

IA Generativa - O novo desafio na e para a Educação

Marco Neves

Agrupamento de Escolas da Batalha



Atualmente o ser humano é desafiado na sua última exclusividade, a Inteligência¹. Nesse sentido, a humanidade enfrenta complexos desafios colocados pela inteligência artificial (IA) através de sistemas cognitivos avançados, acentuados desde o surgimento do ChatGPT em novembro de 2022.

A IA generativa, produzindo textos, imagens e vídeos, desafia a noção de produção exclusivamente humana, provocando reflexões profundas no campo educacional. É crucial entender esta nova realidade e a resposta crítica, urgente e necessária da educação em toda a sua plenitude, preparando devidamente os alunos para estas realidades através da promoção de uma literacia para a IA.

A IA generativa refere-se a um tipo específico de inteligência artificial que pode gerar novos dados, como texto, imagens, vídeo ou música, utilizando dados existentes para produzir novos dados através de uma solicitação (*prompt*) do utilizador. Temos várias abordagens no campo da IA que torna isso possível, como as redes adversárias generativas (GAN) ou os autoencodificadores variacionais (VAE).

No entanto, a atual arquitetura que mais se destaca, sobretudo nos modelos geradores de texto, é designada por “transformer”, que são um tipo específico de arquitetura de rede neural, introduzida no artigo “Attention Is All You Need”² de Vaswani et al. em 2017.

No campo dos modelos de Inteligência Artificial Generativa os que mais se destacam e são mais utilizados, temos dois, os LLM (geração de texto) e os LDM (geração de imagens), sendo LLM (Large Language Models) e os LDM (Large Diffusion Models).

¹ Floridi, L. (2014a). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press.

² Vaswani, A. (2017, June 12). *Attention Is All You Need*. arXiv.org. <https://arxiv.org/abs/1706.03762>

Nos LLM importa destacar o GPT (serve de base ao chatGPT), o LaMDA ,PaLM2 e o LLaMA, relativamente aos LDM (Large Diffusion Models) temos DALLE-2, Stable Diffusion e MidJourney.

Mas qual a razão que nos leva a mencionar e a dar tanta atenção e destaque a estes sistemas suportados por IA, mesmo que comparativamente a outras tantas tecnologias que já foram alvo, no passado, em termos de investigação, análise, integração e estudo no campo educativo?

Estas tecnologias assistidas por IA, ao contrário de anteriores, não são tecnologias passivas. Não se limitam a servir de suporte para e a algo (como a transição de artefactos analógicos para formatos digitais), não se limitam a ser uma interface para criação de recursos e produtos, pelo contrário tem a capacidade própria de gerar conteúdo que se torna muito difícil, e cada vez mais, de distinguir se foi produzido por um humano ou por um destes sistemas.

E acrescente-se, que a cada interação (*prompt*) produz conteúdo, novo original, diferente da interação anterior. Tal não tem uma limitação ou se limita a uma área específica do conhecimento, é, sim, aplicável a qualquer dos domínios temáticos explorados em Educação, desde a História, passando pela Economia ou Filosofia, entrando também no campo das STEM e das artes e criação artística.

Só isto já seria suficiente para estarmos muitos atentos, curiosos, despertos, críticos e prontos a perceber que no campo Educativo enfrentamos desafios que nos obrigarão a desenvolver uma nova mentalidade perante a forma como se desenvolve o processo educativo.

Nesta reflexão vamos focar-nos na componente dos modelos de linguagem, e em particular no caso do chatGPT, pois tem sido sobretudo este que mais questões, dúvidas, inquietudes e sobressaltos tem colocado ao campo educativo³.

O ChatGPT e outros modelos de linguagem de grande dimensão podem ser uma importante e disruptiva tecnologia para a educação. Mas tal só fará sentido se repensarmos as estratégias e abordagens pedagógicas atuais, e reconsiderarmos como se avalia e o que se avalia.

No entanto, existem desafios que as escolas, os professores e os alunos devem ter em atenção:

1. Limitações: estes modelos de suporte estatístico, não tem noção sobre o conteúdo produzido, não tem capacidade de senso comum relativamente às interações que estabelecem e de entendimento do contexto onde operam, bem como uma limitação considerável em termos de memória para recuperar interações anteriores. O seu objetivo maior é prever o próximo *token* (conjunto de caracteres) numa sequência de *tokens*.
2. Preconceito: o ChatGPT e outros modelos de linguagem são treinados num enorme conjunto de dados de texto, o que pode refletir os preconceitos presentes nos dados. Isto pode fazer com que o modelo gere possivelmente texto tendencioso ou até mesmo não credível ou factual, apesar das salvaguardas adicionadas no sistema após o seu treino. Os professores e os alunos devem estar cientes deste facto e devem ser ensinados a avaliar criticamente todo o output gerado pelo modelo.

³ Isso não significa que os restantes modelos de linguagem ou os modelos de difusão não sejam também importantes nos desafios que são colocados à Educação, e lembrando que estamos somente num ponto muito embrionário deste processo de transformação e disrupção.

3. Falta de compreensão: O ChatGPT é um modelo complexo e pode ser difícil para os professores e alunos compreenderem como utilizá-lo o mesmo de uma forma significativa e eficaz. Os professores e os alunos devem dispor de recursos, apoio e suporte adequado para os ajudar a compreender e a saber utilizar e integrar devidamente o modelo.

O ChatGPT, e outros modelos de linguagem, têm potencial para serem ferramentas valiosas para a educação, mas é importante estarmos conscientes dos desafios e tomar as medidas adequadas para os saber integrar de forma contextualizada, responsável e significativa no sentido de potenciar o que mais importa, a aprendizagem dos alunos.

Embora possamos reconhecer que o chatGPT é apenas o início de uma transformação, no que diz respeito ao potencial destas ferramentas assistidas por IA, é importante mencionar que devemos estar cientes de algumas considerações antes da sua integração pedagógica:

- Os alunos devem ser devidamente informados das limitações destas ferramentas e ter sempre uma atitude muito crítica, ponderada e exigente em relação aos resultados obtidos, fazendo sempre o *cross-checking* da informação obtida.

- Os alunos têm de aprender a interagir com estas ferramentas, o chatGPT é o primeiro com este potencial, mas é apenas um tiro de partida neste domínio, e tanto os professores como os alunos estão numa fase de aprendizagem. A proibição é a atitude mais prejudicial que se pode ter para com os alunos. Se a IA generativa é cada vez mais uma ferramenta de trabalho e cada vez mais utilizada, que sentido faz proibi-la nas escolas? Não é um dos pilares básicos da Educação preparar os alunos para o futuro? Se sim, o que nos pode levar simplesmente a proibir a IA Generativa nas escolas?

- A utilização destas ferramentas em determinados momentos e processo educativos, e de acordo com determinados princípios e objetivos de aprendizagem, pode não ser benéfica para a aprendizagem dos alunos, sobretudo em fases em que determinadas capacidades cognitivas ainda não estão devidamente solidificadas.

- A importância da criação de *prompts* não é um processo simples, mas sim complexo, e exige que os alunos tenham conhecimentos de base/experiência sobre a matéria com que vão interagir através destas ferramentas, pelo que é fundamental que os professores trabalhem esta competência com os alunos.

- Os alunos devem compreender que perguntas (*prompts*) sem sentido conduzirão inevitavelmente a respostas e sugestões sem sentido.

- Um sistema como o chatGPT está orientado para a interação em termos de discussão e sugestão de conteúdos textuais que requer competências linguísticas de conversão muito bem estruturadas e fundamentadas. Na essência é importante ter a capacidade de estabelecer um diálogo socrático com estes sistemas de modo a se conseguir obter os melhores resultados. Tal como numa conversa humana, se não soubermos estimular a quem colocamos questões, não obteremos output de qualidade.

- Não se deve esperar que estas ferramentas, treinadas com enormes quantidades de dados de numerosos domínios científicos e outros, sejam “super” especialistas em qualquer área específica, tal como nenhum ser humano é especialista em todas as áreas

do conhecimento. No domínio do ensino e aprendizagem, devem ser encaradas como uma ferramenta de apoio, um co-assistente e um co-autor.

Em conclusão, a IA generativa, representada por modelos como o chatGPT, apresenta um potencial significativo para a educação. No entanto, é fundamental reconhecer e abordar os desafios associados a essa integração. O preconceito nos dados de treino pode resultar em conteúdo tendencioso, exigindo uma avaliação crítica por parte dos professores e alunos. Além disso, os modelos de linguagem têm limitações em termos de compreensão contextual e memória, o que requer uma utilização cuidadosa. É importante fornecer recursos e formação adequados para professores e alunos, a fim de maximizar o uso eficaz dessas ferramentas.

Banir ou proibir é contribuir para o que designamos como “AI Divide”, que comparativamente ao “Digital Divide”, trará efeitos negativos ainda mais fortes e profundos colocando num lado os que sabem interagir e criar com IA, e no outro lado, os que desconhecem a IA ficando na penumbra da ignorância., o que limitará o desenvolvimento integral dos alunos, comparativamente aos que tiverem acesso a esta tecnologia, através da Educação, de uma forma integrada e responsável.

A integração responsável e significativa destes modelos promoverá a aprendizagem dos alunos, mas é necessário um equilíbrio entre a compreensão das limitações e o aproveitamento do potencial que estes oferecem.