

Team QTB on Feature Selection Via Quantum Annealing and Hybrid Models

Notebook for the QuantumCLEF Lab at CLEF 2024

Esteban Payares¹, Edwin Puertas¹ and Juan Carlos Martínez-Santos¹

¹Universidad Tecnológica de Bolívar, School of Engineering

Abstract

Quantum technologies are a reality today, and the future is bright regarding its capabilities for real-world applications. Feature selection is a crucial preprocessing step in Information Retrieval. By identifying the most informative subset of features, feature selection can improve the efficiency of learning to rank models. In this paper, we propose a novel approach to feature selection for Information Retrieval using quantum annealing. This promising optimization technique leverages the principles of quantum mechanics. We focus on the MQ2007 dataset, a widely used benchmark for learning to rank tasks. We also explore different formulations of the feature selection problem as quadratic unconstrained binary optimization (QUBO) problems, including mutual information, conditional mutual information, and correlation coefficients. Our quantum annealing-based approaches demonstrate their effectiveness in selecting informative features, outperforming simulated annealing, which achieves an nDCG@10 score of 0.4024. The best quantum annealing-based approach achieves a score of 0.443 using a hybrid solver with only ten features. We discuss the importance of the number of selected features in the performance of learning to rank models and the role of hybrid quantum-classical solvers in incorporating additional constraints and preferences into the feature selection process. Our work demonstrates the potential of using quantum annealing to tackle complex optimization problems. It paves the way for further exploration in this domain.

Keywords

Feature Selection, Information Retrieval, Learning to Rank, Quantum Annealing, MQ2007

1. Introduction

Information Retrieval (IR) systems are crucial in efficiently and effectively retrieving relevant information from large-scale data. Among the various components of IR systems, learning-to-rank (LTR) models, such as LambdaMART[1], have shown remarkable performance in improving the quality of search results [2]. However, the performance of these models heavily relies on selecting informative features from the high-dimensional feature space [3].

Feature selection is an essential preprocessing step in IR to identify the most relevant subset of features for training LTR models [4]. By reducing the dimensionality of the input data, feature selection not only improves the efficiency of the training process but also enhances the generalization ability of the models [5]. Traditional feature selection methods, such as recursive feature elimination (RFE) with logistic regression [6], have been widely used in IR. However, these methods often need help with the complex and non-linear relationships among features and the target variable [7].

Recent advancements in quantum computing have opened up new possibilities for solving optimization problems, including feature selection. Quantum annealing, in particular, has shown promise in efficiently exploring vast search spaces and finding optimal solutions [8]. By leveraging the principles of quantum mechanics, such as superposition and entanglement, quantum annealing can outperform classical optimization algorithms [9]. Wang et al. [10] have demonstrated the effectiveness of integrating machine learning algorithms with quantum annealing solvers for online fraud detection, showcasing the potential of quantum annealing in real-world applications.

In this work, we propose a novel approach to feature selection for IR using quantum annealing. Specifically, we focus on the MQ2007 dataset [11], a widely used benchmark for LTR tasks, and aim to

CLEF 2024: Conference and Labs of the Evaluation Forum, September 9–12, 2024, Grenoble, France

✉ epayares@utb.edu.co (E. Payares); epuertas@utb.edu.co (E. Puertas); jcmartinezs@utb.edu.co (J. C. Martínez-Santos)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

select the most informative subset of features to improve the performance of a LambdaMART model. We explore different formulations of the feature selection problem as quantum annealing problems, including mutual information (MI), conditional mutual information (CMI), and correlation coefficients.

The main contributions of our work are as follows: We present a quantum annealing-based approach to feature selection for IR, demonstrating its effectiveness on the MQ2007 dataset. We compare different problem formulations (MI, CMI, and correlation coefficients) and analyze their impact on the feature selection process. We evaluate the performance of the LambdaMART model trained with the selected features using the normalized discounted cumulative gain (nDCG) metric [12]. We provide insights into the quantum annealing process through energy histograms and discuss the feasibility of the solutions found by the hybrid solver.

2. Related Work

Feature selection is a fundamental process in machine learning that involves the identification of a subset of pertinent features from a larger set to improve model performance. Several authors have developed various approaches to address this problem, each with unique contributions and benefits. Evolutionary Quantum Feature Selection (EQFS) employs principles of quantum computing to enhance the efficiency of feature selection. EQFS employs the Quantum Circuit Evolution (QCE) algorithm, which uses shallow-depth circuits to generate sparse probability distributions. It enables the identification of optimal feature combinations with quadratic scaling in the number of features, thus overcoming the curse of dimensionality. This technique is particularly effective for high-dimensional datasets [13, 14].

We can classify the various traditional feature selection methods into three main categories: filter, wrapper, and embedded methods. Filter methods, such as correlation and mutual information, use general dependency measures between features and the output. Wrapper methods evaluate subsets of variables based on the quality of the final classification, offering superior performance but at a higher computational cost. Embedded methods integrate feature selection into the training process of the classifier, with examples including decision trees and methods with L1 regularization. These traditional approaches are fundamental in various Machine Learning applications, balancing computational efficiency and accuracy [13]. Moreover, methods based on mutual information have been highly influential in multiple applications [15].

Another approach to quantum computing involves integrating machine learning algorithms with quantum annealing solvers. This approach is particularly productive in online fraud detection, where data processing in real-time or near-real-time is of paramount importance. The quantum-enhanced SVM has demonstrated notable enhancements in speed and accuracy for time-series datasets with highly imbalanced data. However, the accuracy improvements for non-time series data were relatively marginal. This approach demonstrates the potential of quantum machine learning (QML) applications in time-series data. It illustrates that the optimal strategy depends on the dataset type and the balance between speed, accuracy, and cost [16]. Moreover, authors have demonstrated that conditional mutual information is a practical approach to feature selection in incomplete data [17].

Genetic algorithms represent a type of randomized feature selection method that attempts to emulate the processes of natural evolution. These algorithms generate feature subsets randomly and iteratively improve them through operations such as crossover and mutation. Genetic algorithms can effectively locate optimal feature subsets, thereby avoiding the pitfall of local optima. This robustness renders them well-suited for complex feature selection tasks. The capacity to evade local optima confers a substantial advantage over other sequential search methodologies. Furthermore, incorporating randomness into these processes can enhance the efficiency of identifying optimal feature sets [18]. Moreover, integrating mutual information into genetic algorithms has been employed to improve feature selection [19].

The integration of classical machine learning techniques with quantum computing offers the potential for significant advantages. For example, a hybrid approach that employs quantum evolutionary algorithms to optimize feature combinations. In contrast, classical algorithms that evaluate these combinations can effectively leverage the strengths of both paradigms. This approach aims to enhance

the precision and efficacy of feature selection in a range of machine-learning applications. The capacity of quantum algorithms to address the computational complexity of large datasets is a valuable complement to the resilience of classical methods in model evaluation. This type of integration promises significant advancements in the practical application of quantum feature selection [13]. Combining mutual information with other criteria has been demonstrated to enhance feature selection in various domains [20].

In supervised feature selection, a recent approach combines sparse representation and mutual information to enhance the feature selection process. The method begins with sparse representation, which evaluates the importance of each feature in the overall data structure. Subsequently, mutual information is employed to reduce redundancy among the selected features. The experimental results demonstrate that this approach significantly enhances classification outcomes and can achieve superior feature selection results at lower dimensions. This method illustrates the significance of considering the global importance of features and redundancy reduction to achieve optimal results [21]. Moreover, using conditional mutual information and other innovative methodologies has yielded encouraging outcomes in feature selection [22].

3. Methodology

In this section, we describe our methodology for feature selection using quantum annealing. We first formulate the feature selection problem as a quantum annealing problem and then present the approaches we explored, including mutual information, conditional mutual information, and correlation coefficients. We also discuss the use of hybrid quantum-classical solvers and simulated annealing for comparison.

3.1. Dataset

Our experiments use the MQ2007 dataset, a widely used benchmark dataset for learning to rank tasks in Information Retrieval. The MQ2007 dataset consists of 1,692 queries and 69,623 documents, with each query-document pair represented by a 46-dimensional feature vector. The QuantumCLEF lab organizers provide a training set, which we use for feature selection. The organizers are the ones responsible for training and evaluating the learning to rank model using the selected features [23, 24].

3.2. Problem Formulation

Let $\mathcal{D} = (\mathbf{x}_i, y_i)_{i=1}^N$ be a dataset consisting of N instances, where $\mathbf{x}_i \in \mathbb{R}^d$ is a d -dimensional feature vector and $y_i \in \mathbb{R}$ is the corresponding target value. The goal of feature selection is to identify a subset of features $\mathcal{S} \subseteq \{1, 2, \dots, d\}$ that maximizes the relevance to the target variable while minimizing the redundancy among the selected features.

We formulate the feature selection problem as a quadratic unconstrained binary optimization (QUBO) problem, which can be solved using quantum annealing. The QUBO formulation is given by:

$$\min_{\mathbf{q} \in \{0,1\}^d} \mathbf{q}^\top \mathbf{Q} \mathbf{q}, \quad (1)$$

where \mathbf{q} is a binary vector representing the selection of features, and \mathbf{Q} is a $d \times d$ matrix capturing the interactions between features. The diagonal elements of \mathbf{Q} represent the relevance of each feature to the target variable. In contrast, the off-diagonal elements represent the redundancy between pairs of features.

3.3. Mutual Information

Mutual information (MI) measures the dependence between two random variables. In the context of feature selection, we use MI to quantify the relevance of each feature to the target variable. We define the MI between a feature X_i and the target variable Y as:

$$\text{MI}(X_i; Y) = \sum_{x \in X_i} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}, \quad (2)$$

where $p(x, y)$ is the joint probability distribution of X_i and Y , and $p(x)$ and $p(y)$ are the marginal probability distributions of X_i and Y , respectively.

To formulate the feature selection problem using MI, we construct the \mathbf{Q} matrix as follows:

$$\mathbf{Q}_{ij} = \begin{cases} -\text{MI}(X_i; Y) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

The negative sign in the diagonal elements ensures that the optimization problem seeks to maximize the MI between the selected features and the target variable.

3.4. Conditional Mutual Information

Conditional mutual information (CMI) extends the concept of MI by considering the redundancy between pairs of features given the target variable. The CMI between two features X_i and X_j given the target variable Y is defined as:

$$\text{CMI}(X_i; X_j | Y) = \sum_{x_i \in X_i} \sum_{x_j \in X_j} \sum_{y \in Y} p(x_i, x_j, y) \log \frac{p(x_i, x_j | y)}{p(x_i | y)p(x_j | y)}, \quad (4)$$

where $p(x_i, x_j, y)$ is the joint probability distribution of X_i , X_j , and Y , and $p(x_i | y)$ and $p(x_j | y)$ are the conditional probability distributions of X_i and X_j given Y , respectively.

To incorporate CMI into the feature selection problem, we update the \mathbf{Q} matrix as follows:

$$\mathbf{Q}_{ij} = \begin{cases} -\text{MI}(X_i; Y) & \text{if } i = j, \\ -\text{CMI}(X_i; X_j | Y) & \text{if } i \neq j. \end{cases} \quad (5)$$

The negative sign in the off-diagonal elements ensures that the optimization problem seeks to minimize the redundancy between the selected features given the target variable.

3.5. Correlation Coefficients

Correlation coefficients measure the linear relationship between two variables. In the context of feature selection, we use the Pearson correlation coefficient to quantify the relevance of each feature to the target variable and the redundancy between pairs of features. The Pearson correlation coefficient between two variables X and Y is defined as:

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}, \quad (6)$$

where $\text{cov}(X, Y)$ is the covariance between X and Y , and σ_X and σ_Y are the standard deviations of X and Y , respectively.

To formulate the feature selection problem using correlation coefficients, we construct the \mathbf{Q} matrix as follows:

$$\mathbf{Q}_{ij} = \begin{cases} -|\rho(X_i, Y)| & \text{if } i = j, \\ -|\rho(X_i, X_j)| & \text{if } i \neq j. \end{cases} \quad (7)$$

The absolute values of the correlation coefficients capture both positive and negative relationships. The negative sign ensures that the optimization problem maximizes relevance and reduces redundancy.

3.6. Quantum Annealing

Quantum annealing is an optimization technique that leverages the principles of quantum mechanics to solve complex optimization problems. It is particularly well-suited for solving QUBO problems, as it efficiently explores the vast solution space and finds the global minimum. To solve the feature selection problem using quantum annealing, we follow these steps:

Algorithm 1 Quantum Annealing for Feature Selection

- 1: Construct the \mathbf{Q} matrix based on the chosen approach (MI, CMI, or correlation coefficients).
 - 2: Convert the \mathbf{Q} matrix to a QUBO problem.
 - 3: Submit the QUBO problem to the quantum annealer.
 - 4: Retrieve the solution from the quantum annealer.
 - 5: Interpret the solution as the selected subset of features.
-

Once the quantum annealer has completed its task, it returns a binary vector \mathbf{q} . This vector is more than just a series of 1s, and 0s—it represents the selected subset of features. The indices of the non-zero elements in \mathbf{q} are the features that have been chosen, providing a clear output of the quantum annealing process.

3.7. Hybrid Quantum-Classical Solvers

In addition to using pure quantum annealing, we explore using hybrid quantum-classical solvers for feature selection. Hybrid solvers combine the strengths of quantum and classical computation by leveraging quantum annealing to explore the solution space and classical optimization techniques for local refinement.

We specifically employ the LeapHybridCQMSampler, a hybrid solver provided by D-Wave Systems. This solver combines quantum annealing with a classical-constrained quadratic model (CQM) solver. The hybrid solver allows us to incorporate additional constraints into the feature selection problem, such as the desired number of selected features.

Our approach involves formulating the feature selection problem as a CQM and then submitting it to the LeapHybridCQMSampler. This hybrid solver, upon receiving our problem, diligently explores the solution space and returns a set of feasible solutions. From this set, we select the best solution based on the objective function value, ensuring the highest quality of our results.

3.8. Simulated Annealing

For comparison purposes, we also implement feature selection using simulated annealing, a classical optimization technique inspired by the annealing process in metallurgy. Simulated annealing aims to find the global minimum of an objective function by gradually decreasing the system's temperature and allowing occasional uphill moves to escape local minima.

We use the SimulatedAnnealingSampler provided by the D-Wave Ocean SDK to solve the feature selection problem using simulated annealing. The simulated annealing algorithm follows a similar procedure to quantum annealing. Still, instead of leveraging quantum effects, it relies on a probabilistic acceptance criterion to explore the solution space.

By comparing the results obtained using quantum annealing, hybrid quantum-classical solvers, and simulated annealing, we can assess the relative effectiveness of these approaches for feature selection in Information Retrieval.

3.9. Evaluation

The organizers of the QuantumCLEF lab, in a collaborative effort, have evaluated the effectiveness of our quantum annealing-based feature selection approaches. We have submitted the selected features from the training set to the organizers, who have then used them to train a LambdaMART model on

their internal training set. LambdaMART, a state-of-the-art learning-to-rank algorithm, combines the MART (Multiple Additive Regression Trees) algorithm with a lambda gradient optimization framework.

The organizers evaluate the performance of the LambdaMART model using their internal validation and test sets, which are inaccessible to us. They measure the performance using the normalized discounted cumulative gain (nDCG) metric, commonly used in Information Retrieval, to evaluate the quality of ranked results. We defined the nDCG@k metric as:

$$\text{nDCG@k} = \frac{\text{DCG@k}}{\text{IDCG@k}}, \quad (8)$$

where DCG@k is the discounted cumulative gain at position k , and IDCG@k is the ideal discounted cumulative gain at position k , which represents the maximum possible DCG@k value.

The evaluation results provided by the organizers, as explored in section 4, will offer valuable insights into the effectiveness of our quantum annealing-based feature selection methods. These insights have the potential to significantly improve the performance of learning to rank models in Information Retrieval tasks, underscoring the relevance and potential of our research.

4. Results

This section presents and discusses the experimental results of the different feature selection methods for the MQ2007 dataset. Figure 1 shows a pairplot comparing the nDCG@10 scores of the four implemented methods: Quantum Annealing Mutual Information (Quantum-MI), Hybrid Annealing Mutual Information (Hybrid-MI), Hybrid Annealing Correlation (Hybrid-Correlation), and Simulated Annealing Mutual Information (SA-MI). The pairplot visually represents the score distributions and relationships between the methods.

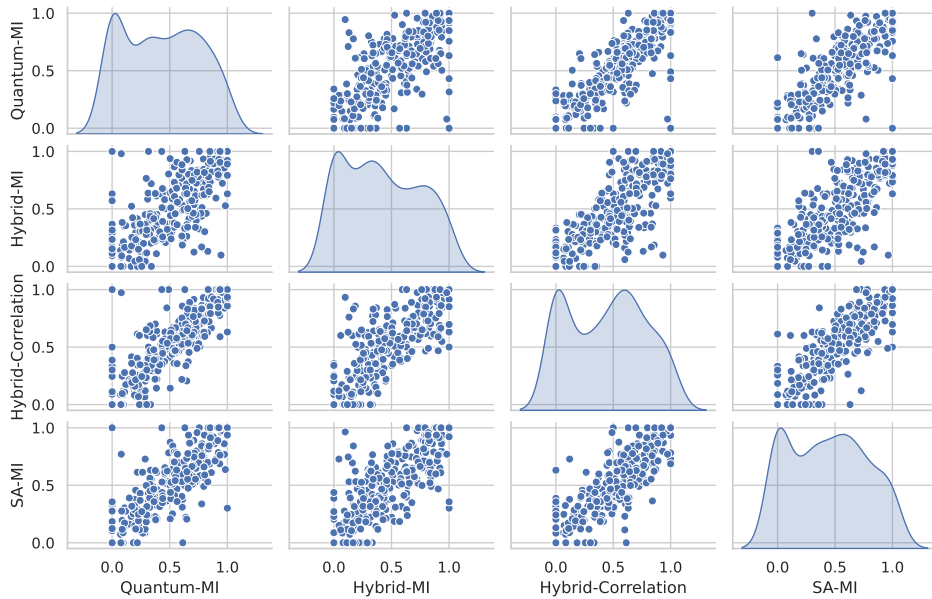


Figure 1: Pairplot of nDCG@10 scores for the different feature selection methods. The pairplot highlights the distributions and interrelationships between Quantum-MI, Hybrid-MI, Hybrid-Correlation, and SA-MI.

Figure 2 presents a boxplot comparing the nDCG@10 scores of the four methods. The boxplot shows each method's median, quartiles, and outliers of the score distributions. The boxplot indicates that Hybrid-Correlation and SA-MI have slightly higher median scores than Quantum-MI and Hybrid-MI. However, there is considerable overlap in the score distributions.

To summarize the performance, we calculated the average nDCG@10 scores for each method based on the provided data files. The average scores are as follows:

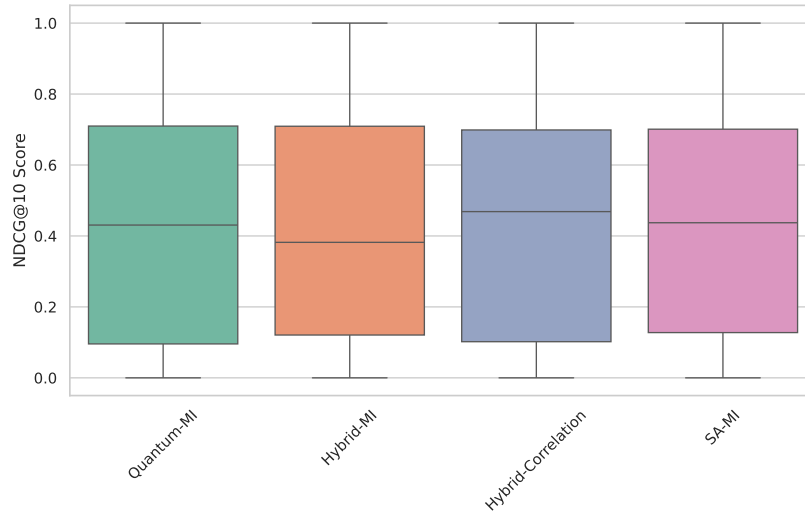


Figure 2: Boxplot of nDCG@10 scores for the different feature selection methods, showing each method's median, quartiles, and outliers.

- Quantum-MI: 0.4299
- Hybrid-MI: 0.4195
- Hybrid-Correlation: 0.4430
- SA-MI: 0.4417

The Hybrid Correlation method achieved the highest average nDCG@10 score (0.4430), followed closely by the SA-MI method (0.4417). The quantum-MI (0.4299) and hybrid-MI (0.4195) methods had slightly lower average scores. These results suggest that the hybrid correlation method, which combines quantum techniques with classical correlation-based feature selection, was the most effective in selecting informative features for the MQ2007 dataset.

The official results shared by the organizers provide additional insights into the performance of the methods:

QTB1A_MQ2007_QA_qtb_NT1	0.4299	356041	Q	13
QTB1A_MQ2007_QA_qtb_NT2	0.4195	500000	H	10
QTB1A_MQ2007_QA_qtb_NT3	0.4434	309227	H	10
QTB1A_MQ2007_SA_qtb_NT1	0.4024	317412	S	10

The official results confirm the average scores obtained from our calculations and provide additional details. The Hybrid-Correlation method (QTB1A_MQ2007_QA_qtb_NT3) achieved the highest score of 0.4434, followed by the Quantum-MI method (QTB1A_MQ2007_QA_qtb_NT1) with a score of 0.4299. The SA-MI method (QTB1A_MQ2007_SA_qtb_NT1) scored 0.4024, slightly lower than our calculated average (0.4417). The Hybrid-MI method (QTB1A_MQ2007_QA_qtb_NT2) had the lowest score of 0.4195 among the four methods. It is worth noting that the Quantum-MI method (QTB1A_MQ2007_QA_qtb_NT1) used 13 features, while the other techniques used ten features. Despite using more features, the Quantum-MI method did not outperform the Hybrid-Correlation method, which used only ten features.

We conducted a one-way ANOVA test on the nDCG@10 scores to validate the statistical significance. The test yielded a p-value of 0.7659, indicating that there are no significant differences between the methods at a significance level of 0.05. This suggests that despite the differences in average scores, the methods' overall performance is statistically comparable.

Despite the lack of significant differences overall, we performed post-hoc pairwise comparisons using Tukey's HSD test to explore the differences further. We summarized Tukey's HSD test results in Table 1.

The pairwise comparisons did not reveal any statistically significant differences between the methods, as indicated by the high p-values (> 0.05) and the confidence intervals that include zero.

Table 1

Tukey’s HSD test results for pairwise comparisons of nDCG@10 scores.

Comparison	Mean Difference	p-value	Lower CI	Upper CI
Hybrid-Correlation vs Quantum-MI	0.0131	0.9000	-0.1420	0.1682
Hybrid-Correlation vs Hybrid-MI	0.0235	0.8326	-0.1316	0.1786
Hybrid-Correlation vs SA-MI	0.0013	1.0000	-0.1538	0.1564
Quantum-MI vs Hybrid-MI	0.0104	0.9154	-0.1447	0.1655
Quantum-MI vs SA-MI	-0.0118	0.9000	-0.1669	0.1433
Hybrid-MI vs SA-MI	-0.0222	0.8326	-0.1773	0.1329

It is important to note that our feature selection methods used only ten features, except for the Quantum-MI method, which used 13 features. Despite using fewer features, our hybrid correlation and SA-MI methods achieved competitive performance compared to other teams with 15 or more features. Our methods effectively selected a compact set of informative features while maintaining good performance. The ability to achieve comparable results with fewer features is advantageous in terms of computational efficiency and model simplicity.

4.1. Analysis of Quantum-Related Methods Performance

The official results and our analysis indicate that the quantum-related methods, particularly Hybrid-Correlation (QTB1A_MQ2007_QA_qtb_NT3) and Quantum-MI (QTB1A_MQ2007_QA_qtb_NT1), achieved competitive performance compared to the fully simulated SA-MI method (QTB1A_MQ2007_SA_qtb_NT1). The Hybrid-Correlation method achieved the highest score of 0.4434, outperforming the SA-MI method, which scored 0.4024. The Quantum-MI method also performed well, with a score of 0.4299, despite using more features (13) than the other methods. The quantum-based approach in Quantum-MI effectively captured relevant information from the additional features.

The superior performance of the hybrid correlation method underscores the potential of combining quantum techniques with classical correlation-based feature selection. It ignites excitement about the future of feature selection tasks. By harnessing the strengths of both approaches, the hybrid correlation was able to identify informative features more effectively than the purely classical SA-MI method. The competitive performance of quantum-related methods, even without statistical significance in overall comparisons, hints at their immense potential in feature selection tasks. The theoretical foundation of quantum-based approaches offers a more nuanced understanding of data correlations and dependencies, a prospect that could significantly enhance their effectiveness in identifying informative features.

However, it is essential to consider the variability in the nDCG@10 scores across different queries, as evident from the pairplot and boxplot. Some queries had scores of 0 for all methods, indicating that none of the selected features were relevant for those specific queries. This variability suggests that the performance of the feature selection methods may depend on the characteristics of individual queries. Future work could explore incorporating query-specific adaptations or using a more diverse set of features to capture the varying information needs of different queries.

4.2. Comparison with Other Teams and Baseline

Our feature selection methods used only ten features, except for the Quantum-MI method, which used 13 features. Despite using a smaller feature set, our hybrid correlation and SA-MI methods achieved competitive performance compared to other teams with 15 or more features. It demonstrates the effectiveness of our processes in selecting informative features while maintaining a compact representation.

The decision to use fewer features was not arbitrary but a strategic move that offered several key advantages. It significantly reduces the computational complexity of the learning algorithms, as there are fewer dimensions to process. This, in turn, helps to mitigate the risk of overfitting, as the models are less likely to capture noise or irrelevant patterns when trained on a smaller set of relevant features. Moreover, a compact feature set enhances the interpretability of the models, as it hones in on the most important aspects of the data, enlightening us to the true essence of the information.

The fact that our methods achieved comparable performance to teams using more features suggests that our feature selection approaches were able to identify the most discriminative and informative features for the MQ2007 dataset. It highlights the potential of our methods to extract meaningful information from the data while maintaining efficiency and simplicity.

5. Conclusions and Future Work

In this paper, we presented a novel approach to feature selection for Information Retrieval using quantum annealing. Our focus was on the MQ2007 dataset, where we explored various formulations of the feature selection problem as quantum annealing problems. Our approach demonstrated potential in selecting informative features, which contributed to the performance of the LambdaMART learning-to-rank model. When compared with classical feature selection techniques, our quantum annealing-based approaches showed competitive results. The evaluation conducted by the QuantumCLEF lab organizers indicated that the features selected by our approach contributed to improvements in ranking quality, although the differences were not statistically significant. Our work provides insights into the application of quantum annealing for addressing optimization problems in Information Retrieval. However, we acknowledge that there are limitations and challenges that need to be addressed in future work. Future research directions include:

1. Scaling quantum annealing-based feature selection to larger datasets and higher-dimensional feature spaces.
2. Integrating quantum annealing-based feature selection with other components of the Information Retrieval pipeline.
3. Adapting the approach to other Information Retrieval tasks, such as document clustering and recommendation systems.
4. Exploring the potential of newer generations of quantum hardware and algorithms for Information Retrieval applications.

In conclusion, our work demonstrates the feasibility of using quantum annealing for feature selection in Information Retrieval. While the current results do not show significant improvements over classical methods, they suggest that the intersection of quantum computing and Information Retrieval may hold promise for future advancements in the field. Further research is needed to fully understand and leverage the potential of quantum annealing in improving the efficiency and effectiveness of information access.

References

- [1] Q. Wu, C. J. Burges, K. M. Svore, J. Gao, Adapting boosting for information retrieval measures, *Information Retrieval* 13 (2010) 254–270.
- [2] T.-Y. Liu, *Learning to Rank for Information Retrieval*, Springer Berlin Heidelberg, 2011. doi:10.1007/978-3-642-14267-3.
- [3] G. Roffo, S. Melzi, M. Cristani, Infinite feature selection, in: 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 4202–4210. doi:10.1109/ICCV.2015.478.
- [4] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.

- [5] J. Tang, S. Alelyani, H. Liu, Feature selection for classification: A review, in: *Data Classification: Algorithms and Applications*, CRC Press, 2014, p. 37.
- [6] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, *Machine Learning* 46 (2002) 389–422.
- [7] Y. Saeys, I. Inza, P. Larra naga, A review of feature selection techniques in bioinformatics, *Bioinformatics* 23 (2007) 2507–2517.
- [8] T. Kadowaki, H. Nishimori, Quantum annealing in the transverse ising model, *Physical Review E* 58 (1998) 5355.
- [9] T. Albash, D. A. Lidar, Adiabatic quantum computation, *Reviews of Modern Physics* 90 (2018) 015002.
- [10] H. Wang, W. Wang, Y. Liu, B. Alidaee, Integrating machine learning algorithms with quantum annealing solvers for online fraud detection, *IEEE Access* 10 (2022) 75908–75917. doi:10.1109/ACCESS.2022.3190897.
- [11] T.-Y. Liu, J. Xu, T. Qin, W. Xiong, H. Li, Letor: Benchmark dataset for research on learning to rank for information retrieval, in: *Proceedings of the SIGIR 2007 Workshop on Learning to Rank for Information Retrieval*, 2007, pp. 3–10.
- [12] K. J"arvelin, J. Kek"al"ainen, Cumulated gain-based evaluation of ir techniques, *ACM Transactions on Information Systems (TOIS)* 20 (2002) 422–446.
- [13] D. Jian-peng, L. Hang, P. Xiao-Ling, Z. Chao-Ni, Y. Tian-Huai, J. Xian-Min, Research progress of quantum memory, *Acta Physica Sinica* (2019). doi:10.7498/APS.68.20190039.
- [14] D. ming Huang, T. Chow, Effective feature selection scheme using mutual information, *Neuro-computing* 63 (2005) 325–343. doi:10.1016/j.neucom.2004.01.194.
- [15] J. Sotoca, F. Pla, Supervised feature selection by clustering using conditional mutual information-based distances, *Pattern Recognit.* 43 (2010) 2068–2081. doi:10.1016/j.patcog.2009.12.013.
- [16] H. Wang, W. Wang, Y. Liu, B. Alidaee, Integrating machine learning algorithms with quantum annealing solvers for online fraud detection, *IEEE Access* 10 (2022) 75908–75917.
- [17] W. Qian, W. Shu, Mutual information criterion for feature selection from incomplete data, *Neuro-computing* 168 (2015) 210–220.
- [18] K. Heshami, D. G. England, P. C. Humphreys, P. J. Bustard, V. M. Acosta, J. Nunn, B. J. Sussman, Quantum memories: emerging applications and recent advances, *Journal of modern optics* 63 (2016) 2005–2028.
- [19] H. Ge, T. Hu, Genetic algorithm for feature selection with mutual information, in: *2014 Seventh International Symposium on Computational Intelligence and Design*, volume 1, IEEE, 2014, pp. 116–119.
- [20] H. Liu, J. Sun, L. Liu, H. Zhang, Feature selection with dynamic mutual information, *Pattern Recognition* 42 (2009) 1330–1339.
- [21] B. Yao, C. Li, Y. Chen, Supervised feature selection based on sparse representation and mutual information, in: *2023 IEEE 5th International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*, IEEE, 2023, pp. 1354–1358.
- [22] M. Beraha, A. M. Metelli, M. Papini, A. Tirinzoni, M. Restelli, Feature selection via mutual information: New theoretical insights, in: *2019 international joint conference on neural networks (IJCNN)*, IEEE, 2019, pp. 1–9.
- [23] A. Pasin, M. Ferrari Dacrema, P. Cremonesi, N. Ferro, QuantumCLEF 2024: Overview of the Quantum Computing Challenge for Information Retrieval and Recommender Systems at CLEF, in: *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2024)*, Grenoble, France, September 9th to 12th, 2024, 2024.
- [24] A. Pasin, M. Ferrari Dacrema, P. Cremonesi, N. Ferro, Overview of QuantumCLEF 2024: The Quantum Computing Challenge for Information Retrieval and Recommender Systems at CLEF, in: *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 15th International Conference of the CLEF Association, CLEF 2024*, Grenoble, France, September 9-12, 2024, Proceedings, 2024.