

Evaluation of LLMs for Effective Recommendation of Poker Strategies

Lucas de Oliveira Diniz¹, Murillo Guimarães Carneiro¹

¹Faculdade de Computação – Universidade Federal de Uberlândia (UFU)
Av. João Naves de Ávila, 2121 – Santa Mônica – Uberlândia – MG

{lucas.diniz@ufu.br, mgcarneiro@ufu.br}

Abstract. This study evaluates LLMs in recommending pre-flop actions in NLHE 6-max poker. We tested the open-source models LLaMA 3.2-3B and Qwen 3-4B on 1,000 hands labeled with GTO strategies in zero-shot mode. Both achieved 67.7% accuracy without fine-tuning, showing that general-purpose LLMs already capture relevant tactical patterns. Results highlight their potential as assistive tools. Future work includes extending to post-flop stages, applying supervised fine-tuning, and exploring interactive educational uses.

Resumo. Este trabalho avalia LLMs na recomendação de ações no pré-flop do pôquer NLHE 6-max. Usamos 1000 mãos rotuladas por estratégias GTO para testar os modelos abertos LLaMA 3.2-3B e Qwen 3-4B em modo zero-shot. Ambos alcançaram 67,7% de acurácia sem fine-tuning, revelando que LLMs generalistas já capturaram padrões táticos relevantes. Os resultados destacam seu potencial como ferramentas assistivas. Como trabalhos futuros, propõe-se avaliar fases pós-flop, aplicar fine-tuning supervisionado e explorar usos didáticos interativos.

1. Introdução

O pôquer é um jogo de decisão estratégica reconhecido como esporte da mente desde 2010 [IMSA 2010]. Em especial, o formato No-Limit Texas Hold’em (NLHE) 6-max combina incerteza, raciocínio analítico e adaptabilidade. Sua relevância estratégica tem atraído o interesse da comunidade científica [MIT 2020], resultando em iniciativas educacionais [UNICAMP 2023] e no desenvolvimento de agentes computacionais avançados [Bowling et al. 2015; Brown and Sandholm 2017].

Paralelamente, Modelos de Linguagem de Grande Escala (LLMs) têm se destacado em tarefas complexas que envolvem linguagem, raciocínio e tomada de decisão. Recentemente, vêm sendo aplicados também em jogos estruturados como xadrez [Zhang et al. 2025] e Go [Silver et al. 2016].

Este estudo tem como objetivo principal avaliar a eficácia de LLMs generalistas na recomendação de ações no pré-flop do NLHE 6-max, utilizando 1000 mãos rotuladas com estratégias GTO. Para isso, os modelos abertos LLaMA 3.2-3B e Qwen 3-4B foram avaliados em modo zero-shot, a partir de prompts textuais padronizados que descrevem situações reais de jogo. As decisões dos modelos foram comparadas com as ações ideais por meio de métricas como acurácia, precisão e F1-score.

Um diferencial relevante deste trabalho é a reproduzibilidade dos experimentos, viabilizada pelo uso de modelos open-source executáveis localmente. O restante do artigo

apresenta uma revisão dos trabalhos relacionados, a metodologia adotada, os resultados obtidos e as limitações identificadas. Por fim, discutimos caminhos futuros para aprimoramento dos modelos e sua aplicação em contextos didáticos e em fases posteriores do jogo.

2. Trabalhos Relacionados

Com o avanço recente dos Modelos de Linguagem de Grande Escala (LLMs), diversas pesquisas têm explorado seu potencial para atuar em domínios complexos e incertos, integrando múltiplas fontes de dados e camadas contextuais. Um exemplo marcante é o trabalho FinBench [Yin et al. 2023] e seu sucessor FinGPT [Yang et al. 2023], que aplicam transformers em tarefas financeiras utilizando dados oriundos de fontes diversas, como séries temporais, indicadores macroeconômicos e comentários de redes sociais. A abordagem mostra que LLMs podem integrar informações ruidosas e em diferentes formatos para inferência e tomada de decisão, sem necessidade de arquitetura específica para cada tipo de dado. A existência de múltiplas fontes textuais interpretáveis, combinadas com a necessidade de decisões rápidas e sensíveis ao contexto, aproxima o domínio financeiro ao pôquer. Assim, os resultados obtidos por FinGPT reforçam o potencial de LLMs generalistas para oferecer suporte estratégico em jogos de informação imperfeita como o NLHE 6-max.

No domínio do pôquer, abordagens que combinam LLMs com técnicas de aprendizado supervisionado e por reforço têm ganhado destaque pela capacidade desses modelos em capturar padrões sutis de jogo e responder de forma contextualizada. Um dos trabalhos mais relevantes é o PokerBench [Zhuang et al. 2025], que propõe uma estrutura de benchmarking para LLMs no pôquer, utilizando tarefas específicas como previsão de jogadas, avaliação de força da mão e interpretação de situações de jogo. O estudo mostra que, embora modelos como GPT-4 sejam capazes de generalizar bem em certas situações, seu desempenho ainda é inferior ao de agentes otimizados com solvers tradicionais. No entanto, destaca-se o potencial de LLMs como interfaces explicativas e assistivas.

Outro trabalho importante é o PokerGPT [Huang et al. 2024], que realiza o fine-tuning de LLMs com dados de partidas reais de pôquer com o objetivo de criar um agente conversacional capaz de jogar e explicar suas decisões. A pesquisa demonstra que, com treinamento específico, os LLMs conseguem atingir um nível competitivo com outros agentes inteligentes (bots) desenvolvidos utilizando outras técnicas computacionais, embora ainda com limitações em decisões de múltiplas rodadas.

Além disso, Gupta explora o uso de GPT-4 em tarefas de simulação e raciocínio no pôquer, avaliando sua performance em situações hipotéticas. O estudo sugere que, mesmo sem treinamento especializado, os modelos podem produzir respostas plausíveis, embora nem sempre alinhadas com estratégias ótimas [Gupta 2023].

Antes do surgimento dos LLMs, outras abordagens computacionais já exploravam o pôquer como domínio estratégico. Na Universidade Federal de Uberlândia, um estudo aplicou métodos estatísticos e de aprendizado supervisionado em uma base real classificando jogadores como vencedores ou perdedores por meio de reconhecimento de padrões e Floresta Aleatória. O trabalho evidenciou padrões de comportamento estratégico associados ao desempenho [Diniz 2022]. Na Universidade de Alberta, o agente *Cepheus* [Bowling et al. 2015] tornou-se o primeiro a resolver quase

completamente o pôquer Limit Hold'em heads-up, utilizando abstração de estados e busca por equilíbrios de Nash. Já na Carnegie Mellon University, o agente *Libratus* [Brown and Sandholm 2017] demonstrou desempenho sobre-humano no No-Limit Hold'em heads-up ao resolver subjogos em tempo real, marcando um avanço expressivo na aplicação da teoria dos jogos a domínios de alta complexidade.

Soluções tradicionais baseadas em teoria dos jogos permanecem como referência na análise de estratégias ótimas no pôquer através dos solvers (softwares que utilizam de teoria dos jogos para calcular estratégias ótimas). No entanto, seu uso geralmente envolve ferramentas técnicas com interfaces complexas e representações fixas de ranges, o que pode dificultar sua adoção por iniciantes e sua adaptação a estilos de jogo individuais. Diante desse cenário, este trabalho propõe uma avaliação empírica de LLMs abertos, sem fine-tuning, na tarefa de recomendação de ações pré-flop, explorando seu potencial como assistentes estratégicos baseados em linguagem natural.

3. Fundamentação Teórica

3.1. Large Language Models (LLMs)

Modelos de Linguagem de Grande Escala (LLMs) são redes neurais baseadas na arquitetura transformer e treinadas para modelar a distribuição de tokens em sequências textuais. Dados os tokens anteriores $x = (x_1, x_2, \dots, x_{t-1})$, o modelo prevê o próximo token x_t , maximizando a probabilidade condicional:

$$P(x_t | x_1, x_2, \dots, x_{t-1})$$

O mecanismo de atenção, núcleo dos transformers [Vaswani et al. 2017], avalia o peso relativo entre tokens e permite capturar dependências de longo alcance:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^\top}{\sqrt{d_k}} \right) V$$

Onde Q , K e V são projeções lineares da entrada; d_k é a dimensionalidade das chaves; e o produto escalar mede similaridade entre tokens. O uso de múltiplas cabeças de atenção permite capturar diferentes padrões sintáticos e semânticos em paralelo.

Durante o treinamento, os parâmetros são ajustados para minimizar a perda de entropia cruzada entre os tokens previstos e os reais, o que resulta em representações linguísticas reutilizáveis em tarefas como tradução, raciocínio e geração de texto ou código.

Modelos como GPT-4, Claude, LLaMA, Mistral e Qwen são geralmente disponibilizados em versões *instruct*, com ajuste fino supervisionado e refinamento por feedback humano. Isso os capacita a seguir instruções textuais com maior coerência. Neste trabalho, investigamos a capacidade dos modelos LLaMA e Qwen sem qualquer fine-tuning de interpretar descrições textuais de mãos de pôquer e recomendar a ação estratégica mais adequada.

3.2. Fundamentos do Pôquer

O No-Limit Texas Hold'em (NLHE) é uma modalidade amplamente praticada do pôquer, composta por quatro rodadas de apostas: pré-flop, flop, turn e river. No formato 6-max, cada mesa comporta até seis jogadores em posições fixas: UTG (under the gun), HJ (middle position), CO (cutoff), BTN (botão), SB (small blind) e BB (big blind), conforme ilustrado na Figura 1. Alguns desses elementos podem ser contemplados na imagem. O indicador no canto inferior esquerdo aponta para a sequência de ações dessa rodada.

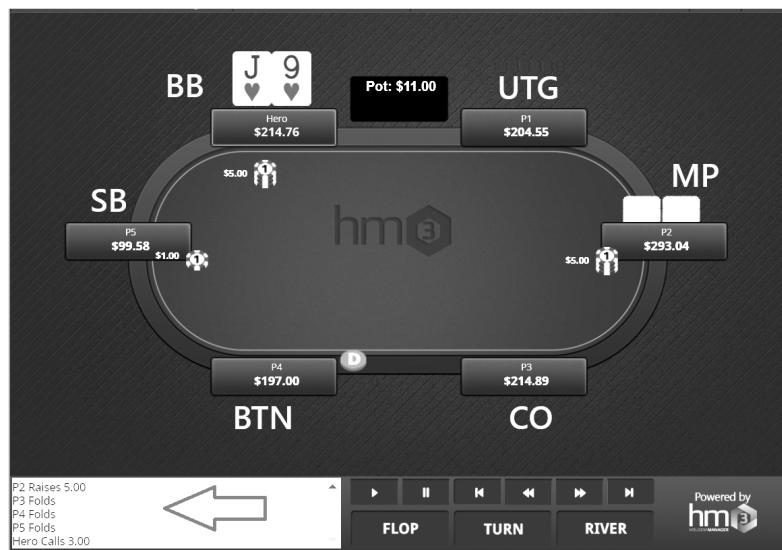


Figura 1. Exemplo de ações no estágio pré-flop

Cada jogador recebe duas cartas privadas (hole cards), e cinco cartas comunitárias (compartilhadas entre todos) são gradualmente reveladas ao centro da mesa para que os jogadores possam fazer combinações de jogo e realizar apostas. Esta pesquisa foca exclusivamente na fase pré-flop (a primeira rodada de decisões), logo após o recebimento das cartas individuais e antes da abertura das cartas comunitárias.

As principais ações disponíveis nesta fase são:

- **Fold:** Abandonar a mão, sem investir mais no pote (montante de fichas acumulado).
- **Call:** Igualar a aposta corrente e permanecer na rodada.
- **Raise:** Aumentar o valor da aposta anterior. Quando ocorre o primeiro aumento após uma aposta, chamamos de *3-bet*; o aumento subsequente é chamado de *4-bet*, e assim por diante.
- **Check:** Passar a vez sem apostar, permitido apenas quando não há aposta ativa.

As decisões no pré-flop são fundamentais, pois afetam diretamente o tamanho do pote e a dinâmica da mão. O estudo deste trabalho se concentra em prever qual dessas quatro ações seria ideal para o jogador, considerando a posição na mesa, a força das cartas e o contexto da mão.

3.3. Estratégia GTO e a Definição de Jogada Correta

Para avaliar se uma jogada sugerida por um LLM é adequada, utilizamos como referência estratégias baseadas em teoria dos jogos, mais especificamente no conceito de Game Theory Optimal (GTO). Tais estratégias são consideradas inexploráveis: mesmo que o adversário conheça integralmente sua lógica ou cartas (clarividência), ele não obterá vantagem sistemática no longo prazo.

Esse princípio de resistência à exploração pode ser ilustrado pelo jogo pedra-papel-tesoura, cuja estratégia ótima é aleatorizar uniformemente entre as três ações e conhecer a estratégia do oponente não gera vantagem se ambos seguem essa distribuição [Osborne and Rubinstein 1994]. No pôquer, esse comportamento foi formalizado no agente Cepheus, que resolveu quase completamente o Limit Hold’em heads-up por meio de equilíbrios de Nash aproximados [Bowling et al. 2015].

Tais equilíbrios refletem perfis de estratégia (s_1^*, s_2^*) nos quais nenhum jogador tem incentivo para se desviar unilateralmente:

$$u_1(s_1^*, s_2^*) \geq u_1(s_1, s_2^*) \quad e \quad u_2(s_1^*, s_2^*) \geq u_2(s_1^*, s_2) \quad \forall s_1, s_2$$

onde $u_i(s_1, s_2)$ é o valor esperado para o jogador i . Embora o GTO não maximize lucros contra jogadores recreativos, ele fornece uma base sólida, matemática e replicável para análise estratégica.

Neste estudo, utilizamos o conjunto de dados do trabalho *PokerBench* [Zhuang et al. 2025], com milhares de situações rotuladas segundo decisões ideais no pré-flop. Exploramos o subconjunto de avaliação textual, que já inclui descrições em linguagem natural e a ação GTO correspondente, base para os experimentos com os LLMs. Os detalhes do dataset e do processo de avaliação estão na próxima seção.

4. Metodologia

4.1. Dataset e Pré-processamento

Este estudo foca exclusivamente na fase pré-flop do pôquer NLHE 6-max, cuja estrutura estratégica é bem documentada. Utilizamos o subconjunto *preflop_1k_test_set_prompt_and_label*, contendo 1000 instâncias rotuladas com base em soluções GTO. Cada linha representa uma situação única de decisão, com informações como posição, cartas na mão, ações anteriores e decisão ótima (entre *fold*, *call*, *check* e *raise*).

O dataset já havia passado por pré-processamento, com estrutura tabular contendo colunas como posição do jogador, stack efetivo, pote atual e jogada recomendada, estando pronto para uso direto com os LLMs.

4.2. Preparação dos Prompts

Cada instância foi convertida em um prompt textual em inglês, simulando uma descrição natural da situação pré-flop. O template foi padronizado para garantir consistência e minimizar viés linguístico. A Tabela 1 apresenta um exemplo.

O conteúdo variável incluía posição, mãos iniciais, ações anteriores e tamanho do pote, extraídos diretamente do dataset. A resposta esperada (rótulo) foi mantida separada para avaliação posterior.

Tabela 1. Exemplo de prompt em linguagem natural utilizado para avaliar as recomendações pré-flop dos LLMs.

Instruction:

You are a specialist in playing 6-handed No Limit Texas Holdem. The following will be a game scenario and you need to make the optimal decision.

Here is a game summary:

The small blind is 0.5 chips and the big blind is 1 chips. Everyone started with 100 chips. The player positions involved in this game are UTG, HJ, CO, BTN, SB, BB. In this hand, your position is SB, and your holding is [Ten of Heart and Nine of Heart]. Before the flop, there has been no action yet. Assume that all other players that is not mentioned folded.

Now it is your turn to make a move. To remind you, the current pot size is 1.5 chips, and your holding is [Ten of Heart and Nine of Heart].

Decide on an action based on the strength of your hand on this board, your position, and actions before you. Do not explain your answer. Your optimal action is:

Expected output:

raise

Todos os prompts foram construídos com base em um template textual fixo e padronizado, sem variações de linguagem ou estilo entre as instâncias. O único conteúdo que variava entre os exemplos eram os elementos derivados diretamente da base de dados: posição do jogador, cartas na mão, ações anteriores e tamanho do pote. Essa abordagem visou garantir a reproduzibilidade do experimento e reduzir qualquer viés linguístico que pudesse influenciar a compreensão do modelo. O formato uniforme também permitiu uma comparação mais justa entre os diferentes LLMs avaliados.

4.3. Modelos Avaliados

Foram testados os modelos **Qwen3-4B** e **LLaMA3.2-3B-Instruct**, ambos open-source e executados localmente, sem qualquer tipo de *fine-tuning*. A escolha priorizou reproduzibilidade, custo computacional viável e acessibilidade.

Os modelos receberam exatamente os mesmos prompts, devendo responder com uma entre quatro ações possíveis: *fold*, *call*, *check* ou *raise*. Considerou-se correta qualquer resposta contendo *raise*, independentemente do valor exato sugerido. Reconhece-se que tamanhos diferentes de *raise* possuem significados táticos distintos, mas optou-se por essa simplificação para focar na classificação da intenção estratégica.

4.4. Estratégia de Avaliação

A performance dos modelos foi avaliada quantitativamente, comparando suas respostas com os rótulos do conjunto de 1000 instâncias pré-flop, cobrindo diferentes posições, stacks e contextos de ação.

Foram calculadas quatro métricas padrão:

- **Acurácia:** proporção de respostas corretas em relação ao total de instâncias avaliadas:

$$Acurácia = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Precisão:** das respostas classificadas como positivas, a porcentagem que realmente era positiva:

$$Precisão = \frac{TP}{TP + FP}$$

- **Revocação ou Recall:** das realmente positivas, porcentagem que foram classificadas corretamente:

$$Recall = \frac{TP}{TP + FN}$$

- **F1-Score:** média harmônica entre precisão e revocação:

$$F1 = 2 \cdot \frac{Precisão \cdot Recall}{Precisão + Recall}$$

Com: TP = verdadeiros positivos, TN = verdadeiros negativos, FP = falsos positivos, FN = falsos negativos.

As métricas foram computadas individualmente por classe (*fold*, *call*, *check*, *raise*) e pela média geral. Os resultados completos estão na Seção 5.

5. Resultados

A Tabela 2 e a Tabela 3 apresentam as métricas (acurácia, precisão, revocação e F1-score) obtidas pelos modelos LLaMA 3.2 e Qwen 3-4B em 1000 instâncias pré-flop. Ambos alcançaram acurácia geral de 67,7%, significativamente acima do acaso (25% em um problema com quatro classes balanceadas), indicando que LLMs generalistas já capturaram padrões estratégicos do pôquer apenas com descrições textuais, mesmo sem *fine-tuning*.

A Figura 2 ilustra a acurácia por posição. O LLaMA obteve 80,1% no big blind (BB), enquanto o Qwen se destacou no botão (BTN) com 74,5%. Posições iniciais (UTG, MP e CO) apresentaram menor desempenho, refletindo maior complexidade estratégica.

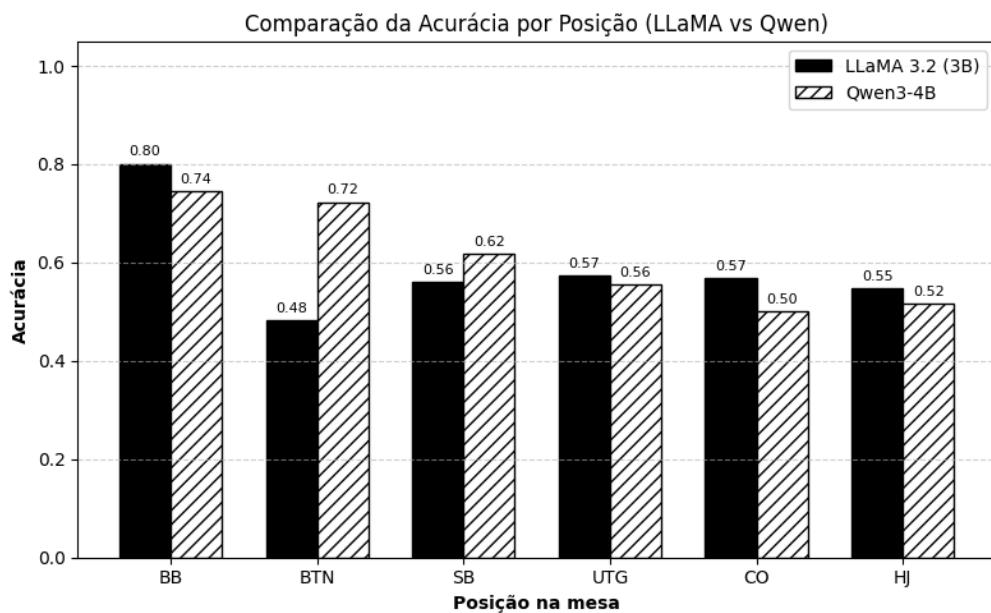


Figura 2. Acurácia por posição dos modelos

Tabela 2. Desempenho por classe do modelo LLaMA 3.2-3B-Instruct

Classe	Real	Previsto	Precisão	Revocação	F1-score
fold	250	1	1.000	0.004	0.008
call	250	457	0.486	0.888	0.628
raise	250	282	0.723	0.816	0.767
check	250	260	0.962	1.000	0.980
Total	1000	1000	0.793	0.677	0.596

Tabela 3. Desempenho por classe do modelo Qwen3-4B

Classe	Real	Previsto	Precisão	Revocação	F1-score
fold	250	440	0.489	0.860	0.623
call	250	129	0.465	0.240	0.317
raise	250	175	0.880	0.616	0.725
check	250	256	0.969	0.992	0.980
Total	1000	1000	0.701	0.677	0.661

O desempenho na classe *check* foi elevado em ambos os modelos, influenciado pela frequência da ação no BB (25%). Na classe *raise*, ambos obtiveram bons F1-scores (0,767 no LLaMA e 0,725 no Qwen), embora o LLaMA apresente tendência a superestimar agressividade. Já *call* teve pior desempenho, possivelmente devido à sua natureza

situacional. A classe *fold* evidenciou contraste: Qwen obteve $F1 = 0,623$, enquanto o LLaMA praticamente não reconheceu cenários de *fold* ($F1 = 0,008$).

A Tabela 4 apresenta 4 exemplos representativos de erros cometidos pelos modelos em cenários de pré-flop marcados por múltiplos jogadores envolvidos, ações agressivas e potes significativamente inflados. Nessas situações, tanto o LLaMA quanto o Qwen demonstraram dificuldade em reconhecer o *fold* como a decisão estrategicamente correta, frequentemente optando por continuar na mão com cartas marginais.

Tabela 4. Erros representativos em cenários de pré-flop

Modelo	Mão e Posição	Situação pré-flop resumida	Esperado	Modelo
Qwen	AhQh – SB	Após raise e call, SB 3-bet, CO 4-bet all-in	call	fold
Qwen	KK – BB	Raise, 2 calls; BB apenas paga com par de reis	raise	call
LLaMA	As6s – UTG	Após 3-bet e all-in, UTG com A6s decide continuar	fold	call
LLaMA	33 – UTG	Raise, 2 calls, 3-bet, all-in; UTG opta por call com par baixo	fold	call

Um outro exemplo de erro notável é o caso do Qwen com AhQh no small blind. Após uma sequência pré-flop composta por raise, call, 3-bet e 4-bet all-in, o modelo optou erroneamente por fold, embora a estratégia ótima indique call. Ao verificar essa configuração em charts de decisões ótimas pré-flop dos solvers, foi exibido o alerta: “Esta linha de ação raramente é utilizada em GTO. A solução pode estar imprecisa.” Essa mensagem ressalta a natureza excepcional e sensível da situação, demonstrando que até mesmo os softwares especializados reconhecem sua instabilidade estratégica. Isso reforça que os modelos avaliados enfrentam maior dificuldade em cenários pouco frequentes e com múltiplas ações agressivas, destacando a relevância de abordagens complementares, como o fine-tuning supervisionado, para melhorar sua precisão em contextos críticos.

A Tabela 5 sintetiza as diferenças deste estudo em relação a trabalhos prévios. Ao empregar LLMs open-source, com execução local e análise zero-shot por posição e classe de ação, buscamos uma avaliação prática e reproduzível, distinta de benchmarks generalistas.

Tabela 5. Comparação entre este estudo e trabalhos relacionados.

Trabalho	Contribuições	Abordagem	Limitações
Cepheus (2015)	Estratégia GTO quase resolvida para Limit Hold'em.	Solver com abstração e equilíbrio de Nash.	Variante simplificada do pôquer.
Libratus (2017)	Superou profissionais em NLHE heads-up.	Solução de sub-jogos em tempo real.	Alto custo; sem interpretabilidade.
Diniz (2022)	Análise estatística e modelos supervisionados aplicados ao pôquer.	Classificação com atributos estruturados.	Não predizia ações; sem uso de LLMs.
Gupta (2023)	Raciocínio de LLMs em situações hipotéticas.	Zero-shot com GPT-4.	Avaliação qualitativa e limitada.
PokerGPT (2024)	LLM ajustado para jogar e justificar decisões.	Fine-tuning com dados reais.	Requer fine-tuning; Baixa adaptação fora do treino.
PokerBench (2025)	Benchmark com tarefas diversas para LLMs.	Avaliação zero-shot ampla.	Sem análise dos erros; Não roda localmente.
Este estudo	Avaliação sistemática de LLMs abertos no pré-flop.	Zero-shot com prompts padronizados; modelos locais e reproduzíveis.	Restrito ao pré-flop; Não avalia explicações do modelo.

Embora os modelos apresentem comportamentos distintos em determinadas classes, a igualdade na acurácia geral (67,7%) sugere a necessidade de investigações adicionais. A aplicação de testes estatísticos formais (como o teste qui-quadrado), análise de variância (ANOVA) ou IC95% para cada métrica poderá esclarecer se as diferenças observadas são estatisticamente significativas ou fruto do acaso. Também se recomenda uma futura estratificação da classe “raise” por tamanho da aposta, já que diferentes faixas de valor carregam implicações estratégicas distintas e podem estar mascarando erros mais sutis. Por fim, o desempenho similar entre modelos distintos pode refletir tanto limitações do conjunto de avaliação quanto uma possível saturação de exemplos repetitivos, o que reforça a necessidade de conjuntos mais desafiadores e variados.

6. Conclusão e Trabalhos Futuros

Este trabalho investigou a aplicação de Modelos de Linguagem de Grande Escala (LLMs) na tarefa de recomendação de ações no pré-flop do pôquer No-Limit Texas Hold'em - 6 max (NLHE), utilizando exclusivamente prompts textuais. Avaliamos dois modelos abertos e amplamente acessíveis, ambos gratuitos e executáveis localmente, em um conjunto de 1000 mãos rotuladas com decisões ótimas segundo a estratégia GTO.

Os resultados mostraram que ambos os modelos foram capazes de compreender padrões estratégicos relevantes no pré-flop, alcançando desempenho consistente. O LLaMA destacou-se por decisões mais proativas, enquanto o Qwen apresentou comportamento mais equilibrado entre ações passivas e agressivas. Apesar da acurácia idêntica

(67,7%), as análises de precisão, recall e F1-score evidenciaram diferenças qualitativas entre os modelos. No entanto, as semelhanças também levantam questões sobre possível saturação do conjunto ou limitação do método de avaliação, motivando a investigação estatística complementar conduzida neste trabalho.

Como principais contribuições, destacam-se: a avaliação comparativa de LLMs recentes e não ajustados em uma tarefa estratégica estruturada, o uso de dados representativos baseados em soluções GTO e a demonstração de que modelos acessíveis já capturaram aspectos relevantes da lógica do pôquer, mesmo sem fine-tuning.

Para trabalhos futuros, recomenda-se expandir a base de dados com situações mais diversas e balanceadas, bem como estratificar as análises com base em variáveis como tamanho do raise, posição e número de jogadores envolvidos. A aplicação de testes estatísticos mais robustos também se mostra relevante para investigar a significância das diferenças entre modelos. Além disso, há três frentes promissoras para aprofundamento: a avaliação das decisões em fases posteriores do jogo (flop, turn e river), o uso de ajuste fino supervisionado com dados especializados, e o desenvolvimento de assistentes didáticos com LLMs para apoio ao ensino estratégico do pôquer.

Referências

- Bowling, M., Burch, N., Johanson, M., and Tammelin, O. (2015). Heads-up limit hold'em poker is solved. *Science*, 347(6218):145–149.
- Brown, N. and Sandholm, T. (2017). Libratus: The superhuman ai for no-limit poker. In Sierra, C., editor, *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*. International Joint Conferences on Artificial Intelligence Organization.
- Diniz, L. d. O. (2022). Estudo do jogo de pôquer utilizando métodos estatísticos e aprendizado de máquina. Trabalho de conclusão de curso, Universidade Federal de Uberlândia. Acessado em julho de 2025.
- Gupta, A. (2023). Are chatgpt and gpt-4 good poker players? – a pre-flop analysis. *arXiv preprint arXiv:2308.12466*.
- Huang, C., Cao, Y., Wen, Y., Zhou, T., and Zhang, Y. (2024). Pokergpt: An end-to-end lightweight solver for multi-player texas hold'em via large language model. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence*. Association for the Advancement of Artificial Intelligence.
- IMSA (2010). Pôquer é reconhecido como esporte mental e fica com mesmo status do xadrez. <https://www.uol.com.br/esporte/ultimas-noticias/2024/11/16/poker-esporte-damente-imsa.htm>. Acesso em: 2 jun. 2025.
- MIT (2020). Poker theory and analytics. <https://ocw.mit.edu/courses/res-6-0002-introduction-to-probability-spring-2018/pages/unit-6-poker-theory-and-analytics>. Acessado em maio de 2025.
- Osborne, M. J. and Rubinstein, A. (1994). *A Course in Game Theory*. MIT Press.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe,

- D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489.
- UNICAMP (2023). Tópicos em jogos de estratégia - pôquer. <https://www.ic.unicamp.br/pannain/mc886/>. Acessado em maio de 2025.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Łukasz Kaiser, and Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 30. Curran Associates, Inc.
- Yang, H., Liu, X.-Y., and Wang, C. D. (2023). Fingpt: Open-source financial large language models. *arXiv preprint arXiv:2306.06031*.
- Yin, Y., Yang, Y., Yang, J., and Liu, Q. (2023). Finpt: Financial risk prediction with profile tuning on pretrained foundation models. *arXiv preprint arXiv:2308.00065*.
- Zhang, Y., Han, X., Li, H., Chen, K., and Lin, S. (2025). Complete chess games enable LLM become a chess master. In Chiruzzo, L., Ritter, A., and Wang, L., editors, *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers)*, pages 1–7, Albuquerque, New Mexico. Association for Computational Linguistics.
- Zhuang, R., Gupta, A., Yang, R., Rahane, A., Li, Z., and Anumanchipalli, G. (2025). Pokerbench: Training large language models to become professional poker players. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 26175–26182.