International Journal of Database and Theory and Application
Vol.9, No.2 (2016), pp.215-226
http://dx.doi.org/10.14257/ijdta.2016.9.2.23

Markov Prediction based on Semi-supervised Kernel Fuzzy Clustering

Li Yi-ran[1] , Zhang Chun-na[2],Guo Sheng-xing[3]

1. College of Applied Technology , University of Science and Technology Liaoning, Anshan Liaoning 114011, China
2. School of Software, University of Science and Technology Liaoning, Anshan Liaoning 114051, China
3. Angang Steel Company Limited, Anshan Liaoning 114021, China
E-mail：lyr7879@163.com

Abstract

This paper proposes an improved Markov prediction, according to the temperature control problem of hot blast stove alternate supply air in the operation of blast furnace operation, namely, implement clustering for waiting to be processed data, used kernel fuzzy c-means clustering based on the pairwise constraints. The supply air temperature of hot blast stove is seen as without aftereffect things in this method, introduce semi-supervised learning mechanism in traditional fuzzy clustering to deal with the basic data, and using the kernel effectiveness index improved the FCM algorithm. Experiments show that the improved clustering algorithm is superior to other algorithms in accuracy and performance, at the same time, the improved prediction model comparison with the traditional values of temperature prediction, which has obvious advantages in defined temperature range and the fit of the temperature value, the guiding significance was significantly enhanced in industrial field.

***Keywords**: hot blast stove; pairwise constraints; FCM; Markov; kernel function*

1. Introduction

In recent years, machine learning and semi-supervised clustering algorithm in data mining got the attention of many scholars, especially in industrial applications, for the vague and incomplete data that can't be directly referenced in the industrial scene, often spend a lot of manpower and time cost[1,2]. Semi-supervised clustering used the prior knowledge implied the data to integrate these unlabeled data in order to realize the clustering.

The hot blast stove is very important status in blast furnace ironmaking, its main function is that the heat generated by gas combustion blast to the blast furnace. In order to ensure the blast furnace obtains continuous high temperature hot blast, need several hot blast stove alternate supply air. Each technological process of hot blast stove includes two stages: heating and supply air. The former burns the gas mixed with air in the combustion chamber to heat the refractory brick. The waste gas is gradually reduced and the heat is stored in the refractory brick in heating, when the vault temperature reaches a certain value, the hot blast stove supply air. Using the differential pressure between the hot blast stove and the blast furnace sent the cold air after being heated into the blast furnace in the process of the supply air, the size of supply air through the cold air valve to regulate, after supply air is sent, the hot blast stove continues to heat, followed by reciprocating. For the blast furnace, the furnace temperature control in burning furnace period and uninterrupted supply air in supply period is the key of normal operation. Here, the accurate prediction of the supply air temperature is focus of this study.

With the deepening of research about the supply air temperature of the hot blast stove, people have accumulated lots of experience and data, this information has a great

ISSN: 1975-0080 IJDTA

influence on the temperature control of the hot blast stove[3]. Some forecasting methods are used to process the information in order to guide operation in the production. When the operating parameters of the blast furnace change, the prediction model of the hot blast stove to predict supply air temperature, then get coal injection rate, is used to evaluate the performance of the hot blast stove. Markov prediction method is adopted in this paper, its idea is: whether it is the social field or natural science, the change of a certain kind of thing is only related to the recent state, and has nothing to do with the past state of things, the focus of this paper is based on that[4]. It should be noted that the complexity of the data collected, data come from industrial field, there is a large number of unlabeled data, and the marked data is limited, that data cannot be effective support to predict, therefore, this article puts forward a kind of semi-supervised kernel fuzzy clustering algorithm based on pairwise constraints to deal with huge amounts of data, the main research direction is improved clustering algorithm, introducing pairwise constraints to deal with excessive unlabeled data, the kernel function is used to analyze the processing of a large amount of data, in this way, the source of the forecast data conforms to the requirements, in order to improve the accuracy of Markov prediction.

2. Mathematical Model of Dome Temperature

2.1. The Mathematical Model of Heat Storage

Burning period of hot blast stove, the heat transfer through the heating surface of the vault in unit time as follows:

$$\Delta Q_1 = S_1 k_1 \Delta T \Delta \mathrm{t} = S_1 K_1 (T_0\text{-}T_1)\Delta \mathrm{t} \tag{1}$$

In the formula, S_1 is heat transfer coefficient of the vault and the flue gas, k_1 is heat transfer area of the vault and the flue gas, T_0 and T_1 are respectively temperature of the flue gas and the vault, ΔT is the temperature difference between the flue gas and the vault, $\Delta \mathrm{t}$ is quantity of unit time.

Meanwhile, the vault is lost heat to the outside world, the loss heat to the outside world through the vault in unit time as follows:

$$\Delta Q_2 = S_2 k_2 \Delta T \Delta \mathrm{t} = S_2 K_2 (T_1\text{-}T_2)\Delta \mathrm{t} \tag{2}$$

In the formula, S_2 is heat transfer coefficient of the vault and the outside world, k_2 is heat transfer area of the vault and the outside world, T_1 and T_2 are respectively temperature of the outside world and the vault.

Therefore, the change of the vault's heat storage as follows:

$$\Delta \mathrm{Q}=\Delta Q_1\text{-}\Delta Q_2 \tag{3}$$

2.2. Optimization of Air Fuel Ratio Mathematical Model

Fast burning furnace method is commonly used in the enhanced combustion period, which requires the optimal air-fuel ratio control, which is the guarantee of obtaining the maximum heat. But no matter what circumstance can't is too much gas, in other words, the air is still excessive when it reaches the optimum air-fuel ratio. It is generally believed that the burning state of the hot blast stove is good when the volume fraction of residual oxygen is maintained at 0.2%~0.8% in flue gas. In this paper select 0.5% as the control target, and 0.2% as the acceptable fluctuation range. That is, in the range of 0.3%~0.7%, the control is not regulated.

When the volume fraction of the residual oxygen in the flue gas is greater than the maximum, the following air-fuel ratio adjustment model is launched:

$$\mathrm{b_k} = \frac{\mathrm{v_k} - \dfrac{(x-0.5)v_y}{0.21}}{v_m} \tag{4}$$

When the oxygen volume fraction in the flue gas is lower than the minimum value, the air-fuel ratio is changed according to the formula (5):

$$\mathrm{b_k} = \frac{\mathrm{v_k} + \dfrac{(0.5-x)v_y}{0.21}}{v_m} \tag{5}$$

In the above formula, $\mathrm{b_k}$ is the air-fuel ratio. x is the oxygen content in the exhaust gas(volume fraction,%). v_y is the amount of the flue gas, m^3/h. v_m is the amount of the gas, m^3/h. $\mathrm{v_k}$ is the amount of the free-air, m^3/h.

3. Markov Operation Prediction

3.1. Basic Principle

Markov prediction method is a kind of highly effective and scientific forecasting method, which is proposed by Markov, a famous Russian mathematician. Basic idea is: for the things in the real world, including natural and social areas, its change is only referring to the recent changes of the thing, and its past status has not been greatly related[5,6]. Therefore, the largest feature of this method is no aftereffect, prediction don't need a large amount of historical data to support, that is, the development of things is only related to the current state, compared to the previous state has nothing to do. At present, the more popular methods include: regression analysis, time series, neural network, grey forecasting, etc., these methods often require a large amount of historical data, it is clear that the Markov prediction method has a stronger adaptability, it mainly makes use of the transfer probability matrix between the state to predict the development and trend of the things.

3.2. Markov and Supply Air tTemperature

With the development of hot blast stove technology, people get a lot of data about supply air temperature, these empirical data are undoubtedly very useful, however, the certain information exist partially missing and incomplete, and cause not reasonably use. Here, the introduction of prediction method is very good to solve the problem. The prediction model of the temperature control in the hot blast stove is a kind of method to predict the change of temperature, calculate the air-fuel ratio, and evaluate current operation when the operation parameters are changed.

The supply air temperature of the hot blast stove recorded as s, the initial state is s_0, through change process: $s_1, s_2, s_3 ..., s_n$. Supply air temperature around the engagement, the final state s_n only related with the previous state s_{n-1}, that is to say, the impact of a change only depends on the last change, namely Markov's no aftereffect. In this, for the change of the supply air temperature of the hot blast stove based on the following two considerations: first, in the process of the evolution of things, the changes are random; second, the structure of things every alteration has no aftereffect of state transitions. The change process of thing s can be regarded as the Markov process. Markov is feasible in theory. Here, there is a problem worthy of consideration, the quality of the Markov prediction results depend on the quality of the source data, high quality temperature control data will significantly increase the accuracy of the prediction. Therefore, this paper intends to introduce clustering algorithm to divide the data of waiting processing, to respond to the complicated temperature state, improve the accuracy of prediction.

4. Kernel Fuzzy Clustering Algorithm based on Pairwise Constraints

4.1. Clustering Performance Index

The performance index is mainly used to investigate the number of clusters, reasonable the number of clusters to make the algorithm more effective, at the same time, it can also cause the distortion of the results if the algorithm select improper for the different data. Therefore, we must find a practical and effective performance function to support clustering.

At present, the effective performance function can be roughly divided into three categories: geometry based on the data set, statistical information based on the data set and fuzzy partition based on the data set[7]. With the increasing amount of information, and its structure are becoming more and more complex, so the corresponding effective performance function is introduced, the effective function of fuzzy c-means clustering algorithm is as follows:

$$V_B = -\frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\ln(u_{ij}) \tag{6}$$

$$V_{KS} = \sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}^{m}\left\|x_j - x_i\right\|^2 - \sum_{j=1}^{n}\sum_{i=1}^{c} u_{ij}^{m}\left\|v_j - \bar{v}\right\|^2 \tag{7}$$

$$V_{KB} = \sum_{i=1}^{c}\left(\frac{\sum_{j=1}^{n} u_{ij}\left(1-K\left(V_i, X_j\right)\right)}{n\sum_{i=1}^{c}\left(1-K\left(V_i, V_j\right)\right)}\right) \tag{8}$$

These three kinds of the effective performance function are respectively proposed by Bezdek、Kayama&Sugeno、Bensald. In the formula, u_{ij} is the membership degree of the j-th data points to the i-th cluster center, the minimum value of the function is the optimal number of clusters.

According to the complicated situation of the industrial scene, and combined with the characteristics of clustering algorithm, this paper intends to use the formula (8), that is, the kernel V_{KB} as an effective function, it can not only solve the problem of setting up the optimal number of clusters, but also because of the upper part of the index is to investigate the compactness of the class, so that the algorithm is not sensitive to the outlier data.

4.2. Fuzzy c-means Clustering Algorithm

Fuzzy c-means clustering algorithm is a common method of data classification[8], according to the principle of Lagrange's least square method, repeated update clustering center and elements in the classification matrix in the iteration, meet the conditions of membership, so that the variance within the cluster is minimized [9-12]. This paper uses the kernel c-means algorithm to cluster analysis.

There are data sets $X = \{x_1, x_2, ..., x_n \mid x_i \in S\}$ in Q dimensional space S, using fuzzy matrix u, the vector is classified into m class $m \in (1, n)$ the clustering center of the initial state is $Z = \{Z_1, Z_2, ..., Z_m\}$ the objective function is defined as follows:

$$J(u, Z) = \sum_{j}\sum_{i} u_{ji}^{r} D_{x_i, Z_j}^{2} \tag{9}$$

In the formula, u_{ji}^r is the membership, that is the i -th data belongs to the membership of the j -th class; $D_{xi,Zj}^2$ is Euclidean distance, that is the distance from the i -th data to the center of the j -th class; r is the weight factor, examine the value of fuzzy degree in the fuzzy matrix $\in [1,3]$.

In the other, the fuzzy matrix element u_{jr}^r should be satisfied in the type (9):

$$\sum_{j=1}^{m} u_{ji}^r = 1, \quad \sum_{i=1}^{m} u_{ji}^r > 1$$

(10)

Among them, $u^r{}_{ji} \in [0,1]$, $0 \le i \le n$, and $0 \le \sum_i u_{ji}^r \le n$.

In the extreme case, u_{ji}^r and $D^2_{x_i,Z_j}$ can get the optimal value when $J(u,Z)$ get the minimum value. Here, according to Lagrange's least squares method, adjust the u、Z in $J(u,Z)$ can refer to the following formula:

$$u_{ji} = \frac{1}{\sum_{k=1}^{m} (\frac{D_{x_i,Z_j}}{D_{x_i,Z_i}}) 2(r-1)^{-1}}$$

(11)

$$Z_j = \frac{\sum_{i=1}^{n} (u_{ji})^r x_i}{\sum_{i=1}^{n} (u_{ji})^r}$$

(12)

Here, if $\| x_i - Z_j \| = 0$, then $u_{ji} = 1$.

Formula (11), (12) continuous iteration, and adjust the value of u_{ji} and Z_j; the clustering results are obtained when $J(u,Z)$ convergence is satisfied.

4.3. FCM Based on the Kernel

The element in the cluster is defined as a three tuple $M(P,E,C)$, The three elements in the formula respectively represent: the vertex set, the edge set, and the weight matrix, and the Gauss kernel function is used to express it, the following is the definition of Gauss kernel:

$$\overline{k_{ij}} = \begin{cases} \exp(-\frac{\| x_i - x_j \|^2}{\sigma^2}) & i \ne j \\ 0 & i = j \end{cases} \quad (13)$$

In the formula, σ is a regulator, which is used to control the decay of the kernel. The row vector of $\overline{K_{ij}}$ will be distributed in the super sphere in K dimension space, the corresponding diagonal matrix is:

$$A_{ij} = \sum_{j=1}^{n} \overline{K}_{ij} \tag{14}$$

Then $\overline{K}_{ij}$ is converted to:

$$K_{ij} = \frac{\overline{K}_{ij}}{\sqrt{A_i A_j}} \tag{15}$$

Processing the matrix K, which extracts in front of the k maximum eigenvalue corresponding to the feature vector, the matrix N is obtained by the transformation:

$$N_{ij} = \frac{K'}{\sqrt{\sum_j K_{ij}^{'2}}} \tag{16}$$

Each row vector in the matrix N can be regarded as the data points in the sample space, and the FCM is used to implement the clustering.

Using kernel function, the sample x_k in Q dimension space is mapped to high dimensional space Q', at the same time, the topological structure of the original spatial data is not changed, cluster center is represented by the function $\psi(x_k)$, then the corresponding clustering function can be expressed as:

$$\begin{aligned} J &= \sum_j \sum_i u_{ji}^r \| \psi(x_i) - \psi(Z_k) \|^2 \\ &= \sum_j \sum_i u_{ji}^r (K(x_i, x_j) - 2K(x_i, Z_j) + K(Z_j, Z_i)) \end{aligned} \tag{17}$$

In the formula, Z_k is the clustering center of the k-th class, and $\psi(Z_k)$ is its image in the kernel space, can be expressed as:

$$\psi(Z_i) = \frac{\sum_{i=1}^{n} u_{ji}^r \psi(x_k)}{\sum_{i=1}^{n} u_{ji}^r} \tag{18}$$

In high dimensional space Q', the transform of the corresponding membership as follows:

$$u_{ji} = \frac{1}{\sum_{k=1}^{m} (\frac{T_{ij}}{T_{ik}})^{\frac{1}{r-1}}} \tag{19}$$

In the formula, $T_{ij} = K(x_i, x_j) - 2K(x_i, Z_j) + K(Z_j, Z_i)$, $K(x_i, x_j)$ is kernel function.

4.4. Target Function Adjustment

In unsupervised learning algorithm, such as image recognition, the method of pairwise constraints is the most reasonable, complex things are not easy to determine the category, but it is not difficult to determine whether the data has similar constraint information. Get constraint information from the guide of Must-link and Cannot-link[13-15]. Must-link agreement, two data samples of the space belong Must-link constraints, both will fall into a category; two data samples of the space belong Cannot-link constraints, the two is divided into different classes.

Set Must-link constraint set is S_m. That is $S_m = \{x_i, x_j\}$, sample data x_i, x_j in the same class; Cannot-link constraint set is S_c ,referred to as $\{x_m, x_n\}$, sample data x_m, x_n in different classes. After introduce the pairwise constraints, the FCM objective function based on kernel adjusted as follows:

$$J = \sum_j \sum_i u_{ij}^r \left\| \psi(x_i) - \psi(z_k) \right\|^2 + \beta \left(\sum_{i=1}^{n} \left(\sum_{(x_i, x_j) \in S_m} \sum_{k=1}^{c} u_{ik} u_{jk} + \sum_{(x_i, x_j) \in S_c} \sum_{k=1}^{c} u_{ik} u_{jk} \right) \right)$$

(20)

$$\beta = \frac{n}{m} \frac{\sum_{k=1}^{c} \sum_{i=1}^{n} u_{ik}^2 d^2(x_i, Z_k)}{\sum_{k=1}^{c} \sum_{i=1}^{n} u_{ik}^2} \tag{21}$$

Among them, add item is a penalty term of violate constraints. This is the degree of violate Must-link and Cannot-link. The target function is adjusted by the membership, and keep small amount of constraint violations, so that the objective function value is minimized. The function of constraint information in the process of clustering is obvious, and β is the regulator, which is used to reflect the influence of violate constraints in clustering, that is the important degree of constraint term in the objective function, the main use of the normalized performance index to adjust the constraint term, the value of β was smaller when the degree of Angelica was better, on the contrary, the value is relatively large when the level of the angelica is poor, usually values are $[0,1]$.

Algorithm description:

Dataset $X = \{x_1, x_2, ..., x_n\}$, number of clusters is c, adjustment factor is β, kernel regulatory factor is b, termination threshold is ϕ, the maximum number of iterations is $\max count$, counter $icount = 0$.

Step 1: add a pair constraint information, construct D;

Step 2: construct the similarity matrix $\overline{k_{ij}}$, get in front of the k-th maximum eigenvalue corresponding feature vector;

Step 3: initial membership;

Step 4: calculate the clustering center;

Step: 5: calculate the objective function according to the formula (20);

Step 6: recalculate membership;

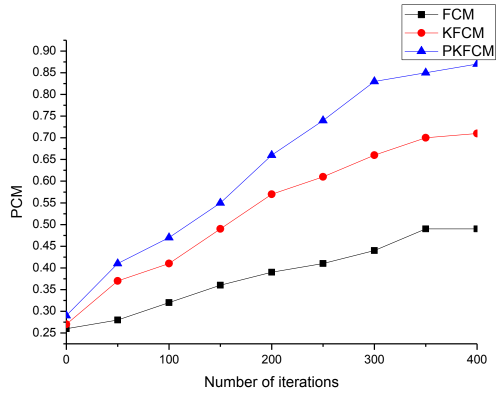
Step 7: update the kernel parameter;

Step 8: determine whether the current iteration reaches the maximum number of iterations, or to meet the termination threshold conditions, if it is then terminated, otherwise return to Step 4.

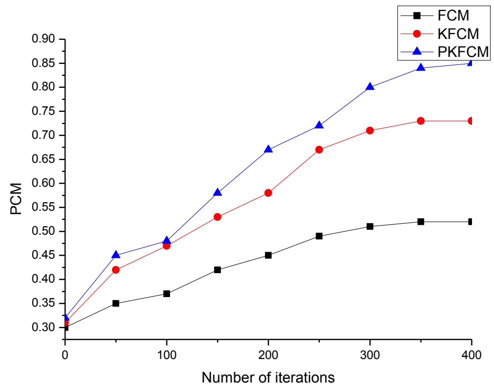
5. Experimental Analysis

Experiment is divided into two parts, the algorithm performance index analysis and Markov prediction results analysis. The data of the former comes from universal database UCI, plan to select four data sets Balance,Wine,Same and Similar, Balance is a low dimensional data set, the dimension is 4, which can be divided into 3 categories, the number samples of per class is 625; Wine is also a low dimensional data set, the dimension is 13, which can be divided into 3 categories, the number samples of per class is 180; Same and Similar are high dimensional data sets, the dimension is 16090, the sample number is the same as 3. The clustering effect and the convergence of the algorithm are verified. For the analysis of clustering effect, the paper adopts the pairwise comprehensive measure (PCM), the higher the value, the better the clustering effect. Convergence mainly studied the stability of the algorithm, with the algorithm is carried out, the more stable the value, the better. In markov prediction analysis, the paper intends to select history data of the average temperature of the storage room, 16 times at the end of the supply air before the next prediction at a factory, in order to obtain the predicted value of the average temperature of the regenerative chamber at the end of the next supply air.

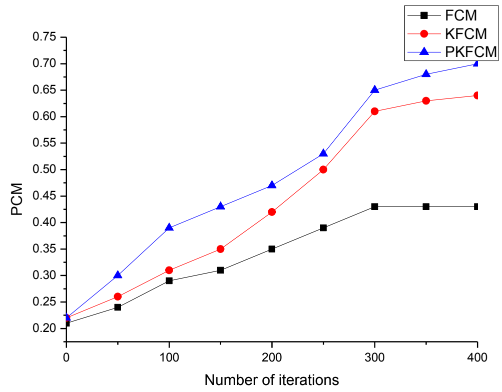
In comparative analysis of clustering algorithms, three algorithms are used in this paper, which are the traditional fuzzy c-means algorithm, denoted as FCM; fuzzy c - means algorithm based on kernel, referred to as KFCM; in this paper, a fuzzy c-means algorithm based on pairwise constraints is proposed, denoted as PKFCM. The following is the corresponding chart analysis.



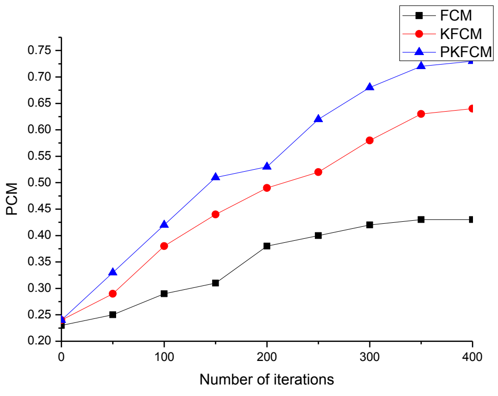
(a) Balance comparison results



(b) Wine comparison results



(c) Same comparison results



(d) Similar comparison results

Figure 1. Comparison of PCM Values in the Data Set

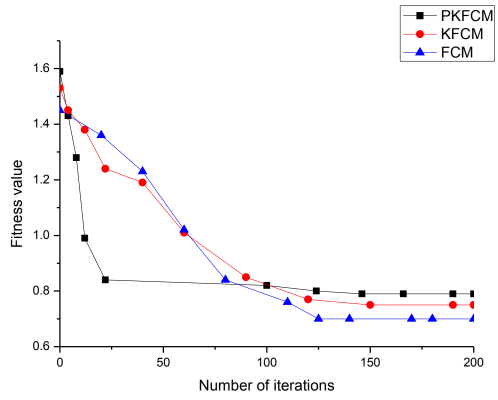


Figure 2. Comparison Results of the Algorithm Convergence

Table 1. Air Supply Record Sheet

Time	1:00	4:00	7:00	10:00	11:00	16:00	19:00	22:00
Temperature	1135	1143	1151	1144	1139	1142	1150	1146
Time	1:00	4:00	7:00	10:00	11:00	16:00	19:00	22:00
Temperature	1145	1144	1154	1142	1146	1138	1153	1145

Table 2. Temperature Status Table

Status	Temperature values	Interval
P1	1135、1139、1138	[1135,1140]
P2	1143、1144、1142、1146、1145、1144、1142、1146、1145	[1142,1146]
P3	1151、1150、1154、1153	[1150,1155]

Table 3. Predictive Results Comparison Table

Time	1:00	4:00	7:00	10:00
Traditional predictive value	1135	1143	1151	1144
Improved predictive value	[1130,1140]	[1140,1150]	[1150,1160]	[1140,1150]
The actual value	[1135,1140]	[1142,1146]	[1150,1155]	[1142,1146]

Figure 1 is a comparison of UCI data sets Balance, Wine, Same, and Similar in PCM values. Experimental results show that the initial state of the algorithm has no space for improvement when the number of iterations is 0, PCM values are basically the same, with the advance of the operation, the advantages of the improved algorithm is gradually reflected, the 4 data sets are reflected after the number of iterations exceeds 150, the rising slope of PKFCM is highlighted, two other algorithms are relatively stable, when the number of iterations reaches 300, the PCM and KPCM show a downward trend, while the PKFCM still maintains a steady rise, The PCM values of the three algorithms respectively is 0.44,0.66,0.83, it can be seen that PKFCM has obvious advantages, with the increase of the amount of information of pairwise constraints, the number of samples of correct clustering gradually increases, while the other two algorithms do not use pairwise constraint information, therefore, when the number of iterations to reach 350, the PCM value barely risen, while PKFCM continues to maintain a steady upward trend, improved results will be reflected.

Figure 2 is the comparison of the convergence of the three algorithms, the experimental results show that the improved PKFCM algorithm is superior to the other two algorithms both in convergence speed and stability in the late, and PKFCM tends to be stable in the middle of the algorithm run, here, based on the introduction of kernel function makes the algorithm enhanced the ability to deal with large data. At the same time, Semi-supervised clustering based on pairwise constraints has the ability to deal with some of the unlabeled data, compared with the other two algorithms, the improved algorithm is more powerful, so it has better convergence.

Table 3 is the result of the Markov prediction for the record of the supply air, the prediction values for the 5 moments are listed in the table, at the same time, compared with the corresponding traditional prediction value, traditional prediction values are derived from production experience, and divided into three intervals, the analysis showed that The temperature range obtained by the improved Markov prediction is more flexible and practical, it mainly depends on the initial prediction data, because the effective clustering is implemented in this paper for the data, which makes the data more standardized, so the predicted value is more close to the actual value.

6. Conclusion

According to the temperature control problem of hot blast stove supply air in industrial field, in this paper, starting from the processing of the data source, in order to obtain a reasonable basic data to support the Markov prediction. Because of the complexity of the temperature control, considering the c - means combined with Markov, at the same time in the clustering algorithm introducing the kernel function to enhance its ability to deal with large data, and make full use of the paired constraint information to make the improved PKFCM algorithm is significantly better than the unsupervised FCM algorithm in clustering effect. The experimental results also fully proved the improved effect of the algorithm, model prediction is closer to the actual value, better solved the problem that the predictive value can't direct production, it also provides a basis for the automatic combustion system control of the hot blast stove.

References

[1] S. Saha and S. Bandyopadhyay, "Semi-GAPS: A Semi-supervised Clustering Method Using Point Symmetry", Fundamenta Informaticae, vol. 96, no. 1-2, **(2009)**, pp. 195-209.
[2] C. Ruiz, M. Spiliopoulou and E. Menasalvas, "Density-based semi-supervised clustering", Data Mining and Knowledge Discovery, vol. 21, no. 3, pp. 345-370.
[3] K. R. Muske, J. W. Howse and G. A. Hansen, "Hot Blast Stove Process Model and Model-based Controller", Iron and Steel Engineer, vol. 76, no. 6, **(1999)**, pp. 56-62.
[4] A. M. S. Barreto and M. D. Fragoso, "Computing the Stationary Distribution of a Finite Markov Chain Through Stochastic Factorization", SIAM Journal on Matrix Analysis and Applications, vol. 32, no. 4, **(2011)**, pp. 1513-1523.
[5] L. Szilagyi, L. Medves and S. M. Szilagyi, "A modified Markov clustering approach to unsupervised classification of protein sequences", Neuro-computing, vol. 73, no. 13-15, **(2010)**, pp. 2332-2345.
[6] A. Sarkar, M. Nikolski and U. Maulik, "Spectral Clustering on Neighborhood Kernels with Modified Symmetry for Remote Homology Detection", Proceedings of the 2011 Second International Conference on Emerging Applications of Information Technology, **(2011)**, pp. 269-272.
[7] A. M. Bensaid, L. O. Hall, J. C. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington and R. F. Murtagh, "Validity-guided (re)clustering with applications to image segmentation", IEEE Transactions on Fuzzy Systems, vol. 4, no. 2, **(1996)**, pp. 112-123.
[8] B. DoğAn and M. KorüRek, "A new ECG beat clustering method based on kernelized fuzzy c-means and hybrid ant colony optimization for continuous domains", Applied Soft Computing, vol. 12, no. 11, **(2012)**, pp. 3442-3451.
[9] B. Sowmya and B. S. Rani, "Colour image segmentation using fuzzy clustering techniques and competitive neural network", Applied Soft Computing, vol. 11, no. 3, **(2011)**, pp. 3170-3178.
[10] K. L. Du, "Clustering : A neural network approach", Neural Networks, vol. 23, no. 1, **(2010)**, pp. 89-107.
[11] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm", IEEE Transactions on Image Processing, vol. 19, no. 5, **(2010)**, pp. 1328-1337.
[12] T. Geweniger, D. Zulke, B. Hammer and T. Villmann, "Median fuzzy c-means for clustering dissimilarity data", Neuro-computing, vol. 73, no. 7-9, **(2010)**, pp. 1109-1116.
[13] K. Wagstaff and C. Cardie, "Clustering with instance-level constraints", Proceedings of the Seventeenth International Conference on Machine Learning, **(2000)**, pp. 103-1110.
[14] C. Ruiz, M. Spiliopoulou and E. Menasalvas, "Density-based semi-supervised clustering", Data Mining and Knowledge Discovery, vol. 3, no. 21, **(2010)**, pp. 345-370.
[15] S. Faußer and F. Schwenker, "Semi-Supervised kernel clustering with sample-to-cluster weights", PSL'11 Proceedings of the First IAPR TC3 conference on Partially Supervised Learning, **(2011)**, pp. 72-81.

Authors

Y.R.LI received the Master's degree in computer application technology from University of Science and Technology Liaoning, in 2008. Currently, he is a lecturer at School of applied technology college at University of Science and Technology Liaoning. His research interests include Distributed computing and data mining.

C.N.ZHANG received the Master's degree in computer application technology from University of Science and Technology Liaoning, in 2007. Currently, she is a lecturer at School of Software Engineering at University of Science and Technology Liaoning. Her research interests include Distributed computing and data mining.

S.X.GUO received the Bachelor's degree in computer science and technology from Liaoning Shihua University, in 2001. Currently, he is An engineer at ANSTEEL GROUP CORPORATION. his research interests include Data Mining and Pattern Recognition.