Journal of Engineering Science and Technology Review 15 (1) (2022) 51 - 57

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Research Article

Complex Power Quality Disturbances Classification Based on Multi-label Active Learning

Chen Zhao*, Haoyi Zhang, Caigen Wei, and Ziying Zeng

College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou 350100, China

Received 5 February 2022; Accepted 9 April 2022

Abstract

Power quality (PQ) disturbances generated during the power grid operation are complicated and volatile in real life. If a large number of complex PQ disturbances (CPQDs) from power grid monitoring devices are all labelled artificially, then it may consume a lot of human resources. To effectively utilize these unlabeled data collected by the monitors to improve the accuracy of the learning model, this study proposed an approach for the recognition of CPQDs using a multi-label active learning strategy. First, the study presented a novel active learning strategy based on label exclusiveness and ranking score (LERS) by analyzing the label relation among different PQ disturbances. Second, the strategy was incorporated into the multi-label extreme learning machine classifier to train and identify CPQDs. Finally, extensive experiments in the study validate the effectiveness of the proposed method. Results indicate that LERS improves the performance of the classification model by adding the most informative sample. As the number of labelled samples increases from 1000 to 8000, the evaluation metric MicroF1 reaches more than 0.7. The corresponding labelling cost of the proposed strategy is reduced by 40% compared with other strategies when obtaining certain accuracy. This study provides a specific reference for recognizing CPQDs and has a bright application prospect.

Keywords: Power Quality Disturbance, Multi-Label Classification, Active Learning, Extreme Learning Machine, Label Exclusion

1. Introduction

With the large-scale integration of new energy sources into the power grid and the booming increase in power load types, especially the wide spread of distributed energy sources, sensitive electronic components, power converters, industrial drives, reactive power devices, and solid-state switches, various power quality (PQ) events are increasing day by day, manifesting as a large number of steady-state and transient PQ problems [1]. Therefore, to protect the power system from adverse effects, the accurate recognition of complex PQ disturbances (CPQDs) is of great importance.

The recognition of CPQD signals usually combines the modern signal processing algorithms for feature extraction and machine learning algorithms for final classification. Scholars have conducted numerous related studies on PQ analysis in the framework [2-4]. In the current study of CPQD classification, the sample data set used to train the classifier generates training samples randomly, in which most are generated by MATLAB within the IEEE 1529 standard. Recorded and broadcast signals from standard power sources or grid fault recorders are used as training datasets for the classifiers. Conventional CPQD training methods usually adopt a typical supervised machine learning framework. Specifically, the supervised model is trained on input data that have been labeled for a particular output. Notably, most PQ monitors in power systems only collect the raw waveforms without corresponding disturbance class information. To accommodate the supervised learning process, these time series unlabeled data need to be labeled artificially with the type of disturbances it probably contains. The efforts require abundant domain knowledge and labeling experience of CPQDs. Therefore, fully utilizing the unlabeled data in the supervised learning scenario of real grid signal analysis is time-consuming and expensive. However, training a supervised learning model with only few labeled data will result in severe overfitting problems when recognizing real signals. Furthermore, the abundant intrinsic information contained in the unlabeled CPQDs is neglected.

To make full use of the labeled and unlabeled data in the training phase of the CPQDs classification model, this study presents a novel multi-label active learning strategy that incorporates label exclusion into the conventional uncertainty measure. In the proposed method, the margin vectors of unknown CPQDs are gained from a multi-label classifier based on extreme learning machine (ELM), and then the Borda method is adopted to aggregate the vectors to output a unified uncertainty measure rank for the unlabeled samples across all labels. The novelty of this study is to develop a label exclusion measure for CPQD samples in selecting the most indeterminate data for the classifier. Experiment results demonstrate that the proposed strategy can achieve better accuracy in selecting the most informative example using the random discriminative projection extreme learning machine for multi-label learning (RDPEML). The comparison results on the synthetic dataset demonstrate that the proposed strategy outperforms the state-of-the-art multi-label active learning strategies.

2. State of the art

Many studies on PQ analysis have been published in the last decade. Liu et al. used the discrete wavelet transform to

*E-mail address: fauzhao@gmail.com
ISSN: 1791-2377
doi:10.25103/jestr.151.07

conduct time-domain analysis on the disturbance signals, and the extracted features were fed to a random forest classifier. This method is relatively simple to implement, but a lot of difficulties is encountered in explaining random forest, which is a black-box model [5]. Li et al. adopted wavelet transform to extract features and then optimized the parameters of the support vector machine (SVM) algorithm by genetic algorithm. Although this method improved speed and accuracy, it neglected the CPQD signals and only identified the PQ event that contained a single disturbance [6]. Asman et al. proposed a PQ recognition method based on discrete wavelet transform and SVM. This method could extract the information of transient signals and increase the recognition rate; however, solving a quadratic problem in SVM added the computational cost of classification [7]. Kumar et al. extracted the statistical features of PQ by using the traditional signal analysis method and then classified them by decision tree (DT). This method verified the robustness of PQ to noise, but its extracted features are relatively traditional, and only eight types of CPQDs were classified, with few classification categories and insufficient consideration [8]. Mahela et al. combined S transform with Hilbert transform to extract the statistical features of signals and finally classified those using DTs. This method was easy to implement because of its simple structure and short operation time, whereas the DT algorithm was prone to overfit when types of CPQDs increased [9]. Swarnkar et al. used the S-transform with the Hilbert transform when conducting feature extraction and then fed them into DT for classification. This method considered various CQPDs, and the experiment results showed its effectiveness. However, the selection of manufacturing features greatly influenced the final performance of DT [10]. Elango et al. hybridized S transform and wavelet transform to extract the time-frequency features of CPQDs and classified them using a backpropagation (BP) neural network. Although the method obtained excellent accuracy, the BP neural network's slow convergence hindered the PQ recognition model [11-12]. Wang et al. decomposed the signals by using modulation wideband mode decomposition algorithm and then extracted the features by multi-scale fuzzy entropy. Finally, the features were classified by a BP neural network, but they considered DC signals. Thus, the analysis was relatively straightforward [13]. Liao et al. combined improved local mean decomposition and Hilbert transform to process signals and then used radial basis function (RBF) neural network for classification. This method had low requirements for PQ disturbance signal conditions and was universal. However, optimizing the hyperparameters of the RBF neural network was time-consuming [14]. Vidhya et al. constructed a learning model that contained ELM and RBF, which reflected the superiority of CPQD classification. Still, the network's complex structure had a risk of poor generalization [15]. Liu et al. combined the traditional signal analysis method with butter worth distribution to extract high-order moment statistical features and used the RBF neural network as the classifier. This method extracted the effective features of CPQDs that were difficult to distinguish, but the classification types of CPQDs were limited [16].

Qu et al. employed sparse autoencoders to extract the features in unsupervised learning framework. A softmax classifier was adopted to learn and classify the CPQDs. This method exhibited a remarkable anti-noise characteristic, whereas a range of only two kinds of compound disturbances hindered the universality of this method [17]. Wu et al. utilized a recurrent neural network to extract the deep features of CQPDs and then connected them to a softmax classifier to output the recognition accuracy, but it had the disadvantage of heavy computational cost [18]. Zheng et al. transformed CPQDs into images by using Gramian angular field and inputted them into convolutional neural networks to extract and classify features automatically. However, the PQ signal was a typical time-series signal. The 2D transformation might neglect the intuitive characteristic [19].

The above studies mainly develop their classification model in supervised learning. So far, works on classifying CPQDs from weakly supervised learning, not even from active learning, are few. This study proposes a new multi-label classification algorithm for CPQDs by combining active learning with ELM and by adding mutual label exclusion. Starting from active learning strategies, the influence of different active learning strategies on the final performance curve is studied, and the superiority of this research method is obtained.

The remainder of this study is organized as follows. The third section mainly introduces the active learning and the active learning strategy proposed in this study and describes the CPQD classification method that combines ELM and active learning. The fourth section combines different active learning strategies with ELM to classify the same data with multiple labels and obtain the final performance curve. The last section summarizes this study and draws relevant conclusions.

3. Methodology

3.1 Active Learning

Since active learning first proposed by Cohn in 1996 [20], it has rapidly become a subfield of machine learning. Its fundamental assumption is whether the learning algorithm is allowed to select the most valuable data for improving the learning model. And if so, compared with the supervised method, the learning algorithm might be trained with less training data when reaching comparable performance. Active learning is often divided into stream-based selective and pool-based learning according to how unlabeled samples are obtained [21]. In the stream-based selective sampling process, unlabeled data flow from the data source and are sequentially submitted to the selection engine. The selection engine decides whether to label the currently offered samples or not. As the samples are handed over to the selection engine one by one, the distribution of the overall dataset is unknown. Thus, evaluating the potential value of the samples becomes difficult. Pooling-based active learning does not rely on a single sample but sorts and selects the most valuable subset from the data pool. Pooling-based active learning is suitable for general application scenarios where a small set of labeled samples and a large set of unlabeled samples $\mathrm{U_s}$ coexist. The CPQD classification problem examined in this study also conforms to this application scenario, so the pool-based active learning method is adopted.

The basic flow of pool-based active learning is shown in Figure 1. Given a basic classifier ϕ used in the active learning process, in the pool-based active learning scenario, a small set of labeled samples $\mathrm{L_s}$, a large set of unlabeled samples $\mathrm{U_s}$ and an active learning strategy γ (selection criteria, such as uncertainty measures) are assumed to exist.

The general process of active learning is presented as follows:

(1) γ selects unlabeled sample data from $\mathrm{U_s}$;

(2) The selected sample data are labeled by human experts;

(3) The selected samples are added to $\mathrm{L_s}$ and removed from $\mathrm{U_s}$;

(4) ϕ is trained by $\mathrm{L_s}$ again;

(5) The performance of the basic classifier ϕ is evaluated;

(6) Return to Step 1 if the stop criterion is not satisfied.

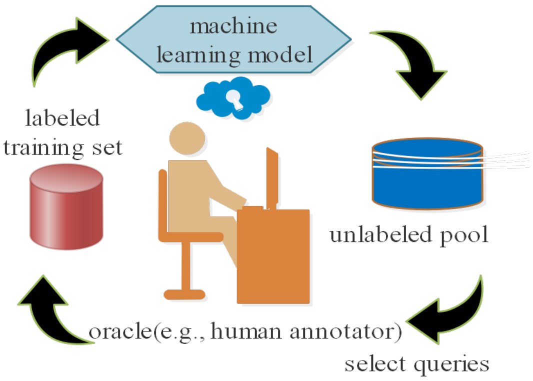


Fig. 1. Pool-based active learning circle

3.2 CPQD classification based on active learning

3.2.1 Measure based on Borda method

Uncertainty sampling is one of the most well-known strategies that active learning adopts in selecting unlabeled samples. Specifically, uncertainty sampling heuristically selects the samples with the most uncertain prediction results by using the current learning model and submits them to human experts for labeling. For multi-class classification problems, least confidence, margin, and entropy are typical schemes for uncertainty measures. Given that identifying CPQDs is a typical multi-label learning problem, human experts have to decide whether the signal is associated with each label when labeling unknown disturbance signals. Therefore, the cost of manual labeling of CPQDs is much higher than that of multi-class classification.

In this study, the margin of sample x_i in l class label can be calculated as:

$$m_{\Phi}^{i,l} = \left|P(l=1|\ f_l(x_i)) - P(l=0|\ f_l(x_i))\right| \tag{1}$$

If the margin is large, then the learning model is less likely to make mistakes when predicting whether the sample x_i contains this label, and x_i is less valuable to improve the classifier's performance. On the contrary, a small margin means that the classifier is not sure enough to predict whether the sample contains a label. Therefore, the smaller margin between the two most likely class labels indicates more informative the instance is for the performance lifting of the model. Given vectors of margins for unlabeled samples, q rankings of samples (q represents the total number of labels contained in the label set) can be calculated to consider the information of all unlabeled samples. Moreover, each sequence represents one ranking of all unlabeled samples. The sorted sequence can be calculated by the formula:

$$\tau_\ell = \left(i_{\pi 1}, i_{\pi 2}, \ldots, i_{\pi |U_s|}\right)|\ m_{\Phi}^{i_{\pi 1},\ell} < m_{\Phi}^{i_{\pi 2},\ell} \ldots < m_{\Phi}^{i_{\pi |U_s|},\ell} \tag{2}$$

The ranking τ_l corresponds to the complete ordering of all unlabeled samples according to the margin on the label l. The uncertainty measure of multi-label samples needs to integrate the ranking information of all samples on all labels; that is, by accumulating all the rankings, the sample placed in the first position of the final ranking contains the most uncertain information. Obtaining the cumulative margin value is a well-known rank aggregation problem. This study adopts an efficient and straightforward rank aggregation method called the Borda method to address this problem. Borda's method is a position-based cumulative voting method that assigns a score to an element based on the position where the element appears in each ranking [22]. As a classic voting method first proposed by Jena-Charles de Borda, voters express their preference for candidates in high to low order. The last candidate is assigned 1 point, the second-to-last candidate is given 2 points, and so on, the first-ranked candidate is given q points. The results of all electors are summed up to obtain the Borda score for each candidate, the candidate with the largest point total is the winner. The advantage of Borda's method is that it is fast and efficient. With the Borda method, the rank value of the unlabeled sample is calculated by Equation (3):

$$s(x_i) = \frac{\sum_{l \in L}(|U_s| - \tau_l(x_i))}{q(|U_s| - 1)} \tag{3}$$

where $\tau_l(x_i)$ is the position of the sample x_i in the ranking τ_l. The value of $s(x_i)$ is proportional to the uncertainty of the sample across all labels. Figure 2 shows the block diagram of multi-label sample uncertainty measure based on Borda method.

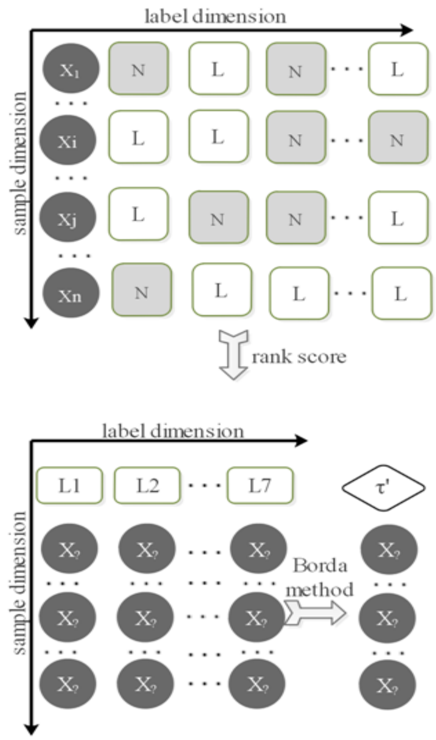


Fig. 2. Borda method based multi-label sample measurement diagram

3.2.2 Measure based on CPQDs label relation

In addition to the uncertainty measure defined based on the Borda ranking aggregation problem, incorporating information from the label space is important. The label–label relation contained in the sample is a crucial problem studied in the multi-label classification [23], and the relation can be divided into correlation and mutual exclusion. Specifically, the multi-label algorithm recognizes that if the label "sea" appears in the picture, the probability that related labels, such as 'blue' and 'ship' will also appear is high. Still, the probability that "high-rise buildings" and "trains" will not occur is high. In the problem of CPQD classification, the label-label relation is mainly reflected in the form of mutual exclusion between class labels in different disturbance groups. According to their definitions and characteristics, this study categorizes seven single disturbance types (e.g., voltage sag, swell, interruption, pulse transient, oscillation transient, harmonic, and flicker). Among them, voltage sags, swells, and interruptions are in the same group because of the characteristics of the amplitude variation. Transient pulses and oscillations are grouped because of the short duration. The harmonic and flicker with long duration and their characteristics are separated [24]. Disturbances in different groups may co-occur, but the disturbance components in the same group cannot co-occur, which is defined as the mutual exclusion of disturbances. Figure 3 shows a schematic of the label prediction of the CPQDs, the relevant label sets the irrelevant label set, the threshold, are the white blocks represent and the black blocks represent, respectively. The predicted class labels contained in the sample cannot be in Groups 1 or 2 simultaneously. By conducting extensive experiments, this study summarizes the following four cases and assigns the corresponding coefficients according to the counts of label exclusions pairs caused by the misclassification. The value of the specific disturbance event evaluation function v is shown in Table 1.

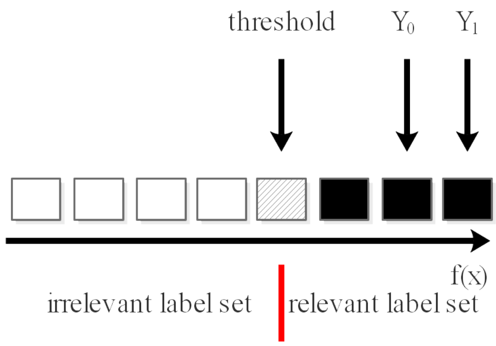


Fig. 3. Prediction of complex PQ disturbances

Table 1. Evaluation score of CPQDs

	Group1			Group2		Coefficient
Disturbance type	Sag	Swell	Interruption	Oscillation transient	Impulsive transient	
Prediction result	✓	✓	✓	✓	✓	0.01
	✓	✓	✓	✓	×	0.1
	✓	✓	✓	×	✓	
	✓	✓	✓	×	×	
	✓	✓	×	✓	✓	0.5
	✓	×	✓	✓	✓	
	×	✓	✓	✓	✓	
	✓	✓	×	×	×	0.8
	✓	×	✓	×	×	
	×	✓	✓	×	×	
	×	×	×	✓	✓	

3.2.3 CPQD active learning strategy

This study adopts the most widely used uncertainty sampling strategy in active learning. Different from conventional uncertainty measure, this study combines the two measure strategies defined in the previous two subsections innovatively. The strategy can be formulated as follows:

$$\underset{x_i \in U_s}{argmax} \quad \frac{s(x_i)}{v(x_i)} \tag{4}$$

where s and v are the measure functions defined in the two previous subsections. This new uncertainty sampling strategy is named label exclusion and ranking scores (LERS). When selecting the most informative sample, LERS not only obtains the ranking score by cumulative voting but also measure the mutual exclusion of disturbance labels in CPQDs.

3.3 CPQD classification based on active ELM

ELM is a novel single-hidden layer feedforward neural network proposed by Professor Huang Guangbin in 2006 [25]. In ELM, the conventional backpropagation process in the neural network was replaced by a simple least square method solution to the output weight. As a sequence, ELM can achieve extremely fast learning speed. Inspired by the successful application of difference vectors of between-class samples in the classification problem [26], we proposed a new multi-label learning algorithm (random discriminative projection extreme learning machine algorithm for multi-label learning, RDPEML) in our previous work [27]. The RDPEML algorithm uses discriminative multi-label inter-class samples to generate a subset of difference vectors and generate hidden layer nodes from them, which improves the random mapping architecture of conventional ELM. RDPEML extends the basic ELM to a multi-label learning domain by incorporating a threshold learning kernel ELM. The input weight and bias in RDPEML are formulated as:

$$\omega = \frac{2(x_{s2} - x_{s1})}{\left\| x_{s2} - x_{s1} \right\|_2^2} \tag{5}$$

$$b = \frac{(x_{s2} + x_{s1})^T (x_{s2} - x_{s1})}{\left\| x_{s2} - x_{s1} \right\|_2^2} \tag{6}$$

The workflow of the CPQD active learning algorithm based on RDPEML is presented as follows:

Input: U_s :unlabeled CPQDs pool; L :label set; **RDPEML:** Φ . j: maximum number of iteration
begin
for iteration= 1 to j do

Step 1: For each sample x_i in U_s , Φ output the corresponding probability vector.

Step 2: Calculate the margin $m_{\Phi}^{i,l}$ on each label l according to the formula (1), and obtain a vector $M_{\Phi}^{i}=\left[m_{\Phi}^{i,1},m_{\Phi}^{i,2},m_{\Phi}^{i,3},\cdots,m_{\Phi}^{i,q}\right]$ of the margin of the sample.

Step 3: Obtain which is the ranking of all unlabeled samples on the label.

Step 4: Calculate the output of measure function by the rank aggregation.

Step 5: Assign the value of according to the predicted results given by and Table 1.

Step 6: Find the most informative sample according to formula (4), add it to the labeled training set, and retrain.

$$M_{\Phi}^{i}=\left[m_{\Phi}^{i,1},m_{\Phi}^{i,2},m_{\Phi}^{i,3},\cdots,m_{\Phi}^{i,q}\right]$$

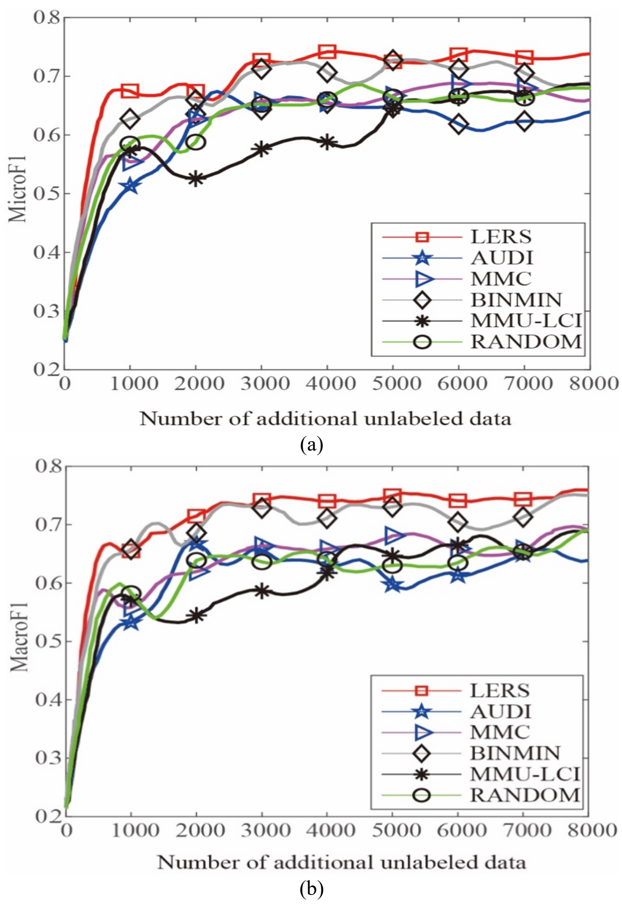
4 Result Analysis and Discussion

4.1 Experimental comparison methods
To the best of our knowledge, this is the first study on an active learning strategy for multi-label problems based on the ELMmodel. Therefore, random sampling (Random), least binary classification (BinMin) [28], maximum marginal prediction uncertainty and label cardinality inconsistency (MMU-LCI) [29], the maximum loss of maximum confidence drop (MMC) [30] and incremental multi-label active learning strategies based on uncertainty and diversity (AUDI) [31] , five of the most relevant active learning strategies were chosen for performance comparison. In order to accomplish a fair comparison, the above strategies are conducted on the same model RDPEML.

4.2 Experimental setup
A total 47 kinds of CPQDs events is generated by MATLAB, and each event is randomly generated 200 times within the range of IEEE1529 parameters. MicroF1 and MacroF1, commonly used multi-label evaluation metrics in active learning literature, are adopted as performance evaluation metrics [32]. At the beginning of the experiment, 5% of the labeled sample set of the training set is randomly selected as the initial labeled pool, and the remaining samples were left as unlabeled pools. In each iteration, the active learning strategy queries and selects the most informative sample to human experts for annotation. In all, 8000 disturbance signal samples are added to the marked sample pool. The remaining 95% of the samples were used to create the unlabeled sample set.

The effectiveness of different active learning strategies can be evaluated by the intuitive vision of the learning curves exhibited by the classifiers running under each strategy. This learning curve is constructed by plotting the evaluation metric as a function of the total number of unlabeled samples queried. A learning strategy outperforms other strategies if it dominates most points of its learning curve.

4.3 Experimental results
Figures 4(a) and Figures 4(b) illustrate the comparison of MicroF1 and comparison of MicroF1 between different active learning strategies. Figures 5 visualize the accuracy curves of the LERS sampling strategy and the other five multi-label sample active learning sampling strategies of Random, BinMin, MMU-LCI, MMC, and AUDI on the Matlab synthetic data set. The horizontal axis of the figure represents the total number of query samples that have been added, and the vertical axis is the evaluation metrics MicroF1 and MacroF1 of RDPEML for unknown CPQDs identification after a number of queries. It can be seen from the figure that after the LERS strategy starts several epochs of query, especially before the number of unlabeled sample queries reaches 2000, the performance learning curves of the two metrics are better than that of the rest strategies. It is obvious that the curve rises more steeply than other sampling strategies. This noticeable improvement is because the LERS strategy incorporates the label exclusion coefficient into the uncertainty measure. The selection of unknown CPQDs that do not comply with the principle of mutual exclusivity plays a crucial role in training a more accurate multi-label classification model. After many rounds of selections, the learning curves of MicroF1 and MacroF1 of LERS also enter a stable stage, this mainly because with the continuous improvement of the learning model, the signals in the unlabeled sample pool that do not meet the mutual exclusion criterion dramatically decrease. From all these results observed, the proposed LERS strategy outperformed the rest of the strategies in almost the entire active learning process, followed by the AUDI method. The BINMIN strategy shows the worst performance in comparing both performance metrics of all strategies.



(a)

(b)

Fig. 4. Comparison of MicroF1 and comparison of MicroF1 between different active learning strategies

To investigate whether the proposed LERS sampling strategy is sensitive to the initial labeled dataset size, the experiments gradually increase the number of initial training data from 1000 to 5000 in steps of 1000. For fixed number of initial label set, the performance of the final classifier after 3000 active learning queries was compared. Figure 5 shows the trend of MicroF1 when initial training labeled datasets size varies. Figure 5 shows that the LERS strategy consistently outperforms nearly all other strategies, except for the initial state of 2000, which is slightly lower than the AUDI strategy. The MicroF1 curve variation of the LERS strategy is minimal for different numbers of initial training labeled datasets, which indicates its characteristics of robustness. Table 2 shows the comparison of the MicroF1 values of each strategy after different iterations. After adding 8000 queried samples, LERS outperforms the MicroF1 values of other algorithms. Table 2 shows that when LERS reaches a MicroF1 value of 0.7271 after adding 3000 unlabeled samples, and AUDI needs to do 5000 queries to obtain a similar performance. This finding shows that under the requirement of this performance, LERS can save approximately 40% manual labeling effort than AUDI.

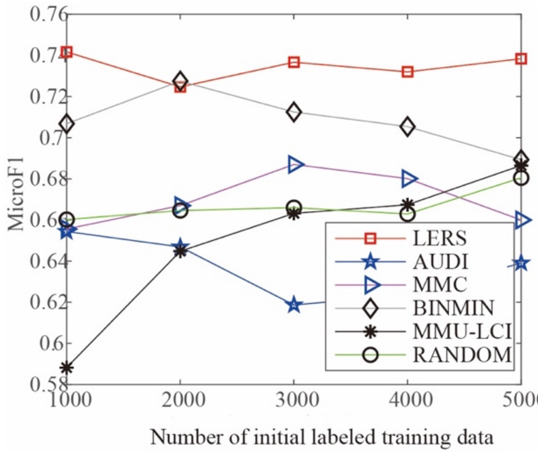


Fig. 5. Comparison of MicroF1 under different numbers of initial training database

Table 2. Experiment results

Additional unlabeled data	Active Learning Strategies					
	LERS	**AUDI**	**MMC**	**BINMIN**	**MMU-LCI**	**RANDOM**
1000	0.6749	0.6275	0.5542	0.5741	0.5127	0.5845
2000	0.6742	0.6595	0.6267	0.5256	0.6354	0.5878
3000	0.7271	0.7126	0.6576	0.5762	0.6436	0.6520
4000	0.7417	0.7068	0.6556	0.5882	0.6544	0.6602
5000	0.7247	0.7275	0.6669	0.6448	0.6469	0.6646
6000	0.7367	0.7125	0.6870	0.6632	0.6187	0.6660
7000	0.7320	0.7054	0.6801	0.6674	0.6235	0.6629
8000	0.7384	0.6893	0.6600	0.6865	0.6391	0.6805

5. Conclusions

To make full use of scarce labeled data with abundant unlabeled data is a challenge when training the CPQDs classification model, a novel active learning strategy LERS strategy was developed to consider label mutual exclusionwas developed to consider label mutual exclusion and ranking score in uncertainty measure process. The LERS was incorporated in the learning phase of RDPEML for further identification. The following conclusions could be drawn as follows:

(1) The LERS sampling strategy that incorporates a label mutual exclusion measure for CPQDs can effectively select the most valuable unknown samples for the classification model improvement and provide them to human experts for labeling when labeling unlabeled CPQDs. Moreover, the additional training data improve the recognition performance of RDPEML for composite disturbances gradually.

(2) When facing different numbers of initial training sets, LERS exhibits a notable robustness. Comparing with other strategies, LERS yields a remarkable reduction in manual labeling cost when reaching certain performance aim.

(3) The experimental results on the synthetic data set reveal that LERS outperforms other strategies in rising trends and obtain the best performance finally.

This study identifies CPQDs based on the active learning strategy LERS, in which the measure of both ranking score and label mutual exclusion are hybridized. The learning model that incorporates LERS strategy has been effectively improved by utilizing unlabeled CPQD samples. This study provides a solution when facing scarce labeled data and abundant unlabeled data in PQ disturbance identification. Considering that LERS select and add unlabeled samples one by one, in future studies, we will focus on researching an efficient active learning strategy for selecting unlabeled samples in batches and using a more effective ranking aggregation method to improve the uncertainty sampling in the active learning process.

References

1. Mahela, O.P., A.G. Shaik, and N. Gupta, "A critical review of detection and classification of power quality events". *Renewable & Sustainable Energy Reviews*, 41, 2015,pp. 495-505.
2. Gao, J., Cui, X., Zou, C.L., and Yang, L., "S-transform based on modified energy concentration and identification of power quality disturbance in random forest". *Electrical Measurement & Instrumentation*, 56(01), 2019,pp.8-14+21.
3. Zhang, W. H., Huang, N. T., Yang, J. C., Yang, Y. J. and Wang, K. X., "Power quality disturbances classification based on generalized S-transform and differential evolution extreme learning machine". *Electrical Measurement & Instrumentation*, 53(20), 2016,pp.50-55.
4. Wang, L. Y., Li, K. C., Xiao, X. Y., Chen, Z., Meng, Q. X., and Cai, D. L., "Classification for multiple power quality disturbances based on sparse decomposition". *Electrical Measurement & Instrumentation*, 55(01), 2018, pp.14-20+33.

5. Liu, J., Song, H., Sun, H. W. and Zhao, H. Y., "High-Precision Identification of Power Quality Disturbances Under Strong Noise Environment Based on FastICA and Random Forest," *IEEE Transactions on Industrial Informatics*, 17(1), 2021, pp. 377-387
6. Li, B., Cao, M., Li, S. L. and Li, C. Y., "Power quality disturbance identification method based on WT and GA-SVM". *Power Electronics*, 54(03), 2020,pp.52-55.
7. Asman, S.H., Abidin, A.F., Yusoh, M.A.T.M and Subiyanto S., "Identification of transient overvoltage using discrete wavelet transform with minimised border distortion effect and support vector machine". *Results in Engineering*, 13, 2022,pp.100311.
8. Kumar, B.N., Manikandan, M.S. and Karanki, S.B., "Integrated Hilbert-Fast Fourier Transform approach for detecting and classifying power quality disturbances". In: *2021 National Power Electronics Conference (NPEC)*, Bhubaneswar, India: IEEE, 2021,pp.01-06.
9. Mahela O.P., Parihar, M., Garg, A.R., Khan, B. and Karmel, S., "A Hybrid Signal Processing Technique for Recognition of Complex Power Quality Disturbances". *Electric Power Systems Research*, 207, 2022,pp. 107865.
10. Swarnkar, N. K., Mahela, O. P., Khan, B. and Lalwani, M., "Identification and Classification of Multiple Power Quality Disturbances Using a Parallel Algorithm and Decision Rules". *Energy Engineering*, 119(2),2022,pp.473–497.
11. Elango, M.K. and Loganathan, K., "Classification of power quality disturbances using Stockwell Transform and Back Propagation algorithm". In: *2016 International Conference on Emerging Technological Trends (ICETT),* Kollam, India: IEEE, 2016,pp.1-7.
12. Elango, M.K., Jagadeesan, A. and Sundaram, K.M., "Power Quality Disturbance Classification Using Wavelet Packet Transform and Back Propagation Algorithm". *Journal of Computational and Theoretical Nanoscience*, 13(10), 2016,pp.7099-7109.
13. Wang Z, Peng Y, Liu Y, Guo, Y., Liu, Y., Geng, H. Y., Li, S. and Chao, F., "Photovoltaic Power Quality Analysis Based on the Modulation Broadband Mode Decomposition Algorithm". *Energies*, 14(23), 2021,pp.7948.
14. Liao, X. H., Xiao, J. B., Wang, Z. G., Wang, H., and Ning, K., "Research on Distributed Power Quality Disturbance Detection Based on ILMD". In: *2019 4th International Conference on Advances in Energy and Environment Research,* Shanghai, China: EDP Sciences, 2019, pp.118.
15. Vidhya S and Kamaraj V., "Complex neural classifiers for power quality data mining". J*ournal of Electrical Engineering and Technology*, 13(4), 2018,pp.1715-1723.
16. Liu, Z. G., Zhang, Q. G., Han, Z. W. and Chen, G., "A new classification method for transient power quality combining spectral kurtosis with neural network". *Neurocomputing*, 125, 2014,pp.95-101.
17. Qu, X. S., Duan, B., Yin, Q. X., Yan, Y. X. and Zhong, Y., "Classification method of power quality disturbances based on deep neural network of sparse auto-encoder". *Electric Power Automation Equipment*, 39(05), 2019,pp.157-162.
18. Wu, S. X., Yang, A. and Zhu, L. J., "Power quality disturbance recognition based on a recurrent neural network". *Power System Protection and Control* ,48(18), 2020,pp.88-94.
19. Zheng, Y., Lin, R. Q., Wang, J. and Li, Z. J., "Power quality disturbance classification based on GAF and a convolutional neural network". *Power System Protection and Control*, 49(11), 2021,pp.97-104.
20. Cohn, D.A., Z. Ghahramani, and M.I. Jordan, "Active learning with statistical models". *Journal of Artificial Intelligence Research*, 4, 1996, pp.129-145.
21. Fu, Y.F., X.Q. Zhu, and B. Li, "A survey on instance selection for active learning". *Knowledge and Information Systems*, 35(2), 2013,pp. 249-283.
22. Ratliff T. C., "A comparison of Dodgson's method and the Borda count". *Econ Theory*, 20, 2002, pp.357–372.
23. Zhang, M.L. and Z.H. Zhou, "A Review on Multi-Label Learning Algorithms". *IEEE Transactions on Knowledge and Data Engineering*, 26(8), 2014,pp.1819-1837.
24. Liu, Z. G., Yan, C., and Li, W. H., "A Classification Method for Complex Power Quality Disturbances Using EEMD and Rank Wavelet SVM". *IEEE Transactions on Smart Grid*, 6(4), 2015, pp.1678-1685.
25. Huang, G.B., Zhu, Q.Y., and Siew, C. K., "Extreme learning machine: Theory and applications". *Neurocomputing*, 70(1-3), 2006,pp. 489-501.
26. Yu, S. , Shan, S. , Chen, X. and Wen, G., "Classifiability-based discriminatory projection pursuit". *IEEE Transactions on Neural Networks*, 22(12), 2011,pp. 2050-2061.
27. Zhao, C., Li, K. C., Li, Y., Wang, L., Luo, Y., Xu, X., Ding, X., and Meng, Q., "Novel Method Based on Variational Mode Decomposition and a Random Discriminative Projection Extreme Learning Machine for Multiple Power Quality Disturbance Recognition". *IEEE Transactions on Industrial Informatics*, 15(5), 2019, pp.2915-2926.
28. Brinker, K, "On Active Learning in Multi-label Classification". In: *Proceeding of From Data and Information Analysis to Knowledge Engineering,* Berlin, Heidelberg: Springer, 2006, pp.206–213.
29. Li, X. and Guo, Y., "Active learning with multi-label SVM classification". In: *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence,* Beijing, China: AAAI Press, 2013, pp.1479-1485.
30. Yang, B. S., Sun, J. T., Wang, T. J. and Chen, Z., "Effective multi-label active learning for text classification". In: *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* Paris, France: ACM, 2009, pp.917-926.
31. Huang, S. and Zhou, Z., "Active Query Driven by Uncertainty and Diversity for Incremental Multi-label Learning". In: *Proceedings of IEEE International Conference on Data Mining,* Dallas, TX, USA: IEEE, 2013, pp.1079-1084.
32. Tsoumakas, G., Katakis, I. and Vlahavas, I., "Mining Multi-label Data". *Data Mining and Knowledge Discovery Handbook*, 2010,pp.667-686.