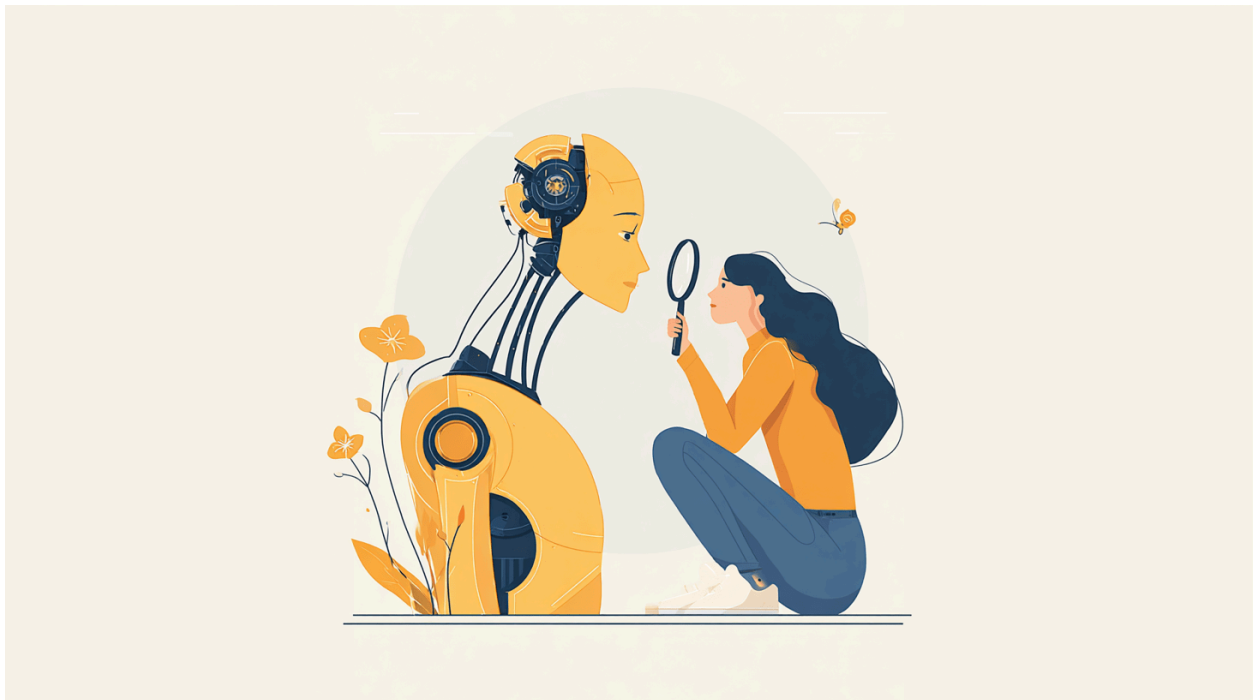


Precisão dos detectores de IA [Atualizado em março de 2026]



Principais conclusões:

Pesquisas acadêmicas sobre a precisão da detecção por IA mostram que o desempenho relatado varia normalmente entre 83% e 100%, com a maioria dos estudos relatando níveis de precisão entre 96% e 99%, indicando que os detectores

modernos de IA podem atingir alto desempenho em condições de teste controladas.

Comparações entre ferramentas individuais mostram variação notável na precisão da detecção de IA, com valores relatados variando de 77% para o Copyleaks a 98% para o Originality.ai, enquanto o Turnitin AI relata 92% de precisão e o GPTZero apresenta 86% de precisão no conjunto de dados avaliado.

A análise da precisão de detecção do Turnitin AI mostra que o sistema detectou texto gerado por IA com 29% de precisão, em comparação com 83% do Originality.ai no mesmo conjunto de dados, enquanto ambas as ferramentas apresentaram precisão relativamente alta na identificação de texto escrito por humanos (93% e 96%, respectivamente).

A precisão de detecção varia dependendo do modelo de IA que gerou o texto, com os sistemas identificando conteúdo do ChatGPT, Grok e Gemini com 100% de precisão, enquanto a precisão de detecção foi ligeiramente menor para o GPT-3.5 (99,7%) e o GPT-4 (98,7%).

Os resultados também mostram que o desempenho da detecção difere dependendo do tipo de texto, com 98% de precisão para texto totalmente gerado por IA, 96% para texto escrito por humanos, 90% para conteúdo humano editado por IA () e 87% para escrita híbrida de IA e humanos, indicando que a autoria mista é mais difícil de classificar pelos detectores.

A análise da taxa de erros demonstra que os sistemas de detecção de IA podem produzir erros de classificação, com taxas de falsos positivos variando de 2% a 38% e taxas de falsos negativos variando de 2% a 20%, enquanto revisores humanos apresentaram uma taxa de falsos negativos de 15% ao identificar conteúdo gerado por IA.

Comparações entre várias plataformas de detecção indicam que a precisão relatada para as principais ferramentas varia de 77% a 99%, com GPTZero, Smodin e Hive relatando 99% de precisão, Turnitin relatando 98% e Quillbot AI Detector relatando 80%, ilustrando uma variação significativa no desempenho entre os sistemas de detecção de IA.

De modo geral, os resultados mostram que a precisão da detecção por IA depende fortemente dos conjuntos de dados, da metodologia de avaliação, da fonte do modelo de IA e do tipo de texto, o que significa que os valores de precisão relatados devem ser interpretados como resultados de pesquisa dependentes do contexto, e não como garantias universais de desempenho.

O rápido crescimento das ferramentas de IA generativa levantou uma questão importante nas áreas de educação, publicação e conteúdo digital: os detectores de IA são precisos ao identificar textos gerados por IA? À medida que os sistemas de redação por IA se tornam mais avançados, muitas organizações dependem de ferramentas de detecção para distinguir entre conteúdo escrito por humanos e gerado por máquinas.

Na prática, avaliar a precisão da detecção por IA é mais complexo do que um único valor percentual. Os pesquisadores geralmente avaliam os detectores usando múltiplas métricas, conjuntos de dados e ambientes de teste. Por causa disso, os resultados relatados podem variar significativamente, dependendo da metodologia utilizada. Essa variação explica por que as discussões sobre a precisão dos detectores de IA frequentemente chegam a conclusões diferentes.

As taxas de precisão dos detectores de IA variam significativamente dependendo do conjunto de dados, do tipo de texto e da metodologia de avaliação. A análise a seguir agrega resultados relatados em diferentes estudos acadêmicos e testes de benchmark.

Os valores de precisão discutidos neste artigo representam resultados relatados de diferentes avaliações, e não um único benchmark padronizado.

Nas seções abaixo, examinamos evidências empíricas sobre a precisão da detecção de IA, incluindo resultados de pesquisas acadêmicas, comparações entre as principais ferramentas de detecção e taxas de erro, como falsos positivos e falsos negativos. Essa análise ajuda a esclarecer se os detectores de IA são precisos em cenários do mundo real e como o desempenho varia entre detectores, conjuntos de dados e tipos de texto.

Após discutir o conceito geral de precisão da detecção de IA, é útil examinar evidências empíricas de estudos acadêmicos que mediram a eficácia com que os detectores de IA identificam textos gerados por IA.

Precisão da detecção de IA em estudos acadêmicos

O gráfico abaixo resume os resultados relatados em vários estudos revisados por pares e de benchmark que avaliam a precisão da detecção de IA em diferentes conjuntos de dados e contextos de pesquisa. Esses estudos abordam a questão comum de se os detectores de IA são precisos, medindo a confiabilidade com que as ferramentas de detecção distinguem entre texto gerado por IA e texto escrito por humanos.

As taxas de precisão dos detectores de IA variam significativamente dependendo do conjunto de dados, do tipo de texto e da metodologia de avaliação. A comparação a seguir agrega resultados relatados em diferentes estudos acadêmicos e testes de benchmark.

De modo geral, os resultados fornecem uma visão geral baseada em dados para os leitores, questionando a precisão dos detectores de IA e se os sistemas atuais são capazes de identificar consistentemente conteúdo gerado por IA.

Estudo



- A maior precisão de detecção de IA relatada atingiu 100% tanto no estudo “Humans vs AI Detectors in Medical Writing” quanto no “Multi-Authored Text Detection Study”.
- A menor precisão no conjunto de dados foi de 83% no “Estudo de Detecção em Contexto Acadêmico”, mostrando que os resultados podem variar significativamente dependendo do ambiente de teste.
- A maioria dos estudos relatou precisão entre 96% e 99%, incluindo 97,6% no “Estudo de Detecção em Publicações Biomédicas” e 98% no “Estudo de Detecção de Redação de Estudantes de STEM”.

Qual é o grau de precisão dos detectores de IA?
Evidências da pesquisa acadêmica

Estudo	Precisão, %
Estudo empírico sobre ferramentas de detecção de textos gerados por IA	97,0%
Detecção de textos gerados por IA (estudo com 16 detectores)	97,0%
Estudo de Referência RAID	85,0%
Detectores humanos vs. detectores de IA em textos médicos	100,0%
Estudo de detecção de resumos oncológicos	96,0%
Estudo sobre o uso de LLMs por estudantes	91,0%
Estudo sobre redação acadêmica impulsionada por IA	96,6%
Estudo sobre detecção de publicações biomédicas	97,6

Revisão sistemática de detectores de plágio baseados em IA	99,0%
Estudo de detecção de plágio em textos de estudantes de ciências, tecnologia, engenharia e matemática	98,0%
Estudo de detecção de textos com múltiplos autores	100,0%
Estudo de detecção de IA em árabe	96,0%
Estudo de detecção de contexto acadêmico	83,0%

Entre os estudos analisados, a maioria dos valores de precisão relatados situa-se na faixa de 90 a 100%, indicando que as ferramentas modernas de detecção de IA podem atingir um desempenho relativamente alto em ambientes de pesquisa controlados. Ao mesmo tempo, a variação entre 83% e 100% de precisão demonstra que a eficácia desses sistemas depende fortemente do conjunto de dados, do domínio do texto e da metodologia de avaliação.

Essas descobertas ajudam a responder a perguntas como: qual é a precisão da IA na identificação de textos gerados por IA e se os detectores de IA são confiáveis em cenários do mundo real. Embora os dados mostrem que muitos sistemas apresentam bom desempenho em benchmarks acadêmicos, a variação entre os estudos sugere que os resultados da detecção por IA devem ser interpretados dentro do contexto, em vez de tratados como uma pontuação de precisão única e universal.

Embora os estudos acadêmicos forneçam uma visão geral da precisão da detecção de IA, uma comparação mais detalhada de ferramentas individuais ajuda a responder a questões práticas, como se o detector de IA do Turnitin é preciso e como seu desempenho se compara a outros sistemas de detecção de IA.

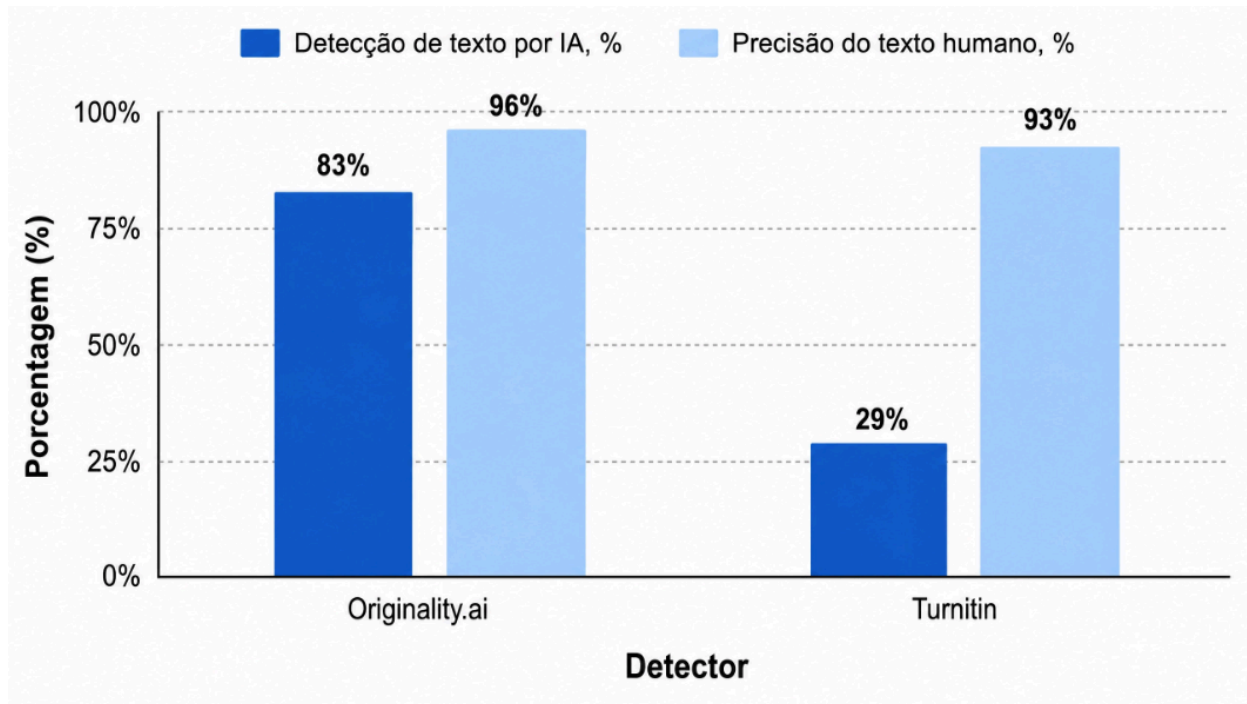
Precisão da detecção de IA do Turnitin em comparação com outros detectores de IA

O gráfico abaixo compara o desempenho de detecção de dois detectores de IA em duas métricas: a capacidade de identificar corretamente textos gerados por IA e a precisão na classificação de conteúdo escrito por humanos.

Essa comparação ajuda a esclarecer questões comuns, como o quão preciso é o detector de IA do Turnitin e se seu desempenho está alinhado com a precisão de detecção de IA do Turnitin relatada em avaliações acadêmicas. Os resultados provêm de um estudo que testou ambos os detectores em conjuntos de dados contendo textos gerados por IA, escritos por humanos e híbridos.

As taxas de precisão dos detectores de IA variam dependendo do conjunto de dados, do tipo de texto e da metodologia de avaliação. A comparação a seguir reflete os resultados relatados em um estudo específico, e não um parâmetro de referência universal.

Turnitin vs. outros detectores de IA



- O Originality.ai detectou textos gerados por IA com 83% de precisão, em comparação com 29% do Turnitin na mesma avaliação.
- O Turnitin identificou corretamente textos escritos por humanos em 93% dos casos, um pouco abaixo do Originality.ai, que atingiu 96%.
- A diferença na capacidade de detecção de IA entre os dois sistemas chegou a 54 pontos percentuais (83% vs 29%) neste conjunto de dados.

Qual é a precisão do detector de IA do Turnitin em comparação com outras ferramentas?

Detector Detecção de texto por IA, % Precisão do texto humano, %

Originality.ai	83%	96%
Turnitin	29%	93%

A comparação destaca como os resultados podem variar significativamente entre as ferramentas de detecção de IA ao avaliar textos gerados por IA. Neste conjunto de dados, a precisão de detecção de IA relatada pelo Turnitin para identificar conteúdo gerado por IA foi consideravelmente menor do que a do sistema alternativo, embora ambas as ferramentas tenham demonstrado precisão relativamente alta ao classificar textos escritos por humanos.

Essas descobertas contribuem para as discussões em andamento sobre se o detector de IA do Turnitin é preciso e ilustram por que a precisão de detecção de IA relatada pelo Turnitin deve ser interpretada dentro do contexto de condições específicas de teste. No geral, os dados sugerem que o desempenho da detecção depende não apenas da ferramenta em si, mas também do conjunto de dados e da metodologia de avaliação utilizados no estudo.

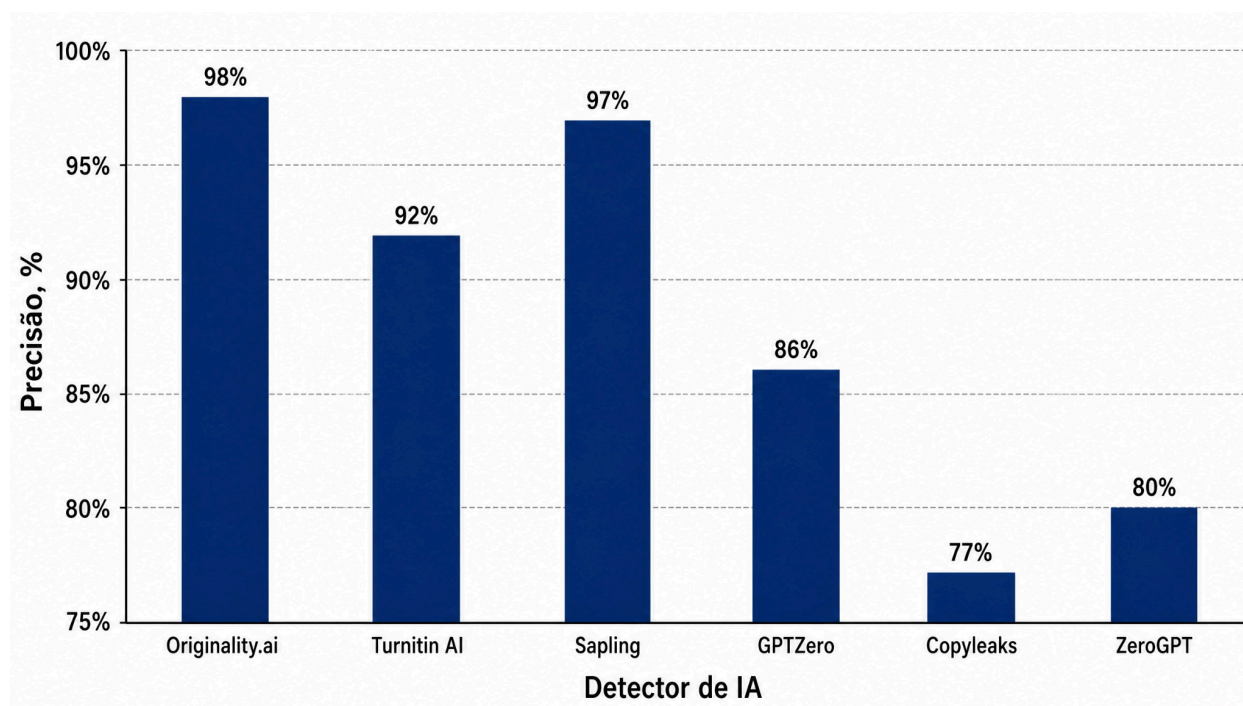
Após examinar a precisão da detecção de IA do Turnitin em comparação com outro detector, o próximo passo é analisar o desempenho de várias ferramentas importantes de detecção de IA lado a lado.

Comparação de precisão dos principais detectores de IA

O gráfico abaixo compara a precisão de detecção de IA relatada por várias ferramentas de detecção de IA amplamente utilizadas. Essa comparação ajuda a responder a perguntas comuns, como qual é o detector de IA mais preciso e qual é o nível de precisão dos detectores de IA quando avaliados em diferentes plataformas.

As taxas de precisão dos detectores de IA variam significativamente dependendo do conjunto de dados, do tipo de texto e da metodologia de avaliação. A comparação a seguir agrega resultados relatados em diferentes estudos acadêmicos e testes de benchmark.

Devido a essas diferenças, o gráfico deve ser interpretado como uma visão geral comparativa, e não como uma classificação definitiva do desempenho dos detectores.



- O Originality.ai apresenta a maior precisão registrada, com 98%, seguido de perto pelo Sapling, com 97%.
- O Turnitin AI apresenta uma precisão de 92%, situando-se entre os detectores de melhor desempenho e as ferramentas de desempenho inferior.
- O Copyleaks e o ZeroGPT apresentam níveis de precisão mais baixos, de 77% e 80%, respectivamente, enquanto o GPTZero apresenta 86% de precisão nos resultados avaliados.

Qual é o detector de IA mais preciso? Comparação de precisão entre as ferramentas

Detector de IA	Precisão, %
Originality.ai	98%
Turnitin AI	92%
Sapling	97%
GPTZero	86%
Copyleaks	77%
ZeroGPT	80%

A comparação ilustra que a precisão de detecção de IA relatada pode variar consideravelmente entre os detectores de IA. Enquanto algumas ferramentas relatam níveis de precisão acima de 95%, outras apresentam desempenho mais próximo de 77-86%, dependendo da avaliação.

Essas diferenças ajudam a explicar por que questões como “qual é o detector de IA mais preciso” permanecem abertas à interpretação. Como os resultados da detecção dependem de conjuntos de dados, métodos de avaliação e dos tipos de texto analisados, os valores de precisão relatados devem ser vistos como indicadores comparativos, e não como garantias fixas de desempenho.

Além da comparação entre detectores individuais, outro fator importante que afeta a precisão da detecção de IA é o tipo de modelo de IA que gerou o texto.

Precisão da detecção de IA por fonte de LLM

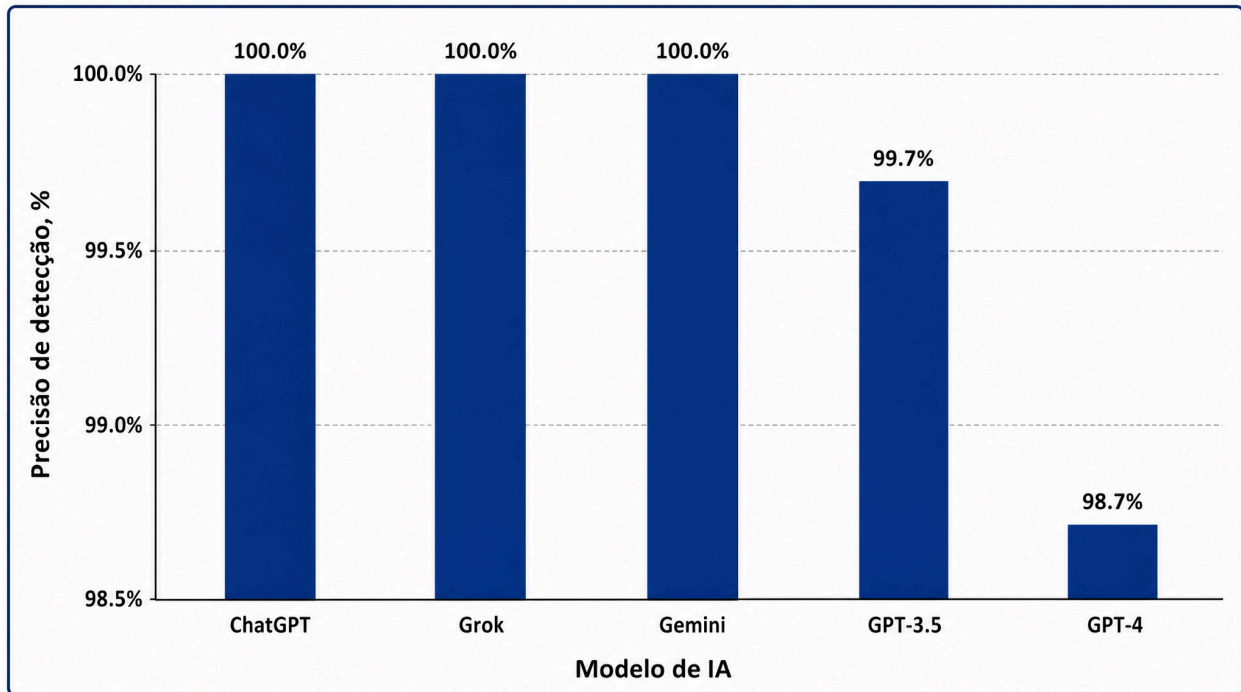
O gráfico abaixo mostra com que precisão os sistemas de detecção de IA identificam textos gerados por diferentes modelos de linguagem de grande porte (LLMs). Avaliar o desempenho de detecção em vários modelos ajuda a abordar questões mais amplas, como a precisão da IA ao identificar textos gerados por IA de diferentes fontes.

A precisão relatada pode variar dependendo do conjunto de dados, da estrutura do texto e da metodologia de avaliação utilizada em cada estudo. Como resultado, o desempenho da detecção pode diferir ao analisar resultados de diferentes modelos de IA.

Os valores no gráfico representam resultados relatados de avaliações específicas, e não um único parâmetro de referência padronizado.

Compreender essas diferenças ajuda a explicar por que as discussões sobre a precisão dos detectores de IA frequentemente dependem do modelo de IA que está sendo testado.

Precisão da detecção de IA por fonte de LLM



- Textos gerados por IA a partir do ChatGPT, Grok e Gemini foram detectados com 100,0% de precisão nos testes avaliados.
- A precisão de detecção do GPT-3.5 atingiu 99,7%, indicando uma identificação quase perfeita de conteúdo gerado por IA.
- O GPT-4 apresentou uma precisão de detecção ligeiramente inferior, de 98,7%, embora ainda tenha permanecido acima do nível de 98%.

Como a precisão de detecção varia entre diferentes modelos de IA

Modelo de IA

Precisão de detecção, %

ChatGPT	100,0%
Grok	100,0%
Gemini	100,0%
GPT-3.5	99,7%
GPT-4	98,7%

Os resultados sugerem que os sistemas de detecção de IA podem atingir alta precisão ao identificar textos gerados pelas principais plataformas de LLM. No conjunto de dados analisado, a precisão da detecção variou de 98,7% a 100%, dependendo do modelo de IA.

Essas descobertas fornecem um contexto adicional para questões como se os detectores de IA são precisos e quão eficazmente os sistemas de detecção distinguem conteúdo gerado por IA. Embora os resultados relatados demonstrem fortes capacidades de detecção em vários LLMs importantes, as variações entre conjuntos de dados e métodos de avaliação significam que o desempenho no mundo real pode diferir dependendo da fonte de texto específica que está sendo analisada.

Além do modelo de IA que gera o texto, outro fator que influencia a precisão da detecção de IA é o tipo de conteúdo que está sendo analisado.

Precisão da detecção por tipo de texto

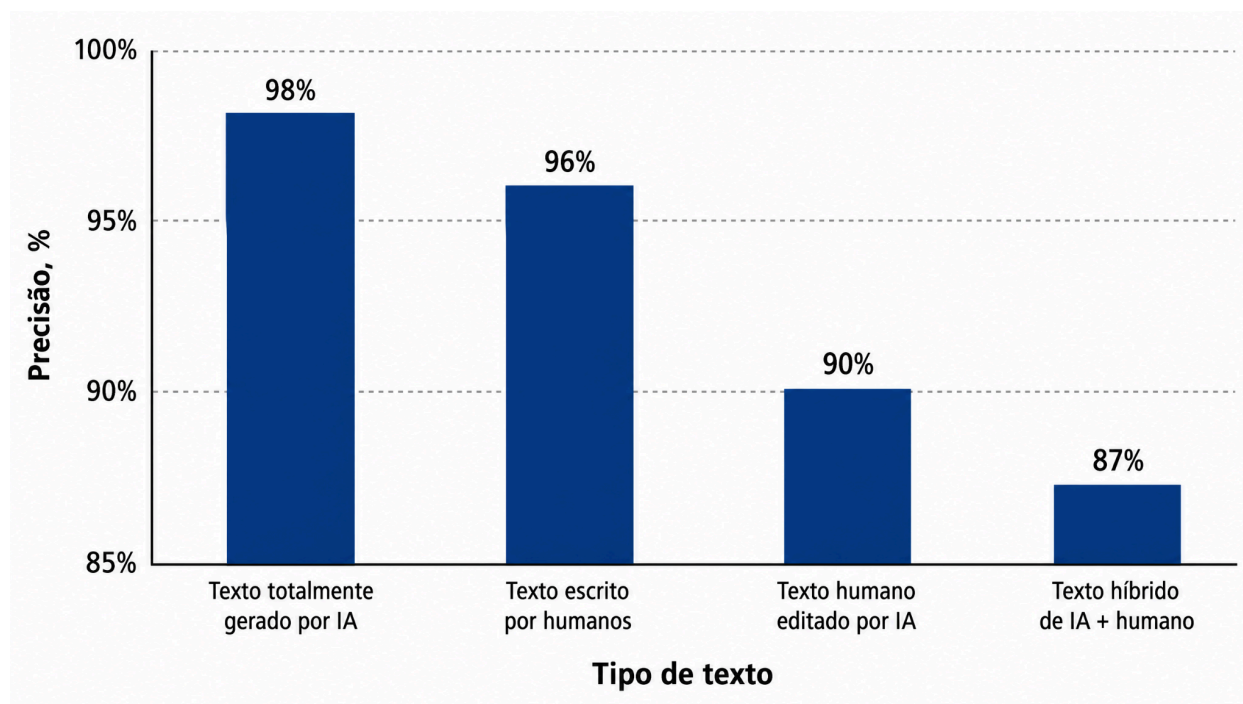
O gráfico abaixo compara o desempenho de detecção em diferentes categorias de texto, incluindo conteúdo totalmente gerado por IA, texto escrito por humanos, texto editado por IA e conteúdo híbrido de IA e humano.

Essas distinções são importantes ao avaliar a precisão dos detectores de IA, pois o nível de envolvimento da IA no texto pode afetar significativamente os resultados da detecção. Por exemplo, identificar texto totalmente gerado por IA é normalmente mais fácil do que detectar conteúdo parcialmente editado ou híbrido.

As taxas de precisão dos detectores de IA variam dependendo do conjunto de dados, da estrutura do texto e da metodologia de avaliação. Os valores mostrados no gráfico refletem resultados relatados em uma avaliação específica, e não um padrão de referência universal.

Compreender essas diferenças também ajuda a explicar as discussões em andamento sobre se os detectores de IA são confiáveis ao analisar textos mistos ou parcialmente assistidos por IA.

Precisão da detecção por tipo de texto



- A precisão de detecção atingiu 98% para textos totalmente gerados por IA, representando o melhor desempenho entre os tipos de texto avaliados.
- Os detectores de IA classificaram corretamente 96% dos textos escritos por humanos, indicando uma taxa relativamente baixa de falsos positivos neste conjunto de dados.
- A precisão caiu para 90% no caso de textos humanos editados por IA e para 87% no caso de conteúdo híbrido (IA e humano), demonstrando que a autoria mista é mais difícil de detectar.

Como o tipo de texto afeta a precisão da detecção de IA

Tipo de texto	Precisão, %
Texto totalmente gerado por IA	98%
Texto escrito por humanos	96%
Texto humano editado por IA	90
Texto híbrido de IA + humano	87%

Os resultados mostram que a precisão da detecção por IA varia dependendo de como o conteúdo foi criado. Os sistemas apresentam melhor desempenho ao analisar textos totalmente gerados por IA, onde a precisão da detecção atingiu 98%, enquanto conteúdos mistos ou parcialmente editados apresentam maiores desafios.

Essas descobertas fornecem um contexto adicional para questões como a precisão dos detectores de IA e se eles são confiáveis ao avaliar textos do mundo real. À medida que o uso de ferramentas de edição assistidas por IA aumenta, distinguir entre conteúdo escrito por humanos, editado por IA e híbrido pode se tornar um fator-chave que influencia o desempenho dos sistemas de detecção de IA.

Embora as métricas de precisão ofereçam uma visão geral da precisão da detecção de IA, avaliar falsos positivos e falsos negativos ajuda a responder a outra questão importante: os detectores de IA podem estar errados?

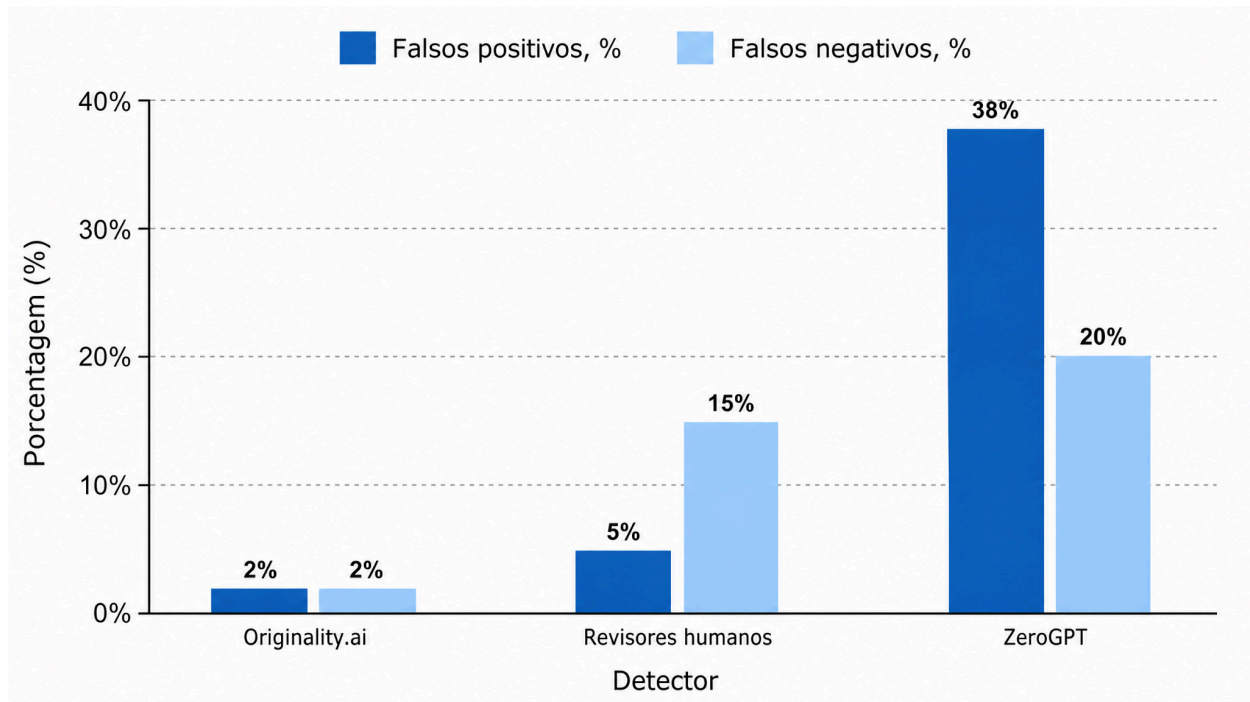
Taxas de falsos positivos e falsos negativos na detecção de IA

O gráfico abaixo compara as taxas de erro de diferentes avaliadores, incluindo detectores de IA automatizados e revisores humanos. Esses resultados são normalmente relatados em pesquisas que avaliam metodologias de artigos sobre detectores de IA e verificadores de IA, nas quais métricas da matriz de confusão, como falsos positivos e falsos negativos, são usadas para medir a confiabilidade.

Os falsos positivos ocorrem quando um texto escrito por humanos é identificado incorretamente como gerado por IA, enquanto os falsos negativos ocorrem quando um conteúdo gerado por IA é classificado como escrito por humanos. Ambos os tipos de erros influenciam a confiabilidade dos detectores de IA em aplicações do mundo real.

As taxas de precisão dos detectores de IA variam dependendo do conjunto de dados e do método de avaliação. Os valores apresentados abaixo refletem resultados relatados em estudos específicos, e não um único benchmark padronizado.

Taxas de falsos positivos e falsos negativos



- O Originality.ai registrou as menores taxas de erro, com 2% de falsos positivos e 2% de falsos negativos no conjunto de dados avaliado.
- Os revisores humanos apresentaram uma taxa de falsos negativos mais alta, de 15%, o que significa que textos gerados por IA foram ignorados em 15% dos casos.
- O ZeroGPT apresentou as taxas de erro mais altas, com 38% de falsos positivos e 20% de falsos negativos na mesma avaliação.

Os detectores de IA podem estar errados? Taxas de erro em sistemas de detecção de IA

Detector	Falsos positivos, %	Falsos negativos, %
Originality.ai	2%	2%
Revisores humanos	5%	15%
ZeroGPT	38%	20%

Os dados mostram que avaliar a precisão da detecção de IA requer mais do que uma única porcentagem de precisão. Falsos positivos e falsos negativos fornecem informações adicionais sobre o desempenho dos sistemas de detecção ao classificar textos gerados por IA e escritos por humanos.

Neste conjunto de dados, os detectores automatizados apresentaram taxas de erro que variaram de 2% a 38%, enquanto os revisores humanos demonstraram uma taxa de falsos negativos de 15%. Essas diferenças ajudam a explicar por que as discussões sobre se os detectores de IA podem estar errados continuam relevantes na pesquisa acadêmica e técnica.

De modo geral, os resultados relatados em vários artigos sobre detectores de IA e estudos sobre verificadores de IA sugerem que tanto as ferramentas automatizadas quanto os revisores humanos podem cometer erros de classificação, reforçando a importância de interpretar os resultados da detecção por IA dentro do contexto de conjuntos de dados e métodos de avaliação específicos.

Após examinar a precisão da detecção, a variação do conjunto de dados e as taxas de erro, é útil analisar como as diferentes ferramentas de detecção de IA se comparam em termos gerais de precisão relatada.

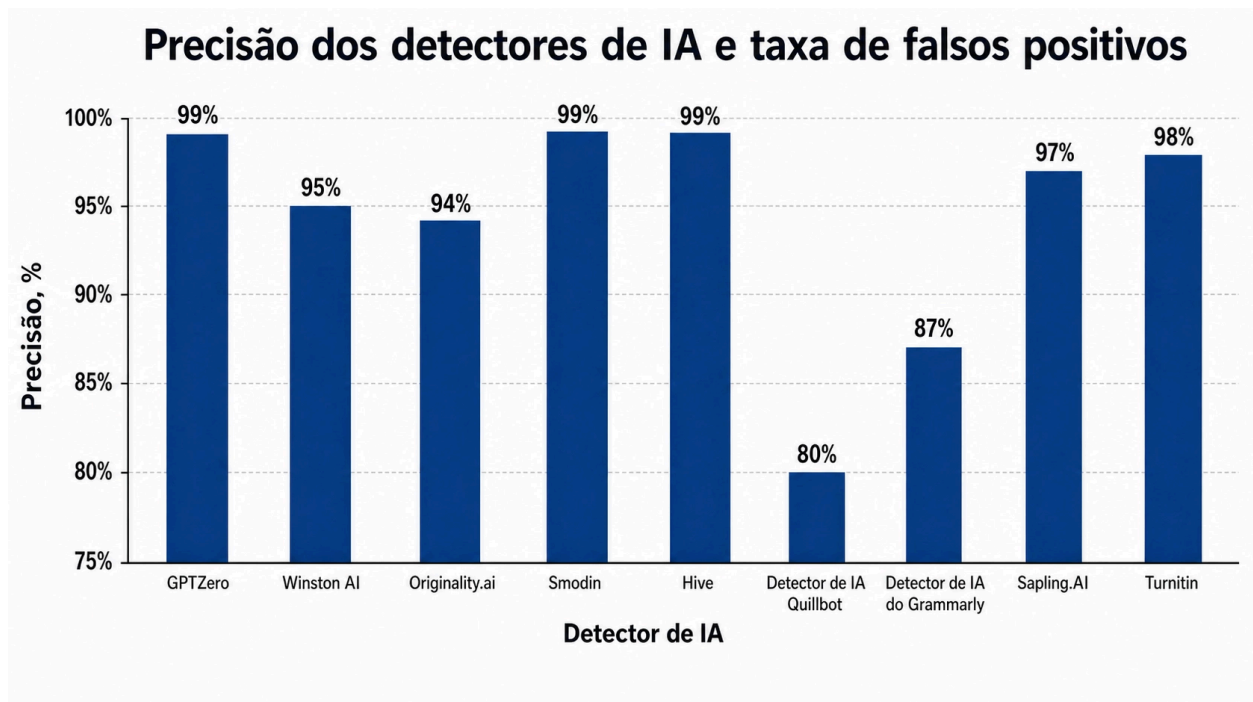
Detectores de conteúdo de IA mais precisos

O gráfico abaixo compara várias ferramentas de detecção de IA amplamente utilizadas com base nos valores de precisão relatados a partir de diferentes conjuntos de dados e avaliações. Essa comparação ajuda a responder à pergunta comum sobre qual é o detector de IA mais preciso, resumindo os indicadores de desempenho relatados em diferentes ferramentas.

As taxas de precisão dos detectores de IA variam significativamente dependendo do conjunto de dados, do tipo de texto e da metodologia de avaliação. A comparação a seguir agrega resultados relatados em diferentes avaliações e testes de benchmark.

Devido a essas diferenças metodológicas, as comparações devem ser interpretadas como uma visão geral, e não como uma classificação definitiva do desempenho dos detectores.

Detectores de conteúdo de IA mais precisos



- O GPTZero, o Smodin e o Hive apresentam os níveis mais altos de precisão, com 99%, de acordo com os dados disponíveis.
- O Turnitin apresenta 98% de precisão, enquanto o Sapling.AI apresenta 97% de precisão com taxas de falsos positivos relativamente baixas.
- Valores de precisão mais baixos aparecem para algumas ferramentas, incluindo o Quillbot AI Detector, com 80%, e o Grammarly AI Detector, com 87%.

Qual é o detector de IA mais preciso? Comparação da precisão das principais ferramentas

Detector de IA	Precisão, %	Taxa de falsos positivos
GPTZero	99%	Baixa
Winston AI	95%	Moderado
Originality.ai	94%	Moderado-alto
Smodin	99%	Moderado

Hive	99%	Muito baixo
Detector de IA Quillbot	80%	Moderado
Detector de IA do Grammarly	87%	Alto
Sapling.AI	97%	Baixo
Turnitin	98%	Baixo

A comparação mostra que a precisão de detecção de IA relatada varia substancialmente entre as plataformas de detecção de IA. Alguns sistemas relatam níveis de precisão próximos a 99%, enquanto outros operam mais próximos da faixa de 80-90%, dependendo das condições de avaliação.

Essas diferenças ajudam a explicar por que perguntas como “qual é o detector de IA mais preciso” continuam difíceis de responder de forma definitiva. O desempenho da detecção depende não apenas da ferramenta em si, mas também do conjunto de dados, da metodologia de teste e do tipo de conteúdo que está sendo analisado.

Como resultado, as comparações entre detectores devem ser interpretadas dentro do contexto dos estudos relatados, em vez de serem tratadas como garantias fixas de desempenho em todos os casos de uso.

Conclusões

- Os dados disponíveis sobre a precisão da detecção de IA indicam que os sistemas modernos de detecção de IA podem atingir um desempenho relativamente alto em ambientes de pesquisa controlados. Nos estudos acadêmicos analisados, os valores de precisão relatados variaram de 83% a 100%, com a maioria dos estudos relatando resultados entre 96% e 99%, sugerindo que muitos detectores de IA são capazes de identificar com confiabilidade textos gerados por IA sob condições específicas de teste.
- Comparações entre ferramentas individuais mostram que a precisão da detecção de IA varia substancialmente entre as plataformas de detecção. Nos conjuntos de dados avaliados, a precisão relatada variou de 77% para o Copyleaks a 98% para o Originality.ai, enquanto o Turnitin AI relatou 92% de precisão e o GPTZero relatou 86% de precisão, indicando que diferentes ferramentas podem produzir resultados visivelmente diferentes.
- Análises adicionais mostram que a precisão de detecção do Turnitin AI pode variar dependendo do conjunto de dados e do método de avaliação. No conjunto de dados examinado neste artigo, o Turnitin detectou texto gerado por IA com 29% de precisão, em comparação com 83% do Originality.ai, enquanto ambas as ferramentas demonstraram precisão relativamente alta na identificação de texto escrito por humanos (93% e 96%, respectivamente).
- O desempenho da detecção também varia dependendo do modelo de IA que gerou o texto. Nos resultados avaliados, os detectores de IA identificaram textos produzidos pelo ChatGPT, Grok e Gemini com 100% de precisão, enquanto taxas de detecção ligeiramente mais baixas foram relatadas para o GPT-3.5 (99,7%) e o

GPT-4 (98,7%), demonstrando que os resultados da detecção podem diferir dependendo da fonte do LLM.

- A análise mostra ainda que a estrutura do texto e o tipo de autoria influenciam o desempenho da detecção. A precisão da detecção atingiu 98% para textos totalmente gerados por IA, 96% para textos escritos por humanos, 90% para conteúdo humano editado por IA e 87% para textos híbridos (IA e humanos), indicando que conteúdos mistos ou parcialmente assistidos por IA podem ser mais difíceis de classificar corretamente pelos detectores.
- A análise da taxa de erros indica que tanto os sistemas automatizados quanto os avaliadores humanos podem produzir erros de classificação. No conjunto de dados analisado, as taxas de falsos positivos variaram de 2% a 38%, enquanto as taxas de falsos negativos variaram de 2% a 20%, e os revisores humanos demonstraram uma taxa de falsos negativos de 15%, ilustrando que os resultados da detecção por IA não estão isentos de erros.
- De modo geral, as descobertas sugerem que a precisão da detecção por IA depende fortemente do conjunto de dados, da metodologia de avaliação, da fonte do modelo de IA e do tipo de texto que está sendo analisado. Como resultado, os valores de precisão relatados devem ser interpretados como resultados dependentes do contexto, e não como garantias universais de desempenho, ao avaliar se os detectores de IA podem identificar de forma confiável o conteúdo gerado por IA.

Fontes

- Barlow, Written. “9 melhores detectores de IA com a maior precisão em 2026.” Recursos de detecção de IA | GPTZero, 2 de janeiro de 2026, <https://gptzero.me/news/best-ai-detectors/>. Acessado em 16 de março de 2026.
- Gillham, Jonathan. “Estudos de precisão na detecção de IA – Meta-análise de 13 estudos – Originality.AI.” Originality.AI, <https://originality.ai/blog/ai-detection-studies-round-up>. Acessado em 16 de março de 2026.
- “Temos 99% de precisão na detecção de IA: Estudo da Originality.AI – Originality.AI.” Originality.AI, <https://originality.ai/blog/ai-accuracy>. Acessado em 16 de março de 2026.