Scholar Publishing Group
Machine Learning Theory and Practice
https://doi.org/10.38007/ML.2020.010303
ISSN 2790-0983 Vol. 1, Issue 3: 20-28

Intelligent Identification of Reservoir Fluid in Daniudi Gas Field Based on AdaBoost Machine Learning Algorithm

Junmin Gull*
Uni de Moncton, Canada
**corresponding author*

Keywords: AdaBoost Algorithm, Machine Learning, Daniudi Gas Field, Reservoir Fluid, Intelligent Identification

Abstract: Due to the large amount of data information in the reservoir, it is difficult to process or obtain accurate required parameters. There is a large amount of untreated gas in Daniudi gas field, and there are problems such as small wellhead, high temperature and rapid pressure change. Therefore, intelligent identification of reservoir fluid in Daniudi gas field is studied in this paper. Its purpose is to improve the recognition ability by using machine learning algorithm. This paper mainly uses the methods of experiment and comparison, selects 5 groups of data from the samples for comparison, and expounds the application of related algorithm models in reservoir fluids. The experimental data show that the error data of velocity density as input set and sensitive parameter are not very different, mostly within 1. But the input of sensitive parameters can get smaller error. Therefore, parameters with high sensitivity can be added for fluid identification.

1. Introduction

With the rapid development and updating of computer technology, people have a deeper understanding of large capacity, real-time computing. Under different wheel diameters, spectral characteristics of Daniudi gas field are established using SVM. The experimental data are preprocessed, and the corresponding neural network model is constructed by identifying the isotropic coefficient and dielectric constant n value of each parameter in the image and obtaining the corresponding relationship matrix KS-MLR representation of the pixel points. The training sample vector is used as the support vector, combined with AdaBoost algorithm to obtain the reservoir structure for the convenience of subsequent research.

There are many theories about machine learning algorithm, adaboost algorithm and gas field reservoir fluid identification. For example, some people put forward that fluid identification is difficult in Daniudi gas field due to complex reservoir, diversified mineral components, complex reservoir space and strong heterogeneity [1-2]. Some scholars have finally established a suitable fluid identification template in combination with the study of reservoir fluid properties and four property relations [3-4]. Some scholars also combined the logging response with the "four characteristics" relationship characteristics, and used core, testing, logging and other data to identify the fluid properties of the volcanic reservoir of Yingcheng Formation in the study area [5-6]. Therefore, it is necessary to detect the complex environment for gas field identification research. The Adaboost algorithm proposed in this paper can play a certain role.

In this paper, the reservoir fluid identification method is first studied, and the reservoir structure is also analyzed accordingly. Secondly, the principle of AdaBoost algorithm is analyzed to highlight its role in gas field identification. Then, a brief description of Daniudi gas field is given to clarify the purpose of the research object in this paper. Finally, the relevant conclusions are drawn through experiments and data analysis.

2. Gas Field Reservoir Fluid Identification Based on AdaBoost Machine Learning Algorithm

2.1. Reservoir Fluid Identification Method

With the exploration and development of low permeability and tight oil and gas fields, a series of new logging technologies have been developed rapidly, such as the "addition" of nuclear magnetic resonance logging calculation, which has improved the ability to use this technology to describe reservoirs in detail and identify fluids. It also provides more new methods for identifying fluid properties in tight reservoirs. Reservoir fluid logging facies is the most important part of reservoir logging facies. The reservoir fluid logging facies, literally, refers to the properties of the fluid contained in the reservoir section according to the profile of a single well. Its types include dry layer, water layer, oil-water layer and oil-gas layer [7-8].

The best method to identify oil and water layers in the work area is to use the cross plot method based on the correlation between parameters and two discriminant vectors of logging multi parameters. It is mainly to collect parameters that can reflect fluid properties more accurately and effectively and adopt reasonable mathematical classification criteria according to the characteristics of logging response curve and production test and development data.

Because the reservoir fluid properties are greatly affected by lithology, pore structure, etc., the resistivity of water layer and oil layer is mixed with each other, and there is no obvious boundary, so it is difficult to identify its oil and water layers using conventional methods. The correlation between lithology, physical property and oil bearing property of oil layer and water layer is used for fluid discrimination. In order to remove the changes in lithology, physical properties and bottom water layer with oilfield development or other geological reasons, which may cause certain deviation or impact on resistivity logging curves, and also to enhance the appearance of oil-bearing characteristics of the reservoir, provide the coefficient derived from the increase of resistivity to judge the fluid properties. According to its sedimentary characteristics, reservoir physical properties, oil testing and production conditions, its oil and gas filling intensity can be inferred. The formation water resistivity in the study area can be calculated by comprehensively using the potential logging curve and the apparent formation water resistivity in each layer to make qualitative and quantitative discrimination of reservoir fluid [9-10].

Because the internal structure of the reservoir is divided into two parts from top to bottom, one is the airway, and the other is the fracture. Here, we mainly realize fluid identification by detecting the change of different side pressure. Reservoir porosity is one of the important parameters to characterize fluid properties. It is related to factors such as pressure difference and volume loss, and has relatively little influence on these factors. So here we use piezoelectric materials for research. In reservoir, the velocity of fluid is related to pressure distribution, velocity field strength and temperature. When the fluid enters, its energy is high. It has good dynamic performance when the pressure difference is large and there is a large amount of turbulence load [11-12].

2.2. AdaBoost Algorithm Principle and Analysis

The development of neural network enables us to process seismic data manually. In Daniudi gas field, there are many complex factors, so there are many parameters. And these large amounts of information are contained in many hidden layers, neurons, topology, external environment and other influence media. If all hidden layers can be taken into account, it shows that neural networks can make good use of this aspect. Boosting method makes the subsequent learning machine more focused on the problem of sample discrimination. Boosting method can enhance the generalization ability of a given algorithm. Secondly, this method may lead to unstable data performance [13-14]. Based on the feature analysis method of neural network, when the training sample data set appears, the model is adaptively learned through artificial neurons. Then output to each region according to the preprocessing results. The algorithm adopts unsupervised updating method. It is different for different information points, and there is a certain deviation, which makes it impossible to identify the required accurate pattern. However, when a large number of data sets exist at the same time and the frequency of the curve smaller than the minimum value is high, it will cause local extremely inaccurate phenomena, resulting in the rate of wrong prediction and correct decision-making.

SVM algorithm has the characteristics of high precision, stable learning performance and strong scalability. Therefore, here we will focus on the Adaboost classifier model and other reservoir fluid identification methods. First, the mathematical model is established and the preprocessing steps are proposed. Secondly, the experiment set of training language alignment is conducted. Feature selection and synthesis are performed again. This paper uses cluster analysis technology. AdaBoost is an iterative algorithm that is implemented by changing the data distribution. The AdaBoost classifier can be used to place the key on the key training data [15-16]. The boosting algorithm flow is shown in Figure 1:

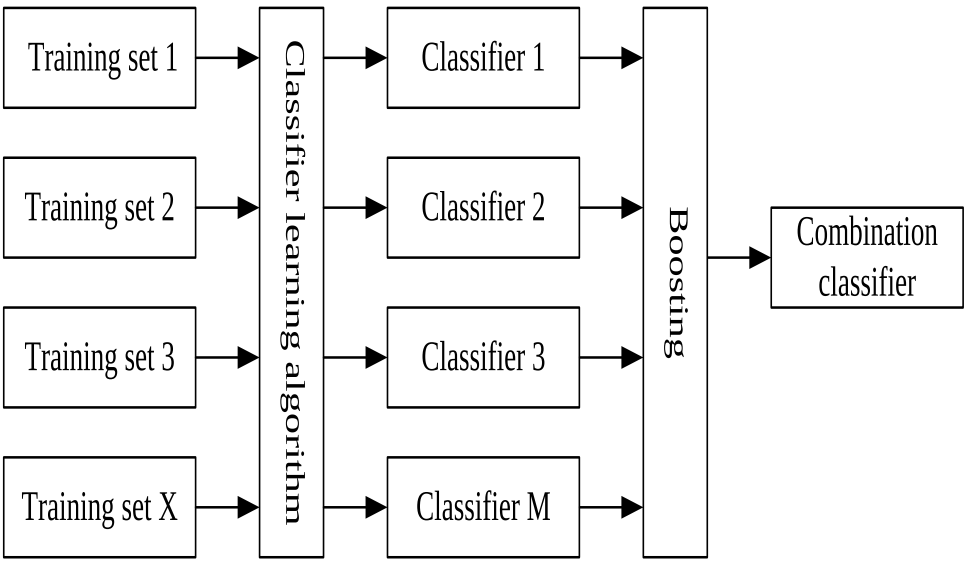

Figure 1. Boosting algorithm process structure

AdaBoost algorithm is an adjusted Boosting algorithm, which can adaptively adjust the errors of weak classifiers obtained from weak learning. The maximum training error of the final classifier generated is:

$$\prod_{s}\left[1\sqrt{\sigma_s(1-\sigma_s)}\right]=\prod_{s}\sqrt{1-4m_s^2}\leq f \tag{1}$$

Among them, σ_s is the training error of g_s. Assuming the number of samples in the training set is x, the dimension of the weak classifier is c, and the number of iterations is s, the maximum generalization error is:

$$\tilde{q}_s(G(m)\neq n)+\tilde{\mu}(\sqrt{\frac{sc}{x}}) \tag{2}$$

If the number of training iterations is too many, it will lead to the phenomenon of over adaptation or degradation. Therefore, the number of iterations must be properly selected.

In AdaBoost method, the information of weak classifier error limit is no longer needed, which undoubtedly has important theoretical significance and practical value. Due to its solid theoretical basis, accuracy of prediction and simplicity of calculation, Adaboost has also been applied in the field of intelligent identification in gas fields. In this paper, a mathematical model is established to identify the training sample data. After construction, we need to combine the eigenvalues and eigenvectors to form a new weight function. According to this implementation, important information such as all element parameters and weights in the whole link is extracted and the corresponding result output is generated. Then we can use this process to complete the input to the specific area contained in the neural network. Reservoir fluid recognition based on AdaBoost is realized by digitizing the input image and feature extraction. The unsupervised learning algorithm (such as neural network) is mainly aimed at the detection system with less damage or insensitive parameters [17-18].

2.3. Daniudi Gas Field

Daniudi Gas Field is located in the north of Ordos Basin. Most reservoirs in the work area are tight and low permeability reservoirs, and these reservoirs have strong heterogeneity. Due to the influence of these characteristics in physical properties, the distribution of gas reservoirs is relatively dispersed. It increases the difficulty of reservoir prediction and improves the requirements of reservoir prediction. Seismic inversion technology is the core technology of reservoir prediction, and high-precision inversion technology becomes the key of reservoir prediction.

In order to make full use of the advantages of logging data in vertical high resolution and seismic data in horizontal continuity, the post stack geostatistics well seismic joint inversion method is used to make the inversion result more accurate, which not only obtains a higher resolution, but also clearly reflects the continuous interface information in seismic data. Then, through comparative analysis with logging drilling data, the macro and micro combination is made, The distribution law and change characteristics of underground rock strata are more finely displayed. This inversion method has a good prediction effect for reservoirs with complex geological conditions. In this paper, a new inversion method is formed by improving the inversion method, which greatly improves the prediction ability of the reservoir and makes it suitable for the prediction of tight and low permeability reservoirs in the work area.

3. Identification of Castagna and Smith Oil-Bearing Sandstone Models

3.1. Model Identification Data

Using 20 data from 30 data sets measured globally released by Castagna and Smith, do two experiments and analyze them. The three elastic parameters of the model are used as the input set of the artificial neural network to identify the fluid properties. After the quantitative intersection technology is applied and the intersection diagram is analyzed, the sensitivity parameters are selected as the input set of the neural network to identify the fluid properties. Finally, the recognition results are compared.

3.2. Velocity and Density Set Input Neural Network Fluid Identification

The first 10 samples of Castagna and Smith models are used as the input of training sets. After network training, the last 12 sample sets are used as unknown properties for neural network recognition. During training, it is assumed that the expected output of mudstone is 001. The expected output of gas reservoir is: 111. The expected output of the water layer is: 011. The minimum error is 0.001 and the learning rate is 0.75.

3.3. Fluid Identification of Sensitive Parameter Set Input Neural Network

Cross plots with different parameters are made, and after the cross plots are analyzed, the highly sensitive elastic parameters are selected. After the calculation of oil-bearing sandstone model established by Castagna and Smith, the shear wave impedance, elastic impedance at 30° incidence angle and wave velocity ratio fluid identification coefficient are the highest. The selected high sensitivity parameters are used as neural network inputs. Comparison and analysis are directly input by compressional wave velocity, shear wave velocity and density, which are different in the process of fluid identification.

4. Analysis of Identification Results

4.1. Statistical Comparison of Mudstone Error

The Castagna and Smith models are used for fluid identification, and the results of density, velocity and sensitive parameters input into the neural network are analyzed and compared. It can be seen that the program compiled by the neural network has a good effect on model identification. The specific error of mudstone is shown in Table 1:

Table 1. Statistical comparison of mudstone errors

	Speed density	Sensitive parameters	Sensitive parameters
1	1.4	1.1	1
2	1.3	1.1	1
3	0.1	1	1
4	1.2	1.8	1
5	2.4	1	1

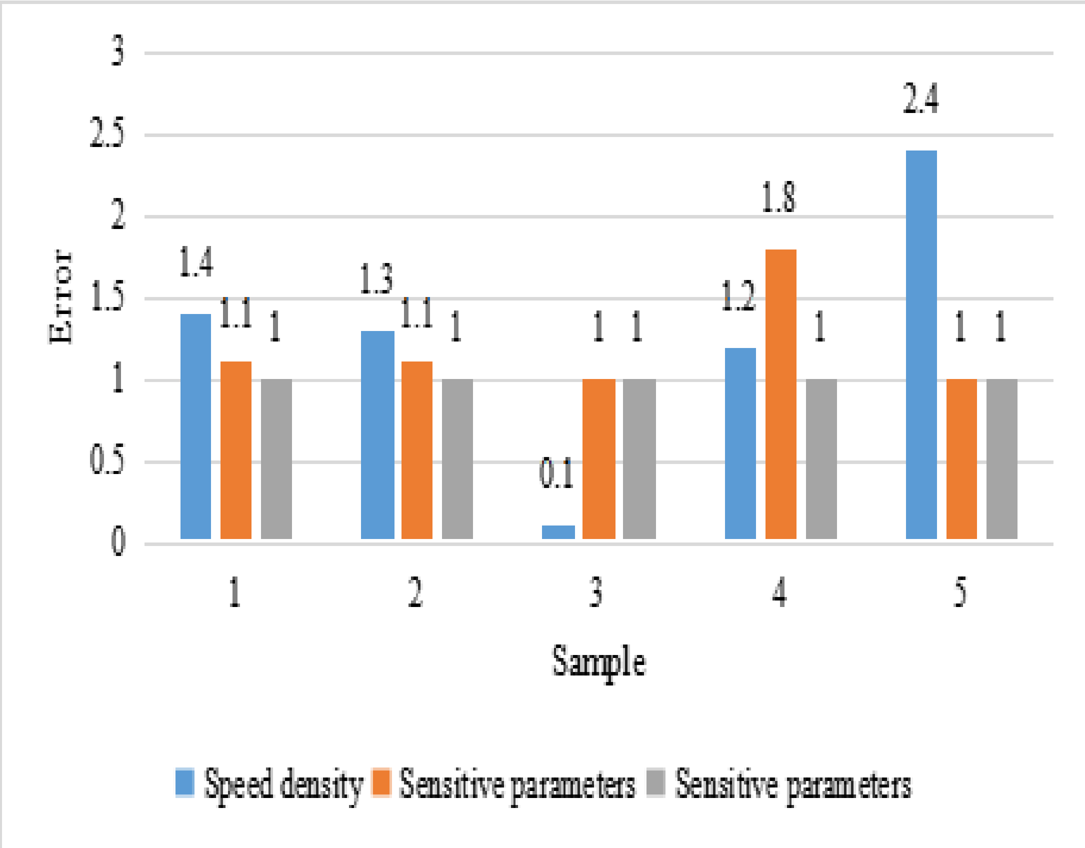

Figure 2. Statistical comparison of mudstone errors

As shown in Figure 2, we can see that the maximum error obtained by using velocity density as the input set is 2.4, and the maximum error obtained by using sensitive parameters as the input set is 1.8. The reference values are all based on 1. There are different data for different samples.

4.2. Statistical Comparison of Gas Reservoir Error

The error statistics are defined as the ratio of the actual output to the expected output, and then the error statistics of the two experiments are calculated respectively. Through the exploration of the experimental results of different samples, the gas reservoir error in Table 2 is obtained:

Table 2. Statistical comparison of air layer error

	Speed density	Sensitive parameters	Sensitive parameters
1	0.98	0.99	1
2	0.7	0.99	1
3	0.99	0.8	1
4	0.18	0.99	1
5	1	1	1

As shown in Figure 3, we can see that the effect of fluid identification with sensitive parameter set as input is slightly better, and its error statistics are closer to the reference value 1, while the effect of fluid identification with velocity density set as input is slightly worse. Therefore, when identifying other unknown layer fluids, the parameters with high sensitivity are selected through quantitative intersection technology, and the sensitive parameter set is used as input, which not only has small error, but also has high accuracy.

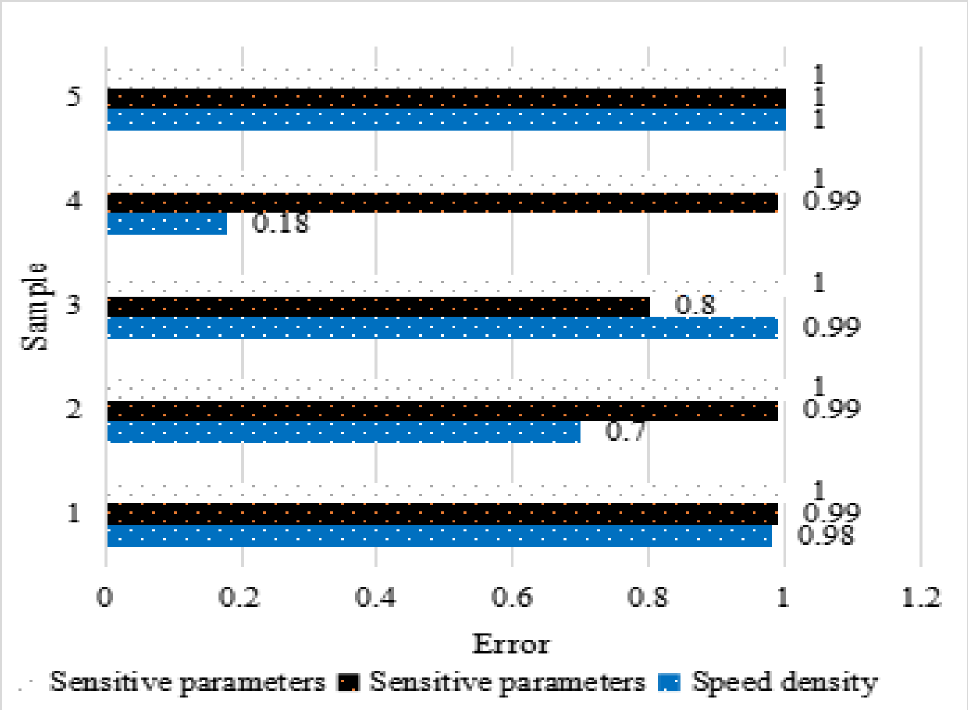

Figure 3. Statistical comparison of air layer error

5. Conclusion

Through the study of neural networks, we can draw relevant conclusions. Based on finite fuzzy theory, better recognition results can be obtained when the model training reaches a certain level. However, satisfactory results cannot be achieved in practical application due to various reasons. Therefore, this paper decided to take Daniudi gas field as the research object to conduct data preprocessing and intelligent identification in order to establish an early warning evaluation system. This paper mainly aims at the gas characteristic value and energy distribution in the region to determine the corresponding threshold to achieve early warning evaluation indicators. The experiment proves that the method can effectively reduce the error rate of detection.

Funding

This article is not supported by any foundation.

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Dohyun Lee, Kyoungok Kim: Improved noise-filtering algorithm for AdaBoost Using the Inter-and Intra-Class Variability of Imbalanced Datasets. J. Intell. Fuzzy Syst. 43(4): 5035-5051 (2020). https://doi.org/10.3233/JIFS-213244

[2] Krithiga R, E. Ilavarasan: Hyperparameter Tuning of AdaBoost Algorithm for Social Spammer Identification. Int. J. Pervasive Comput. Commun. 17(5): 462-482 (2020). https://doi.org/10.1108/IJPCC-09-2020-0130

[3] L. T. Akin Sherly, T. Jaya: An Efficient Indoor Scene Character Recognition Using Bayesian Interactive Search Algorithm-Based Adaboost-CNN Classifier. Neural Comput. Appl. 33(22): 15345-15356 (2020).

[4] Mehrnaz Mazini, Babak Shirazi, Iraj Mahdavi: Anomaly Network-Based Intrusion Detection System Using a Reliable Hybrid Artificial Bee Colony and AdaBoost Algorithms. J. King Saud Univ. Comput. Inf. Sci. 31(4): 541-553 (2019). https://doi.org/10.1016/j.jksuci.2018.03.011

[5] Ashraf Ullah, Nadeem Javaid, Muhammad Asif, Muhammad Umar Javed, Adamu Sani Yahaya: AlexNet, AdaBoost and Artificial Bee Colony Based Hybrid Model for Electricity Theft Detection in Smart Grids. IEEE Access 10: 18681-18694 (2020).

[6] Zohreh Bostanian, Reza Boostani, Malihe Sabeti, Mokhtar Mohammadi: ORBoost: An Orthogonal AdaBoost. Intell. Data Anal. 26(3): 805-818 (2020). https://doi.org/10.3233/IDA-205705

[7] Alpesh M. Patel, Anil Suthar: AdaBoosted Extra Trees Classifier for Object-Based Multispectral Image Classification of Urban Fringe Area. Int. J. Image Graph. 22(3): 2140006:1-2140006:16 (2020).

[8] Serap Aydin: Cross-validated Adaboost Classification of Emotion Regulation Strategies Identified by Spectral Coherence in Resting-State. Neuroinformatics 20(3): 627-639 (2020).

[9] K. M. Padmapriya, B. Anandhi, M. Vijayakumar: MapReduce fuzzy C-means Ensemble Clustering with Gentle Adaboost for Big Data Analytics. Int. J. Bus. Intell. Data Min. 19(2): 170-188 (2020).

[10] Noha A. Hikal, Mahmoud Y. Shams, Hanaa Salem, Marwa Metwally Eid: Detection of black-hole attacks in MANET using adaboost support vector machine. J. Intell. Fuzzy Syst. 41(1): 669-682 (2020). https://doi.org/10.3233/JIFS-202471

[11] Harmandeep Kaur, Munish Kumar: On the Recognition of Offline Handwritten Word Using Holistic Approach and AdaBoost Methodology. Multim. Tools Appl. 80(7): 11155-11175 (2020). https://doi.org/10.1007/s11042-020-10297-7

[12] E. Suganya, C. Rajan: An Adaboost-Modified Classifier Using Particle Swarm Optimization and Stochastic Diffusion Search in Wireless IoT networks. Wirel. Networks 27(4): 2287-2299 (2020). https://doi.org/10.1007/s11276-020-02504-y

[13] Bum-Soo Kim, Sanghyun Seo: Intelligent Digital Human Agent Service With Deep Learning Based-Face Recognition. IEEE Access 10: 72794-72805 (2020).

[14] Merlin Linda George, N. V. S. Sree Rathna Lakshmi, Senthil Murugan Nagarajan, Rajendra Prasad Mahapatra, V. Muthukumaran, M. Sivaram: Intelligent Recognition System for Viewpoint Variations on Gait and Speech Using CNN-CapsNet. Int. J. Intell. Comput. Cybern. 15(3): 363-382 (2020).

[15] Aaron Rasheed Rababaah: A Novel Skin-Inspired Model For Intelligent Object Recognition in Sensor Networks. Int. J. Sens. Networks 39(2): 93-105 (2020).

[16] Nilanjan Banerjee, Samarjeet Borah, Nilambar Sethi: Intelligent Stuttering Speech Recognition: A Succinct Review. Multim. Tools Appl. 81(17): 24145-24166 (2020).

[17] Rajiv Kapoor, Rohini Goel, Avinash Sharma: An Intelligent Railway Surveillance Framework Based on Recognition of Object and Railway Track Using Deep Learning. Multim. Tools Appl. 81(15): 21083-21109 (2020).

[18] Motrenko, Egor Simchuk, Renat Khairullin, Andrey Inyakin, Daniil O. Kashirin, Vadim V. Strijov: Continuous physical Activity Recognition for Intelligent Labour Monitoring. Multim. Tools Appl. 81(4): 4877-4895 (2020).