A Quantum Classifier Based Active Machine Learning for Intelligent Interactive Service

Jiamin Cheng, Lei Chen(✉), and Ping Cui

Jiangsu Province Key Laboratory of Intelligent Industry Control Technology, Xuzhou University of Technology, Xuzhou 221000, China
chenlei@xzit.edu.cn

Abstract. The response time of interactive services depends not only on network latency, but also on computer time. Active learning algorithms are the most important methods. One problem is that these algorithms with uncertain sampling strategies propose an active learning sampling strategy on the basis of sample error correction to ensure that the efficiency and accuracy of interactive information calling are improved, and they have high computational complexity. However, due to computational complexity, this method is only suitable for smaller data sets. This article discusses the use of quantum clusters to accelerate calculations

Keywords: Response time · Quantum classifier · Active learning · Machine learning · Human–computer interaction

1 Introduction

With the acceleration of mobile phone access, the number of Internet users accessing the Internet through mobile devices is increasing rapidly, and the amount of computing is shifting to the cloud [1]. Services such as online translation, voice cloud, and information push are information calls. Based on large fuselage. The limitations of natural language recognition technology make it difficult to accurately retrieve large text databases. It is necessary to strengthen the search target through interactive query. Realizing the above interactive information search is an active learning strategy. Active learning will ask questions instead of passively accepting knowledge, which will increase the applicability of the next interaction according to the reaction tank. Active learning helps to find high-quality samples in the sample room and accurately describe user needs.

Due to the development of communication technology between the core network side and mobile terminal users [2–4], the user experience of traditional services has been improved on the basis of the new scheduling scheme [5–8]. Traffic analysis [9–12], a new planning strategy aimed at improving network performance indicators (such as resource use [13–16], frequency efficiency [17–20], energy efficiency [21, 22]). Measuring the performance of these new planning strategies can determine traffic flow. Therefore, the reconstruction of business flow has become an important topic [23–25]. In dense user communication, the effective capacity of the channel must be improved [26–30]. Within

H. Song and D. Jiang (Eds.): SIMUtools 2020, LNICST 370, pp. 323–332, 2021.
https://doi.org/10.1007/978-3-030-72795-6_26

the framework of the intelligent scheduling scheme, the accuracy of prediction starts from [3, 31, 32], and the traffic volume will seriously affect the effective capacity. Increasing the network performance index can improve the quality of these traditional services. In some emerging interactive services, the quality of human-computer interaction not only depends on communication performance. For these services, users have higher requirements for the delay and accuracy of such information country query services. In order to improve the quality of user experience, certain information request services put forward higher requirements for the accuracy and speed of calls. The limitation of computational cost and interaction time is the two main challenges of active learning in interactive information search.

Active learning is a continuously evolving interactive learning method. Lead to the uncertain sampling strategy in [33]. In their classification experiment, it will be a selection strategy that cannot be determined for each query problem, combined with a probabilistic classifier and satisfactory results. The above-mentioned active learning strategy has become a common learning method. Machine learning strategies are gradually being developed using different learning algorithms in the fields of natural language processing and information retrieval [34]. Common active learning strategies include committee questions, marginal sampling, and verifiability [35, 36]. These methods have produced good results in the field of information acquisition.

However, to obtain a good user experience in the interactive information search, building a high-quality training toolkit and reducing the computational complexity of the text classifier are two key issues. Active learning algorithms are the key method for learning interactive information. In recent years, deep drilling in the field of mechanical engineering has had huge practical advantages. However, Carrier Vector Machine (SVM) provides a solid theoretical foundation. Combined with the support vector machine, based on the analysis of the uncertainty and influence of the sample on the version space, several active learning strategies have been developed, including simple margins, maximum and minimum ranges and quotation ranges [37]. This is a hot topic. Therefore, the influence of labeled samples on the classification model is very important. It is very important to choose the most informative example [38]. Due to the large amount of storage space and time, SVM cannot support online services. Although multiplexing can shorten the training time [39], it will face high computational load under a larger gap.

The method of reinforcement learning was proposed in [40]. Since the latent semantic model is becoming more and more popular in text classification [42, 43], it has been widely used in the field of text classification [41]. Their version 2 of AdaBoost.MHAufgrund is more suitable for online information consulting environment. However, it is necessary to build a huge training set to increase the number of interactions and deteriorate the user's online service experience.

Interactive retrieval services require active learning algorithms with low complexity and high configuration Zision. The key to improving the accuracy of active learning algorithms is to select samples with a large amount of information [44]. The traditional active learning algorithm first selects a representative initial training set from cluster analysis [45], and then marks it. The most uncertain sample. In this case, the learning process will run through the initial set. In the case of limited initial information: In the early learning stage, the classifier often "misunderstands" the retreat target, which requires us to spend a lot of time. A method is proposed on AdaBoost.MHVersion [46]. In this research, we discussed the application of quantum mechanisms to the active learning model proposed in [47]. accelerate. This article is divided into five parts. The second part reviews the interactive retrieval method proposed in Reference [47]. The third part discusses: How to combine the retrieval method with the quantum classifier proposed in Reference [48]. The fourth part is the conclusion.

2 Survey on Interactive Methods for Querying Information

Interactive information retrieval has the following three characteristics: (1) There is no sample selected first. (2) Through interactive selection of high-quality samples, a high-quality training set is established to form a high-precision classifier. (3) Services are sensitive to delays; high-quality training courses must be established within a limited time.

In order to meet these challenges, the interactive retrieval algorithm can be designed as follows:

Step 1) The user sends query information.
Step 2) Calculate the degree of association between the query information and different documents, and create an association model.
Step 3) Use active learning strategies to improve the association model through multiple interactions with users.
Step 4) Sort the documents according to the association model, and output the query results in descending order of association level.

The interactive retrieval system model can be expressed as $A = (C, L, S, Q, U)$, where C is classifier, $L = R_t \cup N_t$, R_t and N_t are relevant and irrelevant document set which are identified through interaction, S is classifier, Q is evaluation function that acquires simplified training set by identifying high-value samples, and U is the unlabeled sample set and evaluation object of Q.

Let C_L be the classifier formed by using L as the training set, and $C_L(x)$ be the classifier calculated value of sample x. In binary classification, classification results are generally judged by sign of values. If $x \in U$, $|C_L(x)|$ is proportional to the certainty degree of categorization of x and samples with low certainty degree are usually chosen into the training set in active learning. Additionally, if $y(x)$ is the real category of sample x, the $\left|\left\{C_L^{'}, \forall C_L^{'}(x) = y(x)\right\}\right|$ is size of the version size (where $C_L^{'}$ is set of all candidate classifiers), the samples that can reduce the version space to the greatest extent can be used as high-value samples.

If these samples are recognized as different categorization by human, namely $y \cdot C_L(x) < 0$, then the high value of $|C_L(x)|$ means the high error-correcting capacity to the classifier C_L. As x is added, the new training set $C_{L\cup\{x\}}$ will be closer to real user demands. Samples that can correct classifiers have higher values when interactions are few and the early cognition of classifier significantly deviates from the retrieval target.

To evaluate the unlabeled samples' capacity to correct classifier, we designed an expression to select the high quality samples:

$$\alpha \cdot po \cdot \left(\frac{(S_{Max} - S_d)}{(S_{Max} - S_{Min})} \right) + \beta \cdot ne \cdot \left(\frac{(S_d - S_{Min})}{(S_{Max} - S_{Min})} \right) \quad (1)$$

where α and β are coefficients, po is expected contribution (correcting capacity) of documents judged as positive sample to the classifier, ne is expected contribution (correcting capacity) of documents judged as negative sample to the categorization, S_d is the score given by classifier to the current document d (a higher score means a higher expectation for the document belonging to positive sample and a lower score translates to lower expectation), S_{Max} and S_{Min} are the highest and lowest scores of the classifier for unlabeled documents.

In Eq. (1), $\frac{S_{Max} - S_d}{S_{Max} - S_{Min}}$ reflects the probability for document to be judged by the current classifier C_L as positive samples and $\frac{S_d - S_{Min}}{S_{Max} - S_{Min}}$ is the probability for document to be judged by the current classifier C_L as negative sample.

To each unlabeled document sample, the calculation formula determining its contribution coefficients (correcting capacity) for current classifier:

$$po = \sum_{\forall w \in W} c(w) \cdot idf(w) \quad (2)$$

and

$$ne = \sum_{\forall w \notin W} tf{-}idf(w, d), \quad (3)$$

where $c(w)$ is the relevancy between term w given by the classifier and the target query document. This score can measure the consistence between sample and the retrieval target. And W is a set of key terms in the current document d. Let D be the document set, $d \in D$ is the current document and $Tr \subset D$ is the labeled document set. Let $|Tr|$ stand for total number of labeled document, $\#Tr(w)$ stand for number of labeled documents containing the word w and $\#(w, d)$ is the frequency of w in document d. The calculation formula of the idf function becomes $idf(w) = \log\left(|Tr| \big/ \#Tr(w)\right)$, so the $tf - idf$ function formula is $tf - idf(w, d) = (\#(w, d)) \cdot idf(w)$.

3 Redesign Method Based on a Quantum Enclosure

This part is explained by the mechanical authentic. By Please write that model of the quantum capsule FOI proposed in [48]. Let A_0 is the hypothesis that the sample does not belong to the wanted documents, and A_1 is the hypothesis that the document contains some wanted content. The choice is represented as a value in the interval [0, 1]. The decision is A_0 for $\Delta = 0$ and A_1 for $\Delta = 1$, respectively.

The detection Δ operator yields 1 under A_0 with the following probability:

$$Q_0 = P(\Delta = 1 | A_0) = tr(\rho_0 \Delta). \quad (4)$$

$$Q_1 = P(\Delta = 1|A_1) = tr(\rho_1 \Delta). \tag{5}$$

The average cost can be written as

$$\bar{K} = \xi K_{00} + (1-\xi)K_{01} - (1-\xi)(K_{01} - K_{11})tr(\rho_1 - \lambda\rho_0)\Delta \tag{6}$$

where

$$\lambda = \frac{\xi(K_{10} - K_{00})}{(1-\xi)(K_{01} - K_{11})}. \tag{7}$$

If $K_{01} > K_{11}$, $\bar{K}$ will be minimum if $tr((\rho_1 - \lambda\rho_0)\Delta)$ can be maximized.

The best detection operator is provided by the eigenstates $|e_l\rangle$ if the operator $\rho_1 - \lambda\rho_0$ corresponding to the positive eigenvaues, where the eigensystem is provided by

$$(\rho_1 - \lambda\rho_0)|e_l\rangle = e_l|e_l\rangle \quad l = 1, ..., (\text{rank of } \rho_1 - \lambda\rho_0) \tag{8}$$

So it is essential to maximize

$$tr(\rho_1 - \lambda\rho_0)\Delta = \sum\nolimits_l e_l\langle e_l|\Delta|e_l\rangle \tag{9}$$

and this can be obtained if

$$e_k\langle e_l|\Delta|e_l\rangle = 1, e_l \geq 0 \quad \text{and} \quad e_l\langle e_l|\Delta|e_l\rangle = 0, e_l < 0 \tag{10}$$

The estimation of the optimal projection operator between A_0 and A_1 can be written as

$$\Delta = \sum\nolimits_{l:e_l \geq 0} |e_l\rangle\langle e_l| \tag{11}$$

So the error probabilities is given in [48]:

$$Q_0 = \sum\nolimits_{l:e_l \geq 0} \langle e_l|\rho_0|e_l\rangle \quad \text{and} \quad Q_1 = 1 - \sum\nolimits_{l:e_l \geq 0} \langle e_l|\rho_1|e_l\rangle \tag{12}$$

In [48], the minimum average cost is formulated as:

$$K^{-}_{min} = \xi K_{00} + (1-\xi)K_{01} - (1-\xi)(K_{01} - K_{11})\sum\nolimits_{l:e_l > 0} e_l \tag{13}$$

Consider the vector $|y\rangle$, which is the input query of an IR system.

In this model, each document can being represented as a word vector $|x\rangle$. The density operator ρ_0 can assess the relevance between unlabelled sample and negative sample. The density operator ρ_1 can assess the relevance between unlabelled sample and positive sample. The quantum SDT projects both the query vector $|y\rangle$ and the document vectors $|x\rangle$ by means of the optimal detection. The ranking is computed by:

$$\langle x|\Delta|y\rangle \tag{14}$$

The operators ρ_0 and ρ_1 are given as [48]:

$$\rho_0 = \frac{|v_0\rangle\langle v_0}{tr(|v_0\rangle\langle v_0)} \qquad \rho_1 = \frac{|v_1\rangle\langle v_1}{tr(|v_1\rangle\langle v_1)} \tag{15}$$

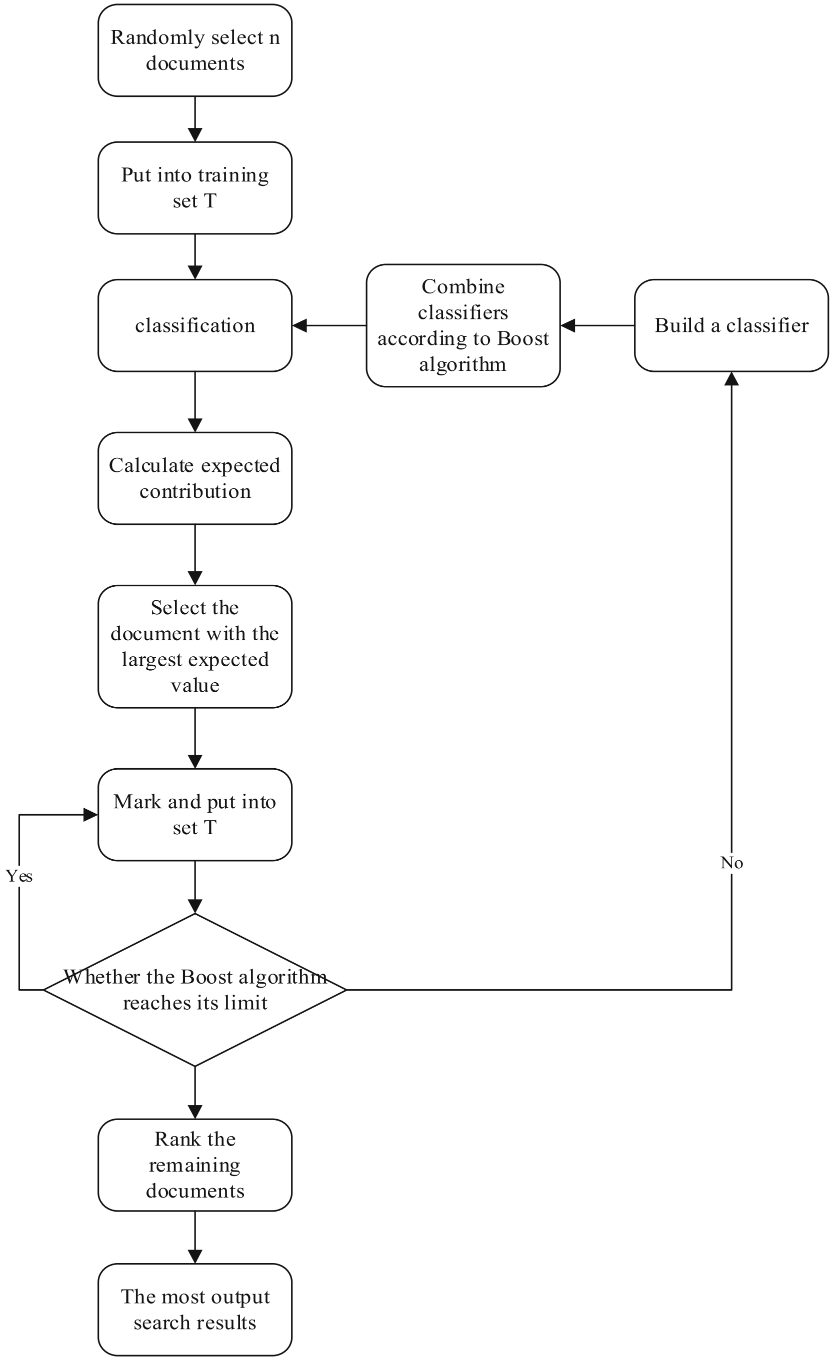

Fig. 1. Flow chart of MH active learning method

In combination with Eq. (1) and Eq. (15), the potential proportion of uncertain samples is defined. Most uncertain sample and the sample determined, if marked in the unexpected class, will provide important information.

An interactive retrieval algorithm is developed by combining the Boost algorithm with the classifier with low computational complexity, using the active learning strategy based on the error correction capability, incrementally increase the selected samples of the users and the performance of the classifier to improve. Algorithms and document sets can be used in the cloud are. To reduce the delay in interaction, adopting AdaBoost. As shown in the Fig. 1, MH Active learning methods can be redesigned as follows:

Step 1) randomly select N documents for the training set t.
Step 2) a classifier by Eq. (15).
Step 3) Combine classifiers by Boost algorithm.
Step 4) Use combination classifier to classify unlabeled documents.
Step 5) calculates the expected contribution of each unmarked document in this loop by formula (1).
Step 6) Select the document with the highest expected contribution
Step 7) provide the selected patterns for identification to the user and insert them into set t.
Step 8) If the Boost algorithm does not reach the iteration limit, return to step 2, otherwise go to step 7.
Step 9) Sort the remaining documents by Eq. (14).
Step 10) publish most search results.

4 Conclusions

Active learning is an effective way to achieve the performance of interactive information search improve. The number of iterations and response delay affect the end user's experience. We are using concept of expected contribution to evaluate the quality of the samples, to the number of iterations to reduce. Quantum-transformation is used to react the classifier to accelerate. Combined these two methods suggest an active learning algorithm.

Acknowledgements. This work is partly supported by Jiangsu technology project of Housing and Urban-Rural Development (No. 2018ZD265) and Jiangsu major natural science research project of College and University (No. 19KJA470002).

References

1. Toumi, H., Brahmi, Z., Benarfa, Z., Gammoudi, M.M.: Server load prediction using stream mining. In: 31st International Conference on Information Networking, Da Nang, Vietnam, pp. 653–661. IEEE (2017)
2. Wang, F., Jiang, D., Qi, S.: An adaptive routing algorithm for integrated information networks. China Commun. **7**(1), 196–207 (2019)

3. Huo, L., Jiang, D., Lv, Z., et al.: An intelligent optimization-based traffic information acquirement approach to software-defined networking. Comput. Intell. **36**, 151–171 (2020)
4. Zhang, K., Chen, L., An, Y., et al.: A QoE test system for vehicular voice cloud services. Mob. Netw. Appl. (2019). https://doi.org/10.1007/s11036-019-01415-3
5. Barakabitze, A.A., et al.: QoE management of multimedia streaming services in future networks: a tutorial and survey. IEEE Commun. Surv. Tutor. **22**(1), 526–565 (2020)
6. Orsolic, I., Skorin-Kapov, L.: A framework for in-network QoE monitoring of encrypted video streaming. IEEE Access **8**, 74691–74706 (2020)
7. Song, E., et al.: Threshold-oblivious on-line web QoE assessment using neural network-based regression model. IET Commun. **14**(12), 2018–2026 (2020)
8. Seufert, M., Wassermann, S., Casas, P.: Considering user behavior in the quality of experience cycle: towards proactive QoE-aware traffic management. IEEE Commun. Lett. **23**(7), 1145–1148 (2019)
9. Chen, L., Jiang, D., Bao, R., Xiong, J., Liu, F., Bei, L.: MIMO scheduling effectiveness analysis for bursty data service from view of QoE. Chin. J. Electron. **26**(5), 1079–1085 (2017)
10. Jiang, D., Wang, Y., Lv, Z., et al.: Big data analysis-based network behavior insight of cellular networks for industry 4.0 applications. IEEE Trans. Ind. Inform. **16**(2), 1310–1320 (2020)
11. Jiang, D., Huo, L., Song, H.: Rethinking behaviors and activities of base stations in mobile cellular networks based on big data analysis. IEEE Trans. Netw. Sci. Eng. **1**(1), 1–12 (2018)
12. Chen, L., Jiang, D., Song, H., Wang, P., Bao, R., Zhang, K., Li, Y.: A lightweight end-side user experience data collection system for quality evaluation of multimedia communications. IEEE Access **6**(1), 15408–15419 (2018)
13. Chen, L., Zhang, L.: Spectral efficiency analysis for massive MIMO system under QoS constraint: an effective capacity perspective. Mob. Netw. Appl. (2020). https://doi.org/10.1007/s11036-019-01414-4
14. Wang, F., Jiang, D., Qi, S., et al.: A dynamic resource scheduling scheme in edge computing satellite networks. Mob. Netw. Appl. (2019). https://doi.org/10.1007/s11036-019-01421-5
15. Jiang, D., Huo, L., Lv, Z., et al.: A joint multi-criteria utility-based network selection approach for vehicle-to-infrastructure networking. IEEE Trans. Intell. Transp. Syst. **19**(10), 3305–3319 (2018)
16. Jiang, D., Zhang, P., Lv, Z., et al.: Energy-efficient multi-constraint routing algorithm with load balancing for smart city applications. IEEE Internet Things J. **3**(6), 1437–1447 (2016)
17. Lee, Y., Kim, Y., Park, S.: A machine learning approach that meets axiomatic properties in probabilistic analysis of LTE spectral efficiency. In: 2019 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, South Korea, pp. 1451–1453 (2019)
18. Ji, H., Sun, C., Shieh, W.: Spectral efficiency comparison between analog and digital RoF for mobile fronthaul transmission link. J. Lightwave Technol. **38**, 5617–5623 (2020)
19. Hayati, M., Kalbkhani, H., Shayesteh, M.G.: Relay selection for spectral-efficient network-coded multi-source D2D communications. In: 2019 27th Iranian Conference on Electrical Engineering (ICEE), Yazd, Iran, pp. 1377–1381 (2019)
20. You, L., Xiong, J., Zappone, A., Wang, W., Gao, X.: Spectral efficiency and energy efficiency tradeoff in massive MIMO downlink transmission with statistical CSIT. IEEE Trans. Signal Process. **68**, 2645–2659 (2020)
21. Jiang, D., Li, W., Lv, H.: An energy-efficient cooperative multicast routing in multi-hop wireless networks for smart medical applications. Neurocomputing **220**, 160–169 (2017)
22. Jiang, D., Wang, Y., Lv, Z., et al.: Intelligent optimization-based reliable energy-efficient networking in cloud services for IIoT networks. IEEE J. Sel. Areas Commun. (2019)
23. Jiang, D., Wang, W., Shi, L., et al.: A compressive sensing-based approach to end-to-end network traffic reconstruction. IEEE Trans. Netw. Sci. Eng. **5**(3), 1–12 (2018)

24. Jiang, D., Huo, L., Li, Y.: Fine-granularity inference and estimations to network traffic for SDN. PLoS ONE **13**(5), 1–23 (2018)
25. Huo, L., Jiang, D., Qi, S., et al.: An AI-based adaptive cognitive modeling and measurement method of network traffic for EIS. Mob. Netw. Appl. (2019). https://doi.org/10.1007/s11036-019-01419-z
26. Guo, C., Liang, L., Li, G.Y.: Resource allocation for low-latency vehicular communications: an effective capacity perspective. IEEE J. Sel. Areas Commun. **37**(4), 905–917 (2019)
27. Shehab, M., Alves, H., Latva-aho, M.: Effective capacity and power allocation for machine-type communication. IEEE Trans. Veh. Technol. **68**(4), 4098–4102 (2019)
28. Cui, Q., Gu, Y., Ni, W., Liu, R.P.: Effective capacity of licensed-assisted access in unlicensed spectrum for 5G: from theory to application. IEEE J. Sel. Areas Commun. **35**(8), 1754–1767 (2017)
29. Xiao, C., Zeng, J., Ni, W., Liu, R.P., Su, X., Wang, J.: Delay guarantee and effective capacity of downlink NOMA fading channels. IEEE J. Sel. Top. Signal Process. **13**(3), 508–523 (2019)
30. Björnson, E., Larsson, E.G., Debbah, M.: Massive MIMO for maximal spectral efficiency: how many users and pilots should be allocated? IEEE Trans. Wirel. Commun. **15**(2), 1293–1308 (2016)
31. Qi, S., Jiang, D., Huo, L.: A prediction approach to end-to-end traffic in space information networks. Mob. Netw. Appl. (2019). https://doi.org/10.1007/s11036-019-01424-2
32. Wang, Y., Jiang, D., Huo, L., et al.: A new traffic prediction algorithm to software defined networking. Mob. Netw. Appl. (2019). https://doi.org/10.1007/s11036-019-01423-3
33. Lewis, D.D., Gale, W.A.: A sequential algorithm for training text classifiers. In: Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, vol. 29, no. 2, pp. 13–19. ACM (1994)
34. Du, B., Wang, Z., Zhang, L., Zhang, L., Liu, W., Shen, J., Tao, D.: Exploring representativeness and informativeness for active learning. IEEE Trans. Cybern. **47**(1), 14–26 (2017)
35. Tuia, D., Ratle, F., Pacifici, F., Kanevski, M.F., Emery, W.J.: Active learning methods for remote sensing image classification. IEEE Trans. Geosci. Remote Sens. **47**(7), 2218–2232 (2009)
36. Liu, K., Qiang, X., Wang, Z.: Survey on active learning algorithm. Comput. Eng. Appl. **48**(34), 1–4 (2012)
37. Tong, S., Koller, D.: Support vector machine active learning with applications to text classification. J. Mach. Learn. Res. **2**(1), 45–66 (2002)
38. Kremer, J., Pedersen, K.S., Igel, C.: Active learning with support vector machines. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. **4**(4), 313–326 (2014)
39. Hu, R., Mac, N.B., Delany, S.J.: Active learning for text classification with reusability. Expert Syst. Appl. **45**(C), 438–449 (2015)
40. Freund, Y., Schapire, Robert E.: A decision-theoretic generalization of on-line learning and an application to boosting. In: Vitányi, P. (ed.) EuroCOLT 1995. LNCS, vol. 904, pp. 23–37. Springer, Heidelberg (1995). https://doi.org/10.1007/3-540-59119-2_166
41. Jiang, Y., Zhou, Z.H.: A text classification method based on term frequency classifier ensemble. J. Comput. Res. Dev. **43**(10), 1681–1687 (2006)
42. Al-Salemi, B., Ab-Aziz, M.J., Noah, S.A.: LDA-AdaBoost.MH: accelerated AdaBoost.MH based on latent Dirichlet allocation for text categorization. J. Inf. Sci. **41**(1), 27–40 (2015)
43. Omar, M., On, B.W., Lee, I., Choi, G.S.: LDA topics: representation and evaluation. J. Inf. Sci. **41**(5), 1–4 (2015)
44. Forestier, G., Wemmert, C.: Semi-supervised learning using multiple clusterings with limited labeled data. Inf. Sci. **361**(C), 48–65 (2016)
45. Zhu, J., Wang, H., Tsou, B.K., Ma, M.: Active learning with sampling by uncertainty and density for data annotations. IEEE Trans. Audio Speech Lang. Process. **18**(6), 1323–1331 (2010)

46. Zhu, J., Wang, H., Yao, T., Tsou, B.K.: Active learning with sampling by uncertainty and density for word sense disambiguation and text classification. In: Proceedings of the 22nd International Conference on Computational Linguistics, Manchester, UK, pp. 1137–1144. ACL (2008)
47. Chen, L., Bao, R., Li, Y., Zhang, K., An, Y., Van, N.N.: An interactive information-retrieval method based on active learning. J. Eng. Sci. Technol. Rev. **10**(3), 1–6 (2017)
48. Tiwari, P., Melucci, M.: Towards a quantum-inspired binary classifier. IEEE Access **7**, 42354–42372 (2019). https://doi.org/10.1109/ACCESS.2019.2904624