Proceedings of Proceedings of the 2018 26th International Conference on Nuclear Engineering
ICONE26
July 22-26, 2018, London, England

ICONE26-81645

A SUSTAINABLE UPDATING NUCLIDE IDENTIFICATION SYSTEM BASED ON ONLINE LEARNING

Minzhong Luo
Department of nuclear technology application
China Institute of Atomic Energy
Beijing,102413,China
China
Email: 879671933@qq.com

Li Liu
Reactor Engineering Technology Research Division
China Institute of Atomic Energy
Beijing,102413,China
China
Email: 1920148271@qq.com

ABSTRACT

This paper introduces an automatic nuclide identification system based on online learning algorithm. Preprocessing the data of nuclide energy spectrum to extract feature vectors, which includes matrix transformation and singular value decomposition, and with which a model is trained to identify nuclides. We upgrade this model with wrong examples and the new label produced by unknown nuclide, on the basis of the weighted voting mechanism, and obtain a sustainable self-correction nuclide identification system. This nuclide identification method overcomes the shortcomings of current ones that can only be based on a group of fixed training set and the untrained nuclides can not be identified, more applicative to various spectrometer systems. Through the verification of large data sets of a variety of nuclides such as Co60 and Ba132 and mix nuclide, the identification accuracy can be over 98%*.*

Keywords: nuclide identification, online learning, weighted voting, sustainable updating.

INTRODUCTION

Online learning is a sustainable training and predicting learning algorithm, which has stronger adaptability to dynamic changes and can be constantly corrected. We combine online learning with energy spectrum analysis and propose a new nuclide identification algorithm with high accuracy and reliability. Unlike the traditional peak search algorithm, this algorithm does not require searching and matching energy spectrum peaks, overcomes the vulnerability of classic methods that miss and misunderstand of nuclides;compared with the full spectrum analysis, the identification is rapid;compared with other onefold machine learning algorithms, online learning can be sustainable updated and no longer just rely on a fixed training set.

At present, it begins common to combine artificial intelligence with energy spectrum to identify nuclides in the world. Using BP neural network to identify nuclides has a good accuracy [4], but the model depends more on the selection of training set;BP Neural network recognition algorithm also tends to overfit the training set [5];the BP neural network can also be optimized by particle swarm optimization [7], neverthelessas a result of the neural network can only be performed one-time off-line training and can not be updated, so the identification accuracy is reduced due to the poorer prediction effect on new databesides, there is also an identification algorithm based on Sequential Bayes method [2], but all are uncorrectable off-line training modes.

We use various generated energy spectrum data to train a plurality of machine learning models KMeans as unit predictors to form an expert set and form a sustainable updating nuclide identification system through the online learning with a voting mechanism.

In practical detection environments, the noise is superimposed over the background for the lower energy resolution of commonly used NaI detectors, which is also the reason why traditional peak search algorithms find it difficult to match the weak and overlapping peaks accurately;so we generate a large number of energy spectrum data with different degrees of interference to simulate complex situations of the reality using Monte Carlo program. Energy spectrum can be regarded as a one-dimensional vector, converting the various generated one-dimensional energy spectrum vectors into two-dimensional energy spectrum matrixs, and using SVD decomposition (singular value decomposition) to extract its eigenvectors, the eigenvectors are as training datas and input to the unit predictor and we can get differential unit predictors. The KMeans model is adopted in the unit predictor. When a separate KMeans model is used as the prediction model, it has the defects of limited prediction accuracy and difficulty converging to the optimal solutiontherefore, we construct an expert set with multiple unit predictors and each expert in the expert set is assigned a certain discourse power. In the actual detection environment, the online learning algorithm with voting mechanism receives the energy spectrum data outside the training set, reading the samples and the corresponding real nuclide labels into the training continuously, solving the optimal discourse power of the expert set with the index reduction method and correcting the misidentified nuclides or increasing the new identifiable nuclide tags with using the updating algorithm, and we can get a real-time optimal nuclides identification system.

Data generation

Since the selection of nuclide and spectrum data used for training has a significant impact on the accuracy and reliability of the model, and machine learning algorithm requires a large amount of data for training, training data are generally generated using program simulations [3]. We used the Geant4 particle simulation package to generate the γ spectrum for a variety of nuclides, including Am_{241}, Co_{60}, Er_{175}, Ra_{226}, Ba_{133}, Cs_{137}, I_{131}, Rn_{222}. Geant4 is a package developed by CERN that simulates particle physics and provides the ability to simulate detectors and simulate decay reactions. By writing a program to use the Geant4 interfaces to define the particle decay model, we use Monte Carlo simulation to generate multiple sets of spectrum data. Each piece of spectrum data is a vector(1500 channels) in the form of $E_{1\times1500} = (e_1, ..., e_{1500})$, which will be feature extracted before inputting the model.

Feature extraction

Feature extraction can minimize the size of the sample and maximize the features of different samples, which is very important for the accuracy of machine learning prediction. As the full spectrum data contains a lot of noise and non-characteristic components, we hope to get the eigenvector with as few dimension as possible but containing high-volume information is used as the input to the K-Means model. The method used in this paper is to obtain eigenvectors by SVD.

SVD decomposition is an important matrix factorization in linear algebra. It is widely used in feature extraction and compression storage of matrices. Defined as follows: Suppose M is a $m \times n$ matrix, which belongs to the domain K:$M \in K^{m\times n}$, that is, real or complex. Then there is a decomposition so that

$$M = U\Sigma V^T \tag{1}$$

Where U is a $m{\times}m$ unitary matrix;Σ is a semi-positive definite $m \times n$ diagonal matrix;and V^T, the conjugate transpose of V, Is $n \times n$ unitary matrices. Such decomposition is called the singular value decomposition of matrix M.

First the spectrum vector $E_{1\times1500}$ is transformed into the matrix form $E_{30\times50}$ and then SVD it:

$$E = U\Sigma V^T \tag{2}$$

The first 10 dimensions of the vector form a 1×10 vector as the final eigenvector $x_{1\times10}$, the first 10 singular values in Σ contain more than %90 energy of this spectrum vector [8]. The eigenvector $x_{1\times10}$ can be thought of as a vector containing 10 features of the γ spectrum.

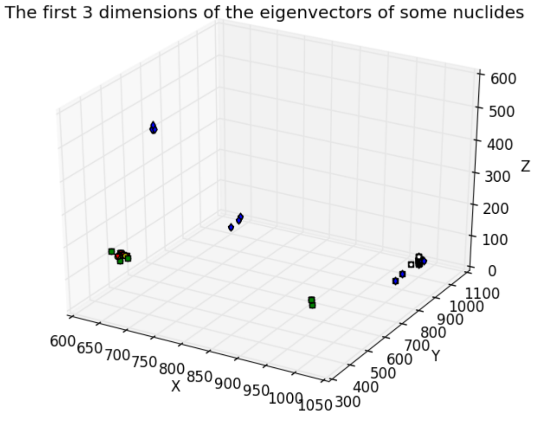

FIGURE 1. The distribution of the first 3 dimensions of the eigenvectors of some nuclides in 3-dimensional space

Now that each eigenvector corresponds to a nuclide lable, we generate 150 group data for each nuclide decay simulation.

Set the set of eigenvectors as X and the set of nuclide labels as y and store the data in csv format as (X_i, y_i) pairs for inputting model training. The visual spatial distribution of the eigenvectors in Figure.1 shows that there is a significant difference in the distribution of eigenvectors from different radionuclides in space.

Unit Predictor

We use KMeans as the unit predictor. The typical kmeans algorithm is a clustering method that classifies the input samples $\{x_i\}$ based on the similarity. For the classification of the radionuclides, the spectrums with similar eigenvectors are clustered to the same group of radionuclide. The kmeans algorithm is described in detail as follows:

Algorithm 1 algorithm Kmeans

```
Input: $X = x_1, x_2 ... x_m$, Number of clusters: $k$
Init:initialize the center point randomly:$U = u_1, u_2 ... u_k$
Output:$U = u_1, u_2 ... u_k$
while Convergence condition was not reached do
    $\forall i \in [k]$:let $C_i = x \in X : i = argmin_j ||x - u_j||$
    $\forall i \in [k]$:update $u_i = \sum_{x \in C_i} / |C_i|$
end while
```

Where k is a user-specified parameter, which is the number of clusters expected. The algorithm selects k initial centers at first and clusters the samples into k groups. In the current nuclide classification task, k is the number of the labels of nuclide.

We train multiple instances of KMeans model, each one called a unit predictor, and each unit predictor uses different batches of training samples to train. The actual KMeans unit predictor is trained as follows:

1) Read in a set of capacity m feature vectors and corresponding (X_i, y_i) data of nuclide labels, which contain k lables;

2) Using KMeans iterative algorithm to cluster, calculate the center vector set $U = \{u_1 ... u_k\}$ of k lables nuclides;

3) Add the trained unit predictor of this round to a set of experts and return to step 1 to train the next unit predictor.

Using the above steps, we trained 10 unit predictors. In order to test their correctness, we make a test dataset in the same format as the training dataset but with a capacity 1/3 of training dataset. Through testing, we found that the accuracy using only KMeans predictor for nuclide identification is about in the 90 or so. In fact, the optimal solution to the KMeans algorithm is NP-hard, so the generally used iterative algorithm can only get approximate solution, which has a negative impact on the accuracy. Especially the problem sensitive to numerical changes like nuclide identification, the approximate solution of using an iterative algorithm deviates from the best center for each class of radionuclides.

We propose two theorems to illustrate that the iterative KMeans algorithm may not converge to the best, and the accuracy of the KMeans algorithm depends on the selection of the initial centers. These two theorems explain why it is difficult to continue to improve the accuracy of the unit predictor.

Theorem 0.1. *The K-means algorithm has suboptimality. Assuming that the optimal objective solution of a K-Means problem is S_{min}, there is an example in which it can find an algorithm with an iterative solution of at least tS_{min}.*

Theorem 0.2. *The convergence point of K-means algorithm is not necessarily a local minimum, which means that K-Means relies on the initial values setting of u, and sometimes can not converge to a local minimum.*

To sum up, if the current unit predictor is used as the final predictor at this moment, there will be the following deficiencies:

1) The current unit predictor employs a strategy that is offline training, which is equivalent to assuming that all predicted dataset will follow the same probability distribution as the training dataset, whereas many detection situations have different background interferences. So this assumption has drawbacks;

2) The unit predictor gives a wrong classification for nuclide lables that are not part of the training set and can not update the nuclide lables that the predictor can recognize;

3) KMeans algorithm has suboptimality, and its accuracy depends on the selection of the initial centers, so it is difficult to get the best solution.

Online Learning System with Experts Voting Mechanism

We construct a set of predictors by training multiple kmeans unit predictors, treating the set of predictors as an abstract set of experts, which predict the nuclide labels through a weighted voting mechanism [6].

In the weighted voting mechanism, each unit predictor in the set of experts is regarded as an abstract expert, who has certain discourse power to express his own prediction suggestions, the finally adopted predictive result depends on comparing the size of each expert's discourse right value;the discourse power is abstracted as weight value, and the weight value is dynamically adjusted according to each expert's correct rate in each round of prediction, that is, the wight value of experts with the higher correct rate is improved, and the wight value of experts with lower correct rate is reduced. we propose the index descent method for weight value adjustment based on the above adjustment principle:

$$w_i = w_i e^{-\eta v_i} \tag{3}$$

w_i is the weight value of the unit predictor f_i, v_i is the penalty value of unit predictor f_i that gives wrong prediction in this round, which measures the deviation from the correct value.

We combine the weighted voting mechanism and the online learning algorithm to construct an optimal prediction system $F(x)$. Online learning is different from the traditional machine learning methods by one-time batch training, which can receive non-quantitative training samples sustainly to update the model dynamically. online learning can be described as follows: suppose the training samples with a certain order as follows:

$$(x_1,y_1),(x_2,y_2)...(x_m,y_m) \quad x_i \in R^d, y_i \in R \tag{4}$$

Where x_i represents the eigenvectors of a sample and y_i represents the nuclide label of a sample. The learning model receives a sample x_i and gives the current predicted value $\hat{y}_i$, and then the model receives the real value y_i to the sample x_i, then the model gives a penalty of false prediction and readjust the model parameters. The whole process is repeated iterative process until all the samples are trained completly. Training samples for batch learning can also be used as training samples for online learning. In the online learning algorithm, model focus on the wrong predicted samples in each prediction round and gives the penalty. A typical online learning algorithm can be described as follows:

Algorithm 2 online Learning

```
Input: $X = x_1, x_2...x_m$, $y = y_1, y_2...y_m$;
Init: initialize the parameters $\theta$ randomly;
Output:parameters $\theta$;
for $t = 1, 2, ...m$ do
    receive $x_t$
    give $\hat{y}_t = f(x_t, \theta)$
    receive real lable $y_t$
    learning model get loss $L(\hat{y}_t, y_t)$
    readjust parameters $\theta$
end for
```

Online learning algorithm is very suitable for the task of spectrum recognition. The task can be described as that a predictor adapts to different decay environments sustainly and improves the classification accuracy. If the model tries to maintain a high classification accuracy under different environmental noise and decay environments, which must have a good generalization. The generalization of the model refers to the accuracy the model $f(x,\theta)$ predicts a sample not in the training dataset. If the model fits too well or too bad for the training sample set X, which will result in poor generalization. If the online learning algorithm was used in the actual decay environments, the $f(x,\theta)$ was able to adjust the parameter θ in real time while receiving the true lable y of nuclide(equivalent to the commissioning phase), predictor $f(x,\theta)$ can be adapt to the current decay environment quickly, can get better generalization quickly.

A common online learning algorithm is with elimination mechanism [1], which can be described as follows. For a set containing multiple predictors $F = \{f_1...f_n\}$, trained by the training samples $(x_1,y_1),(x_2,y_2)...(x_m,y_m)$ iteratively. Each predictor f_i in each round gives the prediction $\hat{y}_i(x_t)$, compared with the true label y_t, if the prediction is incorrect, the predictor will be removed from the set of predictors. Predictors with correctness $\%100$ can be in the set after the iteration completed.

Instead of the eliminate mechanism, we use the online learning algorithm with the voting mechanism because the online learning with the eliminate mechanism can be considered as a special case of the online learning with the voting mechanism(the unit predictor with weight value of zero is equivalent to the unit predictor is eliminated), the online learning algorithm with voting mechanism is more flexible and generalized (the predictor with the correct rate of $\%100$ will be easily over-fitted). The online learning algorithm for nuclide identification with voting mechanism is described as follows:

Algorithm 3 The online learning algorithm for nuclide identification with voting mechanism

```
Input: $X = x_1, x_2...x_T, y = y_1, y_2...y_T$ ;number of experts $d$;
Init: parameters $\eta = \sqrt{2log(d)/T}, \widetilde{w}^{(1)} = (1,...,1)$
for $t = 1, 2, ...T$ do
    let $Z_t = \sum_i \widetilde{w}_i^{(t)}$, $w^{(t)} = \frac{\widetilde{w}^{(t)}}{Z_t}$
    each expert gives the prediction $\hat{y}_i$
    receive the real label $y_t$
    calculate each expert's loss $v_t \in [0,1]^d$
    update the weight value $\widetilde{w}_i^{(t+1)} = \widetilde{w}_i^{(t)} e^{-\eta v_{t,i}}$
end for
```

For the value of η in the algorithm, we propose the theorem:

Theorem 0.3. *In the online learning algorithm for nuclide identification with voting mechanism, $\eta = \sqrt{2log(d)/T}$ is the optimalvalue that minimizes the algorithm error*

To prove this theorem we also need to prove a lemma on the upper bound of loss $\sum_{t=1}^{T} < w^{(t)}, v_t >$ for the weighted voting algorithm:

Lemma 0.1. *When $T > 2log(d)$, the upper bound of the loss of weighted voting algorithm:*

$$\sum_{t=1}^{T} < w^{(t)}, v_t > \leq min_{i\in[d]} \sum_{t=1}^{T} v_{t,i} + \eta T/2 + log(d)/\eta$$

The prediction system can not only correct the wrong cases, but also update the nuclide labels added into the system. Assuming in a decay environment, we find nuclide y that all experts can not recognize(ie, all experts have no lable y), then one of the experts can be updated with the update algorithm to recognize lable y. If the lable y appears in the training set, but all experts predict wrongly, it's also necessary to update it. Algorithm is as follows:

Algorithm 4 update algorithm for experts

```
Input: new sample $(x,y)$;
if $y$ is a new lable then
    select expert $e_i$ with largest $w_i$
    add a new center point $u = x$
    add a new lable $y$
else
    select expert $e_i$ with largest $w_i$
    update the center point $u_y = (u_y|C_y| + x)/(|C_y + 1|)$ for $y$
end if
```

We propose a theorem and prove that it is most reasonable to update the expert who has the largest real-time weight value in the online learning algorithm for nuclide identification:

Theorem 0.4. *For the expert set $E = \{e1...em\}$, if $w_i > w_j$ (that is, the expert e_i has a greater weight value than expert e_j), it's more effective to update e_i.*

The final prediction system is trained by the training samples with different probability distribution as the training samples for unit predictors. The Figure.2 is the error convergence curve of the iterative training of the prediction system, which shows that the online learning system can be adapt to the spectrum data with different probability distribution quickly:

TABLE 1. The prediction system gives the results in the form of probability(Mix nuclide).

Real Nuclide	Predicted Result
Ba133, Cs137	the probability of Er175:[5.238e-07]
	the probability of Ba133, Cs137:[0.999999]
I131, Rn222	the probability of Am241:[4.32144e-08]
	the probability of I131, Rn222 :[0.999999]

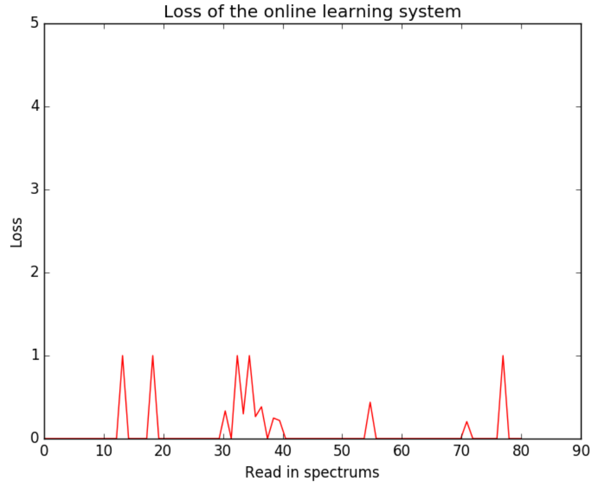

FIGURE 2. The loss of the online learning system in receiving samples and online learning process. Error less than 1 means that some experts give wrong lables. An error of 0 means that all experts predict the lable correctly. An error of 1 means that all experts predict the lable incorrectly. Note that the probability distribution of the sample data during the online learning process is significantly different from the training set when each unit predictor is off-line trained.

The prediction given by the final prediction system is in the form of probability like Table.1, and the probability value of a nuclide y is equal to the sum of the weights of experts giving the same predictive value y. We think that the the lable y with the highest probability gived is the only result of the prediction. The test results from the test dataset has a correct rate of %98.75.

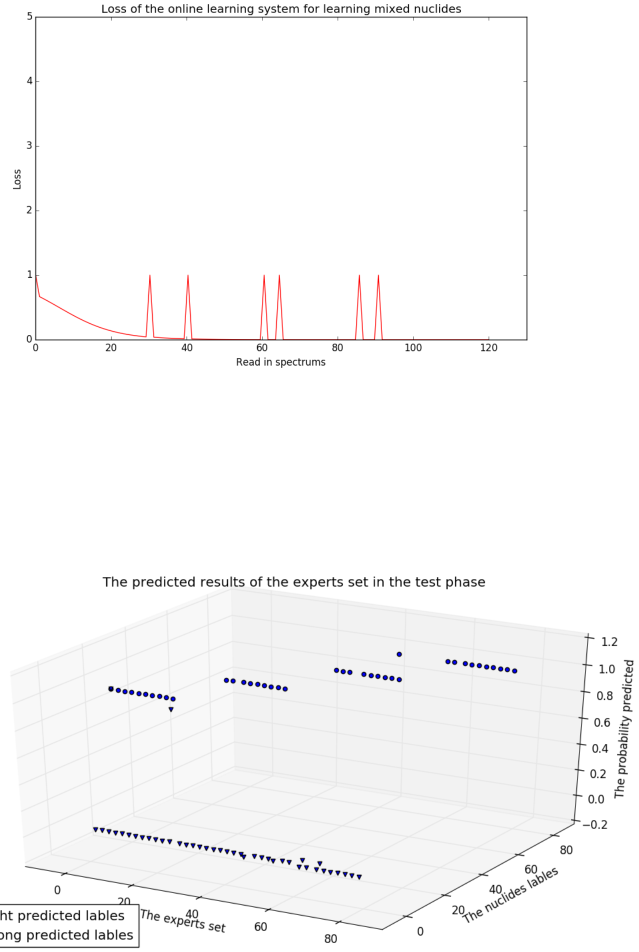

FIGURE 3. Updating and testing of online learning system using mix nuclide data sets. The first figure of Figure.3 shows the loss curve during the online update phase, each loss peak of which corresponding to the mix nuclide lable that can not be recognized by the experts system. After being updated by the update algorithm, the loss descend to 0 gradually, until the next mix nuclide label that system can not recognize is read in. The second figure of Figure.3 shows the prediction of mix radionuclide samples during the testing phase. The prediction given by the system is in the form of probability. It can be found that the probability of wrong samples is close to zero, the probability of correct samples is close to 1. This means that the correctness of the testing is close to 1 after online updating phase.

Conclusion

By feature extraction based on SVD, the eigenvector of spectrum matrix after dimension compression is obtained as feature input for spectrum identification. The low dimensional input reduces the amount of computation. The unit predictor uses the unsupervised learning algorithm Kmeans. We combine multiple unit predictors into a set of experts and use online learning with voting mechanism to overcome the suboptimality and dependence on initial value of the Kmeans iterative algorithm, and the prediction system based on the online learning with real-time correction can immediately adapt to different decay environments, always maintains high accuracy and high reliability. We tried to remove Am_{241} and Ba_{133} from the training dataset while training the unit predictor and then give samples of Am_{241} and Ba_{133} during the online learning phase, to test whether the learning system can learn to identify these two nuclides in real time. The test result shows that after receiving the samples of Am_{241} and Ba_{133}, the learning system updates the unit predictor and can identify the corresponding nuclides correctly. The accuracy rate of the final online learning prediction system is above $\%98$.

We change the dimension of the SVD eigenvector extracted from the spectrum feature, in extreme cases, the accuracy of prediction system is changed by 0.05, 0.05, -0.05 and -0.05 respectively when the dimension 10 is changed to 9, 8, 11 and 12 dimensions. The current deficiency is that there are only 8 types of radionuclides to be used, and more radionuclides can be added to the training set to improve the identification breadth of the prediction system.

For the identification of multiple nuclides, we have produced the spectrum dataset of mix nuclides(Am_{241} and Co_{60} are mixed, Er_{175} and Ra_{226} are mixed, Ba_{133} and Cs_{137} are mixed, I_{131} and Rn_{222} are mixed). Before the online learning phase, the unit predictor does not have the ability to predict the mix nuclides. In the online learning phase, the prediction system receives mix nuclide data and updates the parameters according to the update algorithm. Figure.3 shows the details of the online learning system during the updating phase and the testing phase.

The nuclide identification system updated sustainablely based on online learning uses C++ to develop and compiled under Linux, which is suitable for porting to various embedded devices.

REFERENCES

[1] Jacob Abernethy, Peter Bartlett, Alexander Rakhlin, and Ambuj Tewari. Optimal strategies and minimax lower bounds for online convex games [technical report no. ucb/eecs-2008-19]. 2008.

[2] J. V. Candy, E. Breitfeller, B. L. Guidry, D. Manatt, K. Sale, D. H. Chambers, M. A. Axelrod, and A. M. Meyer. Physics-based detection of radioactive contraband: A sequential bayesian approach. *IEEE Transactions on Nuclear Science*, 56(6):3694–3711, 2009.

[3] R.B.Patterson C.Backhouse. Library event matching event classification algorithm for electron neutrino interactions in the noa detectors. 2015.

[4] K.Shizuma E.Yoshida and T.Oka S.Endo. Nucl. instr. and meth. a 484. 2002.

[5] Han Xuli. Hou Muzhou. Multivariate numerical approximation using constructive rbf neural network. 2012.

[6] Nick Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. 1988.

[7] Song Yongqiang Liu Yicong, Zhu Hongguang. Optimize bp neural network by an improved particle swarm optimization to implement nuclide identification. 2016.

[8] Wang Kunpeng Ren Junsong, Zhang Jiangmei. Radioactive nuclide identification method based on svd and svm. 2017.

Appendix A: The Source Code

All the source code of this paper, including the data-make programm and the algorithm programm, is put on the GitHub;

https://github.com/Luomin1993/ICONE2018-NuclideIdentify.

Appendix B: Proof

Theorem .5. *The K-means algorithm has suboptimality. Assuming that the optimal objective solution of a K-Means problem is S_{min}, there is an example in which it can find an algorithm with an iterative solution of at least tS_{min}.*

Proof. This is an existential problem, and we only need to find an instance that satisfies the conditions described by this theorem. let $X = x_1, x_2, x_3, x_4 \in R^2$, and $x_1 = (0,0), x_2 = (0,2), x_3 = (2a,0), x_4 = (2a,2)$, $a > 1$, using L_2 norm, and let $k = 2$; Now use the iterative KMeans algorithm, which means that it may stop after the first iteration and any subsequent iteration, and in the first-step, let$u_1 = x_1, u_2 = x_2$; the first-step:x_1 and x_3 are clustered, x_2 and x_4 are clustered;then $u_1 = (a,0)$, $u_2 = (a,2)$; now $S = 4a^2$; But the optimal solution:x_1 and x_2 are clustered, x_3 and x_4 are clustered;and $u_1 = (0,1)$, $u_2 = (2a,1)$; now $S = S_{min}a^2$ and $a > 1$, So it's proved.

Theorem .6. *The convergence point of K-means algorithm is not necessarily a local minimum, which means that K-Means rely on the initial values setting of u, and sometimes can not converge to a local minimum.*

Proof. This is an existential problem, and we only need to find an instance that satisfies the conditions described by this theorem. let $X \in R$, That is, the sample is one-dimensional, let $X = \{1,2,3,4\}$ and $k = 2$, using L_1 norm, and x_i is divided into the category:

$$j = argmin_j ||x_i - u_j||$$

We also need to add a limit, when $| \quad argmin_j ||x_i - u_j|| \quad | > 1$, that is, when there is more than one tag that satisfies the minimum distance, we select the tag with the smaller tag value as the final category:

$$j = min\ argmin_j ||x_i - u_j||$$

Now let $u_1 = 2$, $u_2 = 4$, then the final convergence solution is:

$$u_1 = 2 \quad C_1 = \{1,2,3\} \quad u_2 = 4 \quad C_2 = \{4\}$$

Now the target value $S = 2$; But the optimal solution is:

$$u_1 = 1.5 \quad C_1 = \{1,2\} \quad u_2 = 3.5 \quad C_2 = \{3,4\}$$

Now the target value $S_{min} = 2$;

Theorem .7. *For the expert set $E = \{e1...em\}$, if $w_i > w_j$(that is, the expert e_i has a greater weight value than expert e_j), it's more effective to update e_i.*

Proof. Using inverse proof, suppose now that an update to e_j is performed, then e_i cannot identify the sample with y_{new} correctly. So for the next sample y_{new}, the expert e_i cannot recognize correctly, Its weight suffers a loss: $w_i = w_i e^{-\eta v_i}$ Repeat the previous step until $w_i < w_j$, which equates to updating the e_i with greater weight. So it's proved.

For the value of η in the algorithm, we propose the theorem:

Theorem .8. *In the online learning algorithm for nuclide identification with voting mechanism, $\eta = \sqrt{2log(d)/T}$ is the optimal value that minimizes the algorithm error*

To prove this theorem we also need to prove a lemma on the upper bound of loss $\sum_{t=1}^{T} < w^{(t)}, v_t >$ for the weighted voting algorithm:

Lemma .2. *When $T > 2log(d)$, the upper bound of the loss of weighted voting algorithm:*

$$\sum_{t=1}^{T} < w^{(t)}, v_t > \leq min_{i \in [d]} \sum_{t=1}^{T} v_{t,i} + \eta T/2 + log(d)/\eta$$

Proof. let $Z_t = \sum_i \widetilde{w}_i^{(t)}$ Since $\forall i$, $\widetilde{w}_i^{(t+1)} = \widetilde{w}_i^{(t)} e^{-\eta v_{t,i}}$ we know $Z_{t+1} \leq Zt$ so:

$$log(\frac{Z_{t+1}}{Z_t}) = log(\sum_i \frac{\widetilde{w}_i^{(t)}}{Z_t} e^{-\eta v_{t,i}}) = log(\sum_i w_i^{(t)} e^{-\eta v_{t,i}})$$

let $\eta v_{t,i} = s$, and we know $s \in (0,1)$, and the inequality $e^{-s} \leq 1 - s + s^2/2$ and $\sum_i w_i^{(t)} = 1$ we obtain:

$$log(\frac{Z_{t+1}}{Z_t}) = log(\sum_i w_i^{(t)} e^{-s})$$

$$\leq log(\sum_i w_i^{(t)} (1 - s + s^2/2))$$

$$\leq log(1 - \sum_i w_i^{(t)} (s - s^2/2))$$

now let $s - s^2/2 = m$, and according to $m \in (0,1)$, we infer $1 - m < e^{-m}$, so we obtain $log(1-m) \leq -m$, so:

$$log(\frac{Z_{t+1}}{Z_t}) \leq log(1-m) \leq -m$$

$$= -\sum_i w_i^{(t)} (\eta v_{t,i} - \eta^2 v_{t,i}^2/2)$$

$$= -\eta < w^{(t)}, v_t > + \eta^2 \sum_i w_i^{(t)} v_{t,i}^2/2$$

$$\leq -\eta < w^{(t)}, v_t > + \eta^2/2$$

Using the chain rule, we obtain:

$$\sum_{t=1}^{T} log(\frac{Z_{t+1}}{Z_t}) = logZ_{T+1} - logZ_1$$

$$\leq \eta \sum_{t=1}^{T} < w^{(t)}, v_t > + \eta^2 T/2$$

Now we try to find the lower bound of Z_{T+1}, Using the chain rule, we obtain:

$$\widetilde{w}_i^{(t+1)} = \widetilde{w}_i^{(t)} e^{-\eta v_{t,i}} = ... = e^{-\eta \sum_t v_{t,i}}$$

$$logZ_{T+1} = log(\sum_i e^{-\eta \sum_t v_{t,i}})$$

$$\geq log(max_i e^{-\eta \sum_t v_{t,i}}) = -\eta min \sum_t v_{t,i}$$

At the last, according to $logZ_1 = logd$ we obtain:

$$-\eta min \sum_t v_{t,i} - logd \leq -\eta \sum_{t=1}^{T} < w^{(t)}, v_t > + \eta^2 T/2$$

Both sides divided by η. So it's proved.

Proof. for:

$$\sum_{t=1}^{T} < w^{(t)}, v_t > - min_{i \in [d]} \sum_{t=1}^{T} v_{t,i} \leq \eta T/2 + log(d)/\eta$$

$min_{i \in [d]} \sum_{t=1}^{T} v_{t,i}$ is the minimal cost that the algorithm can achieve, Then reduce the upper bound of the right, using the mean inequality:

$$\eta T/2 + log(d)/\eta \geq \eta T$$

and $\eta = \sqrt{2log(d)/T}$. So it's proved.