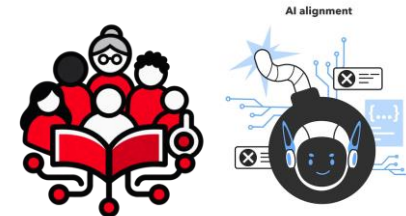




Etapas de Verificação de DeepFake e FakeNews

Equipe de Professores

- Prof. Dr. Ivan Carlos Alcântara de Oliveira (Coordenação) - ivan.oliveira@mackenzie.br
- Prof. Dr. Leandro Zerbinatti- leandro.zerbinatti@mackenzie.br - responsável



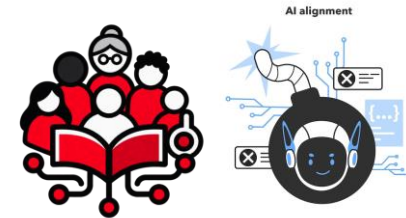
Inteligência Artificial

Ética e Responsabilidade - Impacto Social e Responsabilidade

- **FakeNews e DeepFake:** a IA traz benefícios imensos, mas seu uso na desinformação e criação de deepfakes representa riscos sérios à coesão social, à democracia e à economia. O equilíbrio entre inovação e responsabilidade é essencial para mitigar esses impactos.
- **Ferramentas de IA:** podem auxiliar na checagem de fatos, tais como: **Google Fact Check Explorer, InVID/WeVerify (Extensão de Navegador), Google Imagens (Recurso: "Sobre esta imagem"), TinEye, Sensity (Extensão "Fake Profile Detector"), AI or Not, The Factual (Extensão de Navegador).**

Inteligência Artificial

Ética e Responsabilidade - Impacto Social e Responsabilidade



Cotidiano

'Pequei', diz mulher que contou fake news sobre espancamento do cão Orelha

Do UOL, em São Paulo

08/02/2026 21h49 ⌚ Atualizada em 09/02/2026 11h25



+99



'Pequei', diz mulher que contou fake news sobre espancamento do cão Orelha em SC 💬 99+

Etapas de Verificação de DeepFake e FakeNews



IDENTIFICAÇÃO DA DÚVIDA

O usuário desconfia que uma notícia, imagem ou vídeo possa ser falso e decide verificar sua autenticidade.



ENVIO DO CONTEÚDO

O usuário acessa o site ou aplicativo da IA e envia o material suspeito



PRÉ-PROCESSAMENTO DE CONTEÚDO

A IA remove ruídos e elementos irrelevantes do texto (como emojis), extrai frames e áudio de vídeos e converte imagens e vídeos para formatos adequados à análise.



ANÁLISE DE FAKE NEWS

Um modelo avalia o texto, executando tarefas como: detecção de padrões linguísticos típicos de desinformação, análise do tom emocional, das fontes citadas e da coerência factual, comparação com bases de dados verificadas (sites de checagem e notícias oficiais).



ANÁLISE VISUAL PARA DETECÇÃO DE DEEPPAKES

Caso o conteúdo seja uma imagem ou vídeo, um modelo de Visão Computacional realiza: identificação de padrões anormais de iluminação, piscadas, sincronização labial e textura da pele, detecção de inconsistências em pixels e artefatos gerados por IA.

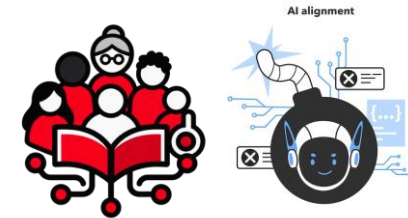


VERIFICAÇÃO DE FONTES CONFIÁVEIS

A IA consulta bancos de dados de agências de checagem e notícias oficiais (Jornal G1, por. ex.) para confirmar a existência de versões reais ou falsas do conteúdo.



Etapas de Verificação de DeepFake e FakeNews



CLASSIFICAÇÃO FINAL E EXPLICAÇÃO

O sistema gera um relatório automático, indicando: o grau de confiabilidade (alta, média ou baixa), motivos da classificação (como uso de linguagem sensacionalista ou padrões visuais artificiais detectado), sugestões de fontes verificadas sobre o mesmo tema.

RETORNO AO USUÁRIO

O usuário recebe a resposta de forma visual e objetiva, por exemplo: “Provável Fake News” – Texto apresenta inconsistências factuais e linguagem manipulativa (escrito em vermelho) ou “Conteúdo Autêntico” – Verificado com fontes oficiais e sem sinais de manipulação. (escrito em verde)

APRENDIZADO CONTÍNUO DO SISTEMA

A IA armazena novos exemplos (de forma anônima e com consentimento) para aprimorar seus modelos ao longo do tempo, tornando-se mais precisa frente a novas técnicas de falsificação.

