



Social emotion classification of Japanese text information based on SVM and KNN

Qiang Zhao¹

Received: 4 December 2020 / Accepted: 2 March 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

How to model multi-modal data and rich context information under the same model framework has become the key to user sentiment analysis in social media. This research uses SVM and KNN to build models for Japanese social sentiment classification and proposes a topic model for user sentiment analysis based on SVM and KNN. Moreover, taking TE process data as an example, this paper selects radial basis kernel function and grid search method for the construction of SVM classifier. In addition, by introducing the correlation variable between the sentiment of the comment and the sentiment of the tweet, the comment is related to the original tweet. Based on the establishment of the model, this research proposes a parameter estimation algorithm for model solving based on the idea of SVM-KNN, and uses the Twitter real data set as the experimental data set to verify the effectiveness of the user sentiment analysis model proposed in this paper. The research results show that the method proposed in this paper has a certain effect.

Keywords Support vector machine · KNN · Japanese social · Emotion classification

1 Introduction

In the past 2 decades, artificial intelligence researchers have endeavored to give machines the ability to recognize and discern emotions in order to recognize and express emotions. This research is called sentiment analysis. Sentiment analysis has become a new trend in social media, which effectively helps users understand the opinions expressed on different platforms (Stefano and Gabriele 2020). In recent years, a large number of researchers have conducted sentiment research through text, and text sentiment analysis has made great progress. With the rise of video social networking apps, people are more inclined to use video rather than text to express opinions when expressing their opinions on products or services. In life, consumers tend to shoot videos to record their comments and opinions on products and upload the videos to social media platforms such as YouTube or Facebook to express their opinions and favorites to users. Some people think that browsing text comments on social media is not as easy to understand and convenient as

videos, and that long texts will take a lot of time for users. At the same time, it is difficult to find authentic and reliable text-based comments on social media, but it is easy to find reliable and authentic video comments. The voice data in the video expresses the tone of the speaker, while the visual data conveys facial expressions, which in turn helps to understand the emotional state of the user. The data obtained from the video can be used as a useful information source for sentiment analysis, but there are some major challenges that need to be resolved. For example, the expression of opinions varies greatly from person to person. Some people express their opinions euphemistically, some people express their opinions intuitively, and some people rely entirely on logic to express emotions. When a person changes his voice to express his opinion, the audio data usually contains most of the information that expresses his opinion. When a person expresses personal opinions through facial expressions, most of the data required for sentiment analysis can often be determined through facial expression analysis. Therefore, these personal differences prompt us to find a general sentiment analysis framework. At present, there are few researches on sentiment analysis of multimodality, but most of multimodality research only considers visual and audio information and ignores text information. On the other hand, how to effectively integrate different forms of features is a

✉ Qiang Zhao
zhaoqiang8208@163.com

¹ School of Foreign Languages, Dalian Polytechnic University, Dalian 116000, China

problem that needs to be solved. Therefore, by developing a multimodal framework, it is possible to cope with all three sensing modes in a human-centric environment: text, audio, and video. It enables people to communicate and express their emotions through different channels, and enables text, audio, and visual forms to be used simultaneously to effectively extract the semantic and emotional information conveyed in the dialogue (Arnold et al. 2011).

The features extracted from traditional emotion recognition methods are based on artificially designed features (Bourdeau et al. 2019). With the increasing popularity of large-scale data sets, deep learning has become a general method of machine learning. In many computer vision and natural language processing tasks, machine learning can get better results. Recently, 3D Convolutional Neural Networks (C3D) have made great progress in handling various video analysis tasks. C3D can model the appearance and motion information at the same time, and the C3D features of the linear classifier can achieve good performance on different video analysis benchmarks (Mary and Arockiam 2017). In the existing video-based emotion recognition task, few researchers use C3D network. Therefore, constructing a novel multi-modal fusion framework combined with C3D network can play a very important role in the development of social sentiment analysis.

This study uses SVM and KNN to build models for Japanese social sentiment classification, which lays the foundation for the further development of the translation system.

2 Related work

The traditional analysis of micro-blog text is mainly carried out independently from the content of the text or the subjective emotion of the text. For example, through text content, Weibo text can be divided into entertainment, society, education, sports and other categories, and through the subjective sentiment in the text, the Weibo text can be divided into two categories: positive and negative. The topic-emotion mixed model combines the two together, and analyzes the objective topics contained in the text and the corresponding subjective emotions. The literature Rabie et al. (2020) proposed an unsupervised topic emotion hybrid algorithm. The algorithm comprehensively considers the characteristics of Chinese microblogs, builds a five-level sentiment lexicon, and recognizes new words through rules and statistical methods. Moreover, the algorithm samples the sentiment labels of each sentence in the text, and the subject label samples of each word, and establishes the corresponding relationship, and finally extracts the theme and emotion elements. However, when the unrelated emotional words in the text are not included in the constructed emotional lexicon, the algorithm's recall rate will be affected. In order to be

able to use different annotation tree banks and sentiment lexicons at the same time, the literature Schuelke-Leech et al. (2015) proposed a mixed syntax processing method. Based on the undirected graph MST syntax analyzer, this method searches the dependent structure in the text through the maximum spanning tree and learns the edge weights through the training algorithm, which eliminates the need to re-develop the syntax analyzer after fusing multiple corpora. The literature Vimalkumar and Radhika (2017) proposed a new topic sentiment classification method. This method first divides the existing sentiment dictionary into core and ordinary levels. Then, this method uses the N-Gram algorithm to segment the words in the Weibo text, and calculates its feature value to determine the level of the emotion word and enhance it with different multiples. The literature Yu et al. (2018) has conducted research on the theme sentiment in product reviews. The literature Cui et al. (2017) proposed the ILDA model, extended PLSI and LDA, and combined with a variety of probabilistic graphical models and product review data to extract relevant product review topics and corresponding ratings. In order to solve the problem of sparseness in short texts, the literature Li et al. (2018) proposed a short-text sentiment-topic model (SSTM) for short texts. Aiming at the characteristics of short text, this model represents the entire corpus as a set of word pairs, which effectively improves the accuracy of topic sentiment classification. In view of the characteristics of social media data such as large volume, few words, and rapid spread, literature Mukherjee et al. (2017) proposed a text sentiment analysis technology based on keyword analysis. The technique first extracts key sentences from three features: location, keywords, and word frequency, and then proposes seven types of part-of-speech collocations that have a greater influence on emotional tendencies and gives corresponding calculation rules. Finally, the technique calculates sentiment values based on the weights of all keywords. In order to reduce the reliance on artificially defined rules, the literature Nasir and El-Ferik (2017) combined machine learning methods with Weibo topic sentiment analysis. Moreover, the literature used a variety of machine learning models to quantitatively evaluate the modeling effect and introduced emoticons into word features, making the analysis results closer to objective facts. The literature Saafan et al. (2017) extended the existing supervised theme sentiment model and proposed a theme model based on multi-label and implicit emotions, and connected the potential theme with the user's induced emotions.

The literature Van (2018) implements a visual query system for e-commerce product data. The system is mainly based on statistical graphs, which is convenient for users to analyze shopping consumption preferences and tap potential value. In order to enable students to better understand the abstract data structure, the literature Xiu et al. (2017) used the JSAV (JavaScript Algorithm Visualization) visualization library to

combine visualization and tutorials, so that students can learn algorithms and practice in the visual display process. The literature Yen et al. (2017) improved the LIC algorithm for the visualization of vector data such as ocean currents and typhoons and implemented parallelization with GPU to realize real-time interactive visualization of vector fields. The literature (Adomavicius and Tuzhilin 2011) improved the data prefetching method and improved the data exchange capability during the visualization of large-scale flow field vector lines. Parallel coordinates (Wang and Zheng 2017) is the most classic visual representation method, which is mostly used to visualize high-dimensional geometry and multivariate data. Each attribute in high-dimensional data can be represented by each axis in parallel coordinates, and the data value gradually increases along the direction of the coordinate axis. The positioning of data points in parallel coordinates is very fast, as long as the data values of each attribute of the data points are connected on the corresponding coordinate axis. However, the classic parallel coordinates have the disadvantage of fixed coordinate axes. In the case of a large amount of data, the coordinates will have a serious edge overlap and cross phenomenon. The literature Celik et al. (2006) proposed a novel radial visual arrangement based on Time Wheel. This method surrounds the attributes into a time axis in the form of hexagons. The mapping relationship between each attribute in the data and the time axis is represented by different color lines. This method enhances the efficiency of multi-dimensional data browsing and analysis. The literature Das et al. (2014) proposed a visual method based on flexible link axes. In this method, each axis defines a related attribute and range. By modifying the parallel coordinate axis, it can help the user to flexibly define various attributes. The literature Eisenman et al. (2009) used hierarchical clustering algorithms to develop multi-resolution views of data, construct hierarchical clustering trees, and hierarchically display the data in parallel coordinates, which can well expand the size of the data set and represent the data from different abstract levels and reduce the visual clutter in parallel coordinates. The literature Gao et al. (2009) proposed a new parallel coordinate visualization clustering method. This method uses spline curves to replace the original parallel coordinates of lines to minimize the curvature and maximize the parallelism of adjacent edges to optimize the spline curves. By adjusting the shape of the edges and maintaining their relative order, the overall clustering effect is improved, and the visual effect is improved.

3 Linear local tangent space arrangement based on orthogonal discrimination

The LLTSA algorithm is obtained by LTSA linearization. This algorithm can obtain a clear mapping relationship, so that when new data points enter the program, dimensionality

reduction processing can be quickly performed to obtain the data after dimensionality reduction. The specific steps of the LLTSA algorithm are as follows:

The n data points in the high-dimensional space R^D are selected, and then the conversion matrix $X = \{x_1, x_2, \dots, x_n\}$, $x_i \in R^D$ is found to make the data points after the projection of the data m become $Y = (y_1, y_2, \dots, y_n)$, $y_i \in R^d, d < D$, and satisfy the linear mapping $Y = A^T X H_n, H_n = I - ee^T/n$. e is an n -dimensional column vector whose elements are all 1. We assume that the k points closest to each data point x_i are $X_i = (x_{i1}, x_{i2}, \dots, x_{ik})$, and k is a selection matrix containing only 0 and 1, and satisfying $Y_i = Y_{si}$. In order to maintain the local linear maximization and its internal local geometric structure, the reconstruction error E_i of the data should be minimized, so that the objective function of LLTSA can be obtained as:

$$\min \sum_i \|E_i\|^2 = \min \|YSW\|^2 = \min tr(YSWW^T S^T Y^T) \quad (1)$$

In the formula,

$$E_i = Y_i w_i = Y_{si} w_i, S = [s_1, \dots, s_n], \\ W = \text{diag}(w_1, \dots, w_n), W_i = H_n(I - V_i V_i^T)$$

V_i is the right singular vector of the d -th largest singular value of $X_i H_n$. To determine the unique Y , we set $Y Y^T = I_d$, and I_d is the identity matrix. Considering that the mapping is $Y = A^T X H_n$, the objective function is as follows:

$$\begin{cases} \min_Y tr(A^T X H_n B H_n X^T A) \\ A^T X H_n X^T A = I_d \end{cases} \quad (2)$$

From the above formula, we can see that the arrangement matrix is:

$$B = S W W^T S^T \quad (3)$$

In order to obtain the optimal projection vector, we can introduce the Lagrange multiplier and then solve the following generalized eigenvalue problem:

$$X H_n B H_n X^T a = \lambda X H_n X^T a \quad (4)$$

If the eigenvector corresponding to eigenvalue $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_d$ is a_1, a_2, \dots, a_d , then the conversion matrix of LLTSA is:

$$A = [a_1, a_2, \dots, a_d] \quad (5)$$

Matrix A is the mapping matrix. By multiplying the data matrix and matrix A , the feature vector of each data point can be obtained to realize the dimensionality reduction of the data, and then it can be used for data analysis. Although LLTSA has obtained a clear mapping relationship, it only maintains the neighbor relationship and

does not use the category information between the data to increase the spatial distance of different categories of data. Therefore, it needs to be improved.

In order to solve the above-mentioned shortcomings of LLTSA, this paper applies the ODLLTSA algorithm to complex chemical process fault diagnosis to extract feature data from process data. It can form a discriminant linear local tangent space arrangement algorithm (DLLTSA) by using the inter-class divergence matrix to obtain a new objective function, thereby obtaining the mapping matrix. Then, the mapping matrix is orthogonalized to obtain the mapping matrix of ODLLTSA. In this way, the category information can be fully utilized to increase the distance between different categories of data.

We assume that the data point x_i belongs to a class in $\{x_1, x_2, \dots, x_c\}$, and

$$\begin{cases} u = \left(\frac{1}{n}\right) \sum_{i=1}^n y_i \\ u_i = \left(\frac{1}{m_i}\right) \sum_{y_i \in X_i} y_i \end{cases} \tag{6}$$

Among them, u is the mean vector of all the data after projection, u_i is the mean vector of the i -th data after projection, and m_i is the number of the i -th data. If $Z = (z_1, z_2, \dots, z_n) = XH_n$ and $Y = (y_1, y_2, \dots, y_n) = A^T XH_n$, then $y_i = A^T z_i$. Then, the inter-class divergence matrix T_B of the ODLLTSA algorithm is:

$$\begin{aligned} & \sum_{i=1}^c m_i (u_i - u)(u_i - u)^T \\ &= \sum_{i=1}^c m_i \left\| \left(\frac{1}{m_i} \sum_{x_k \in X_i} A^T z_k - \frac{1}{n} \sum_{i=1}^n A^T z_i \right) \right\|^2 \\ &= \sum_{i=1}^c m_i \operatorname{tr} \left\{ \begin{pmatrix} \left(\frac{1}{m_i} \sum_{x_k \in X_i} A^T z_k - \frac{1}{n} \sum_{i=1}^n A^T z_i \right) \\ \left(\frac{1}{m_i} \sum_{x_k \in X_i} A^T z_k - \frac{1}{n} \sum_{i=1}^n A^T z_i \right)^T \end{pmatrix} \right\} \tag{7} \\ &= \operatorname{tr} \left\{ A^T \sum_{i=1}^c \begin{pmatrix} \left(\frac{1}{m_i} \sum_{x_k \in X_i} z_k - \frac{1}{n} \sum_{i=1}^n z_i \right) \\ \left(\frac{1}{m_i} \sum_{x_k \in X_i} z_k - \frac{1}{n} \sum_{i=1}^n z_i \right)^T \end{pmatrix} A \right\} \end{aligned}$$

In order to be able to better identify different types of data, the above value needs to be maximized, that is:

$$\begin{aligned} & \max \operatorname{tr} \left\{ A^T \sum_{i=1}^c \begin{pmatrix} \left(\frac{1}{m_i} \sum_{x_k \in X_i} z_k - \frac{1}{n} \sum_{i=1}^n z_i \right) \\ \left(\frac{1}{m_i} \sum_{x_k \in X_i} z_k - \frac{1}{n} \sum_{i=1}^n z_i \right)^T \end{pmatrix} A \right\} \tag{8} \\ &= \max \operatorname{tr} (A^T T_B A) \end{aligned}$$

In order to make the local geometric structure unchanged and satisfy the above formula, the optimization problem of DLLTSA needs to be calculated:

$$\begin{cases} \min \operatorname{tr} (A^T XH_n B H_n X^T A) \\ \max \operatorname{tr} (A^T T_B A) \end{cases} \tag{9}$$

The objective function of the DLLTSA algorithm can be obtained from the above formula:

$$J(A) = \frac{\operatorname{tr} (A^T XH_n B H_n X^T A)}{\operatorname{tr} (A^T T_B A)} \tag{10}$$

The above formula can be transformed into the following generalized eigenvalue problem:

$$\begin{aligned} XH_n B H_n X^T a_i &= \lambda_i T_B a_i, \\ \lambda_1 &\leq \lambda_2 \leq \dots \leq \lambda_d \end{aligned} \tag{11}$$

The conversion matrix of DLLTSA can be obtained from the above formula:

$$A = (a_1, a_2, \dots, a_d) \tag{12}$$

The projection vector of A obtained in the above formula is non-orthogonal, so in order to make the subspaces orthogonal, the Gram-Schmidt orthogonal method is used to orthogonalize A . If $c_1 = a_1$

$$c_i = a_i - \sum_{j=1}^{i-1} \frac{c_j^T a_i}{c_j^T c_j} c_j \quad (i = 2, 3, \dots, d) \tag{13}$$

$$\text{If } h_{j,i} = \frac{c_j^T a_i}{c_j^T c_j}, \text{ then } c_i = a_i - \sum_{j=1}^{i-1} h_{j,i} c_j.$$

If $C = AL$, and L is the upper triangular matrix, then the objective function of the ODLLTSA algorithm is:

$$J(C) = J(AL) = \frac{\operatorname{tr} (L^T A^T XH_n B H_n X^T AL)}{\operatorname{tr} (L^T A^T T_B AL)} \tag{14}$$

The following results can be obtained from the above formula:

$$C = (c_1, c_2, \dots, c_d) \tag{15}$$

The C obtained by the above formula is the mapping matrix of the ODLLTSA algorithm. The matrix after the

dimensionality reduction can be obtained by multiplying the data matrix and the mapping matrix.

From the above description, the flow of the ODLLTSA algorithm can be obtained as shown in Fig. 1.

The specific steps of the ODLLTSA algorithm:

Input: data $X = \{x_1, x_2, \dots, x_n\}$, number of nearest neighbors k , dimension d after dimensionality reduction.

Output: conversion matrix $C = (c_1, c_2, \dots, c_d)$.

1. According to the given value of k , the k nearest neighbors of point $x_i (i = 1, 2, \dots, n)$ are found to obtain the nearest neighbour matrix X_i .
2. $X_i H_n$ is subjected to singular value decomposition to obtain V_i , and then the weight $W_i = H_n (I - V_i V_i^T)$ is calculated.
3. According to formula (3), the permutation matrix B is calculated.
4. According to formula (7), the inter-class divergence matrix T_B of the ODLLTSA algorithm is calculated.
5. According to formula (11), the generalized eigenvalues are solved, and matrix A is obtained, and the non-orthogonal mapping matrix of DLLTSA is obtained;
6. According to formula (13) and formula (14), matrix A is orthogonalized, and finally the mapping matrix C of ODLLTSA is obtained.

After the mapping matrix C is obtained, the data matrix and matrix C are multiplied to obtain the feature vector of each data to complete the dimensionality reduction of the data, and then the data after the dimensionality reduction is input into the chemical process fault monitoring and diagnosis. ODLLTSA can effectively deal with nonlinear data. Reducing the m -dimensional matrix to the n -dimensional matrix ($m > n$) can not only retain the local neighbor relationship between data, but also reduce or maintain the distance between data under normal working conditions. The distance between the fault condition data and the normal condition data is increased as much as possible, thereby improving the accuracy of fault monitoring and diagnosis.

4 Chemical process fault monitoring based on KPCA

The kernel principal component analysis method is a nonlinear form of principal component analysis. It uses a kernel function to map the original space of nonlinear data into a high-dimensional space to make the data linearly separable, and then implement PCA in the high-dimensional space. The mapping diagram is shown in Figs. 2 and 3.

The specific analysis of KPCA is as follows:

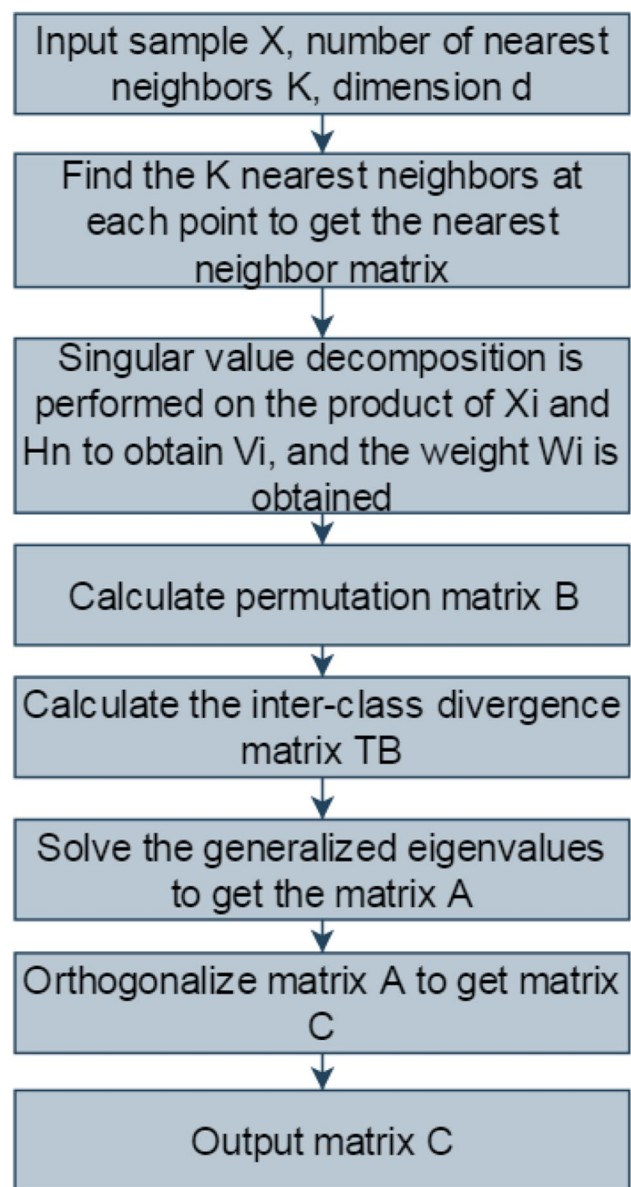


Fig. 1 ODLLTSA algorithm flow chart

We set the data point to n , and a positive definite kernel matrix K can be obtained from the kernel function calculation, as follows:

$$K = k(x_i, x_j) (i, j = 1, 2, \dots, n) \quad (16)$$

In order to ensure that the input data vector of the feature space is 0 mean, the data is standardized, and $\sum_{i=1}^N \varphi(x_i) = 0$. Then, the covariance matrix of the feature space is:

$$C^H = \frac{1}{n} \sum_{i=1}^N \varphi(x_i) \varphi(x_i)^T \quad (17)$$

Among them, $\phi(X_i)$ is the function after data mapping, C^H is the eigenvalue decomposition. If we assume that λ is the eigenvalue and v is the eigenvector, then

$$\lambda v = C^H v \quad (18)$$

After multiplying both sides of formula (18) by the mapped point $\phi(X_k)$, the following result is obtained:

$$\lambda(\phi(X_k) \cdot v) = (\phi(X_k) \cdot C^H v) \quad (19)$$

For the eigenvector v of any $\lambda \neq 0$, there is a coefficient matrix $\alpha_i (i = 1, 2, \dots, N)$.

$$v = \sum_{i=1}^N \alpha_i \phi(X_i) \quad (20)$$

After formulas (16), (17), (20) are brought into formula (19), we obtain:

$$\begin{aligned} & \lambda \sum_{i=1}^N \alpha_i \langle \phi(X_k), \phi(X_i) \rangle \\ &= \frac{1}{N} \sum_{i=1}^N \alpha_i \left\langle \phi(X_k), \sum_{j=1}^N \phi(X_j) \right\rangle \langle \phi(X_i), \phi(X_j) \rangle \end{aligned} \quad (21)$$

In the formula: $k = 1, 2, \dots, N, j = 1, 2, \dots, N$.

After formula (16) is brought into formula (19), we can obtain:

$$\lambda \sum_{i=1}^N \alpha_i K_{ki} = \frac{1}{N} \sum_{i=1}^N \alpha_i \sum_{j=1}^N K_{kj} K_{ji} \quad (22)$$

Fig. 2 Before mapping

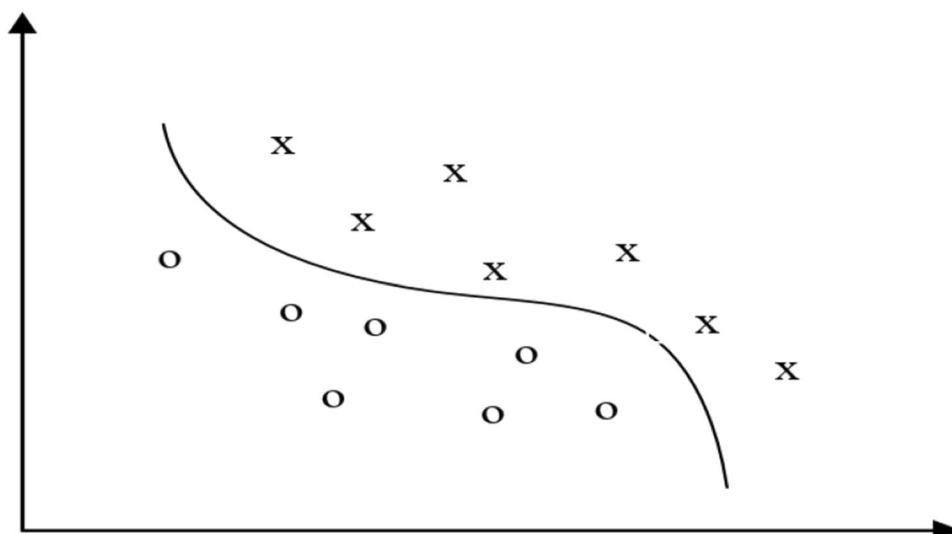
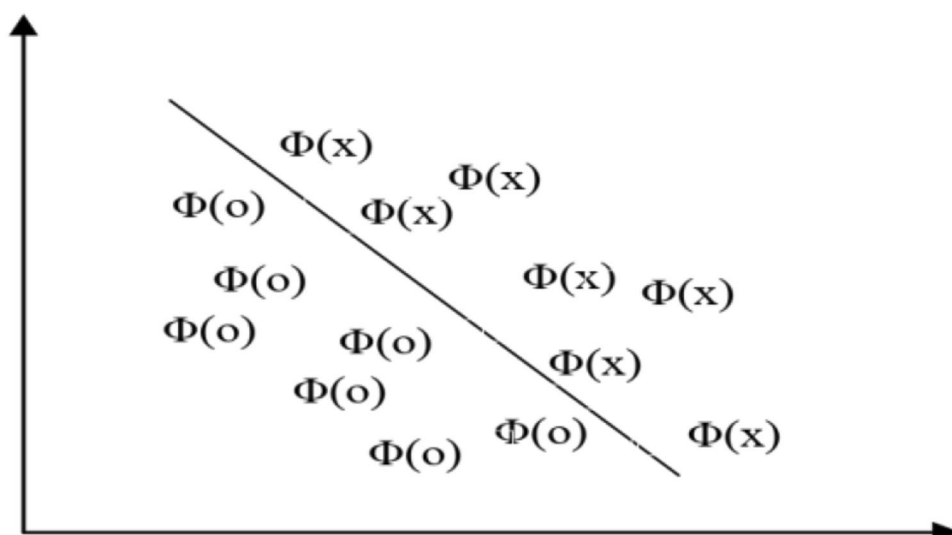


Fig. 3 After mapping



Formula (22) is converted into the following characteristic equation:

$$\lambda NK\alpha = K^2\alpha \tag{23}$$

The kernel matrix K obtained by formula (16) needs to be centralized in the feature space. The process is as follows:

$$\tilde{K} = K - 1_N K - K 1_N + 1_N K 1_N \tag{24}$$

Among them: 1_N refers to a matrix that all elements are $\frac{1}{N}$. Then, formula (23) can be expressed as:

$$\lambda N\alpha = \tilde{K}\alpha, \alpha = [\alpha_1, \dots, \alpha_N] \tag{25}$$

α is standardized to make α satisfy $\|\alpha\|^2 = \frac{1}{N}\lambda$. Then, KPCA can obtain the components of each principal component by solving the above eigenvalue problem.

When the system generates new data $\phi(X)$, it needs to project under the feature vector $v_k, k = 1, \dots, p$. p is the number of principal elements of KPCA. Then, the projection t_k is as follows:

$$t_k = \sum_{i=1}^N \alpha_i^k [\bar{\phi}(X_i), \bar{\phi}(X)] = \sum_{j=1}^N \alpha_j^k \bar{k}(x_j, x) \tag{26}$$

KPCA includes two parts: offline modeling and online monitoring. The specific steps are as follows:

(a) Offline modeling:

1. The data matrix Z under normal operating conditions is normalized to obtain a mapping matrix and a standardized matrix X.
2. The Gaussian kernel function is selected to calculate the kernel matrix K, and centralized processing is performed according to formula (24).
3. According to formula (25), the eigenvalues and eigenvectors are calculated, and the eigenvectors are normalized.
4. According to formula (26), the principal component t_k is calculated in the feature space.
5. Under normal operating conditions, the statistics of T^2 and SPE are calculated and defined as follows:

The statistic of T^2 is:

$$T^2 = t\Lambda^{-1}t^T = [t_1, t_2, \dots, t_p]\Lambda^{-1}[t_1, t_2, \dots, t_p]^T \tag{27}$$

In the formula: $t_k(k = 1, 2, \dots, p)$ represents the principal element after normal working condition data mapping, and Λ^{-1} represents the inverse matrix of the diagonal matrix of eigenvalues.

The statistic of SPE is:

$$SPE = \|\phi(x) - \phi_p(x)\|^2 = \sum_{i=1}^N t_i^2 - \sum_{i=1}^p t_i^2 \tag{28}$$

6. The control limits of T^2 and SPE statistics are solved, the formula is as follows:

The control limits of T^2 is:

$$T^2_{p,N,\alpha} = \frac{p(N-1)}{N-p} F_\alpha(p, N-p) \tag{29}$$

In the formula: N represents the number of input data points, and p represents the number of principal elements of KPCA.

The control limits of SPE is:

$$SPE_{lim} = g\chi_h^2 \tag{30}$$

In the formula: g and h are constant coefficients related to the mean and variance of SPE.

(b) Online monitoring

1. The real-time working condition data x_i is collected, and the mapping matrix obtained by normalizing the normal working condition data is used to standardize the data x_i ;
2. The new kernel matrix K_i is constructed and centrally processed.
3. The nonlinear component t_i is extracted in the feature space.
4. The T^2 and SPE statistics of real-time operating condition data are calculated.
5. The control limits of T^2 and SPE statistics are obtained. If both statistics exceed or only the SPE statistics exceed the control limit, the system detects the fault and enters it into the fault diagnosis. However, if it is otherwise, the operator needs further analysis to determine whether there is a fault.

5 Model building

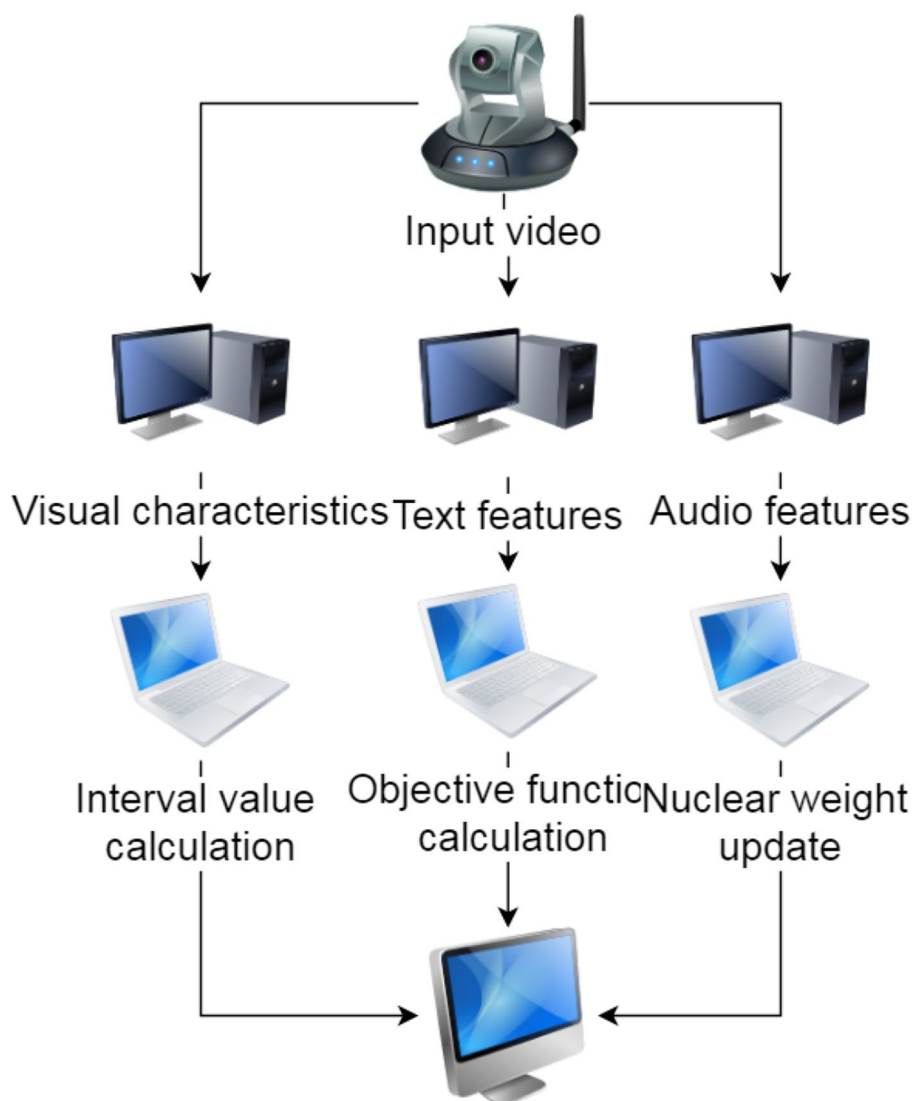
Feature level fusion is to extract features from all patterns, combine them to form a long feature vector, and then feed the feature vector to MKL to perform the classification task. Multi-core model is a kind of core-based learning model with strong flexibility. Recently, theory and application have proved that using multi-core instead of single-core can achieve better classification results and can obtain better performance than single-core model or single-core machine combination model. The simplest and most commonly used method for constructing multi-core models is to consider the combination of multiple basic kernel functions. After proposing features from visual, text, and audio information,

we input the features into the kernel space, extend MKL using interval constraints, and calculate its target cost function. According to the recognition ability of different base features, the weights of different base features are learned, and the gradient weighting method is used to update the feature weights. Moreover, the kernel is dimensionally normalized to maximize the maximum discriminative power of its basic features. The multi-core learning model based on interval dimension constraints proposed in this paper is shown in Fig. 4:

The latest developments in multi-core learning make it an attractive technology in the field of kernel machine learning. By merging multiple cores into a unified optimization framework, MKL strives to jointly learn the optimal combination of multi-core and related predictors in supervised or semi-supervised learning settings. It does not require single-core learning, and can meet the practical needs of heterogeneous information in sample features, irregular data

in feature space, uneven data distribution, and large problem size. Due to the automatic adjustment of kernel parameters, various features of data representation are described in detail, so that multi-source and heterogeneous data sets are processed flexibly and stably. In addition, using multiple kernels can enhance the interpretability of the model and improve the generalization performance of the classifier. The latest research on multi-core learning shows that this method can effectively fuse multiple basic features in target detection and recognition. However, MKL tends to select only the most discriminatory basic features and ignore other less discriminatory basic features that may provide supplementary information. In addition, MKL usually uses Gaussian RBF kernels to transform each basic feature into a high-dimensional space. Generally, the basic characteristics of different modes require different kernel parameters to get the best performance. Therefore, MKL may not be able to utilize the maximum discriminative power of all

Fig. 4 Multi-core learning model based on interval dimension constraints



basic features from multiple modes simultaneously. In order to solve these problems, we propose a multi-core learning MDMKL method based on interval dimension constraints. This method extends MKL with interval constraints and applies dimension normalized RBF kernel to multimodal feature fusion. Moreover, the proposed MDMKL method learns the weights of different base features based on their recognition capabilities. Unlike traditional MKL, when constructing the optimal combination kernel, MDMKL merges and discriminates basic features by allocating smaller weights, so as to take full advantage of the complementary features of different modes.

6 Model performance analysis

In this study, after extracting features from audio, text, and visual information, the proposed MDMKL is used to perform feature-level fusion. Figure 5 shows the experimental results obtained by using different fusion algorithms to classify sentiment of multimodal data information MOUD. It can be seen from the figure that as the training data set increases, the classification results after fusion become better and better. In the initial stage of training sample growth, the classification accuracy is rapidly improved. When the amount of data increases to a certain extent, the growth rate will slow down or even occasionally show negative growth. The MDMKL model proposed by this research has the highest classification accuracy. The accuracy rate obtained by MDMKL is 97.25%, and the accuracy rates of SVM and MKL are 88.90% and 94.34%, respectively.

When conducting experiments on the MOUD dataset, we conduct an analysis of the convergence of the algorithm, that is, the f value of the objective function changes as the number of iterations increases. In Fig. 6, we compare the

convergence performance of the MDMKL and MKL methods. The stopping criteria for both methods is that the change between two consecutive steps is less than the given threshold 0.001. Compared with traditional MKL, it is found that within a limited range of iterations, as the number of iterations increases, MDMKL converges faster than MKL. In addition, we can also observe that the target value of MDMKL can converge to a stable value in less than 5 iterations.

Table 1 and Fig. 7 show the results of using MDMKL to fuse different modalities. In the single-mode classifier, the visual mode provides the classification results with the highest accuracy. For visual and text modal fusion, the accuracy rate of V + T fusion modal is up to 96.97%. For visual and audio modal fusion, the accuracy rate of V + A fusion modal is up to 91.84%.

For text and audio modal fusion, the accuracy of T + A modal fusion is up to 96.56%. We found that when visual emotion is integrated with audio and text emotion, the visual text frame is superior to the visual audio frame. When the three modes of audio, text and visual are fused, the accuracy is higher than that of any two modes. The accuracy of the three modal fusion is up to 97.25%, which is higher than the accuracy of the existing framework.

Finally, the performance analysis of the Japanese social sentiment classification of this research model is conducted. The data of this research comes from Twitter, which has many Japanese users, and selects the Japanese part of the platform. A total of 60 sets of data are collected for experiments, and sentiment classification is performed on these data to determine the classification accuracy. The system requires a recognition rate of 95% to qualify. Therefore, this study takes 95% as the bottom line, and the results are shown in Table 2 and Fig. 8.

Fig. 5 Experimental comparison diagram of different fusion algorithms

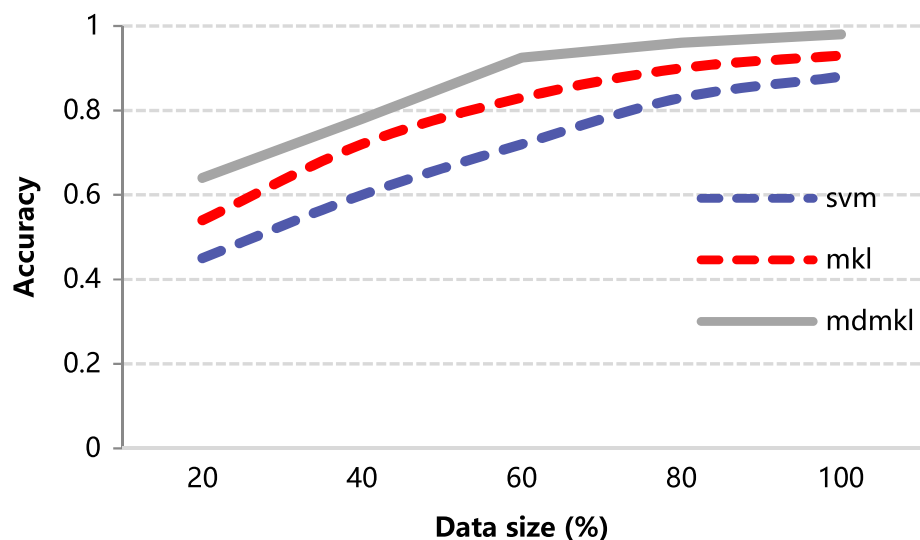


Fig. 6 Comparison diagram of target convergence and iteration times

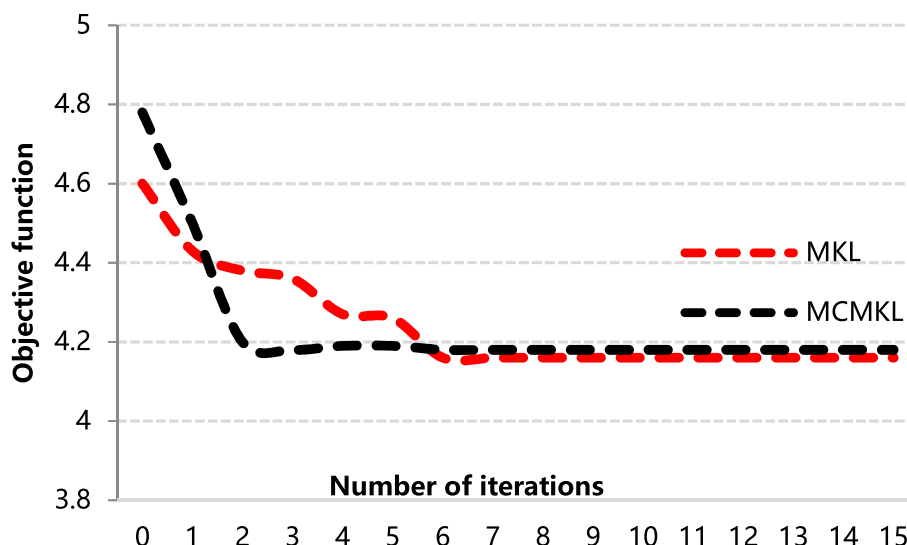


Table 1 Comparison table of single-mode and multi-mode feature level fusion

Modal	CNN (%)	CRMKL (%)	The method mentioned in this article (%)
Audio Mode A	74.96	79.93	77.06
Visual modality V	77.14	95.45	96.20
Text modality T	80.57	75.23	82.49
V + T modal fusion	86.31	96.71	97.94
V + A modal fusion	84.53	84.69	92.76
A + T modal fusion	84.96	96.33	97.53
V + T + A modal fusion	89.49	97.08	98.22

requirements of the system for this research algorithm. This result shows that the model proposed in this study has a certain stability, which meets the system’s requirements for algorithm stability.

Based on the real data set of Twitter, according to the research needs of this article, a reasonable filtering strategy is designed to select the experimental data set, and this study supplements the tweet pictures and comment information through web crawling. On this basis, the tweet text is preprocessed to reduce its irregularity. Secondly, from a psychological point of view, this study extracts the underlying visual features related to emotional factors from tweet pictures, and uses the visual bag of words model to quantify the underlying visual features into visual words. Finally, this study verifies the validity of the data set by analyzing the relevant statistical characteristics of the data set.

As can be seen from Fig. 8 and Table 2 above, the accuracy rate of this research model on the Japanese social sentiment classification results exceeds 95%, which meets the

Using the Twitter real data set as the experimental data set, the effectiveness of the model algorithm proposed in

Fig. 7 Comparison diagram of single-mode and multi-mode feature level fusion

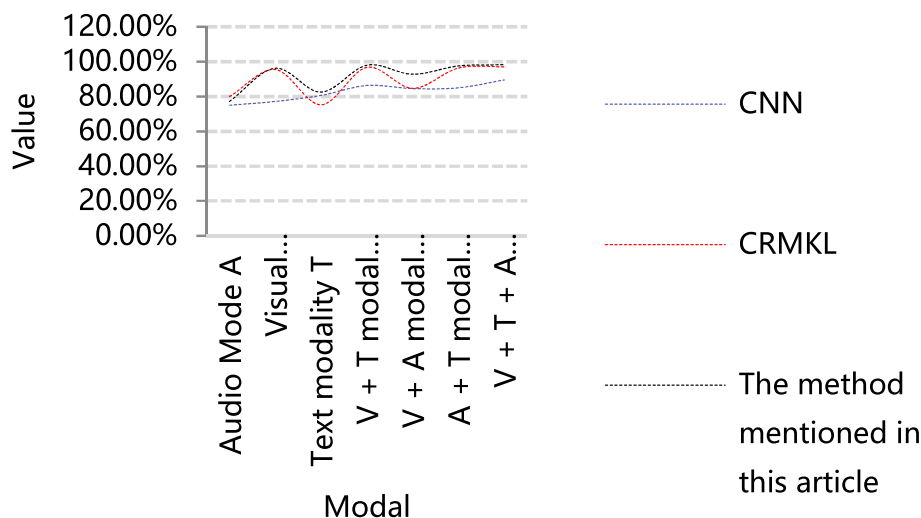


Table 2 Statistical table of the accuracy of the Japanese social emotion classification of the model

	Model recogni- tion rate	Set limits	Model recogni- tion rate	Set limits
1	98	95	31	96
2	97	95	32	98
3	97	95	33	96
4	98	95	34	97
5	99	95	35	96
6	99	95	36	95
7	98	95	37	98
8	99	95	38	96
9	97	95	39	99
10	99	95	40	98
11	96	95	41	99
12	100	95	42	99
13	96	95	43	99
14	96	95	44	98
15	99	95	45	99
16	97	95	46	95
17	99	95	47	95
18	99	95	48	97
19	96	95	49	99
20	96	95	50	99
21	96	95	51	97
22	99	95	52	98
23	97	95	53	97
24	99	95	54	96
25	100	95	55	99
26	100	95	56	100
27	95	95	57	95
28	99	95	58	97
29	96	95	59	97
30	100	95	60	95

this paper is verified through experiments. On this basis, this research designs and implements a user sentiment analysis prototype system.

7 Conclusion

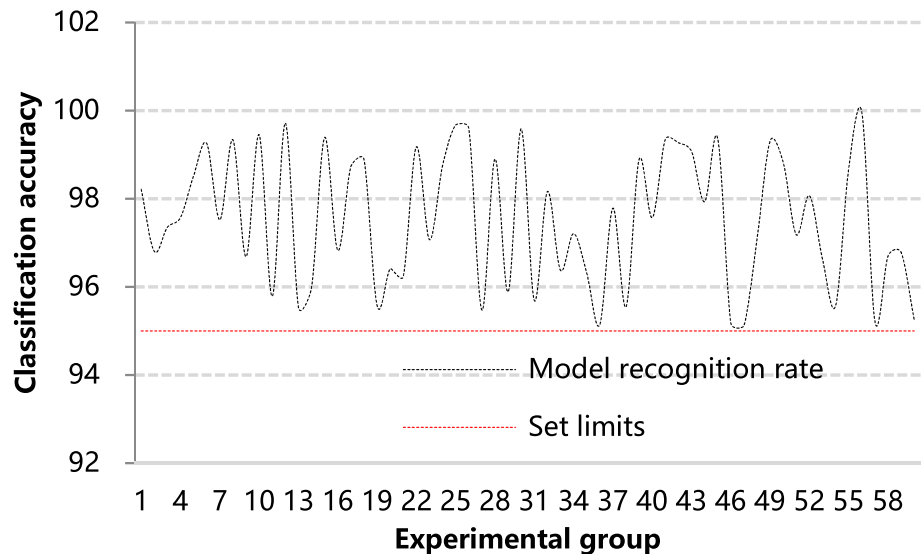
Aiming at the problem of user sentiment analysis in Japanese social media, this paper proposes a topic model of user sentiment analysis based on SVM and KNN and uses real social media data sets to conduct experiments to verify the effectiveness of the model algorithm.

Taking TE process data as an example, this paper selects radial basis kernel function and grid search method for the construction of SVM classifier.

Moreover, this study combines the ODLLTSA algorithm with the KNN algorithm to propose KNN-based feature recognition. The ODLLTSA algorithm can increase the distance between heterogeneous points, reduce or maintain the distance between similar points, and ensure that the local neighbor points remain unchanged. The KNN algorithm is a distance-based monitoring algorithm and can handle non-linear and non-Gaussian data very well. Therefore, the two can be well matched to improve the speed and accuracy of fault monitoring.

In addition, this study uses the Twitter real data set as the experimental data set to verify the effectiveness of the user sentiment analysis model proposed in this article. Through experimental research, we know that the model proposed in this study has certain stability, and the accuracy rate of emotion classification is above 95%, which meets the system's requirements for algorithm stability and accuracy.

Fig. 8 Statistical diagram of the accuracy of the Japanese social sentiment classification of the model



References

- Adomavicius G, Tuzhilin A (2011) Context-aware recommender systems. In: Ricci F, Rokach L, Shapira B, Kantor P (eds) Recommender systems handbook. Springer, pp 217–253
- Arnold M, Rui H, Wellssow W (2011) An approach to smart grid metrics. In: Proceedings of the 2011 2nd IEEE PES International Conference and exhibition on innovative smart grid technologies, Manchester, pp 1–7
- Bourdeau M, Zhai X, Nefzaoui E, Guo X, Chatellier P (2019) Modeling and forecasting building energy consumption: a review of data-driven techniques. *Sustain Cities Soc* 48:1–27
- Celik T, Demirel H, Ozkaramanli H (2006) Automatic fire detection in video sequences. In: 14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4–8, 2006. pp 1–5
- Cui R, Chen L, Yang C et al (2017) Extended state observer-based integral sliding mode control for an underwater robot with unknown disturbances and uncertain nonlinearities. *IEEE Trans Industr Electron* 64(8):6785–6795
- Das J, Mukherjee P, Majumder S, Gupta P (2014) Clustering-based recommender system using principles of voting theory. In: 2014 International Conference on contemporary computing and informatics (IC3I), Mysore, India, 27–29 November 2014. pp 230–235
- Eisenman SB, Miluzzo E, Lane ND, Peterson RA, Ahn G-S, Campbell AT (2009) Bikenet: a mobile sensing system for cyclist experience mapping. *ACM Trans Sens Netw (TOSN)* 6(1):6
- Gao C, Kong F, Tan J (2009) Healthaware: tackling obesity with health aware smart phone systems. In: 2009 IEEE International Conference on robotics and biomimetics (robio), Guilin, China, 19–23 December 2009. pp 1549–1554
- Li P, Bao G, Fang X et al (2018) Adaptive robust sliding mode trajectory tracking control for 6 degree-of-freedom industrial assembly robot with disturbances. *Assem Autom* 38(3):259–267
- Mary I, Arockiam L (2017) Detection of outliers in the IoT data using the STCPOD model. *Int J Eng Res Comput Sci Eng* 4(10):1–6
- Mukherjee J, Mukherjee S, Kar IN (2017) Sliding mode control of planar snake robot with uncertainty using virtual holonomic constraints. *IEEE Robot Autom Lett* 2(2):1077–1084
- Nasir MT, El-Ferik S (2017) Adaptive sliding-mode cluster space control of a non-holonomic multi-robot system with applications. *IET Control Theory Appl* 11(8):1264–1273
- Rabie A, Ali S, Saleh S, Ali H (2020) A fog based load forecasting strategy based on multi-ensemble classification for smart grids. *J Ambient Intell Human Comput* 11(1):209–236
- Saafan MM, Abdelsalam MM, Elksas MS et al (2017) An adaptive neuro-fuzzy sliding mode controller for MIMO systems with disturbance. *Chin J Chem Eng* 25(04):87–100
- Schuelke-Leech B, Barr B, Muratori M, Yurkovich B (2015) Big Data issues and opportunities for electric utilities. *Renew Sustain Energy Rev* 52:937–947
- Stefano F, Gabriele DA (2020) On the ethereum blockchain structure: a complex networks theory perspective. *Concurr Comput Pract Exp* 32(12):e5493
- Van M (2018) An enhanced robust fault tolerant control based on an adaptive fuzzy PID-nonsingular fast terminal sliding mode control for uncertain nonlinear systems. *IEEE/ASME Trans Mech* 23(3):1362–1371
- Vimalkumar K, Radhika N (2017) A big data framework for intrusion detection in smart grids using Apache Spark. In: Proceedings of the 2017 International Conference on advances in computing, communications and informatics (ICACCI). IEEE, pp 198–204
- Wang H, Zheng TQ (2017) RBF network adaptive control based on SMC compensation for six-axis manipulator. *J Northeast Univ* 38(11):1601–1606
- Xiu C, Hou J, Zang Y et al (2017) Synchronous control of hysteretic creep chaotic neural network. *IEEE Access* 4:8617–8624
- Yen VT, Nan WY, Cuong PV et al (2017) Robust adaptive sliding mode control for industrial robot manipulator using fuzzy wavelet neural networks. *Int J Control Autom Syst* 15(6):2930–2941
- Yu W, Liang F, He X, Hatcher W, Lu C, Lin J, Yang X (2018) A survey on the edge computing for the internet of things. *IEEE Access* 6:6900–6919

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.