. Bluetooth upload data up to 20Byte, 1Byte packet header, 1Byte flag bit and 18Byte specific values (include: acceleration X, Y, Z, angular velocity X, Y, Z, angle X, Y, Z, each dimension data includes 1Byte low byte and 1Byte high byte in turn), calculate as follows:

X axis acceleration \(ax=((axH<<8)|axL)/32768*16g\) (\(g\) is the acceleration of gravity, 9.8\(m/s^2\))

Y axis acceleration \(ay=((ayH<<8)|ayL)/32768*16g\) (\(g\) is the acceleration of gravity, 9.8\(m/s^2\))

Z axis acceleration \(az=((azH<<8)|azL)/32768*16g\) (\(g\) is the acceleration of gravity, 9.8\(m/s^2\))

X-axis angular velocity \(wx=((wxH<<8)|wxL)/32768*2000(°/s)\)

Y-axis angular velocity \(wy=((wyH<<8)|wyL)/32768*2000(°/s)\)

Z-axis angular velocity \(wz=((wzH<<8)|wzL)/32768*2000(°/s)\)

Roll angle (x-axis) \(Roll=((RollH<<8)|RollL)/32768*180(°)\)

Roll angle (y-axis) \(Pitch=((PitchH<<8)|PitchL)/32768*180(°)\)

Roll angle (z-axis) \(Yaw=((YawH<<8)|YawL)/32768*180(°)\)

In order to be able to present the data, Android apps use a broadcast mechanism. Bluetooth rack data obtain from Bluetooth is sent out through custom broadcast in Bluetooth service class, broadcast is receive in DeviceControlActivity class, and data is displayed. The following four pictures show all the processes of the shelf:
2.2. Attitude Recognition Classification

In this paper, three different classification algorithms in weka are selected for attitude recognition. The results are from weka’s seeker interface, where all algorithm parameters are weka’s default parameters and are cross-validated using a 10-fold cross-validation (i.e., the data is divided into 10 copies, nine of which are used as training data and one as test data to be validated in turn). The results are as follows:

2.2.1. Naive Bayesian

Use weka.Classifiers.Bayes.NaiveBayes. This class was trained to correctly classify 90.2905% of the sample shelves, with other parameters as shown in the figure below. Fig. 5 and Fig.6 show a fusion matrix showing that actual classification of all feature tag.

![Figure 5. Naive Bayesian Algorithm result](image)
2.2.2. SVM

Using the weka.classifiers.functions.SMO class for training, the percentage of correctly classified instances is 88.0688%, and the other parameters are shown in the following figure. Fig.7 and Fig.8 show a fusion matrix showing that actual classification of all feature tag.

![Confusion Matrix](image)

**Figure 6. Naive Bayesian Confusion Matrix**

**Figure 7. SVM Algorithm result**
2.2.3. J48

Using the weka.classifiers.trees.J48 class for training, the percentage of correctly categorized instances is 99.6155%, and the other parameters are shown in the figure below. Fig.9 and Fig.10 show a fusion matrix showing that actual classification of all feature tag.

Time taken to build model: 5.93 seconds

--- Stratified cross-validation ---

--- Summary ---

Correctly Classified Instances 135468 99.6155 %
Incorrectly Classified Instances 623 0.3845 %
Kappa statistic 0.9959
Mean absolute error 0.0007
Root mean squared error 0.0217
Relative absolute error 0.5662 %
Root relative squared error 8.9858 %
Total Number of Instances 136022
Ignored Class Unknown Instances 74

Figure 8. SVM Confusion Matrix

Figure 9. J48 Algorithm result
3. Results
This paper collects the data of human activity through the bracelet, and classifies and recognizes actions through three different classification algorithms of machine learning, the accuracy of identification is close to 90% at least, as high as 99 per cent. These data fully show that it is feasible to use the data obtained from the bracelet to judge the active state of human body, and that it is feasible to use the data obtained from the bracelet to judge the active state of human body. Applying this conclusion to the real situation of guardianship of the elderly and children, the invention can judge whether the old children are in a dangerous state such as falling or not, and further can report to the people around through a mobile phone short message or an alarm, so that the old children can get timely help, thus avoiding the situation that the old children are in a dangerous state without timely notification, and also avoiding the state of constant nursing by a guardian.

4. Discussion
Many problems were also found in the experiment. First, in the initial phase of data acquisition, Mobile phones need to select label buttons in order to collect labeled data, in real life this operation is not realistic, the solution is to collect a sufficient number of data, according to different types of people classification training, to obtain training model, in order to better promote the application. Secondly, according to the results of the training and testing of the above three algorithms, Naive Bayesian training test results of the shortest time, using SVM results of the longest time, using J48 to obtain the best training effect, in real life should be achieved in the case of ensuring the accuracy of the results of the shortest time, in order to achieve the effect of timely monitoring. Third, look closely at the Confusion matrix, Although most action classification results are accurate, But there are still a small number of movements are recognized as other movements, especially similar movements, such as upstairs and downstairs, sweeping and mopping, drinking water and brushing teeth, these similar movements in the identification process caused a large error, in the future research can be carried out in depth. Fourth, in real life, human actions are continuous, how to obtain effective actions from the collected continuous data is also a part of the future research needs to be explored in depth.

5. Conclusions
In this paper, the data collection and motion recognition in attitude recognition are introduced in detail. Firstly, the background of the research topic and the related algorithms are introduced. Secondly, it introduces how to collect data and use the relevant technology. Then, the specific tools and results of identification are introduced. Finally, according to the experimental results, the application and limita-
tion in practice are expounded. The experimental results show that the elderly can be monitored by the bracelet data collection and machine learning classification algorithm to achieve the goal of value return to society through scientific and technological means.

Acknowledgements
The author of this paper and Professor Wang Botao of Northeastern University made a useful discussion. This work is supported in part by the National Innovation and Entrepreneurship Undergraduate Development Programme.

References
[1] Jonathan Lester, Tanzeem Choudhury, and Gaetano" Borriello. Pervasive computing: 4th international conference, pervasive 2006, dublin, ireland, may 7-10, 2006. proceedings. In A Practical Approach to Recognizing Physical Activities, pages 1–16. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
[2] Jhun-Ying Yang, Jeen-Shing Wang, and Chen Yen-Ping. Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers. Pattern Recognition Letters, 29(16):2213–2220, 2008.
[3] Wan-Yu Deng, Qing-Hua Zheng, and Zhong-Min Wang. Cross-person activity recognition using reduced kernel extreme learning machine. Neural Networks, 53:1–7, 2014.
[4] Moataz Kilany; Aboul Ella Hassanien; Amr Badr,"Accelerometer-based human activity classification using Water Wave Optimization approach",International Computer Engineering Conference,IEEE,175 - 180,29-30 Dec.2015.
[5] A.Reiss,D.Stricker,Introducing a new benchmarked dataset for activity monitoring, in:International Symposiumon Wearable Computers,2012.