Soccer Movement Classification

2020.3.21 武子越

CONTEXT

- 1 Goals and Background
- 2 Related Work
- 3 Data Description
- 4 Modeling and Result
- 5 Summary and Prospect

1 Goals and Background

Assisting sports training using data modeling

Action Identification

Using sensors to collect data from athletes can help us analyze the performance of a team during the race.

Athletes can have a better understanding of their physical state and tactical usage.

Data-based Sports guidance

Integrating the motion analysis algorithm into the application products and comparing it with some standard motion data can help people improve their movements' qualities and analyze their health status in the exercise.

1 Goals and Background

Existing Products

小米 www.mi.com

keep

1 Goals and Background

Movements in Soccer Training

Goals: Classify movements base on the sensor data

Some actions are frequently appear in the soccer training, by analyzing when these actions appears in the race can help us manage the performance of the athletes and improve the tactics in the race.

The most important actions in the soccer race are the following:

- Walking
- Jogging & Rushing
- Passing the ball
- Turn-back
- Shooting

2 Related Work

[Bulling *et al.*] A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys 2014 (CSUR)

- ✓ Discuss the framework of HAR systems
- ✓ Discuss the challenge for HAR and the principle of Activity Recognition Chain (ARC)

[Hammerla *et al.*] Let's (not) sticktogether: pair wise similarity biases cross-validation in activity recognition. Ubicomp 2015.

- ✓ Illustrate that the standard approach to model selection: (stratified) cross validation, is unfit for segmented time-series data.
- ✓ Present a novel approach to cross-validation that explicitly circumvents neighbourhood bias.

2 Related Work

[Y.Hammerla *et al.*] Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables, IJCAI 2016

- ✓ Discuss the data characteristics in Human Activity Recognition (HAR)
- ✓ The effects of different deep learning models were evaluated on existing data sets
- ✓ It is believed that different models have different applicable scenarios and it is impossible to directly determine which is better
- ✓ Although a comprehensive model comparison is provided, the basis for data partitioning is not specified, and only a few types of motion data are used to test the model.

3 Data Description

Data Sources

Sensor Data From Chu Kochen Honors College Soccer Team.

Sensor: Six-axis sensor from VETTE INTELLIGENCE company

Data collection: Bluetooth signal transmission

The data includes four movements: Walking, Jogging, Turning and Passing the ball. The training data comes from 16 members who plays soccer well.

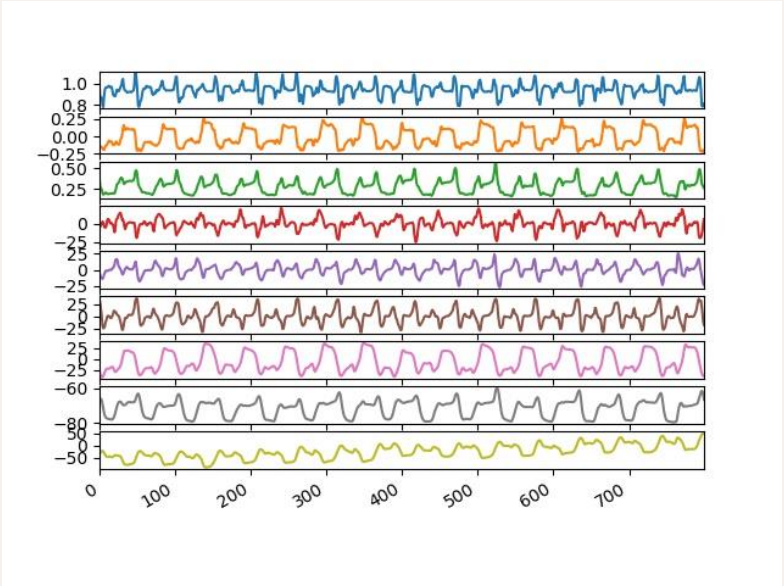
* In order to eliminate the irrelevant influence, the sensor was tied in the same position of left hip joint for each volunteer.

3 Data Description

Exploratory Data Analysis

The sensor can provide different physical data records.

Nine columns include angles, angular speed and acceleration in three dimensions.

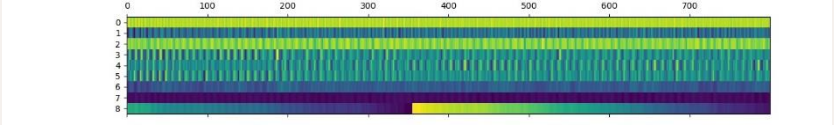


3 Data Description

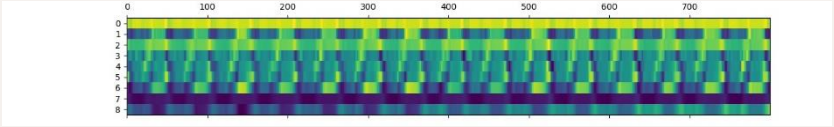
Exploratory Data Analysis

After the normalization, it is easy to see that different types of actions show different patterns.

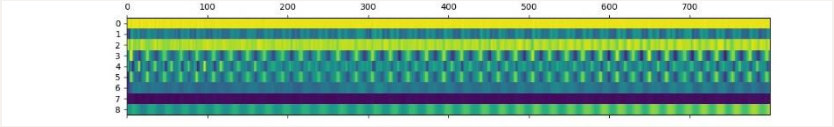
Running



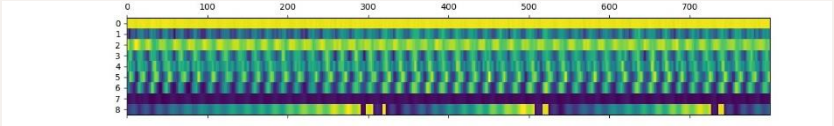
Passing the Ball



Turning



Jogging

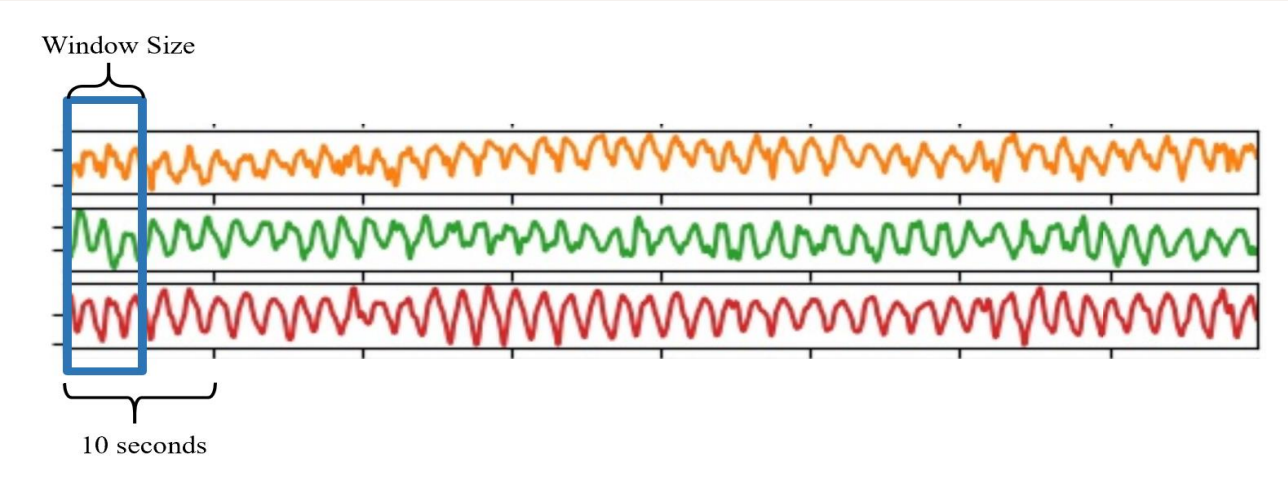


4 Modeling and Result

Data Processing

Sampling and Data Framing

- ✓ The sampling rate of the sensor is 10Hz, which means the sensor records the status ten times a second.
- ✓ The lasting of a movement in football is usually less than 2.5 seconds, so in order to record a complete period of movement, we created sliding windows to split the data to 5-second duration.



4 Modeling and Result

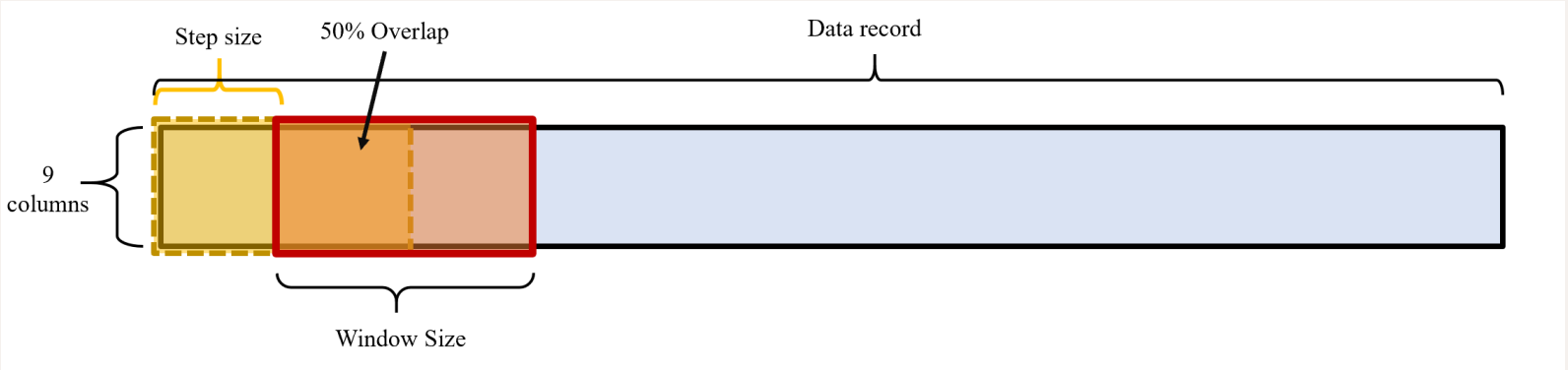
Data Processing

Two different framing method for the split of data.

- ✓ Sequence Framing
- ✓ Stochastic Framing

Sequence Framing

The sliding window with 5-second length moves over the data set from beginning to the end with 50% of overlap.

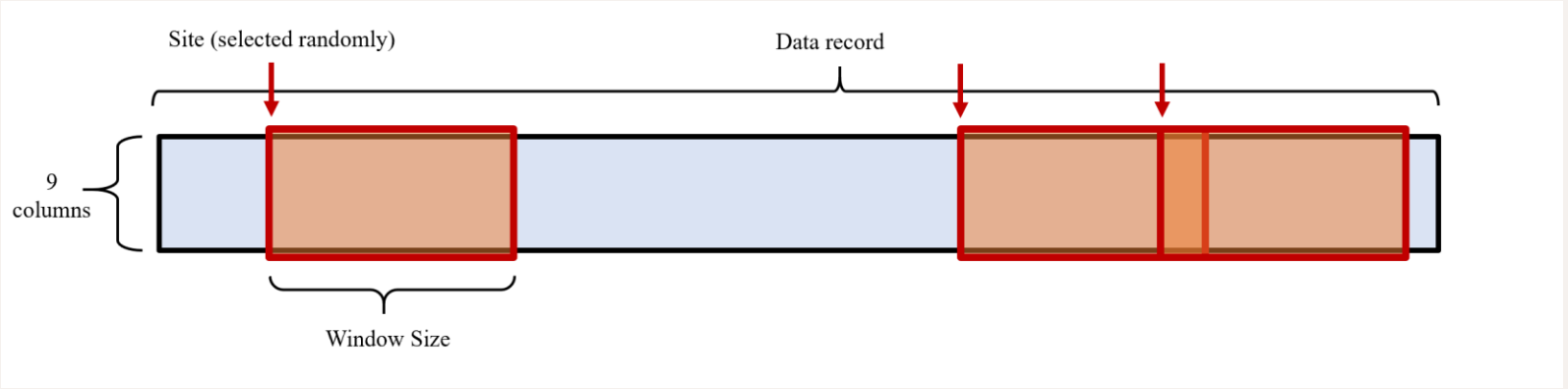


4 Modeling and Result

Data Processing

Stochastic Framing

Random sites were selected on the data set, and 5 seconds back from the site was selected as the data window.



4 Modeling and Result

Models and Outcome

How representative is the performance across the models compared during exploration? These issues are important for who would like to apply HAR technique on soccer race.

We used several different data mining models to compare the performance on soccer movement dataset. The models include MLP, Logistic Regression, Gradient Boost Decision Tree and SVM. Some deep learning approach are also applied, such as CNN, LSTM and Bi-LSTM.

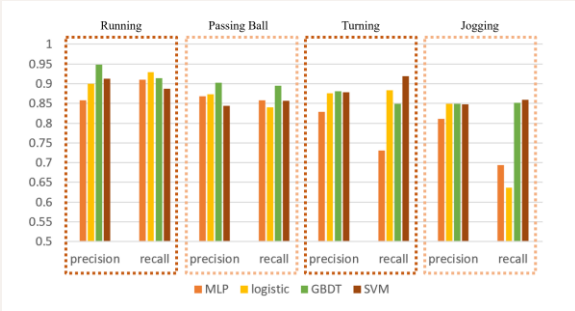
4 Modeling and Result

Performance of Traditional Models

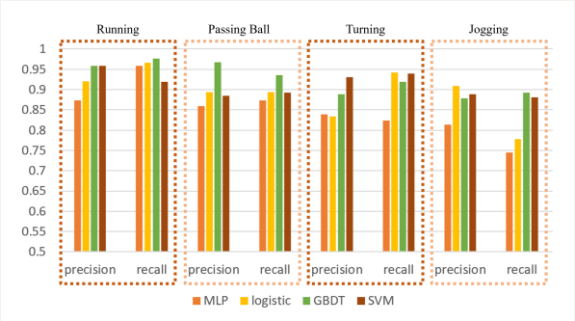
	Sequence Framing		Stochastic Framing	
	Overall Precision	Overall Recall	Overall Precision	Overall Recall
MLP	0.8467	0.8194	0.8535	0.8734
logistic	0.8807	0.8629	0.8854	0.9213
GBDT	0.9058	0.8828	0.9314	0.9399
SVM	0.8777	0.8867	0.9241	0.9148

- In the problem of soccer movement classification, some traditional models show good results.
- By analyzing the confusion matrix, models such as GBDT and SVM show high precision and recall rate.

Sequence Framing



Stochastic Framing



4 Modeling and Result

Performance of Deep Neural Network

Sequence Framing	Running		Passing Ball		Turning		Jogging	
	precision	recall	precision	recall	precision	recall	precision	recall
CNN	0.9234	0.9588	0.8939	0.9183	0.9184	0.9013	0.9184	0.8943
LSTM	0.9210	0.9355	0.9190	0.8922	0.9433	0.8928	0.8244	0.8123
Bi-LSTM	0.9567	0.9220	0.8899	0.8749	0.9459	0.9334	0.9284	0.8746

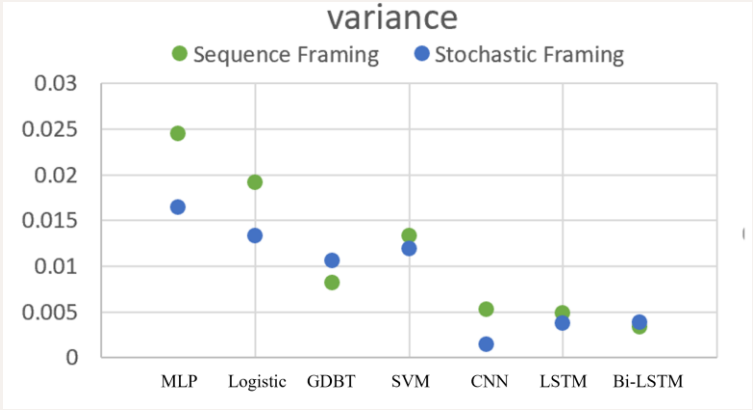
Stochastic Framing	Running		Passing Ball		Turning		Jogging	
	precision	recall	precision	recall	precision	recall	precision	recall
CNN	0.9677	0.9609	0.9770	0.9779	0.9684	0.9483	0.9360	0.9580
LSTM	0.9239	0.9404	0.9137	0.9039	0.9450	0.9300	0.8577	0.8602
Bi-LSTM	0.9644	0.9835	0.8938	0.9133	0.9872	0.9584	0.9213	0.8937

4 Modeling and Result

Models and Outcome

However, deep learning approach such as LSTM have a better and robust performance during the experiment.

The model take the influence of context into account, which is useful in action recognition task. This result is similar to studies on other motion datasets.



Similar with NLP problems, bi-directional semantic dependency also exists for motion data. In the experiment, adding a bi-directional structure can improve the performance of the model, not very much, though.

4 Modeling and Result

Models and Outcome

CNN is also a good method to classify movements. Every sliding windows can be seen as a picture. In this scenario, CNN capture the local features in each movement pattern and enable translational invariance, which is useful to distinguish different periodic movements with multi-dimensional time series data.

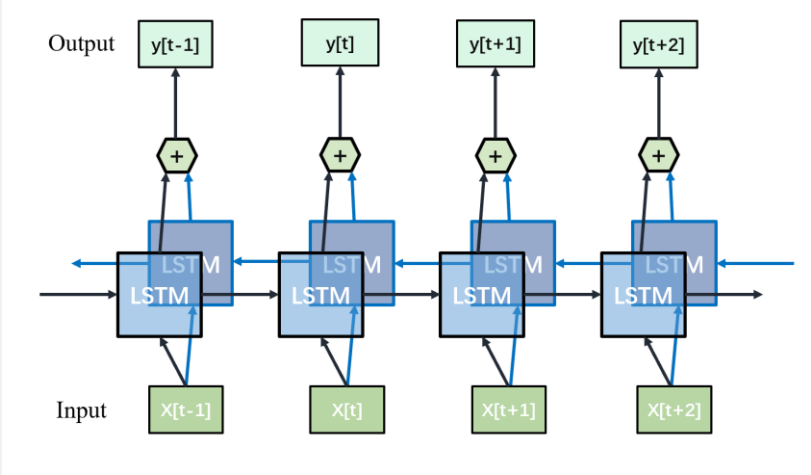
Data partition has a great influence on action classification. In the experimental work it demonstrates that Stochastic Framing can provide a better average performance in most cases.

This phenomenon may due to the data collection process. As time goes by, the movement frequency and range may also change due to physical strength consumption. So Sequence Framing may introduce bias to the modeling process.

4 Modeling and Result

Models and Outcome

The model with the best comprehensive performance is bi-LSTM in the case of motion recognition and motion data classification.



The model receives 9 dimensions data input at each time step and catenates with 18 dimensions internal state. It contains 1 hidden layers with two parallel tracks in both direction.

5 Summary and Prospect

Conclusion

Movement classification task can draw lessons from image recognition, text classification or speech recognition, some general deep learning models can also achieve positive results on soccer movement data.

However, movement classification problem have differences and unique challenge.

- Due to the temporal correlation of sensor data, general data partition is no longer applicable. It is necessary to divide training set and test set randomly and select proper framing size to prevent over fitting and reduce deviation.
- Two-way semantic dependency can also be considered to improve the performance.
- Even if the same action is completed by the same person at different times, or by different people, the data patterns will differ a lot.

5 Summary and Prospect

Prospect

Football sports recognition will bring a lot of convenience for sports training, at the same time, it can be used to develop a variety of application products. Here’re some example:

Action correction

Compare the personal data with the sensor data of the athletes, and carry out the action comparison and correction.

Health condition detection

Using sports data to track athletes' performance in real time and find health problems in time.

Team tactical evaluation

Compare the performance of each race and the collected sensor data, find problems and make tactical adjustments

Acknowledgement

Project Code

https://github.com/Tinky2013/Movements-classificaiton

Project Participants

Guidance: Professor Xi Chen, Department of data science and management engineering, Zhejiang University.

Data acquisition assistants: Chengbo Sun, Linen Yu

Volunteers: Hao Yang, Mali Luo, Wanhang Ma, Weibiao Song, Siyuan Qin, Yuan Sun, Chengbo Sun, Yizhou Lu, Haiwei Wang, Linen Yu, Linlin Ma, Anqi Lu, Haitian Yun, Yufei Chen, Jie Zhang.