Active Learning in New Generation Optical Networks

Stanisław Kozdrowski[1,3][0000−0001−6647−5189],
Sławomir Sujecki[2][0000−0003−4588−6741], and
Paweł Cichosz[1][0000−0002−8049−7410]

[1] Department of Computer Science,
Warsaw University of Technology,
Nowowiejska 15/19, 00-665 Warsaw, Poland.
[2] Faculty of Electronics, Military University of Technology,
S. Kaliskiego 2, 00-908 Warsaw, Poland.
[3] stanislaw.kozdrowski@pw.edu.pl

Abstract. This article presents a study of the application of active learning for the classification of optical paths with respect to transmission quality in real-world optical networks. This approach iteratively selects the most informative unlabeled paths from a large pool, thereby minimizing the labeling effort while maximizing classification quality. Both a realistic simulation study on a real dataset is performed, with class labels initially hidden but then selectively revealed in reply to labeling queries, and a real-world application study, with queries submitted to the operator to verify path feasibility in the actual real network. The results show substantial predictive performance improvement occurring within a small number of active learning iterations, with a limited number of labeled paths used for model creation. This provides strong evidence of the benefits of active learning for cost-effective optical path classification.

Keywords: machine learning · active learning · optical network · real data · quality of transmission

1 Introduction

Quality of transmission (QoT) estimation in optical networks can be formulated as a binary classification task when predicting for a given path connecting two network nodes whether a QoT measure exceeds or lies below an acceptability threshold. Paths for which the selected QoT measure threshold is met are classed as feasible, while those for which this is not the case are considered infeasible. Therefore, given a labeled training dataset, a classification model can be created to provide such class predictions. Consequently, the application of machine learning methods for QoT estimation is an obvious and natural methodology. Hence, in this contribution we use machine learning methods for QoT estimation in optical networks with a particular focus on the application of active learning.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

1.1 Motivation

The availability of labeled training data is a severe bottleneck of using machine learning for QoT estimation. Real datasets gathered network operators tend to be small and heavily imbalanced, with very few or no infeasible path examples. This clearly limits the level of predictive quality possible to obtain [10]. While collecting bigger datasets with more infeasible paths would be possible in the long term, this would be a costly and time consuming process. Network simulators can provide synthetic data that are big enough and balanced but the simulation may not be sufficiently realistic to permit high quality predictions when applied to optical channels from a real network.

Active learning [11] offers an effective strategy to significantly reduce the expense and labor involved in assigning class labels to training data, while maintaining high predictive accuracy. The learning process can be initiated with only a handful of labeled instances used to create an imperfect initial model, provided that there is access to a pool of unlabeled data from which further instances can be drawn and labeled as needed. Each time a labeling query is submitted, it augments the training set with new data instances, and with this augmented training set, a new model is created. The choice of which instances from the pool to label next is guided by query selection strategies [14, 20].

Through the repeated selection of a few appropriately chosen instances for labeling, active learning can produce a high-quality model using far fewer labeled instances than would be required by conventional passive learning, thereby optimizing the use of scarce data annotation resources. This has been confirmed in tasks such as text classification [27] and image classification [5]. Under class imbalance, active learning may not only lessen the burden of labeling but also enhance the model's predictive accuracy [13].

1.2 Related Work

Most previous studies on machine learning applied to optical path classification have adopted a conventional passive learning methodology, presuming that labeled training instances are readily accessible, e.g., [4, 18, 24, 21, 16, 10, 12, 19, 25]. Many of them rely on synthetic data generated through simulation models rather than real-world measurements.

Only limited exploration of active learning QoT estimation in optical networks has been performed [2, 3, 17, 9]. These existing studies demonstrate that it can substantially reduce the amount of labeled data required to achieve a given prediction performance level. Of those, only [3] and [9] compare multiple active learning setups and only [9] uses real network data to simulate the active learning process. This work extends the latter by not only using real data but actually applying active learning in a real telecom network. To our best knowledge this is the first attempt reported in the literature in which iterative query selection, labeling, and model refinement is performed in an actually operating optical network, in cooperation with the operator.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

1.3 Contributions

More specifically, this paper presents the results of the following experimental studies:

- a realistic simulation of active learning for optical channel classification using a real dataset, supposed to verify the utility of different classification algorithms and query selection strategies,
- a fully real-world application of active learning optical channel classification, with labeling queries submitted to the network operator and labels determined by the network operator's optical communication engineering team.

They confirm the utility of active learning for QoT estimation, making it possibly to obtain high-quality models with reduced labeled data requirements and workload.

2 Methods

This section describes data representation, classification algorithms, and the active learning model creation and evaluation process.

2.1 Data representation

A real optical network with 85 nodes and 75 edges is considered, maintained by one of major network operators in Poland. An optical network route (optical channel) consists of several edges, known as *hops*, where each hop represents the transfer of a signal between two nodes. The quality of transmission is affected by the distribution of these jumps and their specific characteristics, such as the distance the signal travels in each hop and the degree of signal attenuation. Optical channels connecting different nodes may overlap in some sections, leading to common hops and affecting the overall transmission quality. The set of available attributes describing optical channels includes both properties common to the entire route (number of hops, transponder modulation, and transponder transmission rate) and properties of individual hops (hop length, hop loss, and number of optical channels sharing the hop).

Following [16], network paths are represented as attribute value vectors, using an aggregation approach to vectorize the three hop properties:

1. hop length (distance),
2. hop loss,
3. number of paths sharing the hop.

They are aggregated across all hops in a path using the following set of aggregation functions:

1. mean,
2. standard deviation,

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

3. minimum,
4. maximum,
5. median,
6. first quartile,
7. third quartile,
8. linear correlation coefficient with the ordinal number of the hop in the path.

This results in 8 attributes for each of the 3 hop properties, totaling 24 attributes derived from hop properties. With the additional 3 path-level attributes not tied to individual hops:

1. number of hops,
2. transponder modulation,
3. transponder bitrate,

this yields 27 attributes for the vector path representation.

2.2 Classification Algorithms

Conventional classification algorithms for tabular data, which have demonstrated effectiveness in optical path classification in previous studies [16], are the most promising candidates for using within the active learning framework. Of those, the random forest (RF) [6] and extreme gradient boosting (xgboost) [7] algorithms have been adopted for this work as potentially the most useful. They are relatively resistant to overfitting and can generalize effectively even with small training sets. Additionally, they are not overly sensitive to the choice of hyperparameters within a reasonable range, which is convenient when working with small datasets for which hyperparameter tuning is problematic. Furthermore, during prediction they can output class probability estimates, which is desirable for active learning query selection.

2.3 Active Learning

Active learning is an iterative model creation process with training information provided by answering labeling queries [11]. The process creates the first model on the initial training set and then subsequent models on enlarged training sets obtained by adding query instances with their class labels. With an appropriate query selection method, model quality would improve in the course of learning and a satisfactory model may be reached after using a relatively small number of labeled instances.

This work adopts the following active learning scenario.

1. There is a pool of unlabeled instances, denoted as $\mathcal{P}$.
2. An initial training set $\mathcal{T}$ is selected from $\mathcal{P}$.
3. In each iteration:
   (a) the current labeled training set $\mathcal{T}$ is used to train a classification model h,

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

(b) based on the predictions of h, a subset of query instances is selected from the pool $\mathcal{P}$,
(c) the true class labels for these selected instances are requested and received,
(d) the newly labeled instances are added to $\mathcal{T}$ and removed from $\mathcal{P}$,
(e) this process repeats for a specified maximum number of iterations, until the pool $\mathcal{P}$ is exhausted, or until another stopping criterion is met.

Initial Training Set Selection A simple random initial training set selection method is used for this work. If after obtaining class labels for the initially selected instances there are no representatives of the minority class, additional instances are drawn at random until both the classes are represented. While more refined selection methods are possible [28], random selection tends to handle extreme class imbalance better.

Query Selection The most useful instances for a class label query are those that would make the next model the most improved over the current model. Several selection strategies have been proposed in the literature, both general-purpose and specific to some model representations and classification algorithms [14]. Two general-purpose strategies are considered for this work: uncertainty sampling and diversity sampling. Additionally a hybrid strategy combining these two is used as well as random sampling serving as a comparison baseline.

Uncertainty sampling operates on the premise that instances which the current model finds most challenging to classify are likely to be the most informative. These high-uncertainty instances are identified using the models's class probability predictions. For binary classification uncertainty sampling reduces to choosing those instances from the pool for which the predicted probability of the more likely class is the smallest (closest to 0.5), i.e.:

$$\underset{x \in \mathcal{P}}{\text{minimize}} \ \max_{c} p_c(x), \tag{1}$$

where $p_c(x)$ is the probability of class c for instance x.

The idea behind diversity sampling is that the most useful new training instances should be maximally different from those selected previously. Therefore instances from the pool that maximize minimum dissimilarity to instances in the training set are selected:

$$\text{maximize for } x \in P: \quad \min_{x' \in T} \delta(x, x'), \tag{2}$$

where $\delta(x, x')$ denotes the L2 dissimilarity between instances x and x'.

The hybrid strategy randomly mixes uncertainty and diversity sampling in a fixed proportion, i.e., each query instance is obtained using either one or the other strategy with a specified probability.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

2.4 Model Evaluation

On each active learning iteration the model is evaluated on a test set using the precision-recall analysis. It provides a reliable and useful assessment of the model's predictive power under class imbalance and evaluates predicted class probabilities taking into account the whole range of thresholds for actual feasible or infeasible class predictions.

A precision-recall curve represents possible tradeoff points between precision – the proportion of positive class predictions that are correct – and recall – the proportion of actual positive class instances correctly predicted as positive. The infeasible class is considered positive, since the primary purpose of the model is to predict which paths cannot achieve the required transmission quality level. Thus, precision is the proportion of paths predicted to be bad by the model that are actually bad, and recall is the proportion of actually bad paths that are correctly predicted to be bad by the model.

Precision-recall tradeoff points are associated with varying settings of the decision threshold, which is used to compare predicted class probabilities to assign class labels. Lowering the threshold boosts recall, though this often comes at the expense of precision; conversely, raising the threshold typically enhances precision but may lower recall.

To capture model performance across the full spectrum of these tradeoffs, the area under the precision-recall curve (PR AUC) is employed. This metric reflects the average precision achieved over all recall levels. In contrast to single-threshold measures like accuracy, precision, recall, or F1 score, PR AUC does not evaluate performance at just one threshold, making it a more robust and insightful indicator. Notably, PR AUC is far more responsive to false positives than the commonly used ROC AUC, which can give an overly optimistic impression when classes are highly imbalanced. This is because, in datasets with an abundance of true negatives, a considerable number of false positives has minimal impact on the false positive rate but significantly lowers precision.

3 Results

The utility of active learning for optical path classification has been verified experimentally using real network data. This section describes the experimental procedure and presents the results.

3.1 Implementation and Configuration

The implementation of the random forest algorithm provided by the scikit-learn library [22] and the implementation of the extreme gradient boosting algorithm provided by the `xgboost` library [8] were used. No hyperparameter tuning was performed to avoid an optimistic bias resulting from reusing the tuning data for other purposes, which would be unavoidable due to the small number of labeled paths. The random forest is used with default settings except for the number of

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

trees which is set to 500. The xgboost algorithm is used with 50 trees, minority class weight set to the square root of the imbalance ratio, and default values of other hyperparameters.

3.2 Experimental Studies

Two experimental studies have been performed:

realistic simulation: using data from the real network to simulate the active learning process,

real-world application: actually performing the active learning process in the real network.

The former is supposed to evaluate the utility of different classification algorithms and active learning query selection strategies. The latter applies the identified best setup to perform several real active learning iterations.

3.3 Study 1: Realistic Simulation

In this study, the real dataset provided by the network operator is used to simulate the active learning process by treating it as unlabeled pool, i.e., initially hiding all class labels except from a randomly selected small initial training set and then unhiding them for paths selected as query instances during active learning iterations. Therefore, the simulation aspect does not consist in using synthetic data from network simulation, and it just concerns labeling – instead of testing query paths in the field by issuing work orders, their class labels are being revealed upon request. These are the same assumptions as those adopted in the experiments reported by [9], but the experiment is performed on a more up-to-date, considerably larger dataset.

For this study, the initial training set size is set to 10 paths. If the initially selected set does not contain examples of infeasible paths, then additional instances are drawn at random in increments of 2 until both classes are represented. At each active learning iteration 5 query paths are selected. The random forest and xgboost classification algorithms are used with each of the four query selection strategies described in Section 2.3.

To evaluate model quality, the stratified 10×5-fold cross validation procedure is used, applied to the whole labeled set of 229 paths. For each fold, the current cross-validation training subset serves as a pool of unlabeled instances for active learning, from which the initial training set and instances for class label queries are selected. At each active learning iteration, the predictive performance of the current model is evaluated on the cross-validation test set.

Figure 1 presents the obtained learning curves, obtained by plotting the PR AUC values versus the percentage of the pool labeled so far. Horizontal dashed lines present the baseline passive learning performance level, corresponding to the situation when all class labels are available.

It can be seen that the xgboost algorithm outperforms random forest both in passive and in active learning mode, reaching the PR AUC level of more than

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

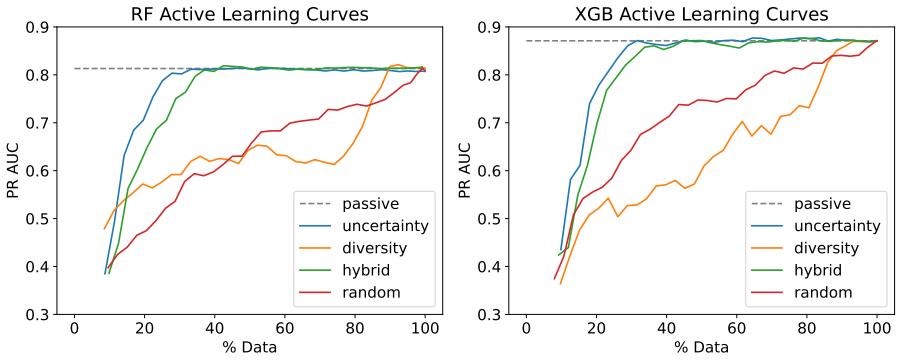

Fig. 1. Learning curves for the realistic simulation study.

0.85 whereas the latter just exceeds 0.8. Both the algorithms exhibit quite rapid performance improvement during active learning, reaching the passive performance level after labeling about 30% of the pool. However, the query selection strategy is crucial for this learning speed: random sampling permits only a slow PR AUC increase, proportional to the labeled data percentage, and diversity sampling, while slightly better at the start, is not better in the longer run. Uncertainty sampling clearly performs the best, followed by the hybrid strategy. Whereas these findings generally confirm the results from [9], the actual performance levels observed here are considerably higher. This is due to an extended dataset available for this experiment, containing about twice as many paths, including 5 times more infeasible ones.

3.4 Study 2: Real-World Application

This study investigates a fully real-world active learning process. The available labeled dataset on which the previous simulation experiment was performed here is only used to randomly select an initial training set of 20 paths, including 5 infeasible ones.

An unlabeled active learning pool contains 500 paths generated based on the network topology. Class labels for these paths are initially truly unknown and can only be determined by the network operator's engineering team. With this caveat, the active learning process works in the usual way, with a classification model created at each iteration using the current training set and evaluated on the test set. The latter is composed of 150 paths generated in the same way as the pool and labeled by the operator, one-third of which are infeasible ones.

The best active learning setup identified in the previous study is used, i.e., the extreme gradient boosting algorithm combined with uncertainty sampling for query selection. At each active learning iteration 5 paths from the pool are selected and the corresponding work orders are submitted to the network operator, to determine whether they do or do not provide an acceptable level of

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

transmission quality. These paths with the obtained class labels are added to the training set, which completes the current active learning iteration.

During the timeframe available for this study 30 active learning iterations were performed. The test set performance of the models from each of these iterations is presented in Figure 2 by the PR AUC learning curve. Iteration 0 corresponds to the first model created based on the initial training set and then for $i = 1, 2, \ldots, 30$ iteration i corresponds to the model obtained after the i-th labeling query. The plot also presents the minimum confidence in the unlabeled pool at each iteration, measured as the difference between the probability of the more likely class and 0.5, scaled to the $[0, 1]$ range.

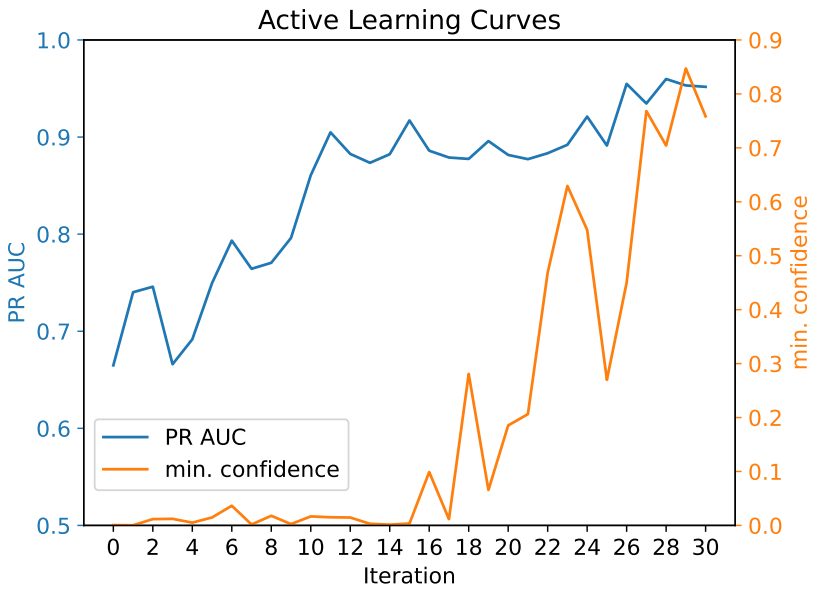

Fig. 2. Learning curves for the real-world application study.

One can observe that PR AUC, starting at slightly above 0.65, reaches 0.9 within 12 iterations and, after a period of stabilization, jumps to 0.95 in the last three iterations. This exceeds the values from the previous study because classes obtained for pool paths are less imbalanced. The minimum confidence over the unlabeled pool is a useful indicator of the increasing certainty of the model's prediction. It starts to visibly increase to about 0.1 at iteration 16, then goes through two peaks at iterations 18 and 24, and reaches a maximum of nearly 0.85 at iteration 29. This is a strong stopping signal for the active learning process, suggesting that little or no further improvement is likely.

Table 1 presents the row-normalized confusion matrices obtained at iterations 12 and 30 of the active learning process. The entries of these matrices

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

represent the percentage of the actual class with given predicted classes, obtained at the default probability threshold of 0.5. While both these matrices demonstrate useful class discrimination power, the model from iteration 30 is much better at detecting the positive class, corresponding to infeasible paths. Adjusting the decision threshold might identify more preferred model operating points.

Table 1. Confusion matrices at different training iterations.

Iteration 12			**Iteration 30**		
	Predicted			Predicted	
Actual	0	1	Actual	0	1
0	100.0%	0.0%	0	85.0%	15.0%
1	43.8%	56.2%	1	4.2%	95.8%

4 Conclusions

In this paper, we have investigated the utility of active learning for optical channel classification by performing both a simulation study using real data and a fully real-world application study. The former was used to select the classification algorithm and the query selection strategy. The latter applied the identified setup to an actually working network, with labeling queries handled by the operator's engineering team. To the best of our knowledge, this is the first experiment of this kind described in the literature.

The results confirm the utility of active learning for QoT estimation, making it possible to obtain high-quality models with reduced labeled data requirements. The simulation study identified the extreme gradient boosting classification algorithm with the uncertainty sampling query selection strategy as the best active learning setup. In the real-world study the predictive performance of the classification model reached a satisfactory level after processing the first 50–60 query paths, with an additional improvement observed at a later stage. Additionally, the minimum prediction confidence over the unlabeled pool, rapidly raising at final iterations, was found to provide a useful stopping signal for active learning.

While strongly encouraging, the reported study leaves several issues to be addressed by future work. It would be worthwhile to extend the scope of active learning configurations considered, including more classification algorithms and data representation variations as well query selection strategies, possibly incorporating class imbalance compensation techniques [13, 30, 1]. Exploring non-random initial training set selection methods would be also interesting [28, 15]. Another, possibly more important direction, is related to detecting and compensating a possible distribution shift between the source domain of training instances, selected from the pool, and the target data of real network paths to which the final model would be eventually applied. While in this study the pool and the test set were generated from the same distribution, the distribution of

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

paths sent to an actually deployed model for prediction is likely to be different. Incorporating domain adaptation techniques to the active learning process would be necessary to make it robust with respect to such a distribution shift [26, 23]. Finally, it remains to be verified whether semi-supervised learning [29], which takes an even more radical approach to reducing the need for labeled training data, may be useful for optical channel classification. This could lead to a hybrid method combining active learning with such semi-supervised techniques as self-training or label propagation.

References

1. Aggarwal, U., Popescu, A., Hudelot, C.: Minority class oriented active learning for imbalanced datasets. arXiv preprint arXiv:2202.00390 (2022)
2. Azzimonti, D., Rbottondi, C., Tornatore, M.: Using active learning to decrease probes for QoT estimation in optical networks. In: Proceedings of the 2019 Optical Fiber Communications Conference and Exhibition. OFC-2019, IEEE (2019)
3. Azzimonti, D., Rottondi, C., Tornatore, M.: Reducing probes for qot estimation in optical networks with active learning. IEEE Transactions on Network and Service Management **17**(2), 1022–1035 (2020). https://doi.org/10.1109/TNSM.2020.2972058
4. Barletta, L., Giusti, A., Rottondi, C., Tornatore, M.: Qot estimation for unestablished lighpaths using machine learning. In: 2017 Optical Fiber Communications Conference and Exhibition (OFC). pp. 1–3 (2017)
5. Beluch, W.H., Genewein, T., Nurnberger, A., Kohler, J.M.: The power of ensembles for active learning in image classification. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR-2018, IEEE Press (2018)
6. Breiman, L.: Random forests. Machine Learning **45**, 5–32 (2001)
7. Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: Proceedings of the Twenty-Second ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM Press (2016)
8. Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y.: XGBoost Python Package (2021), https://github.com/dmlc/xgboost, library version 1.5.1
9. Cichosz, P.: Active learning of optical path classification. Engineering Applications of Artificial Intelligence **151**, 110582 (2025)
10. Cichosz, P., Kozdrowski, S., Sujecki, S.: Learning to classify dwdm optical channels from tiny and imbalanced data. Entropy **23**(11) (2021). https://doi.org/10.3390/e23111504, https://www.mdpi.com/1099-4300/23/11/1504
11. Cohn, D., Atlas, L., Ladner, R.: Improving generalization with active learning. Machine Learning **15**, 201–221 (1994)
12. Côté, O., Pesic, M., et al.: Quality of transmission estimation using random forest classification for optical networks. Journal of Optical Communications and Networking **15**(3), B63–B73 (2023). https://doi.org/10.1364/JOCN.471165
13. Ertekin, S., Huang, J., Bottou, L., Giles, L.: Learning on the border: Active learning in imbalanced data classification. In: Proceedings of the Sixteenth ACM on Information and Knowledge Management. CIKM-2007, ACM Press (2007)

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16

14. Fu, Y., Zhu, X., Li, B.: A survey on instance selection for active learning. Knowledge and Information Systems **35**, 249–283 (2013)
15. Kottke, D., König, G., Stein, B., Sick, B.: To actively initialize active learning. Pattern Recognition **128**, 108662 (2022). https://doi.org/10.1016/j.patcog.2022.108662
16. Kozdrowski, S., Cichosz, P., Paziewski, P., Sujecki, S.: Machine learning algorithms for prediction of the quality of transmission in optical networks. Entropy (Basel, Switzerland) **23**(1) (January 2021). https://doi.org/10.3390/e23010007
17. Li, Z., Gu, Z., Zhang, J., Zhou, Y., Ji, Y.: Predictive uncertainty aware active learning for regression-based QoT estimation in optical networks. In: Proceedings of the 2021 Asia Communications and Photonics Conference. Optica Publishing Group (2021)
18. Morais, R.M., Pedro, J.: Machine learning models for estimating quality of transmission in dwdm networks. IEEE/OSA Journal of Optical Communications and Networking **10**(10), D84–D99 (2018)
19. Morais, R.M., da Silva, E.P.: Machine learning-based qot estimation using support vector machines in optical transport networks. IEEE Access **11**, 62901–62912 (2023). https://doi.org/10.1109/ACCESS.2023.3287421
20. Nguyen, V.L., Shaker, M.H., Hüllermeier, E.: How to measure uncertainty in uncertainty sampling for active learning. Machine Learning **111**, 89–122 (2022)
21. Panayiotou, T., Manousakis, K., Chatzis, S.P., Ellinas, G.: A data-driven bandwidth allocation framework with qos considerations for eons. Journal of Lightwave Technology **37**(9), 1853–1864 (2019)
22. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011), library version 1.1.3
23. Rai, P., Saha, A., Daumé III, H., Venkatasubramanian, S.: Domain-adaptive active learning. In: Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI). pp. 1086–1092. AAAI Press, Atlanta, GA, USA (2010)
24. Rottondi, C., Barletta, L., Giusti, A., Tornatore, M.: Machine-learning method for quality of transmission prediction of unestablished lightpaths. IEEE/OSA Journal of Optical Communications and Networking **10**(2), A286–A297 (2018)
25. Sahu, R., Clement, Y.: Deep learning techniques for qot estimation in optical networks. Optics Communications **560**, 129992 (2025). https://doi.org/10.1016/j.optcom.2024.129992
26. Sugiyama, M., Nakajima, S., Kashima, H., von Bünau, P., Kawanabe, M.: Active learning for covariate shift. In: Proceedings of the 2008 Advances in Neural Information Processing Systems (NeurIPS). pp. 1089–1096. Curran Associates, Inc. (2008)
27. Tong, S., Koller, D.: Support vector machine active learning with applications to text classification. Journal of Machine Learning Research **2**, 45–66 (2001)
28. Yuan, W., Han, Y., Guan, D., Lee, S., Lee, Y.K.: Initial training data selection for active learning. In: Proceedings of the Fifth International Conference on Ubiquitous Information Management and Communication. ICUIMC-2011, ACM Press (2011)
29. Zhu, X., Goldberg, A.: Introduction to Semi-Supervised Learning. Morgan & Claypool (2009)
30. Zhu, X., Hovy, E.: Active learning with imbalanced data. Proceedings of the 2007 IEEE International Conference on Data Mining Workshops (ICDMW) pp. 192–201 (2007). https://doi.org/10.1109/ICDMW.2007.83

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_16