



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

SÉRGIO DE VASCONCELOS FILHO

Expressividade e emaranhamento de circuitos quânticos variacionais: uma análise em bases
de sentimento

Recife

2025

SÉRGIO DE VASCONCELOS FILHO

Expressividade e emaranhamento de circuitos quânticos variacionais: uma análise em bases de sentimento

Trabalho apresentado ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Mestre Profissional em Ciência da Computação.

Área de Concentração: inteligência computacional

Orientador (a): Fernando Maciano de Paula Neto

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Filho, Sérgio de Vasconcelos.

Expressividade e emaranhamento de circuitos quânticos
variacionais: uma análise em bases de sentimento / Sérgio de
Vasconcelos Filho. - Recife, 2025.

162f.: il.

Dissertação (Mestrado)- Universidade Federal de Pernambuco,
Centro de Informática, Programa de Pós-graduação Acadêmica em
Ciência da Computação, 2025.

Orientação: Fernando Maciano de Paula Neto.

1. Aprendizagem de máquina quântica; 2. Circuitos quânticos
variacionais; 3. Análise de sentimentos; 4. Processamento
quântico de linguagem natural; 5. Emaranhamento quântico. I.
Paula Neto, Fernando Maciano de. II. Título.

UFPE-Biblioteca Central

Sérgio de Vasconcelos Filho

“Expressividade e emaranhamento de circuitos quânticos variacionais: uma análise em bases de sentimento”

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação. Área de Concentração: Inteligência Computacional.

Aprovado em: 30/07/2025.

BANCA EXAMINADORA

Prof. Dr. Paulo Salgado Gomes de Mattos Neto
Centro de Informática / UFPE

Prof. Dr. Wilson Rosa de Oliveira Junior
Departamento de Estatística e Informática / UFRPE

Prof. Dr. Fernando Maciano de Paula Neto
Centro de Informática / UFPE
(orientador)

Dedico ao Sérgio do passado, do passado aquele que ele teve que superar para chegar até aqui.

AGRADECIMENTOS

À minha mãe, Ana, e à minha irmã, Camilla, pelas suas orações incessantes e apoio incondicional que me dão forças para não desistir dos meus sonhos. É muito bom saber que há um lugar para voltar caso tudo falhe.

Aos meus amigos antigos, aos novos e todos aqueles entre esse tempo, pela compreensão da minha ausência justificada para a realização deste trabalho e pelo envio das boas vibrações. Sinto-me reconfortado ao ver na pergunta sobre a minha defesa a animação da minha conquista como também com a ideia de me ver finalmente.

À Bruna, no qual seu silêncio me fez me escutar mais do que eu já havia feito na minha vida inteira.

A Fernando, por sempre acreditar na minha capacidade e na sua incomensurável paciência para comigo, me mostrando que a resiliência sempre é a resposta quando a adversidade é apresentada.

Ao Centro de Informática, por ser um centro de referência nacional e ofertar o melhor ensino possível perto de onde nasci.

À Universidade Federal de Pernambuco, por me remover da miséria e me permitir oferecer uma melhor qualidade de vida para mim e para minha família.

"Estou convencido de que Ele não joga dados." (EINSTEIN; BORN, 2005).

RESUMO

Circuitos Quânticos Variacionais (CQV) são uma abordagem promissora para o Processamento de Linguagem Natural Quântico. Contudo, persiste uma lacuna entre as propriedades teóricas dos circuitos, como expressabilidade e emaranhamento, e seu desempenho empírico. Esta dissertação investiga sistematicamente essa relação na tarefa de análise de sentimentos, avaliando o impacto da representação de entrada e comparando com *benchmarks* clássicos. Para tal, foram avaliadas 15 arquiteturas de CQV em baixa ($L = 1$) e alta ($L = 10$) profundidade, quantificando-se sua expressabilidade e poder de emaranhamento. Os modelos foram testados em quatro bases de dados (três sintéticas e a pública *Stanford Sentiment Treebank* - SST) utilizando diferentes *embeddings* e dimensionalidades. A performance (métrica F1) foi comparada a 7 modelos clássicos e 26 configurações de *ensembles* com validação estatística. Os resultados confirmam que emaranhamento e profundidade adequada são requisitos para o desempenho em cenários complexos; o aumento da profundidade de $L = 1$ para $L = 10$ foi fundamental para a performance dos circuitos emaranhadores, que também demonstraram maior robustez à redução de dimensionalidade. Contudo, os modelos clássicos de referência, notadamente as Máquinas de Vetores de Suporte, apresentaram desempenho superior na base SST, não sendo observada uma vantagem quântica. Em contrapartida, nas bases sintéticas, diversos modelos quânticos profundos alcançaram desempenho estatisticamente equivalente aos melhores *benchmarks* clássicos, evidenciando sua competitividade. Este trabalho estabelece, portanto, uma ponte empírica entre teoria e prática no projeto de CQVs, validando a relevância do emaranhamento e da profundidade e fornecendo diretrizes para arquiteturas quânticas mais eficazes.

Palavras-chaves: aprendizagem de máquina quântica; circuitos quânticos variacionais; processamento quântico de linguagem natural; análise de sentimentos; emaranhamento quântico.

ABSTRACT

Variational Quantum Circuits (VQCs) are a promising approach for Quantum Natural Language Processing. However, a gap persists between the theoretical properties of circuits, such as expressibility and entanglement, and their empirical performance. This dissertation systematically investigates this relationship in the sentiment analysis task, evaluating the impact of the input representation and comparing against classical benchmarks. To this end, 15 VQC architectures were evaluated at low ($L = 1$) and high ($L = 10$) depth, quantifying their expressibility and entanglement power. The models were tested on four datasets (three synthetic and the public *Stanford Sentiment Treebank* - SST) using different embeddings and dimensionalities. Performance (F1-score) was compared to 7 classical models and 26 ensemble configurations with rigorous statistical validation. The results confirm that entanglement and adequate depth are requirements for performance in complex scenarios; the increase in depth from $L = 1$ to $L = 10$ was fundamental for the performance of entangling circuits, which also demonstrated greater robustness to dimensionality reduction. However, the classical reference models, notably Support Vector Machines, showed superior performance on the SST dataset, and no quantum advantage was observed. In contrast, on the synthetic datasets, several deep quantum models achieved a performance statistically equivalent to the best classical benchmarks, showcasing their competitiveness. This work, therefore, establishes an empirical bridge between theory and practice in VQC design, validating the relevance of entanglement and depth and providing guidelines for more effective quantum architectures.

Keywords: quantum machine learning; variational quantum circuits. quantum natural language processing; sentiment analysis. quantum entanglement.

LISTA DE FIGURAS

- Figura 1 – Representação geométrica do estado de um *qubit* na Esfera de Bloch. Os polos norte e sul correspondem aos estados da base computacional $|0\rangle$ e $|1\rangle$, respectivamente, enquanto os pontos na superfície representam os estados de superposição. 36
- Figura 2 – Exemplo de um circuito quântico elementar que gera um estado de Bell, um dos estados maximamente emaranhados de dois *qubits*. O circuito opera sobre dois *qubits*, $|q_0\rangle$ e $|q_1\rangle$, e um bit clássico, c_0 . Uma porta de Hadamard (H) é aplicada ao primeiro *qubit* para criar uma superposição, seguida por uma porta CNOT, que emaranha os dois *qubits*. Por fim, uma medição na base do observável Z é realizada no primeiro *qubit*, e o resultado clássico é armazenado no bit c_0 39
- Figura 3 – Template de circuito com preparação de estado por amplitude (PE), o bloco de *ansatz* variável através das possibilidades da Figura 4 e medição pelo observável Z. 58
- Figura 4 – Conjunto de *ansatzes* utilizados com seus códigos identificação única. . . . 61
- Figura 5 – Fluxo geral do método experimental, detalhando as etapas sequenciais: (a) seleção e preparação das bases de dados; (b) conversão do texto em vetores numéricos (embeddings); (c) análise da dimensionalidade do vetor de entrada; e (d) treinamento e avaliação dos modelos de classificação quânticos e clássicos, incluindo seus *ensembles*. 63
- Figura 6 – Resultados das medidas de expressabilidade e emaranhamento e a métrica F1 na base *Stanford Sentiment Treebank* (SST). Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados. Cada subíndice no canto inferior direito refere-se ao código do circuito equivalente. 68
- Figura 7 – Resultados do desempenho F1 dos circuitos 1 ao 6 na base SST. Esses algoritmos classificadores demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon. 70

Figura 8 – Resultados do desempenho F1 dos circuitos 7 ao 11 na base SST. Esses algoritmos classificadores implementam diferentes formas de emaranhamento quântico. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	71
Figura 9 – Resultados do desempenho F1 dos circuitos 12 ao 15 na base SST. Esses algoritmos classificadores implementam diferentes tipos de conectividade entre <i>qubits</i> . As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	72
Figura 10 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 na base SST. Esses algoritmos classificadores implementam modelos clássicos lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	73
Figura 11 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 na base SST. Esses algoritmos classificadores implementam modelos clássicos não-lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	74
Figura 12 – Resultados do desempenho F1 dos circuitos 16 ao 21 na base SST. Esses algoritmos classificadores implementam modelos de ensemble através de <i>AdaBoost</i> ou <i>Bagging</i> de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	75
Figura 13 – Resultados do desempenho F1 dos circuitos 22 ao 31 na base SST. Esses algoritmos classificadores implementam modelos de ensemble através de <i>hard</i> ou <i>soft voting</i> de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	76
Figura 14 – Resultados do desempenho F1 dos circuitos 39 ao 44 na base SST. Esses algoritmos classificadores implementam modelos de ensemble clássicos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	77
Figura 15 – Resultados das medidas de expressabilidade e emaranhamento e a métrica F1 na base <i>ChatGPT</i> - Fácil. Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados. Cada subíndice no canto inferior direito refere-se ao código do circuito equivalente.	78

Figura 16 – Resultados das medidas de expressabilidade e emaranhamento e a métrica F1 na base <i>ChatGPT</i> - Médio. Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados. Cada subíndice no canto inferior direito refere-se ao código do circuito equivalente.	79
Figura 17 – Resultados das medidas de expressabilidade e emaranhamento e a métrica F1 na base <i>ChatGPT</i> - Difícil. Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados. Cada subíndice no canto inferior direito refere-se ao código do circuito equivalente.	80
Figura 18 – Resultados da medida de expressabilidade dos circuitos e o impacto da interferência da Preparação de Estados (PE). Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados.	81
Figura 19 – Resultados da medida de emaranhamento dos circuitos e o impacto da interferência da PE.	82
Figura 20 – Resultados do desempenho F1 dos circuitos 1 ao 6 com diferentes tamanhos de vetor de entrada na base <i>ChatGPT</i> - Fácil. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.	85
Figura 21 – Resultados do desempenho F1 dos circuitos 7 ao 11 com diferentes tamanhos de vetor de entrada na base <i>ChatGPT</i> - Fácil. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.	86

- Figura 22 – Resultados do desempenho F1 dos circuitos 12 ao 15 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Fácil. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 87
- Figura 23 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Fácil. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 88
- Figura 24 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Fácil. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 89
- Figura 25 – Resultados do desempenho F1 dos circuitos 1 ao 6 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 90

- Figura 26 – Resultados do desempenho F1 dos circuitos 7 ao 11 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 91
- Figura 27 – Resultados do desempenho F1 dos circuitos 12 ao 15 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 92
- Figura 28 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 93
- Figura 29 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 94

- Figura 30 – Resultados do desempenho F1 dos circuitos 1 ao 6 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 95
- Figura 31 – Resultados do desempenho F1 dos circuitos 7 ao 11 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 96
- Figura 32 – Resultados do desempenho F1 dos circuitos 12 ao 15 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. . . . 97
- Figura 33 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 98

- Figura 34 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 99
- Figura 35 – Resultados do desempenho F1 dos circuitos 1 ao 6 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 103
- Figura 36 – Resultados do desempenho F1 dos circuitos 7 ao 11 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 104

- Figura 37 – Resultados do desempenho F1 dos circuitos 12 ao 15 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 105
- Figura 38 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. . 106
- Figura 39 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 107

- Figura 40 – Resultados do desempenho F1 dos circuitos 1 ao 6 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 108
- Figura 41 – Resultados do desempenho F1 dos circuitos 7 ao 11 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 109
- Figura 42 – Resultados do desempenho F1 dos circuitos 12 ao 15 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 110

Figura 43 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. . 111

Figura 44 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 112

Figura 45 – Resultados do desempenho F1 dos circuitos 1 ao 6 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 113

- Figura 46 – Resultados do desempenho F1 dos circuitos 7 ao 11 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 114
- Figura 47 – Resultados do desempenho F1 dos circuitos 12 ao 15 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. 115
- Figura 48 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm. . 116

Figura 49 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com os modelos geradores de <i>embeddings</i> na base <i>ChatGPT</i> - Difícil. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo <i>Matryoshka</i> , divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.	117
Figura 50 – Resultados do desempenho F1 dos circuitos 1 ao 6 na base <i>ChatGPT</i> - Fácil. Esses algoritmos classificadores demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	121
Figura 51 – Resultados do desempenho F1 dos circuitos 7 ao 11 na base <i>ChatGPT</i> - Fácil. Esses algoritmos classificadores implementam diferentes formas de emaranhamento quântico. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	122
Figura 52 – Resultados do desempenho F1 dos circuitos 12 ao 15 na base <i>ChatGPT</i> - Fácil. Esses algoritmos classificadores implementam diferentes tipos de conectividade entre <i>qubits</i> . As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	123
Figura 53 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 na base <i>ChatGPT</i> - Fácil. Esses algoritmos classificadores implementam modelos clássicos lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	124
Figura 54 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 na base <i>ChatGPT</i> - Fácil. Esses algoritmos classificadores implementam modelos clássicos não-lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	125
Figura 55 – Resultados do desempenho F1 dos circuitos 16 ao 21 na base <i>ChatGPT</i> - Fácil. Esses algoritmos classificadores implementam modelos de ensemble através de <i>AdaBoost</i> ou <i>Bagging</i> de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	126

Figura 56 – Resultados do desempenho F1 dos circuitos 22 ao 31 na base <i>ChatGPT</i> - Fácil. Esses algoritmos classificadores implementam modelos de ensemble através de <i>hard</i> ou <i>soft voting</i> de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	127
Figura 57 – Resultados do desempenho F1 dos circuitos 39 ao 44 na base <i>ChatGPT</i> - Fácil. Esses algoritmos classificadores implementam modelos de ensemble clássicos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	128
Figura 58 – Resultados do desempenho F1 dos circuitos 1 ao 6 na base <i>ChatGPT</i> - Médio. Esses algoritmos classificadores demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	129
Figura 59 – Resultados do desempenho F1 dos circuitos 7 ao 11 na base <i>ChatGPT</i> - Médio. Esses algoritmos classificadores implementam diferentes formas de emaranhamento quântico. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	130
Figura 60 – Resultados do desempenho F1 dos circuitos 12 ao 15 na base <i>ChatGPT</i> - Médio. Esses algoritmos classificadores implementam diferentes tipos de conectividade entre <i>qubits</i> . As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	131
Figura 61 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 na base <i>ChatGPT</i> - Médio. Esses algoritmos classificadores implementam modelos clássicos lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	132
Figura 62 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 na base <i>ChatGPT</i> - Médio. Esses algoritmos classificadores implementam modelos clássicos não-lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	133
Figura 63 – Resultados do desempenho F1 dos circuitos 16 ao 21 na base <i>ChatGPT</i> - Médio. Esses algoritmos classificadores implementam modelos de ensemble através de <i>AdaBoost</i> ou <i>Bagging</i> de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	134

Figura 64 – Resultados do desempenho F1 dos circuitos 22 ao 31 na base <i>ChatGPT</i> - Médio. Esses algoritmos classificadores implementam modelos de ensemble através de <i>hard</i> ou <i>soft voting</i> de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	135
Figura 65 – Resultados do desempenho F1 dos circuitos 39 ao 44 na base <i>ChatGPT</i> - Médio. Esses algoritmos classificadores implementam modelos de ensemble clássicos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	136
Figura 66 – Resultados do desempenho F1 dos circuitos 1 ao 6 na base <i>ChatGPT</i> - Difícil. Esses algoritmos classificadores demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	137
Figura 67 – Resultados do desempenho F1 dos circuitos 7 ao 11 na base <i>ChatGPT</i> - Difícil. Esses algoritmos classificadores implementam diferentes formas de emaranhamento quântico. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	138
Figura 68 – Resultados do desempenho F1 dos circuitos 12 ao 15 na base <i>ChatGPT</i> - Difícil. Esses algoritmos classificadores implementam diferentes tipos de conectividade entre <i>qubits</i> . As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	139
Figura 69 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 na base <i>ChatGPT</i> - Difícil. Esses algoritmos classificadores implementam modelos clássicos lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	140
Figura 70 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 na base <i>ChatGPT</i> - Difícil. Esses algoritmos classificadores implementam modelos clássicos não-lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	141
Figura 71 – Resultados do desempenho F1 dos circuitos 16 ao 21 na base <i>ChatGPT</i> - Difícil. Esses algoritmos classificadores implementam modelos de ensemble através de <i>AdaBoost</i> ou <i>Bagging</i> de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	142

Figura 72 – Resultados do desempenho F1 dos circuitos 22 ao 31 na base <i>ChatGPT</i> - Difícil. Esses algoritmos classificadores implementam modelos de ensemble através de <i>hard</i> ou <i>soft voting</i> de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	143
Figura 73 – Resultados do desempenho F1 dos circuitos 39 ao 44 na base <i>ChatGPT</i> - Difícil. Esses algoritmos classificadores implementam modelos de ensemble clássicos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.	144
Figura 74 – Diagrama de diferença crítica da performance F1 na base <i>ChatGPT</i> - Fácil. Foram exibidos apenas cliques que contenham modelos clássicos, de forma <i>ensemble</i> ou isolada, e os modelos quânticos envolvidos (nas suas versões $L = 1$ e $L = 10$).	147
Figura 75 – Diagrama de diferença crítica da performance F1 na base <i>ChatGPT</i> - Médio. Foram exibidos apenas cliques que contenham modelos clássicos, de forma <i>ensemble</i> ou isolada, e os modelos quânticos envolvidos (nas suas versões $L = 1$ e $L = 10$).	148
Figura 76 – Diagrama de diferença crítica da performance F1 na base <i>ChatGPT</i> - Difícil. Foram exibidos apenas cliques que contenham modelos clássicos, de forma <i>ensemble</i> ou isolada, e os modelos quânticos envolvidos (nas suas versões $L = 1$ e $L = 10$).	149
Figura 77 – Diagrama de diferença crítica da performance F1 na base SST. Foram exibidos apenas cliques que contenham modelos clássicos, de forma <i>ensemble</i> ou isolada, e os modelos quânticos envolvidos (nas suas versões $L = 1$ e $L = 10$).	150

LISTA DE QUADROS

Quadro 1 – <i>Ensemble</i> de modelos quânticos utilizados	62
Quadro 2 – Modelos clássicos utilizados	65
Quadro 3 – <i>Ensemble</i> de modelos clássicos utilizados	65
Quadro 4 – Exemplo de frases criadas através do <i>ChatGPT</i> . É possível perceber um aumento na complexidade lexical das palavras utilizadas.	78

LISTA DE TABELAS

Tabela 1 – Número de exemplos e média de palavras por exemplo por base de dados . 146

LISTA DE ABREVIATURAS E SIGLAS

BERT	<i>Bidirectional Encoder Representations from Transformers</i>
CQP	Circuito Quântico Parametrizado
CQV	Circuito Quântico Variacional
KNN	<i>K-Nearest Neighbors</i>
MRL	<i>Matryoshka Representation Learning</i>
NISQ	<i>Noisy Intermediate-Scale Quantum</i>
PE	Preparação de Estados
PLN	Processamento de Linguagem Natural
PQLN	Processamento Quântico de Linguagem Natural
QML	<i>Quantum Machine Learning</i>
SBERT	<i>Sentence-BERT</i>
SST	<i>Stanford Sentiment Treebank</i>
SVM	<i>Support Vector Machines</i>
WSL	<i>Windows Subsystem for Linux</i>

SUMÁRIO

1	INTRODUÇÃO	29
1.1	CONTEXTUALIZAÇÃO E MOTIVAÇÃO	30
1.2	OBJETIVOS	31
1.2.1	Objetivo Geral	32
1.2.2	Objetivos Específicos	32
1.3	ESTRUTURA DA DISSERTAÇÃO	33
2	FUNDAMENTAÇÃO TEÓRICA	34
2.1	COMPUTAÇÃO QUÂNTICA	34
2.1.1	O <i>qubit</i>: superposição e a esfera de Bloch	35
2.1.2	Portas quânticas	36
2.1.3	Emaranhamento quântico	37
2.1.4	Medição	38
2.2	APRENDIZAGEM DE MÁQUINA QUÂNTICA	39
2.2.1	Circuitos quânticos Variacionais	40
2.2.2	Codificação de dados clássicos em estados quânticos	42
2.2.3	Otimização e a paisagem de custo	43
2.3	MÉTRICAS DE CAPACIDADE DE CIRCUITOS QUÂNTICOS	44
2.3.1	Expressabilidade	44
2.3.2	Emaranhamento	45
2.3.3	Arquitetura e profundidade do <i>ansatz</i>	46
2.4	PROCESSAMENTO DE LINGUAGEM NATURAL E ANÁLISE DE SENTI- MENTO	47
2.4.1	Representação vetorial de texto (<i>embeddings</i>)	47
2.4.2	Modelos <i>Transformer</i> e Sentence-BERT	48
2.4.3	<i>Matryoshka representation learning</i>	49
2.5	MODELOS DE APRENDIZAGEM DE MÁQUINA CLÁSSICOS	50
2.5.1	Modelos lineares	50
2.5.2	K-vizinhos Mais Próximos	51
2.5.3	Máquinas de Vetores de Suporte	51
2.5.4	Métodos de <i>ensemble</i>	52

2.5.4.1	<i>Bagging e Florestas Randômicas</i>	52
2.5.4.2	<i>Boosting e o AdaBoost</i>	53
2.5.5	Agregação por votação	53
2.6	AVALIAÇÃO DE MODELOS E VALIDAÇÃO ESTATÍSTICA	53
2.6.1	Métrica F1	54
2.6.2	Testes de hipótese para comparação de classificadores	55
3	MÉTODO PROPOSTO E DESENHO EXPERIMENTAL	57
3.1	MODELO HÍBRIDO QUÂNTICO-CLÁSSICO PROPOSTO	57
3.1.1	Arquitetura geral do classificador	57
3.1.2	Catálogo de <i>ansatze</i> investigados	58
3.1.3	Agregação de modelos quânticos via <i>ensemble</i>	60
3.2	DESENHO EXPERIMENTAL PARA VALIDAÇÃO	62
3.2.1	Configuração de ambiente	62
3.2.2	Base de dados	63
3.2.3	Processamento e Representação dos Dados	64
3.2.4	Modelos Clássicos para Comparação	65
3.2.5	Estratégia de Otimização e Treinamento	65
3.2.6	Análise de Propriedades Intrínsecas e Desempenho	66
3.2.7	Validação Estatística	66
4	RESULTADOS E DISCUSSÃO	67
4.1	PAPEL DO EMARANHAMENTO, PROFUNDIDADE E EXPRESSABILIDADE NA CAPACIDADE REPRESENTACIONAL	67
4.2	INFLUÊNCIA DA REPRESENTAÇÃO VETORIAL DE ENTRADA	83
4.2.1	Resiliência dos modelos classificatórios e redução de dimensionalidade	84
4.2.2	Sensibilidade do desempenho classificatória à escolha do modelo gerador de <i>embeddings</i>	102
4.3	COMPARAÇÃO ENTRE MODELOS QUÂNTICOS E CLÁSSICOS E SEUS ENSEMBLES	120
5	CONCLUSÃO	151
5.1	SÍNTESE DOS RESULTADOS E VALIDAÇÃO DOS OBJETIVOS	151
5.2	LIMITAÇÕES DA PESQUISA	153
5.3	TRABALHOS FUTUROS	154
	REFERÊNCIAS	156

1 INTRODUÇÃO

A ascensão da inteligência artificial, impulsionada por avanços exponenciais no poder computacional e pela disponibilidade de vastos volumes de dados, redefiniu as fronteiras da tecnologia e da ciência (GOODFELLOW; BENGIO; COURVILLE, 2016). No cerne dessa revolução, o campo do Processamento de Linguagem Natural (PLN) emergiu como uma das áreas de maior impacto, capacitando máquinas a interpretar, processar e gerar a linguagem humana com uma sofisticação crescente. Arquiteturas de redes neurais profundas, notadamente os modelos *Transformer* e seus derivados, estabeleceram novos paradigmas de desempenho em uma gama diversificada de tarefas, desde a tradução automática até a análise de sentimentos, tornando-se ferramentas indispensáveis em aplicações comerciais, sociais e científicas (VASWANI et al., 2023).

Paralelamente a essa evolução no domínio clássico, um novo paradigma computacional, fundamentado nos princípios da mecânica quântica, vem ganhando tração e promete redefinir os limites do que é computacionalmente tratável (ARUTE et al., 2019). A computação quântica, que explora fenômenos como a superposição e o emaranhamento, oferece o potencial para resolver classes específicas de problemas que são intratáveis para os supercomputadores mais avançados da atualidade (NIELSEN; CHUANG, 2012). A intersecção desses dois campos de vanguarda dá origem à Aprendizagem de Máquina Quântica – do inglês, *Quantum Machine Learning* (QML) –, uma área de pesquisa que investiga como os recursos quânticos podem ser alavancados para aprimorar algoritmos de aprendizagem de máquina (BIAMONTE et al., 2017).

Dentro do escopo da *Quantum Machine Learning* (QML), o Circuito Quântico Variacional (CQV) destaca-se como uma das abordagens mais promissoras (PERUZZO et al., 2014) para a era atual de computadores quânticos de escala intermediária e suscetíveis a ruído, a era *Noisy Intermediate-Scale Quantum* (NISQ) (PRESKILL, 2018). Esses algoritmos híbridos, que combinam um processador quântico com um otimizador clássico, são análogos às redes neurais clássicas, onde um circuito quântico parametrizado, ou *ansatz*, é treinado para aprender padrões nos dados. A aplicação desses modelos a tarefas de PLN, um campo emergente conhecido como Processamento Quântico de Linguagem Natural (PQLN), abre uma nova e fascinante fronteira de investigação, com o potencial de explorar os vastos espaços de Hilbert para representar e processar a informação linguística de maneiras fundamentalmente novas (CEREZO et al., 2021).

Contudo, o desenvolvimento de CQV eficazes para tarefas de PQLN ainda é um campo em aberto e repleto de desafios, como o fenômeno dos *barren plateaus* que pode dificultar o treinamento de tais modelos (MCCLEAN et al., 2018). O projeto de um *ansatz* é, em grande parte, um processo heurístico, e a comunidade científica ainda busca compreender quais são os princípios arquitetônicos que governam o desempenho de um classificador quântico. Questões sobre como a topologia do circuito, a escolha das portas quânticas e sua profundidade impactam a capacidade do modelo de aprender e generalizar a partir de dados complexos, como os *embeddings* de texto, permanecem sem respostas definitivas. É nesse contexto que a presente dissertação se insere, buscando investigar a relação entre as propriedades teóricas intrínsecas de um circuito, como sua expressabilidade e poder de emaranhamento (SIM; JOHNSON; ASPURU-GUZI, 2019), e seu desempenho empírico.

1.1 CONTEXTUALIZAÇÃO E MOTIVAÇÃO

A investigação de algoritmos quânticos para tarefas de aprendizagem de máquina não parte de uma premissa puramente teórica, mas é motivada por evidências concretas de que a computação quântica já demonstrou potencial para superar as melhores abordagens clássicas conhecidas em domínios específicos. Algoritmos seminais, como o de Shor para a fatoração de inteiros (SHOR, 1994) e o de Grover para busca em bases de dados não estruturadas (GROVER, 1996), estabeleceram-se como provas de conceito fundamentais. O fato de que problemas de relevância prática podem ser resolvidos de forma mais eficiente em um computador quântico do que com os métodos clássicos atuais serve como o principal catalisador que inspira a comunidade científica a explorar se vantagens análogas podem ser encontradas em outras áreas complexas, como o Processamento de Linguagem Natural.

Desta forma, a motivação central deste trabalho reside na necessidade de se estabelecer uma ponte entre a teoria e a prática no projeto de CQV para análise de sentimentos. Enquanto a literatura teórica sugere que propriedades como a expressabilidade – a capacidade do *ansatz* de explorar o espaço de Hilbert – e o poder de emaranhamento – sua habilidade de gerar correlações quânticas complexas – são indicadores-chave do potencial computacional de um circuito, poucas pesquisas realizaram uma validação empírica sistemática dessas métricas em um problema concreto de PLN (SIM; JOHNSON; ASPURU-GUZI, 2019). A presente pesquisa é motivada pela hipótese de que uma análise quantitativa dessas propriedades pode fornecer diretrizes valiosas para o projeto de arquiteturas de redes neurais quânticas mais robustas e

eficazes.

Adicionalmente, um aspecto fundamental e frequentemente subestimado no desenvolvimento de modelos de PQLN é a interface entre os dados clássicos e o processador quântico (SCHULD; PETRUCCIONE, 2021). A maneira como um vetor de *embedding* de texto é codificado em um estado quântico, bem como a escolha do próprio modelo gerador de *embeddings* e a dimensionalidade desse vetor, são fatores que podem influenciar de forma relevante o desempenho do classificador (HALLER et al., 2022). Este trabalho é motivado, portanto, pela necessidade de se investigar a sensibilidade dos modelos quânticos a essas escolhas, avaliando como a resiliência à redução de dimensionalidade e a escolha do modelo *Transformer* impactam a capacidade de aprendizado, especialmente quando comparados a *benchmarks* clássicos consolidados.

A pesquisa justifica-se, assim, pela sua contribuição em três frentes principais. Primeiramente, ela oferece uma análise experimental abrangente, correlacionando métricas teóricas (expressabilidade e emaranhamento) com o desempenho prático (métrica F1) de uma vasta gama de arquiteturas de circuitos. Essa métrica é escolhida tendo em vista sua robustez em cenários com desbalanceamento de classes, comuns na análise de sentimentos. Por ser a média harmônica entre precisão e revocação, ela oferece uma avaliação de desempenho mais fidedigna do que a acurácia isolada. Em segundo lugar, investiga de forma pormenorizada o impacto da representação de dados clássicos no desempenho quântico, um passo fundamental para o desenvolvimento de aplicações práticas. Por fim, ao realizar uma comparação estatisticamente validada com modelos clássicos, o trabalho busca situar o estado atual dos CQV para análise de sentimentos, identificando tanto seu potencial quanto os desafios que ainda precisam ser superados na busca por uma eventual vantagem quântica.

1.2 OBJETIVOS

A presente dissertação delinea um conjunto de metas claras e interdependentes, estruturadas em um objetivo geral que norteia a investigação e em objetivos específicos que detalham as etapas necessárias para alcançá-lo.

1.2.1 Objetivo Geral

O objetivo geral desta dissertação é realizar uma investigação sistemática e experimental sobre a relação entre as propriedades teóricas de expressabilidade e emaranhamento de diferentes arquiteturas de circuitos quânticos variacionais e seu desempenho empírico na tarefa de classificação de sentimentos, avaliando também o impacto da representação vetorial dos dados de entrada e estabelecendo uma comparação rigorosa com modelos de aprendizagem de máquina clássicos.

1.2.2 Objetivos Específicos

Para alcançar o objetivo geral supracitado, foram definidos os seguintes objetivos específicos:

1. Projetar e implementar um conjunto diversificado de 15 arquiteturas de circuitos quânticos variacionais, com diferentes topologias, tipos de portas quânticas e capacidades de emaranhamento, avaliando cada uma em configurações de baixa ($L = 1$) e alta ($L = 10$) profundidade. A avaliação foca deliberadamente nestes dois extremos para contrastar uma linha de base de recursos mínimos com o limite superior da capacidade representacional investigada. Essa análise nos extremos é a mais informativa para entender o impacto fundamental da profundidade na performance.
2. Quantificar, para cada arquitetura de circuito proposta, as métricas teóricas de expressabilidade, por meio da divergência de Kullback-Leibler, e de poder de emaranhamento, utilizando a medida de Meyer-Wallach.
3. Investigar o impacto da representação de dados de entrada no desempenho dos classificadores, comparando três modelos geradores de *embeddings* (*Matryoshka*, *Nomic* e *MPNet*) e avaliando a resiliência dos modelos à redução de dimensionalidade dos vetores (de 768 para 32 e 16 atributos).
4. Avaliar o desempenho dos modelos quânticos e clássicos em quatro bases de dados de análise de sentimentos: três bases sintéticas de complexidade crescente e a base de dados pública SST, utilizando a métrica F1 como principal indicador de desempenho.

5. Estabelecer uma linha de base robusta por meio da implementação e avaliação de um conjunto de 7 modelos de aprendizagem de máquina clássicos consolidados, além de 10 configurações de *ensembles* clássicos, para uma comparação direta com as abordagens quânticas, incluindo 16 modelos quânticos em formato *ensemble*.

1.3 ESTRUTURA DA DISSERTAÇÃO

A fim de apresentar a pesquisa de forma clara, lógica e progressiva, esta dissertação foi organizada na seguinte estrutura de capítulos. O Capítulo 2 estabelece a **Fundamentação Teórica**, introduzindo os conceitos essenciais da computação quântica, aprendizagem de máquina quântica, as métricas de capacidade de circuitos, as técnicas de processamento de linguagem natural e os métodos de validação estatística que sustentam o trabalho. O Capítulo 3 descreve a **Metodologia**, detalhando o desenho experimental, a configuração do ambiente, as bases de dados, os modelos quânticos e clássicos utilizados e os procedimentos de avaliação e comparação. O Capítulo 4 discorre sobre os **Resultados e Discussão** dos achados, interpretando os resultados à luz da teoria e correlacionando as diferentes variáveis do estudo. Por fim, o Capítulo 5 apresenta a **Conclusão**, sintetizando os principais resultados, validando as hipóteses levantadas, reconhecendo as limitações da pesquisa e apontando direções para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo tem como objetivo apresentar os conceitos teóricos que fundamentam o desenvolvimento desta dissertação. A estrutura do capítulo foi projetada para construir o conhecimento de forma progressiva, partindo dos princípios da computação quântica e da aprendizagem de máquina quântica, passando pelas métricas utilizadas para avaliar os circuitos e pelas técnicas de processamento de linguagem natural, até chegar aos modelos clássicos que servem como referência de comparação e aos métodos de validação estatística empregados.

2.1 COMPUTAÇÃO QUÂNTICA

A computação clássica, que sustenta a tecnologia digital contemporânea, baseia-se em princípios da física clássica para processar informações codificadas em bits, unidades que assumem estados discretos de 0 ou 1. Em contrapartida, a computação quântica emerge como um paradigma computacional que explora os fenômenos da mecânica quântica, como a superposição e o emaranhamento, para processar informações de maneiras fundamentalmente distintas.

Proposta inicialmente por físicos como Richard Feynman, que vislumbrou a utilização de sistemas quânticos para simular outros sistemas quânticos de forma mais eficiente (FEYNMAN, 1982), esta área promete soluções para problemas intratáveis aos computadores mais poderosos da atualidade.

Ganhos de performance teóricos já foram demonstrados em algoritmos canônicos que fornecem fortes indícios dessa vantagem. O algoritmo de Shor, por exemplo, oferece um ganho exponencial, em relação aos melhores algoritmos clássicos conhecidos, para a fatoração de números primos, um problema no cerne da segurança de dados e da criptografia moderna (SHOR, 1994). De forma análoga, o algoritmo de Grover provê um ganho quadrático frente às soluções clássicas na busca em bancos de dados não estruturados (GROVER, 1996). Tais algoritmos fornecem a evidência fundamental que motiva a investigação do potencial quântico em outras áreas, como a simulação de sistemas moleculares, a otimização de larga escala e, como foco deste trabalho, a aprendizagem de máquina (NIELSEN; CHUANG, 2012; ARUTE et al., 2019).

A seguir, serão introduzidos os conceitos elementares da computação quântica que são indispensáveis para a compreensão deste trabalho.

2.1.1 O *qubit*: superposição e a esfera de Bloch

A unidade fundamental de informação na computação quântica é o bit quântico, ou *qubit*. Diferentemente de um bit clássico, que está restrito aos estados 0 ou 1, um *qubit* pode existir em uma combinação linear, ou superposição, desses dois estados. Matematicamente, o estado de um *qubit*, denotado por $|\psi\rangle$ na notação de Dirac, pode ser descrito como um vetor em um espaço de Hilbert complexo e bidimensional, \mathbb{C}^2 , munido do produto interno canônico (NIELSEN; CHUANG, 2012). Neste formalismo, a base computacional é formada pelos estados $|0\rangle$ e $|1\rangle$, que são definidos como ortonormais em relação a este produto interno. Qualquer estado de um *qubit* pode, então, ser expresso como uma combinação linear nesta base:

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \quad (2.1)$$

no qual α e β são números complexos conhecidos como amplitudes de probabilidade. O princípio da superposição é um dos pilares que distinguem a computação quântica, permitindo que um *qubit* explore um contínuo de estados, e não apenas dois valores discretos. O quadrado dos módulos dessas amplitudes, $\|\alpha\|^2$ e $\|\beta\|^2$, representa, respectivamente, a probabilidade de se obter o resultado 0 ou 1 ao se realizar uma medição do *qubit* na base computacional, conforme a regra de Born (BORN, 1926). Uma condição fundamental é que a soma dessas probabilidades deve ser igual a 1, o que impõe a restrição de normalização ao estado do *qubit*: $\|\alpha\|^2 + \|\beta\|^2 = 1$. Esta propriedade permite que um único *qubit* codifique uma quantidade de informação potencialmente maior do que um bit clássico, uma capacidade que se expande exponencialmente com o aumento do número de *qubits* em um sistema.

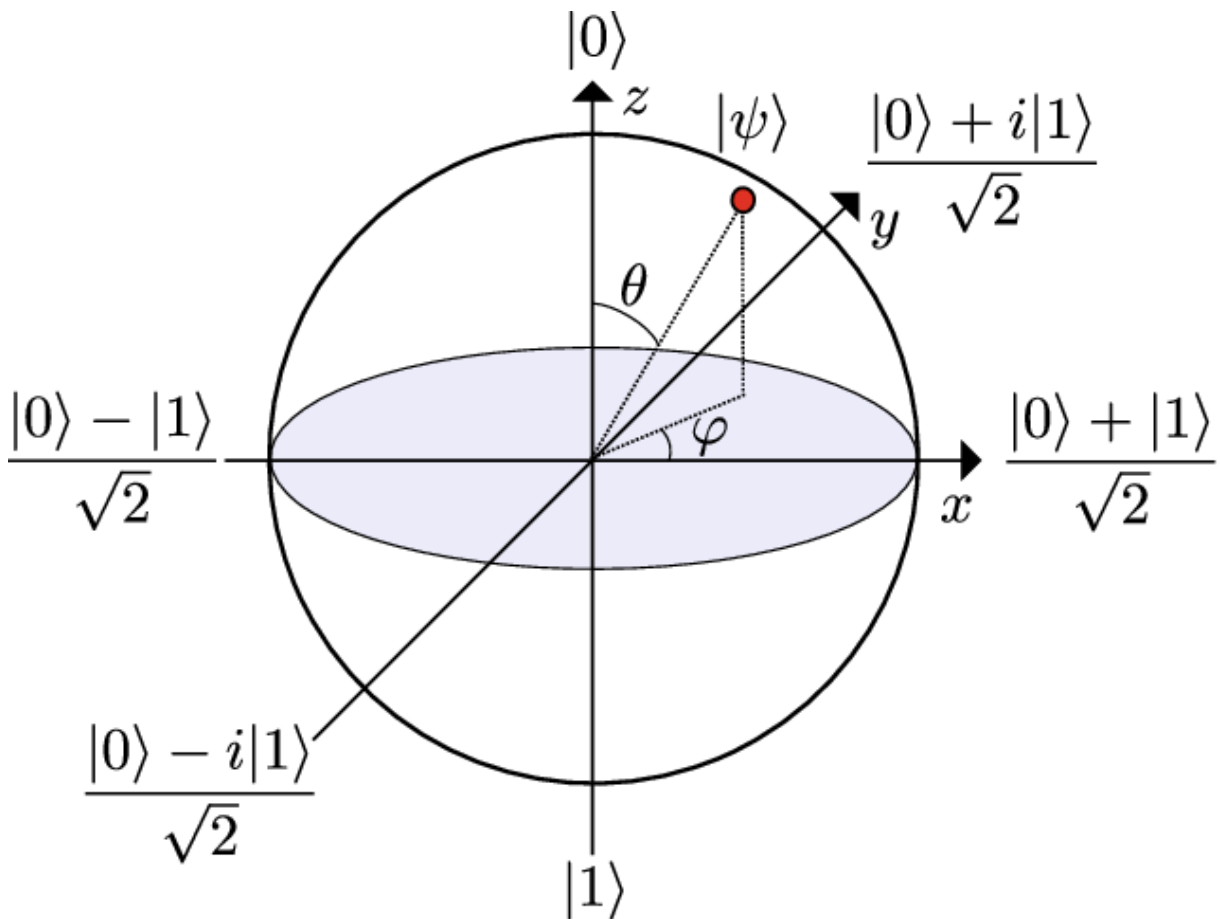
Para a visualização geométrica do estado de um único *qubit*, utiliza-se a Esfera de Bloch, uma representação onde qualquer estado puro pode ser mapeado a um ponto na superfície de uma esfera de raio unitário (NIELSEN; CHUANG, 2012), no qual uma representação gráfica pode ser vista na Figura 1. Dessa forma, o estado de um *qubit* pode ser parametrizado por dois ângulos reais, θ ($0 \leq \theta < \pi$) e ϕ ($0 \leq \phi < 2\pi$), da seguinte forma:

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |1\rangle, \quad (2.2)$$

dessa maneira, o polo norte da esfera ($\theta = 0$) corresponde ao estado $|0\rangle$, enquanto o polo sul ($\theta = \pi$) corresponde ao estado $|1\rangle$. Os demais pontos na superfície da esfera representam todos os possíveis estados de superposição do *qubit*. O ângulo ϕ representa a fase relativa entre os

estados $|0\rangle$ e $|1\rangle$, um grau de liberdade que não possui análogo direto na computação clássica e que é fundamental para os fenômenos de interferência quântica (SCHULD; PETRUCCIONE, 2018).

Figura 1 – Representação geométrica do estado de um *qubit* na Esfera de Bloch. Os polos norte e sul correspondem aos estados da base computacional $|0\rangle$ e $|1\rangle$, respectivamente, enquanto os pontos na superfície representam os estados de superposição.



Fonte: KOCKUM; NORI (2019)

2.1.2 Portas quânticas

De forma análoga às portas lógicas na computação clássica, as portas quânticas são as operações elementares que manipulam o estado dos *qubits*. Matematicamente, uma porta quântica que atua sobre n *qubits* é representada por uma matriz unitária U de dimensão $2n \times 2n$. Sua aplicação sobre um estado $|\psi\rangle$ resulta em um novo estado $|\psi'\rangle = U|\psi\rangle$. A condição de unitariedade ($U^\dagger U = U U^\dagger = I$, onde U^\dagger é a transposta conjugada de U e I é a matriz identidade) garante que a operação seja reversível e preserve a norma do vetor de estado, uma característica fundamental da evolução quântica, conforme descrito pela equação

de Schrödinger (NIELSEN; CHUANG, 2012). Foi demonstrado por Deutsch que um conjunto finito de portas quânticas, conhecido como conjunto universal, é suficiente para aproximar qualquer operação unitária com precisão arbitrária (NADA, 1992).

As portas utilizadas neste trabalho podem ser divididas em portas de um e de dois *qubits*. Dentre as portas de um *qubit*, a porta de Hadamard (H) é uma das mais importantes, sendo responsável por criar superposições uniformes a partir dos estados da base. As portas de Pauli (X , Y , Z) correspondem a rotações de π radianos em torno dos eixos x , y e z da Esfera de Bloch, respectivamente, sendo a porta X o análogo quântico da porta NOT clássica.

As portas de rotação ($R_X(\theta)$, $R_Y(\theta)$, $R_Z(\theta)$) generalizam as portas de Pauli, permitindo rotações de um ângulo arbitrário θ em torno dos respectivos eixos. Além destas, a porta de rotação genérica ($R(\theta, \phi, \lambda)$) permite uma transformação unitária arbitrária sobre um único *qubit*. A combinação de portas de rotação é de particular relevância para os circuitos quânticos variacionais, pois os seus ângulos podem ser parametrizados e otimizados classicamente (BERGHOLM et al., 2022).

Para que a computação quântica seja universal, são necessárias também portas que atuem sobre múltiplos *qubits*, pois estas são as responsáveis pela criação de emaranhamento. A porta CNOT (Controlled-NOT) é um exemplo canônico de porta de dois *qubits*: ela possui um *qubit* de controle e um *qubit* alvo, e aplica uma porta X no alvo se, e somente se, o controle estiver no estado $|1\rangle$. De forma análoga, a porta CZ (Controlled-Z) aplica uma porta Z no alvo se o controle for $|1\rangle$, introduzindo uma fase condicional no sistema. A capacidade de criar correlações condicionais por meio destas portas é o que permite a construção de algoritmos quânticos complexos e o aproveitamento do emaranhamento como recurso computacional.

2.1.3 Emaranhamento quântico

O emaranhamento é um dos fenômenos mais característicos da mecânica quântica, descrito por Schrödinger como a propriedade definidora que a distingue da física clássica (SCHRÖDINGER, 1935). Um sistema de múltiplos *qubits* é dito emaranhado quando o estado quântico do sistema como um todo não pode ser fatorado como um produto tensorial dos estados individuais de seus *qubits* constituintes (NIELSEN; CHUANG, 2012). Em outras palavras, os *qubits* perdem sua descrição individual e passam a existir em um estado correlacionado, onde o resultado da medição de um deles está intrinsecamente ligado ao resultado da medição dos outros.

Estas correlações, que Einstein, Podolsky e Rosen questionaram em seu famoso para-

doxó Einstein-Podolsky-Rosen (EPR) como uma “ação fantasmagórica à distância” (EINSTEIN; PODOLSKY; ROSEN, 1935), foram posteriormente confirmadas como uma propriedade fundamental da natureza pelos experimentos baseados nas desigualdades de Bell (BELL, 1964). Hoje, entende-se que o emaranhamento não viola a causalidade, mas representa correlações não-locais que são um recurso computacional essencial. Um exemplo canônico de um estado emaranhado de dois *qubits* é o estado de Bell $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, onde a medição de um *qubit* determina instantaneamente o estado do outro, independentemente da distância física que os separe.

No contexto da aprendizagem de máquina quântica, o emaranhamento é fundamental para a capacidade dos circuitos de aprender padrões complexos nos dados. A aplicação de portas de múltiplos *qubits*, como CNOT e CZ, é o mecanismo pelo qual o emaranhamento é gerado e manipulado dentro de um circuito. A capacidade de um *ansatz* variacional de explorar o vasto espaço de Hilbert de um sistema de múltiplos *qubits* e de criar correlações complexas que podem não estar explícitas nos dados de entrada está intrinsecamente ligada à sua capacidade de gerar emaranhamento. Esta propriedade é uma das hipóteses centrais para explicar as diferenças de desempenho entre as arquiteturas de circuitos quânticos analisadas neste trabalho.

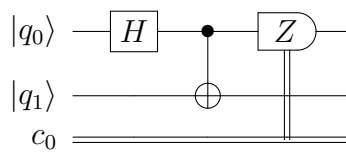
2.1.4 Medição

O processo de medição constitui a interface entre o domínio quântico, onde a informação é processada, e o mundo clássico, onde os resultados são observados. De acordo com os postulados da mecânica quântica, a medição é um processo inerentemente probabilístico e irreversível que projeta o estado quântico sobre um conjunto de estados de base ortogonais (NIELSEN; CHUANG, 2012). Quando um *qubit* em um estado de superposição $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ é medido na base computacional ($|0\rangle, |1\rangle$), seu estado quântico colapsa para um dos estados base.

Este colapso da função de onda ocorre com probabilidades definidas pela Regra de Born, ao estabelecer que a probabilidade de se obter o resultado clássico 0 ser de $P(0) = \|\alpha\|^2$, e a probabilidade de se obter o resultado 1 é $P(1) = \|\beta\|^2$ (BORN, 1926). Uma vez que a medição é realizada, a informação sobre as amplitudes α e β é perdida, e o *qubit* permanece no estado clássico para o qual colapsou. Este processo é fundamental para a extração de resultados de um algoritmo quântico e representa a fonte de aleatoriedade intrínseca à computação quântica.

No contexto dos circuitos quânticos variacionais, é comum não apenas medir os estados finais, mas calcular o valor esperado de um observável. Um observável é uma propriedade física do sistema representada por um operador Hermitiano, como os operadores de Pauli (X , Y , Z). O valor esperado de um observável \hat{O} para um sistema no estado $|\psi\rangle$ é dado por $\hat{O} = \langle\psi|\hat{O}|\psi\rangle$. Este valor esperado é um número real clássico, que pode ser estimado por meio da amostragem repetida do circuito (preparando o estado e medindo múltiplas vezes) e do cálculo da média dos resultados. Na aprendizagem de máquina quântica, este valor é frequentemente utilizado como a saída do modelo, que é então processada por uma função de custo clássica para guiar o processo de otimização dos parâmetros do circuito. Para consolidar os conceitos de *qubit*, portas quânticas e medição, a Figura 2 ilustra um circuito elementar que integra essas operações para gerar e medir um estado emaranhado.

Figura 2 – Exemplo de um circuito quântico elementar que gera um estado de Bell, um dos estados maximamente emaranhados de dois *qubits*. O circuito opera sobre dois *qubits*, $|q_0\rangle$ e $|q_1\rangle$, e um bit clássico, c_0 . Uma porta de Hadamard (H) é aplicada ao primeiro *qubit* para criar uma superposição, seguida por uma porta CNOT, que emaranha os dois *qubits*. Por fim, uma medição na base do observável Z é realizada no primeiro *qubit*, e o resultado clássico é armazenado no bit c_0 .



Fonte: elaborada pelo autor (2025)

2.2 APRENDIZAGEM DE MÁQUINA QUÂNTICA

A intersecção entre a computação quântica e a aprendizagem de máquina dá origem ao campo do QML. Este campo investiga como os princípios quânticos podem ser utilizados para aprimorar ou desenvolver novos algoritmos de aprendizagem, bem como a aplicação de técnicas de aprendizagem de máquina clássica para controlar e analisar sistemas quânticos (BIAMONTE et al., 2017). Na era atual de computadores quânticos com um número intermediário de *qubits* e suscetíveis a ruído (a era NISQ), uma das abordagens mais promissoras e ativamente pesquisadas é a que se baseia em algoritmos quânticos variacionais (CEREZO et al., 2021).

2.2.1 Circuitos quânticos Variacionais

Os CQV, também conhecidos como algoritmos quânticos híbridos, representam uma classe de algoritmos projetados para alavancar o poder computacional dos processadores quânticos atuais em conjunto com os otimizadores clássicos (PERUZZO et al., 2014). A arquitetura de um CQV é inerentemente híbrida e opera em um ciclo de otimização iterativo, conforme ilustrado a seguir:

1. Preparação de parâmetros: um computador clássico inicializa um conjunto de parâmetros θ .
2. Execução quântica: estes parâmetros são enviados a um processador quântico para configurar um Circuito Quântico Parametrizado (CQP), também conhecido como *ansatz*. O circuito é executado, geralmente com um estado de entrada que codifica os dados do problema. Também é possível simular esta execução quântica utilizando computadores clássicos.
3. Medição: o estado de saída do circuito quântico é medido, gerando resultados clássicos (bits). Este processo é repetido múltiplas vezes para se obter uma estimativa estatística do valor esperado de um ou mais observáveis.
4. Cálculo da função de custo: o computador clássico utiliza os resultados das medições para calcular o valor de uma função de custo, $C(\theta)$, que quantifica o quão bem o circuito está resolvendo a tarefa desejada.
5. Otimização clássica: um algoritmo de otimização clássico utiliza o valor da função de custo para propor um novo conjunto de parâmetros θ' com o objetivo de minimizar $C(\theta)$.
6. Iteração: os passos 2 a 5 são repetidos até que um critério de convergência seja atingido.

O coração da parte quântica deste ciclo é o *ansatz*, ou CQP. Trata-se de um circuito quântico de topologia fixa, composto por portas fixas (como CNOT e Hadamard) e portas parametrizadas (como as de rotação $R_X(\theta_i)$, $R_Y(\theta_i)$ e $R_Z(\theta_i)$). O conjunto de parâmetros $\theta = \{\theta_1, \theta_2, \dots, \theta_M\}$ constitui os graus de liberdade do modelo, análogos aos pesos e vieses de uma rede neural clássica (BENEDETTI et al., 2019). A capacidade de um CQV de aprender um determinado padrão de dados está diretamente ligada à capacidade representacional de seu *ansatz*, que, por sua vez, é determinada pela escolha das portas, sua conectividade e sua

profundidade. A concepção de *ansatze* eficazes é um dos principais focos de pesquisa na área, sendo o objeto central da investigação experimental desta dissertação.

A notável flexibilidade dos CQV reside em sua capacidade de atuar como aproximadores de funções, de forma análoga às redes neurais clássicas. Foi demonstrado que os CQP podem ser expressos como uma forma de Série de Fourier generalizada, o que lhes confere o potencial para modelar funções complexas e não-lineares (SCHULD; PETRUCCIONE, 2021). De fato, investigações sobre o poder de expressividade desses modelos sugerem que, para um mesmo número de parâmetros, as redes neurais quânticas podem alcançar uma "dimensão efetiva" maior do que suas contrapartes clássicas, indicando uma capacidade representacional potencialmente superior (ABBAS et al., 2021).

Uma das mais promissoras vias para se obter uma vantagem quântica com estes modelos é através da perspectiva dos *kernels* quânticos. Proposto em trabalhos seminais como o de Havlíček et al. (2019), este método utiliza um circuito quântico para mapear os dados de entrada clássicos para um espaço de características quântico de alta dimensionalidade. Se este mapeamento for difícil de ser simulado classicamente, um classificador linear nesse novo espaço pode superar os melhores classificadores clássicos. A natureza e as condições exatas para que essa vantagem quântica se manifeste, contudo, são um tópico de intensa pesquisa e debate na comunidade (SCHULD, 2021).

Apesar de seu potencial teórico, o sucesso prático de um CQV depende crucialmente do projeto de sua arquitetura e de sua treinabilidade. Pesquisas recentes indicam que mesmo *ansatze* de profundidade rasa podem ser poderosos, contanto que possuam capacidade de emaranhamento suficiente (GILL et al., 2023), um achado que ressoa diretamente com os objetivos desta dissertação. No entanto, a treinabilidade desses circuitos é desafiada pelo fenômeno dos *barren plateaus*. Foi demonstrado que a topografia da paisagem de custo, incluindo a presença de platôs, é fortemente dependente não apenas da arquitetura, mas também da escolha da função de custo, com custos "locais" oferecendo um caminho para mitigar o problema (CEREZO et al., 2020).

Essa interação entre expressividade, capacidade de emaranhamento e treinabilidade motivou o desenvolvimento de diversas arquiteturas e metodologias. No contexto de classificadores, propostas pioneiras como a de Farhi e Neven (2018) estabeleceram as bases para a aplicação de CQV em problemas de aprendizagem supervisionada. O avanço da área tem inspirado também trabalhos que exploram, por exemplo, o uso de otimização evolucionária para os parâmetros de classificadores quânticos variacionais (COSTA; NETO, 2025), o desenvolvimento de neurônios

quânticos parametrizados de profundidade constante (CARVALHO; NETO, 2023), a recomendação de arquiteturas de circuitos com base em medidas de complexidade dos dados (NETO, 2025) e a aplicação de classificadores baseados em sistemas quânticos abertos (BRITO; NETO; BERNARDES, 2024). Tais avanços reforçam a posição dos CQV como uma das mais ativas e promissoras frentes da QML.

2.2.2 Codificação de dados clássicos em estados quânticos

Um pré-requisito fundamental para a aplicação de algoritmos quânticos a problemas de aprendizagem de máquina é a capacidade de codificar dados clássicos em estados quânticos. Este processo, conhecido como preparação de estado ou mapeamento de características quânticas, consiste em mapear um vetor de características clássico $x \in \mathbb{R}^n$ para um estado quântico $|\psi(x)\rangle$ em um espaço de Hilbert (SCHULD; PETRUCCIONE, 2021). A escolha da estratégia de codificação é de relevância central, pois ela define como a informação do problema é apresentada ao processador quântico, influenciando diretamente a geometria do espaço de estados e, consequentemente, a capacidade do modelo de aprender fronteiras de decisão eficazes.

Existem diversas estratégias de codificação, cada uma com suas próprias vantagens e desvantagens em termos de recursos (número de *qubits* e profundidade do circuito) e capacidade expressiva. Exemplos incluem o *basis embedding*, que associa cada vetor de entrada a um estado da base computacional, e o *angle embedding*, que codifica as características nos ângulos de rotação de portas quânticas.

Neste trabalho, adota-se a estratégia de *amplitude embedding*. Este método é particularmente eficiente em termos de número de *qubits*, pois permite codificar um vetor de características x com N elementos no vetor de amplitudes de um estado quântico de apenas $n = \lceil \log_2(N) \rceil$ *qubits*. O estado resultante é da forma:

$$|\psi(x)\rangle = \frac{1}{\|x\|} \sum_{i=0}^{N-1} x_i |i\rangle \quad (2.3)$$

onde $\|x\|$ é a norma euclidiana do vetor, que garante a normalização do estado quântico, e $|i\rangle$ representa os estados da base computacional. Embora o *Amplitude Embedding* seja exponencialmente eficiente no uso de *qubits*, a preparação de um estado arbitrário com esta técnica pode, em geral, exigir um circuito de profundidade exponencial (NIELSEN; CHUANG, 2012). Contudo, existem métodos eficientes para a construção destes circuitos. A implemen-

tação utilizada nesta dissertação se baseia no trabalho de Mottonen et al. (2004), que propõe um método sistemático para decompor a preparação de um estado quântico arbitrário em uma sequência de portas de rotação uniformemente controladas, otimizando a construção do circuito de codificação.

2.2.3 Otimização e a paisagem de custo

O treinamento de um CQV é, em sua essência, um problema de otimização: encontrar o conjunto de parâmetros θ^* que minimiza uma função de custo $C(\theta)$ definida classicamente. A função de custo é tipicamente avaliada a partir das medições do circuito quântico, o que introduz um ruído de amostragem inerente ao processo. Este ruído, combinado com a complexidade da própria arquitetura quântica, pode tornar a otimização uma tarefa desafiadora (CEREZO et al., 2021).

A caracterização da função de custo em relação aos parâmetros do circuito define a paisagem de custo. O sucesso de um algoritmo de otimização, especialmente os baseados em gradiente, como o Adam, depende da topografia desta paisagem. Um dos maiores desafios identificados no treinamento de CQV é o fenômeno dos *barren plateaus*. Formalizado por McClean et al. (2018), este fenômeno descreve como, em vastas regiões da paisagem de custo, a magnitude esperada dos gradientes diminui exponencialmente à medida que mais *qubits* são adicionados ao circuito. Na prática, isso significa que para um sistema com um número ainda modesto de *qubits*, os gradientes já são tão pequenos que se tornam indistinguíveis do ruído estatístico, impedindo que um otimizador como o Adam encontre uma direção para aprimorar os parâmetros do modelo.

A presença de um *barren plateau* implica que a paisagem de custo é extremamente plana em quase toda a sua extensão. Para um otimizador baseado em gradiente, a ausência de uma “descida” clara torna a busca por um mínimo local uma tarefa intratável, pois as atualizações dos parâmetros se tornam aleatórias e ineficazes. Foi demonstrado que a ocorrência de *barren plateaus* está correlacionada com diversos fatores, incluindo a profundidade do *ansatz* (circuitos muito profundos tendem a apresentar o problema), a natureza global da função de custo e o nível de emaranhamento gerado aleatoriamente pelo circuito. Compreender e mitigar este fenômeno é um dos campos de pesquisa mais ativos em QML, sendo crucial para o desenvolvimento de algoritmos quânticos variacionais escaláveis e treináveis.

2.3 MÉTRICAS DE CAPACIDADE DE CIRCUITOS QUÂNTICOS

O desempenho de um CQV em uma tarefa de aprendizagem de máquina não depende apenas do algoritmo de otimização clássico, mas é intrinsecamente governado pelo projeto de seu *ansatz*. A topologia do circuito, a escolha das portas quânticas e sua profundidade determinam as propriedades fundamentais do modelo. A fim de caracterizar e comparar diferentes *ansatze* de uma maneira que transcenda o seu desempenho em uma única tarefa, foram propostas métricas quantitativas para avaliar sua capacidade computacional intrínseca. Dentre estas, a expressabilidade e o poder de emaranhamento se destacam como indicadores-chave do potencial de um circuito. Estas métricas, sistematicamente estudadas por Sim, Johnson e Aspuru-Guzik (2019), oferecem uma metodologia para analisar e projetar arquiteturas de CQV de forma mais fundamentada, uma abordagem que tem sido explorada em trabalhos recentes, como o de Maouaki et al. (2025), para o desenvolvimento de modelos quânticos robustos.

2.3.1 Expressabilidade

A expressabilidade de um CQP quantifica a sua capacidade de gerar um conjunto de estados que seja representativo do espaço de Hilbert. Em outras palavras, um circuito com alta expressabilidade é capaz de explorar uma porção maior e mais diversificada do espaço de todos os possíveis estados quânticos (SIM; JOHNSON; ASPURU-GUZIK, 2019). Esta propriedade é desejável, pois, em teoria, um modelo mais expressivo tem o potencial de aproximar uma classe mais ampla de funções e, conseqüentemente, de aprender fronteiras de decisão mais complexas.

Para quantificar a expressabilidade, (SIM; JOHNSON; ASPURU-GUZIK, 2019) propuseram uma metodologia que compara a distribuição de estados gerados pelo CQP com a distribuição de estados uniformemente aleatórios, conhecida como distribuição de Haar. O procedimento consiste em:

1. Gerar um conjunto de estados $|\psi_i\rangle$ aplicando o CQP $U(i)$ com parâmetros i amostrados aleatoriamente a partir de uma distribuição uniforme. O estado inicial é tipicamente o estado $|0\rangle^{\otimes n}$.
2. Gerar um segundo conjunto de estados $|\psi_j\rangle$ amostrados diretamente da distribuição de Haar, que representa o conjunto dos estados quânticos de forma também uniforme.

3. Calcular a fidelidade, $F = \|\langle \psi_i | \psi_j \rangle\|^2$, entre pares de estados, um de cada conjunto, para construir uma distribuição de probabilidades de fidelidade, $P_{CQP}(F)$.
4. Comparar esta distribuição com a distribuição teórica de fidelidade para estados aleatórios de Haar, que para um sistema de n *qubits* é dada por $P_{Haar}(F) = (2n+1)(1F)^{2n+2}$.

A métrica de expressabilidade, denotada por Expr , é então definida como a divergência de Kullback-Leibler (KL) entre a distribuição gerada pelo CQP e a distribuição de Haar:

$$\text{Expr} = D_{KL}(P_{CQP} \| P_{Haar}) = \int P_{CQP}(F) \log \left(\frac{P_{CQP}(F)}{P_{Haar}(F)} \right) dF. \quad (2.4)$$

Um valor de D_{KL} próximo de zero indica que a distribuição de estados gerada pelo circuito é muito similar à distribuição uniforme de Haar, significando que o circuito possui alta expressabilidade. Por outro lado, um valor elevado de D_{KL} sugere que o circuito só consegue gerar estados em uma sub-região restrita do espaço de Hilbert, possuindo baixa expressabilidade. É relevante notar, contudo, que uma expressabilidade excessivamente alta nem sempre é benéfica, pois pode estar associada a paisagens de custo mais complexas e ao fenômeno dos *barren plateaus*, tornando a otimização do circuito mais desafiadora (HOLMES et al., 2022).

2.3.2 Emaranhamento

O poder de emaranhamento, ou capacidade de emaranhamento, de um circuito quantifica a sua habilidade de gerar estados emaranhados a partir de estados iniciais que não são emaranhados (estados produto). Uma vez que o emaranhamento é um recurso quântico fundamental e uma propriedade que não possui análogo clássico, a capacidade de um circuito de gerá-lo é considerada um indicador crucial de seu potencial para alcançar uma vantagem computacional (SIM; JOHNSON; ASPURU-GUZIK, 2019).

Para quantificar esta propriedade, Sim, Johnson e Aspuru-Guzik (2019) utilizam a medida de emaranhamento de Meyer-Wallach, que oferece uma forma computacionalmente tratável de avaliar o grau de emaranhamento de um estado puro de múltiplos *qubits*. Para um estado puro $|\psi\rangle$ de um sistema de n *qubits*, a medida de Meyer-Wallach, $Q(|\psi\rangle)$, é definida a partir da pureza média dos subsistemas de um único *qubit*:

$$Q(|\psi\rangle) = 2 \left(1 - \frac{1}{n} \sum_{k=1}^n \text{Tr}(\rho_k^2) \right), \quad (2.5)$$

onde $\rho_k = \text{Tr}_{\neq k}(|\psi\rangle\langle\psi|)$ é a matriz de densidade reduzida do k -ésimo *qubit*, obtida ao se traçar sobre todos os outros *qubits* do sistema. O valor de Q varia de 0, para um estado completamente separável (produto), a 1, para um estado que é, em média, maximamente emaranhado, como os estados de Grafo (HEIN; EISERT; BRIEGEL, 2004).

A capacidade de emaranhamento do CQP, denotada por Ent , é então calculada como o valor médio da medida de Meyer-Wallach sobre o conjunto de estados de saída, gerados a partir da aplicação do circuito $U(\theta)$ sobre o estado inicial não emaranhado $|0\rangle^{\otimes n}$, para diferentes conjuntos de parâmetros θ amostrados aleatoriamente. Matematicamente, a capacidade de emaranhamento é a média sobre o espaço de parâmetros:

$$\text{Ent} = \int Q(U(\theta)|0\rangle^{\otimes n}) p(\theta) d\theta, \quad (2.6)$$

no qual $p(\theta)$ é a distribuição de probabilidade dos parâmetros. Um valor de Ent próximo de 1 indica que o circuito é um emaranhador poderoso, enquanto um valor próximo de 0 indica que ele possui pouca ou nenhuma capacidade de gerar correlações quânticas. Como destacado por Maouaki et al. (2025), a análise conjunta da expressabilidade e do poder de emaranhamento oferece um panorama mais completo da capacidade de um circuito, sendo uma ferramenta valiosa para o projeto de arquiteturas de redes neurais quânticas mais robustas e eficazes.

2.3.3 Arquitetura e profundidade do *ansatz*

O projeto do *ansatz* é um fator crítico que determina a capacidade de um CQV. Duas de suas características mais importantes são a profundidade e a topologia. A profundidade do circuito, aqui denotada por L , refere-se ao número de vezes que um bloco fundamental de portas é repetido sequencialmente. Um circuito com $L = 1$ é considerado “raso”, enquanto um com L maior (por exemplo, $L = 10$) é considerado “profundo”. Em geral, o aumento da profundidade eleva a capacidade expressiva do circuito, permitindo-lhe, em teoria, aproximar funções mais complexas. Contudo, esta vantagem vem com um compromisso: circuitos mais profundos são mais suscetíveis aos efeitos do ruído em hardware quântico real e, como será discutido, podem ser mais propensos ao fenômeno dos *barren plateaus*, dificultando o seu treinamento (CEREZO et al., 2021). A análise comparativa entre configurações de circuitos rasos ($L = 1$) e profundos ($L = 10$) é, portanto, uma investigação central desta dissertação.

A topologia do *ansatz* diz respeito ao padrão de conectividade estabelecido pelas portas

de múltiplos *qubits* (emaranhadoras). Esta arquitetura define quais *qubits* interagem entre si, influenciando diretamente a forma como as correlações são criadas e propagadas pelo circuito. Diferentes topologias podem ser mais adequadas para diferentes estruturas de problemas. As estratégias de conectividade investigadas neste trabalho incluem a de vizinhos próximos (linear), onde cada *qubit* interage apenas com seus adjacentes; a de anel, onde o *qubit* controle pode estar não-adjacente ao *qubit* alvo; e a de todos-para-todos, onde cada *qubit* interage com todos os outros. A escolha da topologia é um fator determinante para o poder de emaranhamento do circuito.

2.4 PROCESSAMENTO DE LINGUAGEM NATURAL E ANÁLISE DE SENTIMENTO

O PLN é um campo da inteligência artificial e da ciência da computação que se dedica a capacitar as máquinas a compreender, interpretar e gerar a linguagem humana de forma útil e significativa (JURAFSKY; MARTIN, 2008). Dentre as inúmeras tarefas abrangidas pelo PLN, a análise de sentimentos se destaca como uma das mais relevantes para aplicações comerciais e sociais. Esta tarefa consiste na identificação e extração de informações subjetivas em textos, como opiniões e emoções, classificando-as tipicamente em categorias de polaridade: positiva, negativa ou neutra (LIU, 2012). Para que um algoritmo de aprendizagem de máquina, seja ele clássico ou quântico, possa realizar tal tarefa, é indispensável que o texto, uma forma de dado não estruturado, seja primeiramente convertido em uma representação numérica estruturada que possa ser processada computacionalmente.

2.4.1 Representação vetorial de texto (*embeddings*)

Os modelos de aprendizagem de máquina não operam diretamente sobre o texto em sua forma bruta. Portanto, o passo inicial e fundamental em qualquer pipeline de PLN moderno é a transformação do texto em vetores de números reais, um processo conhecido como representação vetorial de texto ou, mais comumente, *embedding* de texto. As abordagens iniciais para esta tarefa, como *Bag-of-Words* ou TF-IDF, representavam documentos como vetores esparsos e de alta dimensionalidade, onde cada dimensão correspondia a uma palavra do vocabulário. Embora úteis, essas representações falhavam em capturar as relações semânticas e sintáticas entre as palavras (JURAFSKY; MARTIN, 2008).

A revolução no PLN veio com a introdução de *embeddings* de palavras densos, populariza-

dos por modelos como o Word2Vec (MIKOLOV et al., 2013). Estes modelos aprendem a mapear cada palavra para um vetor denso de baixa dimensionalidade, com a propriedade fundamental de que palavras com significados contextuais similares são representadas por vetores próximos no espaço vetorial. Esta noção de que a semântica pode ser capturada pela geometria do espaço de *embeddings* permitiu que os modelos de aprendizagem de máquina generalizassem o conhecimento a partir das relações entre as palavras, em vez de tratá-las como símbolos isolados. Esta mesma lógica se estende da representação de palavras para a de sentenças e documentos inteiros, que é o foco deste trabalho.

2.4.2 Modelos *Transformer* e Sentence-BERT

A arquitetura *Transformer*, introduzida por Vaswani et al. (2023), representou um marco no campo do PLN, superando as limitações das arquiteturas de redes neurais recorrentes (RNNs) e convolucionais (CNNs) que dominavam a área até então. O componente central e inovador do *Transformer* é o mecanismo de autoatenção, que permite ao modelo ponderar a importância de todas as palavras em uma sequência de entrada ao processar cada palavra individualmente. Isso capacita o modelo a capturar dependências de longo alcance e relações contextuais complexas de forma mais eficaz e paralelizável do que as abordagens sequenciais anteriores. Modelos pré-treinados de larga escala baseados na arquitetura *Transformer*, como o *Bidirectional Encoder Representations from Transformers* (BERT), estabeleceram novos estados da arte em uma vasta gama de tarefas de PLN.

Contudo, modelos como o BERT foram primariamente projetados para produzir *embeddings* de *tokens* (palavras ou subpalavras) que são sensíveis ao contexto. Derivar uma representação vetorial de alta qualidade para uma sentença inteira a partir destes *embeddings* de *token* não é uma tarefa trivial; estratégias simples, como a média dos vetores dos *tokens*, frequentemente produzem resultados subótimos para tarefas que exigem uma compreensão semântica fina, como a comparação de similaridade entre sentenças. Para solucionar esta questão, Reimers e Gurevych (2019a) propuseram o *Sentence-BERT* (SBERT), uma modificação da arquitetura BERT. O SBERT utiliza uma rede siamesa para refinar um modelo *Transformer* pré-treinado em tarefas de comparação de pares de sentenças. Este processo otimiza o modelo para gerar *embeddings* de sentenças em um espaço vetorial onde sentenças semanticamente similares possuam uma alta similaridade de cosseno, tornando-as adequadas para tarefas de classificação, agrupamento e busca semântica. O modelo *all-mpnet-base-v2*,

utilizado nesta dissertação, é um exemplo de SBERT que se baseia na robusta arquitetura *MPNet*, uma técnica de pré-treinamento que combina as vantagens do BERT e do *XLNet* para uma compreensão de linguagem mais aprofundada (SONG et al., 2020).

2.4.3 *Matryoshka representation learning*

Apesar do sucesso dos grandes modelos de linguagem na geração de *embeddings* de alta qualidade, a sua dimensionalidade elevada (tipicamente 768, 1024 ou mais dimensões) pode impor desafios práticos em termos de armazenamento, custo computacional e latência em aplicações do mundo real. Para endereçar esta questão, Kusupati et al. (2024) introduziram o conceito de *Matryoshka Representation Learning* (MRL). A técnica se inspira nas bonecas-russas, onde uma boneca menor se aninha dentro de uma maior. De forma análoga, o MRL treina um único vetor de *embedding* de alta dimensionalidade de tal forma que seus prefixos (as primeiras 32, 64, 128, ... dimensões) também funcionem como representações de menor dimensão e de alta qualidade para o mesmo dado.

Este aninhamento de representações é alcançado por meio de uma modificação na função de custo durante o treinamento do modelo. Em vez de otimizar o desempenho apenas para a dimensionalidade final, a função de custo do MRL é uma soma ponderada das perdas calculadas para um conjunto de dimensionalidades pré-definidas. Isso força o modelo a aprender a concentrar a informação mais relevante nas dimensões iniciais do vetor, enquanto as dimensões subsequentes adicionam refinamentos progressivos.

A principal vantagem do MRL é a flexibilidade e a eficiência adaptativa que ele oferece. Com um único modelo treinado, é possível extrair *embeddings* de diferentes tamanhos, permitindo que o usuário escolha a dimensionalidade que oferece o melhor equilíbrio entre desempenho e custo computacional para uma determinada aplicação, sem a necessidade de treinar múltiplos modelos. O modelo *Nomic Embed Text*, da *Nomic AI*, bem como o modelo *tomaarsen/mpnet-base-nli-matryoshka*, são exemplos práticos de arquiteturas que implementam os princípios do MRL (NUSSBAUM et al., 2024). A capacidade destes modelos de fornecer vetores de dimensionalidade variável é uma característica central explorada nos experimentos desta dissertação, especialmente na análise do impacto da redução da dimensionalidade do vetor de entrada no desempenho dos classificadores.

2.5 MODELOS DE APRENDIZAGEM DE MÁQUINA CLÁSSICOS

Para avaliar de forma criteriosa o desempenho e o potencial dos , é indispensável estabelecer uma linha de base sólida a partir de modelos de aprendizagem de máquina clássicos consolidados. A comparação com estes modelos, que representam o estado da prática em diversas tarefas de classificação, permite contextualizar os resultados obtidos e identificar cenários onde as abordagens quânticas podem, eventualmente, oferecer uma vantagem. Esta seção descreve brevemente os algoritmos clássicos utilizados como *benchmarks* nesta dissertação, abrangendo modelos lineares, métodos baseados em instâncias, máquinas de vetores de suporte e técnicas de ensemble.

2.5.1 Modelos lineares

Modelos lineares constituem uma das classes mais fundamentais de algoritmos em aprendizagem de máquina. Sua premissa central é a de que a relação entre as características de entrada e a saída pode ser modelada por meio de uma combinação linear. Para tarefas de classificação, isso se traduz na busca por uma fronteira de decisão linear (um hiperplano) que separe as classes no espaço de características (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

A Regressão Logística, apesar de seu nome, é um modelo de classificação binária. Ela utiliza a função sigmoide para mapear a saída de uma função linear dos dados de entrada para um valor entre 0 e 1, que é interpretado como a probabilidade de a amostra pertencer à classe positiva. Uma vez que as probabilidades são obtidas, um limiar de decisão (tipicamente 0,5) é utilizado para atribuir a classe final. Sua simplicidade, interpretabilidade e eficiência computacional a tornam um dos modelos de referência mais utilizados (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

O *Perceptron*, proposto por Rosenblatt (1958), é um dos algoritmos de aprendizagem supervisionada mais antigos e representa a forma mais simples de uma rede neural artificial. Ele calcula uma soma ponderada das características de entrada e aplica uma função de ativação do tipo degrau para produzir uma saída binária. O algoritmo de aprendizado do *Perceptron* é iterativo e atualiza os pesos apenas quando uma amostra é classificada incorretamente. Sua principal limitação teórica é a garantia de convergência apenas para casos em que os dados são linearmente separáveis.

2.5.2 K-vizinhos Mais Próximos

O algoritmo dos K-vizinhos Mais Próximos - do inglês, *K-Nearest Neighbors* (KNN) - é um método de aprendizagem não-paramétrico e baseado em instâncias. Diferentemente de modelos como a Regressão Logística, que aprendem uma função de mapeamento explícita, o KNN é considerado um algoritmo de “aprendizagem preguiçosa”, pois não constrói um modelo geral durante a fase de treinamento, mas simplesmente armazena todas as amostras do conjunto de dados (ALTMAN, 1992).

O processo de classificação para uma nova amostra é direto: o algoritmo identifica os k pontos de dados do conjunto de treinamento que estão mais próximos da nova amostra no espaço de características, utilizando uma métrica de distância (comumente a distância euclidiana). A classe da nova amostra é então determinada pela votação majoritária entre as classes de seus k vizinhos. A simplicidade de seu princípio, aliada à sua capacidade de formar fronteiras de decisão altamente não-lineares e locais, o torna um *benchmark* eficaz, embora seu custo computacional na fase de predição possa ser elevado para grandes conjuntos de dados.

2.5.3 Máquinas de Vetores de Suporte

As Máquinas de Vetores de Suporte - do inglês, *Support Vector Machines* (SVM) - são uma classe de algoritmos de aprendizagem supervisionada particularmente poderosos para tarefas de classificação. Propostos por Cortes e Vapnik (1995), o princípio fundamental de um SVM, no caso linearmente separável, é encontrar o hiperplano que não apenas separa as classes, mas que o faz com a maior margem de separação possível. Esta margem é a distância entre o hiperplano e os pontos de dados mais próximos de cada classe, conhecidos como vetores de suporte. A maximização da margem confere ao modelo uma boa capacidade de generalização.

Para lidar com dados que não são linearmente separáveis, os SVM utilizam a “artimanha do *kernel*”. Esta técnica permite que o algoritmo opere em um espaço de características de alta dimensionalidade, onde os dados podem se tornar linearmente separáveis, sem a necessidade de calcular explicitamente as coordenadas dos dados neste novo espaço. Isso é feito por meio de uma função de *kernel*, $K(x_i, x_j)$, que calcula o produto escalar entre as amostras em um espaço transformado. Os *kernels* utilizados neste trabalho incluem:

- Linear: o caso mais simples, que resulta em uma fronteira de decisão linear no espaço original.
- Polinomial: permite a criação de fronteiras de decisão polinomiais, cuja complexidade é controlada pelo grau do polinômio.
- Rede de base basal: Um *kernel* popular e flexível, que mapeia as amostras para um espaço de dimensão infinita, sendo capaz de criar fronteiras de decisão não-lineares e complexas.

2.5.4 Métodos de *ensemble*

Métodos de *ensemble* são técnicas que combinam as previsões de múltiplos modelos de aprendizagem (estimadores base) para produzir um classificador final mais robusto e preciso do que qualquer um de seus componentes individuais. A premissa central é que, ao agregar as “opiniões” de diversos modelos, os erros de um estimador podem ser compensados pelos acertos de outros.

2.5.4.1 *Bagging e Florestas Randômicas*

O *Bootstrap Aggregating* (ou *Bagging*), é uma técnica de *ensemble* projetada principalmente para reduzir a variância de um estimador (BREIMAN, 1996). O processo consiste em criar múltiplos subconjuntos de dados a partir do conjunto de treinamento original por meio de amostragem com reposição (*bootstrap*). Um estimador base é treinado de forma independente em cada um desses subconjuntos. A previsão final é obtida pela agregação das previsões de todos os estimadores, geralmente por votação majoritária (para classificação) ou pela média (para regressão).

As Florestas Randômicas são uma aplicação específica e poderosa do *Bagging*, utilizando árvores de decisão como estimadores base (BREIMAN, 2001). Além da amostragem de dados do *Bagging*, as Florestas Randômicas introduzem uma segunda camada de aleatoriedade: ao construir cada árvore, em cada nó de decisão, apenas um subconjunto aleatório de características é considerado para determinar a melhor divisão. Este processo descorrelaciona as árvores do *ensemble*, reduzindo ainda mais a variância e tornando o modelo final robusto ao sobreajuste.

2.5.4.2 Boosting e o AdaBoost

Diferentemente do *Bagging*, onde os modelos são treinados em paralelo, o *Boosting* é uma técnica de ensemble sequencial. Os estimadores base são construídos um após o outro, e cada novo modelo é treinado para corrigir os erros cometidos pelos seus predecessores. O objetivo principal do *Boosting* é reduzir o viés do modelo final.

O *Adaptive Boosting* (ou *AdaBoost*), proposto por Freund e Schapire (1997), é um dos algoritmos de *Boosting* mais conhecidos. Ele funciona ajustando iterativamente os pesos das amostras de treinamento. Inicialmente, todas as amostras têm o mesmo peso. Após cada iteração, o algoritmo aumenta o peso das amostras que foram classificadas incorretamente pelo estimador atual, fazendo com que o próximo estimador foque mais nestes casos errôneos. A predição final é uma votação ponderada de todos os estimadores, onde os modelos com menor taxa de erro recebem um peso maior.

2.5.5 Agregação por votação

A agregação por votação é uma maneira simples e eficaz de combinar as predições de um conjunto de classificadores diversos. Existem duas abordagens principais:

- Votação majoritária (*hard voting*): A predição final é a classe que recebe a maioria dos votos dos classificadores base. Cada modelo tem um voto, e a classe mais votada é a escolhida.
- Votação ponderada (*soft Voting*): Este método agrega as probabilidades preditas por cada classificador para cada classe. A classe final é aquela que possui a maior probabilidade média. Este tipo de votação é frequentemente preferível quando os classificadores individuais são bem calibrados e fornecem estimativas de probabilidade confiáveis.

2.6 AVALIAÇÃO DE MODELOS E VALIDAÇÃO ESTATÍSTICA

A construção e o treinamento de modelos de aprendizagem de máquina, sejam eles clássicos ou quânticos, representam apenas uma parte do ciclo de desenvolvimento. Uma etapa igualmente crucial consiste na avaliação rigorosa de seu desempenho e na validação estatística das comparações entre diferentes arquiteturas. Não é suficiente apenas observar os valores de

uma métrica de desempenho; é necessário garantir que a métrica escolhida seja apropriada para o problema em questão e que as diferenças de desempenho observadas sejam estatisticamente significativas, e não meros artefatos de flutuações aleatórias no processo de treinamento e teste. Esta seção descreve o arcabouço metodológico utilizado neste trabalho para a avaliação e a comparação dos modelos propostos.

2.6.1 Métrica F1

A escolha de uma métrica de avaliação adequada é fundamental para se obter uma compreensão fidedigna do desempenho de um classificador. Embora a acurácia — a proporção de predições corretas — seja uma métrica intuitiva, ela pode ser enganosa em cenários com classes desbalanceadas. Por exemplo, em um conjunto de dados com 95% de amostras da classe A e 5% da classe B, um classificador trivial que sempre prediz a classe A alcançaria uma acurácia de 95%, embora seja completamente inútil para identificar a classe minoritária. Para contornar esta limitação, utilizam-se métricas baseadas na matriz de confusão, como a precisão e a revocação.

A precisão mede a exatidão das predições positivas, sendo definida como a razão entre os verdadeiros positivos (TP) e o total de predições positivas (TP + FP, onde FP são os falsos positivos). Ela responde à pergunta: “Dentre todas as instâncias que o modelo classificou como positivas, quantas eram de fato positivas?”.

$$\text{Precisão} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2.7)$$

A revocação, também conhecida como sensibilidade, mede a completude das predições, sendo definida como a razão entre os verdadeiros positivos (TP) e o total de instâncias que são de fato positivas (TP + FN, onde FN são os falsos negativos). Ela responde à pergunta: “Dentre todas as instâncias que eram de fato positivas, quantas o modelo conseguiu identificar?”.

$$\text{Revocação} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2.8)$$

Frequentemente, existe um *trade-off* entre a precisão e a revocação. A métrica F1 foi concebida para agregar estas duas medidas em um único valor, sendo definida como a média harmônica entre elas. A média harmônica penaliza valores extremos de forma mais severa que a média aritmética, de modo que um F1 elevado só é alcançado quando tanto a precisão quanto a revocação são altas (RIJSBERGEN, 1979).

$$F_1 = 2 \cdot \frac{\text{Precisão} \cdot \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (2.9)$$

Devido à sua robustez em cenários com desbalanceamento de classes, a métrica F1 é a métrica principal utilizada nesta dissertação para avaliar e comparar o desempenho dos modelos de classificação de sentimentos.

2.6.2 Testes de hipótese para comparação de classificadores

Ao comparar o desempenho de dois ou mais classificadores, observar uma diferença no valor médio de uma métrica como a métrica F1 não é suficiente para declarar a superioridade de um modelo sobre o outro. Essa diferença pode ser resultado de flutuações estocásticas inerentes ao processo experimental (como inicialização aleatória de parâmetros ou amostragem de dados). Para determinar se a diferença observada é estatisticamente significativa, recorre-se a testes de hipótese (DEMsAR, 2006).

Neste trabalho, adota-se o teste pareado de Wilcoxon. Trata-se de um teste de hipótese não-paramétrico utilizado para comparar duas amostras pareadas e determinar se suas medianas populacionais diferem. A escolha de um teste não-paramétrico é particularmente adequada para a comparação de classificadores, uma vez que as distribuições de suas métricas de desempenho frequentemente não seguem uma distribuição normal, uma suposição fundamental para testes paramétricos como o teste t-pareado (DEMsAR, 2006). No contexto desta pesquisa, as amostras pareadas são os conjuntos de 30 valores de F1 obtidos para cada par de modelos sob as mesmas condições experimentais (mesmas 30 *seeds* iniciais).

Quando múltiplas comparações de hipóteses são realizadas simultaneamente, surge o problema das comparações múltiplas. Se cada teste individual for realizado com um nível de significância α , a probabilidade de se cometer ao menos um erro do Tipo I (rejeitar incorretamente uma hipótese nula verdadeira) em toda a família de testes aumenta consideravelmente. Para controlar esta taxa de erro familiar, é necessário aplicar um procedimento de correção dos p-valores.

Para este fim, emprega-se a correção de Holm-Bonferroni, proposta por Holm (1979). Este método é um procedimento sequencial que ajusta o nível de significância para cada teste de forma menos conservadora que a correção de Bonferroni tradicional, oferecendo maior poder estatístico. O método ordena os p-valores do menor para o maior e os compara com

níveis de significância sequencialmente ajustados ($\frac{\alpha}{k}, \frac{\alpha}{k-1}, \dots, \alpha$; sendo k o número de p-valores avaliados), proporcionando um controle rigoroso da taxa de erro familiar supracitada. A utilização do teste de Wilcoxon em conjunto com a correção de Holm-Bonferroni estabelece, portanto, um arcabouço estatístico robusto para validar as conclusões sobre as diferenças de desempenho entre os modelos analisados nesta dissertação.

3 MÉTODO PROPOSTO E DESENHO EXPERIMENTAL

Este capítulo é dedicado à apresentação do método central investigado nesta dissertação e ao detalhamento do desenho experimental concebido para sua validação e análise. Primeiramente, na seção 3.1, descreve-se o modelo de classificação quântica proposto, detalhando sua arquitetura, os componentes variacionais e as estratégias de treinamento. Em seguida, na seção 3.2, é apresentado o desenho experimental, que engloba a configuração do ambiente, as bases de dados utilizadas, os modelos clássicos de comparação e o arcabouço estatístico empregado para garantir a robustez das conclusões.

3.1 MODELO HÍBRIDO QUÂNTICO-CLÁSSICO PROPOSTO

Esta seção descreve a arquitetura e os componentes do classificador quântico variacional que constitui o objeto de estudo desta pesquisa. O modelo segue um fluxo híbrido quântico-clássico, onde dados textuais são pré-processados e codificados em estados quânticos, que são então manipulados por um circuito parametrizado para realizar a tarefa de classificação de sentimentos.

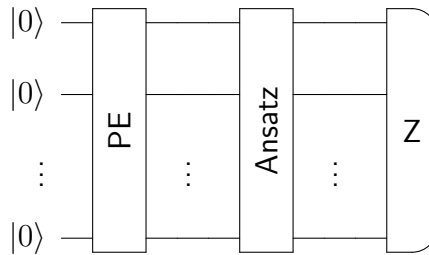
3.1.1 Arquitetura geral do classificador

O classificador quântico proposto opera em três estágios fundamentais, conforme o esquemático exibido na Figura 3.

1. PE: o vetor de características clássico, proveniente do *embedding* de uma sentença, é codificado em um estado quântico. Neste trabalho, utiliza-se a técnica de *amplitude embedding* (MOTTONEN et al., 2004), que codifica um vetor normalizado de N características nas amplitudes de um estado de $\log_2(N)$ *qubits*.
2. *Ansatz* variacional: O estado quântico preparado é então processado por um CQP, ou *ansatz*. Este circuito é composto por uma sequência de portas quânticas cujas rotações são definidas por um conjunto de parâmetros clássicos otimizáveis, θ . O projeto deste *ansatz* é o foco central da investigação.
3. Medição: Ao final do circuito, uma medição é realizada sobre os *qubits*. Para esta tarefa

de classificação, mede-se o valor esperado do observável Pauli Z em todos os *qubits*, o que produz uma saída clássica no intervalo $[-1,1]$. Este valor é então mapeado para as classes do problema, onde -1 representa o sentimento negativo e $+1$, o positivo.

Figura 3 – Template de circuito com preparação de estado por amplitude (PE), o bloco de *ansatz* variável através das possibilidades da Figura 4 e medição pelo observável Z .



Fonte: elaborada pelo autor (2025)

3.1.2 Catálogo de *ansatzes* investigados

Para investigar sistematicamente a relação entre a arquitetura do CQP e seu desempenho, este trabalho propõe e analisa 15 *ansatzes* distintos, cujos diagramas são apresentados na Figura 4. A seleção destes circuitos foi projetada para cobrir uma gama diversificada de topologias e capacidades de emaranhamento, permitindo uma análise incremental de sua complexidade. Diversas arquiteturas foram inspiradas em propostas da literatura, notavelmente do estudo comparativo de (SIM; JOHNSON; ASPURU-GUZIK, 2019). A justificativa para a proposição de cada circuito é detalhada a seguir:

- **Circuitos 1, 2 e 3:** Representam as arquiteturas mais fundamentais, servindo como linhas de base não-emaranhadoras. Cada circuito utiliza uma única camada de portas de rotação em torno de um dos eixos cartesianos da Esfera de Bloch: R_X para o circuito 1, R_Y para o circuito 2 e R_Z para o circuito 3.
- **Circuito 4:** Propõe um incremento em relação aos anteriores ao combinar duas camadas de rotação em eixos não comutáveis (R_X e R_Z). Esta arquitetura é inspirada no “Circuito 1” do estudo de Sim, Johnson e Aspuru-Guzik (2019) e visa testar se a capacidade de gerar rotações arbitrárias sobre um *qubit* melhora o desempenho.

- **Circuito 5:** Utiliza uma camada de portas de rotação genéricas (R), que são, por definição, universais para um único *qubit*. Este circuito é proposto para uma comparação direta com o circuito 4, avaliando se uma parametrização mais explícita da rotação oferece vantagens sobre a composição de rotações em eixos fixos.
- **Circuitos 6:** Adiciona uma camada de portas CNOT em topologia circular ao circuito 5. O objetivo é introduzir emaranhamento de forma explícita e avaliar o impacto da adição de correlações quânticas a um *ansatz* já expressivo em termos de operações de um *qubit*.
- **Circuitos 7 a 11:** Formam um bloco de estudo para uma análise incremental do emaranhamento.
 - O **circuito 7** inicia com uma estrutura simples, combinando portas Hadamard (H) e rotações R_X .
 - O **circuito 8** adiciona portas de emaranhamento CZ ao circuito 7. Sua arquitetura é inspirada no “Circuito 9” de Sim, Johnson e Aspuru-Guzik (2019), que por sua vez se baseia no conceito de *Quantum Kitchen Sinks* de Wilson et al. (2018).
 - O **circuito 9** substitui as portas R_X do circuito 8 por rotações genéricas R , para testar o efeito de uma parametrização mais rica.
 - O **circuito 10** e o **circuito 11** são análogos aos circuitos 8 e 9, respectivamente, mas substituem as portas CZ por CNOTs, permitindo uma comparação direta entre diferentes tipos de portas de controle.
- **Circuitos 12 e 13:** Investigam arquiteturas com conectividade de vizinhos próximos. O **circuito 12** implementa um padrão simples de rotações controladas. O **circuito 13**, inspirado no “Circuito 7” de (SIM; JOHNSON; ASPURU-GUZZIK, 2019) e utilizado no algoritmo QVECTOR (JOHNSON et al., 2017), intercala camadas de rotações não controladas entre as portas de emaranhamento.
- **Circuitos 14 e 15:** São propostos para avaliar o impacto de topologias de conectividade mais complexas.
 - O **circuito 14** emprega uma construção em blocos de circuito, conforme proposto por Schuld et al. (2020), que combina interações de vizinhos próximos com in-

terações de longo alcance, sendo análogo ao “Circuito 15” de Sim, Johnson e Aspuru-Guzik (2019).

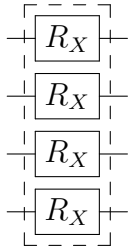
- O **circuito 15**, inspirado no “Circuito 6” de Sim, Johnson e Aspuru-Guzik (2019) e originalmente proposto por Sousa e Ramos (2006), implementa uma conectividade do tipo “todos-para-todos”, representando o mais alto grau de interação entre os *qubits* dentre os modelos testados.

3.1.3 Agregação de modelos quânticos via *ensemble*

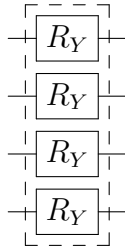
Além da análise de circuitos individuais, este trabalho propõe a construção e avaliação de ensembles de modelos quânticos. O objetivo é investigar se a combinação de múltiplos classificadores pode levar a um desempenho mais robusto e preciso. As estratégias de ensemble e os circuitos base utilizados estão listados no Quadro 1. Para os ensembles do tipo *AdaBoost* e *Bagging*, foram selecionados os circuitos 6, 9 e 14 devido às suas topologias que, embora relativamente enxutas, apresentam um alto potencial de emaranhamento e expressabilidade. Adicionalmente, são propostos ensembles baseados em votação (*hard* e *soft voting*) para analisar o comportamento agregado de circuitos com diferentes níveis de complexidade e topologias de conectividade.

Figura 4 – Conjunto de *ansatze* utilizados com seus códigos identificação única.

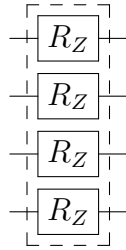
(a) Circuito 1



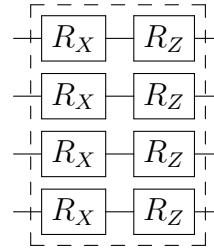
(b) Circuito 2



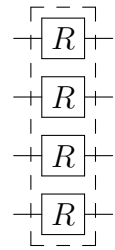
(c) Circuito 3



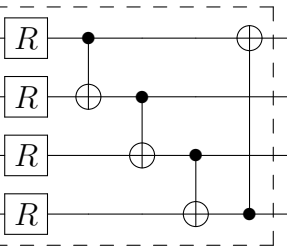
(d) Circuito 4



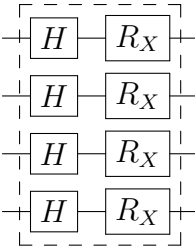
(e) Circuito 5



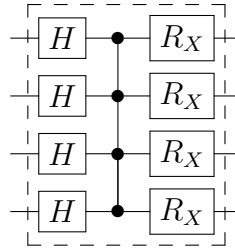
(f) Circuito 6



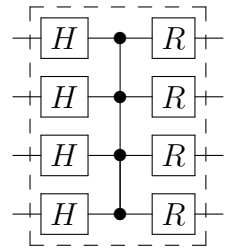
(g) Circuito 7



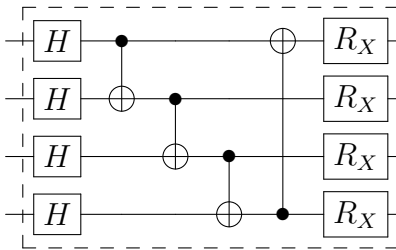
(h) Circuito 8



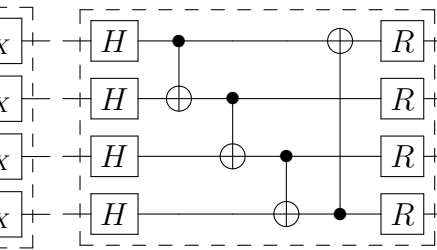
(i) Circuito 9



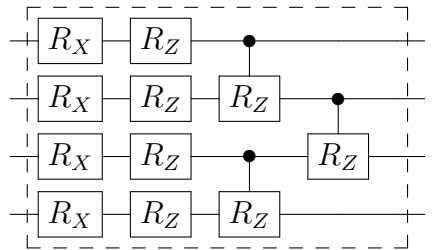
(j) Circuito 10



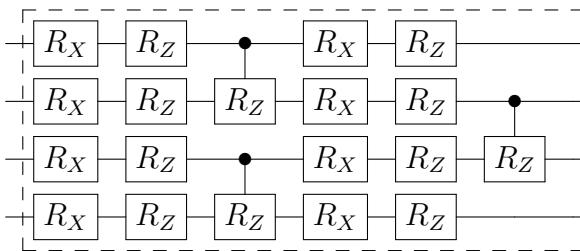
(k) Circuito 11



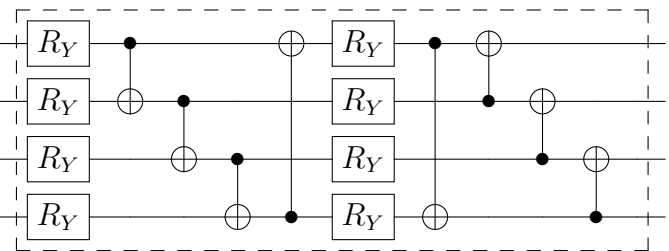
(l) Circuito 12



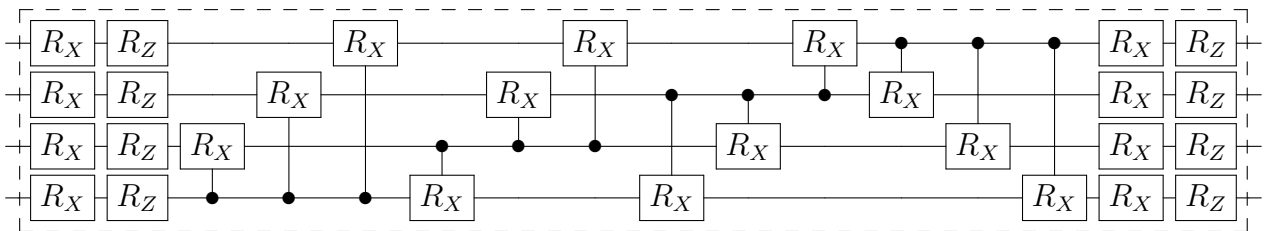
(m) Circuito 13



(n) Circuito 14



(o) Circuito 15



Fonte: elaborada pelo autor (2025)

Quadro 1 – *Ensemble* de modelos quânticos utilizados

ID	Descrição
16	<i>AdaBoost</i> do circuito 6
17	<i>Bagging</i> do circuito 6
18	<i>AdaBoost</i> do circuito 9
19	<i>Bagging</i> do circuito 9
20	<i>AdaBoost</i> do circuito 14
21	<i>Bagging</i> do circuito 14
22	<i>Soft voting</i> dos circuitos 1, 2 e 3
23	<i>Hard voting</i> dos circuitos 1, 2 e 3
24	<i>Soft voting</i> dos circuitos 1, 2, 3 e 5
25	<i>Hard voting</i> dos circuitos 1, 2, 3 e 5
26	<i>Soft voting</i> dos circuitos 1, 2, 3, 5 e 6
27	<i>Hard voting</i> dos circuitos 1, 2, 3, 5 e 6
28	<i>Soft voting</i> dos circuitos 7–11
29	<i>Hard voting</i> dos circuitos 7–11
30	<i>Soft voting</i> dos circuitos 12, 14 e 15
31	<i>Hard voting</i> dos circuitos 12, 14 e 15

Fonte: Elaborada pelo autor (2025)

3.2 DESENHO EXPERIMENTAL PARA VALIDAÇÃO

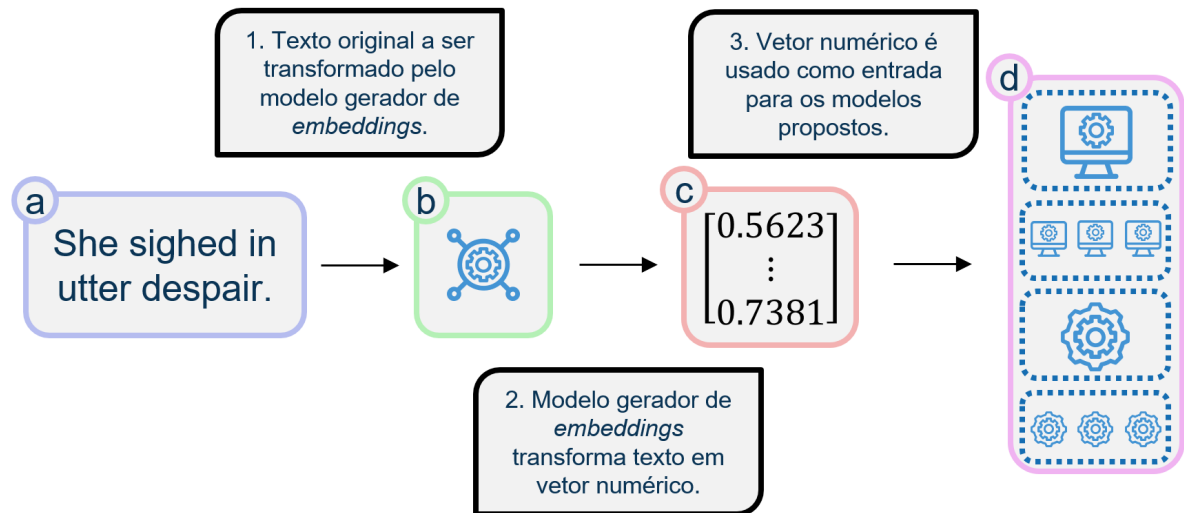
Esta seção descreve o desenho experimental utilizado para treinar, avaliar e comparar os modelos quânticos propostos na seção anterior, bem como os modelos clássicos de referência. O fluxo geral do processo experimental é apresentado no diagrama da Figura 5.

3.2.1 Configuração de ambiente

Esta pesquisa foi executada em um ambiente Linux com sistema operacional Ubuntu 24.04 dentro do Windows 11 utilizando o *Windows Subsystem for Linux* (WSL). O computador no qual este ambiente foi montado possui uma CPU Intel® Core™ i9-10900F e 128 gigabytes de RAM. Somado a isso, utilizou-se a linguagem de programação Python (ROSSUM; JR, 1995) na versão 3.12.8 com as seguintes bibliotecas:

- Aeon (MIDDLEHURST et al., 2024) versão 1.1.0;
- Matplotlib (HUNTER, 2007) versão 3.10.3;

Figura 5 – Fluxo geral do método experimental, detalhando as etapas sequenciais: (a) seleção e preparação das bases de dados; (b) conversão do texto em vetores numéricos (*embeddings*); (c) análise da dimensionalidade do vetor de entrada; e (d) treinamento e avaliação dos modelos de classificação quânticos e clássicos, incluindo seus *ensembles*.



Fonte: elaborada pelo autor (2025)

- Numpy (HARRIS et al., 2020) versão 2.0.2;
- Pandas (TEAM, 2024) versão 2.2.3;
- PennyLane (BERGHOLM et al., 2022) versão 0.41.1;
- PyTreeBank (RAIMAN, 2020) versão 0.2.7;
- Scikit-Learn (PEDREGOSA et al., 2018) versão 1.6.1;
- Sentence Transformers (REIMERS; GUREVYCH, 2019b) versão 4.1.0;
- Statsmodels (SEABOLD; PERKTOLD, 2010) versão 0.14.4;
- Weights and Biases (BIEWALD, 2020) versão 0.19.11.

Todos os códigos pertinentes a esta pesquisa podem ser encontrados em (FILHO, 2025).

3.2.2 Base de dados

Quatro bases de dados sustentam esta pesquisa, sendo três delas sintéticas a partir de Kart-saklis et al. (2021) para distinguir frases de tom alegre ou triste, e a quarta correspondendo ao corpus público SST (SOCHER et al., 2013) através da biblioteca PyTreeBank (RAIMAN, 2020).

Para criar os conjuntos sintéticos, empregou-se o *ChatGPT* (OPENAI, 2025) com um *prompt* desenvolvido especificamente para este trabalho. Obtiveram-se três níveis de complexidade:

- Fácil: frases curtas (5–7 palavras), sem nomes próprios;
- Médio: maior número de termos e grau de ambiguidade lexical;
- Difícil: inclusão de nomes próprios, coletivos e outros elementos que elevam a ambiguidade.

Cada conjunto contém 100 frases por classe (alegre/triste) para treinamento e 25 por classe para teste, todas geradas em inglês para garantir compatibilidade com as ferramentas de pré-processamento e com os modelos de *embeddings*.

O SST complementa esses dados, reunindo 11.855 sentenças de críticas de filmes pré-processadas pelo parser da Stanford e anotadas por três avaliadores humanos, totalizando 215.154 frases únicas. Neste estudo, adota-se a classificação binária, agrupando as sentenças negativas (ou levemente negativas) em uma classe e as levemente positivas (ou positivas) em outra, descartando-se as sentenças neutras.

3.2.3 Processamento e Representação dos Dados

A conversão do texto em vetores numéricos é uma etapa fundamental do pipeline. Para esta tarefa, serão empregados três modelos geradores de *embeddings*: sentence-transformers/all-mpnet-base-v2, a ser chamado apenas de modelo *MPNet*, tomaarsen/mpnet-base-nli-matryoshka, modelo *Matryoshka*, e nomic-ai/nomic-embed-text-v1.5, modelo *Nomic*. A utilização de múltiplos geradores visa avaliar a sensibilidade dos classificadores à representação de entrada.

Adicionalmente, será conduzido um estudo sobre o impacto da dimensionalidade do vetor de entrada no desempenho dos modelos. Utilizando o modelo *Matryoshka*, que permite a geração de vetores de tamanhos variados, serão testadas as dimensionalidades de 768, 32 e 16 atributos. O objetivo principal desta etapa é investigar a resiliência dos modelos à redução de dimensionalidade e justificar a escolha de uma representação de entrada mais compacta (16 atributos) para os experimentos principais, o que, no caso quântico, impacta diretamente na largura (número de *qubits*) do circuito.

3.2.4 Modelos Clássicos para Comparação

Para estabelecer uma linha de base de desempenho e contextualizar os resultados dos modelos quânticos, serão utilizados sete algoritmos de aprendizagem de máquina clássicos, conforme listados no Quadro 2. Estes modelos foram obtidos através da biblioteca Scikit-Learn e utilizados com seus parâmetros padrão, a fim de garantir um benchmark justo e reprodutível. Adicionalmente, serão construídos e avaliados ensembles de modelos clássicos, detalhados no Quadro 3, para permitir uma comparação direta com as estratégias de ensemble quânticas.

Quadro 2 – Modelos clássicos utilizados

ID	Descrição
32	Regressão logística
33	Perceptron
34	K-vizinhos mais próximos
35	Floresta randômica
36	SVM com kernel linear
37	SVM com kernel polinomial
38	SVM com kernel RBF

Fonte: Elaborada pelo autor (2025)

Quadro 3 – *Ensemble* de modelos clássicos utilizados

ID	Descrição
39	<i>AdaBoost</i> da regressão logística
40	<i>Bagging</i> da regressão logística
41	<i>Soft voting</i> de SVM (linear, polinomial, RBF)
42	<i>Hard voting</i> de SVM (linear, polinomial, RBF)
43	<i>Soft voting</i> de regressão logística, perceptron e KNN
44	<i>Hard voting</i> de regressão logística, perceptron e KNN

Fonte: Elaborada pelo autor (2025)

3.2.5 Estratégia de Otimização e Treinamento

Para todos os modelos quânticos, a otimização dos parâmetros θ do *ansatz* é realizada pelo otimizador Adam (KINGMA; BA, 2017). O treinamento se utiliza de um tamanho de *batch*

de 20 amostras para as bases sintéticas e 512 para a base SST, sendo executado por 40 épocas em todos os cenários. A métrica final para avaliação de desempenho é a F1.

3.2.6 Análise de Propriedades Intrínsecas e Desempenho

Parte central deste trabalho é a investigação da relação entre as propriedades intrínsecas dos circuitos e seu desempenho prático. Para tanto, serão calculadas as métricas de expressabilidade e poder de emaranhamento para cada um dos 15 *ansatze* propostos, tanto com uma ($L = 1$) quanto com dez ($L = 10$) camadas de profundidade. Adicionalmente, investiga-se como a etapa de preparação de estados (*amplitude embedding*) afeta estas mesmas medidas. Os valores obtidos para estas métricas serão então correlacionados com o desempenho F1 alcançada pelos respectivos circuitos em cada base de dados, buscando identificar padrões e dependências.

3.2.7 Validação Estatística

Para garantir a robustez das conclusões, cada experimento será repetido 30 vezes com diferentes sementes de aleatoriedade (seeds). A comparação de desempenho entre os modelos será realizada por meio do teste pareado de Wilcoxon (WILCOXON, 1945), aplicado sobre as distribuições dos 30 valores da métrica F1. Para controlar o erro do Tipo I decorrente das múltiplas comparações, os p-valores obtidos serão corrigidos utilizando o método de Holm-Bonferroni (HOLM, 1979). Um resultado será considerado estatisticamente significativo se o p-valor ajustado for inferior a 0.05. Adicionalmente, serão construídos diagramas de diferença crítica para visualização dos agrupamentos de modelos com desempenho estatisticamente equivalente.

4 RESULTADOS E DISCUSSÃO

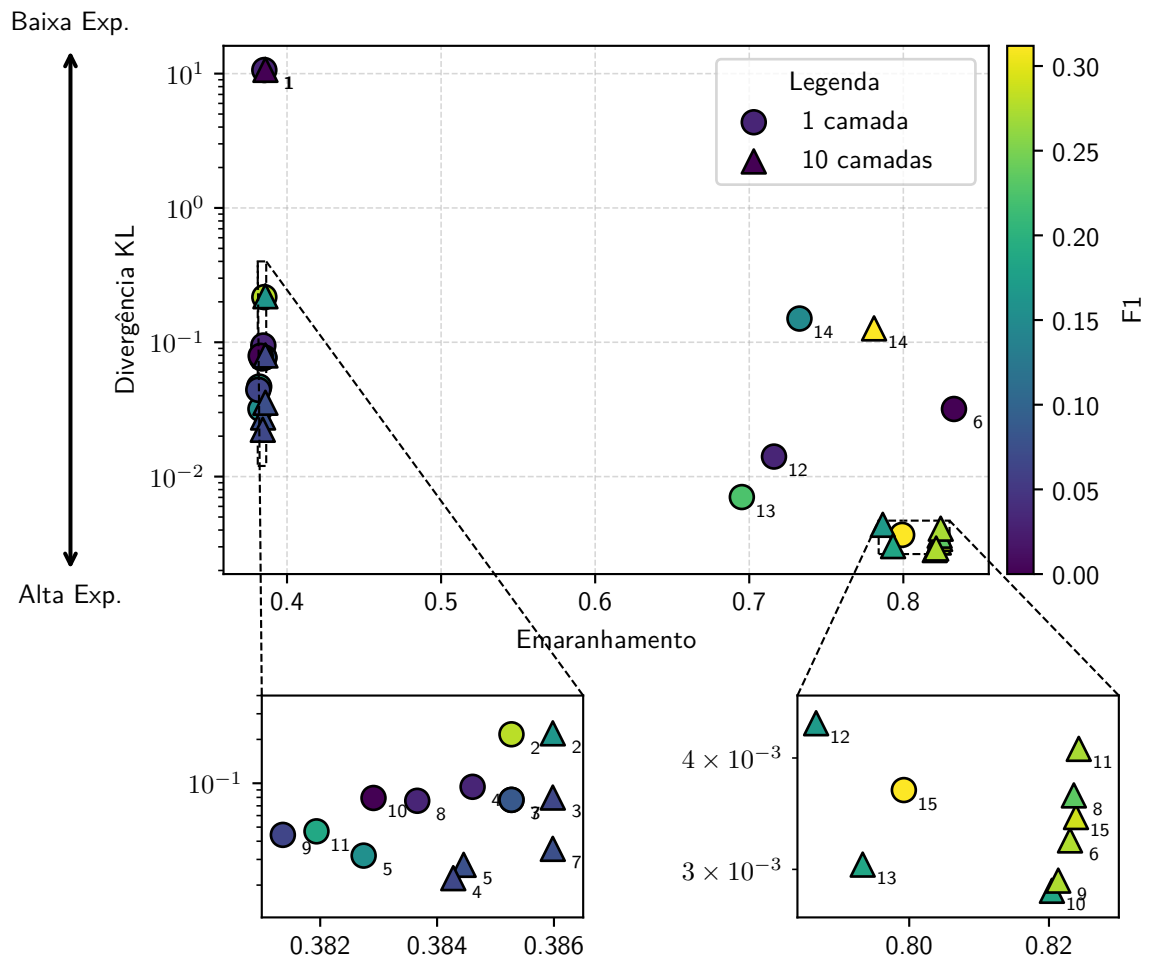
Neste capítulo, procede-se à análise e interpretação dos resultados experimentais obtidos à luz da metodologia delineada no capítulo 3. O objetivo é contextualizar os achados, correlacionando o desempenho empírico dos modelos com suas propriedades teóricas e arquitetônicas. A análise subsequente está estruturada em seções que buscam, de forma progressiva, destrinchar os fatores que influenciaram o desempenho dos classificadores. Inicialmente, a discussão se aprofunda no papel das propriedades quânticas intrínsecas dos circuitos, como o emaranhamento e a expressabilidade. Em seguida, investiga-se a influência da representação vetorial de entrada, abordando a sensibilidade dos modelos à escolha do gerador de *embeddings* e à redução de dimensionalidade. Por fim, realiza-se uma análise comparativa detalhada entre os modelos quânticos e os *benchmarks* clássicos, situando os resultados obtidos no panorama atual da área e discutindo as implicações para a busca por uma vantagem quântica em tarefas de processamento de linguagem natural.

4.1 PAPEL DO EMARANHAMENTO, PROFUNDIDADE E EXPRESSABILIDADE NA CAPACIDADE REPRESENTACIONAL

A análise minuciosa dos resultados obtidos nos experimentos revela uma hierarquia clara de fatores que governam o desempenho dos classificadores quânticos variacionais na tarefa de análise de sentimentos. Dentre estes, a capacidade intrínseca do *ansatz* de gerar emaranhamento, a profundidade do circuito e a sua expressabilidade emergem não como meros otimizadores de desempenho, mas como elementos fundamentais que ditam a capacidade representacional do modelo. Esta seção dedica-se a interpretar os resultados sob a ótica dessas propriedades, demonstrando como sua interação é determinante para o sucesso da classificação, especialmente ao transitar de cenários sintéticos para dados do mundo real.

Observa-se, primeiramente, um agrupamento natural de circuitos com desempenho limitado, notadamente os de natureza não-emaranhadora. Através da Figura 6, percebe-se que os circuitos 1 a 5 e 7, cujas arquiteturas se baseiam exclusivamente em portas de rotação de um único *qubit*, compartilham uma limitação fundamental: sua incapacidade de criar ou manipular ativamente as correlações quânticas entre os *qubits*. Embora demonstrem desempenhos variáveis nas bases sintéticas, frequentemente influenciados pela qualidade do *embedding* de

Figura 6 – Resultados das medidas de expressabilidade e emaranhamento e a métrica F1 na base SST. Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados. Cada subíndice no canto inferior direito refere-se ao código do circuito equivalente.



Fonte: Elaborada pelo autor (2025)

entrada, seu desempenho na base SST é consistentemente modesta. Esta falha considerável em um cenário de dados complexo e ruidoso sugere que a capacidade de processamento local, ainda que sofisticada, é insuficiente para a tarefa.

Para quantificar esta limitação, nota-se que mesmo os *ansatze* não-emaranhadores mais performáticos neste grupo não alcançam resultados expressivos na base SST, tendo em vista os resultados consolidados através das Figuras 7, 8, 9, 10, 11, 12, 13 e 14. O circuito 5, que utiliza uma porta de rotação genérica (R), atinge uma mediana de F1 de aproximadamente 0.16 com $L = 1$. De forma similar, o circuito 2 (R_Y) alcança uma mediana de F1 perto de 0.28 com $L = 1$. Em contrapartida, o circuito 1 (R_X) e o circuito 4 (R_X, R_Z) apresentam um desempenho residual, com medianas de F1 de aproximadamente 0.03 para $L = 1$. O aumento da profundidade para $L = 10$ nestes circuitos não produziu ganhos estatisticamente relevantes na maioria dos casos, e em nenhuma situação foi capaz de elevá-los a um patamar

competitivo. Este comportamento reforça a hipótese de que a mera adição de parâmetros em uma arquitetura não-emaranhadora não expande sua capacidade representacional de forma a capturar a complexidade semântica de dados do mundo real.

Em contraposição, os circuitos projetados com capacidade de emaranhamento intrínseco (6, 8, 9, 11, 12, 13, 14 e 15) exibem uma dinâmica fundamentalmente distinta. Para este grupo, a profundidade do circuito atua como um fator habilitador. A transição da profundidade de $L = 1$ para $L = 10$ resulta, quase invariavelmente, em um salto de desempenho expressivo e estatisticamente relevante na base SST. O caso mais emblemático é o do circuito 6, que passa de uma mediana de F1 nula para aproximadamente 0.45. De forma análoga, o circuito 14 eleva seu desempenho de uma mediana de F1 de cerca de 0.15 para perto de 0.50, estabelecendo-se como um dos melhores estimadores individuais. O circuito 9, por sua vez, transita de uma mediana de 0.04 para aproximadamente 0.44.

Esta melhora de desempenho evidencia que, para *ansatze* emaranhadores, o aumento do número de camadas não é um simples refinamento, mas um requisito para destravar seu potencial computacional. As camadas adicionais de portas de múltiplos *qubits*, como CNOT ou CZ, permitem ao modelo construir correlações mais complexas e explorar um subespaço de Hilbert mais vasto, uma capacidade que se mostra indispensável para navegar a paisagem de otimização de um problema com dados do mundo real. Enquanto a configuração com $L = 1$ destes circuitos muitas vezes falha em convergir para uma solução útil, a configuração com $L = 10$ demonstra uma capacidade de aprendizado robusta.

Em contrapartida, nota-se um comportamento destoante em comparação com bases sintéticas. Com um exemplo de frases sintéticas visto pelo Quadro 4, é possível compreender os comportamentos da métrica F1 nas bases *ChatGPT* Fácil, Médio e Difícil pelas Figuras 15, 16 e 17. Apesar da escala variar conforme as bases, entende-se um comportamento similar entre elas: circuitos emaranhadores, no geral, obtiveram melhores resultados em comparação com os circuitos sem tal capacidade.

Uma análise teórica das propriedades dos circuitos, apresentada nas Figuras 18 e 19, oferece um embasamento para estes achados empíricos. A etapa de preparação de estados (PE) por amplitude, por si só, já introduz um emaranhamento inicial no sistema, com um valor medido de aproximadamente 0.39. Os circuitos não-emaranhadores (1 a 5 e 7) meramente herdaram este emaranhamento, sendo incapazes de manipulá-lo ou ampliá-lo. Seus valores de emaranhamento, quando combinados com a PE, permanecem em um patamar próximo ao da preparação isolada. Isso indica que seu processamento subsequente não explora ativamente as

Figura 7 – Resultados do desempenho F1 dos circuitos 1 ao 6 na base SST. Esses algoritmos classificadores demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

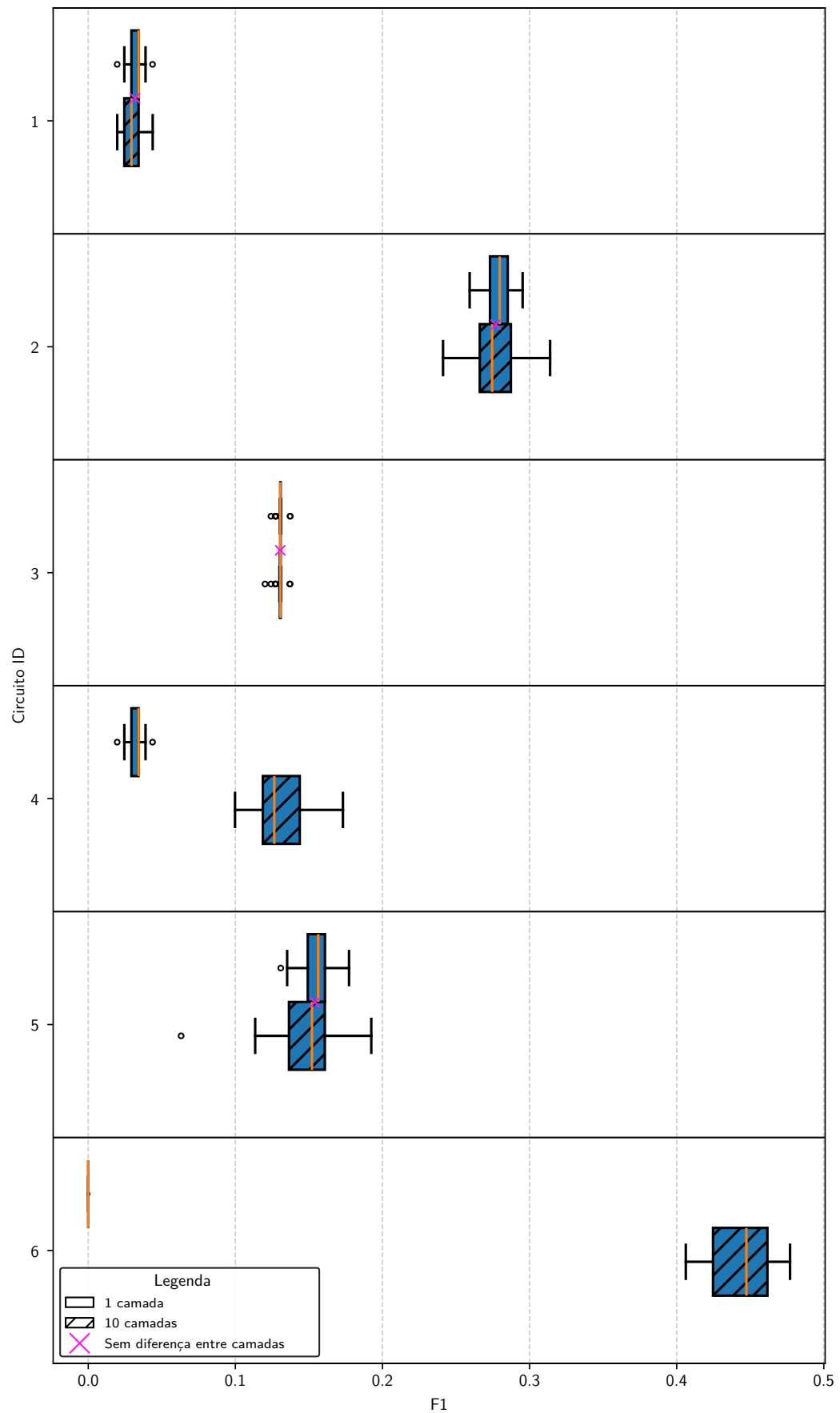
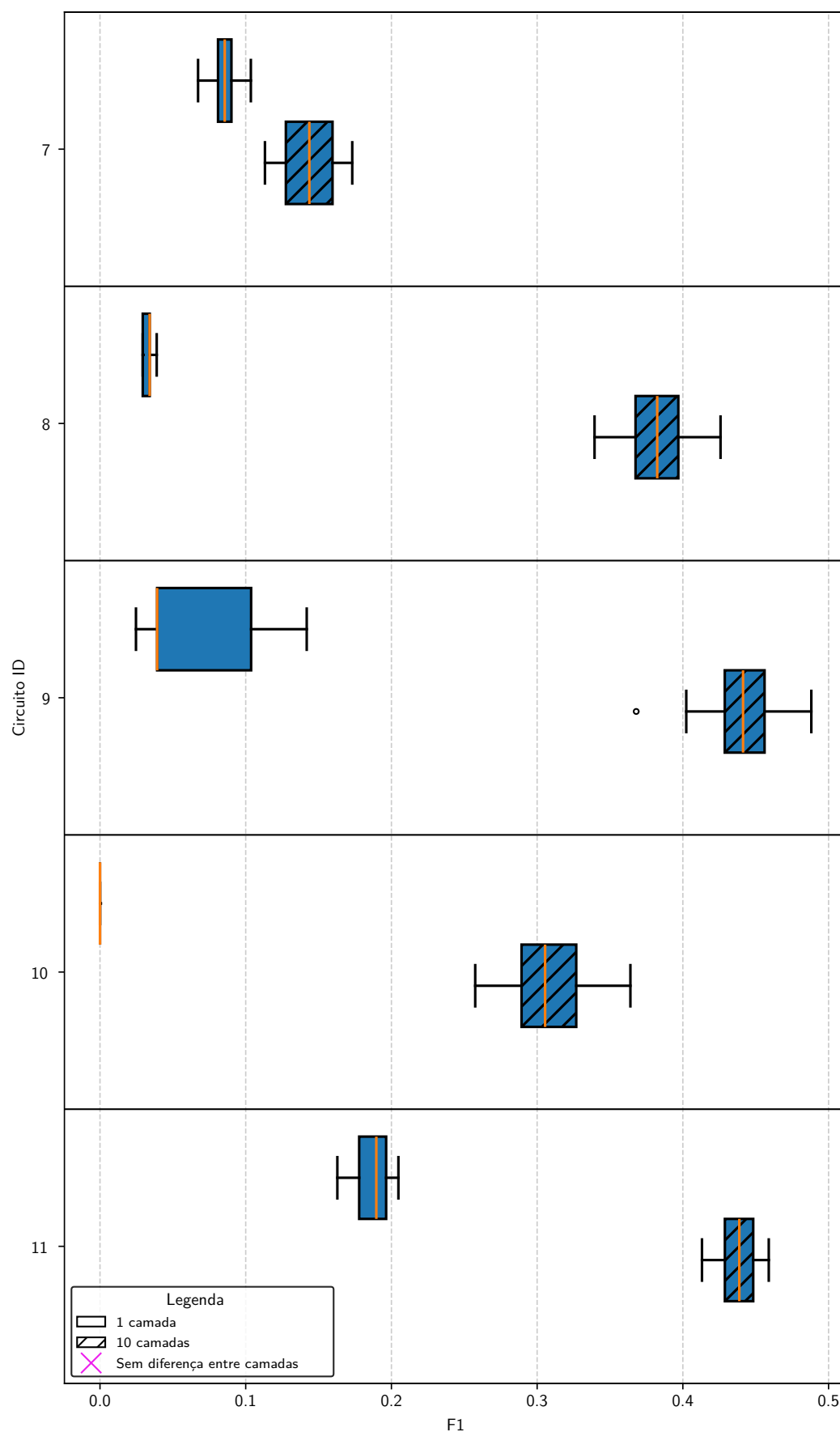
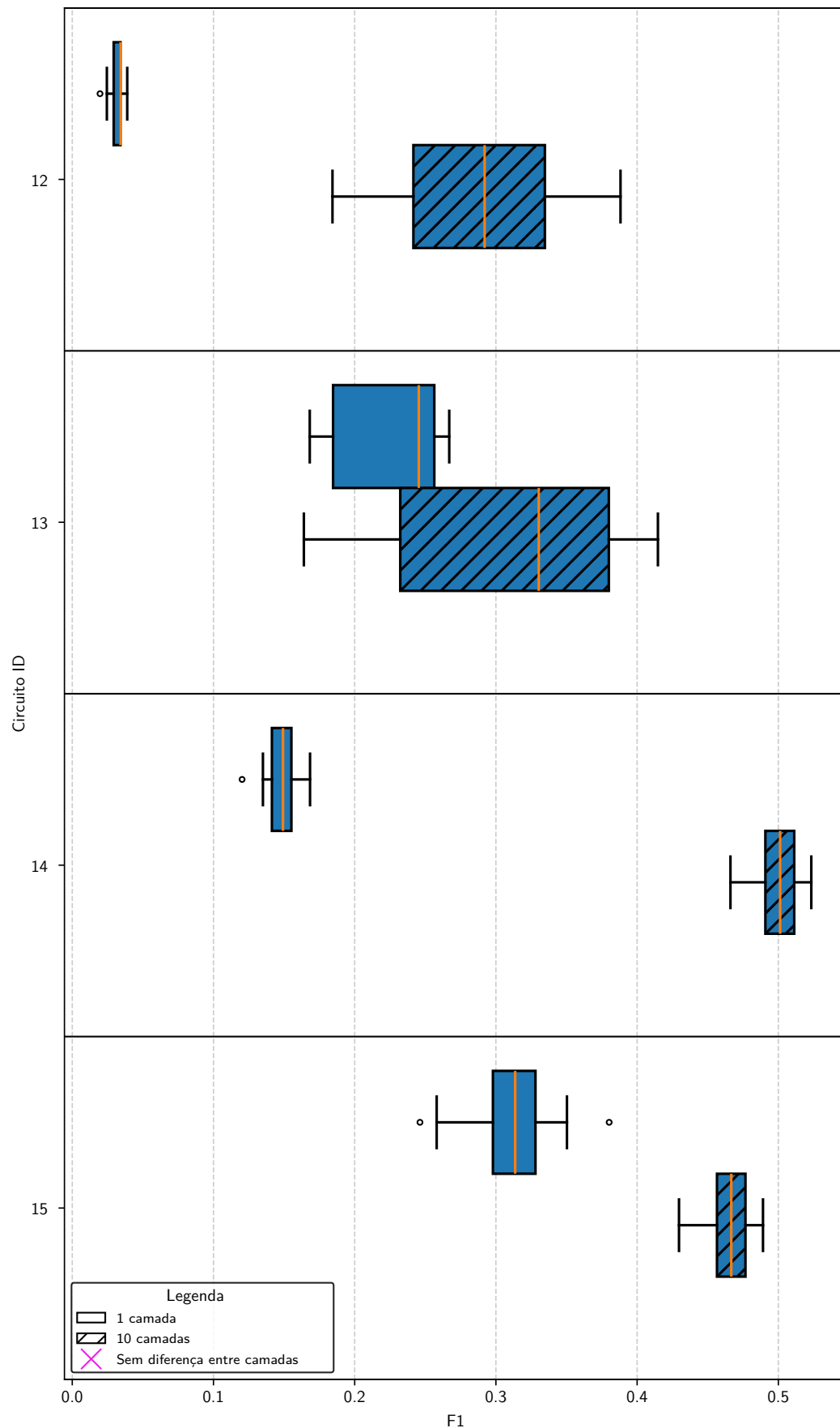


Figura 8 – Resultados do desempenho F1 dos circuitos 7 ao 11 na base SST. Esses algoritmos classificadores implementam diferentes formas de emaranhamento quântico. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

Figura 9 – Resultados do desempenho F1 dos circuitos 12 ao 15 na base SST. Esses algoritmos classificadores implementam diferentes tipos de conectividade entre *qubits*. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

Figura 10 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 na base SST. Esses algoritmos classificadores implementam modelos clássicos lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

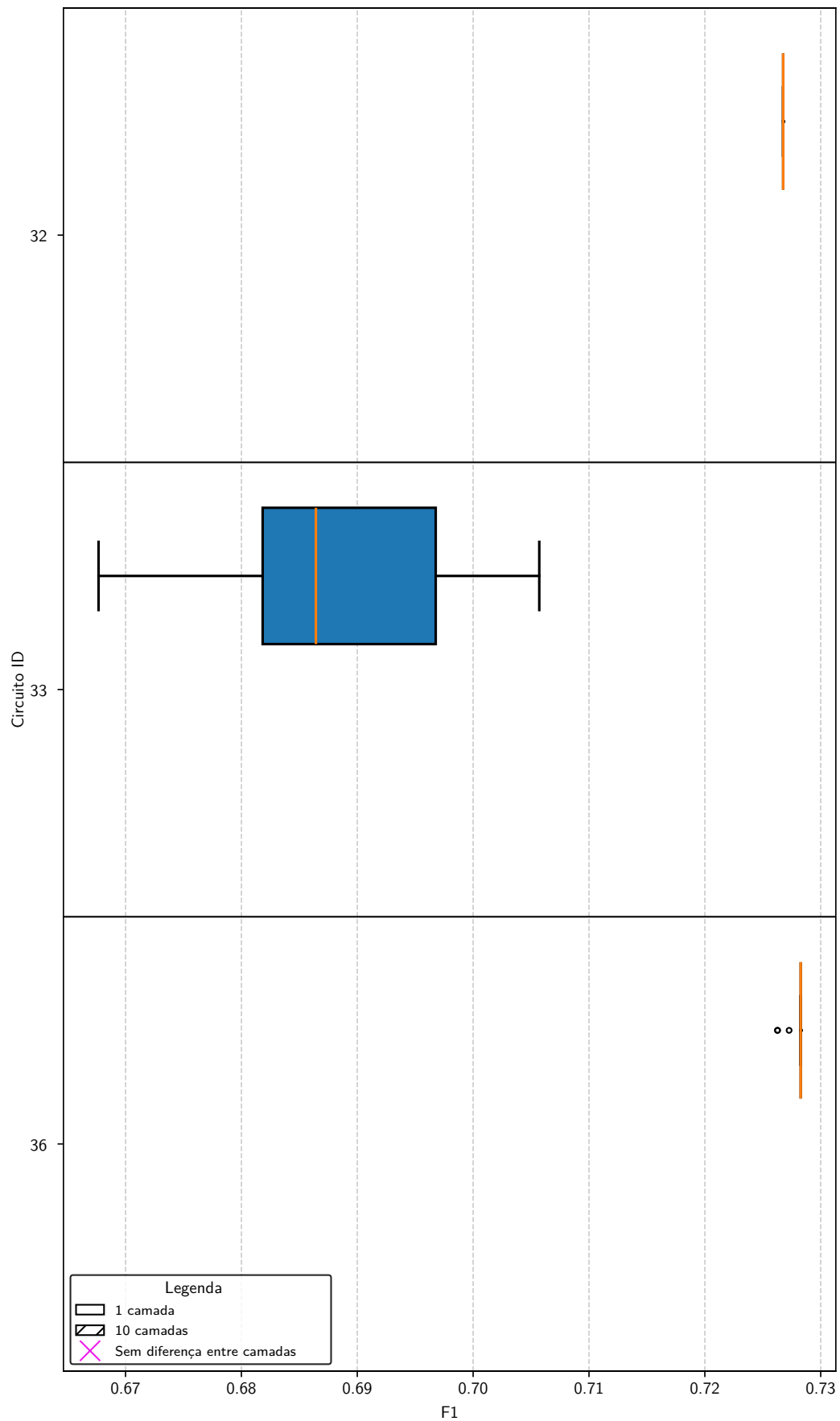
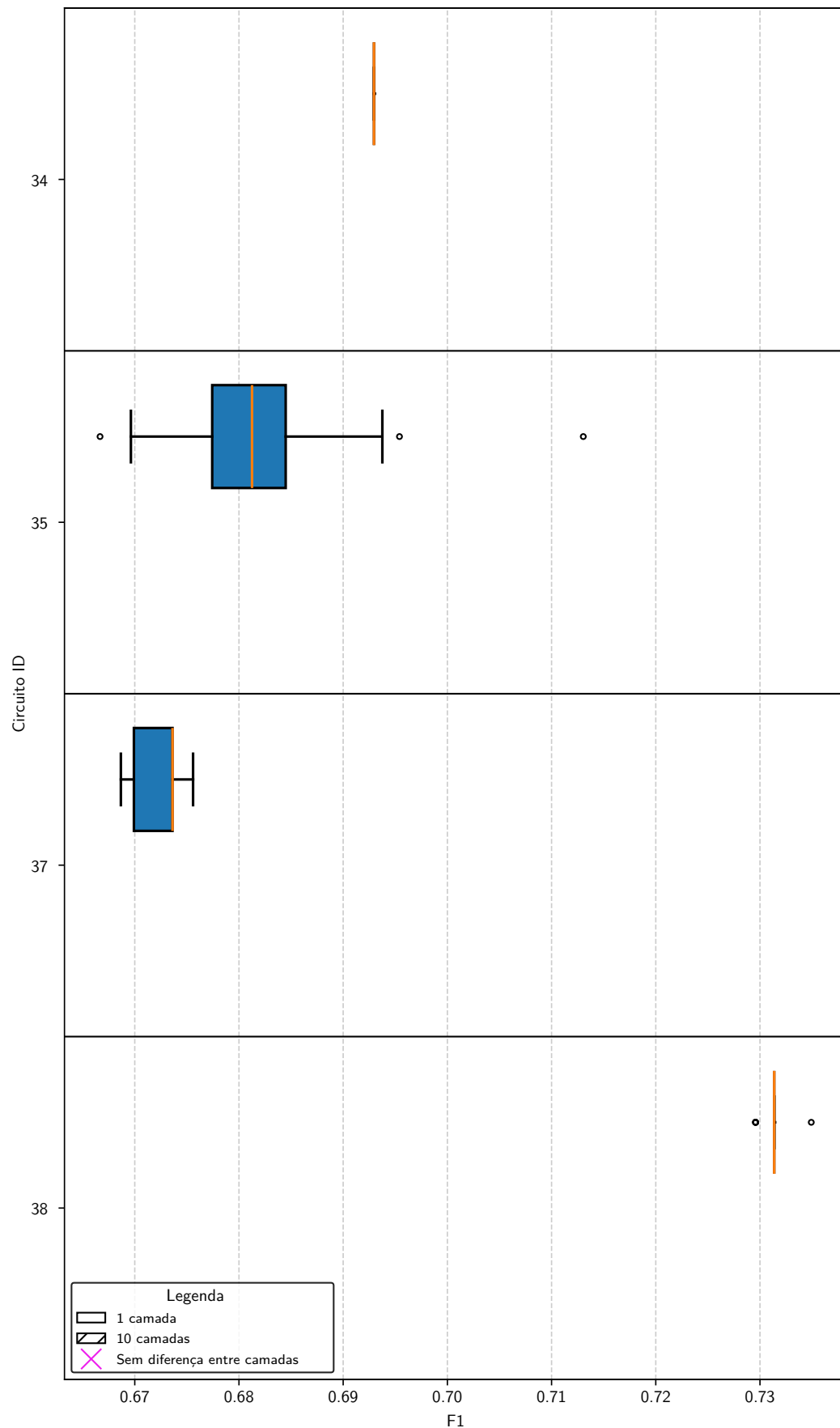
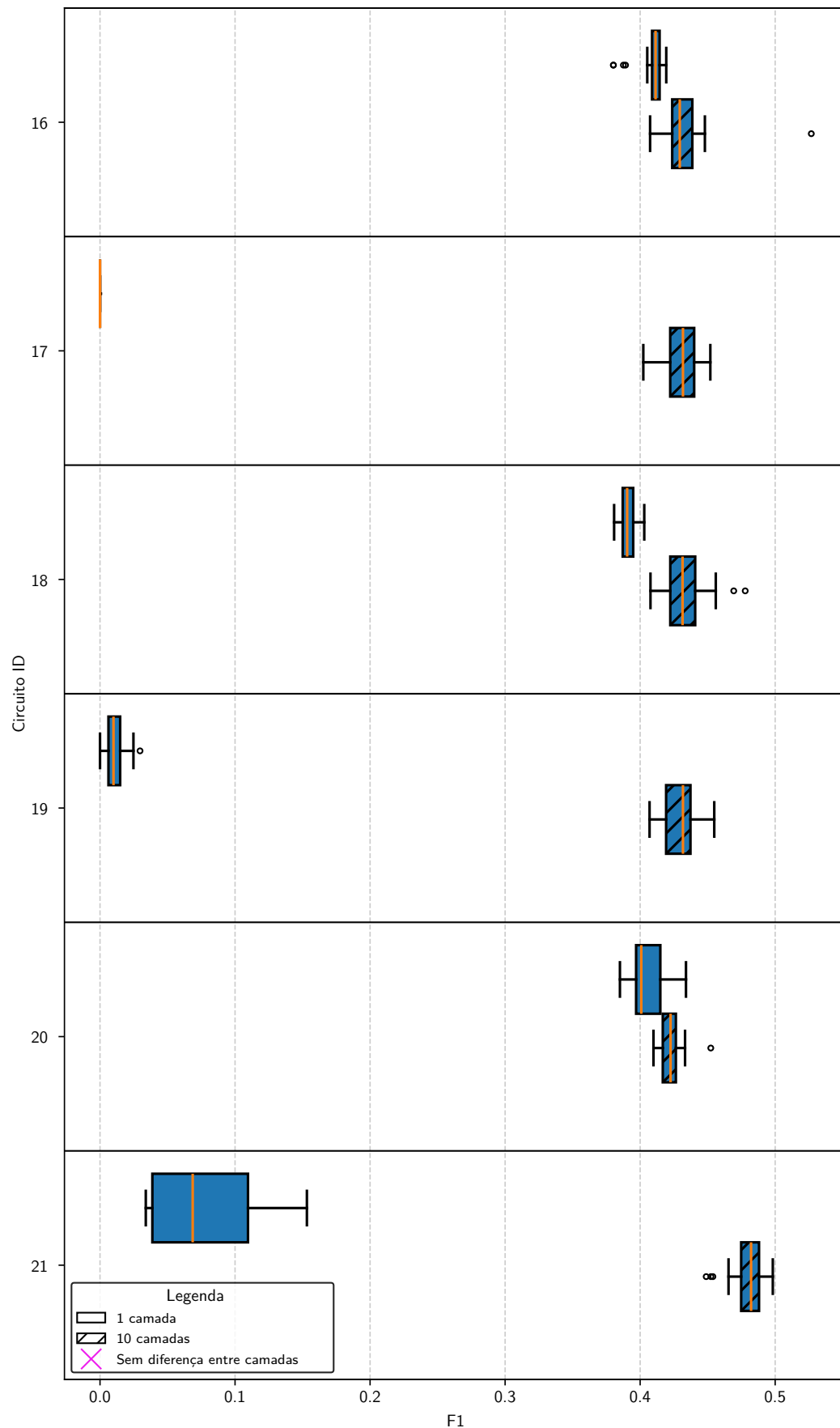


Figura 11 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 na base SST. Esses algoritmos classificadores implementam modelos clássicos não-lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



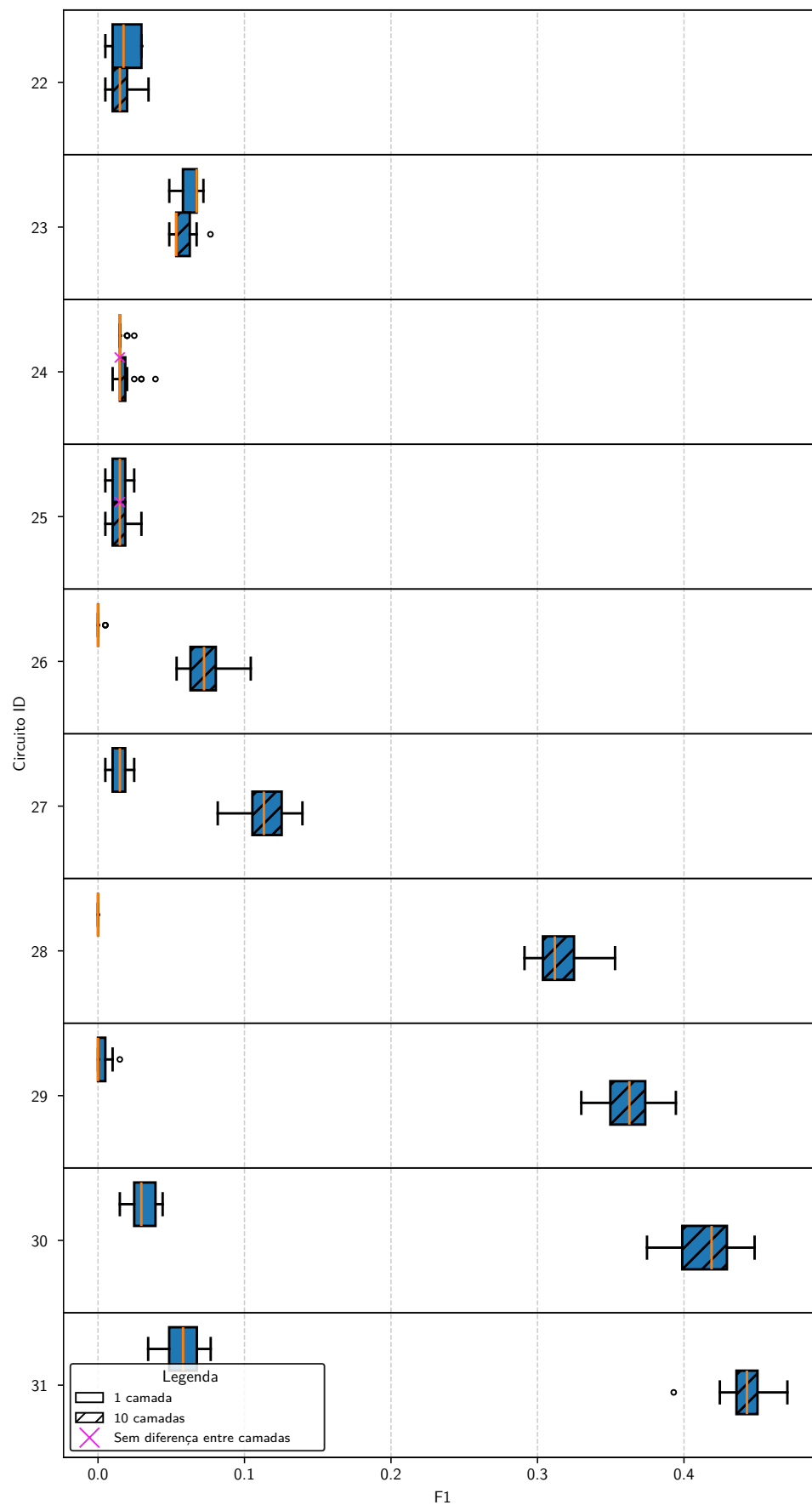
Fonte: Elaborada pelo autor (2025)

Figura 12 – Resultados do desempenho F1 dos circuitos 16 ao 21 na base SST. Esses algoritmos classificadores implementam modelos de ensemble através de *AdaBoost* ou *Bagging* de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



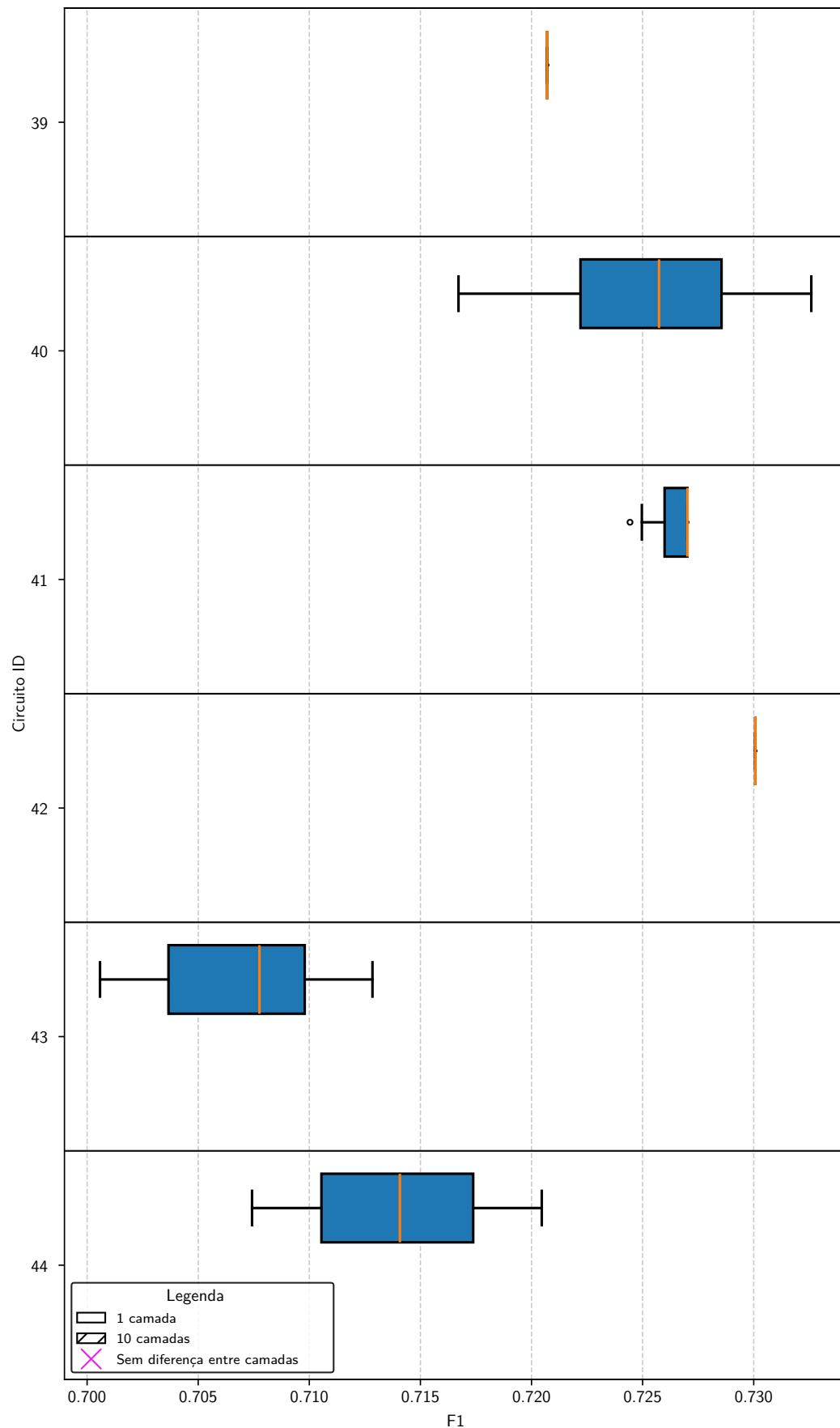
Fonte: Elaborada pelo autor (2025)

Figura 13 – Resultados do desempenho F1 dos circuitos 22 ao 31 na base SST. Esses algoritmos classificadores implementam modelos de ensemble através de *hard* ou *soft voting* de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

Figura 14 – Resultados do desempenho F1 dos circuitos 39 ao 44 na base SST. Esses algoritmos classificadores implementam modelos de ensemble clássicos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



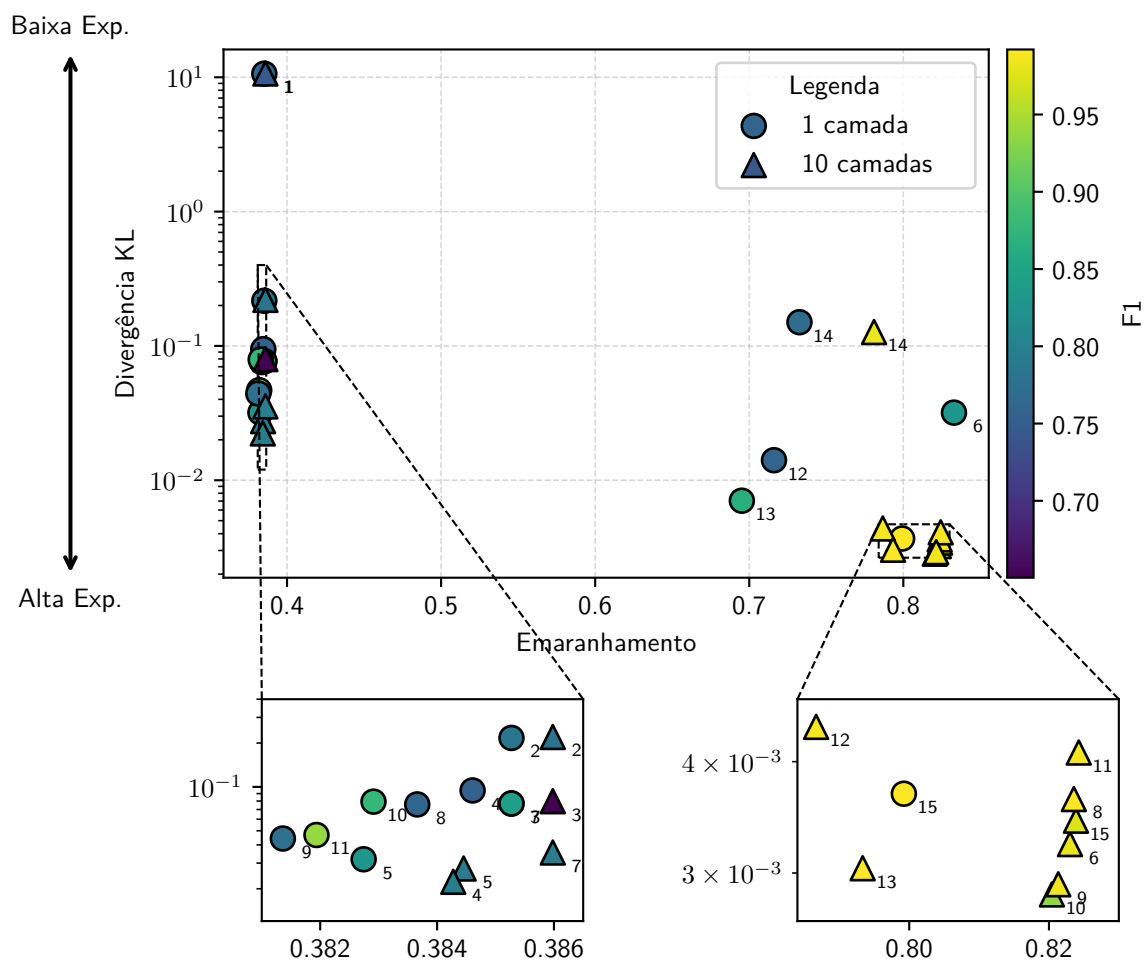
Fonte: Elaborada pelo autor (2025)

Quadro 4 – Exemplo de frases criadas através do *ChatGPT*. É possível perceber um aumento na complexidade lexical das palavras utilizadas.

Nível	Classe	Frase
Fácil	Positiva	She was glad to help her colleagues.
	Negativa	She was mournful after the farewell.
Médio	Positiva	They were optimistic about the future.
	Negativa	He felt disillusioned by the broken promises.
Difícil	Positiva	Jessica relished the challenge of the demanding role.
	Negativa	William pondered the meaning behind the cryptic message.

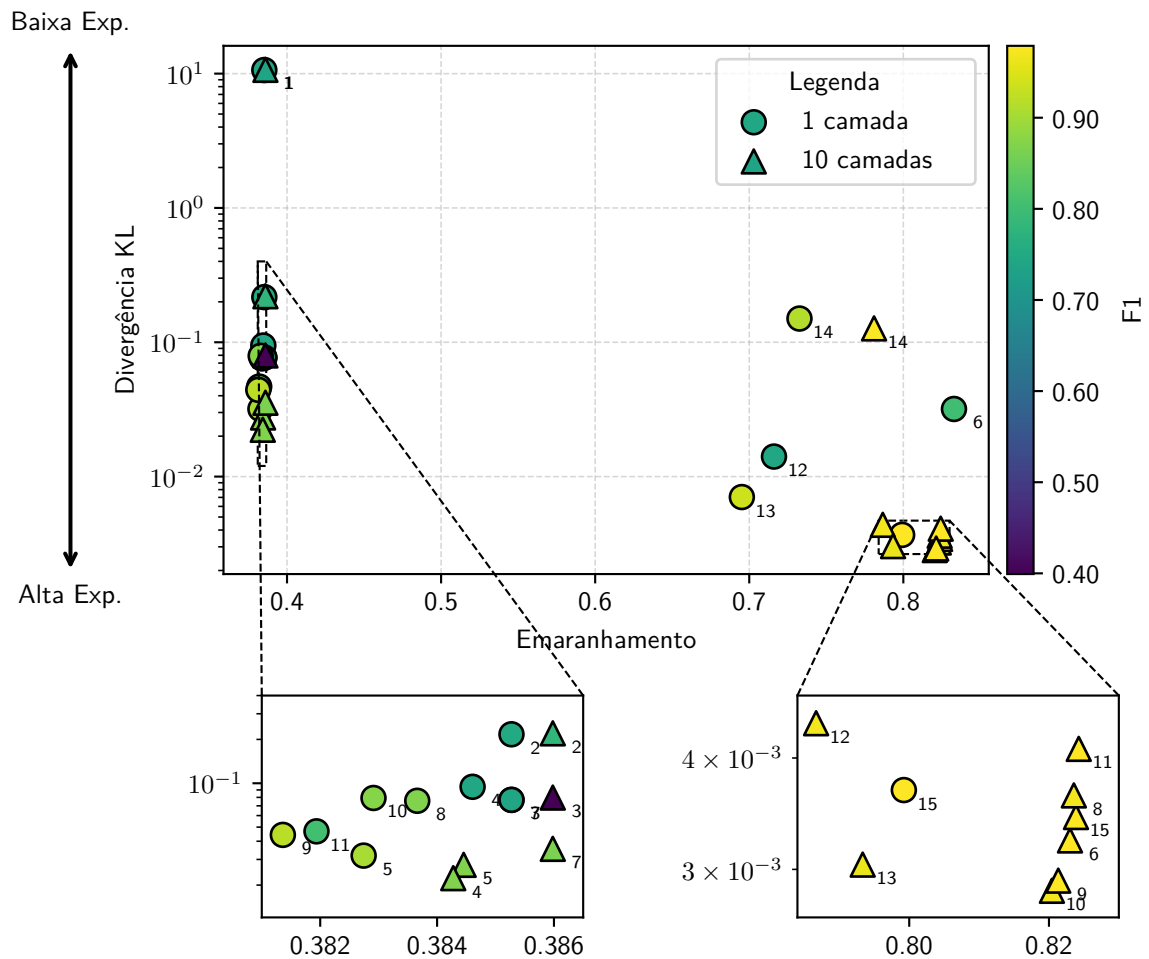
Fonte: Elaborada pelo autor (2025)

Figura 15 – Resultados das medidas de expressabilidade e emaranhamento e a métrica F1 na base *ChatGPT* - Fácil. Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados. Cada subíndice no canto inferior direito refere-se ao código do circuito equivalente.



Fonte: Elaborada pelo autor (2025)

Figura 16 – Resultados das medidas de expressabilidade e emaranhamento e a métrica F1 na base *ChatGPT* - Médio. Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados. Cada subíndice no canto inferior direito refere-se ao código do circuito equivalente.



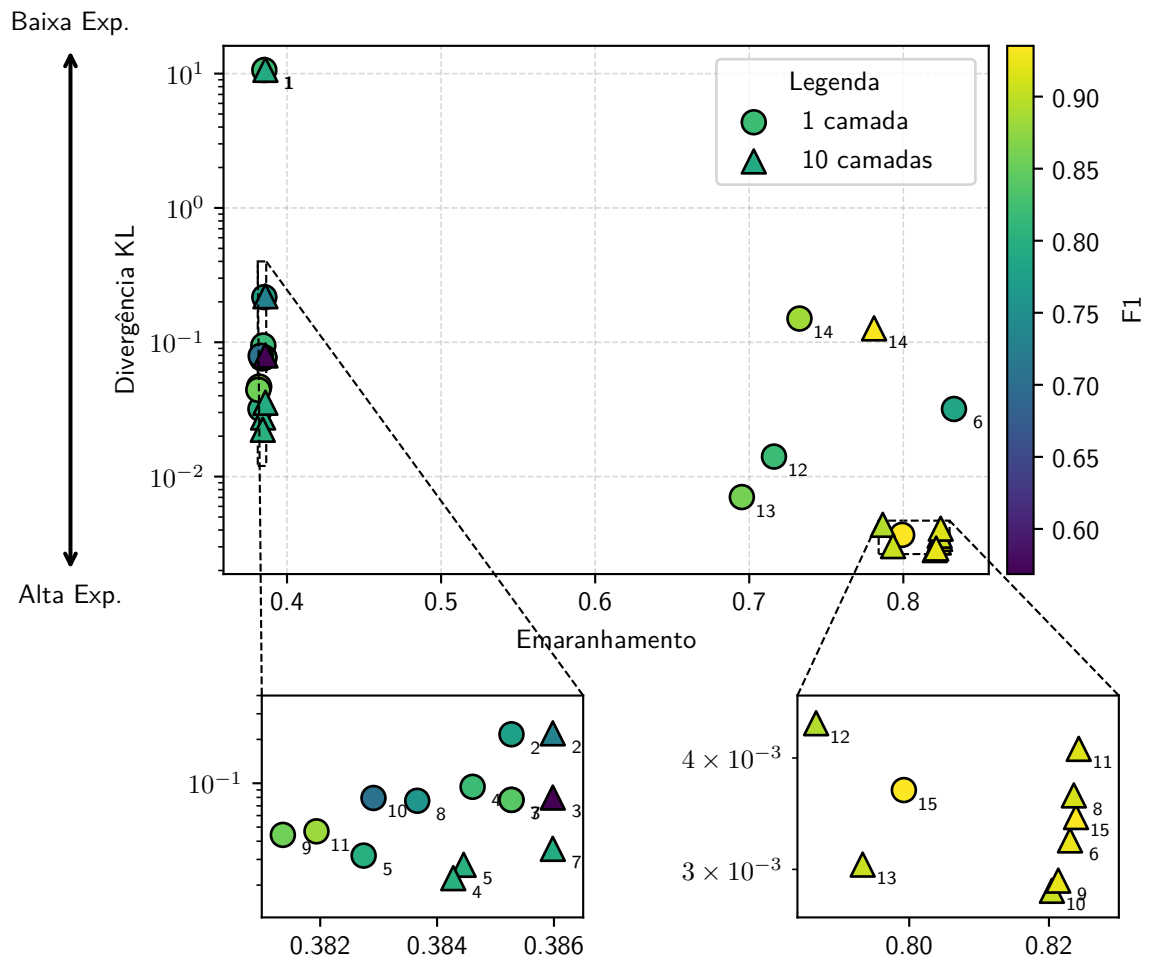
Fonte: Elaborada pelo autor (2025)

correlações quânticas presentes no estado de entrada.

Em contraste, os circuitos com capacidade de emaranhamento intrínseco demonstram um comportamento distinto. Arquiteturas como as dos circuitos 6, 9, 11 e 15, especialmente na configuração com $L = 10$, apresentam valores de emaranhamento próximos, em torno de 0.82, que são substancialmente superiores ao valor herdado da PE. Isso comprova que o *ansatz* está ativamente gerando e acumulando correlações entre os *qubits*, uma característica que se correlaciona diretamente com seu desempenho superior. A capacidade de gerar emaranhamento, e não apenas de recebê-lo como entrada, parece ser a propriedade distintiva que confere poder de classificação a estas arquiteturas.

A expressabilidade, embora seja uma métrica relevante, parece atuar como um fator secundário quando não acompanhada da capacidade de emaranhamento. O circuito 5, por exemplo, quando combinado com a PE e com $L = 1$, exibe uma das melhores métricas de expressa-

Figura 17 – Resultados das medidas de expressabilidade e emaranhamento e a métrica F1 na base *ChatGPT* - Difícil. Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados. Cada subíndice no canto inferior direito refere-se ao código do circuito equivalente.

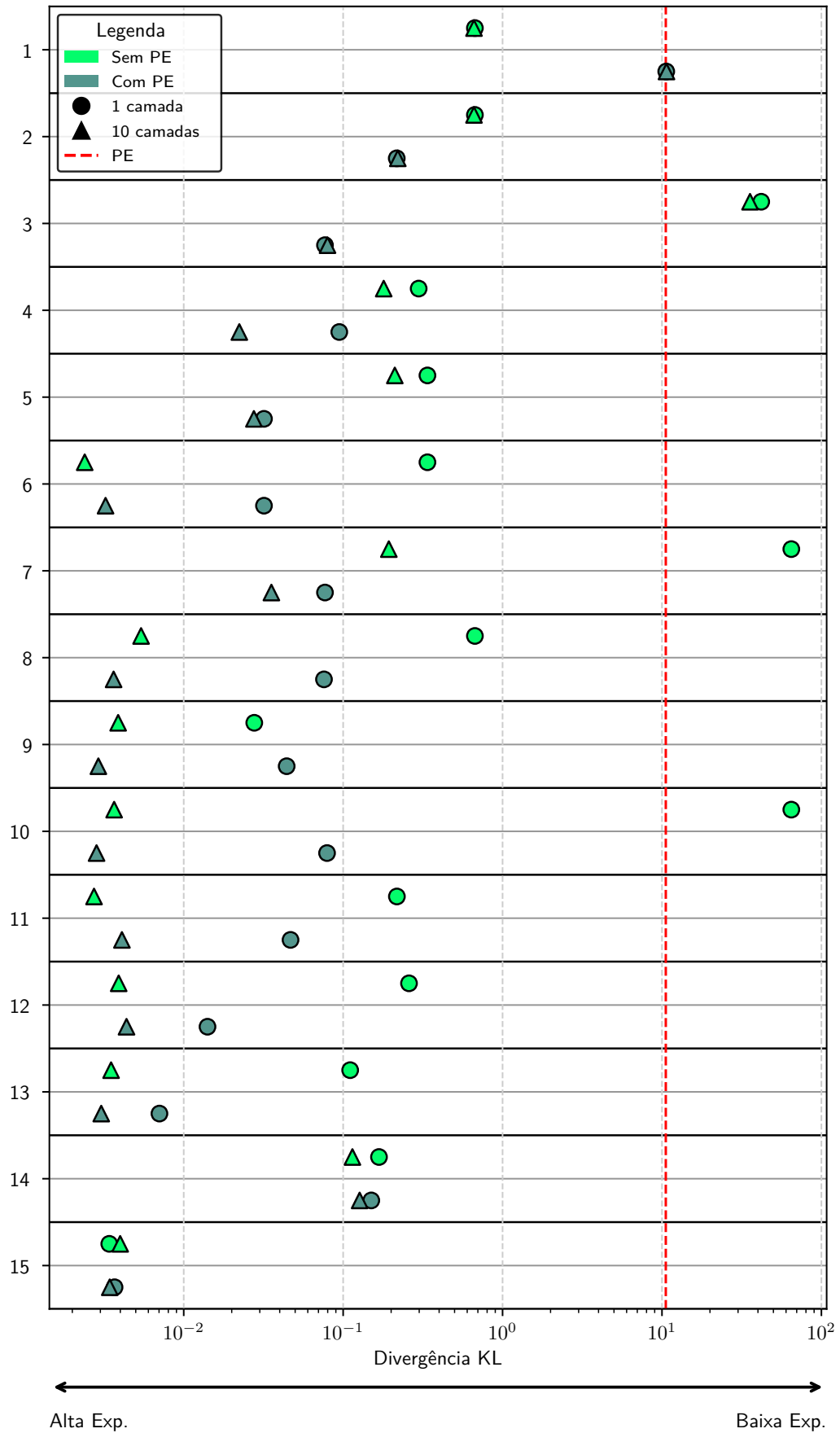


Fonte: Elaborada pelo autor (2025)

bilidade (divergência KL de aproximadamente 0.03), superando, em teoria, outros modelos. Contudo, seu desempenho na base SST (mediana de F1 em torno de 0.16) é modesto. Esta divergência sugere que uma alta capacidade de explorar o espaço de estados de um único *qubit*, por si só, é insuficiente. Uma hipótese para este fenômeno é que uma expressabilidade muito elevada, sem a estrutura conferida por um emaranhamento robusto, pode levar a paisagens de custo complexas e suscetíveis ao fenômeno de *barren plateaus*, dificultando a otimização em cenários com dados ruidosos.

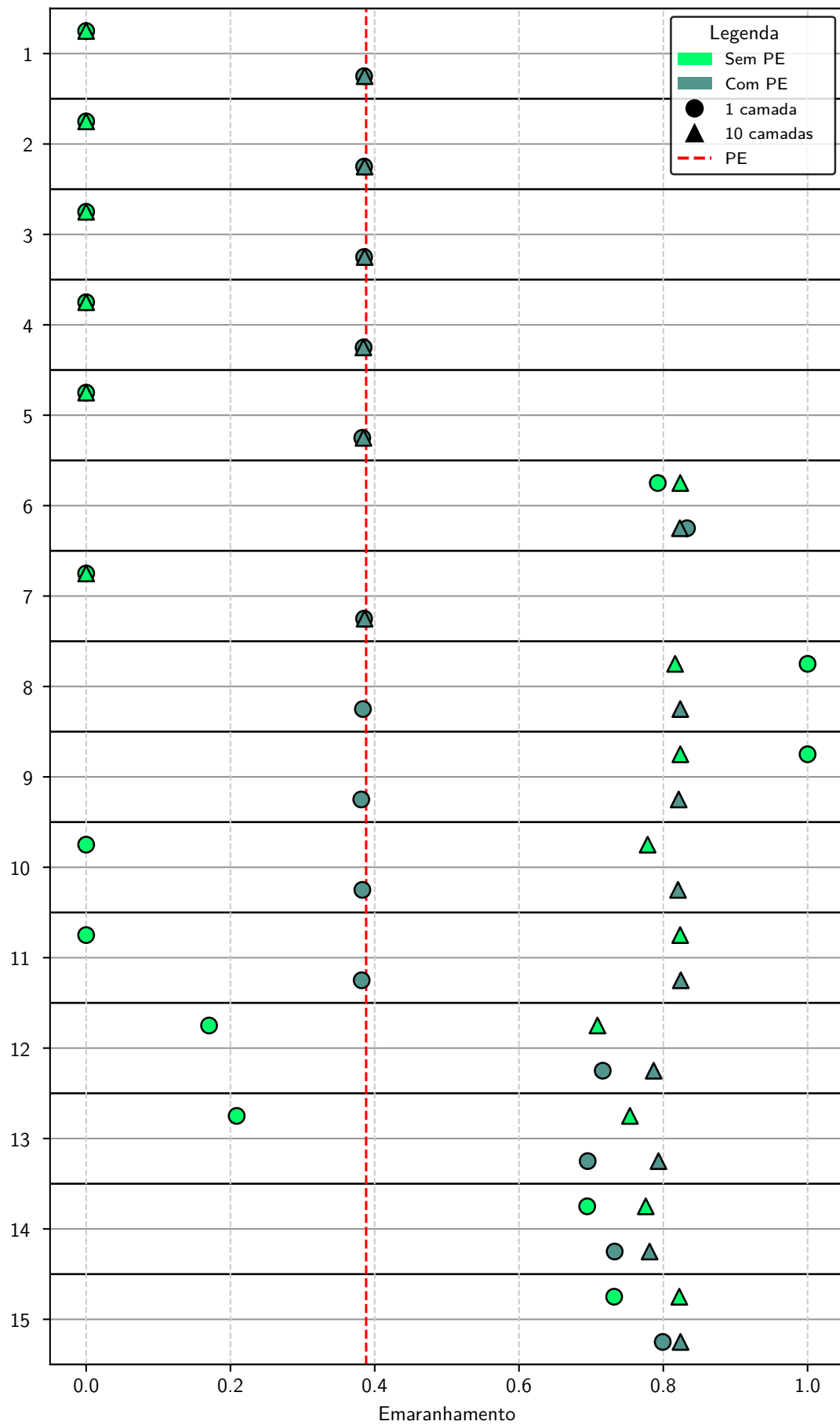
A expressabilidade, portanto, torna-se mais eficaz quando associada a um *ansatz* que pode gerar correlações. O circuito 15 com $L = 1$ ilustra este ponto: ele possui a melhor expressabilidade entre todos os circuitos (KL de próximo de 0.0037) e já possui um poder de emaranhamento relevante (0.73), o que se traduz em um desempenho superior na base SST (mediana de F1 perto de 0.31) em comparação ao circuito 5. Isso indica que a capacidade de

Figura 18 – Resultados da medida de expressabilidade dos circuitos e o impacto da interferência da PE. Utilizou-se o eixo logarítmico da Divergência KL para melhor visualização dos dados.



Fonte: Elaborada pelo autor (2025)

Figura 19 – Resultados da medida de emaranhamento dos circuitos e o impacto da interferência da PE.



Fonte: Elaborada pelo autor (2025)

explorar o espaço de Hilbert é mais bem aproveitada quando o modelo também pode estruturar a informação de forma complexa através das correlações quânticas.

A análise do impacto da profundidade também reforça esta conclusão. Para os circuitos não-entrelaçados, como o 1 e o 2, o aumento de $L = 1$ para $L = 10$ não gera diferença estatística na base SST. Teoricamente, isso ocorre porque múltiplas rotações sequenciais no mesmo eixo são redutíveis a uma única rotação, não expandindo a capacidade fundamental do circuito. Por outro lado, para os circuitos entrelaçados, cada camada adicional introduz uma nova oportunidade de criar correlações, o que se traduz em um ganho de desempenho relevante. O circuito 8, por exemplo, salta de uma mediana de F1 de aproximadamente 0.03 para 0.38 ao passar de $L = 1$ para $L = 10$.

Em síntese, os resultados convergem para uma conclusão central: a combinação de um *ansatz* com capacidade de entrelaçamento intrínseco e uma profundidade adequada é um requisito para o sucesso na tarefa de classificação de sentimentos em cenários de dados do mundo real. Os circuitos não-entrelaçados, independentemente de sua complexidade em termos de rotações de um único *qubit*, mostraram-se fundamentalmente limitados, com seu desempenho sendo dominado pela qualidade dos *embeddings* de entrada e pela simplicidade da tarefa. Em contrapartida, os circuitos entrelaçados, ao serem dotados de profundidade suficiente, demonstraram a capacidade de superar as limitações de uma representação de entrada menos otimizada e de aprender as fronteiras de decisão complexas exigidas pela base SST. Este achado não apenas valida a importância teórica do entrelaçamento, mas também oferece uma diretriz prática para o desenho de futuros classificadores quânticos para tarefas de processamento de linguagem natural, indicando que o investimento em arquiteturas profundas e com alta capacidade de gerar correlações é um caminho promissor. Somado a isso, a expressabilidade, embora relevante, não é um preditor suficiente de sucesso, podendo levar a desafios de otimização quando não acompanhada de outros fatores.

4.2 INFLUÊNCIA DA REPRESENTAÇÃO VETORIAL DE ENTRADA

Uma vez estabelecida a relevância das propriedades intrínsecas dos circuitos quânticos na seção anterior, a análise volta-se para a interface entre o domínio clássico e o quântico: a representação vetorial de entrada. Esta etapa, responsável por traduzir a informação textual em um formato numérico, constitui um pilar fundamental cujo impacto no desempenho dos classificadores, tanto quânticos quanto clássicos, é de relevância central para os objetivos desta

pesquisa. A maneira como os dados são estruturados no espaço de atributos pode facilitar ou dificultar a tarefa de aprendizado, influenciando diretamente a capacidade dos modelos de encontrar fronteiras de decisão eficazes. Desta forma, as subseções a seguir dedicam-se a dissecar, de maneira pormenorizada, a influência de duas facetas desta representação: a resiliência dos modelos à redução de dimensionalidade do vetor de *embeddings* e a sensibilidade de seu desempenho à escolha do modelo gerador.

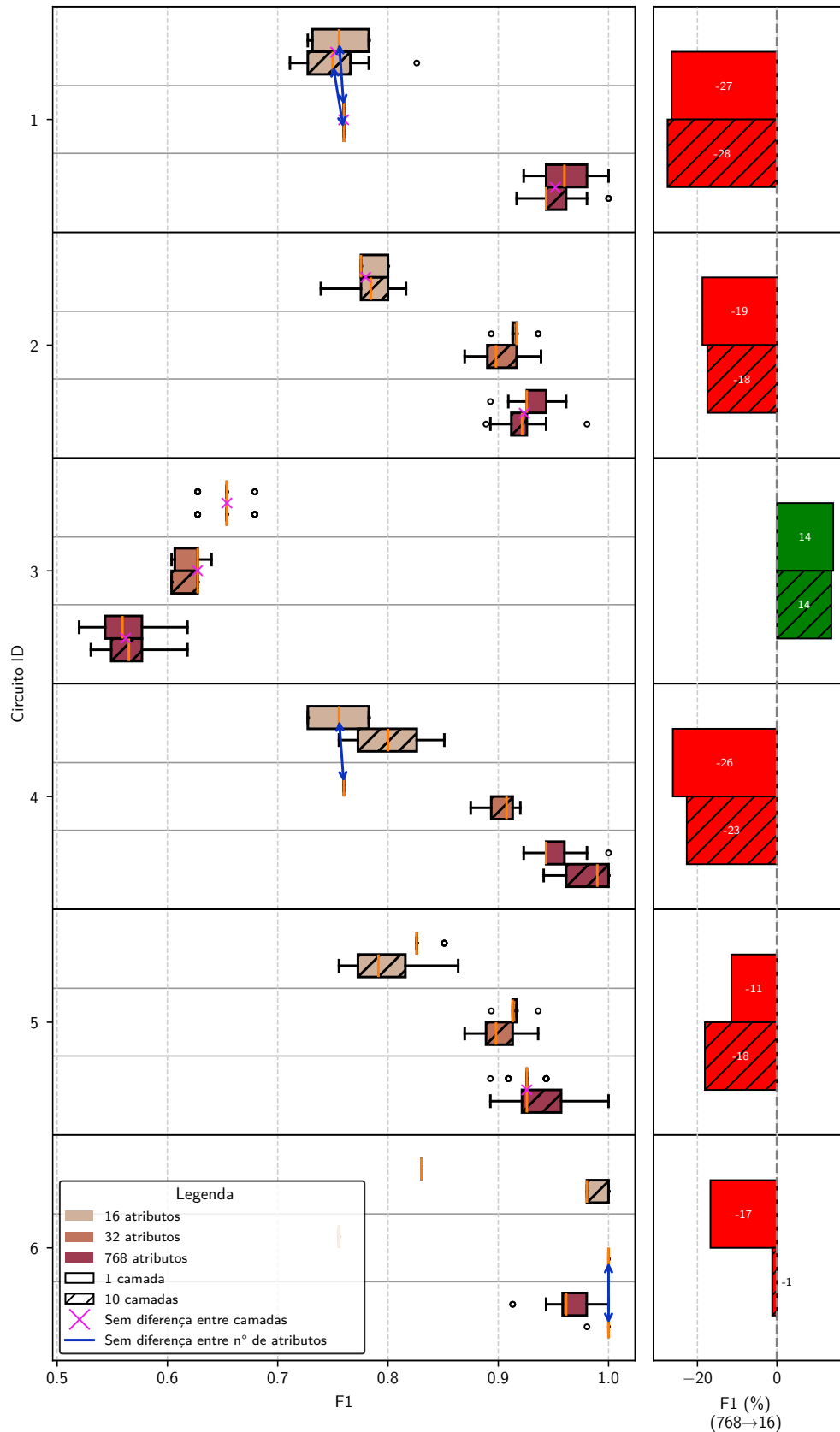
4.2.1 Resiliência dos modelos classificatórios e redução de dimensionalidade

A transição de vetores de *embeddings* de alta dimensionalidade para representações mais compactas é uma etapa de relevância prática e teórica, especialmente no contexto da computação quântica na era NISQ, onde o número de *qubits* disponíveis constitui uma limitação fundamental. A investigação de como diferentes arquiteturas, tanto clássicas quanto quânticas, respondem a essa perda de informação explícita oferece *insights* valiosos sobre a robustez, a eficiência e a capacidade de aprendizado de cada modelo. Esta subseção dedica-se a analisar comparativamente a resiliência dos classificadores propostos, interpretando a variação de seu desempenho F1 quando a dimensionalidade do vetor de entrada, gerado pelo modelo *Matrioshka*, é reduzida de 768 para 16 atributos. Tendo isso em vista, tais interpretações são oriundas das Figuras 20, 21, 22, 23 e 24 para a base *ChatGPT* - Fácil, Figuras 25, 26, 27, 28 e 29 para a base *ChatGPT* - Médio e Figuras 30, 31, 32, 33 e 34 para a base *ChatGPT* - Difícil que exibem o comportamento dos circuitos ao utilizarem diferentes tamanhos de vetores de entrada.

Uma das observações mais diretas que emergem da análise é a notável robustez exibida pelos principais modelos clássicos. Classificadores como a Regressão Logística (modelo 32) e as SVM com *kernels* linear (modelo 36) e RBF (modelo 38) demonstram uma resiliência à redução de dimensionalidade nas bases sintéticas. Na base *ChatGPT* Difícil, por exemplo, o SVM com *kernel* RBF apresentou uma queda de desempenho de apenas 2% ao transitar de 768 para 16 atributos, passando de uma mediana de F1 de 1.0 para 0.98. De forma similar, a Regressão Logística, na mesma base, registrou uma redução de desempenho de aproximadamente 2%, com a mediana caindo de 1.0 para 0.98.

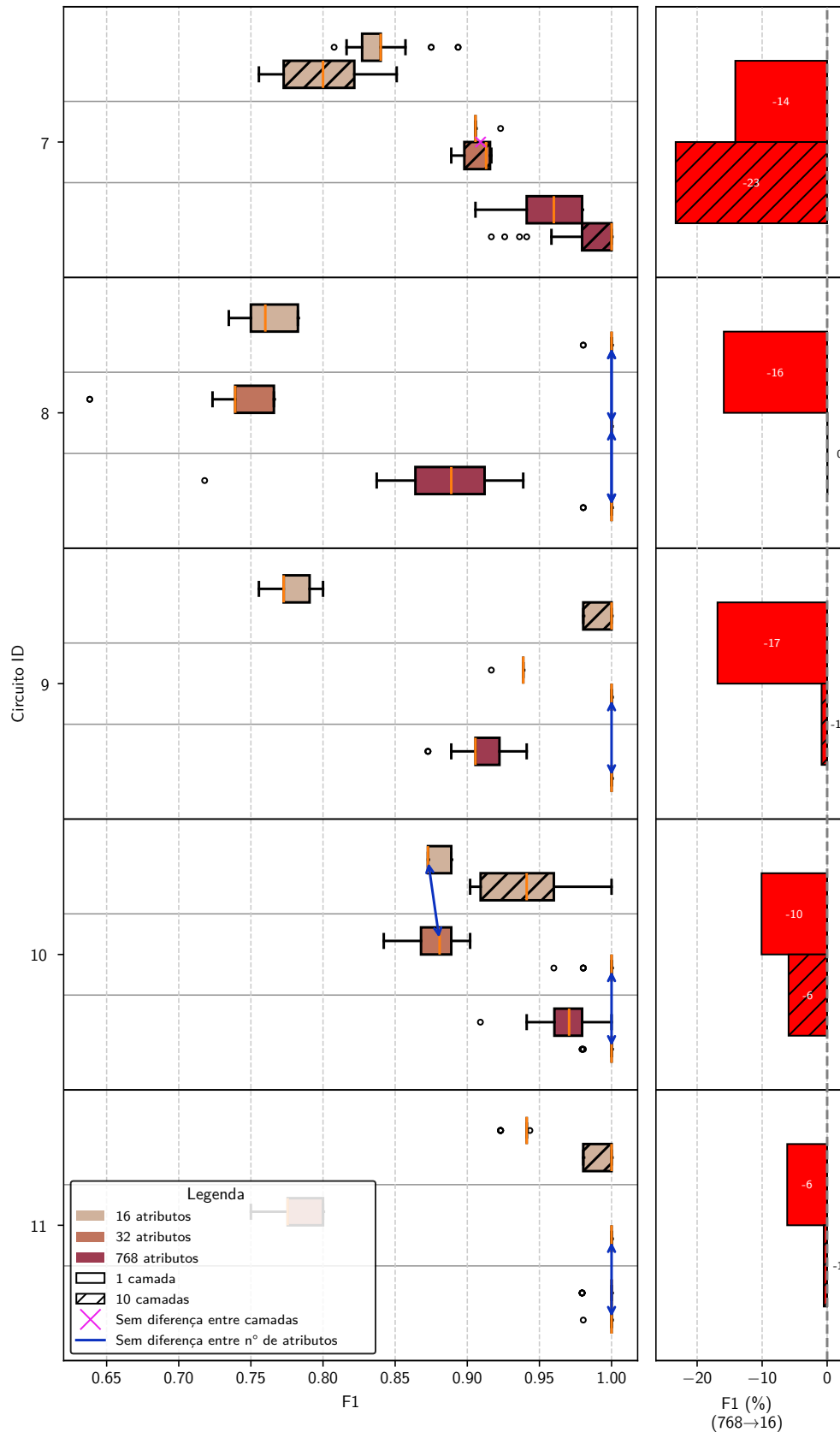
Este comportamento, consistente também nas bases *ChatGPT* Fácil e Médio, onde a degradação de desempenho foi frequentemente nula ou estatisticamente irrelevante, sugere uma conclusão importante sobre a natureza da representação de entrada: a informação essencial

Figura 20 – Resultados do desempenho F1 dos circuitos 1 ao 6 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Fácil. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



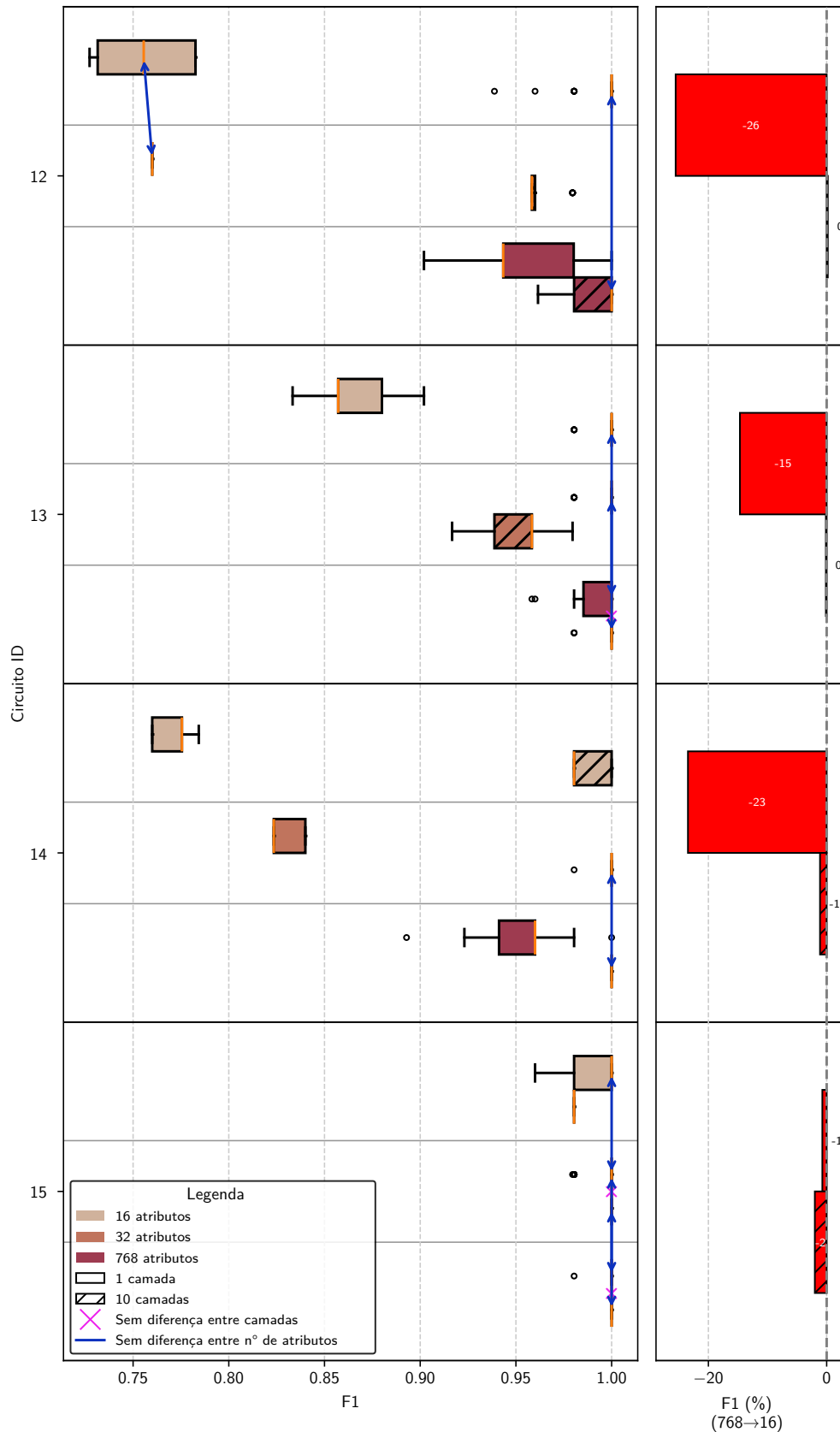
Fonte: Elaborada pelo autor (2025)

Figura 21 – Resultados do desempenho F1 dos circuitos 7 ao 11 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Fácil. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 22 – Resultados do desempenho F1 dos circuitos 12 ao 15 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Fácil. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 23 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Fácil. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

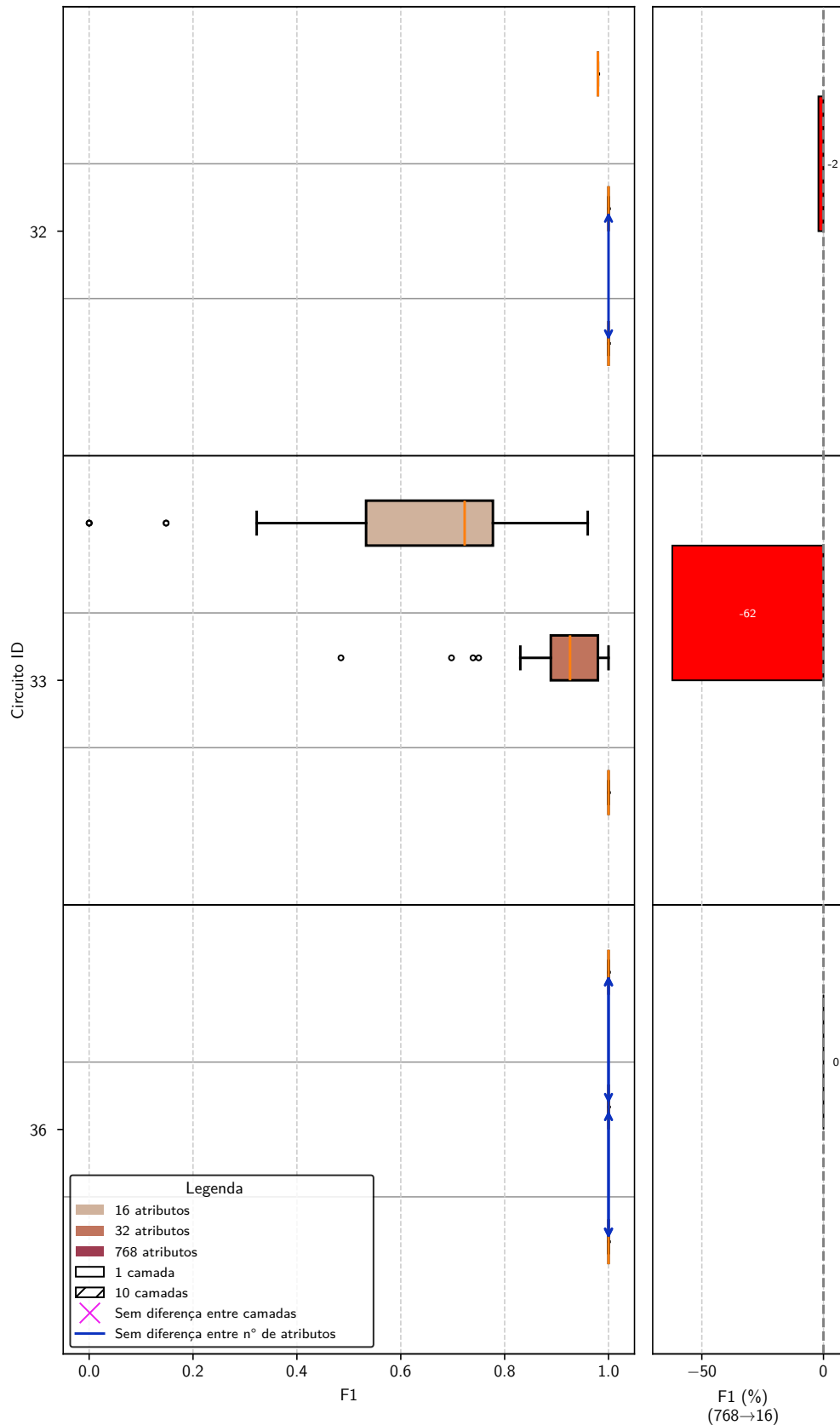
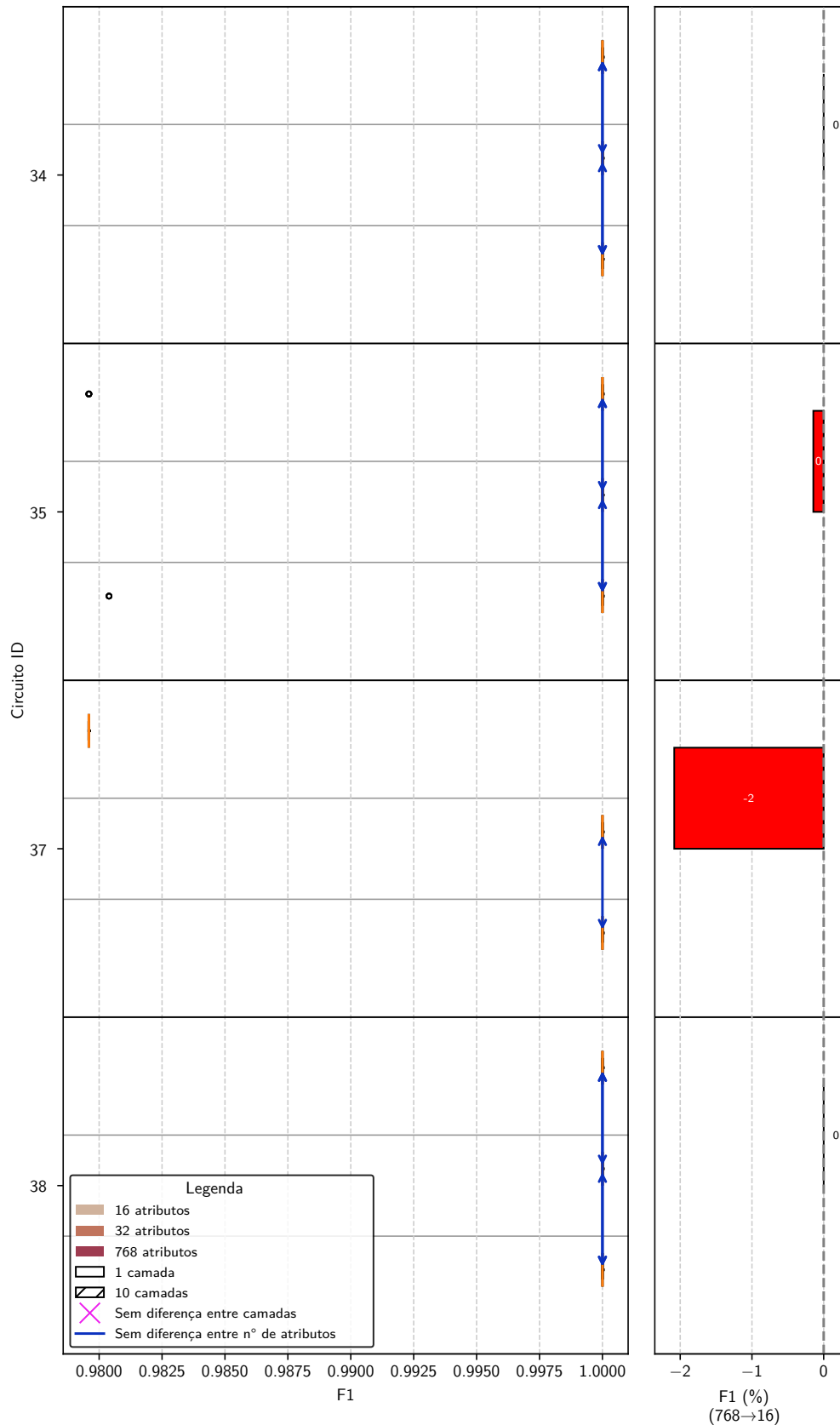
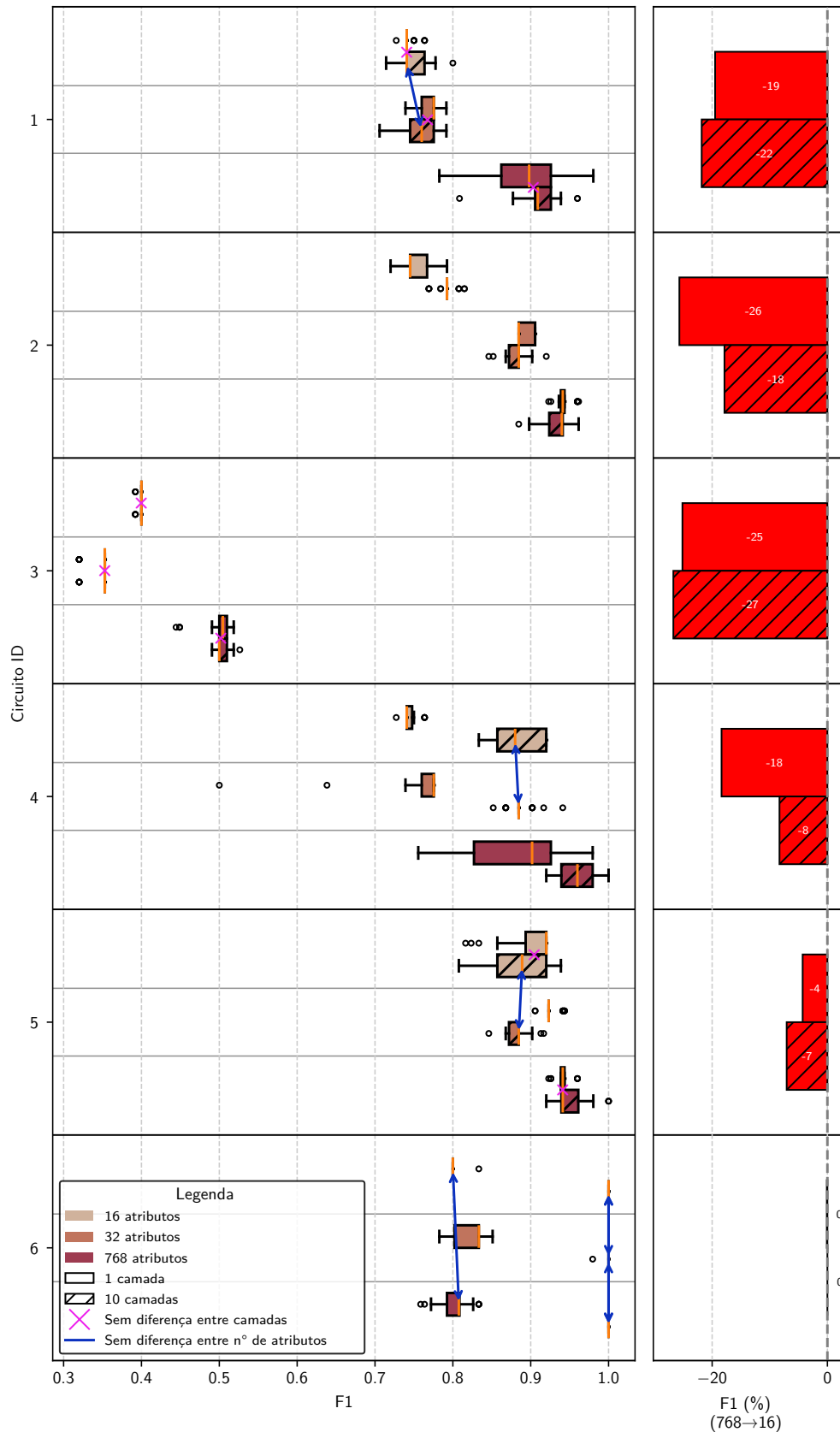


Figura 24 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Fácil. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 25 – Resultados do desempenho F1 dos circuitos 1 ao 6 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 26 – Resultados do desempenho F1 dos circuitos 7 ao 11 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

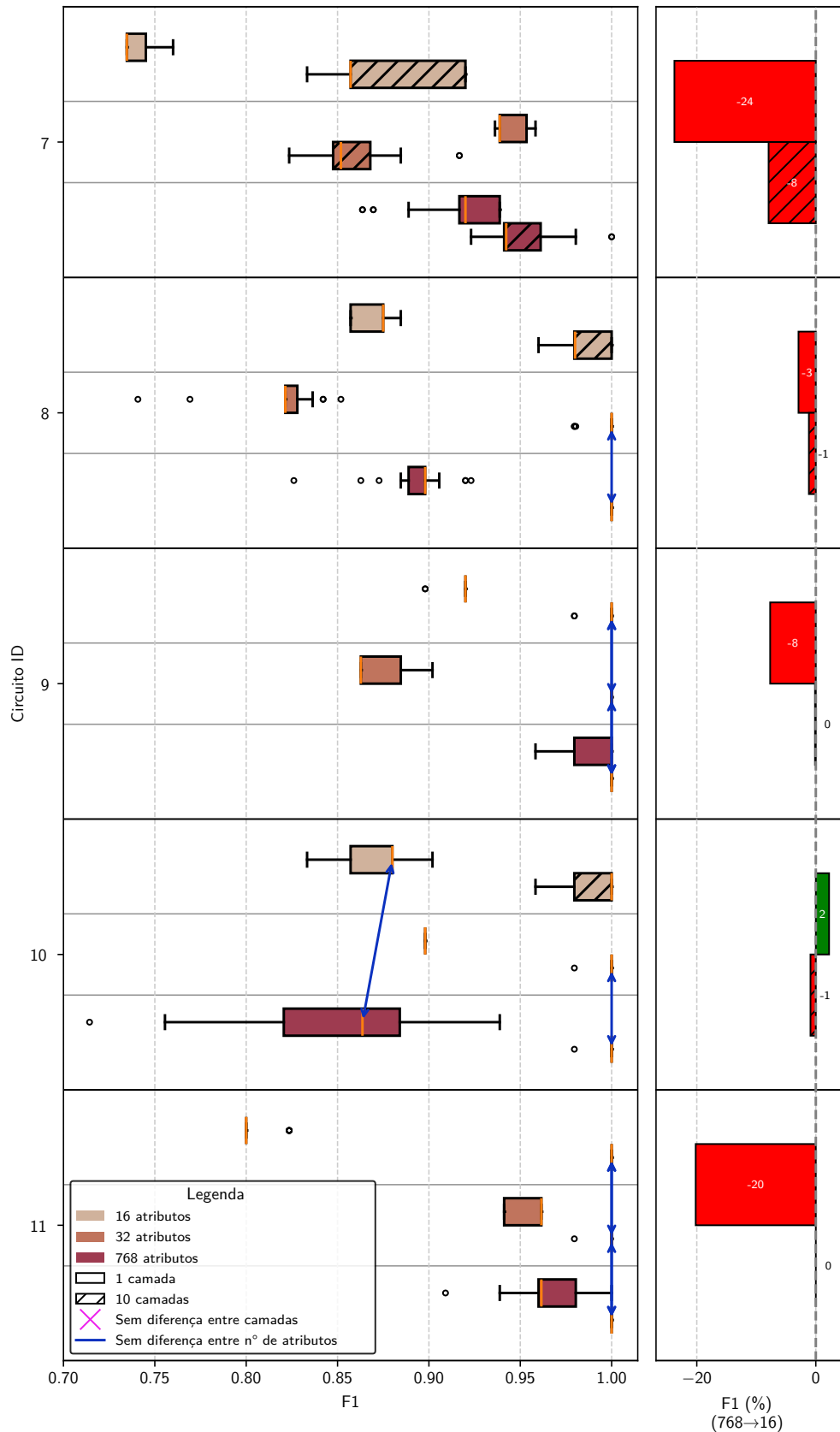
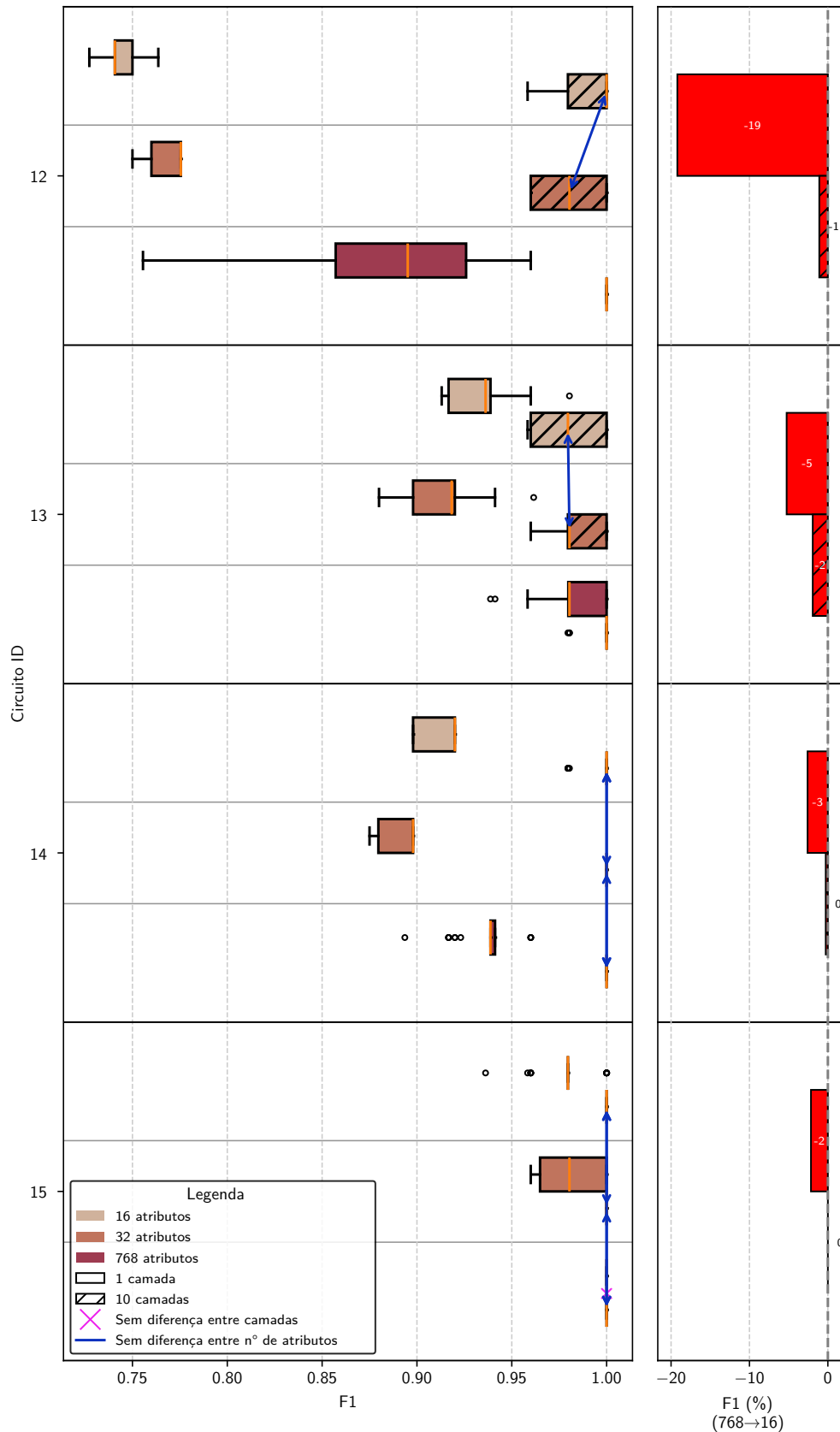


Figura 27 – Resultados do desempenho F1 dos circuitos 12 ao 15 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 28 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

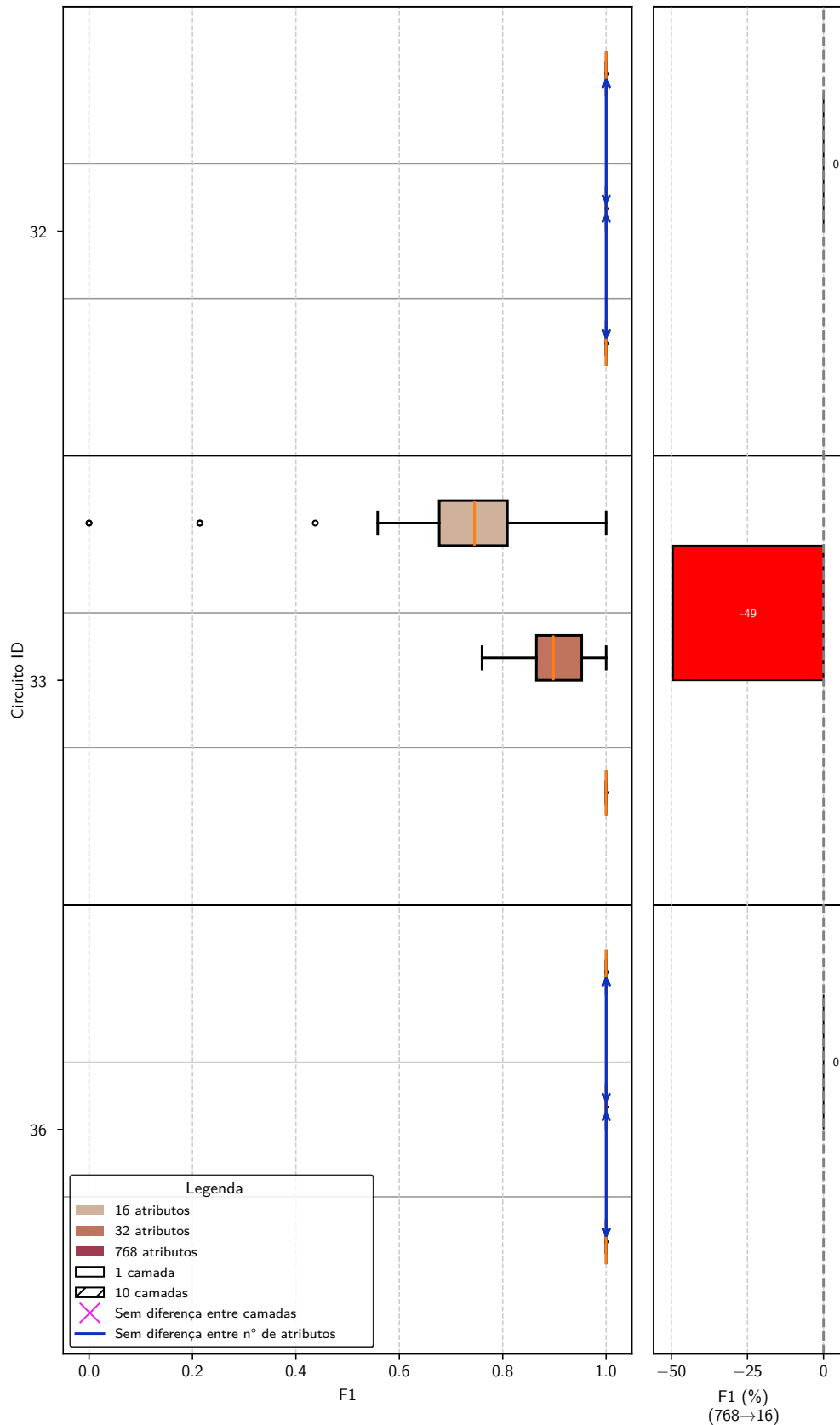
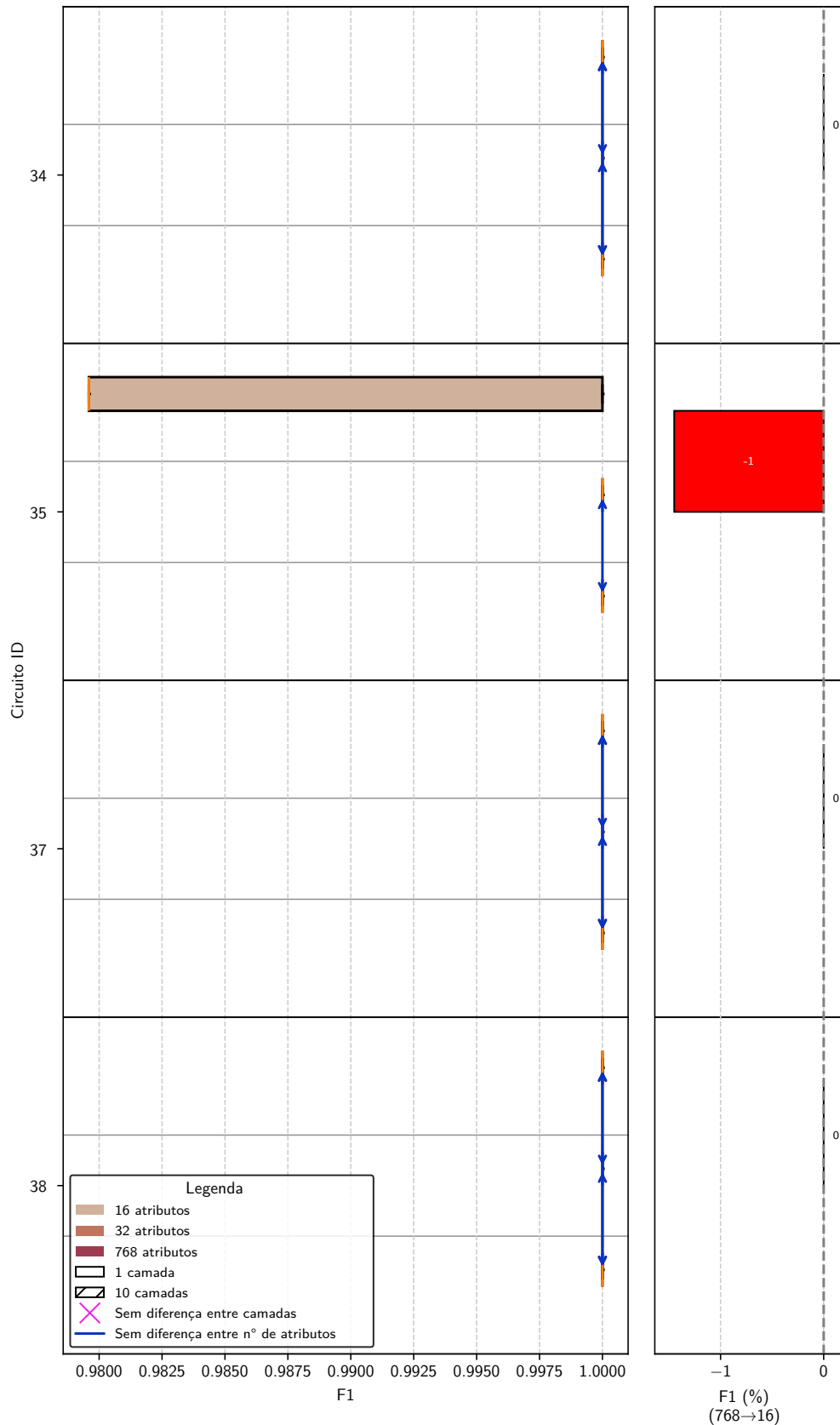
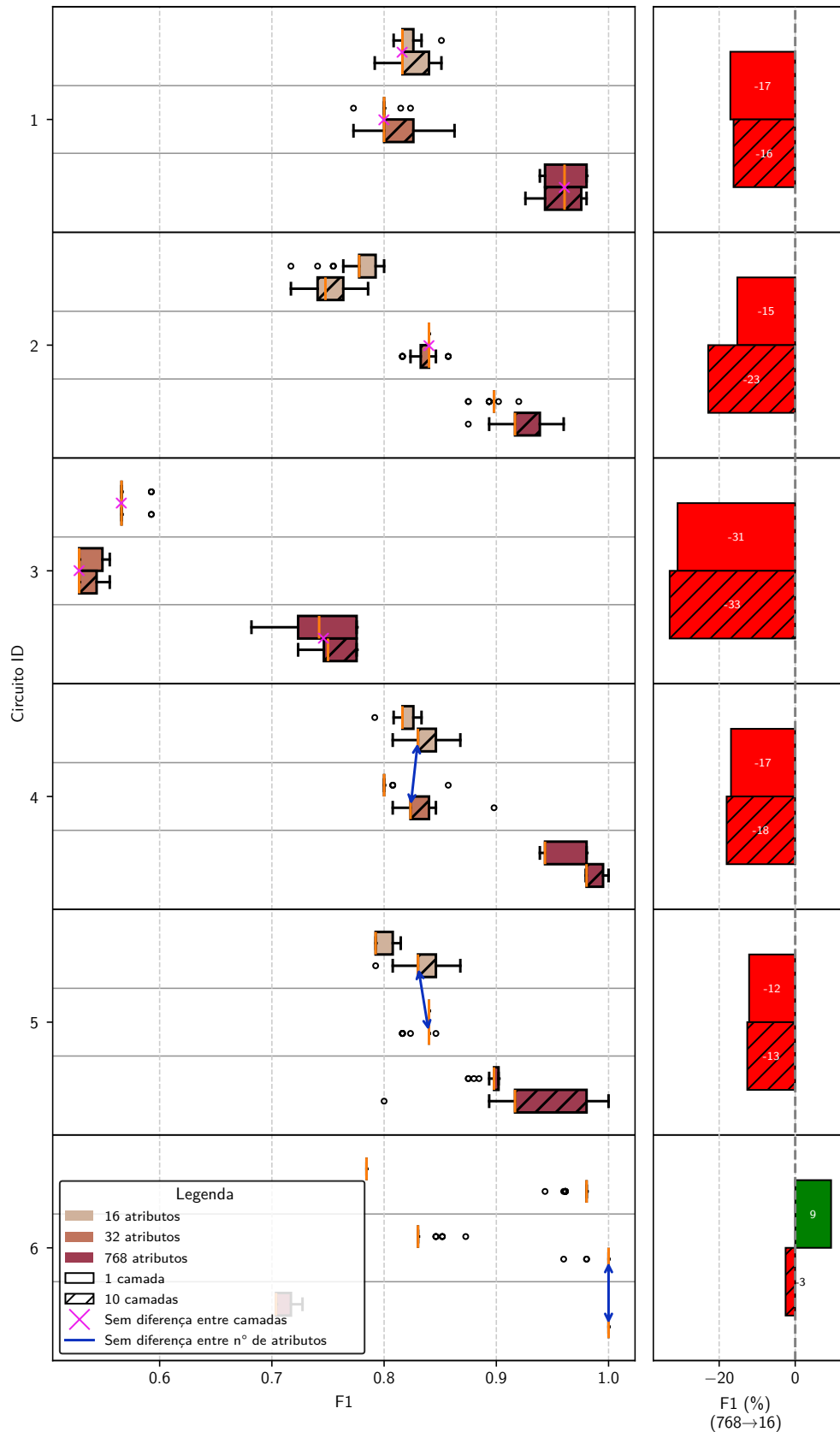


Figura 29 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Médio. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



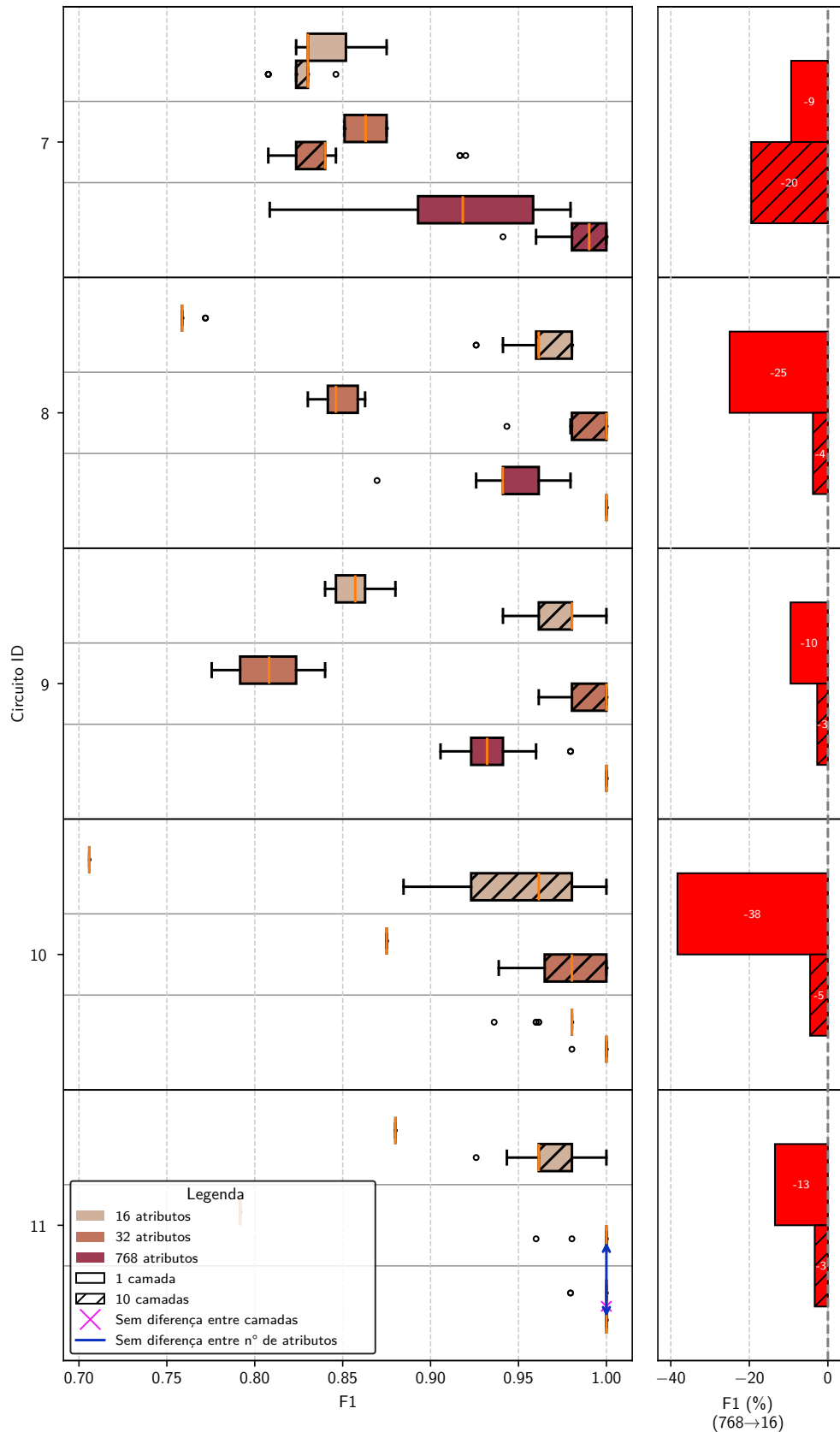
Fonte: Elaborada pelo autor (2025)

Figura 30 – Resultados do desempenho F1 dos circuitos 1 ao 6 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 31 – Resultados do desempenho F1 dos circuitos 7 ao 11 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 32 – Resultados do desempenho F1 dos circuitos 12 ao 15 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

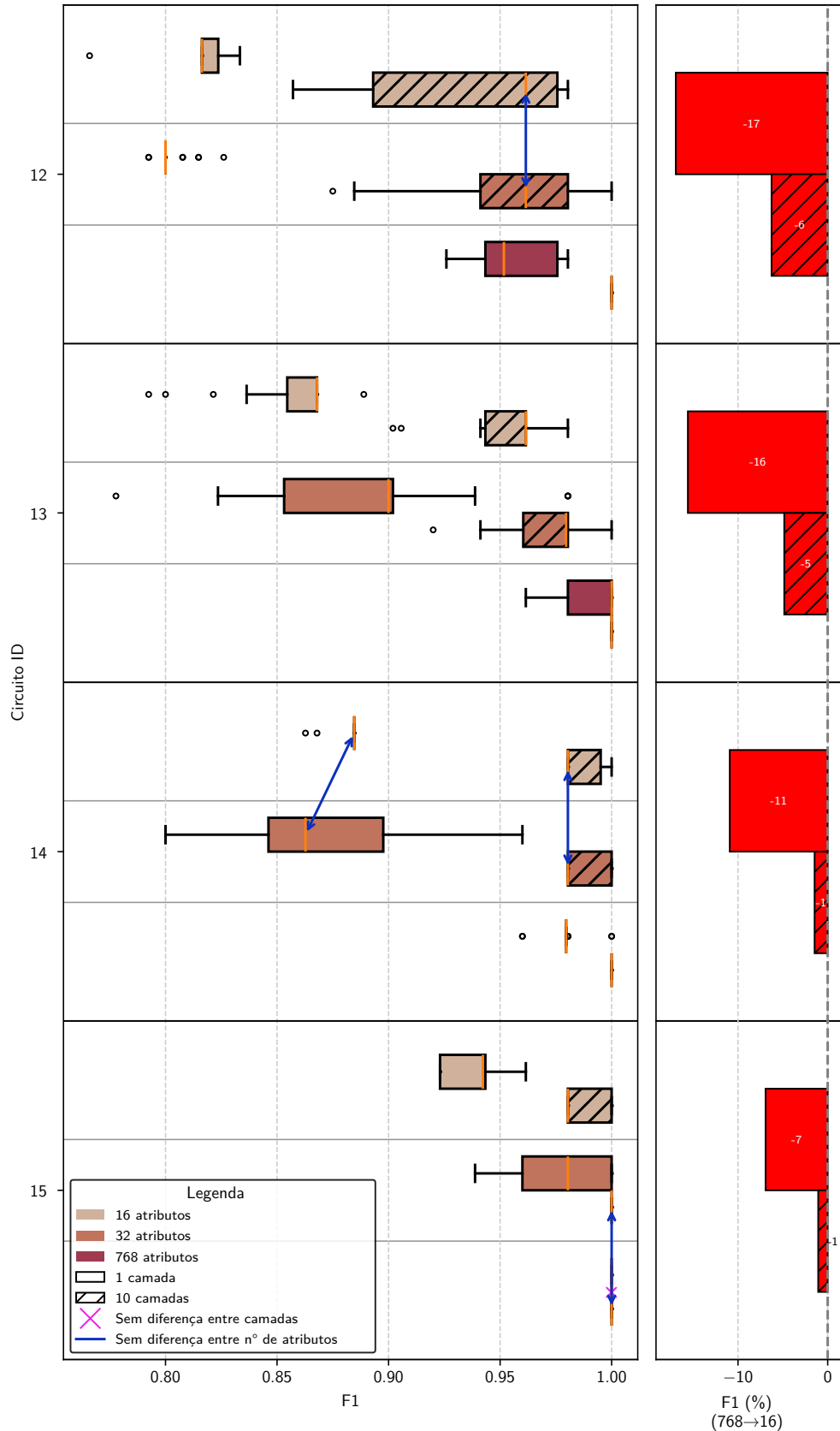
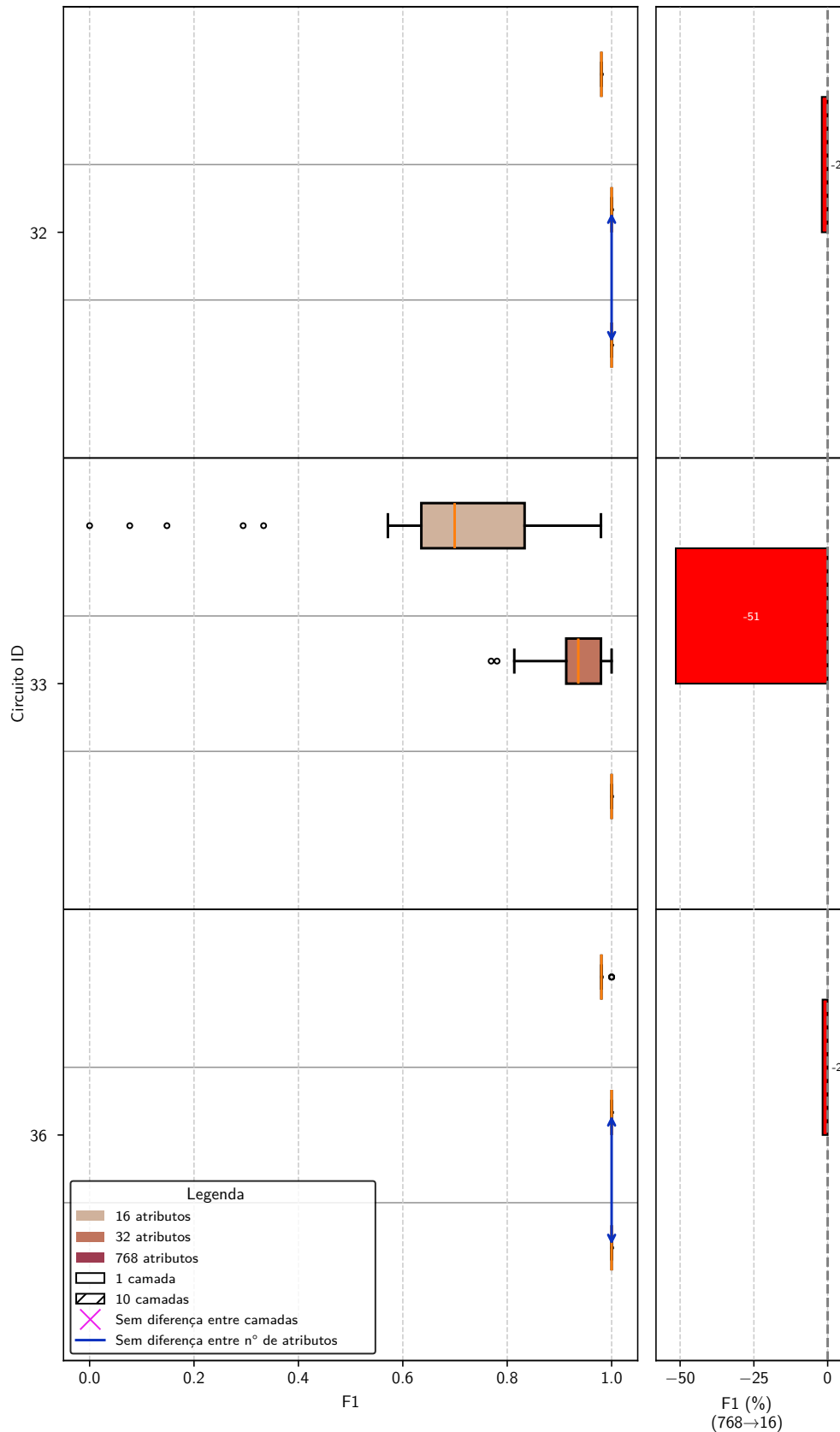
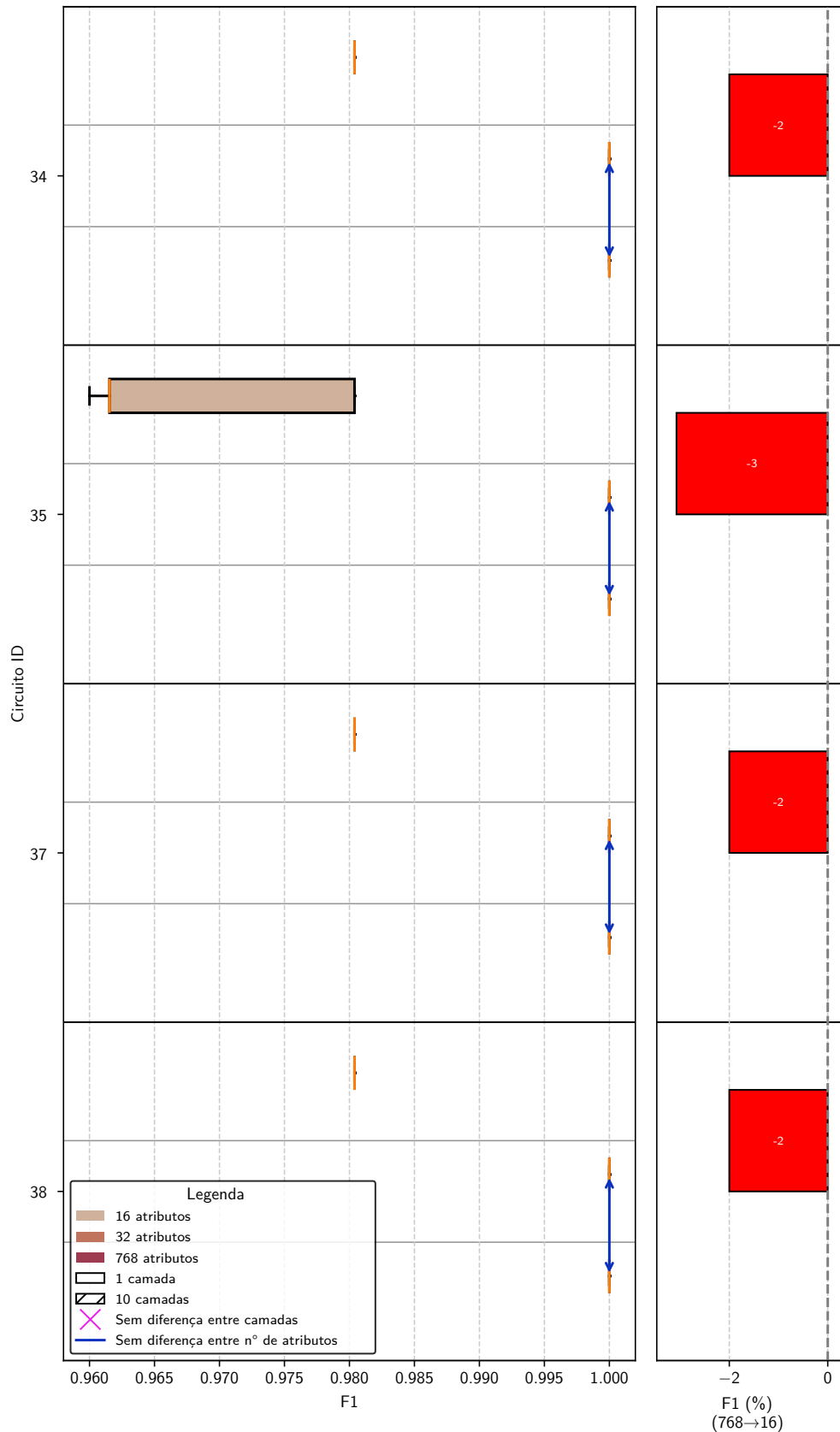


Figura 33 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 34 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com diferentes tamanhos de vetor de entrada na base *ChatGPT* - Difícil. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 ao se reduzir de 768 atributos para 16 atributos, divididos pela profundidade de cada circuito. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

para a tarefa de classificação, ao menos nos cenários sintéticos, parece estar altamente concentrada nos primeiros 16 atributos gerados pelo embedding *Matryoshka*. A eficiência dos modelos clássicos em capturar este sinal concentrado, seja por meio de uma fronteira de decisão linear ou de um mapeamento não-linear para um espaço de maior dimensão, evidencia sua adequação para este tipo de problema quando a informação classificatória é bem definida.

Em contrapartida, os circuitos quânticos não-emaranhadores demonstram uma sensibilidade consideravelmente maior à perda de informação. A análise do circuito 1, que emprega apenas portas R_X , revela uma queda de desempenho relevante em todas as bases sintéticas. Na base *ChatGPT* Fácil, a redução de dimensionalidade de 768 para 16 atributos resultou em uma diminuição da mediana de F1 de 0.96 para 0.76 ($L = 1$), o que representa uma perda de desempenho de aproximadamente 27%. De forma análoga, para o circuito 2, baseado em portas R_Y , a mesma transição na base *ChatGPT* Médio provocou uma redução de cerca de 26% em seu desempenho, com a mediana de F1 caindo de 0.94 para 0.75.

Este comportamento indica que as arquiteturas quânticas mais simples, desprovidas de capacidade de emaranhamento intrínseco, são fortemente dependentes da informação explícita contida nos vetores de alta dimensão. Sua limitada capacidade representacional, que se restringe a operações locais em cada *qubit*, mostra-se insuficiente para aprender as correlações implícitas entre os atributos restantes e, assim, compensar a informação descartada. O aumento da profundidade para $L = 10$ nestes circuitos não foi capaz de mitigar de forma consistente essa perda, reforçando a ideia de que, para arquiteturas não-emaranhadoras, a mera adição de parâmetros não expande sua capacidade de aprendizado de forma a superar uma representação de entrada mais pobre.

Um comportamento peculiar e revelador é observado em alguns circuitos quânticos rasos, que parecem ser suscetíveis a um certo “ruído” introduzido por um excesso de atributos. O caso mais emblemático é o do circuito 6 na base *ChatGPT* Difícil, que, em sua configuração com $L = 1$, apresentou uma melhora de desempenho com a redução da dimensionalidade. Ao passar de 768 para 16 atributos, sua mediana de F1 aumentou de 0.70 para 0.78, um ganho de aproximadamente 10%. Este fenômeno contraintuitivo sugere que, para um *ansatz* com capacidade de aprendizado limitada pela baixa profundidade, um número excessivo de atributos pode complicar a paisagem de otimização, introduzindo ruído que o modelo não consegue filtrar eficazmente. A apresentação de um sinal mais conciso, com 16 atributos, parece ter facilitado a convergência para uma solução de melhor qualidade, evidenciando uma interação sutil entre a capacidade do modelo e a complexidade da representação dos dados.

A profundidade, contudo, emerge como um poderoso mecanismo de mitigação da perda de informação, especialmente nos circuitos com capacidade de emaranhamento. A análise comparativa entre as configurações com $L = 1$ e $L = 10$ revela que as arquiteturas quânticas profundas podem, de fato, aprender a compensar a ausência de atributos explícitos. O circuito 10, na base *ChatGPT* Difícil, ilustra este ponto de forma clara. Com $L = 1$, a redução de 768 para 16 atributos provoca uma queda de desempenho de cerca de 39% (mediana de F1 de 0.98 para 0.71). No entanto, com $L = 10$, essa mesma redução resulta em uma queda de apenas 5% (de 1.0 para 0.96).

De forma ainda mais expressiva, o circuito 8 na base *ChatGPT* Fácil, ao passar de 768 para 16 atributos, sofre uma queda de desempenho de aproximadamente 16% em sua configuração com $L = 1$ (de 0.89 para 0.76). Contudo, na configuração com $L = 10$, essa diferença é completamente eliminada, com o modelo alcançando o desempenho máxima (mediana de 1.0) em ambas as dimensionalidades. Este comportamento se repete em outros circuitos emaranhadores, como o 11 e o 14, e fornece uma forte evidência de que a profundidade adequada permite a estas arquiteturas aprender as correlações implícitas e complexas entre os atributos remanescentes.

Esta capacidade de compensação é uma característica fundamental que distingue os circuitos quânticos profundos e emaranhadores. Enquanto os circuitos quânticos rasos dependem da informação explícita presente nos atributos, os *ansatze* profundos demonstram uma capacidade de aprendizado mais abstrata. Eles não apenas processam a informação de entrada, mas são capazes de inferir relações latentes para reconstruir, de forma efetiva, o sinal classificatório que foi perdido na etapa de redução dimensional.

Em síntese, a análise da resiliência à redução de dimensionalidade revela uma clara hierarquia de capacidade entre os modelos. Os classificadores clássicos de ponta demonstram uma robustez excepcional, indicando que, para as bases sintéticas, a tarefa pode ser resolvida com um sinal concentrado em poucos atributos. Os circuitos quânticos não-emaranhadores, por sua vez, mostram-se frágeis, com seu desempenho sendo degradado pela perda de informação. Finalmente, os circuitos quânticos emaranhadores e profundos se destacam por sua capacidade única de mitigar essa perda, utilizando a profundidade como um mecanismo para aprender correlações complexas e compensar a ausência de atributos. Este achado, embora reforce a dependência do desempenho quântico da profundidade do circuito, aponta para um potencial de aprendizado mais sofisticado, sugerindo que arquiteturas quânticas com capacidade representacional suficiente podem ser particularmente adequadas para problemas onde

o sinal classificatório reside não apenas nos atributos explícitos, mas nas intrincadas relações entre eles.

4.2.2 Sensibilidade do desempenho classificatória à escolha do modelo gerador de *embeddings*

A etapa de geração de *embeddings* constitui a interface fundamental entre o domínio da linguagem natural e o espaço vetorial no qual os modelos classificatórios, tanto clássicos quanto quânticos, operam. A escolha do modelo responsável por esta transformação não é uma mera técnica de pré-processamento, mas uma decisão que pode influenciar de forma relevante a estrutura do espaço de características e, consequentemente, a capacidade dos classificadores de encontrar uma fronteira de decisão eficaz. Esta subseção se dedica a analisar a sensibilidade do desempenho dos modelos propostos à escolha entre os três geradores de *embeddings* investigados — *Matryoshka*, *Nomic* e *MPNet* —, utilizando os 16 atributos iniciais como base de comparação. Tal análise é feita se debruçando sobre as Figuras 35, 36, 37, 38 e 39 para a base *ChatGPT* - Fácil, Figuras 40, 41, 42, 43 e 44 para a base *ChatGPT* - Médio e Figuras 45, 46, 47, 48 e 49 para a base *ChatGPT* - Difícil, no qual é possível compreender o desempenho dos circuitos tendo em vista diferentes modelos geradores de *embeddings*.

Uma das observações mais consistentes, ao se analisar as arquiteturas quânticas mais simples, é sua acentuada dependência da qualidade da representação de entrada. Os circuitos não-emaranhadores, cuja capacidade representacional é inerentemente limitada a operações locais, mostram-se particularmente suscetíveis às sutilezas de cada modelo gerador de *embeddings*. O circuito 2 (R_Y), com $L = 1$, na base *ChatGPT* Médio, exemplifica esta dependência: sua mediana de F1 com o *embedding* *Matryoshka* foi de 0.75, um resultado cerca de 17% inferior ao obtido com o *MPNet*, que alcançou uma mediana de 0.90. Tal disparidade sugere que, na ausência de mecanismos quânticos para criar correlações complexas, o desempenho do classificador torna-se um reflexo direto de quão linearmente separáveis os dados são apresentados pelo *embedding*.

Esta vulnerabilidade não é um fenômeno isolado. Na base *ChatGPT* Difícil, o circuito 1 (R_X) com $L = 1$ obteve uma mediana de F1 de 0.92 com o *embedding* do modelo *Nomic*, enquanto seu desempenho com o *Matryoshka* foi de 0.82, aproximadamente 11% inferior. De forma similar, o circuito 4 (R_X, R_Z), na mesma base e com $L = 1$, também apresentou um desempenho cerca de 11% inferior ao utilizar o *Matryoshka* em comparação com o *Nomic*. Estes

Figura 35 – Resultados do desempenho F1 dos circuitos 1 ao 6 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

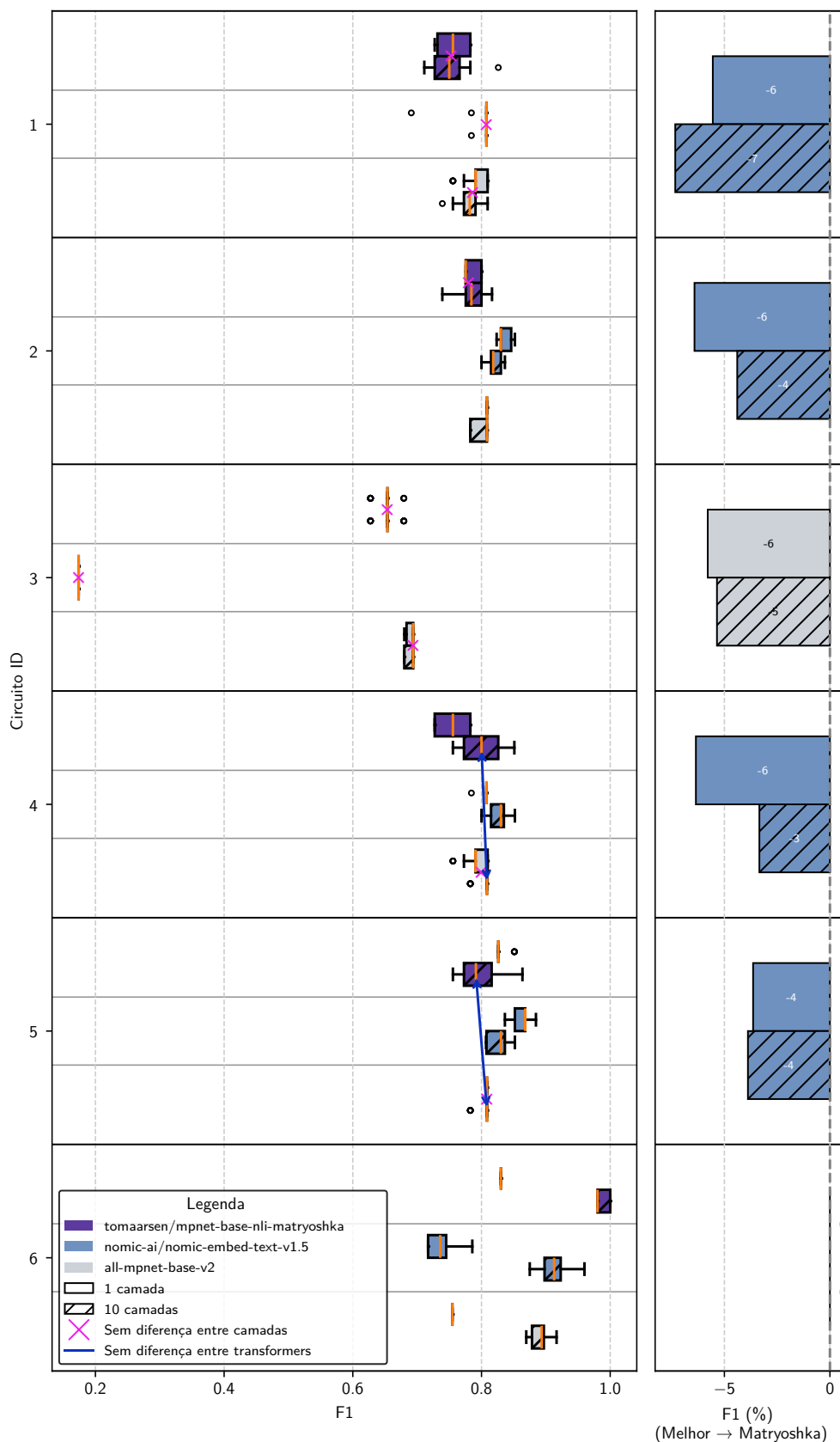


Figura 36 – Resultados do desempenho F1 dos circuitos 7 ao 11 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

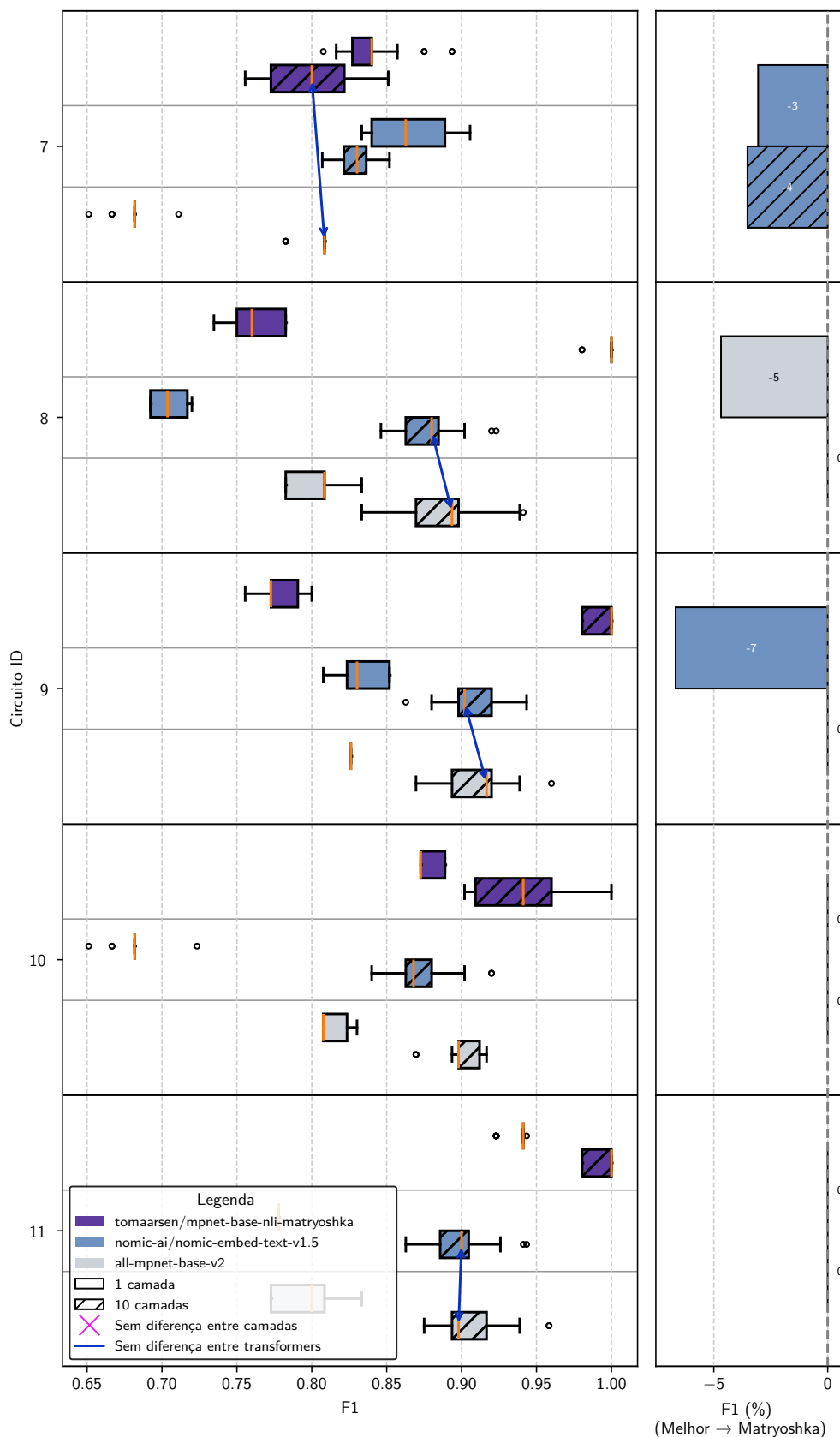


Figura 37 – Resultados do desempenho F1 dos circuitos 12 ao 15 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

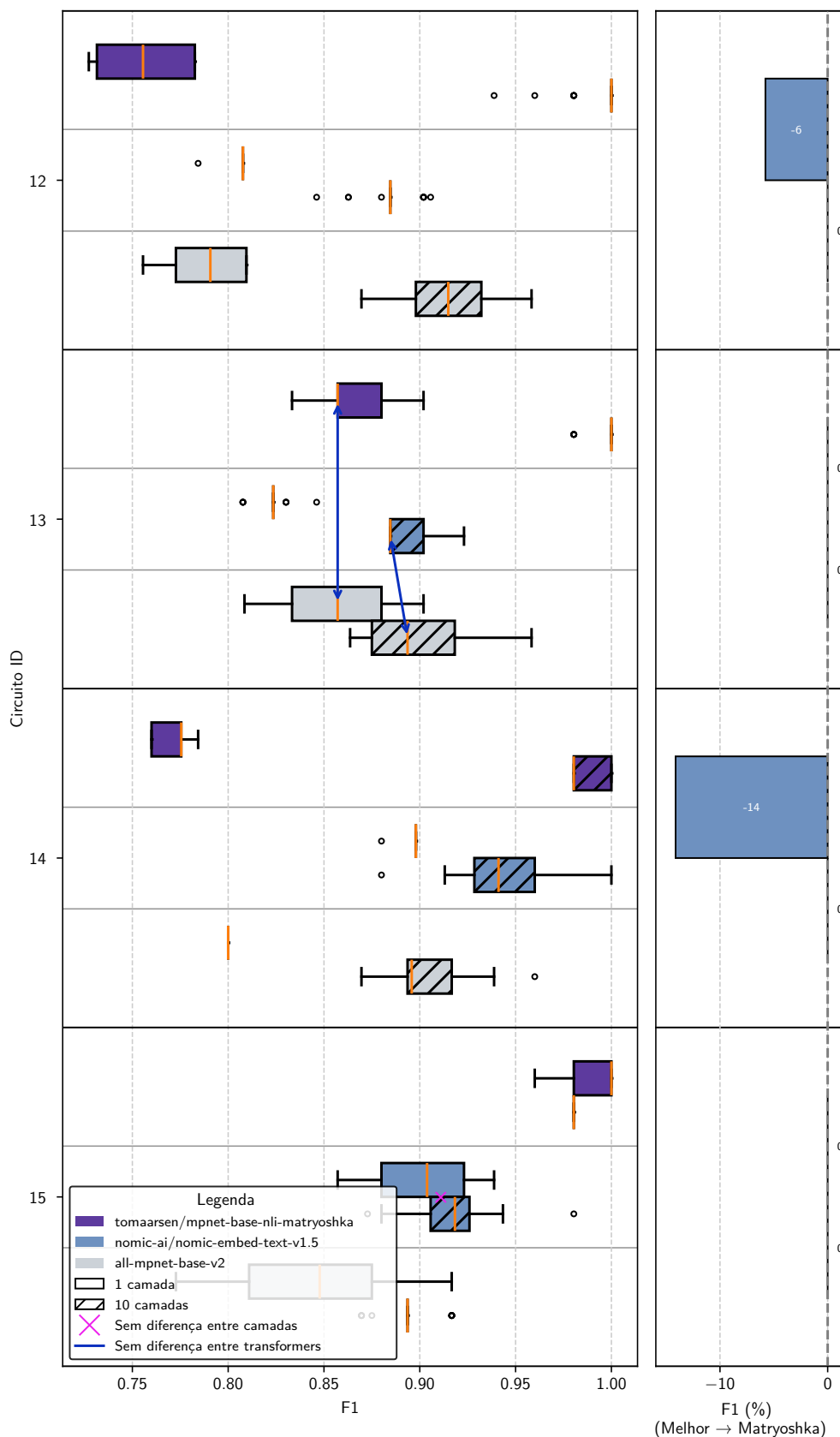


Figura 38 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

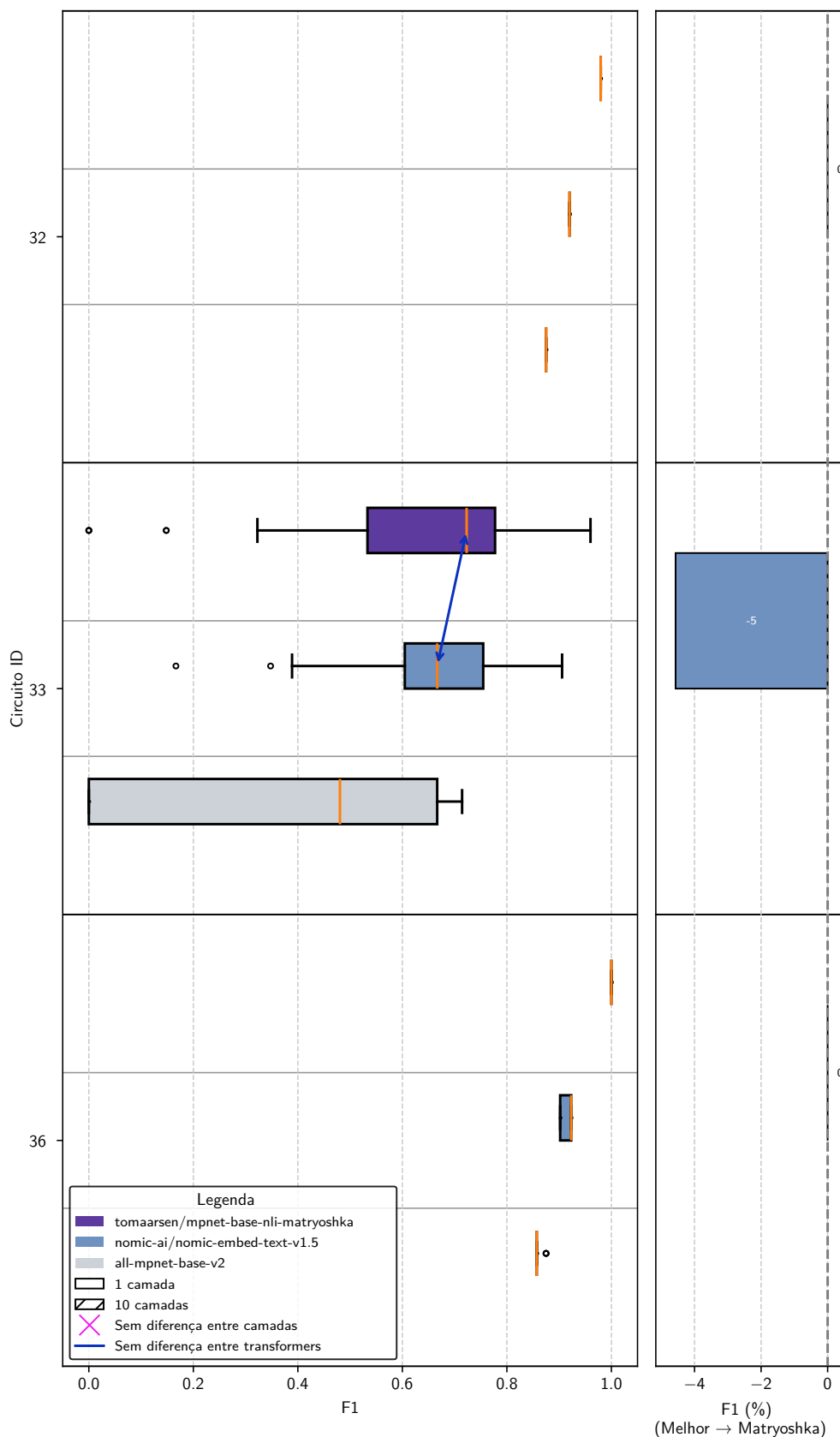


Figura 39 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com os modelos geradores de *embeddings* na base *ChatGPT* - Fácil. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

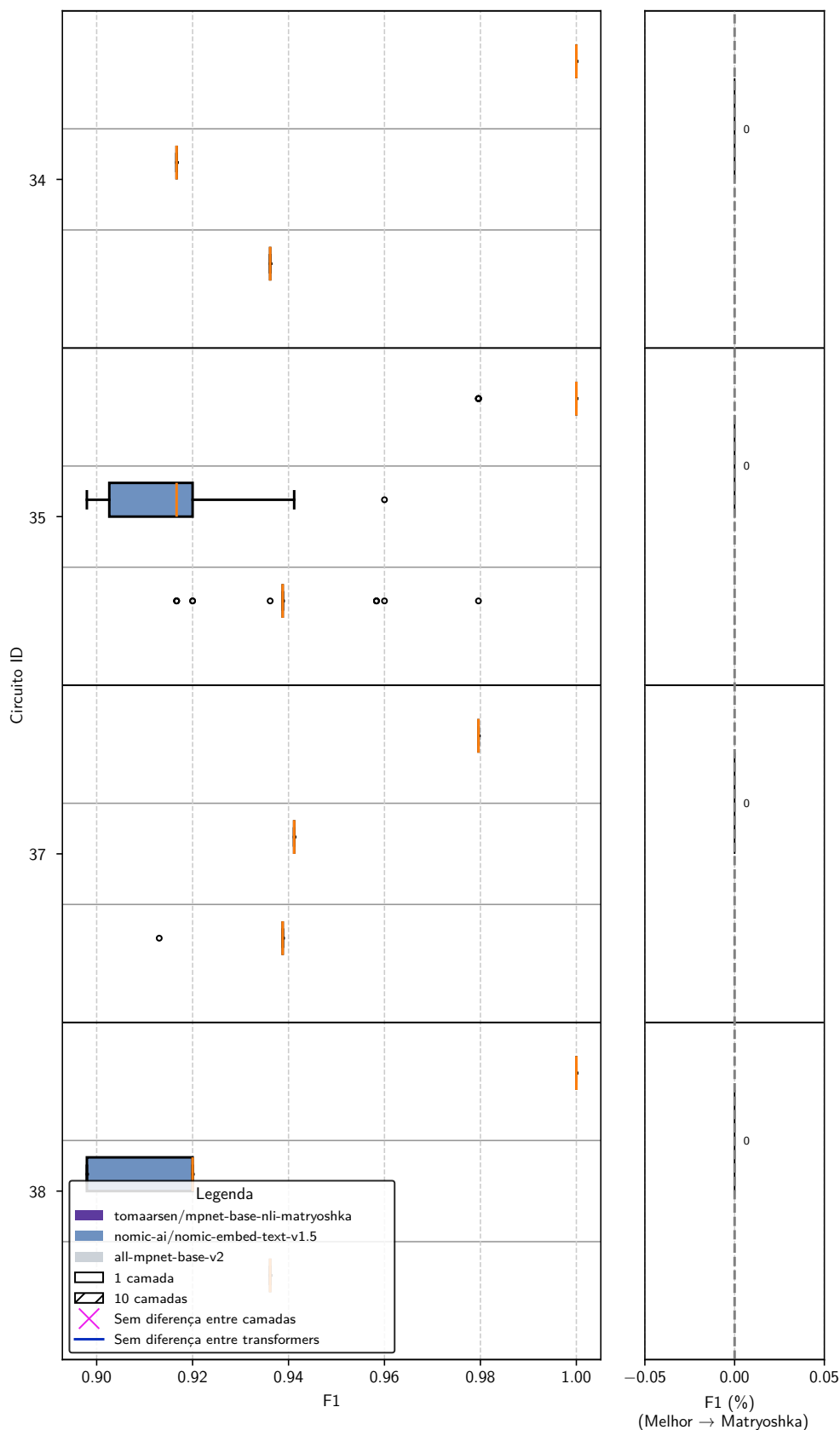
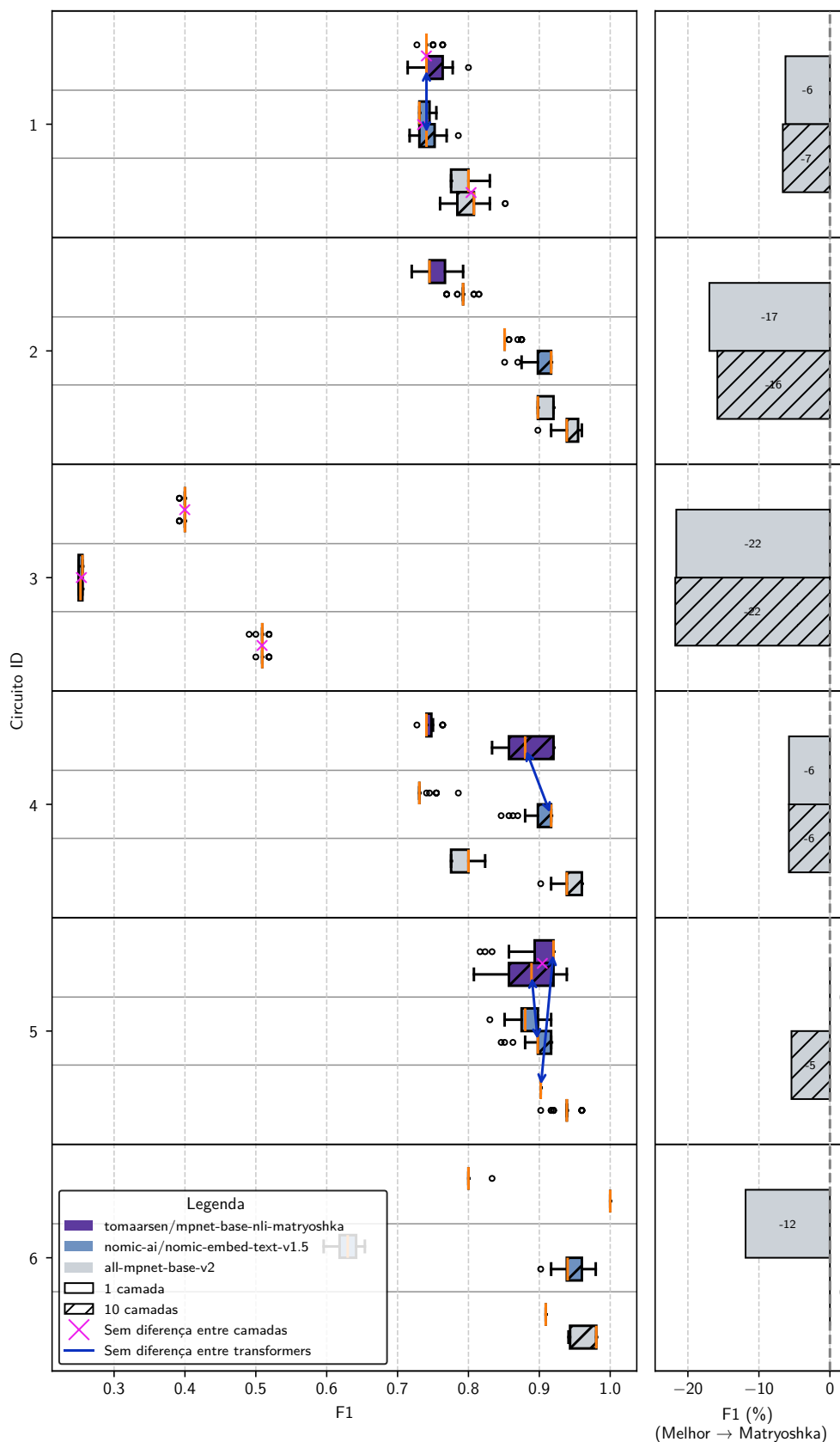


Figura 40 – Resultados do desempenho F1 dos circuitos 1 ao 6 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



Fonte: Elaborada pelo autor (2025)

Figura 41 – Resultados do desempenho F1 dos circuitos 7 ao 11 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

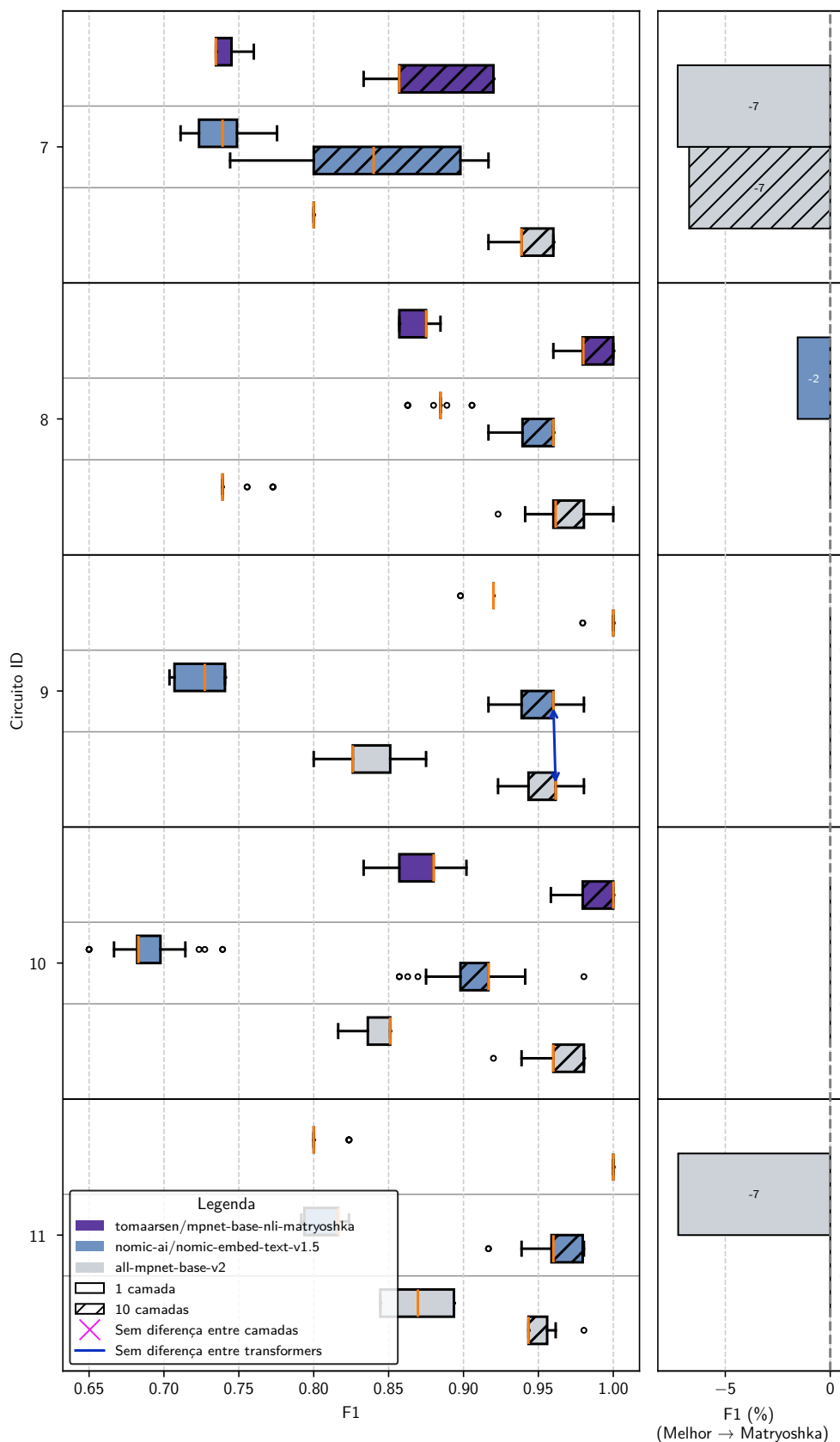


Figura 42 – Resultados do desempenho F1 dos circuitos 12 ao 15 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

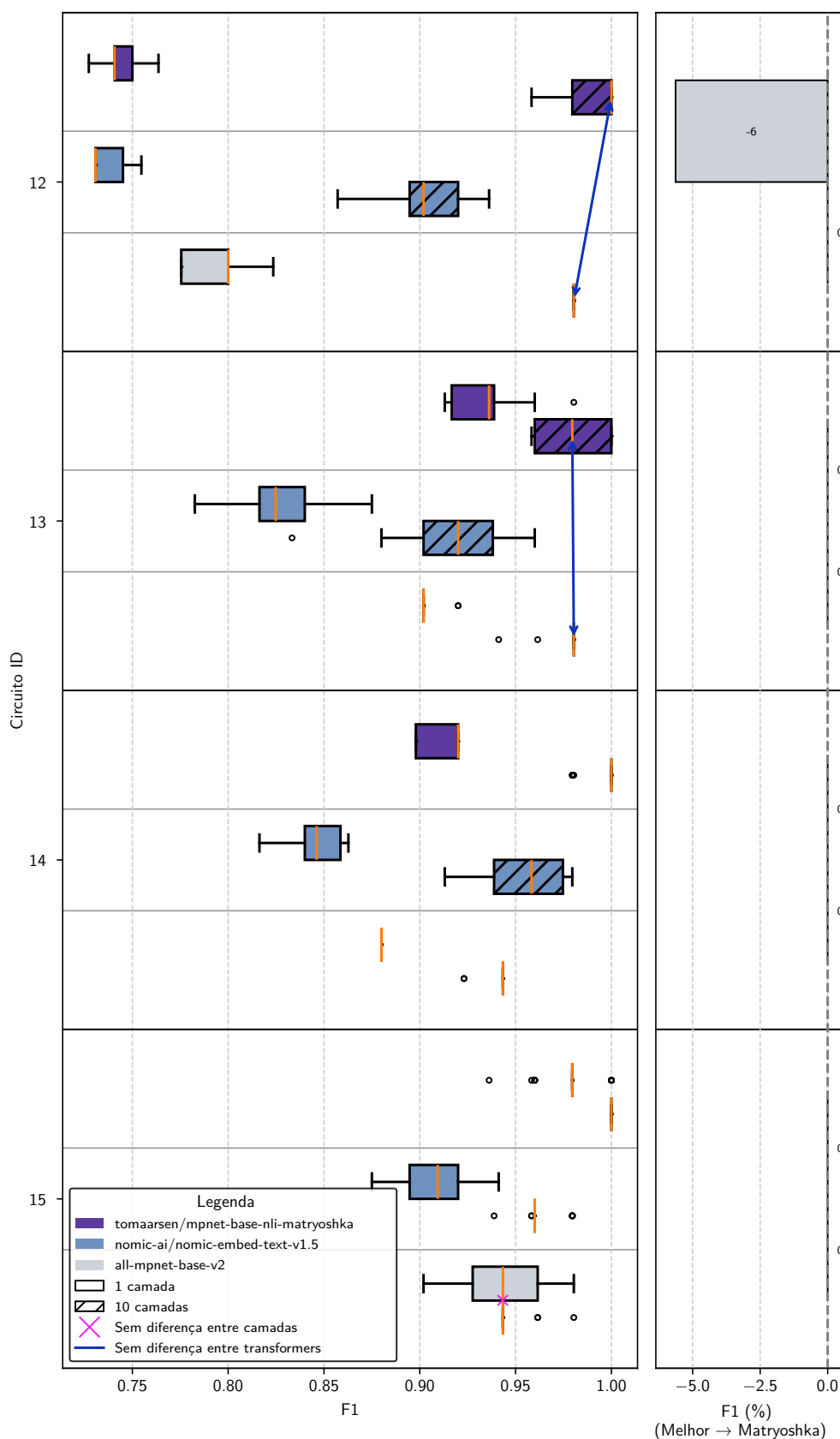


Figura 43 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

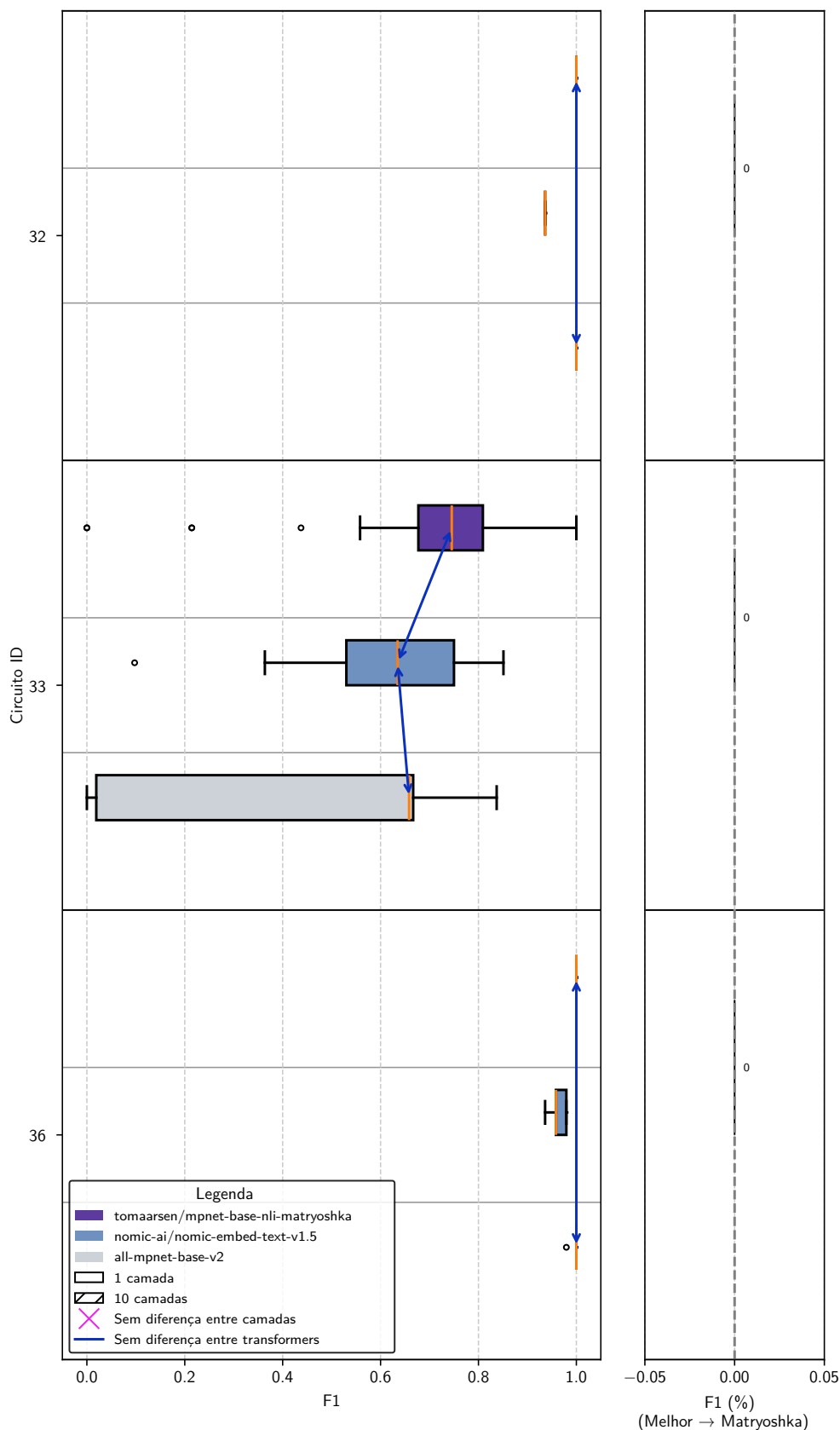


Figura 44 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com os modelos geradores de *embeddings* na base *ChatGPT* - Médio. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

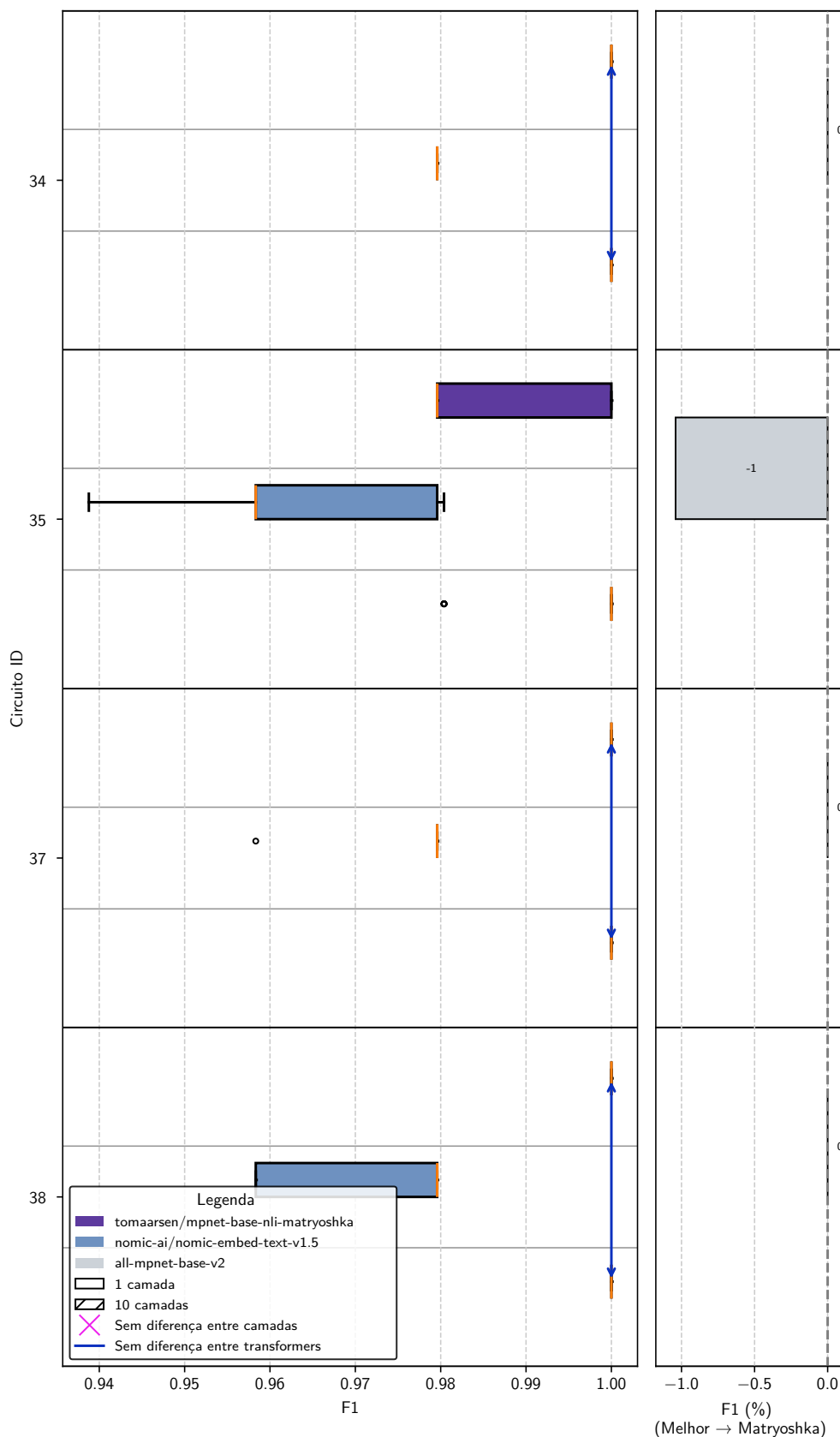


Figura 45 – Resultados do desempenho F1 dos circuitos 1 ao 6 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

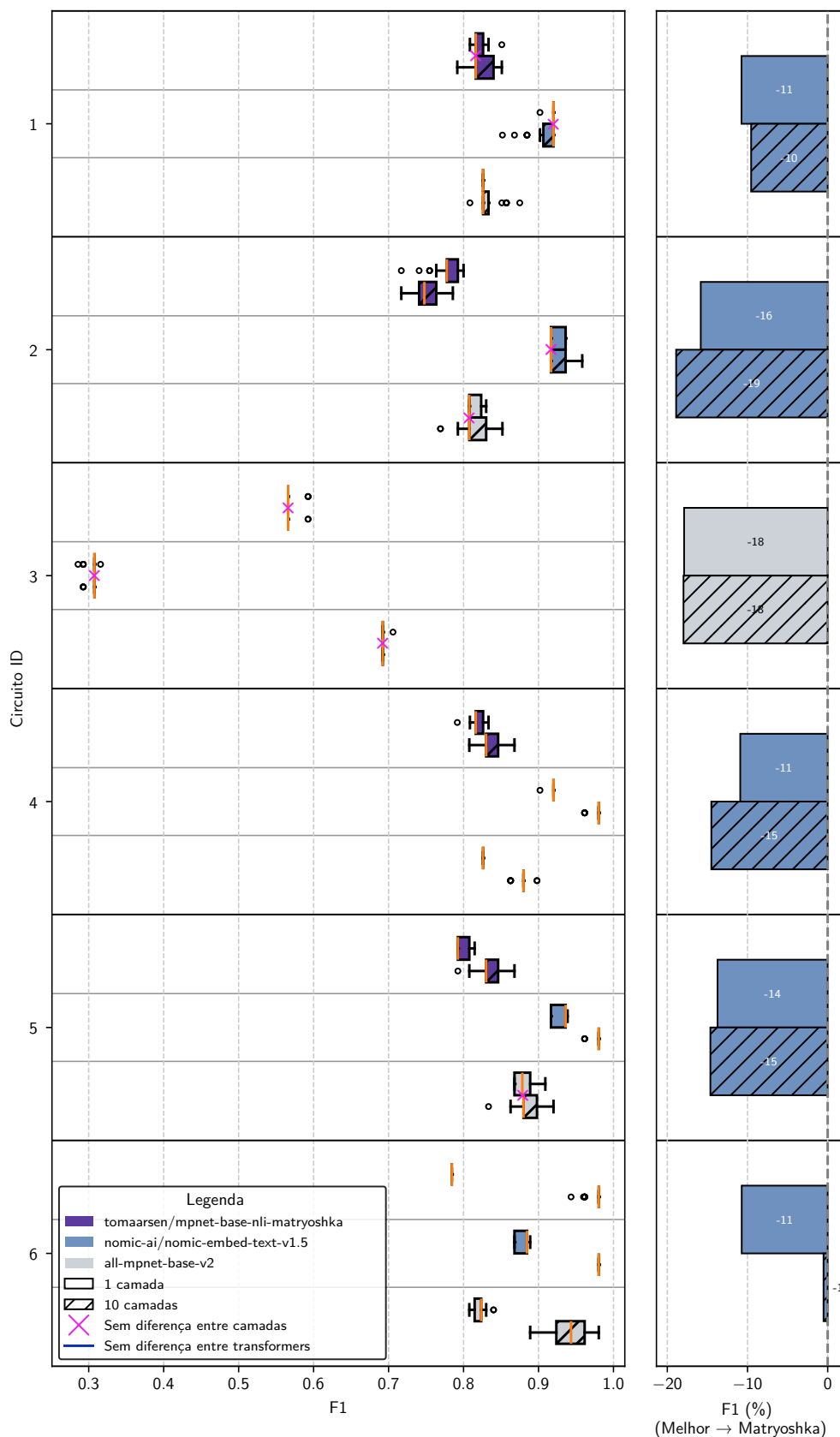


Figura 46 – Resultados do desempenho F1 dos circuitos 7 ao 11 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos implementam diferentes formas de emaranhamento quântico. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

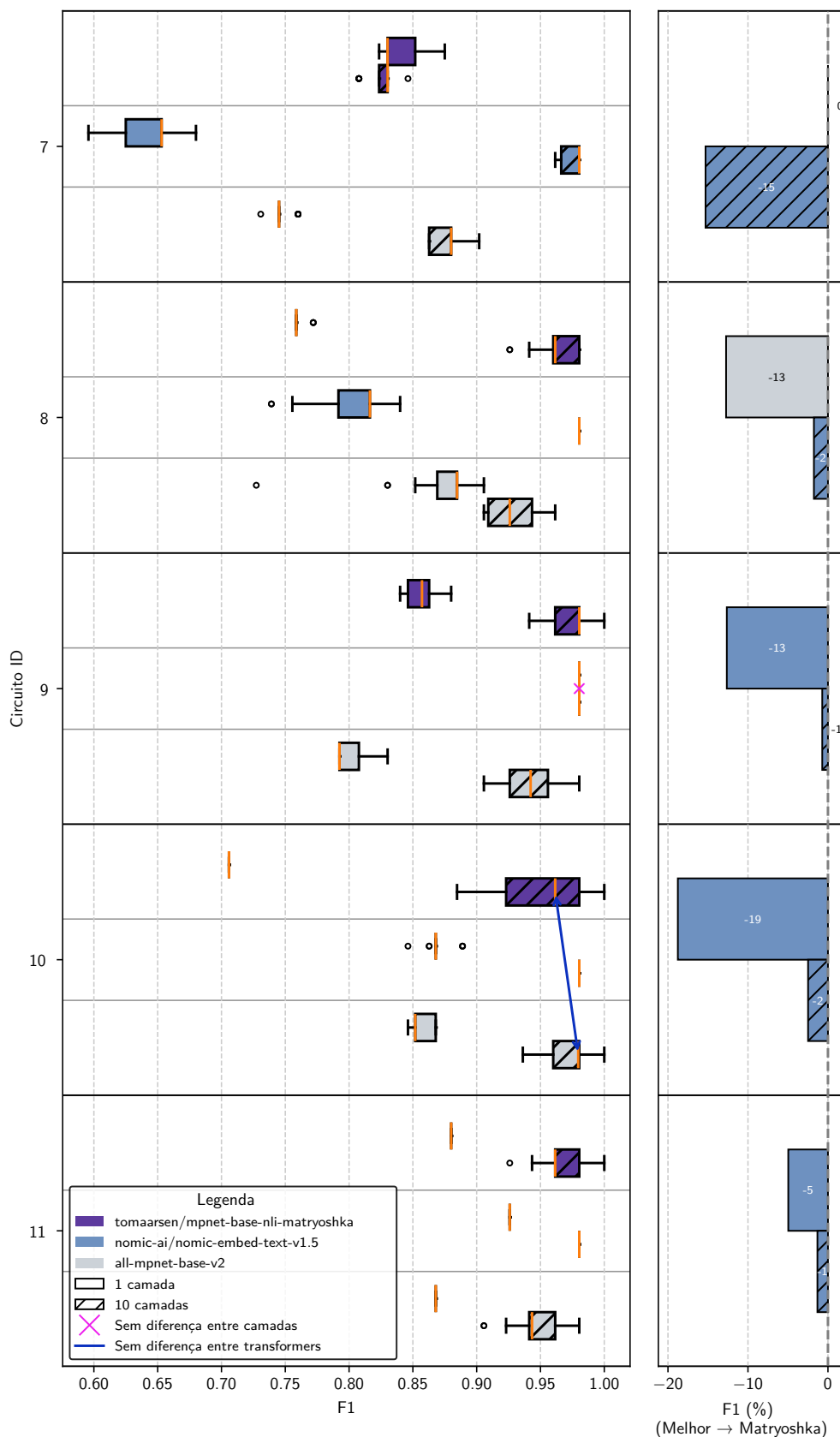


Figura 47 – Resultados do desempenho F1 dos circuitos 12 ao 15 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos implementam diferentes tipos de conectividade entre *qubits*. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

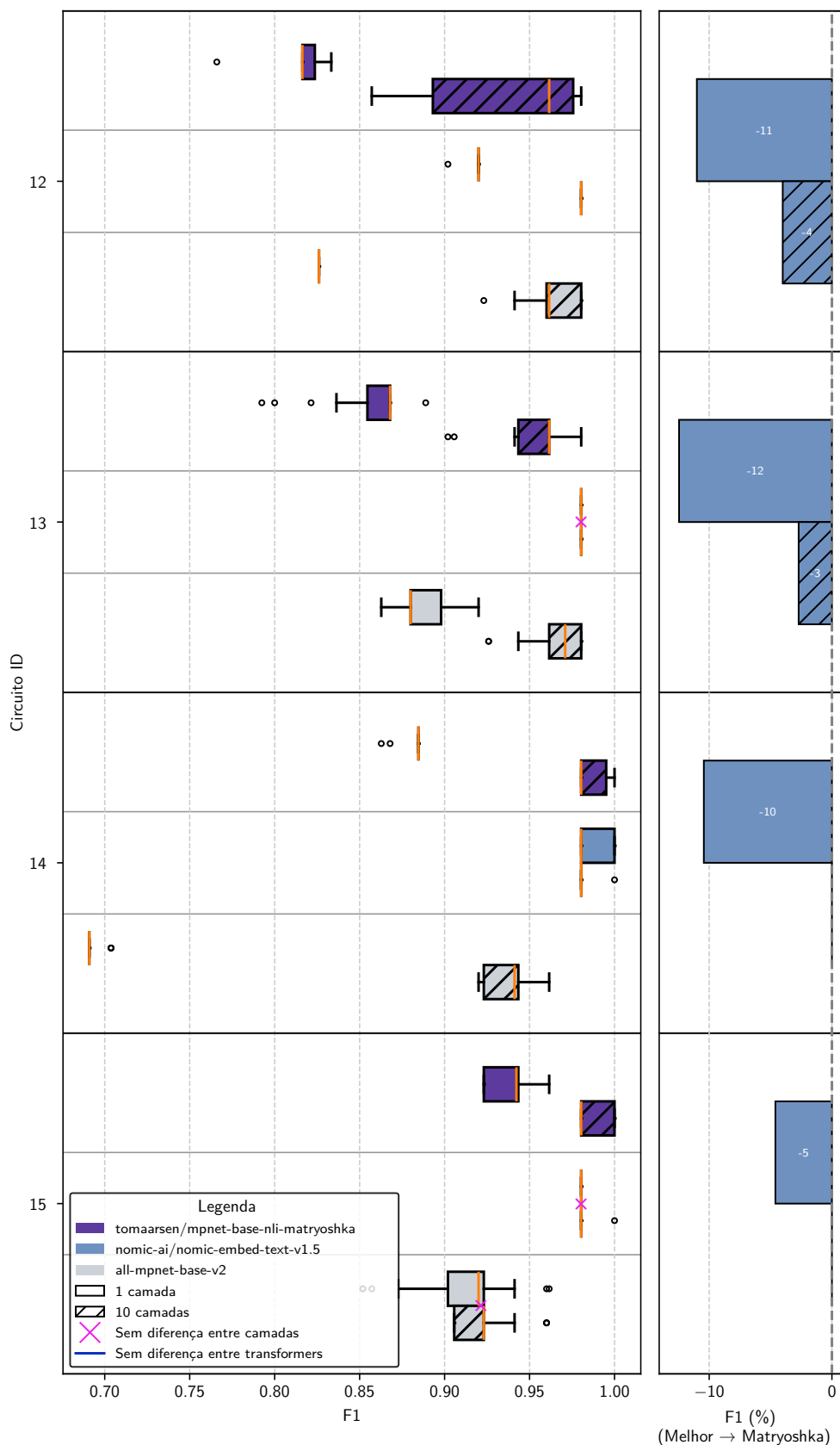


Figura 48 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos implementam modelos clássicos lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.

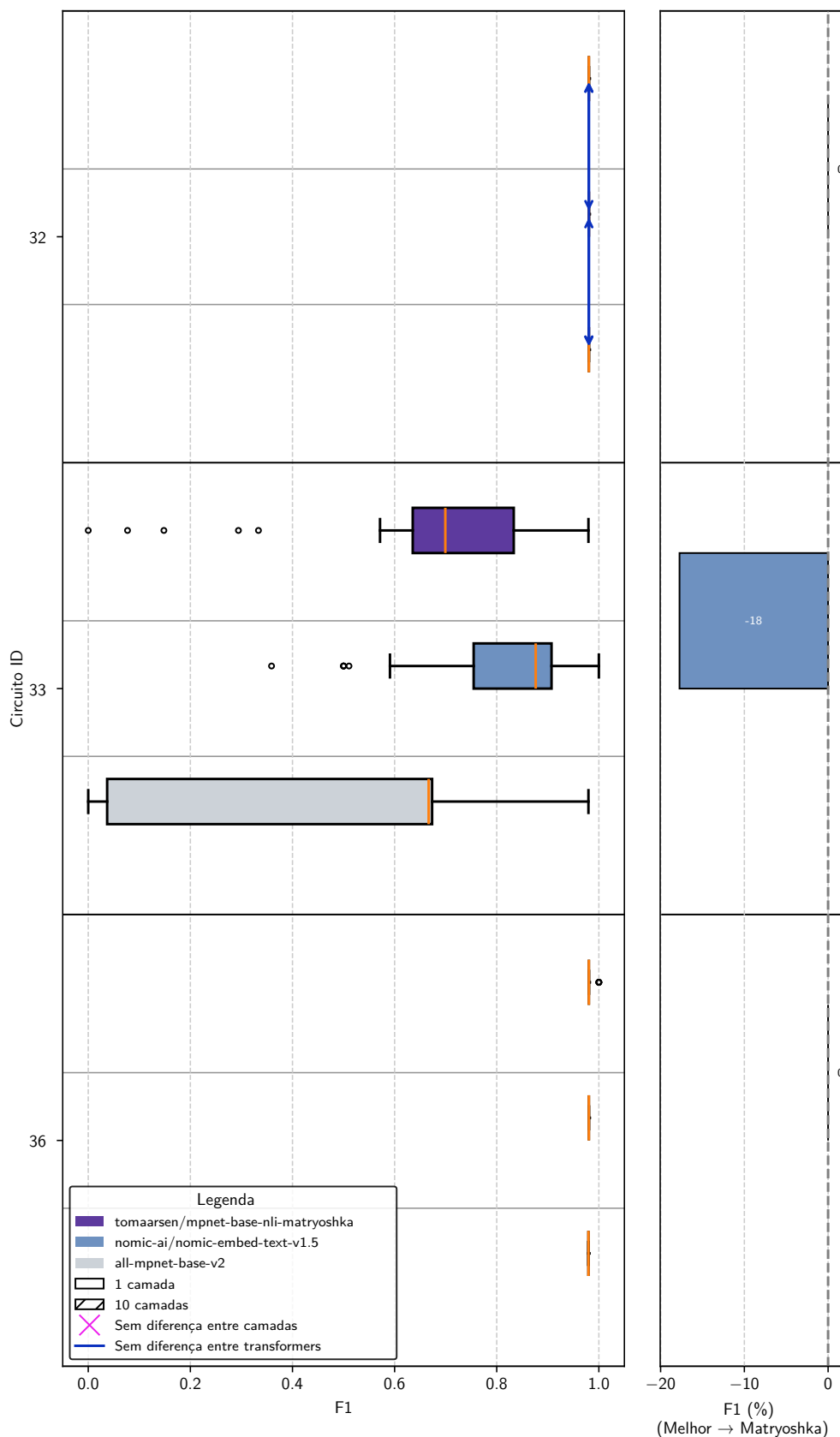
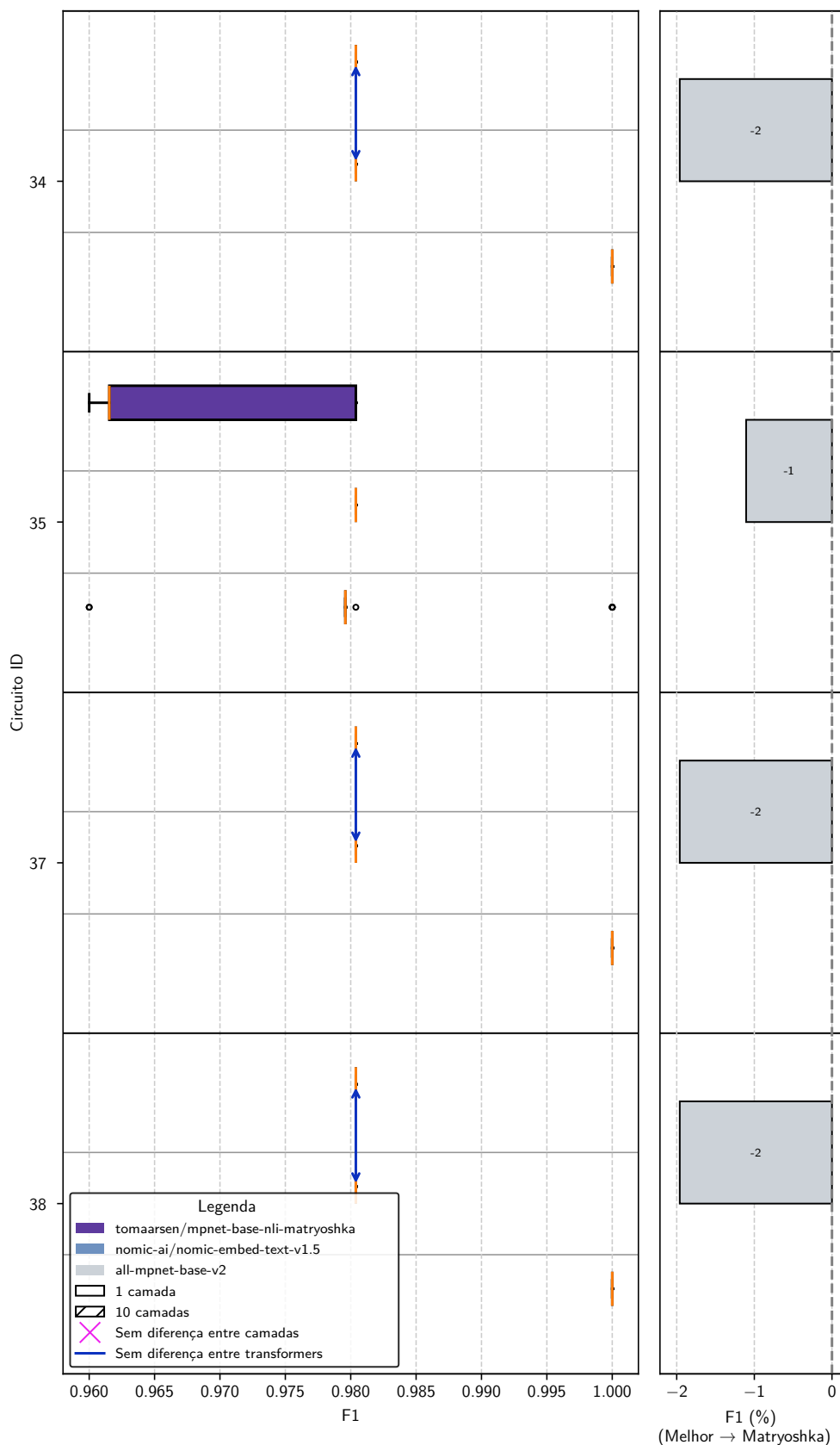


Figura 49 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 com os modelos geradores de *embeddings* na base *ChatGPT* - Difícil. Esses modelos implementam modelos clássicos não-lineares. As barras comparativas do lado direito exibem a diferença percentual da métrica F1 entre o modelo com a maior média F1 e o modelo *Matryoshka*, divididos pela profundidade de cada circuito. As barras de diferença são coloridas com a mesma cor do melhor modelo em questão. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon com correção de p-valor de Holm.



resultados reforçam a conclusão de que, para *ansatze* de baixa capacidade representacional, a escolha do modelo gerador de *embeddings* não é apenas um fator de otimização, mas um determinante crítico do desempenho final.

Em contraposição a este cenário, os modelos clássicos de ponta demonstram uma robustez notável. A Regressão Logística (modelo 32) e as Máquinas de Vetores de Suporte com *kernels* linear (modelo 36) e RBF (modelo 38) exibem um desempenho consistentemente elevada em todas as bases sintéticas, com uma sensibilidade reduzida à escolha do *embedding*. Na base *ChatGPT* Difícil, por exemplo, o SVM com *kernel* linear alcançou uma mediana de F1 de 0.98 com o *Matryoshka*, um resultado estatisticamente equivalente ao obtido com os demais *embeddings*.

A estabilidade destes modelos clássicos sugere que, para a tarefa proposta e com os dados sintéticos, a informação classificatória fundamental é preservada em todas as representações vetoriais testadas e é acessível por meio de fronteiras de decisão (lineares ou não-lineares) que estes modelos são eficientes em encontrar. O Perceptron (modelo 33), no entanto, atua como um contraponto, exibindo uma volatilidade considerável e reforçando que nem todos os modelos clássicos compartilham da mesma robustez, sendo seu desempenho mais dependente de uma representação de entrada que favoreça uma separação linear ideal.

A introdução de profundidade e, de forma mais relevante, de capacidade de emaranhamento intrínseco nos *ansatze* quânticos atua como um poderoso mecanismo mitigador desta dependência. Para os circuitos emaranhadores, o aumento do número de camadas de $L = 1$ para $L = 10$ não apenas aprimora o desempenho, mas também confere ao modelo uma maior robustez em relação à escolha do *embedding*. O circuito 6, na base *ChatGPT* Difícil, é um exemplo emblemático desta dinâmica. Com $L = 1$, o desempenho com o *embedding* *Nomic* (mediana de 0.88) é superior à obtida com o *Matryoshka* (mediana de 0.78). Contudo, com $L = 10$, ambos os modelos convergem para um desempenho de excelência, com medianas de 0.98, eliminando a disparidade inicial.

De forma análoga, o circuito 8, na base *ChatGPT* Fácil, demonstra uma inversão de desempenho. Com $L = 1$, o *MPNet* se mostra a melhor opção, mas com $L = 10$, é o *Matryoshka* que alcança o desempenho máxima (mediana de 1.0), superando os demais. Este comportamento sugere que a capacidade representacional expandida do circuito profundo permite-lhe aprender a extrair o sinal classificatório de forma eficaz, mesmo a partir de uma representação de entrada que, para um *ansatz* mais simples, seria subótima.

O circuito 9, na base *ChatGPT* Difícil, reforça esta conclusão. o desempenho com o

embedding Matryoshka para $L = 1$ (mediana de 0.86) é aproximadamente 13% inferior à do *Nomic* (mediana de 0.98). No entanto, ao se aumentar a profundidade para $L = 10$, o modelo com *Matryoshka* eleva sua mediana para 0.98, alcançando um desempenho estatisticamente equivalente ao do *Nomic*. Fica evidente que a profundidade e o emaranhamento conferem ao classificador quântico uma relativa robustez, tornando-o menos suscetível às variações na etapa de pré-processamento clássico.

Esta robustez adquirida pelos circuitos quânticos profundos e emaranhadores os aproxima, em comportamento, dos modelos clássicos mais estáveis. Enquanto um circuito quântico raso atua como um processador sensível, cujo sucesso depende de uma entrada de dados “ideal”, um circuito profundo torna-se um aprendiz mais autônomo, capaz de inferir e construir as correlações necessárias para a tarefa, mesmo a partir de uma representação de entrada menos otimizada.

A interação entre a complexidade do problema e a escolha do *embedding* também se mostra relevante. Nas bases *ChatGPT* Fácil e Difícil, o modelo *Nomic* frequentemente favorece os circuitos mais simples, enquanto na base *ChatGPT* Médio, o *MPNet* por vezes assume a liderança. O *Matryoshka*, por sua vez, embora nem sempre seja o melhor para os circuitos rasos, demonstra ser uma escolha de excelência quando combinado com arquiteturas profundas e emaranhadoras, como os circuitos 6 e 11, que alcançam desempenho máxima nesta configuração.

Esta variação sugere que não existe um modelo gerador de *embeddings* universalmente superior, mas sim uma interação complexa entre a forma como cada modelo estrutura o espaço vetorial e a capacidade de cada classificador de explorar essa estrutura. A escolha ótima do *embedding* é, portanto, dependente tanto da arquitetura do classificador quanto da natureza intrínseca dos dados.

Em síntese, a análise da sensibilidade dos classificadores à escolha do modelo gerador de *embeddings* revela uma clara hierarquia. Os circuitos quânticos não-emaranhadores posicionam-se como os mais frágeis, com seu desempenho sendo fortemente condicionado pela qualidade da representação de entrada. Os modelos clássicos de ponta, em contrapartida, exibem uma notável estabilidade, alcançando alto desempenho de forma consistente. Por fim, os circuitos quânticos profundos e emaranhadores emergem como as arquiteturas mais sofisticadas, que utilizam sua capacidade representacional expandida não apenas para aprimorar o desempenho, mas também para mitigar a dependência de uma representação de entrada específica. Este achado sublinha uma conclusão fundamental: o investimento em arquiteturas quânticas com

alta capacidade de emaranhamento e profundidade adequada é um caminho promissor não apenas para alcançar um desempenho superior, mas também para desenvolver classificadores mais robustos e versáteis, menos suscetíveis às variações nas etapas de pré-processamento clássico.

4.3 COMPARAÇÃO ENTRE MODELOS QUÂNTICOS E CLÁSSICOS E SEUS ENSEMBLES

Uma análise criteriosa do desempenho dos modelos propostos revela uma distinção clara entre os resultados obtidos nas bases de dados sintéticas e no cenário de maior complexidade da base SST. A comparação direta com os *benchmarks* clássicos estabelece um panorama factual sobre o estado atual dos classificadores quânticos variacionais para a tarefa de análise de sentimentos, permitindo uma discussão aprofundada sobre suas potencialidades e os desafios que ainda persistem na busca por uma vantagem quântica. Esta seção se dedica a dissecar essa comparação, contextualizando os achados e explorando as razões por trás das lacunas de desempenho observadas. Sendo assim, utiliza-se as Figuras 50, 51, 52, 53, 54, 55, 56 e 57 para a base *ChatGPT* - Fácil, Figuras 58, 59, 60, 61, 62, 63, 64 e 65 para a base *ChatGPT* - Médio e Figuras 66, 67, 68, 69, 70, 71, 72 e 73 para a base *ChatGPT* - Difícil de modo a criar uma compreensão conjunta dos modelos quânticos e clássicos dentro dos experimentos com 16 atributos e modelo *Matryoshka* como gerador de *embeddings*.

Observa-se, primeiramente, uma superioridade notável dos modelos clássicos mais robustos na base de dados SST. Classificadores como as SVM com kernels linear (modelo 36) e RBF (modelo 38), bem como a Regressão Logística (modelo 32), estabeleceram uma linha de base de alto desempenho, alcançando medianas da métrica F1 em torno de 0.73. Este patamar de desempenho supera consistentemente todos os circuitos quânticos individuais testados, incluindo as arquiteturas mais complexas e profundas. Este achado é de relevância central, pois evidencia que, para a tarefa de análise de sentimentos com os 16 atributos de *embedding* extraídos do modelo *Matryoshka*, uma solução clássica e relativamente simples não apenas é viável, mas também mais eficaz, indicando que uma possível vantagem quântica não foi alcançada neste cenário experimental específico.

A existência desta lacuna de desempenho merece uma análise multifatorial, cujas raízes podem ser traçadas a partir de desafios teóricos e práticos inerentes aos modelos quânticos variacionais. Uma das hipóteses centrais reside na dificuldade de otimização desses circuitos.

Figura 50 – Resultados do desempenho F1 dos circuitos 1 ao 6 na base *ChatGPT* - Fácil. Esses algoritmos classificadores demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

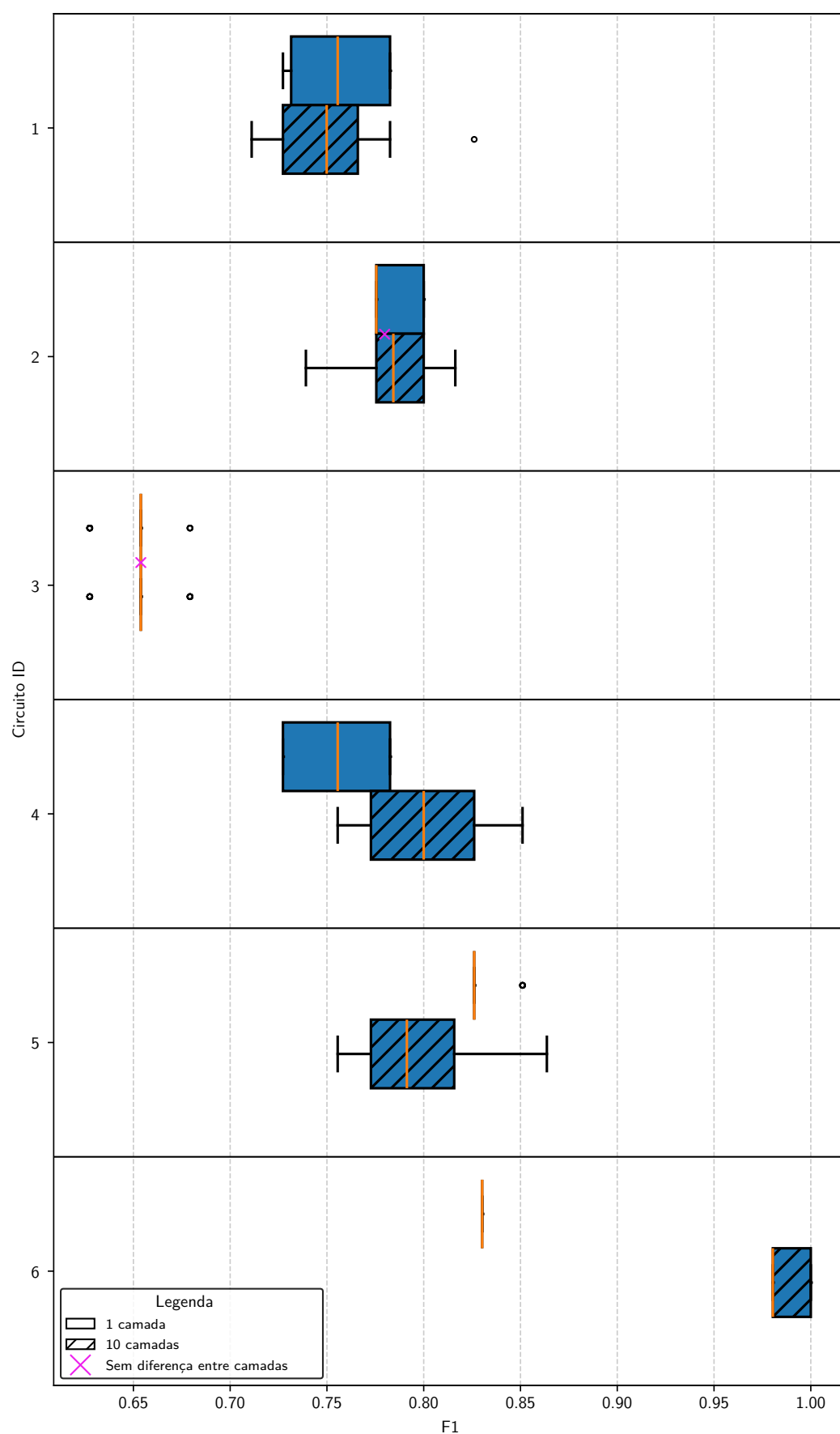
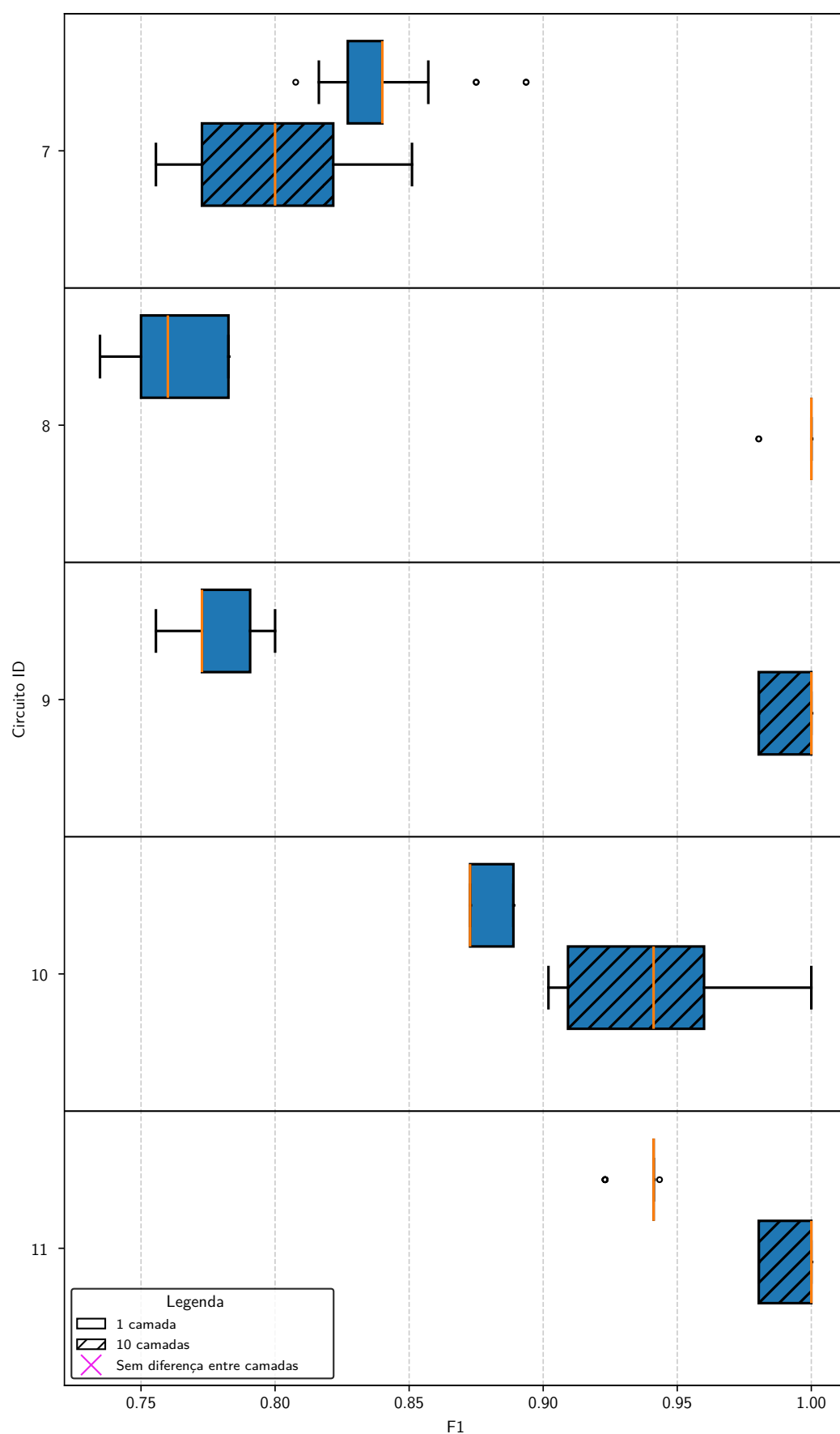


Figura 51 – Resultados do desempenho F1 dos circuitos 7 ao 11 na base *ChatGPT* - Fácil. Esses algoritmos classificadores implementam diferentes formas de emaranhamento quântico. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

Figura 52 – Resultados do desempenho F1 dos circuitos 12 ao 15 na base *ChatGPT* - Fácil. Esses algoritmos classificadores implementam diferentes tipos de conectividade entre *qubits*. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

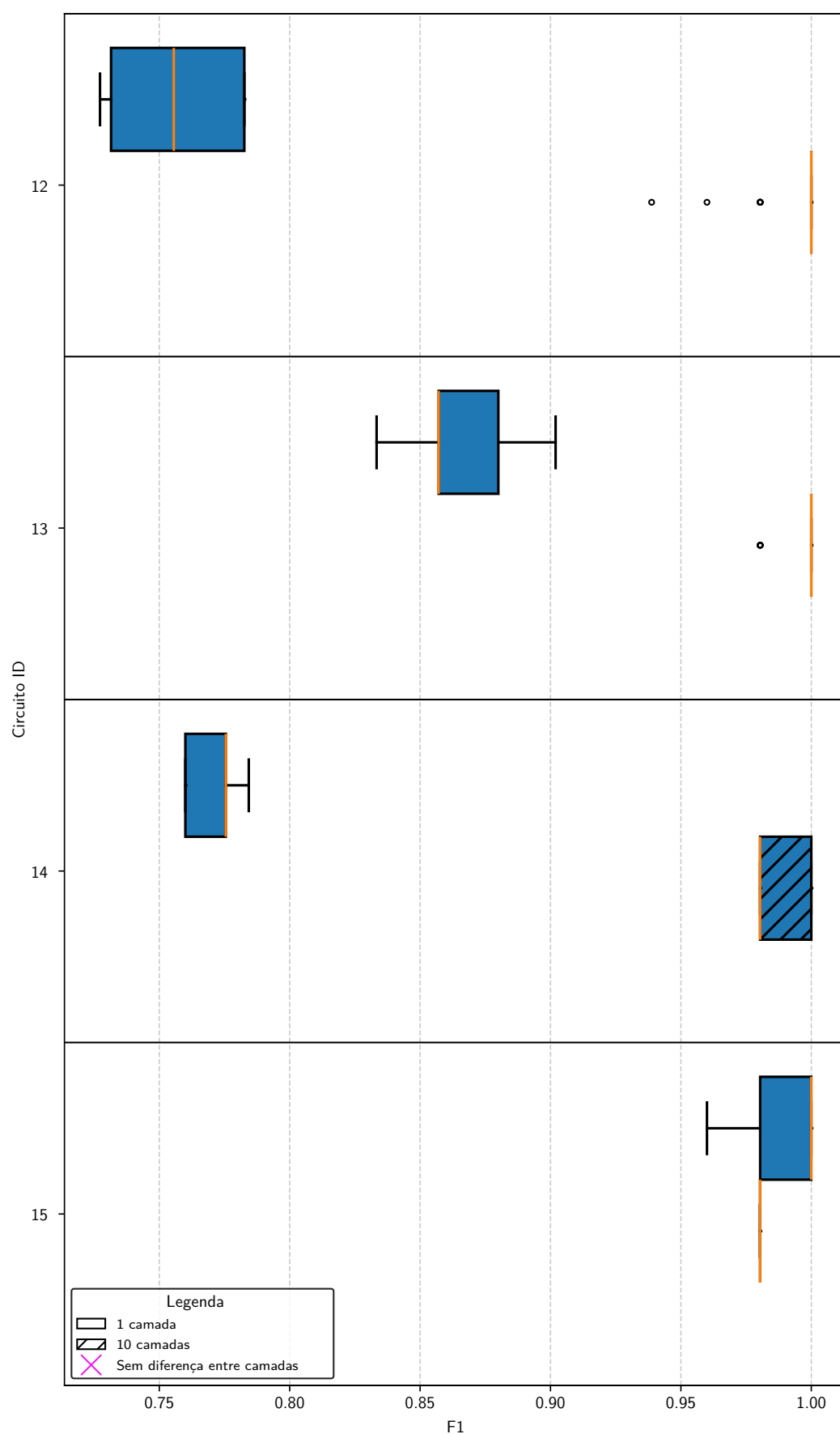


Figura 53 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 na base *ChatGPT* - Fácil. Esses algoritmos classificadores implementam modelos clássicos lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

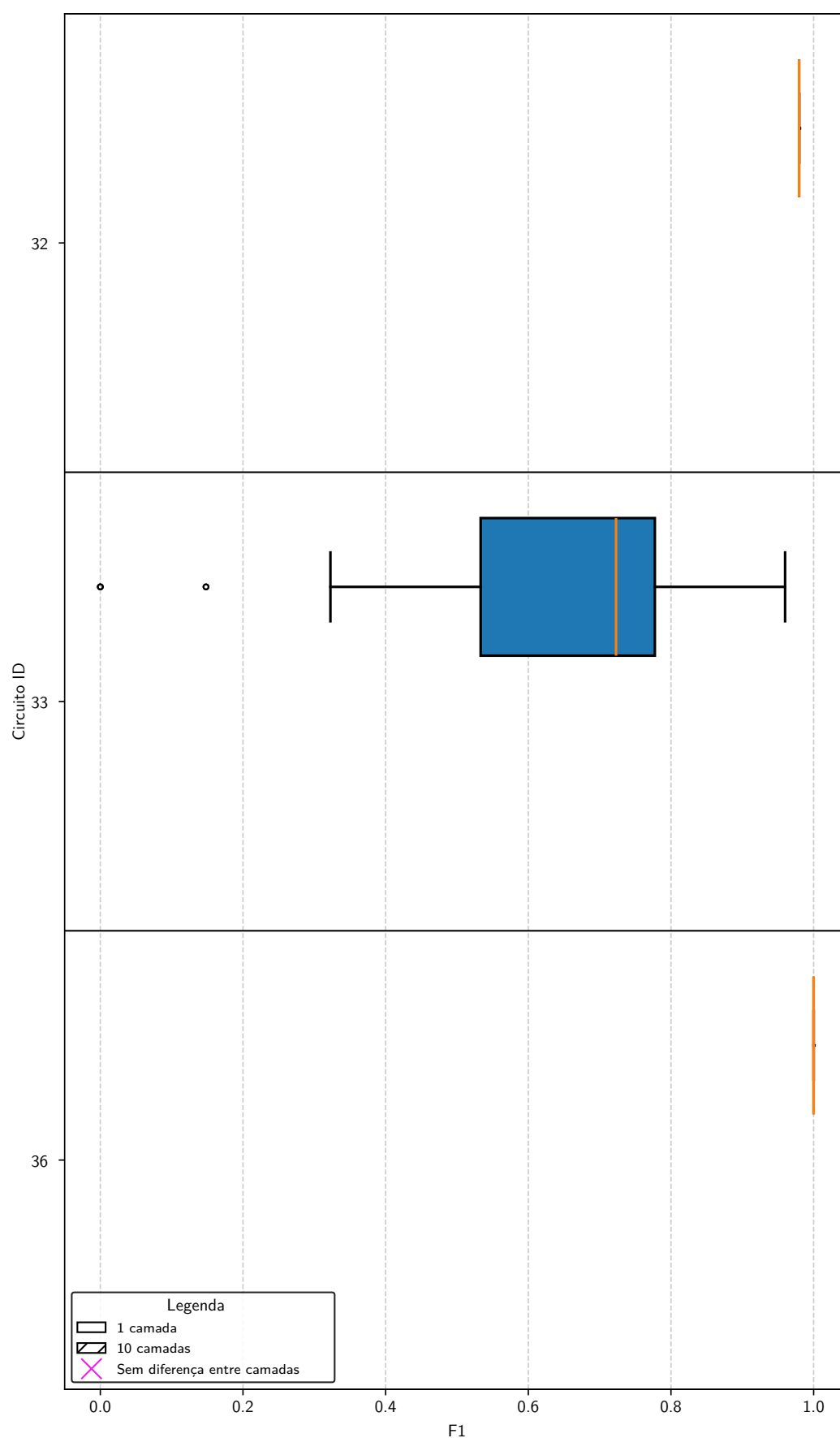


Figura 54 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 na base *ChatGPT* - Fácil. Esses algoritmos classificadores implementam modelos clássicos não-lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

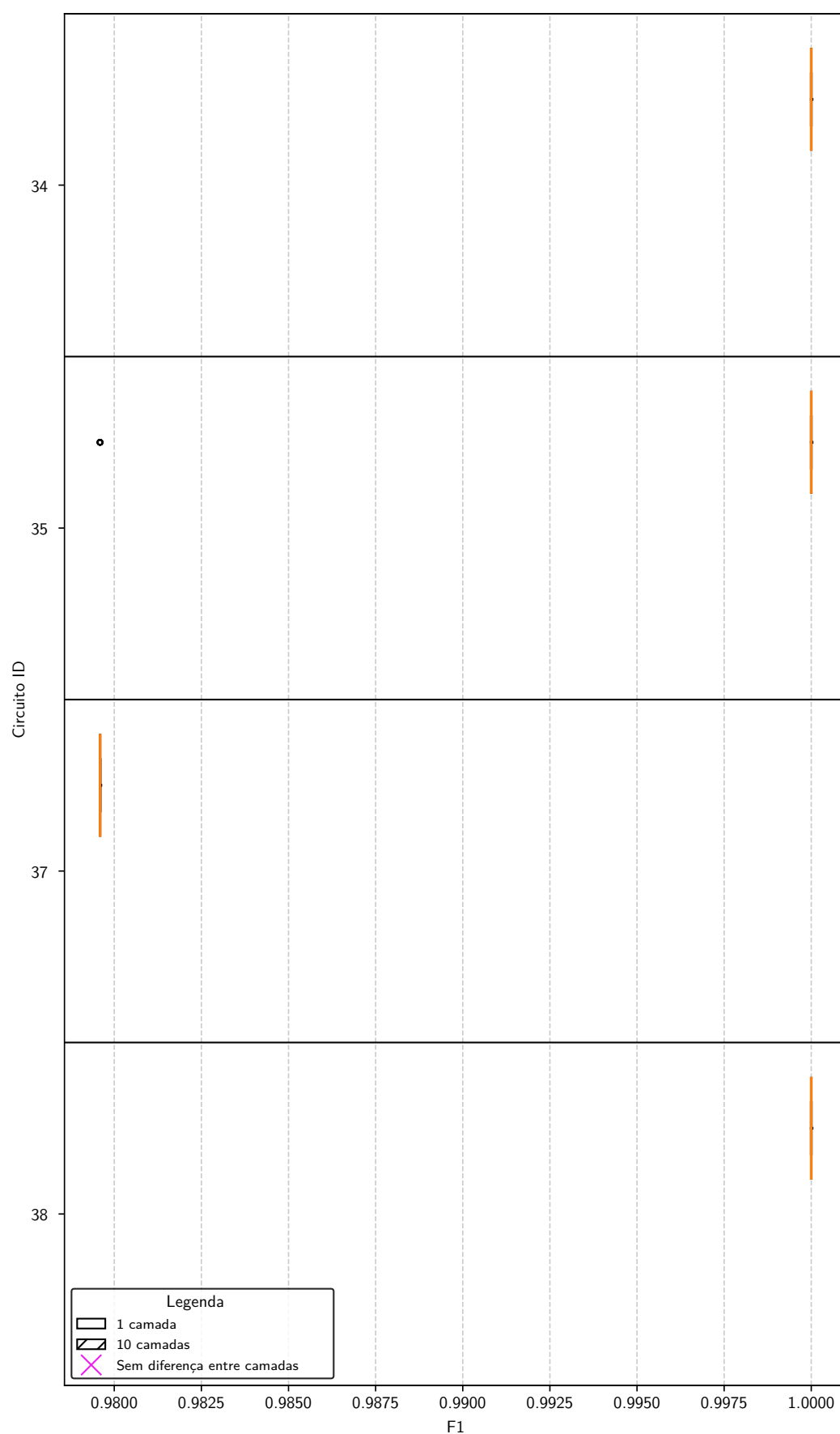


Figura 55 – Resultados do desempenho F1 dos circuitos 16 ao 21 na base *ChatGPT* - Fácil. Esses algoritmos classificadores implementam modelos de ensemble através de *AdaBoost* ou *Bagging* de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

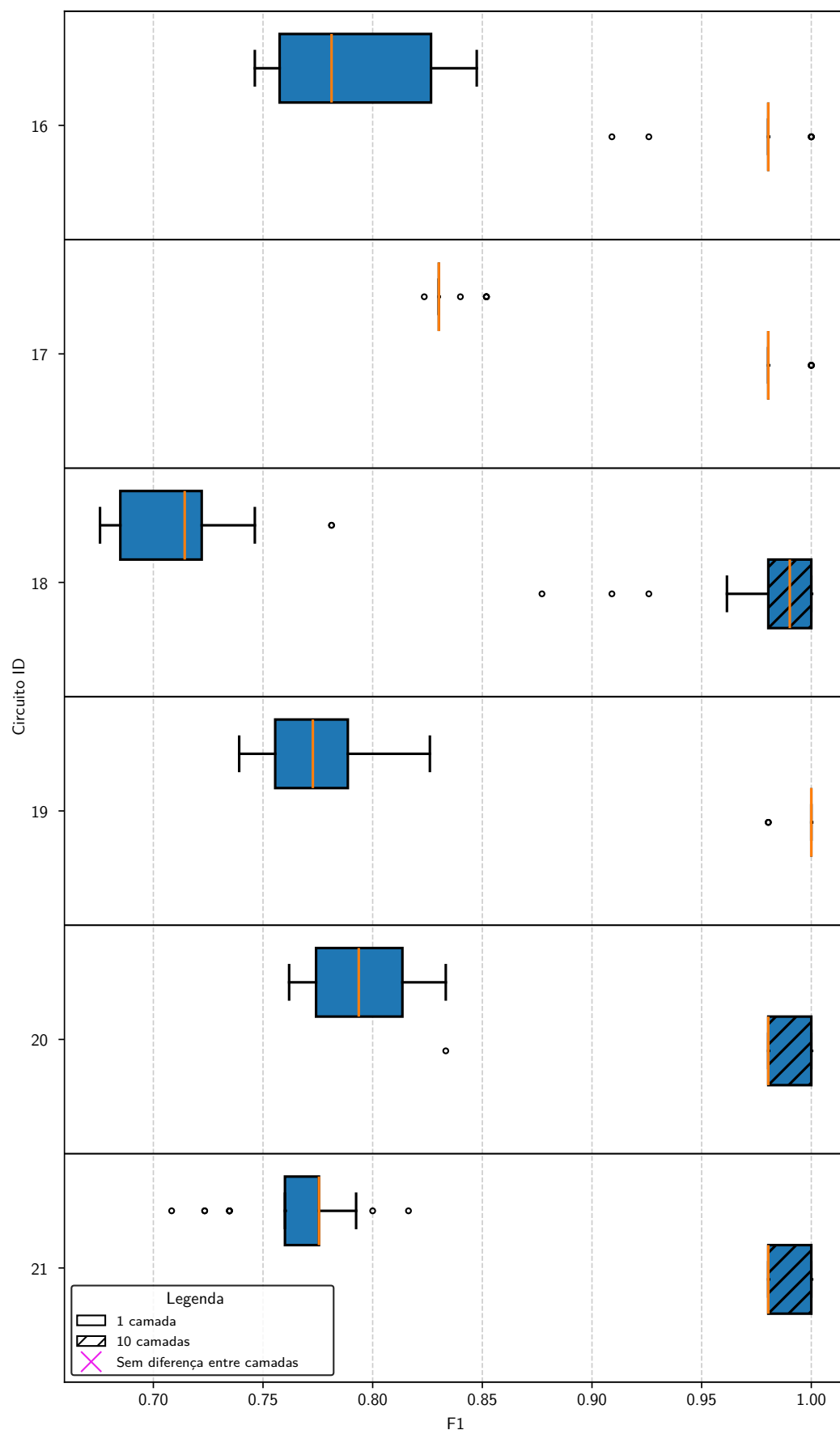


Figura 56 – Resultados do desempenho F1 dos circuitos 22 ao 31 na base *ChatGPT* - Fácil. Esses algoritmos classificadores implementam modelos de ensemble através de *hard* ou *soft voting* de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

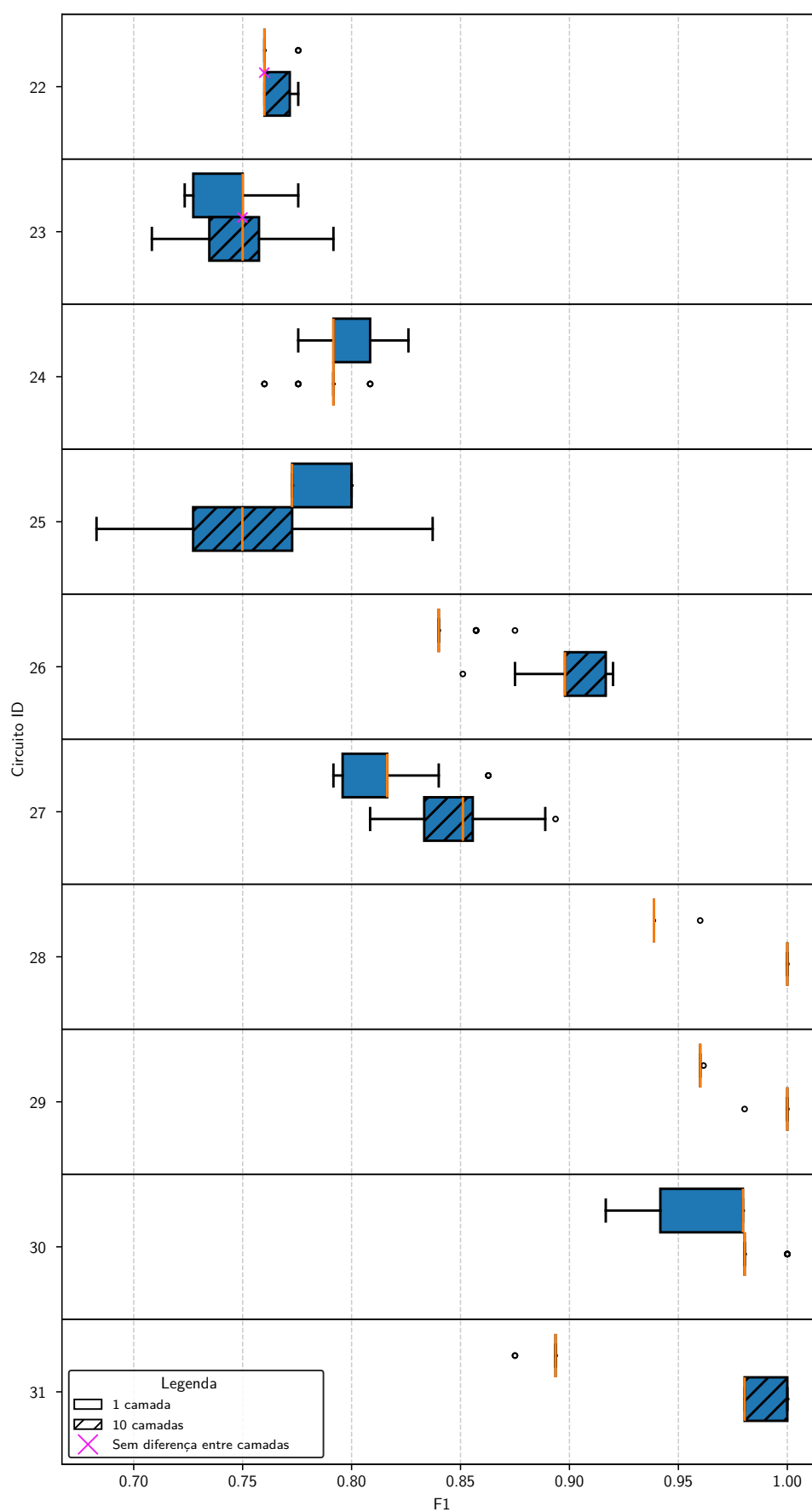
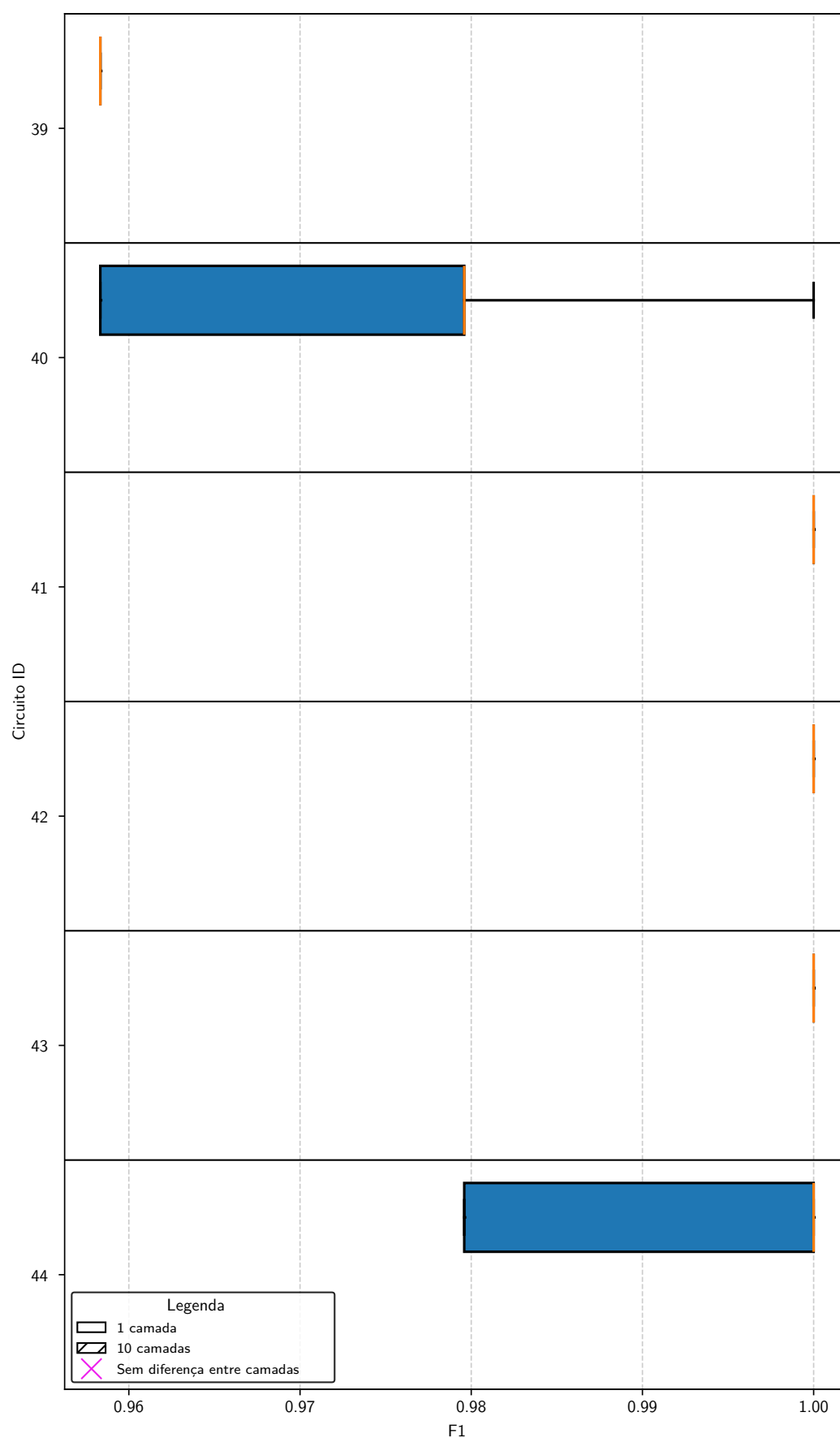
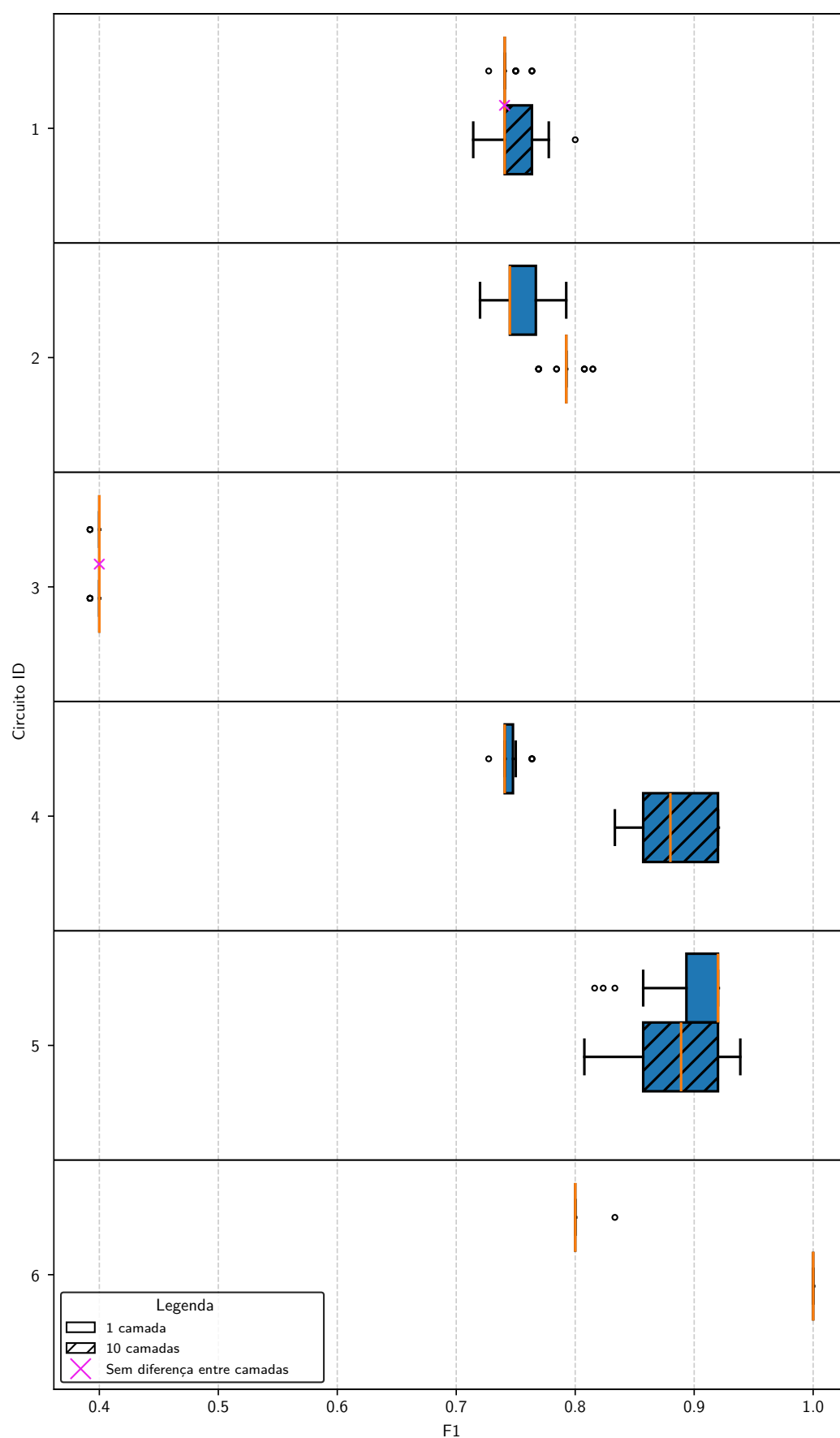


Figura 57 – Resultados do desempenho F1 dos circuitos 39 ao 44 na base *ChatGPT* - Fácil. Esses algoritmos classificadores implementam modelos de ensemble clássicos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



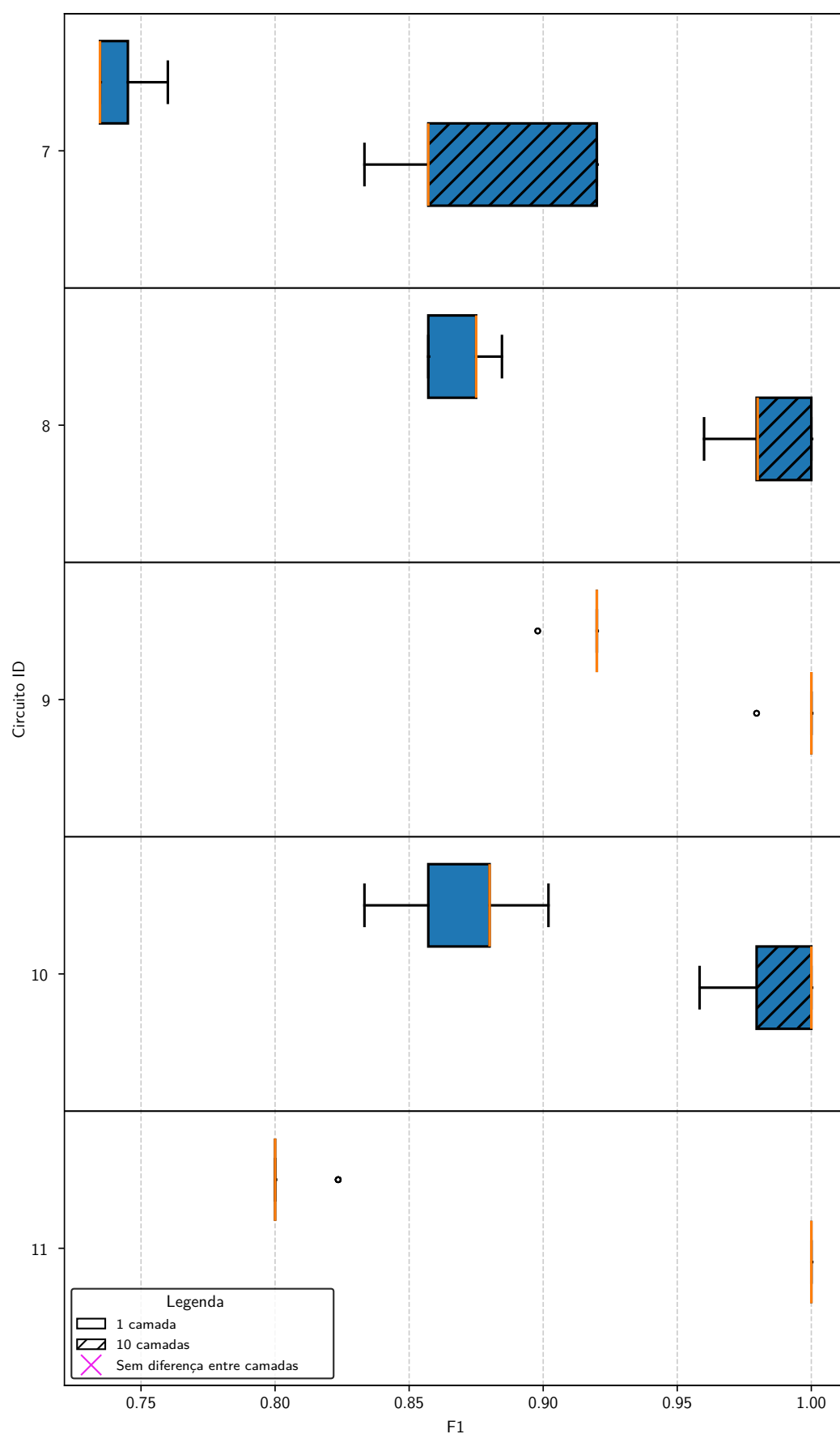
Fonte: Elaborada pelo autor (2025)

Figura 58 – Resultados do desempenho F1 dos circuitos 1 ao 6 na base *ChatGPT* - Médio. Esses algoritmos classificadores demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

Figura 59 – Resultados do desempenho F1 dos circuitos 7 ao 11 na base *ChatGPT* - Médio. Esses algoritmos classificadores implementam diferentes formas de emaranhamento quântico. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

Figura 60 – Resultados do desempenho F1 dos circuitos 12 ao 15 na base *ChatGPT* - Médio. Esses algoritmos classificadores implementam diferentes tipos de conectividade entre *qubits*. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

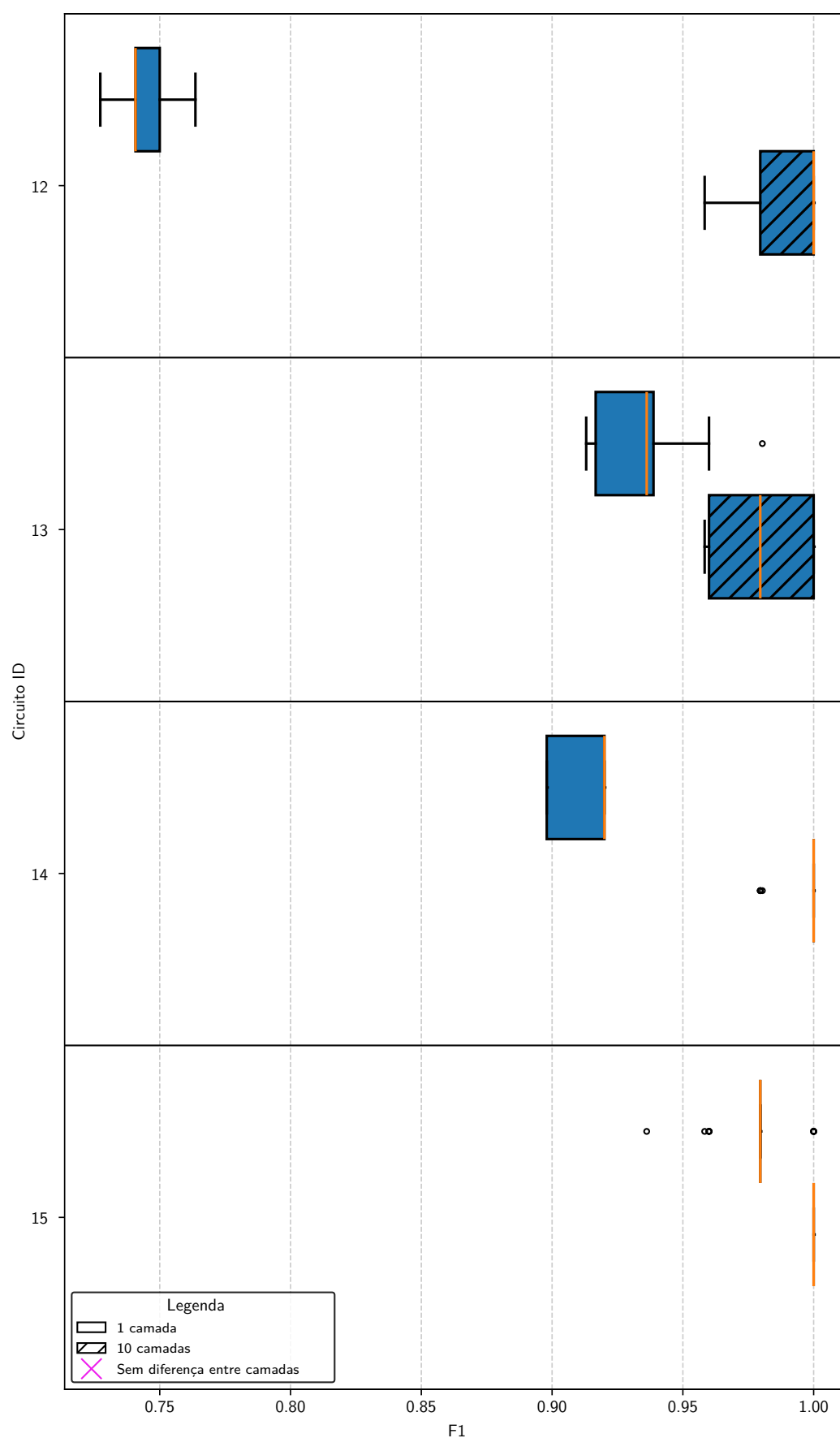


Figura 61 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 na base *ChatGPT* - Médio. Esses algoritmos classificadores implementam modelos clássicos lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

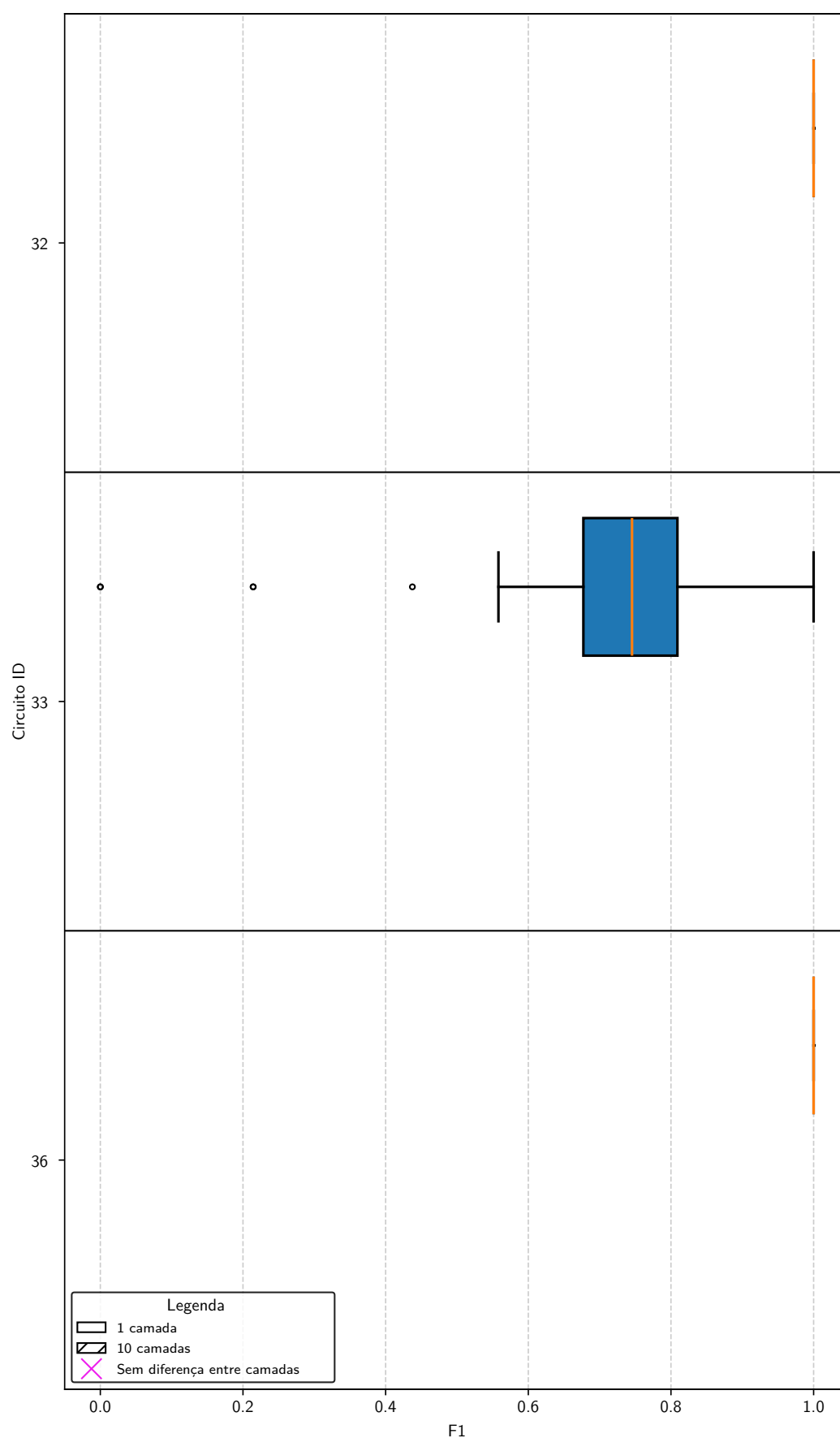
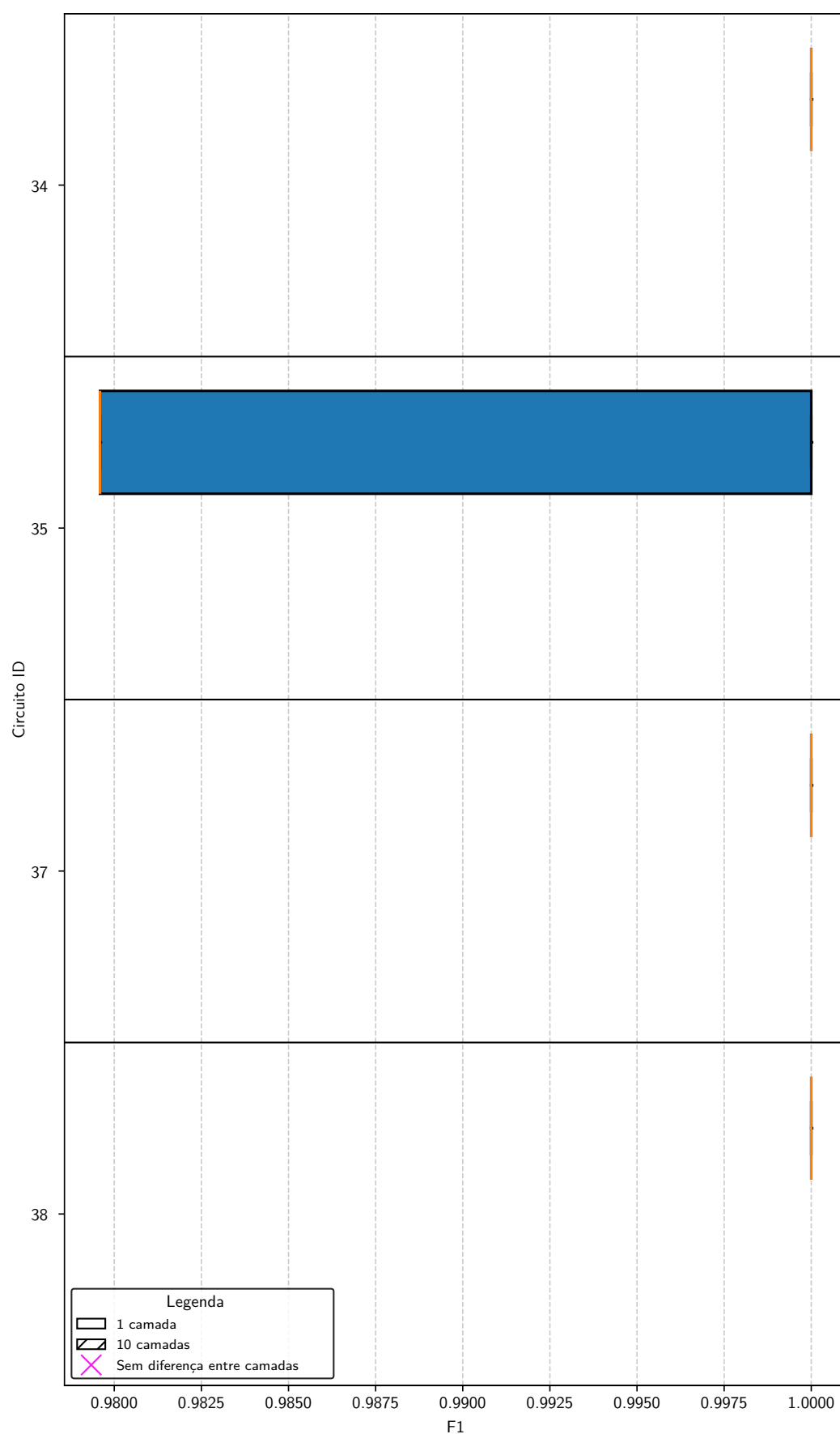


Figura 62 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 na base *ChatGPT* - Médio. Esses algoritmos classificadores implementam modelos clássicos não-lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

Figura 63 – Resultados do desempenho F1 dos circuitos 16 ao 21 na base *ChatGPT* - Médio. Esses algoritmos classificadores implementam modelos de ensemble através de *AdaBoost* ou *Bagging* de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

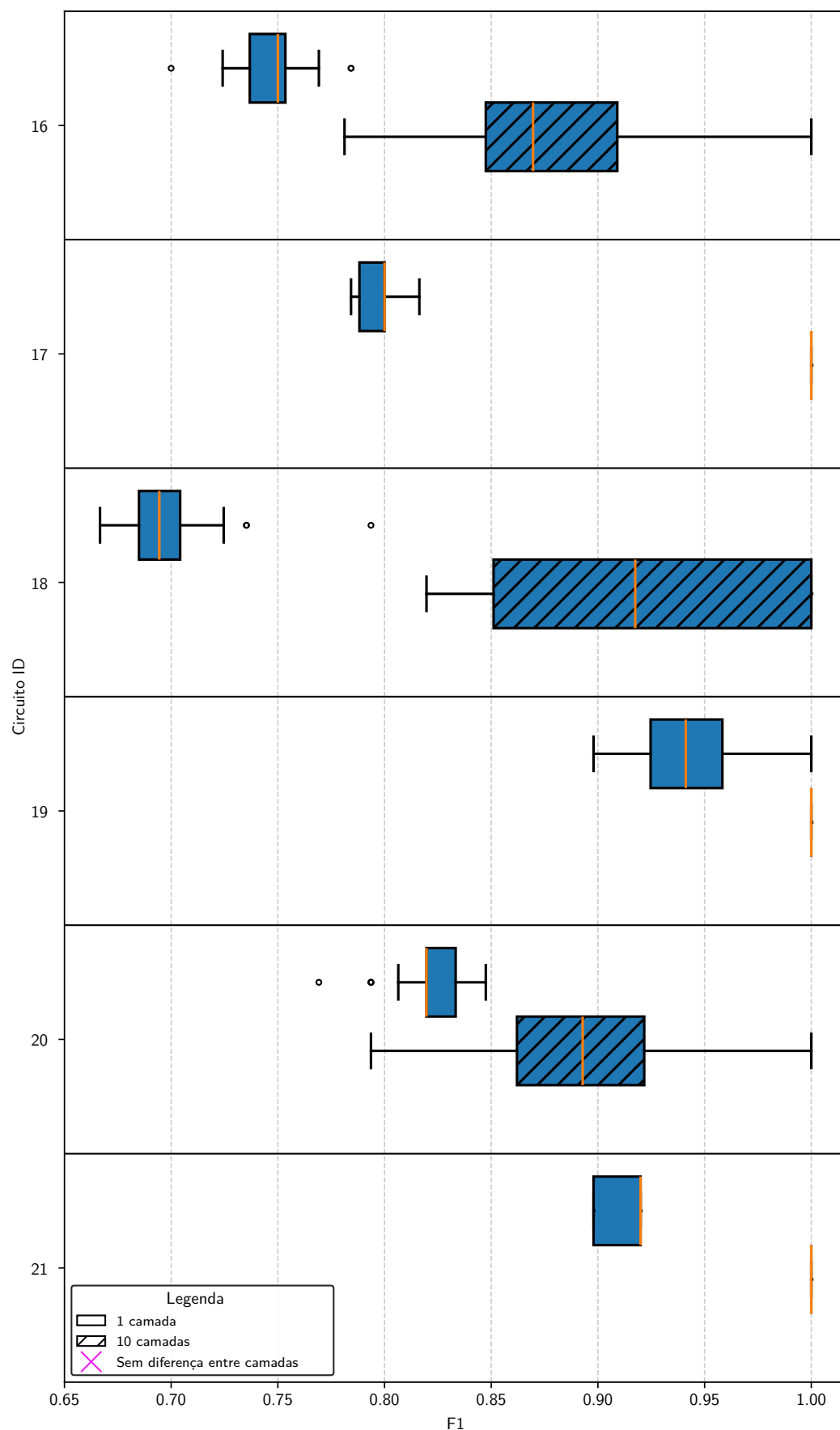


Figura 64 – Resultados do desempenho F1 dos circuitos 22 ao 31 na base *ChatGPT* - Médio. Esses algoritmos classificadores implementam modelos de ensemble através de *hard* ou *soft voting* de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

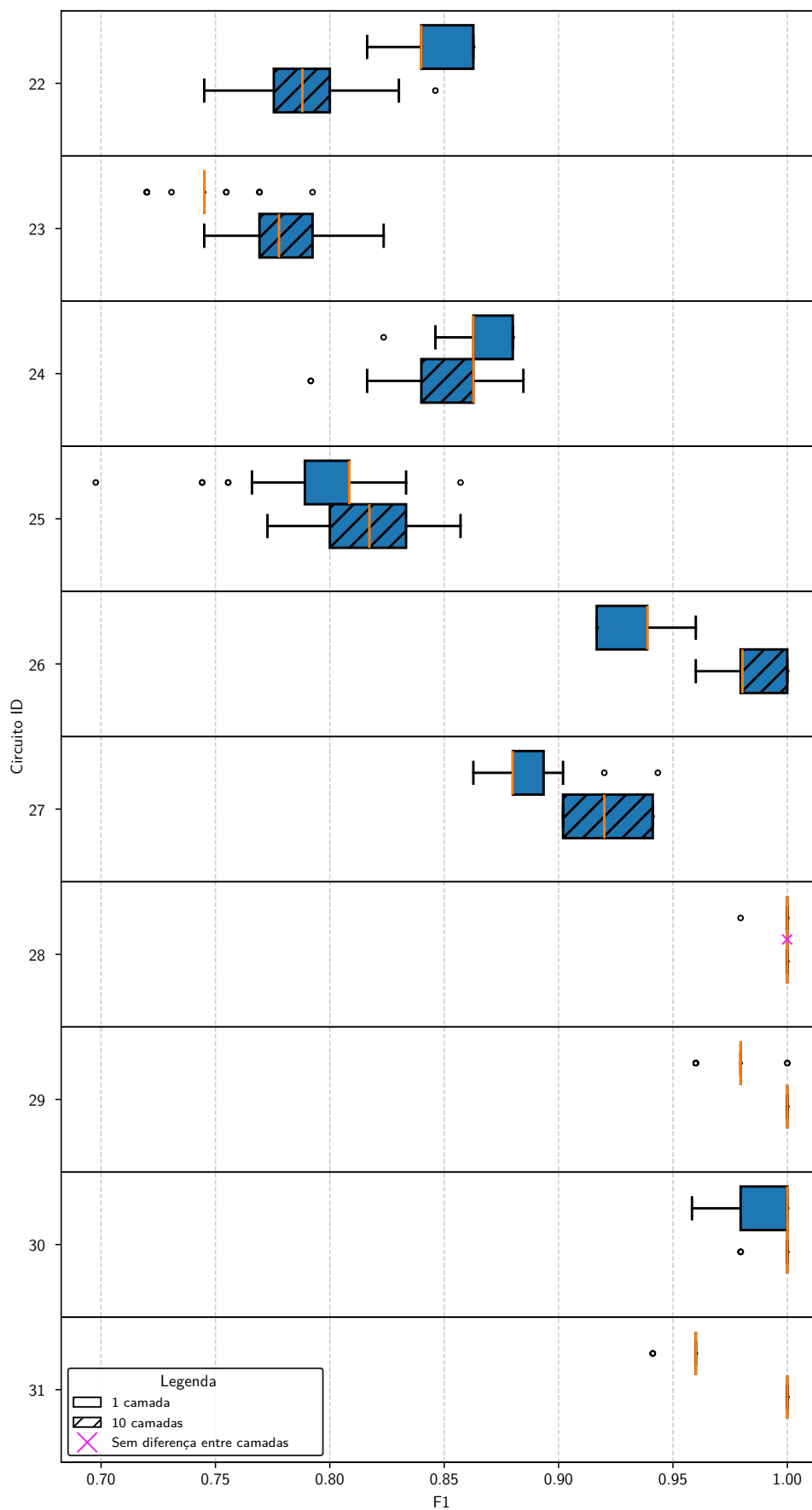


Figura 65 – Resultados do desempenho F1 dos circuitos 39 ao 44 na base *ChatGPT* - Médio. Esses algoritmos classificadores implementam modelos de ensemble clássicos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

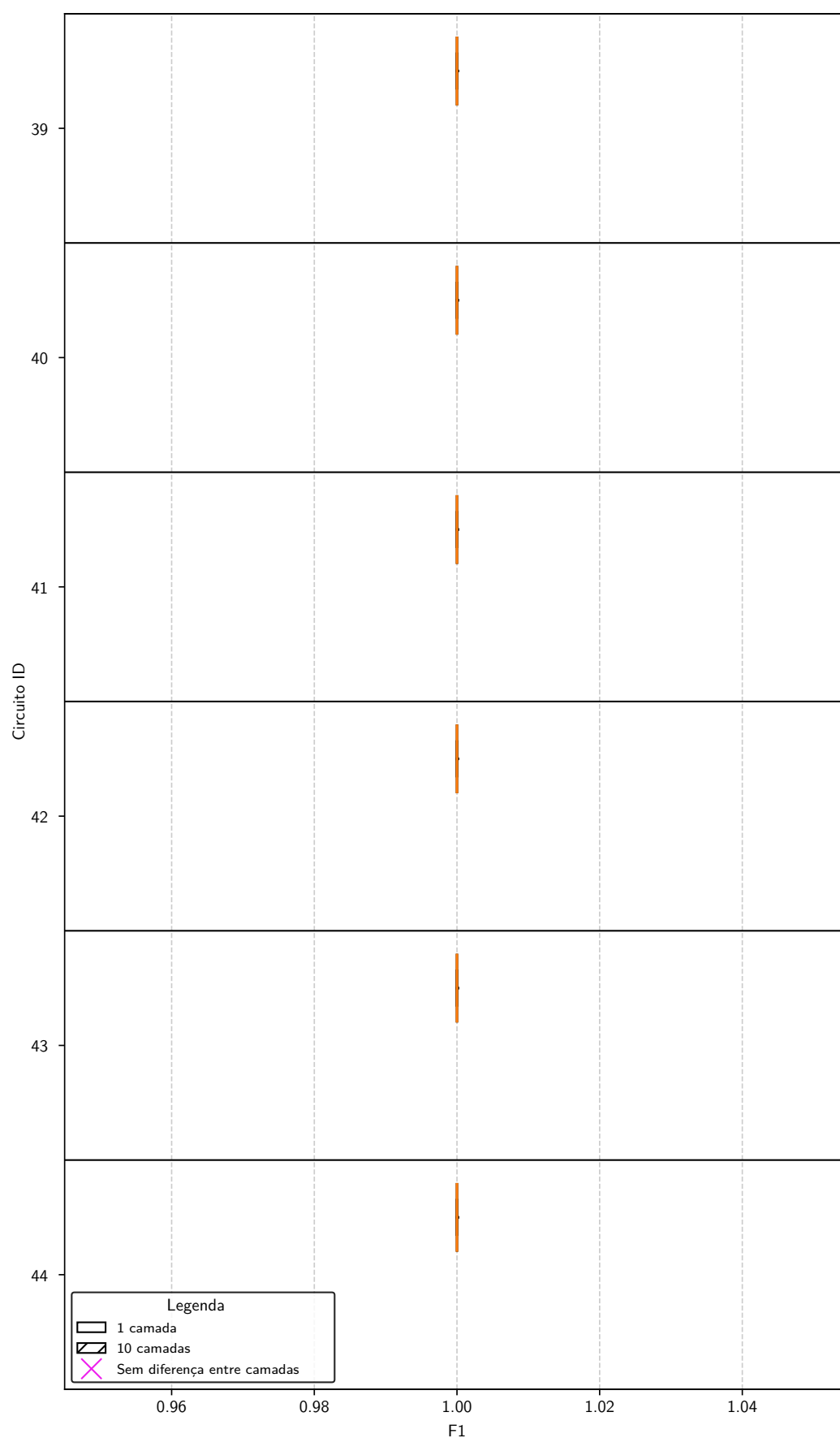


Figura 66 – Resultados do desempenho F1 dos circuitos 1 ao 6 na base *ChatGPT* - Difícil. Esses algoritmos classificadores demonstram a crescente complexidade ao utilizar portas quânticas e portas CNOT. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

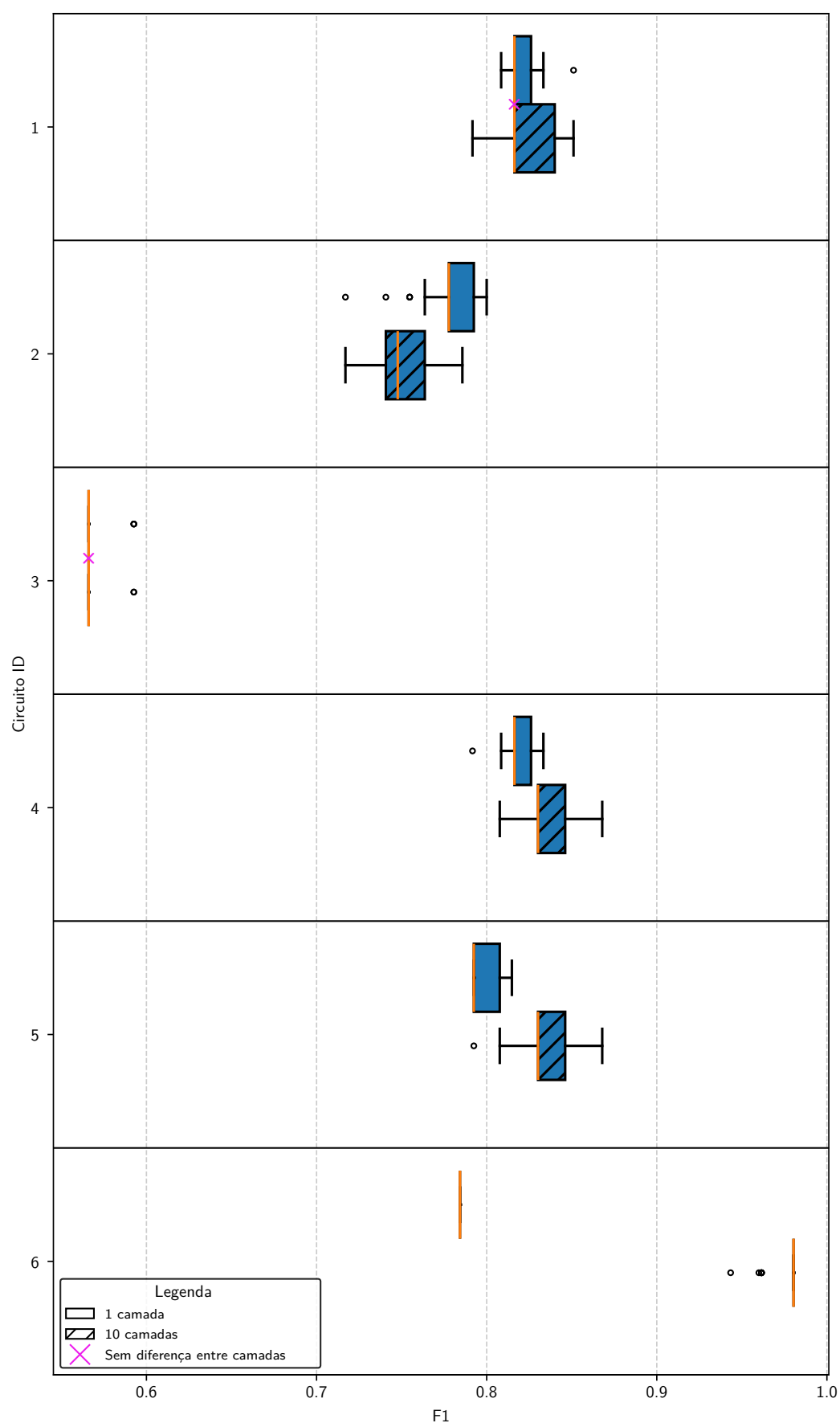
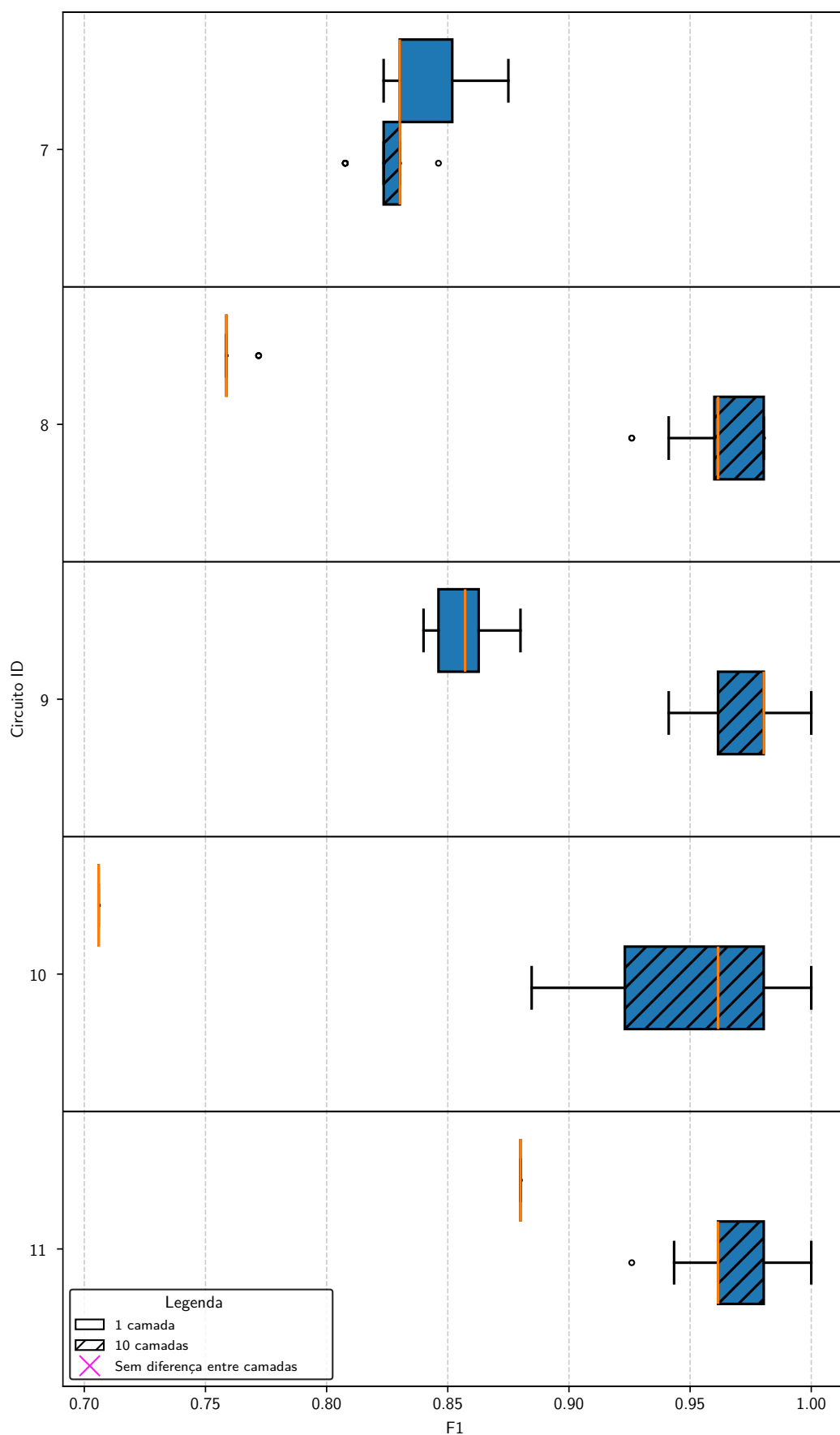


Figura 67 – Resultados do desempenho F1 dos circuitos 7 ao 11 na base *ChatGPT* - Difícil. Esses algoritmos classificadores implementam diferentes formas de emaranhamento quântico. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

Figura 68 – Resultados do desempenho F1 dos circuitos 12 ao 15 na base *ChatGPT* - Difícil. Esses algoritmos classificadores implementam diferentes tipos de conectividade entre *qubits*. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

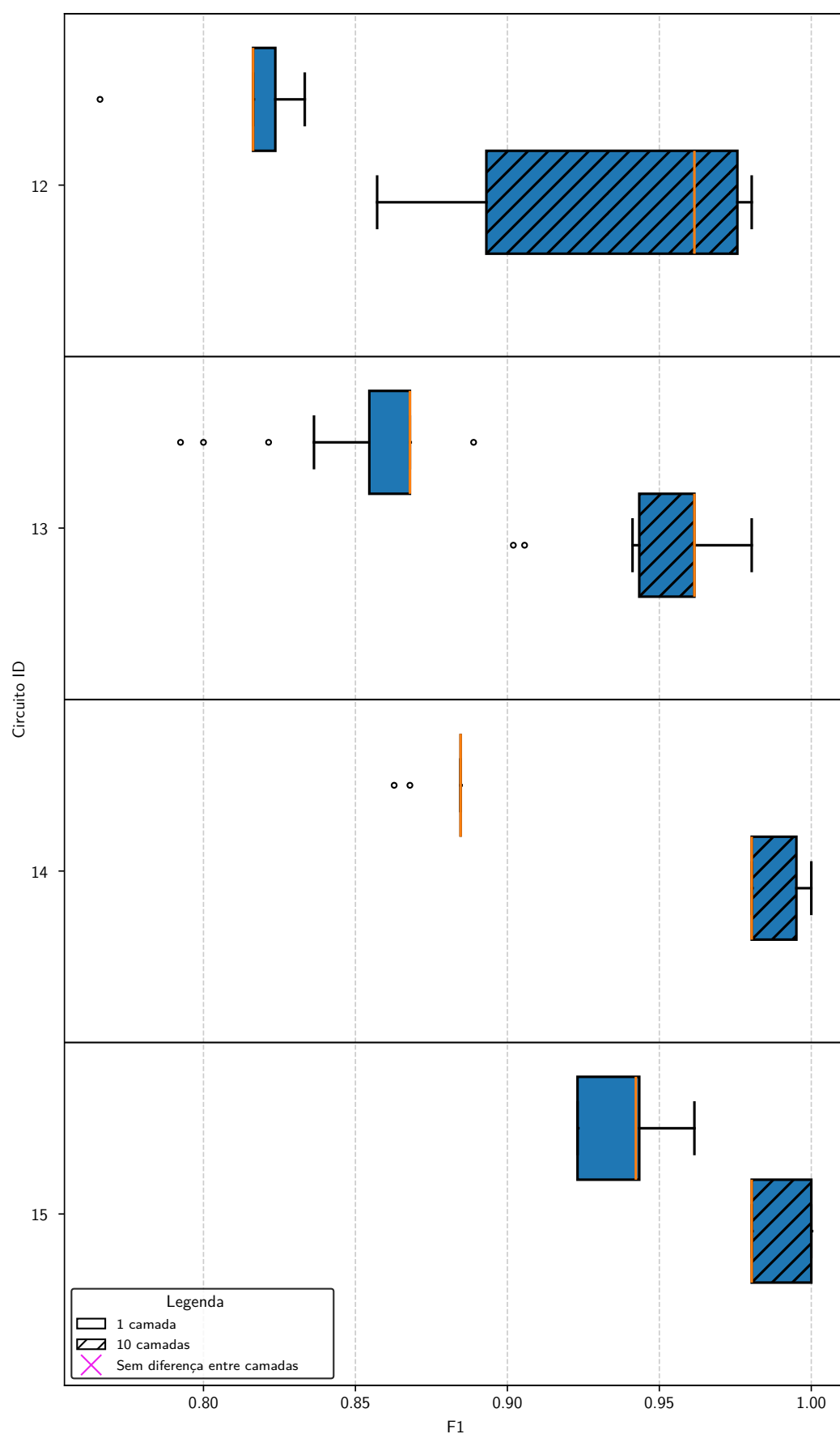


Figura 69 – Resultados do desempenho F1 dos circuitos 32, 33 e 36 na base *ChatGPT* - Difícil. Esses algoritmos classificadores implementam modelos clássicos lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

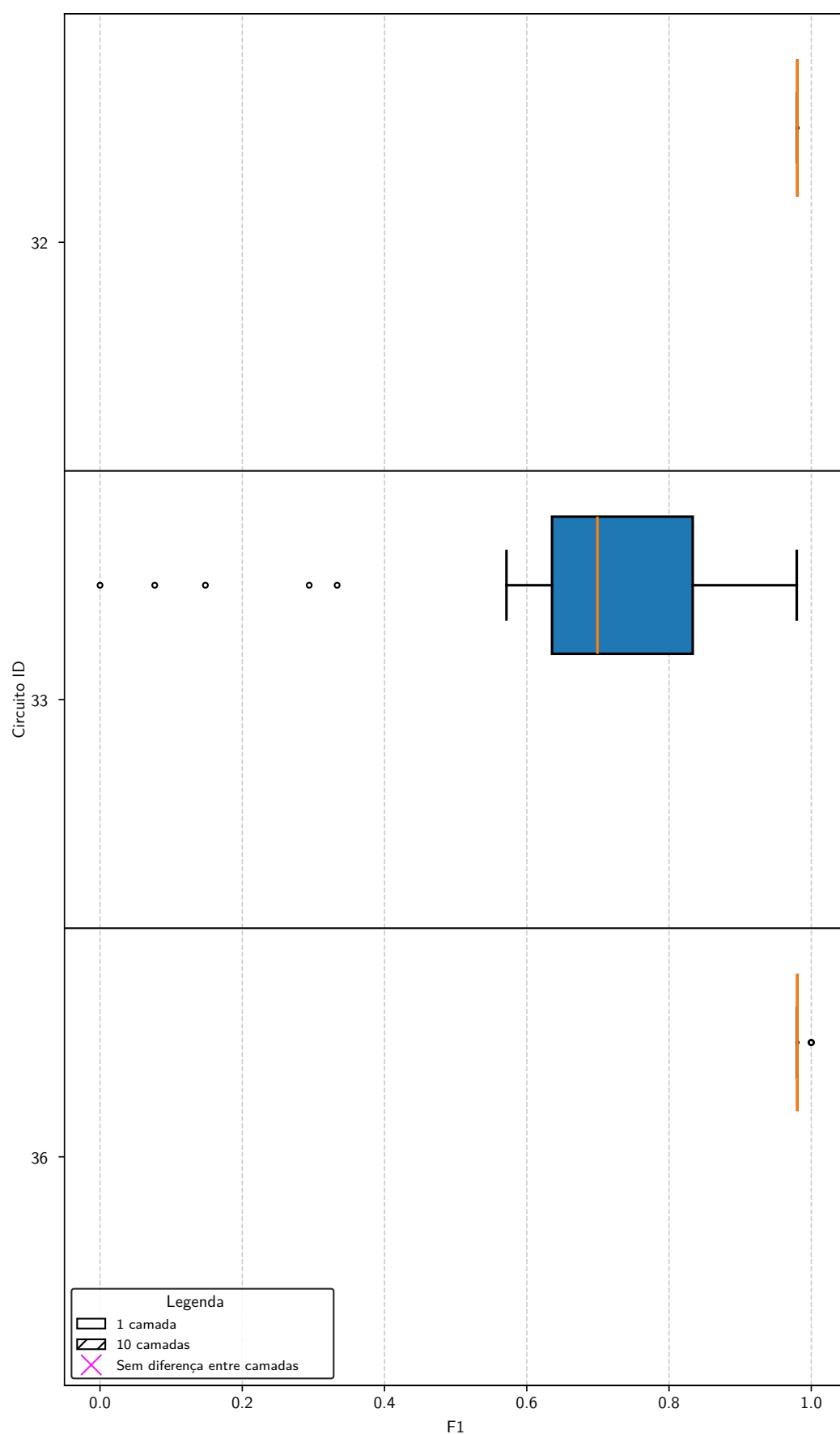


Figura 70 – Resultados do desempenho F1 dos circuitos 34, 35, 37 e 38 na base *ChatGPT* - Difícil. Esses algoritmos classificadores implementam modelos clássicos não-lineares. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

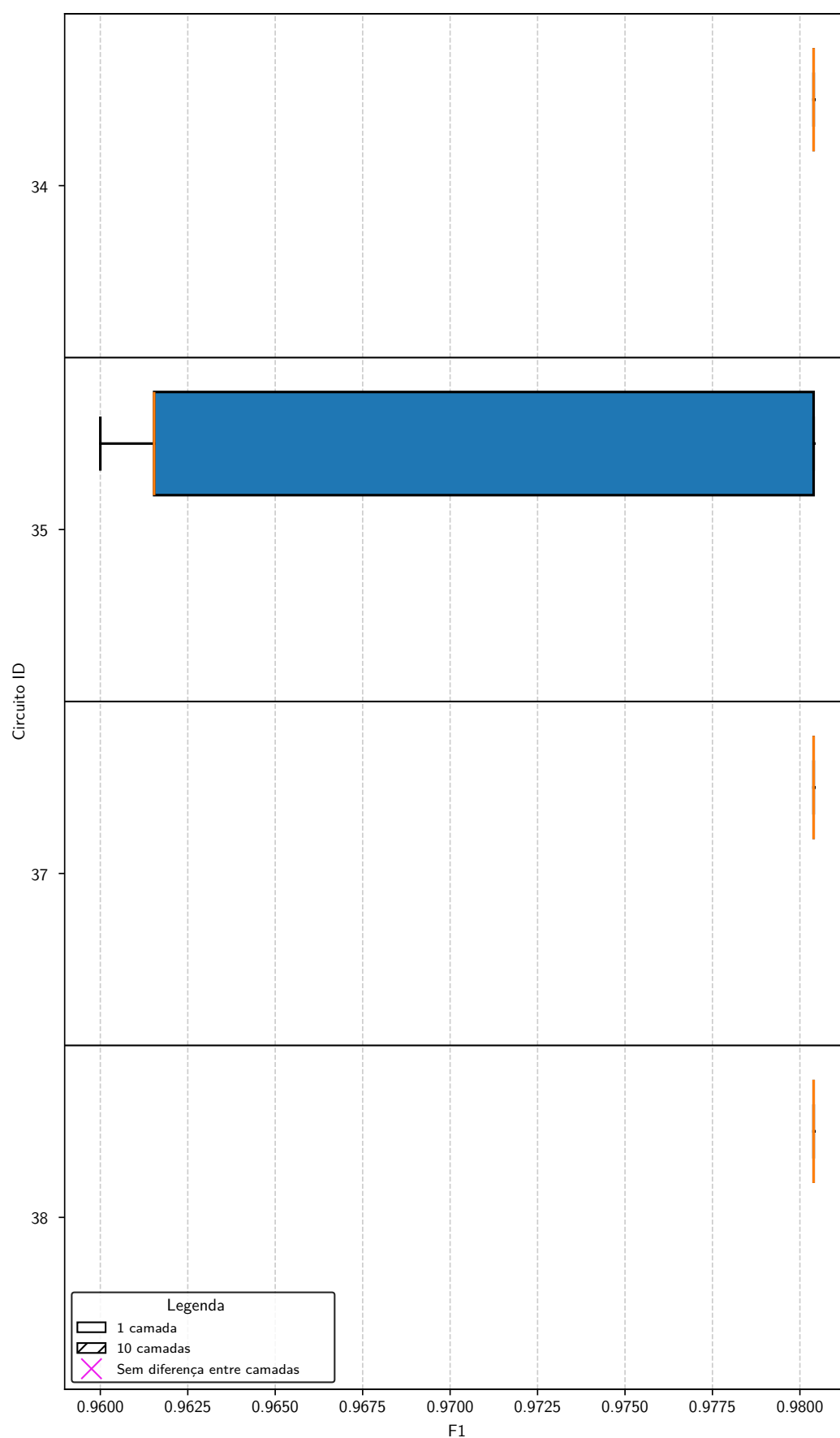


Figura 71 – Resultados do desempenho F1 dos circuitos 16 ao 21 na base *ChatGPT* - Difícil. Esses algoritmos classificadores implementam modelos de ensemble através de *AdaBoost* ou *Bagging* de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

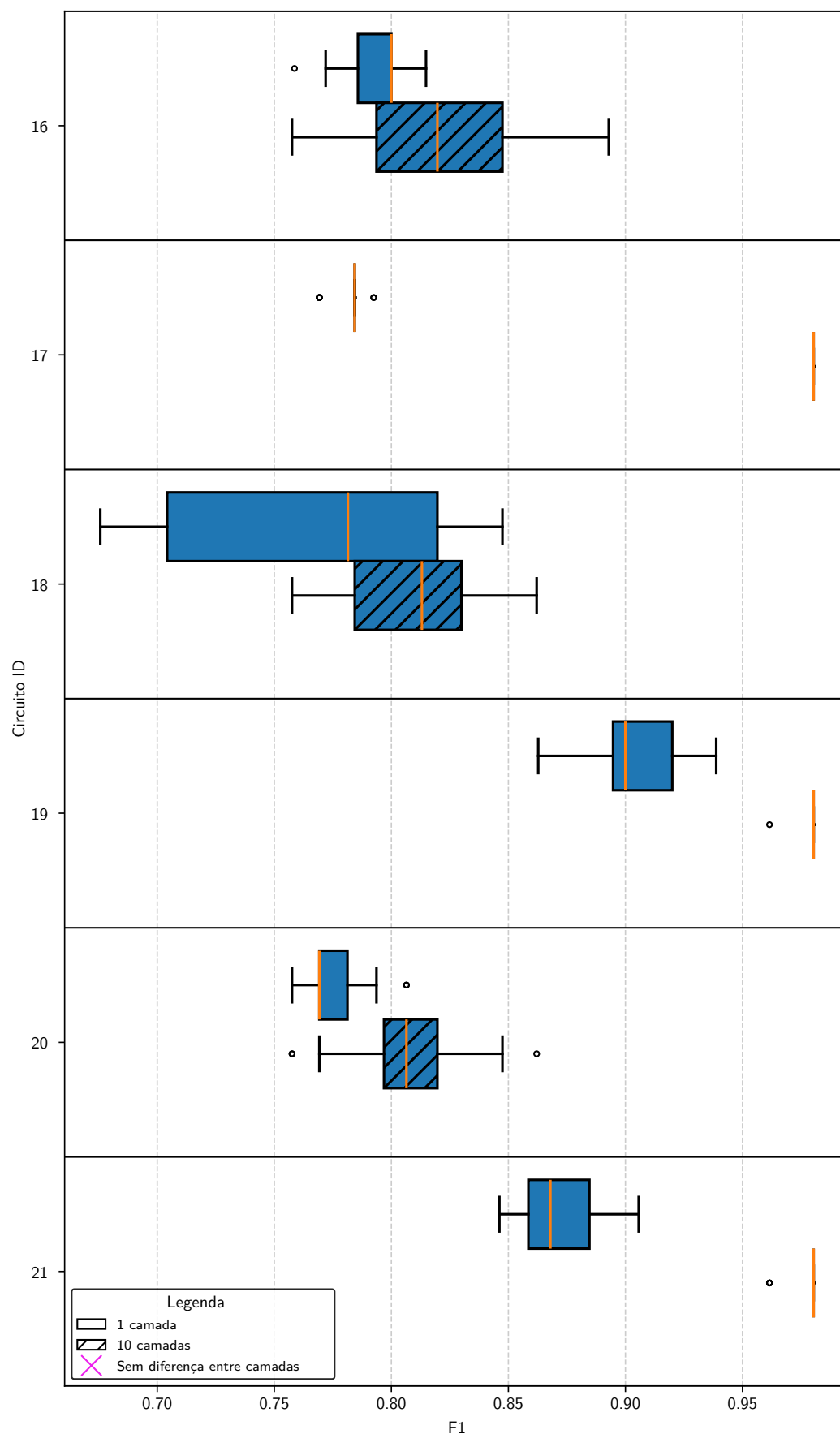


Figura 72 – Resultados do desempenho F1 dos circuitos 22 ao 31 na base *ChatGPT* - Difícil. Esses algoritmos classificadores implementam modelos de ensemble através de *hard* ou *soft voting* de circuitos quânticos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.

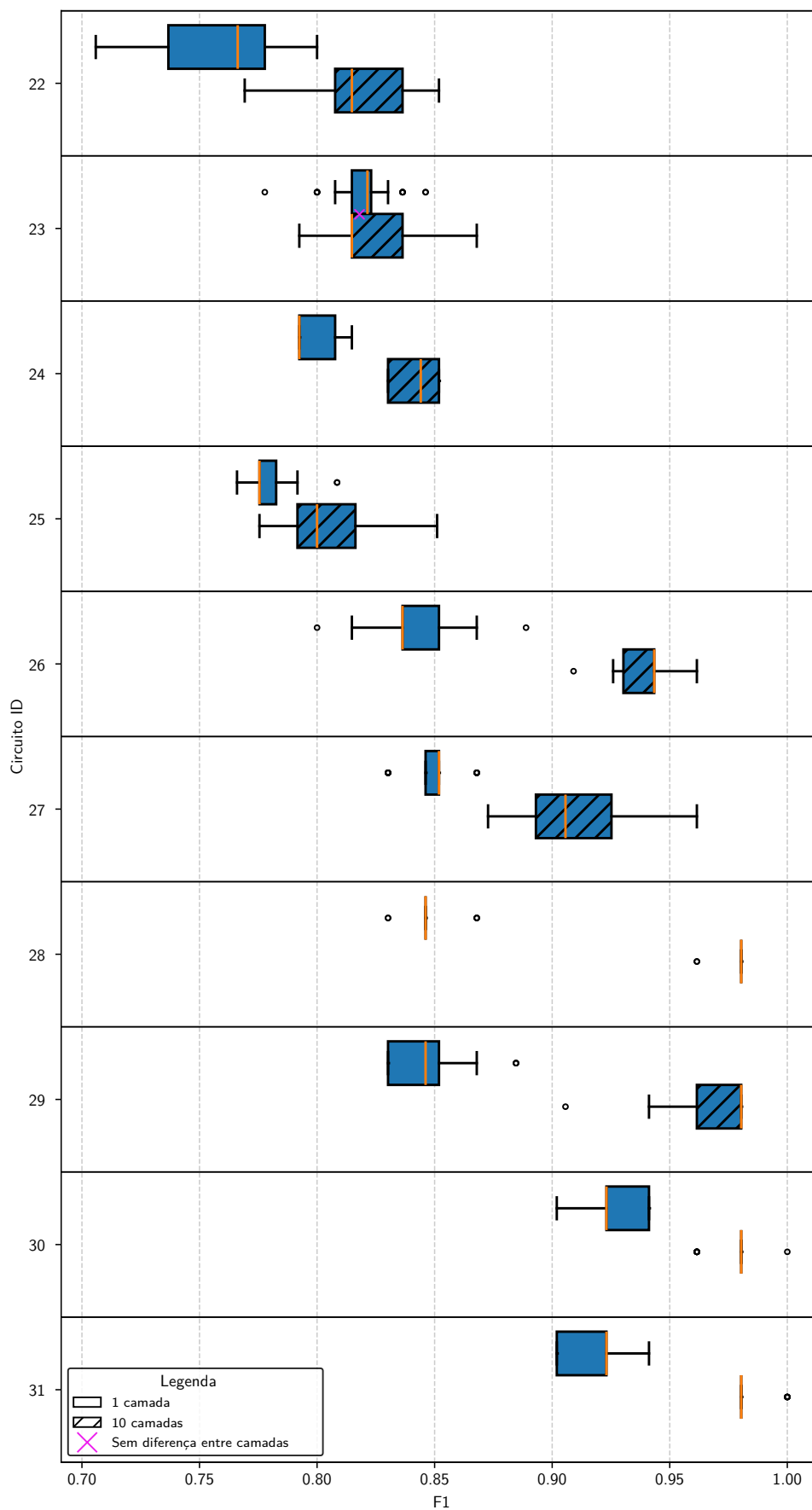
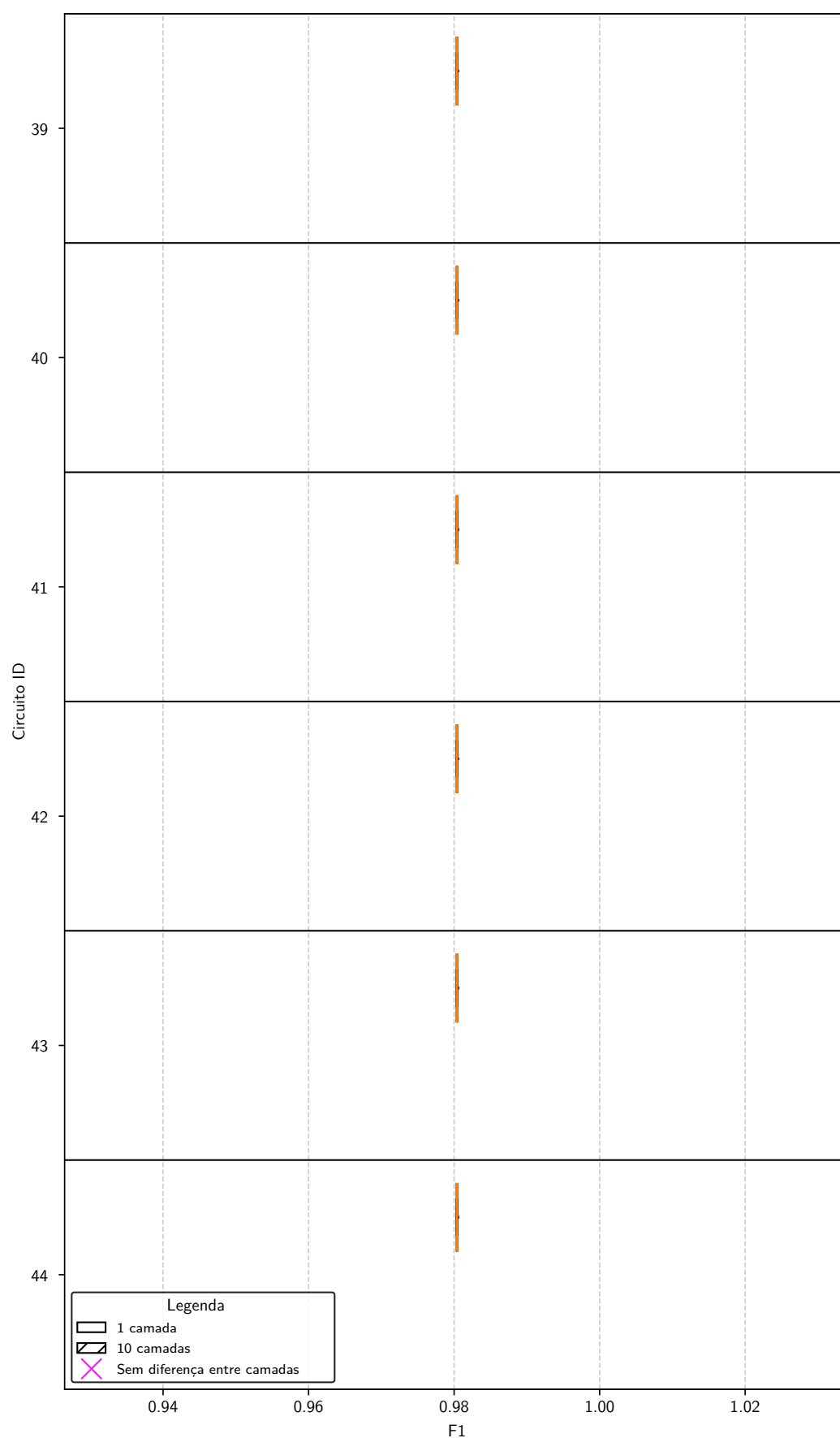


Figura 73 – Resultados do desempenho F1 dos circuitos 39 ao 44 na base *ChatGPT* - Difícil. Esses algoritmos classificadores implementam modelos de ensemble clássicos. As marcações que indicam semelhança estatística são advindas do teste pareado de Wilcoxon.



Fonte: Elaborada pelo autor (2025)

As paisagens de custo associadas aos são frequentemente complexas, repletas de mínimos locais e regiões de gradientes evanescentes, os supracitados *barren plateaus*. Embora o uso de otimizadores avançados como o Adam e o treinamento em múltiplos *seeds* busquem mitigar esse problema, a sua presença pode impedir que os circuitos convirjam para uma solução ótima, uma dificuldade que é exacerbada pela maior complexidade e pelo ruído intrínseco aos dados do mundo real da base SST.

Adicionalmente, a forma como a informação clássica é codificada no estado quântico é um fator que não pode ser subestimado. A estratégia de *amplitude embedding* utilizada neste trabalho, embora eficiente em termos de número de *qubits*, pode não ser a ideal para capturar as nuances semânticas contidas nos vetores de *embeddings*. É possível que esta codificação não estruture os dados de uma forma que permita ao circuito quântico explorar plenamente fenômenos como a superposição e o emaranhamento para a tarefa de classificação. A busca por métodos de codificação de dados mais intrinsecamente “quânticos”, que mapeiem o problema de forma a maximizar o potencial computacional do algoritmo, permanece como uma área de investigação ativa e relevante.

As características intrínsecas da base de dados SST também se apresentam como um fator contribuinte relevante para a dificuldade enfrentada pelos modelos quânticos. Conforme se pode observar na Tabela 1, a base SST possui um desbalanceamento de classes considerável, com um número de exemplos da classe negativa sendo aproximadamente quatro vezes maior que o da classe positiva. Modelos sensíveis à paisagem de otimização, como os circuitos quânticos variacionais, podem ter sua convergência prejudicada por esse desequilíbrio, tendendo a favorecer a classe majoritária e, consequentemente, resultando em valores mais baixos para a métrica F1, que é sensível tanto à precisão quanto à revocação. Os modelos clássicos, por sua vez, frequentemente incorporam mecanismos de regularização ou são inerentemente mais robustos a este tipo de desafio.

Em contrapartida ao desempenho na base SST, os resultados obtidos nas bases sintéticas *ChatGPT* oferecem uma perspectiva mais otimista e revelam que os modelos quânticos, sob certas condições, são altamente competitivos. Os diagramas de diferença crítica demonstram que, nestes cenários com dados mais estruturados e balanceados, a lacuna de desempenho entre os melhores modelos quânticos e os melhores modelos clássicos se estreita consideravelmente, e em muitos casos, desaparece.

Na base *ChatGPT* Fácil, por exemplo, o diagrama de diferença crítica posiciona o ensemble quântico 28 (*soft voting* dos circuitos 7 ao 11) com $L = 10$ na mesma clique de desempenho

Tabela 1 – Número de exemplos e média de palavras por exemplo por base de dados

Base	Tipo de base	Classe	Nº de exemplos	Média do nº de palavras
<i>ChatGPT</i> - Fácil	Treino	Positiva	100	6,95
<i>ChatGPT</i> - Fácil	Treino	Negativa	100	7,02
<i>ChatGPT</i> - Fácil	Teste	Positiva	25	7,48
<i>ChatGPT</i> - Fácil	Teste	Negativa	25	8,00
<i>ChatGPT</i> - Médio	Treino	Positiva	100	7,80
<i>ChatGPT</i> - Médio	Treino	Negativa	100	7,79
<i>ChatGPT</i> - Médio	Teste	Positiva	25	8,00
<i>ChatGPT</i> - Médio	Teste	Negativa	25	7,68
<i>ChatGPT</i> - Difícil	Treino	Positiva	100	9,48
<i>ChatGPT</i> - Difícil	Treino	Negativa	100	10,11
<i>ChatGPT</i> - Difícil	Teste	Positiva	25	9,88
<i>ChatGPT</i> - Difícil	Teste	Negativa	25	10,24
SST	Treino	Positiva	1.288	18,84
SST	Treino	Negativa	4.934	18,95
SST	Teste	Positiva	399	18,74
SST	Teste	Negativa	1.301	19,04

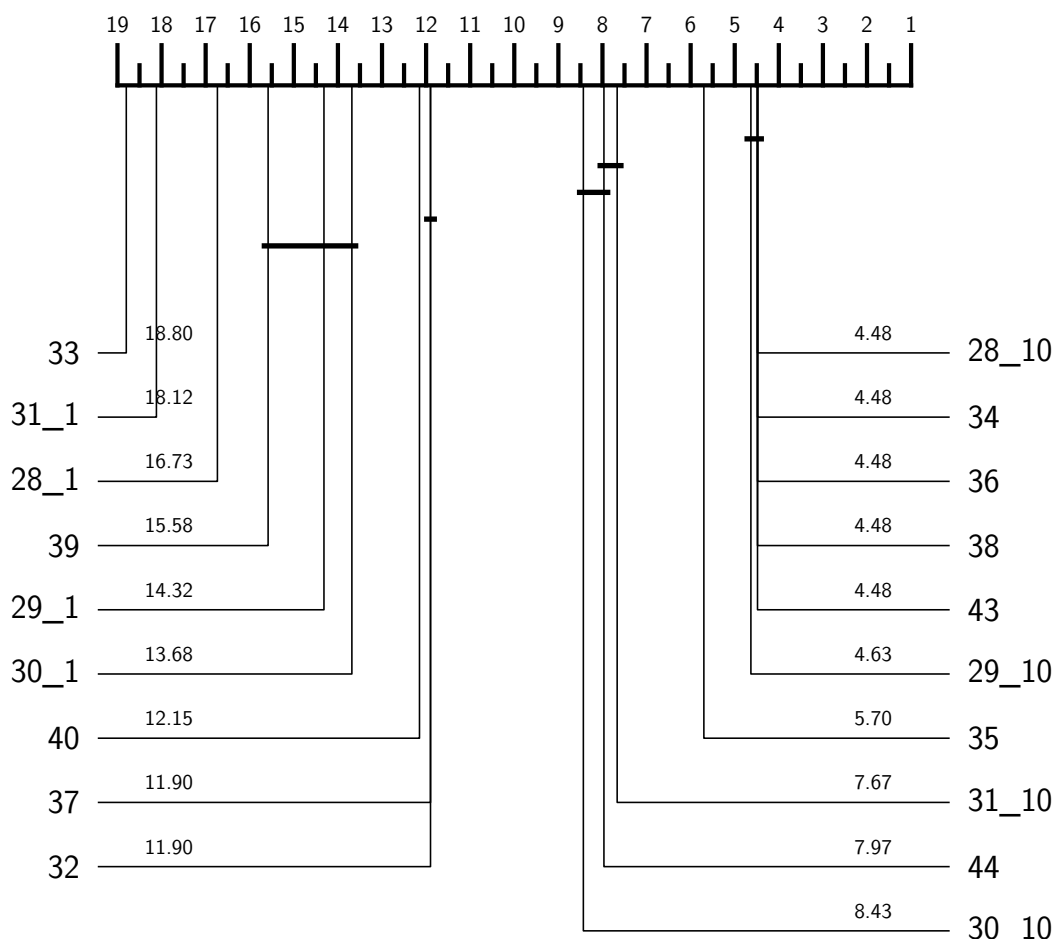
Fonte: Elaborada pelo autor (2025)

que os modelos clássicos de ponta, como o KNN (34), o SVM Linear (36) e o SVM RBF (38), indicando que seus desempenhos são estatisticamente indistinguíveis. De forma similar, na base *ChatGPT* Médio, diversos circuitos quânticos profundos, como o modelo 28, 29 (*hard voting* dos circuitos 7 ao 11) e 31 (*hard voting* dos circuitos 12, 14 e 15), todos com 10 camadas, compartilham a clique de melhor desempenho com os modelos clássicos mais robustos, incluindo a Regressão Logística (32) e as diferentes variações do SVM. Essas e outras constatações são advindas das Figuras 74, 75, 76 e 77.

Este padrão de competitividade se estende à base *ChatGPT* Difícil, onde o ensemble quântico 31 e $L = 10$ e o SVM Linear (36) se encontram no grupo de melhor desempenho, superando outros modelos clássicos e quânticos. A capacidade dos circuitos emaranhadores e profundos de igualar, e por vezes superar, modelos clássicos consolidados em cenários de dados sintéticos sugere que sua capacidade representacional é, de fato, elevada. Isso indica que a dificuldade observada na base SST não advém de uma incapacidade fundamental de aprendizado, mas sim de uma maior sensibilidade a fatores como ruído, desbalanceamento e a complexidade da fronteira de decisão de dados do mundo real.

Apesar da superioridade geral dos *benchmarks* clássicos na base SST, seria um equívoco

Figura 74 – Diagrama de diferença crítica da performance F1 na base *ChatGPT* - Fácil. Foram exibidos apenas cliques que contenham modelos clássicos, de forma *ensemble* ou isolada, e os modelos quânticos envolvidos (nas suas versões $L = 1$ e $L = 10$).

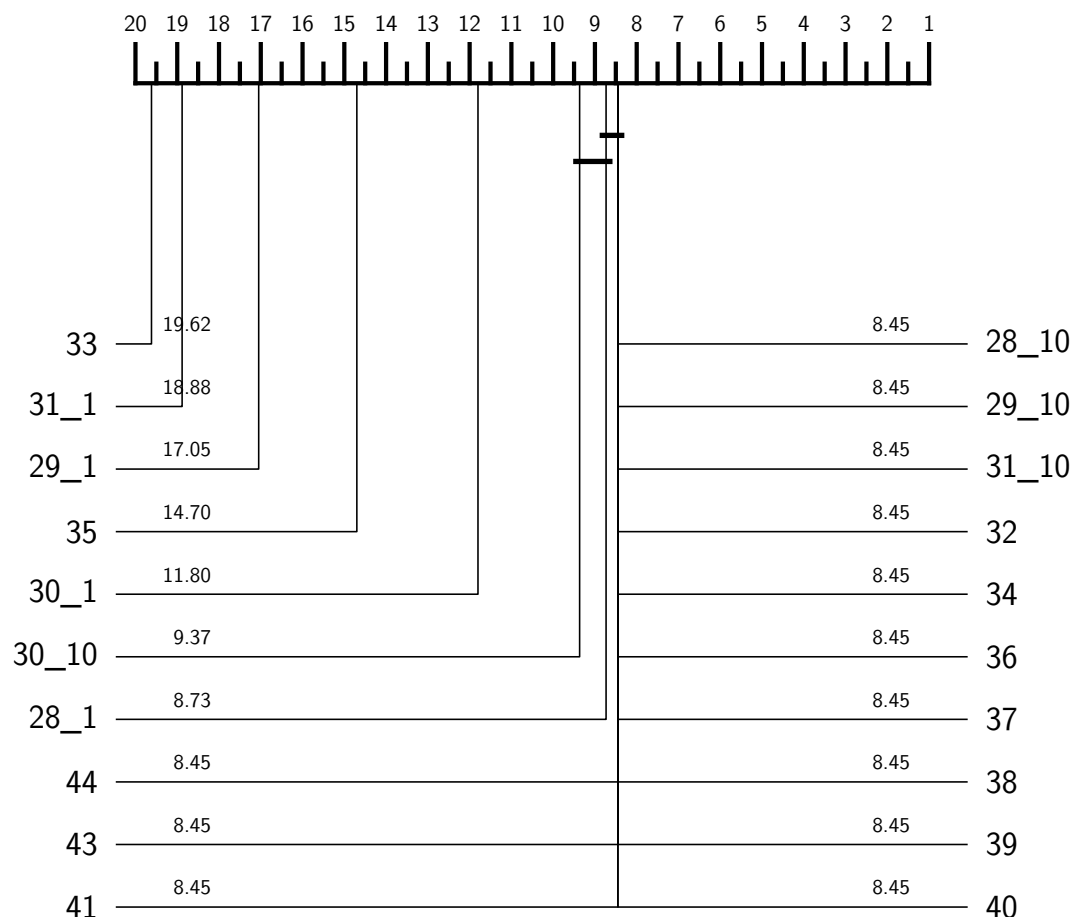


Fonte: Elaborada pelo autor (2025)

desconsiderar o potencial dos melhores modelos quânticos. Circuitos individuais como o 14 e o 15, na sua configuração com $L = 10$, alcançaram medianas de F1 de aproximadamente 0.50 e 0.47, respectivamente. De forma análoga, o ensemble quântico 21 (*bagging* do circuito 14) obteve uma mediana de F1 próxima a 0.48. Embora estes valores sejam inferiores aos obtidos pelos SVMs, eles são relevantes e se aproximam do desempenho de outros modelos clássicos, como o Perceptron (modelo 33, mediana de 0.69) e o KNN (modelo 34, mediana de 0.69). Isso indica que a abordagem quântica não é inviável, mas sim que exige arquiteturas sofisticadas e recursos computacionais (representados aqui pela profundidade) para se tornar competitiva.

Este desempenho dos melhores circuitos quânticos na base SST reforça um dos temas centrais desta dissertação, já explorado na seção 4.1: a combinação de um *ansatz* com ca-

Figura 75 – Diagrama de diferença crítica da performance F1 na base *ChatGPT* - Médio. Foram exibidos apenas cliques que contenham modelos clássicos, de forma *ensemble* ou isolada, e os modelos quânticos envolvidos (nas suas versões $L = 1$ e $L = 10$).

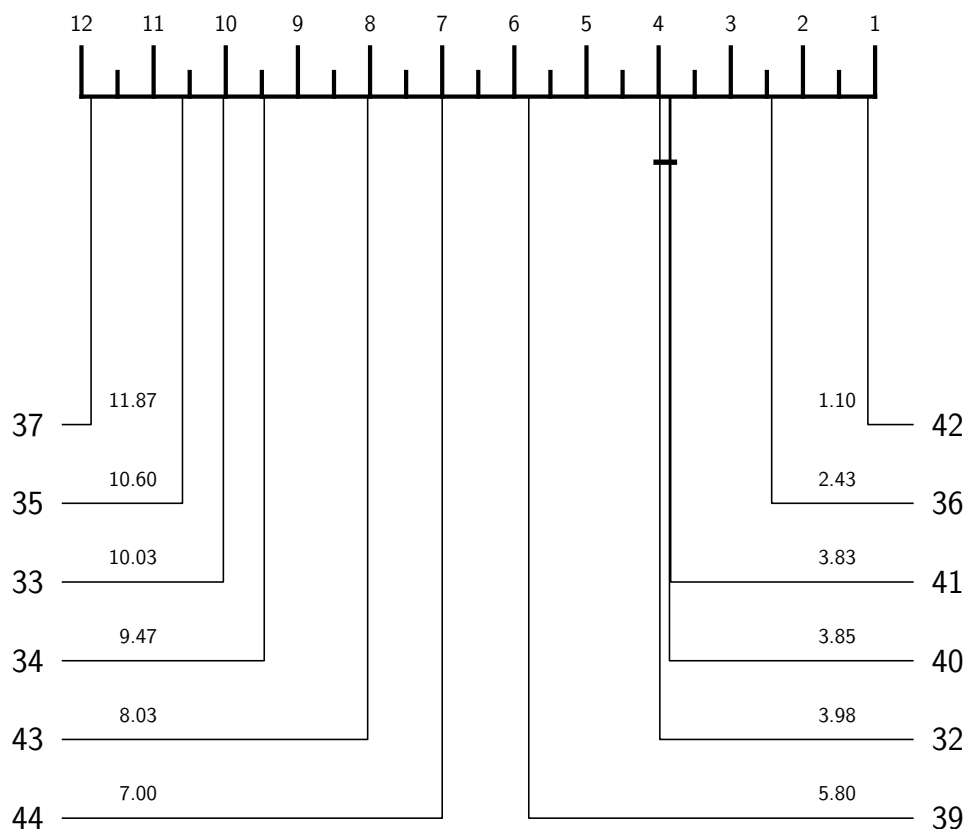


Fonte: Elaborada pelo autor (2025)

pacidade intrínseca de gerar emaranhamento e uma profundidade adequada é um requisito indispensável para o sucesso em tarefas de classificação de sentimentos com dados do mundo real. Os modelos que se destacaram (14, 15, 6, 9 e 11) são precisamente aqueles que, em suas configurações com $L = 10$, exibem as melhores métricas de expressabilidade e emaranhamento. O seu desempenho, ainda que não supere a dos melhores modelos clássicos, está em uma ordem de magnitude superior à dos circuitos não-emaranhadores, validando a hipótese de que estas propriedades quânticas são cruciais para o aprendizado de padrões complexos.

A análise comparativa com os modelos clássicos que utilizam técnicas de ensemble também oferece *insights* valiosos. Modelos como o 41 e o 42, que combinam diferentes kernels de SVM, alcançaram um desempenho de ponta na base SST (mediana de 0.73), igualando-se aos seus melhores componentes individuais. Isso demonstra a eficácia de agregar modelos clássicos

Figura 77 – Diagrama de diferença crítica da performance F1 na base SST. Foram exibidos apenas cliques que contenham modelos clássicos, de forma *ensemble* ou isolada, e os modelos quânticos envolvidos (nas suas versões $L = 1$ e $L = 10$).



Fonte: Elaborada pelo autor (2025)

dos circuitos quânticos, a natureza da codificação dos dados e as características da própria base SST. Contudo, o desempenho competitivo dos modelos quânticos em bases sintéticas e o desempenho relevante das arquiteturas quânticas mais complexas e profundas no cenário real indicam um caminho promissor. O trabalho posiciona o emaranhamento e a profundidade como condições necessárias, embora não suficientes, para que os modelos quânticos possam, futuramente, se tornar competitivos e, eventualmente, superar as abordagens clássicas em domínios de problemas de crescente complexidade.

5 CONCLUSÃO

Este trabalho se propôs a realizar uma investigação sistemática e experimental sobre a expressividade e o emaranhamento de circuitos quânticos variacionais aplicados à tarefa de classificação de sentimentos. Diante do crescente interesse no campo do Processamento de Linguagem Natural Quântico, buscou-se preencher uma lacuna na literatura, transitando de análises puramente teóricas para uma avaliação empírica que correlaciona as propriedades arquitetônicas dos circuitos — como topologia, profundidade e capacidade de emaranhamento — com seu desempenho prático em cenários de complexidade crescente.

Para alcançar este objetivo, delineou-se uma metodologia abrangente, detalhada no capítulo 3, que envolveu a análise de 15 arquiteturas de circuitos distintas, comparadas a 7 modelos clássicos e avaliadas em 22 configurações de *ensembles*. O estudo foi fundamentado em quatro bases de dados: três conjuntos de dados sintéticos de dificuldade crescente (*ChatGPT* Fácil, Médio e Difícil) e uma base de dados do mundo real, o SST. A investigação foi aprofundada pela análise do impacto de três modelos geradores de *embeddings* (*Matryoshka*, *Nomic* e *MPNet*) e pela avaliação da resiliência dos classificadores à redução da dimensionalidade do vetor de entrada. Todos os resultados foram validados por meio de testes estatísticos rigorosos, garantindo a robustez das conclusões subsequentes, que foram interpretadas em detalhe no Capítulo 4.

5.1 SÍNTESE DOS RESULTADOS E VALIDAÇÃO DOS OBJETIVOS

A análise agregada dos resultados, apresentada e discutida no Capítulo 4, permite validar o cumprimento dos objetivos delineados para esta dissertação. A seguir, detalha-se como cada meta específica foi alcançada, sintetizando os principais achados do estudo.

O primeiro objetivo, referente ao projeto e implementação de um conjunto diversificado de 15 arquiteturas de circuitos quânticos variacionais, foi integralmente cumprido através da metodologia descrita na seção 3.1.2. Cada uma das arquiteturas, que variaram em topologia, tipo de portas e capacidade de emaranhamento, foi avaliada em configurações de baixa ($L = 1$) e alta ($L = 10$) profundidade, formando a base para toda a análise empírica subsequente.

Em atendimento ao segundo objetivo, quantificaram-se as métricas teóricas para cada uma das arquiteturas propostas. A expressabilidade foi medida por meio da divergência de Kullback-

Leibler em relação à distribuição de Haar, enquanto o poder de emaranhamento foi avaliado utilizando a medida de Meyer-Wallach, conforme detalhado na seção 2.3.2. Estes cálculos foram fundamentais para a posterior correlação entre a teoria e o desempenho prático dos circuitos.

O terceiro objetivo, que visava investigar o impacto da representação de dados de entrada, foi sistematicamente investigado. Compararam-se três modelos geradores de *embeddings* (*Matryoshka*, *Nomic* e *MPNet*) e avaliou-se a resiliência dos modelos à redução de dimensionalidade (de 768 para 32 e 16 atributos), conforme a análise na seção 4.2.1. A análise demonstrou que, embora modelos clássicos e circuitos quânticos profundos exibam relevante resiliência à redução de dimensionalidade, os circuitos quânticos rasos mostraram-se mais sensíveis a essa perda de informação.

Para o quarto objetivo, avaliou-se o desempenho dos 44 modelos (quânticos e clássicos) em quatro bases de dados: as três bases sintéticas de complexidade crescente (*ChatGPT* Fácil, Médio e Difícil) e a base de dados pública SST. A métrica F1, utilizada como principal indicador de desempenho, permitiu uma comparação quantitativa e rigorosa do poder de classificação de cada abordagem nos diferentes cenários propostos.

Em cumprimento ao quinto objetivo, estabeleceu-se uma linha de base robusta por meio da implementação e avaliação de 7 modelos de aprendizagem de máquina clássicos e 10 configurações de *ensembles* clássicos. A comparação direta com as 15 arquiteturas quânticas individuais e os 16 *ensembles* quânticos, discutida na seção 4.3, revelou que os modelos clássicos, notadamente as Máquinas de Vetores de Suporte (modelos 36 e 38) e a Regressão Logística (modelo 32), apresentaram um desempenho superior na base SST, estabelecendo um *benchmark* de desempenho que não foi superado pelas abordagens quânticas investigadas.

A correlação entre o desempenho empírico e as métricas teóricas, cerne do sexto objetivo, revelou uma dependência fundamental do emaranhamento e da profundidade para o sucesso dos classificadores quânticos, conforme a discussão na seção 4.1. Circuitos não-emaranhadores (1 a 5 e 7) tiveram performance limitada, especialmente na base SST. Em contrapartida, os circuitos emaranhadores (como os modelos 6, 9, 11, 14 e 15), especialmente na configuração com $L = 10$, foram os únicos a alcançar resultados relevantes, validando a hipótese de que a capacidade de gerar correlações quânticas complexas é um requisito para a classificação de dados do mundo real.

Finalmente, o sétimo objetivo, que demandava uma validação estatística rigorosa, foi cumprido por meio da repetição de cada experimento com 30 sementes de aleatoriedade e da apli-

cação do teste pareado de Wilcoxon com a correção de Holm-Bonferroni. Esta metodologia, utilizada em todas as análises comparativas, garantiu que as conclusões sobre as diferenças de desempenho entre os modelos fossem estatisticamente significativas e confiáveis, conferindo robustez aos achados da dissertação.

5.2 LIMITAÇÕES DA PESQUISA

O reconhecimento das limitações de um estudo é um passo fundamental para a contextualização de seus achados e para o delineamento de futuras investigações. A presente pesquisa, embora abrangente em seu escopo, possui certas restrições que merecem ser explicitadas.

Primeiramente, no que tange à representação dos dados, a análise principal de desempenho e o estudo de redução de dimensionalidade concentraram-se nos vetores gerados pelo modelo *Matryoshka*. Embora a comparação com os modelos *Nomic* e *MPNet* tenha fornecido *insights* valiosos sobre a sensibilidade dos classificadores, uma investigação mais aprofundada com outros modelos de *embeddings* de ponta poderia revelar dinâmicas distintas. Adicionalmente, o trabalho se limitou ao uso da codificação por *amplitude embedding*, sendo esta apenas uma das múltiplas estratégias possíveis para mapear dados clássicos em estados quânticos. Métodos alternativos poderiam interagir de maneiras diferentes com os *ansatze* propostos e, conseqüentemente, alterar os resultados.

Outra limitação relevante reside nas características da base de dados SST, que possui um desbalanceamento considerável entre as classes positiva e negativa, conforme detalhado na Tabela 1. Embora a métrica F1 seja adequada para lidar com tal situação, sabe-se que o desbalanceamento pode dificultar o processo de treinamento e a convergência de modelos sensíveis à paisagem de custo, como é o caso dos circuitos variacionais. A aplicação de técnicas de reamostragem, como *SMOTE* ou *undersampling*, não foi contemplada no escopo deste trabalho, mas poderia potencialmente alterar o desempenho dos modelos.

Adicionalmente, todos os experimentos foram conduzidos em um simulador quântico ideal, que não leva em consideração os efeitos do ruído. O desempenho dos circuitos em um hardware quântico da era NISQ, sujeito a erros de portas, decoerência e outros fenômenos, representaria um desafio adicional e, provavelmente, resultaria em um desempenho inferior. A implementação e a avaliação de técnicas de mitigação de erro, embora cruciais para a aplicação prática, fugiram ao escopo desta dissertação, que se concentrou na capacidade representacional intrínseca das arquiteturas.

Por fim, a otimização de hiperparâmetros, tanto para os modelos quânticos (por exemplo, taxa de aprendizado, otimizador, número de épocas) quanto para os modelos clássicos (parâmetros de regularização, grau do polinômio do *kernel*, entre outros), foi mantida com configurações fixas ou padrão para garantir uma comparação justa entre as arquiteturas. Uma busca exaustiva por hiperparâmetros ótimos para cada modelo poderia, potencialmente, alterar os rankings de desempenho, mas representaria um custo computacional custoso e desviaria o foco da análise arquitetônica central deste trabalho.

5.3 TRABALHOS FUTUROS

As limitações identificadas nesta pesquisa abrem caminho para diversas e promissoras avenidas de investigação futura, que podem aprofundar e expandir os achados aqui apresentados. Uma direção natural consiste na validação experimental dos resultados em hardware quântico real. A execução dos circuitos mais promissores, como os modelos 14 e 15 com $L = 10$, em diferentes plataformas de computadores quânticos, aliada à incorporação de técnicas de mitigação de erro, permitiria uma avaliação mais realista de sua viabilidade e desempenho em um ambiente com ruído.

Outra vertente de pesquisa relevante reside na exploração de outras estratégias de codificação de dados. A investigação de métodos alternativos, como o *Angle Embedding* ou o *Dense Angle Embedding*, e a análise de sua interação com as métricas de expressabilidade e emaranhamento poderiam revelar combinações de codificação e *ansatz* mais eficazes para tarefas de processamento quântico de análise de sentimento, potencialmente alterando a hierarquia de desempenho observada neste trabalho.

Além disso, a análise dos *ensembles* revelou que as abordagens de votação simples podem ser prejudicadas pela presença de estimadores de menor desempenho. Isso sugere a necessidade de desenvolver métodos de *ensemble* quânticos mais sofisticados. Uma proposta de trabalho futuro seria o desenvolvimento de esquemas de votação ponderada, nos quais a contribuição de cada estimador base, seja no *Soft* ou no *Hard Voting*, é ajustada por seu desempenho individual ou por uma métrica de confiança, a fim de evitar a diluição do desempenho final pelos modelos mais fracos. Além disso, uma análise da quantidade de atributos também é interessante, tendo em vista modelos quânticos com performances similares a modelos clássicos. Isso pode indicar uma superioridade quântica no que tange à necessidade de complexidade, sabendo que esses algoritmos quânticos requisitaram menos parâmetros treináveis para atingir o mesmo

desempenho.

Por fim, a metodologia de análise sistemática de circuitos aqui desenvolvida pode ser aplicada a um escopo mais amplo de problemas de processamento de texto. Trabalhos futuros poderiam investigar o desempenho destas e de outras arquiteturas quânticas em tarefas mais complexas, como a classificação de textos com múltiplas classes, o reconhecimento de entidade nomeada ou a inferência de linguagem natural. Tais estudos seriam fundamentais para se obter uma compreensão mais completa do domínio de problemas em que os classificadores quânticos variacionais podem, eventualmente, oferecer uma vantagem em relação às abordagens clássicas.

REFERÊNCIAS

WILCOXON, Frank. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, The Royal Society, v. 439, n. 1907, p. 553–558, dez. 1992. ISSN 2053-9177. Disponível em: <<http://dx.doi.org/10.1098/rspa.1992.0167>>.

ABBAS, A.; SUTTER, D.; ZOUFAL, C.; LUCCHI, A.; FIGALLI, A.; STEFANINI, S. The power of quantum neural networks. *Nature Computational Science*, Nature Publishing Group, v. 1, n. 6, p. 403–409, 2021.

ALTMAN, N. S. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, Informa UK Limited, v. 46, n. 3, p. 175–185, ago. 1992. ISSN 1537-2731. Disponível em: <<http://dx.doi.org/10.1080/00031305.1992.10475879>>.

ARUTE, F.; ARYA, K.; BABBUS, R.; BACON, D.; BARDIN, J. C.; BAREND, R.; BISWAS, R.; BOIXO, S.; BRANDAO, F. G. S. L.; BUELL, D. A.; BURKETT, B.; CHEN, Y.; CHEN, Z.; CHIARO, B.; COLLINS, R.; COURTNEY, W.; DUNSWORTH, A.; FARHI, E.; FOXEN, B.; FOWLER, A.; GIDNEY, C.; GIUSTINA, M.; GRAFF, R.; GUERIN, K.; HABEGGER, S.; HARRIGAN, M. P.; HARTMANN, M. J.; HO, A.; HOFFMANN, M.; HUANG, T.; HUMBLE, T. S.; ISAKOV, S. V.; JEFFREY, E.; JIANG, Z.; KAFRI, D.; KECHEDZHI, K.; KELLY, J.; KLIMOV, P. V.; KNYSH, S.; KOROTKOV, A.; KOSTRITSKA, F.; LANDHUIS, D.; LINDMARK, M.; LUCERO, E.; LYAKH, D.; MANDRÀ, S.; MCCLEAN, J. R.; MCEWEN, M.; MEGRANT, A.; MI, X.; MICHELSEN, K.; MOHSENI, M.; MUTUS, J.; NAAMAN, O.; NEELEY, M.; NEILL, C.; NIU, M. Y.; OSTBY, E.; PETUKHOV, A.; PLATT, J. C.; QUINTANA, C.; RIEFFEL, E. G.; ROUSHAN, P.; RUBIN, N. C.; SANK, D.; SATZINGER, K. J.; SMELYANSKIY, V.; SUNG, K. J.; TREVITHICK, M. D.; VAINSENER, A.; VILLALONGA, B.; WHITE, T.; YAO, Z. J.; YEH, P.; ZALCMAN, A.; NEVEN, H.; MARTINIS, J. M. Quantum supremacy using a programmable superconducting processor. *Nature*, Springer Science and Business Media LLC, v. 574, n. 7779, p. 505–510, out. 2019. ISSN 1476-4687. Disponível em: <<http://dx.doi.org/10.1038/s41586-019-1666-5>>.

BELL, J. S. On the einstein podolsky rosen paradox. *Physics Physique*, American Physical Society (APS), v. 1, n. 3, p. 195–200, nov. 1964. ISSN 0554-128X. Disponível em: <<http://dx.doi.org/10.1103/PhysicsPhysiqueFizika.1.195>>.

BENEDETTI, M.; LLOYD, E.; SACK, S.; FIORENTINI, M. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, IOP Publishing, v. 4, n. 4, p. 043001, nov. 2019. ISSN 2058-9565. Disponível em: <<http://dx.doi.org/10.1088/2058-9565/ab4eb5>>.

BERGHOLM, V.; IZAAC, J.; SCHULD, M.; GOGOLIN, C.; AHMED, S.; AJITH, V.; ALAM, M. S.; ALONSO-LINAJE, G.; AKASHNARAYANAN, B.; ASADI, A.; ARRAZOLA, J. M.; AZAD, U.; BANNING, S.; BLANK, C.; BROMLEY, T. R.; CORDIER, B. A.; CERONI, J.; DELGADO, A.; MATTEO, O. D.; DUSKO, A.; GARG, T.; GUALA, D.; HAYES, A.; HILL, R.; IJAZ, A.; ISACSSON, T.; ITTAH, D.; JAHANGIRI, S.; JAIN, P.; JIANG, E.; KHANDELWAL, A.; KOTTMANN, K.; LANG, R. A.; LEE, C.; LOKE, T.; LOWE, A.; MCKIERNAN, K.; MEYER, J. J.; MONTAÑEZ-BARRERA, J. A.; MOYARD, R.; NIU, Z.; O'RIORDAN, L. J.; OUD, S.; PANIGRAHI, A.; PARK, C.-Y.; POLATAJKO, D.; QUESADA, N.; ROBERTS, C.; SÁ, N.; SCHOCH, I.; SHI, B.; SHU, S.; SIM, S.; SINGH, A.; STRANDBERG, I.; SONI, J.; SZÁVA, A.; THABET, S.; VARGAS-HERNÁNDEZ, R. A.; VINCENT, T.; VITUCCI, N.;

WEBER, M.; WIERICHS, D.; WIERSEMA, R.; WILLMANN, M.; WONG, V.; ZHANG, S.; KILLORAN, N. *PennyLane: Automatic differentiation of hybrid quantum-classical computations*. 2022. Disponível em: <<https://arxiv.org/abs/1811.04968>>.

BIAMONTE, J.; WITTEK, P.; PANCOTTI, N.; REBENTROST, P.; WIEBE, N.; LLOYD, S. Quantum machine learning. *Nature*, Springer Science and Business Media LLC, v. 549, n. 7671, p. 195–202, set. 2017. ISSN 1476-4687. Disponível em: <<http://dx.doi.org/10.1038/nature23474>>.

BIEWALD, L. *Experiment Tracking with Weights and Biases*. 2020. Software available from wandb.com. Disponível em: <<https://www.wandb.com/>>.

BORN, M. Zur quantenmechanik der stoßvorgänge. *Zeitschrift für Physik*, Springer Science and Business Media LLC, v. 37, n. 12, p. 863–867, dez. 1926. ISSN 1434-601X. Disponível em: <<http://dx.doi.org/10.1007/BF01397477>>.

BREIMAN, L. Bagging predictors. *Machine Learning*, Springer Science and Business Media LLC, v. 24, n. 2, p. 123–140, ago. 1996. ISSN 1573-0565. Disponível em: <<http://dx.doi.org/10.1007/BF00058655>>.

BREIMAN, L. Random forests. *Machine Learning*, Springer Science and Business Media LLC, v. 45, n. 1, p. 5–32, out. 2001. ISSN 1573-0565. Disponível em: <<http://dx.doi.org/10.1023/A:1010933404324>>.

BRITO, E. B.; NETO, F. M. de P.; BERNARDES, N. K. Quantum classifier based on open quantum systems with amplitude information loading. *Quantum Information Processing*, Springer, v. 23, n. 4, p. 139, 2024.

CARVALHO, J. H. A. de; NETO, F. M. de P. Parametrized constant-depth quantum neuron. *IEEE Transactions on Quantum Engineering*, IEEE, v. 4, p. 1–11, 2023.

CEREZO, M.; ARRASMITH, A.; BABBUSH, R.; BENJAMIN, S. C.; ENDO, S.; FUJII, K.; MCCLEAN, J. R.; MITARAI, K.; YUAN, X.; CINCIO, L.; COLES, P. J. Variational quantum algorithms. *Nature Reviews Physics*, Springer Science and Business Media LLC, v. 3, n. 9, p. 625–644, ago. 2021. ISSN 2522-5820. Disponível em: <<http://dx.doi.org/10.1038/s42254-021-00348-9>>.

CEREZO, M.; SONE, A.; VOLKOFF, T.; CINCIO, L.; COLES, P. J. Cost function dependent barren plateaus in shallow quantum neural networks. *Physical Review Research*, APS, v. 2, n. 3, p. 033125, 2020.

CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, Springer Science and Business Media LLC, v. 20, n. 3, p. 273–297, set. 1995. ISSN 1573-0565. Disponível em: <<http://dx.doi.org/10.1007/BF00994018>>.

COSTA, M. H. J. da; NETO, F. M. de P. Evolutionary weight optimization for variational quantum classifiers. In: IEEE. *2025 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2025.

DEMsAR, J. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.*, JMLR.org, v. 7, p. 1–30, dez. 2006. ISSN 1532-4435.

EINSTEIN, A.; BORN, M. *A Correspondência de Born e Einstein: Amizade, política e física quântica*. Rio de Janeiro: Zahar, 2005. A citação específica é de uma carta de Einstein para Born (4 de dezembro de 1926), contida nesta correspondência compilada.

EINSTEIN, A.; PODOLSKY, B.; ROSEN, N. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, American Physical Society (APS), v. 47, n. 10, p. 777–780, maio 1935. ISSN 0031-899X. Disponível em: <<http://dx.doi.org/10.1103/PhysRev.47.777>>.

FARHI, E.; NEVEN, H. Classification with quantum neural networks on near term processors. *arXiv preprint arXiv:1802.06002*, 2018.

FEYNMAN, R. P. Simulating physics with computers. *International Journal of Theoretical Physics*, Springer Science and Business Media LLC, v. 21, n. 6–7, p. 467–488, jun. 1982. ISSN 1572-9575. Disponível em: <<http://dx.doi.org/10.1007/BF02650179>>.

FILHO, S. d. V. *sdvfh/qnlp: Quantum Natural Language Processing*. 2025. <<https://github.com/sdvfh/qnlp>>. GitHub repository; última atualização (commit) em 26 Apr 2025; acesso em 16 May 2025.

FREUND, Y.; SCHAPIRE, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, Elsevier BV, v. 55, n. 1, p. 119–139, ago. 1997. ISSN 0022-0000. Disponível em: <<http://dx.doi.org/10.1006/jcss.1997.1504>>.

GILI, K.; FOULDS, J. R.; MARTI, H.; GIROLAMI, M. Do we need deep quantum neural networks? *arXiv preprint arXiv:2312.02279*, 2023.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016.

GROVER, L. K. A fast quantum mechanical algorithm for database search. In: *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*. [S.l.: s.n.], 1996. p. 212–219.

HALLER, S.; ALDEA, A.; SEIFERT, C.; STRISCIUGLIO, N. *Survey on Automated Short Answer Grading with Deep Learning: from Word Embeddings to Transformers*. 2022. Disponível em: <<https://arxiv.org/abs/2204.03503>>.

HARRIS, C. R.; MILLMAN, K. J.; WALT, S. J. van der; GOMMERS, R.; VIRTANEN, P.; COUNAPEAU, D.; WIESER, E.; TAYLOR, J.; BERG, S.; SMITH, N. J.; KERN, R.; PICUS, M.; HOYER, S.; KERKWIJK, M. H. van; BRETT, M.; HALDANE, A.; RÍO, J. F. del; WIEBE, M.; PETERSON, P.; GÉRARD-MARCHANT, P.; SHEPPARD, K.; REDDY, T.; WECKESSER, W.; ABBASI, H.; GOHLKE, C.; OLIPHANT, T. E. Array programming with NumPy. *Nature*, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020. Disponível em: <<https://doi.org/10.1038/s41586-020-2649-2>>.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning*. Springer New York, 2009. ISSN 2197-568X. ISBN 9780387848587. Disponível em: <<http://dx.doi.org/10.1007/978-0-387-84858-7>>.

HAVLÍČEK, V.; CÓRCOLES, A. D.; TEMME, K.; HARROW, A. W.; KANDALA, A.; CHOW, J. M.; GAMBETTA, J. M. Supervised learning with quantum-enhanced feature spaces. *Nature*, Nature Publishing Group, v. 567, n. 7747, p. 209–212, 2019.

- HEIN, M.; EISERT, J.; BRIEGEL, H. J. Multiparty entanglement in graph states. *Physical Review A*, American Physical Society (APS), v. 69, n. 6, jun. 2004. ISSN 1094-1622. Disponível em: <<http://dx.doi.org/10.1103/PhysRevA.69.062311>>.
- HOLM, S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, v. 6, p. 65–70, 1979. Disponível em: <<https://api.semanticscholar.org/CorpusID:122415379>>.
- HOLMES, Z.; SHARMA, K.; CERESO, M.; COLES, P. J. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum*, American Physical Society (APS), v. 3, n. 1, jan. 2022. ISSN 2691-3399. Disponível em: <<http://dx.doi.org/10.1103/PRXQuantum.3.010313>>.
- HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007.
- JOHNSON, P. D.; ROMERO, J.; OLSON, J.; CAO, Y.; ASPURU-GUZI, A. QVECTOR: an algorithm for device-tailored quantum error correction. *arXiv preprint arXiv:1711.02249*, 2017.
- JURAFSKY, D.; MARTIN, J. *Speech and Language Processing*. 2. ed. Upper Saddle River, NJ: Pearson, 2008.
- KARTSAKLIS, D.; FAN, I.; YEUNG, R.; PEARSON, A.; LORENZ, R.; TOUMI, A.; FELICE, G. de; MEICHANETZIDIS, K.; CLARK, S.; COECKE, B. *lambeq: An Efficient High-Level Python Library for Quantum NLP*. 2021. Disponível em: <<https://arxiv.org/abs/2110.04236>>.
- KINGMA, D. P.; BA, J. *Adam: A Method for Stochastic Optimization*. 2017. Disponível em: <<https://arxiv.org/abs/1412.6980>>.
- KOCKUM, A. F.; NORI, F. Quantum bits with josephson junctions. In: _____. *Fundamentals and Frontiers of the Josephson Effect*. Springer International Publishing, 2019. p. 703–741. ISBN 9783030207267. Disponível em: <http://dx.doi.org/10.1007/978-3-030-20726-7_17>.
- KUSUPATI, A.; BHATT, G.; REGE, A.; WALLINGFORD, M.; SINHA, A.; RAMANUJAN, V.; HOWARD-SNYDER, W.; CHEN, K.; KAKADE, S.; JAIN, P.; FARHADI, A. *Matryoshka Representation Learning*. 2024. Disponível em: <<https://arxiv.org/abs/2205.13147>>.
- LIU, B. *Sentiment Analysis and Opinion Mining*. Springer International Publishing, 2012. ISSN 1947-4059. ISBN 9783031021459. Disponível em: <<http://dx.doi.org/10.1007/978-3-031-02145-9>>.
- MAOUAKI, W. E.; MARCHISIO, A.; SAID, T.; SHAFIQUE, M.; BENNAI, M. *Designing Robust Quantum Neural Networks via Optimized Circuit Metrics*. 2025. Disponível em: <<https://arxiv.org/abs/2411.11870>>.
- MCCLEAN, J. R.; BOIXO, S.; SMELYANSKIY, V. N.; BABUSH, R.; NEVEN, H. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, Springer Science and Business Media LLC, v. 9, n. 1, nov. 2018. ISSN 2041-1723. Disponível em: <<http://dx.doi.org/10.1038/s41467-018-07090-4>>.

- MIDDLEHURST, M.; ISMAIL-FAWAZ, A.; GUILLAUME, A.; HOLDER, C.; RUBIO, D. G.; BULATOVA, G.; TSAPROUNIS, L.; MENDEL, L.; WALTER, M.; SCHÄFER, P.; BAGNALL, A. *aeon: a Python toolkit for learning from time series*. 2024. Disponível em: <<https://arxiv.org/abs/2406.14231>>.
- MIKOLOV, T.; CHEN, K.; CORRADO, G.; DEAN, J. *Efficient Estimation of Word Representations in Vector Space*. 2013. Disponível em: <<https://arxiv.org/abs/1301.3781>>.
- MOTTONEN, M.; VARTIAINEN, J. J.; BERGHOLM, V.; SALOMAA, M. M. *Transformation of quantum states using uniformly controlled rotations*. 2004. Disponível em: <<https://arxiv.org/abs/quant-ph/0407010>>.
- NETO, F. M. de P. *Data Complexity Measures for Quantum Circuits Architecture Recommendation*. 2025. Preprint or in preparation.
- NIELSEN, M. A.; CHUANG, I. L. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2012. ISBN 9780511976667. Disponível em: <<http://dx.doi.org/10.1017/CBO9780511976667>>.
- NUSSBAUM, Z.; MORRIS, J. X.; DUDERSTADT, B.; MULYAR, A. Nomic embed: Training a reproducible long context text embedder. *arXiv preprint arXiv:2402.01613*, 2024. Disponível em: <<https://arxiv.org/abs/2402.01613>>.
- OPENAI. *ChatGPT [Large language model, versão fevereiro 15, 2025]*. 2025. <<https://chat.openai.com>>. Acesso em 15 de fevereiro de 2025.
- PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; MÜLLER, A.; NOTHMAN, J.; LOUPPE, G.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY Édouard. *Scikit-learn: Machine Learning in Python*. 2018. Disponível em: <<https://arxiv.org/abs/1201.0490>>.
- PERUZZO, A.; MCCLEAN, J.; SHADBOLT, P.; YUNG, M.-H.; ZHOU, X.-Q.; LOVE, P. J.; ASPURU-GUZZIK, A.; O'BRIEN, J. L. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, Springer Science and Business Media LLC, v. 5, n. 1, jul. 2014. ISSN 2041-1723. Disponível em: <<http://dx.doi.org/10.1038/ncomms5213>>.
- PRESKILL, J. Quantum computing in the nisq era and beyond. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 2, p. 79, 2018.
- RAIMAN, J. *pytreebank: Stanford Sentiment Treebank loader in Python*. 2020. <<https://github.com/JonathanRaiman/pytreebank>>. Último commit em 18 de fevereiro de 2020. Acesso em 15 de maio de 2025.
- REIMERS, N.; GUREVYCH, I. Sentence-bert: Sentence embeddings using siamese bert-networks. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Association for Computational Linguistics, 2019. Disponível em: <<http://dx.doi.org/10.18653/v1/D19-1410>>.
- REIMERS, N.; GUREVYCH, I. Sentence-bert: Sentence embeddings using siamese bert-networks. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2019. Disponível em: <<http://arxiv.org/abs/1908.10084>>.

RIJSBERGEN, C. J. V. *Information Retrieval*. 2. ed. Oxford, England: Butterworth-Heinemann, 1979.

ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, American Psychological Association (APA), v. 65, n. 6, p. 386–408, 1958. ISSN 0033-295X. Disponível em: <<http://dx.doi.org/10.1037/h0042519>>.

ROSSUM, G. V.; JR, F. L. D. *Python reference manual*. [S.l.]: Centrum voor Wiskunde en Informatica Amsterdam, 1995.

SCHRÖDINGER, E. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press (CUP), v. 31, n. 4, p. 555–563, out. 1935. ISSN 1469-8064. Disponível em: <<http://dx.doi.org/10.1017/S0305004100013554>>.

SCHULD, M. What is the computational advantage of quantum machine learning? *Nature Reviews Physics*, Nature Publishing Group, v. 3, n. 9, p. 609–611, 2021.

SCHULD, M.; BOCHAROV, A.; SVORE, K. M.; WIEBE, N. Circuit-centric quantum classifiers. *Physical Review A*, APS, v. 101, n. 3, p. 032308, 2020.

SCHULD, M.; PETRUCCIONE, F. *Supervised Learning with Quantum Computers*. Springer International Publishing, 2018. ISSN 2364-9062. ISBN 9783319964249. Disponível em: <<http://dx.doi.org/10.1007/978-3-319-96424-9>>.

SCHULD, M.; PETRUCCIONE, F. *Machine Learning with Quantum Computers*. Springer International Publishing, 2021. ISSN 2364-9062. ISBN 9783030830984. Disponível em: <<http://dx.doi.org/10.1007/978-3-030-83098-4>>.

SEABOLD, S.; PERKTOLD, J. statsmodels: Econometric and statistical modeling with python. In: *9th Python in Science Conference*. [S.l.: s.n.], 2010.

SHOR, P. W. Algorithms for quantum computation: discrete logarithms and factoring. In: IEEE. *Proceedings 35th Annual Symposium on Foundations of Computer Science*. [S.l.], 1994. p. 124–134.

SIM, S.; JOHNSON, P. D.; ASPURU-GUZZIK, A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, Wiley, v. 2, n. 12, out. 2019. ISSN 2511-9044. Disponível em: <<http://dx.doi.org/10.1002/qute.201900070>>.

SOCHER, R.; PERELYGIN, A.; WU, J.; CHUANG, J.; MANNING, C.; NG, A.; POTTS, C. Recursive deep models for semantic compositionality over a sentiment treebank. *EMNLP*, v. 1631, p. 1631–1642, 01 2013.

SONG, K.; TAN, X.; QIN, T.; LU, J.; LIU, T.-Y. *MPNet: Masked and Permuted Pre-training for Language Understanding*. 2020. Disponível em: <<https://arxiv.org/abs/2004.09297>>.

SOUSA, P. B. M.; RAMOS, R. V. Programmable quantum circuits. *arXiv preprint quant-ph/0602174*, 2006.

TEAM, T. pandas development. *pandas-dev/pandas: Pandas*. Zenodo, 2024. Disponível em: <<https://doi.org/10.5281/zenodo.13819579>>.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. *Attention Is All You Need*. 2023. Disponível em: <<https://arxiv.org/abs/1706.03762>>.

WILCOXON, F. Individual comparisons by ranking methods. *Biometrics Bulletin*, v. 1, n. 6, p. 80–83, dez. 1945.

WILSON, C. M.; OTTERBACH, J. S.; TEZAK, N.; SMITH, R. S.; CROOKS, G. E.; SILVA, M. P. da. Quantum kitchen sinks: An algorithm for machine learning on near-term quantum computers. *arXiv preprint arXiv:1806.08321*, 2018.