VETOR

Resolução de Problemas de Seleção de Características com Algoritmos Quânticos em Dados Financeiros Reais☆

Solving Feature Selection Problems with Quantum Algorithms on Real Financial Data

Lucas Q. Galvão [1, 2, 3,†], Otto Pires [1, 2, 3], Yan Alef Chagas [1, 2, 3], Maria Heloísa Fraga [1, 2], Marcelo A. Moret [3]

[1] *LAQCC - Latin American Quantum Computing Center, SENAI CIMATEC, Salvador, Brasil.*
[2] *QuIIN - Quantum Industrial Innovation, Centro de Competência Embrapii Cimatec. SENAI CIMATEC, Av. Orlando Gomes, 1845, Salvador, BA, Brazil CEP 41850-010*
[3] *Centro Universitário SENAI CIMATEC, Salvador, Brasil*
[†] **Autor correspondente:** lucas.queiroz@fbter.org.br

Resumo

O setor financeiro enfrenta desafios significativos ao lidar com conjuntos de dados de alta dimensionalidade e um número limitado de amostras, dificultando a construção de modelos preditivos robustos. Técnicas tradicionais de aprendizado de máquina ajudam a mitigar esses problemas, mas a presença de características irrelevantes e redundantes aumenta a complexidade computacional. Este artigo apresenta a aplicação de algoritmos quânticos na seleção de características usando dados reais do setor financeiro, demonstrando que esses algoritmos podem melhorar a eficiência e a precisão dos modelos preditivos. A abordagem envolve a formulação do problema em termos de Otimização Binária Quadrática Irrestrita (QUBO), e sua solução é implementada em simuladores de um annealer quântico. Os experimentos mostram resultados promissores, que são analisados adotando-se a métrica do Critério de Informação de Akaike. Os resultados sugerem que os algoritmos quânticos variacionais têm grande potencial de aplicação se comparados a técnicas tradicionais de seleção de características.

Palavras-chave

Seleção de características • Dados financeiros • Quantum Annealing • QUBO • Aprendizado de Máquina

Abstract

The financial sector faces significant challenges when dealing with high-dimensional datasets and a limited number of samples, making it difficult to build robust predictive models. Traditional machine learning techniques help mitigate these problems, but the presence of irrelevant and redundant features increases computational complexity. This article presents the application of quantum algorithms in feature selection using real data from the financial sector, demonstrating that these algorithms can improve the efficiency and accuracy of predictive models. The approach involves formulating the problem in terms of Unconstrained Quadratic Binary Optimization (QUBO), and its solution implemented in quantum annealer simulators. The experiments show promising results, which are analyzed using the Akaike Information Criterion metric. The results suggest that variational quantum algorithms have great application potential compared to traditional feature selection techniques.

☆ Este artigo é uma versão estendida do trabalho apresentado no XXVII ENMC Encontro Nacional de Modelagem Computacional e XV ECTM Encontro de Ciência e Tecnologia de Materiais, ocorridos em Ilhéus – BA, de 1 a 4 de outubro de 2024.

Keywords

Feature selection • Financial data • Quantum Annealing • QUBO • Machine learning

1 Introdução

Em se tratando de bancos de dados reais, o setor financeiro enfrenta desafios significativos relacionados a conjuntos de dados caracterizados pela alta dimensionalidade e um número limitado de amostras. Essa complexidade aumenta a dificuldade na construção de modelos preditivos robustos e eficazes, essenciais para tarefas críticas como previsão de inadimplência, avaliação de risco de crédito e detecção de fraudes [1]. A utilização de técnicas tradicionais de aprendizado de máquina tem sido fundamental para lidar com esses desafios, permitindo a avaliação da importância das características presentes nos bancos de dados e contribuindo para a redução do *overfitting* e melhoria da interpretabilidade dos modelos [2, 3].

Contudo, a alta dimensionalidade dos dados e a presença de características irrelevantes ou redundantes têm sido um obstáculo significativo. Técnicas de redução de dimensionalidade, como a seleção de características, surgem como uma abordagem crucial para mitigar esses problemas [1]. A seleção de características envolve o processo de escolher um subconjunto ótimo de variáveis de entrada que preservem informações essenciais para a classificação [4, 5, 6]. Este problema é comumente formulado como uma busca exaustiva em um espaço de combinações de características, tornando-se NP-Difícil devido ao seu elevado custo computacional [7, 8].

Na última década, algoritmos quânticos têm surgido como uma alternativa promissora para resolver problemas de otimização combinatória complexos [9]. A formulação desses problemas em termos de Otimização Binária Quadrática Irrestrita (QUBO) possibilita sua solução em hardware quântico, como computadores quânticos e *annealers* [10, 11, 12]. Estudos têm demonstrado que algoritmos quânticos variacionais também podem ser eficazes, oferecendo resultados competitivos e, em certos casos, superiores às técnicas tradicionais de seleção de características [8].

Nesse contexto, este artigo apresenta resultados da aplicação de algoritmos quânticos na seleção de características utilizando dados financeiros reais, uma vez que essa proposta representa uma oportunidade significativa para melhorar a eficiência e a precisão dos modelos preditivos. Estruturalmente, este trabalho está organizado da seguinte forma: na Seção 2, é fornecida uma visão geral sobre o problema de seleção de características e os métodos convencionais para sua resolução; a Seção 3 descreve a abordagem proposta pelos autores; a Seção 4 apresenta os experimentos realizados, cujos resultados foram discutidos na Seção 4.1 e, por fim, as considerações finais e sugestões para pesquisas futuras estão dispostos na Seção 5.

2 Seleção de Características

O custo de treinar um modelo de aprendizado de máquina em um conjunto de dados de treinamento com milhares de características pode se tornar desafiador devido ao aumento da complexidade computacional envolvida. Nesse sentido, o uso de estratégias para redução de dimensionalidade é importante para reduzir os custos de memória durante o treinamento e melhorar o desempenho do preditor [13, 3].

A seleção de características reduz a dimensão do problema ao selecionar um subconjunto de variáveis que podem descrever eficientemente os dados de entrada sem perder informações [7]. Essa abordagem funciona porque a maioria das fontes de dados produz características redundantes ou irrelevantes. Consequentemente, pode-se também ajudar a aumentar a interpretabilidade e a generalização do modelo [14].

Em geral, o processo de seleção de características consiste em quatro etapas principais: geração de subconjuntos, avaliação de subconjuntos, critérios de parada e validação dos resultados [15]. Durante o processo de busca, um subconjunto de características é gerado a partir do conjunto original e é avaliado seguindo um critério escolhido. Esses passos são repetidos até que um ponto de parada predeterminado seja alcançado. Ao final desse processo, o melhor subconjunto das características selecionadas é escolhido e, em seguida, validado no conjunto de teste. No entanto, encontrar o subconjunto ótimo de características é um problema NP-difícil [8]. Portanto, heurísticas e procedimentos subótimos devem ser usados para manter a computação viável, ao mesmo tempo em que produzem bons resultados.

Os métodos de seleção de características são amplamente classificados em métodos de filtro, embutidos ou *wrappers* [16]. Os métodos de filtro classificam as características com base em sua correlação ou dependência e utilizam testes estatísticos que filtram as características abaixo de um certo limiar. O critério de avaliação desempenha um papel vital no modelo de filtro. Técnicas embutidas ou baseadas em modelos incorporam a seleção de características durante o processo de treinamento do modelo de aprendizado, e os resultados dessa seleção são gerados automaticamente quando o treinamento é concluído. Os métodos wrappers utilizam o desempenho do modelo de treinamento como critério para avaliar o subconjunto de características. Isso significa que o modelo é envolto em um algoritmo de busca que encontrará um subconjunto que oferece o maior desempenho. Os métodos wrappers demonstram a

capacidade de alcançar uma maior precisão de classificação e tendem a resultar em um subconjunto menor. No entanto, eles podem sofrer de baixa capacidade de generalização e maior complexidade temporal.

Dentre esses métodos, técnicas de filtragem que envolvem a modelagem do problema de seleção características em um QUBO têm demonstrado potencialidades [17]. Adicionalmente, resultados do campo têm demonstrado vantagens de aplicar técnicas de algoritmos quânticos para solução do QUBO em um *annealer* quântico, visando resolver esses problemas com maior eficiência [18].

2.1 Quantum annealing

Quantum annealing é uma meta-heurística inspirada na computação quântica adiabática [19]. Essa abordagem consiste em preparar uma Hamiltoniana H_p, cujo estado fundamental equivale a solução do problema de interesse [20]. Nesse processo, uma evolução adiabática é usada considerando intervalos de tempo muito curtos ($t \in [0, T]$), evoluindo de uma Hamiltoniana inicial H_0, cujo estado fundamental é conhecido, para a Hamiltoniana de interesse H_p [20, 21]. Essa evolução pode ser visualizada na Hamiltoniana abaixo:

$$H(t) = \left(1 - \frac{t}{T}\right) H_0 + \frac{t}{T} H_p \tag{1}$$

Note que para $t = 0$ a equação (1) se torna $H(0) = H_0$, enquanto que em $t = T$, tem-se $H(T) = H_p$, conforme descrito acima. Essas características tornam a computação quântica adiabática universal com complexidade em $T = O[\exp(\alpha N^{\beta})]$, sendo N o tamanho do problema e α e β constantes positivas características [22]. Contudo, quantum annealing se diferencia da computação quântica adiabática por duas características fundamentais: i) a Hamiltoniana final deve pertencer a uma classe específica, o que torna restringe a universalidade do quantum annealing [21]; ii) não se tem uma garantia de evolução genuinamente adiabática, visto que o tempo de execução do quantum annealing seria muito longo [19].

Referente ao primeiro ponto, uma classe de Hamiltonianas amplamente usadas em quantum annealing é o Modelo de Ising, estudado na Física para descrever a interação ferromagnética de uma cadeia de spins [21, 23]. Esse modelo matemático é descrito em termos da interação spin-spin e a interação individual dos spins com um campo magnético externo, conforme equação abaixo:

$$-\sum_{j,k} J_{jk} \sigma_z^{(j)} \sigma_z^{(k)} - \sum_j h_j \sigma_z^{(j)} \tag{2}$$

Em que $\sigma_z^{(j)}$ é a matriz de Pauli Z aplicada no qubit j, J_{jk} é o coeficiente de interação entre o spin j e o spin k, e o coeficientes h_j representa a influência de um campo magnético externo na partícula j [21]. Já a Hamiltoniana inicial usualmente é representada como $H_0 = -\sum_j^{n-1} \sigma_x^{(j)}$, em que $\sigma_x^{(j)}$ é a atuação da porta de Pauli X no qubit j, e n é o número de qubits usados [21, 20]. Nesse caso, o estado fundamental da Hamiltoniana H_0 pode ser facilmente verificado como $\bigotimes_{i=0}^{n-1} |+\rangle$. Nesse sentido, a Hamiltoniana usada no quantum annealing é da forma:

$$H(t) = -\left(1 - \frac{t}{T}\right) \sum_j^{n-1} \sigma_x^j - \frac{t}{T} \left(\sum_{j,k} J_{jk} z_j z_k + \sum_j h_j z_j \right) \tag{3}$$

Nesse caso, a Hamiltoniana de interesse pode ser representada por uma instância QUBO, descrita por uma matriz simétrica de dimensão $N \times N$. Nesta matriz, os coeficientes h_j são descritos na diagonal principal, enquanto os coeficientes de acoplamento J_{jk} correspondem aos elementos fora da diagonal principal [18]. Portanto, um resultado importante da computação quântica é a possibilidade de mapear problemas de otimização binária para o quantum annealing [21].

3 Método de Seleção de Características

Para a construção da instancia QUBO implementada na seleção de características, foram usados dois algoritmos quânticos: Quantum-Boosting (Q-Boosting) e Quantum-Correlation (Q-Correlation) [18]. Em ambos os casos, a ideia central é verificar o quão correlacionadas as características estão com a variável que indica se o cliente paga ou não a dívida com o banco. Para tanto, usa-se a correlação de Pearson como métrica para essa avaliação, conforme equação abaixo:

$$\text{Corr}(L, M) = \frac{\text{Cov}(L, M)}{\sigma_L \sigma_M} \tag{4}$$

onde $\text{Cov}(L, M)$ é a covariância entre as variáveis aleatórias L e M, e σ_L e σ_M são os seus respectivos desvios padrão [24]. O Algoritmo Q-Boosting aplica a técnica de boosting, que envolve a combinação de modelos simples e

fracos, cada um treinado com um conjunto limitado de características, para formar um modelo forte que emprega apenas as características mais relevantes [17]. Nesse caso, a matriz do QUBO é construída se acordo com a seguinte simetria:

$$Q_{ij} = \begin{cases} \text{Corr}(h_i, h_j) & \text{se } i \neq j \\ \frac{S}{N^2} + \lambda - 2 \cdot \text{Corr}(h_i, y) & \text{se } i = j \end{cases} \tag{5}$$

Em que h_j coincide com a estimativa de um classificador de Máquina de Vetores de Suporte (SVM) treinado apenas com a i-ésima característica do conjunto de dados analisado, N é o número de features, S é o número de samples dentro do dataset, λ é o hiperparâmetro do algoritmo e y a variável alvo [18]. Note que a correlação de Perason é empregada em sua totalidade somente nos elementos fora da diagonal principal

O Q-Correlation é um algoritmo de seleção de características QUBO que calcula a correlação de Pearson (4) entre as características e a variável alvo, bem como entre as próprias características [25]. A construção da matriz QUBO é denotada por:

$$Q = \begin{cases} Q_{ij} = \text{Corr}(f_i, f_j) & \text{se } i \neq j \\ Q_{ij} = \text{Corr}(f_i, y) & \text{se } i = j \end{cases} \tag{6}$$

Construída a instância QUBO com ambas as matrizes, o passo seguinte do algoritmo é codificá-la no modelo de Ising (2) para, em seguida, encaminhá-la para o quantum annealing (3). Conforme discutido, a relação entre os elementos da diagonal principal da matriz QUBO estão diretamente relacionados com os coeficientes de acoplamento J_{jk} da Hamiltoniana, enquanto que o coeficiente h_j se relaciona com os elementos da diagonal principal [18]. Com isso, é possível encontrar a partir da evolução (quase) adiabática do quantum annealing a solução para a instância QUBO que realiza a seleção de características.

3.1 Desenho Experimental

Para nossos experimentos, utilizamos um dataset contendo um total de 5889 características. Esse dataset incluía dados financeiros, que foram empregados para treinar modelos de *machine learning* com o objetivo de prever se um cliente pagaria ou não a dívida com o banco. Para avaliar a qualidade dos modelos, foi usado o método de validação cruzada para calcular a acurácia das predições.

Para a execução do algoritmo, usou-se simuladores da plataforma da D-Wave, amplamente conhecida por dar suporte a códigos que necessitam de *quantum annealing* para serem compilados [21]. Todas as simulações foram executadas no simulador quântico Kuatomu fornecido pelo SENAI-CIMATEC. Kuatomu compreende 192 núcleos de processamento (CPU) com 384 threads, 3 TB de memória RAM e 4 placas aceleradoras GPU NVidia V100 de 32 GB.

4 Resultados

O método descrito na seção anterior possibilitou a obtenção de 392 subconjuntos de características que foram usados para treinar diferentes modelos de aprendizado de máquina. A Tabela 1 apresenta a acurácia dos diferentes modelos de aprendizado de máquina treinados com os menores subconjuntos de características obtidas pelo QUBOCorrelation ou QUBOSVCBoosting. Cada modelo é identificado por um nome (M1 a M10) e possui um número específico de características (k). A acurácia média de cada modelo é acompanhada por seu desvio padrão (σ) obtidos comparando os valores da validação cruzada e pelo algoritmo usado para a seleção de características. Em contrapartida, a Tabela 2 apresenta os 10 modelos com as melhores acurácias observadas, nomeados de M11 a M20.

4.1 Discussão

A análise estatística de dados é um processo de fundamental importância para para a tomada de decisões informadas e baseadas em evidências disponíveis nas características dos próprios dados. Como o objetivo central da pesquisa foi a seleção de características relevantes para o banco, de modo a reduzir a redundância na base de dados, uma análise em primeiro plano é observar os modelos com menor quantidade de características selecionadas listados na Tabela 1.

É possível notar que os três primeiros modelos tiveram uma drástica redução no número de características indo de 5766 para somente 1. Outro ponto importante é o algoritmo usado, em que o QUBOSVCBoosting parece ser mais adequado para a seleção de menos características. Todavia, conforme discutido anteriormente, é preciso se atentar não somente à quantidade de características selecionadas, mas à validação desses modelos.

Nota-se que os modelos com maiores validações têm um grande número de características selecionas, tendo o algoritmo QUBOCorrelation como predominante. Isso pode indicar que esse algoritmo consegue métricas melhores,

Tabela 1: Desempenho de modelos de machine learning com menos características, medido pela acurácia e desvio padrão, utilizando algoritmos QUBOCorrelation e QUBOSVCBoosting.

Modelo	**k**	**Acurácia** $\pm\sigma$	**Algoritmo**
M1	1	0.912 ± 0.021	QUBOCorrelation
M2	1	0.919 ± 0.017	QUBOSVCBoosting
M3	1	0.912 ± 0.019	QUBOSVCBoosting
M4	101	0.912 ± 0.021	QUBOSVCBoosting
M5	121	0.921 ± 0.015	QUBOSVCBoosting
M6	126	0.917 ± 0.012	QUBOCorrelation
M7	198	0.919 ± 0.023	QUBOSVCBoosting
M8	241	0.917 ± 0.016	QUBOSVCBoosting
M9	249	0.924 ± 0.013	QUBOCorrelation
M10	294	0.921 ± 0.021	QUBOSVCBoosting

Tabela 2: Desempenho de modelos de machine learning com maior acurácia, medido pela acurácia e desvio padrão, utilizando algoritmos QUBOCorrelation e QUBOSVCBoosting.

Modelo	**k**	**Acurácia** $\pm\sigma$	**Algoritmo**
M11	3048	0.931 ± 0.012	QUBOCorrelation
M12	2875	0.929 ± 0.017	QUBOCorrelation
M13	2867	0.929 ± 0.011	QUBOCorrelation
M14	2879	0.929 ± 0.011	QUBOCorrelation
M15	2866	0.929 ± 0.018	QUBOCorrelation
M16	2909	0.929 ± 0.011	QUBOCorrelation
M17	3046	0.929 ± 0.011	QUBOCorrelation
M18	5208	0.929 ± 0.017	QUBOCorrelation
M19	2828	0.929 ± 0.015	QUBOSVCBoosting
M20	2834	0.929 ± 0.011	QUBOSVCBoosting

mas seleciona mais características. No entanto, a validação é extremamente próxima dos modelos da Tabela 1. Se considerarmos o desvio padrão, é possível visualizar que os modelos estão no mesmo intervalo de confiança, o que torna difícil a escolha por um melhor modelo. Nesse caso, é necessário o uso métricas estatísticas capazes de validar a escolha do modelo adequado

Para essa análise, foi considerado o Critério de informação de Akaike (AIC), uma métrica usada para a seleção de modelos estatísticos, balanceando a qualidade do ajuste do modelo e sua complexidade. Introduzido por Hirotsugu Akaike em 1974, o AIC estima a quantidade de informação perdida ao usar um modelo para representar o processo que gerou os dados [26]. Modelos com menor valor de AIC são preferíveis, pois indicam um melhor equilíbrio entre ajuste e simplicidade, penalizando o número de parâmetros do modelo e bonificando a sua verossimilhança [27], conforme equação 7. Ele é particularmente útil na comparação de modelos diferentes e é amplamente usado em campos como econometria, biologia, e aprendizado de máquina.

$$AIC = 2k - 2log(\mathcal{L}) \tag{7}$$

Em resumo, essa métrica foi aplicada aos elementos do banco de dados com os resultados das seleções de características para a escolha dos modelos mais adequados. Para tanto, considerou-se o número de parâmetros do sistema como equivalente ao número de características e a verossimilhança como equivalente à validação dos modelos. Na Tabela 3, é possível observar os 10 modelos com menores AIC escolhidos no sistema.

Nesse caso, é possível visualizar que o AIC teve preferência por dados com menor número de características. De

Tabela 3: Tabela de melhores valores de AIC.

Modelo	N	Validação	Test_Ac	Algoritmo	AIC
M2	1	0.919 ± 0.017	0.903	QUBOSVCBoosting	0.169
M1	1	0.912 ± 0.021	0.910	QUBOCorrelation	0.184
M3	1	0.912 ± 0.019	0.897	QUBOSVCBoosting	0.184
M4	101	0.912 ± 0.021	0.917	QUBOSVCBoosting	0.207
M5	121	0.921 ± 0.015	0.910	QUBOSVCBoosting	0.216
M6	126	0.917 ± 0.012	0.903	QUBOCorrelation	0.219
M7	198	0.919 ± 0.023	0.913	QUBOSVCBoosting	0.238
M8	241	0.917 ± 0.016	0.903	QUBOSVCBoosting	0.244
M9	249	0.924 ± 0.013	0.903	QUBOCorrelation	0.256
M10	294	0.921 ± 0.021	0.910	QUBOSVCBoosting	0.267

fato, se os modelos têm verossimilhanças muito próximas, é preferível a escolha por modelos com menor número de parâmetro. Essa observação está fortemente relacionada à navalha de Occan, em que modelos mais simples e fidedignos melhor descrevem algum fenômeno [28].

Para além dessas observações, uma outra métrica importante para a avaliação dos modelo escolhidos pelo critério de informação Akaike é a acurácia dos dados de teste, representado na Tabela 3 pela coluna *Teste_Ac*. Nesse caso, observa-se que M4 é o modelo com melhor teste de acurácia dentre todos os encontrados considerando o critério de Akaike, enquanto o M3 é o pior modelo.

5 Considerações finais

A partir do presente trabalho, conseguimos explorar a aplicação de otimização combinatória baseados em QUBO para seleção de características que são relevantes para o setor financeiro, mais especificamente tratando questões de previsão de inadimplência. Através de métricas usadas nos algoritmos QBooosting e QCorrelation, foi possível medir a relação entre as características e as variáveis trabalhadas, além de avaliar a performance dos modelos de aprendizado de máquina implementando em um simulador de *annealer* quântico.

Considerando os resultados, nota-se que ambos algoritmos reduzem as quantidades de features, tendo o Q-Boosting utilizado menos features, porém com Q-Correlation apresentando melhores métricas de acurácia. A análise estatística demonstrou a importância da validação cruzada e o Critério de Informação de Akaike (AIC) para trazer um equilíbrio entre a simplicidade do modelo e a sua precisão. Os modelos gerados apresentaram um desempenho variado, com os melhores alcançando uma acurácia média superior a 0.92. No entanto, o fato de a métrica AIC ter optado por modelos com menos features pode indicar pouca sensibilidade a pequenas variações de verossimilhança, sendo necessária a análise interpretativa de profissionais da área de finanças capazes de validar a quantidade de features selecionadas.

Apesar dos resultados, foram identificados algumas limitações referentes à dificuldade de o modelo aprender padrões complexos em dados financeiros. Como perspectivas futuras é possível explorar novos métodos de abordagens híbridos ou totalmente quânticos para identificação desses padrões. Outrossim, a otimização dos parâmetros dos algoritmos e busca por outras métricas são questões que devem ser consideradas. A contínua evolução dos algoritmos quânticos e o aumento da capacidade computacional das plataformas de *quantum annealing* são promissores para a obtenção de modelos mais robustos e eficientes na seleção de características e previsão de padrões financeiros.

Agradecimentos

Esse artigo foi parcialmente financiado pelo Projeto de formação iNOVATeQ Lato Senso Especialização em Computação Quântica, apoiado pelo QuIIN - Centro de Competência EMBRAPII CIMATEC em Tecnologias Quânticas, com recursos financeiros do PPI IoT/Manufatura 4.0/PPI HardwareBR da bolsa MCTI número 053/2023, assinado com EMBRAPII.

Referências

[1] D. Liang, C.-F. Tsai, e H.-T. Wu, "The effect of feature selection on financial distress prediction," *Knowl. Based Syst.*, vol. 73, pp. 289–297, 2015.

[2] J. Cai, J. Luo, S. Wang, e S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, 2018.

[3] S. Velliangiri, S. Alagumuthukrishnan, e S. I. T. Joseph, "A review of dimensionality reduction techniques for efficient computation," *Procedia Computer Science*, 2019. Disponível em: https://doi.org/10.1016/j.procs.2020.01.079

[4] L. van der Maaten, E. O. Postma, e J. van den Herik, "Dimensionality reduction: A comparative review," 2009. Disponível em: https://api.semanticscholar.org/CorpusID:12051918

[5] C. Sorzano, J. Vargas, e A. Montano, "A survey of dimensionality reduction techniques," 03 2014.

[6] S. Ayesha, M. K. Hanif, e R. Talib, "Overview and comparative study of dimensionality reduction techniques for high dimensional data," *Information Fusion*, vol. 59, pp. 44–58, 2020. Disponível em: https://doi.org/10.1016/j.inffus.2020.01.005

[7] S. Khalid, T. Khalil, e S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," em *2014 Science and Information Conference*, 2014, pp. 372–378.

[8] C. Zoufal, R. V. Mishmash, N. Sharma, N. Kumar, A. Sheshadri, A. Deshmukh, N. Ibrahim, J. Gacon, e S. Woerner, "Variational quantum algorithm for unconstrained black box binary optimization: Application to feature selection," *Quantum*, vol. 7, p. 909, 2023. Disponível em: https://doi.org/10.22331/q-2023-01-26-909

[9] A. Abbas, A. Ambainis, B. Augustino, A. Bärtschi, H. Buhrman, C. Coffrin, G. Cortiana, V. Dunjko, D. J. Egger, B. G. Elmegreen *et al.*, "Challenges and opportunities in quantum optimization," *Nature Reviews Physics*, pp. 1–18, 2024. Disponível em: https://doi.org/10.1038/s42254-024-00770-9

[10] S. Mücke, R. Heese, S. Müller, M. Wolter, e N. Piatkowski, "Feature selection on quantum computers," *Quantum Machine Intelligence*, vol. 5, no. 1, feb 2023. Disponível em: https://doi.org/10.1007%2Fs42484-023-00099-z

[11] G. Hellstern, V. Dehn, e M. Zaefferer, "Quantum computer based feature selection in machine learning," *IET Quantum Communication*, vol. 5, no. 3, p. 232–252, 2024. Disponível em: https://doi.org/10.1049/qtc2.12086

[12] R. Nembrini, M. Ferrari Dacrema, e P. Cremonesi, "Feature selection for recommender systems with quantum computing," *Entropy*, vol. 23, no. 8, 2021. Disponível em: https://www.mdpi.com/1099-4300/23/8/970

[13] K. Zhang, Y. Zheng, e Z. Li, "Dimension reduction for efficient data-enabled predictive control," *IEEE Control Systems Letters*, vol. 7, pp. 3277–3282, 2022. Disponível em: https://doi.org/10.1109/LCSYS.2023.3322965

[14] M. I. Jordan e T. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, pp. 255 – 260, 2015. Disponível em: https://doi.org/10.1126/science.aaa8415

[15] M. Rostami, K. Berahmand, E. Nasiri, e S. Forouzandeh, "Review of swarm intelligence-based feature selection methods," *Engineering Applications of Artificial Intelligence*, vol. 100, p. 104210, 2021. Disponível em: https://www.sciencedirect.com/science/article/pii/S0952197621000579

[16] J. Cai, J. Luo, S. Wang, e S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, 2018. Disponível em: https://www.sciencedirect.com/science/article/pii/S0925231218302911

[17] A. J. Ferreira e M. A. T. Figueiredo, *Boosting Algorithms: A Review of Methods, Theory, and Applications*. New York, NY: Springer New York, 2012, pp. 35–85. Disponível em: https://doi.org/10.1007/978-1-4419-9326-7_2

[18] F. Moroni, "A survey on feature selection methods for classification solved with quantum annealing," Dissertação de mestrado, Politecnico di Milano, 2021. Disponível em: https://www.politesi.polimi.it/retrieve/04b2f778-f24a-46a2-9532-8aba12e74652/2021_12_Moroni.pdf

[19] E. K. Grant e T. S. Humble, "Adiabatic quantum computing and quantum annealing," em *Oxford Research Encyclopedia of Physics*, 2020.

[20] L. Andrew, "Ising formulations of many NP problems," *Frontiers in physics*, vol. 2, p. 5, 2014. Disponível em: https://doi.org/10.3389/fphy.2014.00005

[21] E. F. Combarro, S. González-Castillo, e A. Di Meglio, *A Practical Guide to Quantum Machine Learning and Quantum Optimization: Hands-on Approach to Modern Quantum Algorithms*. Packt Publishing Ltd, 2023.

[22] W. Van Dam, M. Mosca, e U. Vazirani, "How powerful is adiabatic quantum computation?" em *Proceedings 42nd IEEE symposium on foundations of computer science*. IEEE, 2001, pp. 279–287.

[23] A. Cervera-Lierta, "Exact ising model simulation on a quantum computer," *Quantum*, vol. 2, p. 114, 2018. Disponível em: https://doi.org/10.22331/q-2018-12-21-114

[24] J. Xu, B. Tang, H. He, e H. Man, "Semisupervised feature selection based on relevance and redundancy criteria," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, pp. 1974–1984, 2017. Disponível em: https://doi.org/10.1109/TNNLS.2016.2562670

[25] R. K. Nath, H. Thapliyal, e T. S. Humble, "Quantum annealing for automated feature selection in stress detection," em *2021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*. IEEE, 2021, pp. 453–457.

[26] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974. Disponível em: https://doi.org/10.1109/TAC.1974.1100705

[27] S. Portet, "A primer on model selection using the Akaike information criterion," *Infectious Disease Modelling*, vol. 5, pp. 111 – 128, 2020. Disponível em: https://doi.org/10.1016/j.idm.2019.12.010

[28] A. Blumer, A. Ehrenfeucht, D. Haussler, e M. K. Warmuth, "Occam's razor," *Information Processing Letters*, vol. 24, no. 6, pp. 377–380, 1987. Disponível em: https://doi.org/10.1016/0020-0190(87)90114-1