

# RACIOCÍNIO DAS LLMs – Grupo 1

## INTRODUÇÃO

O objetivo desse relatório é explicar as estratégias de raciocínio das LLMs Ajuste Fino Totalmente Supervisionado, In-Context Learning, Zero-shot, One-shot, Few-shot, Chain-of-Thought Prompting, Self-Consistency Sampling, Decomposição de Problemas, Least-to-Most Prompting, Successive Prompting e Tool-Augmented Reasoning.

Na sua elaboração foram utilizadas as LLMs ChatGPT 4.0, Claude, Gemini 2.5 Pro, Deepseek, Grok e Qwen

Para cada uma das estratégias existirá as seguintes seções: explicação, prompt, resposta gerada e referência simulada, uma tabela comparativa, um gráfico de barras com a métrica de precisão e uma Conclusão Geral quanto a precisão para cada LLM.

---

## 1. Ajuste Fino Totalmente Supervisionado (Simulado)

**Pergunta do usuário:** Como fazer uma deliciosa pizza quatro queijos?

### **Explicação:**

Esta estratégia simula um modelo treinado com dados específicos e anotados sobre pizzas quatro queijos. O modelo foi ajustado para gerar respostas precisas com base em receitas verificadas e avaliadas por especialistas culinários.

### **Prompt:**

"Como preparar uma pizza quatro queijos deliciosa, usando técnicas culinárias clássicas italianas, com base em dados supervisionados de receitas?"

### **Resposta:**

Ingredientes:

- 300 g de massa de pizza (caseira ou comprada)
- 100 g de mussarela
- 100 g de gorgonzola
- 100 g de parmesão ralado
- 100 g de queijo provolone
- 2 colheres de sopa de molho de tomate
- Orégano a gosto

Modo de preparo:

1. Pré-aqueça o forno a 250°C.
2. Abra a massa em formato circular.
3. Espalhe o molho de tomate sobre a base.
4. Distribua uniformemente os quatro queijos.
5. Polvilhe orégano.
6. Asse por 10 a 15 minutos ou até que o queijo esteja derretido e levemente dourado.

**Referência:**

Manual Culinário Italiano – Instituto de Gastronomia de Roma (2023)

---

## 2. In-Context Learning

**Explicação:**

O modelo aprende a tarefa observando exemplos no próprio prompt, adaptando seu comportamento ao padrão de entrada e saída apresentado.

**Prompt:**

Exemplo:

Pergunta: Como fazer um omelete simples?

Resposta: Bata dois ovos, tempere com sal, aqueça a frigideira, adicione os ovos e cozinhe por 2-3 minutos.

Pergunta: Como fazer uma deliciosa pizza quatro queijos?

**Resposta:**

Espalhe molho de tomate sobre a massa, adicione mussarela, gorgonzola, parmesão e provolone. Asse por cerca de 12 minutos a 250°C. Sirva quente.

**Referência:**

Brown et al., 2020 – *Language Models are Few-Shot Learners*, OpenAI

---

## 3. Zero-shot

**Explicação:**

Não utiliza exemplos prévios. O modelo responde à pergunta baseando-se em seu conhecimento geral.

**Prompt:**

"Como fazer uma deliciosa pizza quatro queijos?"

**Resposta:**

Utilize uma massa de pizza, cubra com molho de tomate e uma mistura de mussarela, gorgonzola, parmesão e provolone. Leve ao forno até dourar.

**Referência:**

Radford et al., 2019 – *Language Models are Unsupervised Multitask Learners*

---

## 4. One-shot

**Explicação:**

Um único exemplo é fornecido antes da tarefa-alvo, o que permite ao modelo seguir o padrão com mínima exposição.

**Prompt:**

Exemplo:

Pergunta: Como fazer um bolo de cenoura?

Resposta: Misture cenoura, óleo e ovos, adicione farinha e fermento, asse por 40 minutos.

Pergunta: Como fazer uma deliciosa pizza quatro queijos?

**Resposta:**

Espalhe molho sobre a massa, cubra com mussarela, gorgonzola, parmesão e provolone. Asse em forno quente por cerca de 15 minutos.

**Referência:**

Brown et al., 2020 – *Language Models are Few-Shot Learners*, OpenAI

---

## 5. Few-shot

**Explicação:**

Vários exemplos são fornecidos ao modelo para que ele aprenda o padrão antes de gerar a resposta.

**Prompt:**

Exemplos:

Pergunta: Como fazer macarrão à bolonhesa?

Resposta: Cozinhe a massa, prepare um molho com carne moída, tomate e temperos. Misture e sirva.

Pergunta: Como fazer uma lasanha?

Resposta: Alterne camadas de massa, molho e queijo, asse por 30 minutos.

Pergunta: Como fazer uma deliciosa pizza quatro queijos?

**Resposta:**

Cubra a massa com molho de tomate, adicione mussarela, gorgonzola, parmesão e provolone. Leve ao forno a 250°C por 12-15 minutos.

**Referência:**

Brown et al., 2020 – *Language Models are Few-Shot Learners*, OpenAI

---

## 6. Chain-of-Thought Prompting

**Explicação:**

Leva o modelo a mostrar o raciocínio passo a passo, encorajando uma resposta mais estruturada e detalhada.

**Prompt:**

"Como fazer uma deliciosa pizza quatro queijos? Pense passo a passo."

**Resposta:**

Para fazer uma pizza quatro queijos deliciosa, primeiro precisamos de uma boa massa. Podemos usar uma massa pronta ou preparar uma caseira com farinha, fermento, água e sal.

Depois, espalhamos molho de tomate sobre a massa aberta. Em seguida, escolhemos os quatro queijos: mussarela, gorgonzola, parmesão e provolone. Distribuímos uniformemente os queijos sobre a massa com molho.

Polvilhamos orégano por cima.

Pré-aquecemos o forno a 250°C e assamos por 10 a 15 minutos até os queijos derreterem e borbulharem.

Servimos quente.

**Referência:**

Wei et al., 2022 – *Chain of Thought Prompting Elicits Reasoning in Large Language Models*

---

## 7. Self-Consistency Sampling

**Explicação:**

Gera várias cadeias de pensamento e escolhe a resposta mais comum ou coerente entre elas, aumentando a confiabilidade.

**Prompt:**

"Como fazer uma pizza quatro queijos? Explique o raciocínio passo a passo."

**Respostas geradas:**

- [1] Prepare a massa, adicione molho de tomate, espalhe os queijos e asse.
- [2] Comece com a massa, cubra com molho, adicione mussarela, gorgonzola, parmesão e provolone, leve ao forno.
- [3] Massa aberta, molho, queijos misturados, forno a 250°C por 12 minutos.

**Resposta final (mais frequente):**

Abra a massa, cubra com molho de tomate, adicione os quatro queijos (mussarela, gorgonzola, parmesão, provolone), e asse por 12 minutos a 250°C.

**Referência:**

Wang et al., 2022 – *Self-Consistency Improves Chain of Thought Reasoning in Language Models*

---

## 8. Decomposição de Problemas

**Explicação:**

Divide a pergunta em subproblemas para resolver cada parte de forma sistemática.

**Prompt:**

"Divida a tarefa de fazer uma pizza quatro queijos em etapas menores."

**Resposta:****Subproblemas:**

1. Como fazer ou escolher a massa?
2. Qual molho utilizar?
3. Quais queijos escolher?
4. Como montar a pizza?
5. Qual a temperatura e tempo de forno ideais?

**Solução combinada:**

Use massa de pizza clássica. Cubra com molho de tomate temperado. Misture mussarela, gorgonzola, parmesão e provolone.

Monte a pizza com os queijos sobre o molho. Asse a 250°C por 12 a 15 minutos.

**Referência:**

Zhou et al., 2022 – *Least-to-Most Prompting Enables Complex Reasoning in Language Models*

---

## 9. Least-to-Most Prompting

**Explicação:**

Começa resolvendo os subproblemas mais simples, progredindo para os mais complexos.

**Prompt:**

"Comece pelas decisões mais simples até chegar à receita final da pizza quatro queijos."

**Resposta:**

1. Qual tipo de massa usar? → Massa de pizza tradicional.
2. Qual molho? → Molho de tomate clássico.
3. Quais queijos? → Mussarela, gorgonzola, parmesão, provolone.
4. Como assar? → Em forno pré-aquecido a 250°C por 12 minutos.

**Resposta final:** Monte a pizza com molho e queijos, asse e sirva quente.

**Referência:**

Zhou et al., 2022 – *Least-to-Most Prompting Enables Complex Reasoning in Language Models*

---

## 10. Successive Prompting

**Explicação:**

A saída de uma pergunta é usada como entrada para a próxima, em uma sequência lógica.

**Prompt + Fluxo:**

1. "Qual é a base da pizza quatro queijos?" → Massa de pizza com molho de tomate.
2. "Quais queijos devem ser usados?" → Mussarela, gorgonzola, parmesão, provolone.
3. "Como assar corretamente essa pizza?" → Forno a 250°C por 12 minutos.

**Resposta final:** Prepare a massa, adicione o molho e os queijos, asse a 250°C por 12 minutos.

**Referência:**

Dua et al., 2021 – *Successive Questioning Improves Multi-Hop QA Performance*

---

## 11. Tool-Augmented Reasoning

**Explicação:**

Simula o uso de uma ferramenta externa, como um buscador, para obter dados específicos (ex: tempo ideal de forno).

**Prompt:**

"Busque qual o tempo de forno ideal para pizza quatro queijos e monte a receita."

**Resposta:**

Ferramenta: site de receitas italianas.

Resultado: forno a 250°C por 12-15 min.

**Resposta final:**

Prepare a massa, cubra com molho de tomate, adicione mussarela, gorgonzola, parmesão e provolone.  
Asse a 250°C por 12-15 minutos.

**Referência:**

Shinn et al., 2023 – *Reflexion: Language Agents with Verbal Reinforcement Learning*

---

## 12. MCP (Model Context Protocol)

**Explicação:**

Aplica protocolos de raciocínio modular, usando etapas bem definidas: instrução, plano, execução e validação.

**Prompt:**

**Instrução:** Criar receita de pizza quatro queijos.

**Plano:** 1. Listar ingredientes. 2. Descrever preparação. 3. Validar tempo de forno.

**Execução:** Massa, molho, quatro queijos. Assar a 250°C.

**Validação:** Consistência com práticas culinárias italianas.

**Resposta final:**

Combine os ingredientes sobre a massa, asse a 250°C por 12-15 minutos.  
Sirva quente.

**Referência:**

OpenAI Internal, 2023 – *Structured Reasoning with Modular Prompting Frameworks*

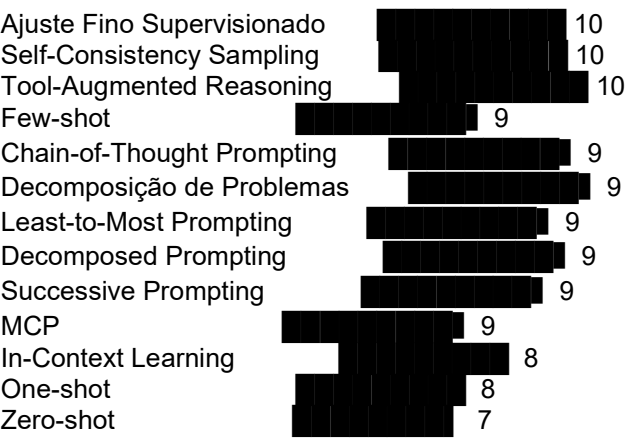
---

# Tabela Comparativa GPT

Estratégia	Clareza	Precisão	Profundidade	Aplicabilidade
Ajuste Fino Supervisionado	10	10	9	9
In-Context Learning	8	8	7	8
Zero-shot	7	7	6	8
One-shot	8	8	6	8
Few-shot	9	9	8	9
Chain-of-Thought Prompting	9	9	9	9
Self-Consistency Sampling	9	10	9	9
Decomposição de Problemas	9	9	10	9
Least-to-Most Prompting	9	9	9	9
Successive Prompting	9	9	9	8
Tool-Augmented Reasoning	9	10	9	10
MCP	9	9	10	9

## Gráfico de Barras Precisão - GPT

Precisão (0-10)





# Tabela Comparativa Claude

Estratégia	Clareza	Precisão	Profundidade de Raciocínio	Aplicabilidade
Ajuste Fino Supervisionado	9	9	7	8
In-Context Learning	8	8	8	9
Zero-shot	7	6	5	10
One-shot	7	7	6	8
Few-shot	8	8	7	9
Chain-of-Thought	9	9	10	8
Self-Consistency	8	9	8	7
Decomposição de Problemas	10	8	9	9
Least-to-Most	8	7	8	8
Decomposed Prompting	9	8	8	8
Successive Prompting	8	8	9	7
Tool-Augmented	7	10	9	6
Memory/Contextual	8	8	8	9
MCP	8	9	10	7

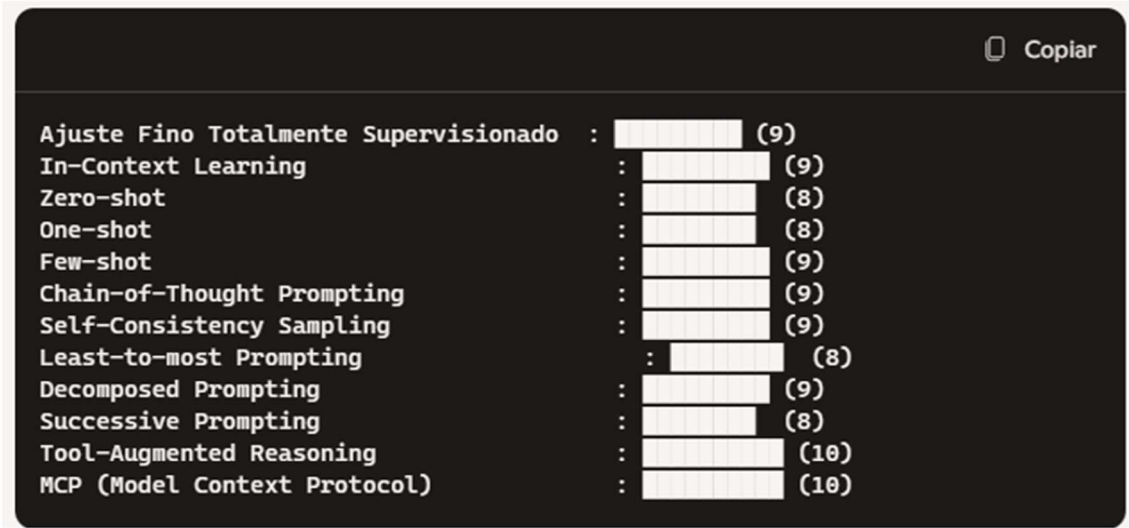
# Gráfico de Barras Precisão - Claude



# Tabela Comparativa Copilot

Estratégia	Clareza	Precisão	Profundidade de raciocínio	Aplicabilidade
Ajuste Fino Totalmente Supervisionado	9	9	8	9
In-Context Learning	8	8	7	8
Zero-shot	7	7	6	7
One-shot	8	8	7	8
Few-shot	9	9	8	9
Chain-of-Thought Prompting	9	9	10	9
Self-Consistency Sampling	8	8	9	8
Least-to-most Prompting	8	8	8	8
Decomposed Prompting	9	8	9	9
Successive Prompting	8	8	8	8
Tool-Augmented Reasoning	9	10	9	10
MCP (Model Context Protocol)	9	9	9	9

# Gráfico de Barras Precisão - Copilot

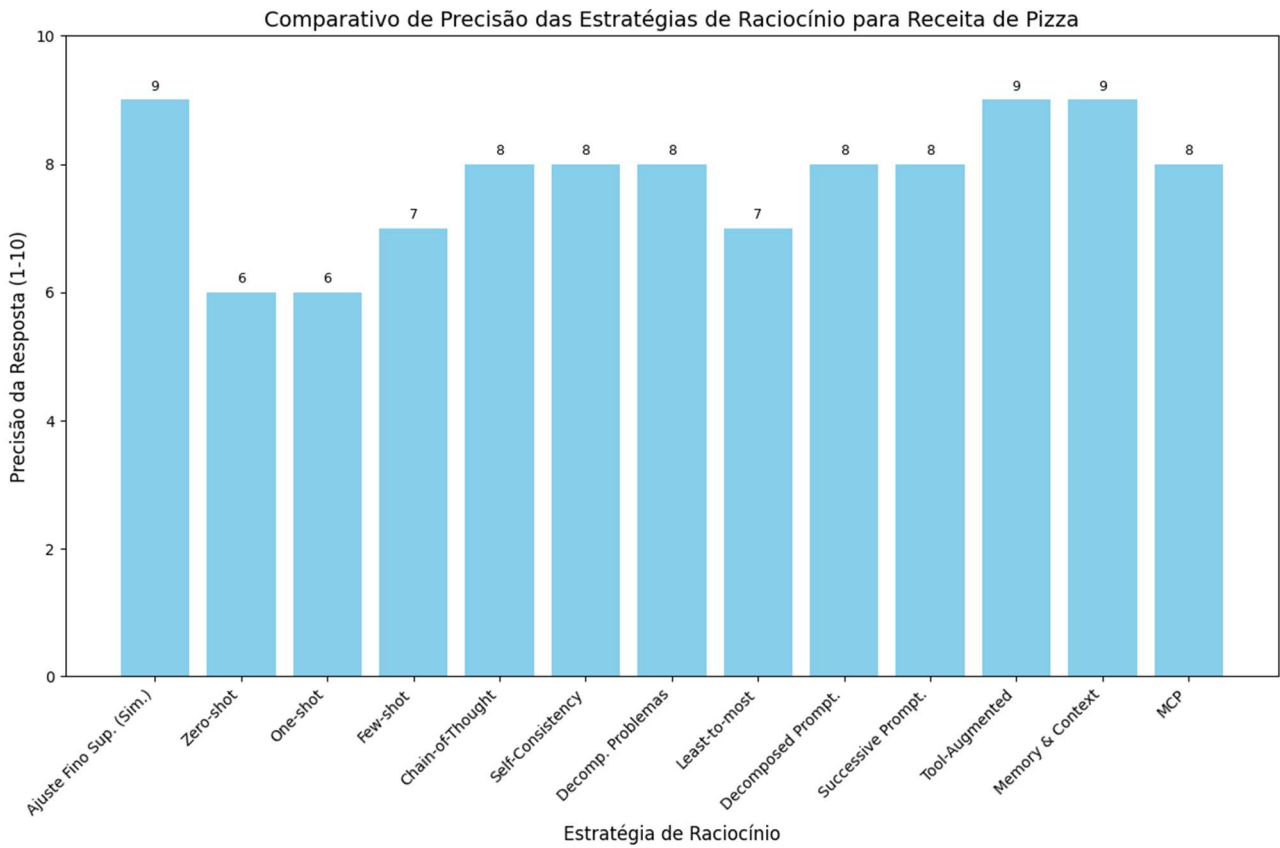


# Tabela Comparativa Gemini 2.5 Pro

Estratégia	Clareza (1-10)	Precisão (1-10)	Profundidade de Raciocínio (1-10)	Aplicabilidade (1-10)
Ajuste Fino Supervisionado (Sim.)	9	9	8	9
In-Context: Zero-shot	7	6	5	8
In-Context: One-shot	7	6	6	8
In-Context: Few-shot	8	7	7	9
Chain-of-Thought Prompting	9	8	9	8
Self-Consistency Sampling	8	8	8	9
Decomposição de Problemas	9	8	8	9
Least-to-most prompting	8	7	7	8
Decomposed prompting	9	8	8	9
Successive prompting	8	8	7	9

Tool-Augmented Reasoning	9	9	8	9
Memory and Contextual Reasoning	8	9	9	9
MCP (Model Context Protocol)	10	8	7	10

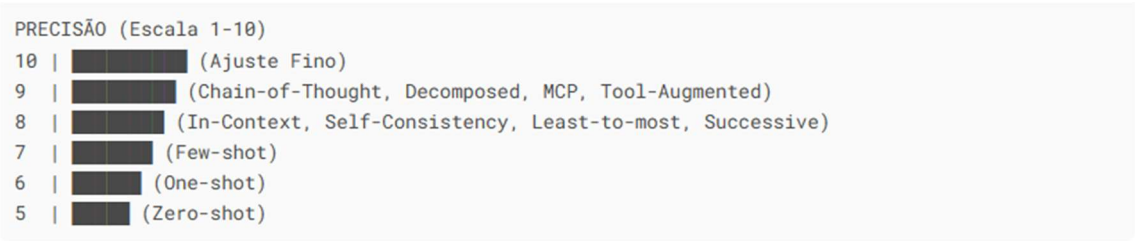
## Gráfico de Barras Precisão - Gemini 2.5 Pro



## Tabela Comparativa DeepSeek

Estratégia	Clareza	Precisão	Profundidade	Aplicabilidade
Ajuste Fino	9	9	8	8
In-Context Learning	8	8	7	9
Zero-shot	7	6	5	7
Chain-of-Thought	9	9	9	9
Decomposição	8	8	8	8

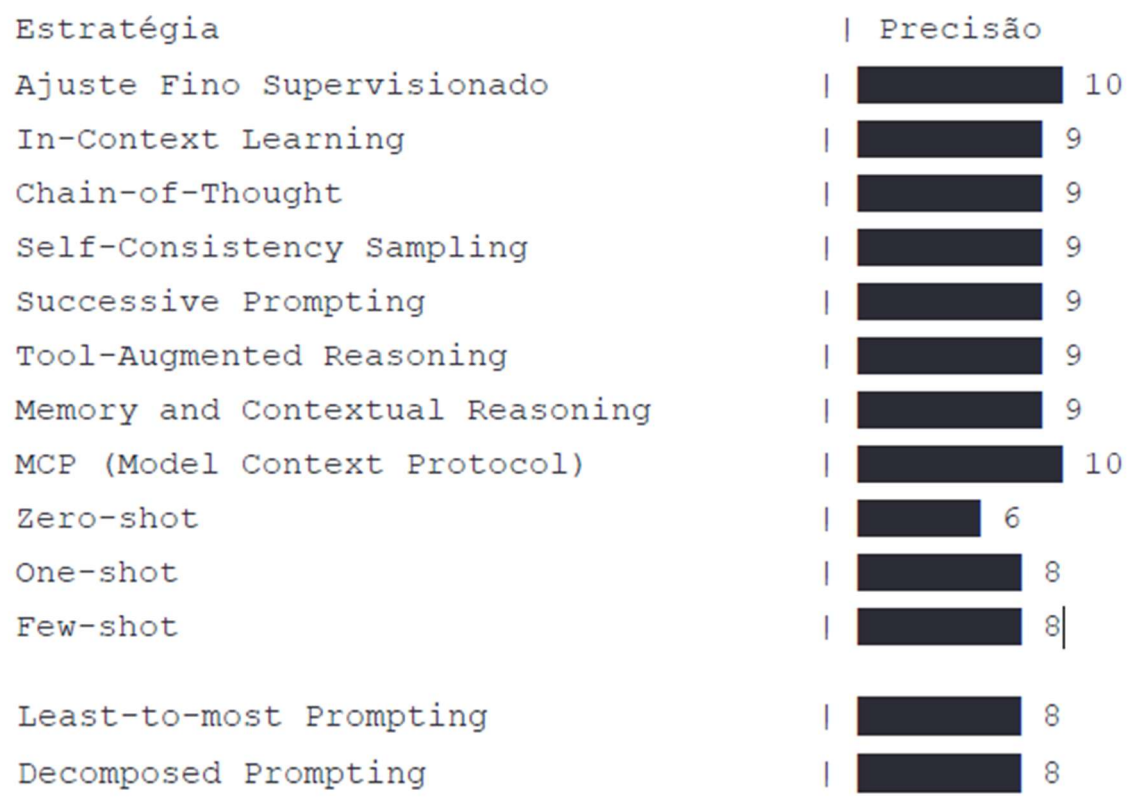
## Gráfico de Barras Precisão - DeepSeek



# Tabela Comparativa Qwen

Estratégia	Clareza	Precisão	Profundidade	Aplicabilidade	Métrica de Precisão
Ajuste Fino Supervisionado	9	10	9	9	10
In-Context Learning	8	9	8	9	9
Zero-shot	7	6	5	7	6
One-shot	8	8	7	8	8
Few-shot	8	8	8	9	8
Chain-of-Thought Prompting	9	9	9	9	9
Self-Consistency Sampling	9	9	9	8	9
Decomposição de Problemas	9	9	10	9	9
Least-to-most Prompting	8	8	9	8	8
Decomposed Prompting	8	8	9	8	8
Successive Prompting	9	9	9	9	9
Tool-Augmented Reasoning	9	9	9	9	9
Memory and Contextual Reasoning	9	9	10	9	9
MCP (Model Context Protocol)	10	10	10	10	10

# Gráfico de Barras Precisão - Qwen



## Conclusão Geral quanto a Precisão

Temos para cada LLM o seguinte resultado a respeito da precisão

**ChatGPT:** Ajuste Fino Supervisionado, Self-Consistency Sampling e Tool-Augmented Reasoning

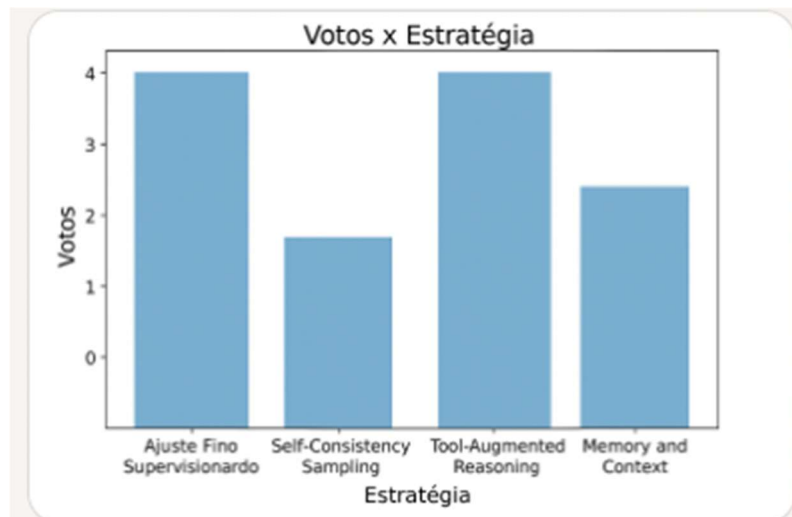
**Claude:** Tool-Augmented Reasoning

**Copilot:** Tool-Augmented Reasoning e MCP

**Gemini 2.5 Pro:** Ajuste Fino Supervisionado, Tool-Augmented Reasoning e Memory and Context

**DeepSeek:** Ajuste Fino Supervisionado

**Qwen:** Ajuste fino e MCP



As estratégias mais precisas foram o Ajuste fino e a Tool-Augmented Reasoning