AICVF

L'INTELLIGENCE ARTIFIICELLE : VERS UNE NOUVELLE ÈRE DES POSSIBLES

D'UN PROCESSUS « *MAGIQUE* » À UN MODÈLE STATISTIQUE ET PROBABILISTE, POUVANT ÊTRE SOURCE DE CONNAISSANCE, DE CRÉATIVITÉ ET DE CROISSANCE !

Partie 1 : Dîtes nous, comment ça marche…

Partie 2 : Montrez-nous des exemples !

Partie 3 : Et après ?

8 novembre 2024

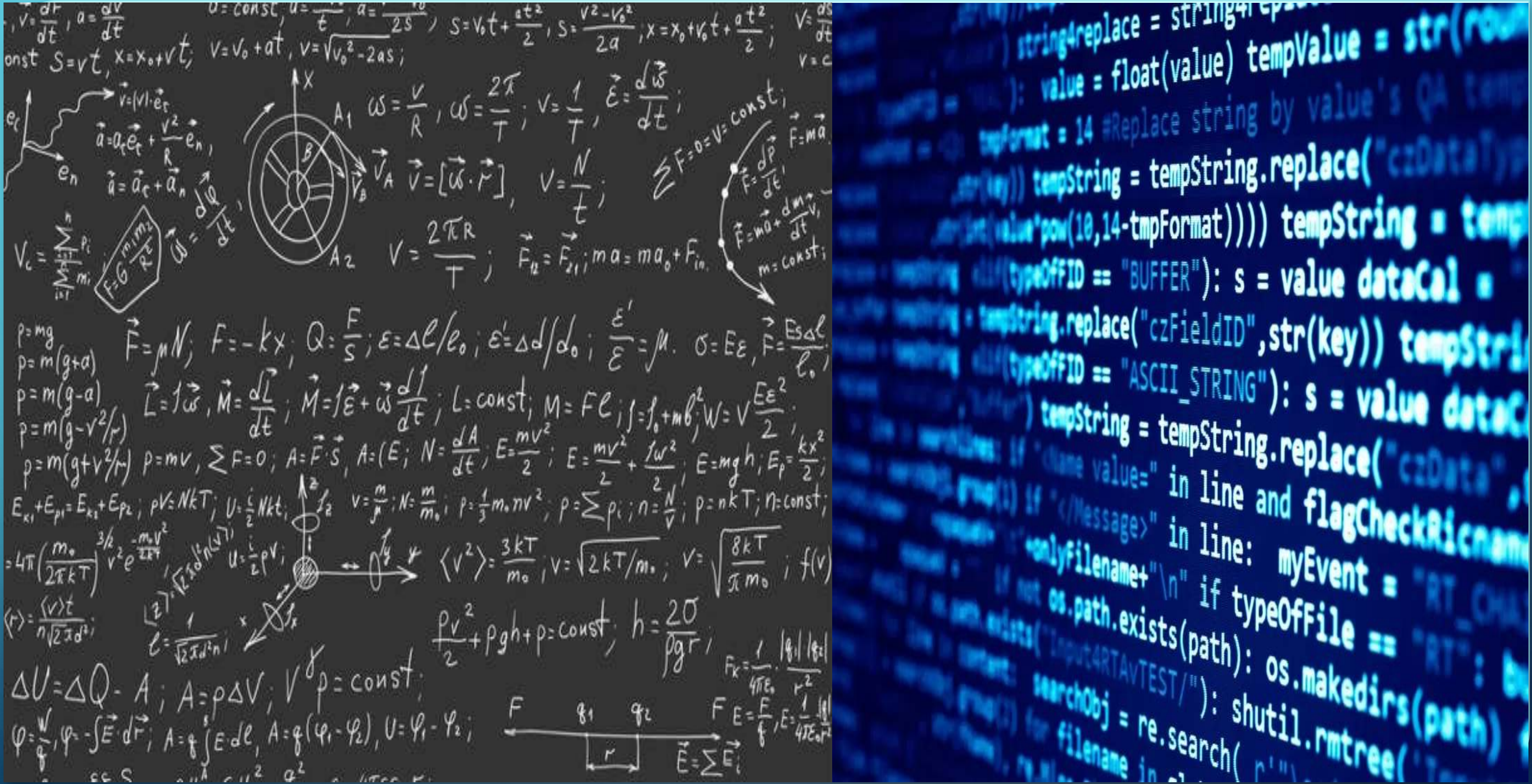
Vincent JOLIVET

L’IA : UN PROCESSUS “*MAGIQUE*”

ENTRÉES

SORTIES

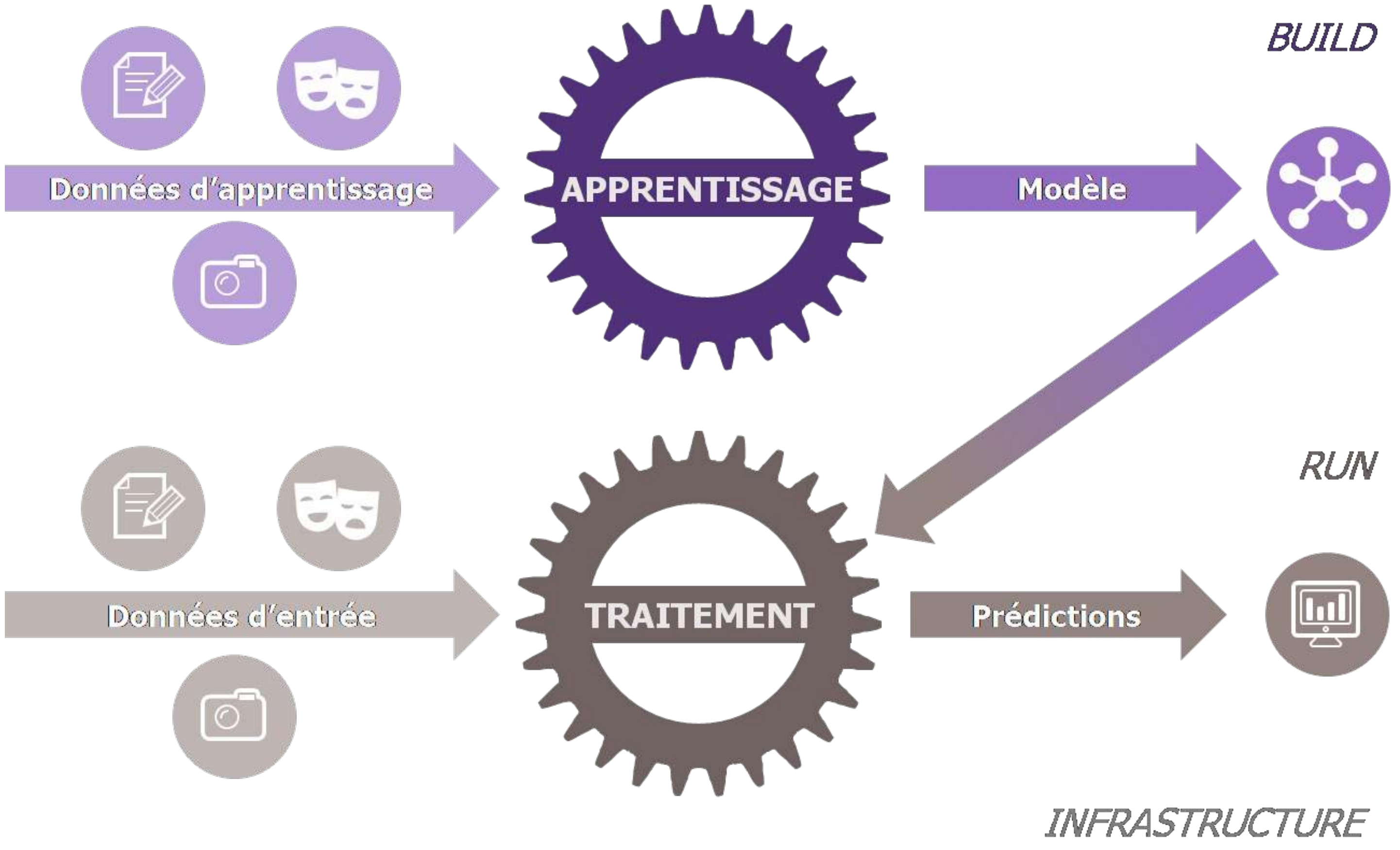
QU'Y A-T-IL DANS LA *BOÎTE MAGIQUE* ?

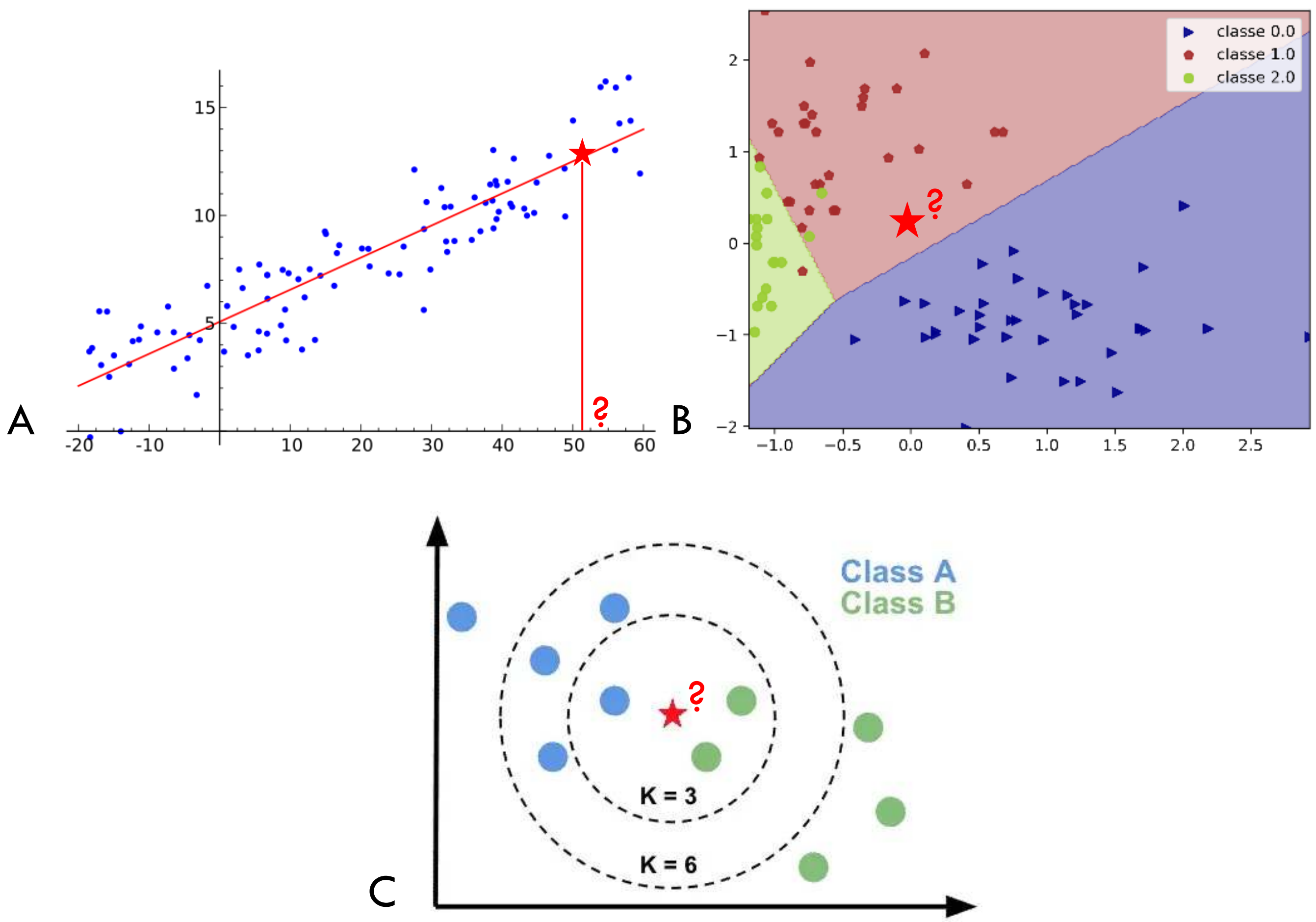


APPRENTISSAGE AUTOMATIQUE (MACHINE LEARNING) (ML)

- Le Machine Learning (apprentissage automatique) est une branche de l'Intelligence Artificielle qui permet aux ordinateurs d'apprendre à partir de données sans être explicitement programmés.

- Le processus peut-être supervisé ou non supervisé.



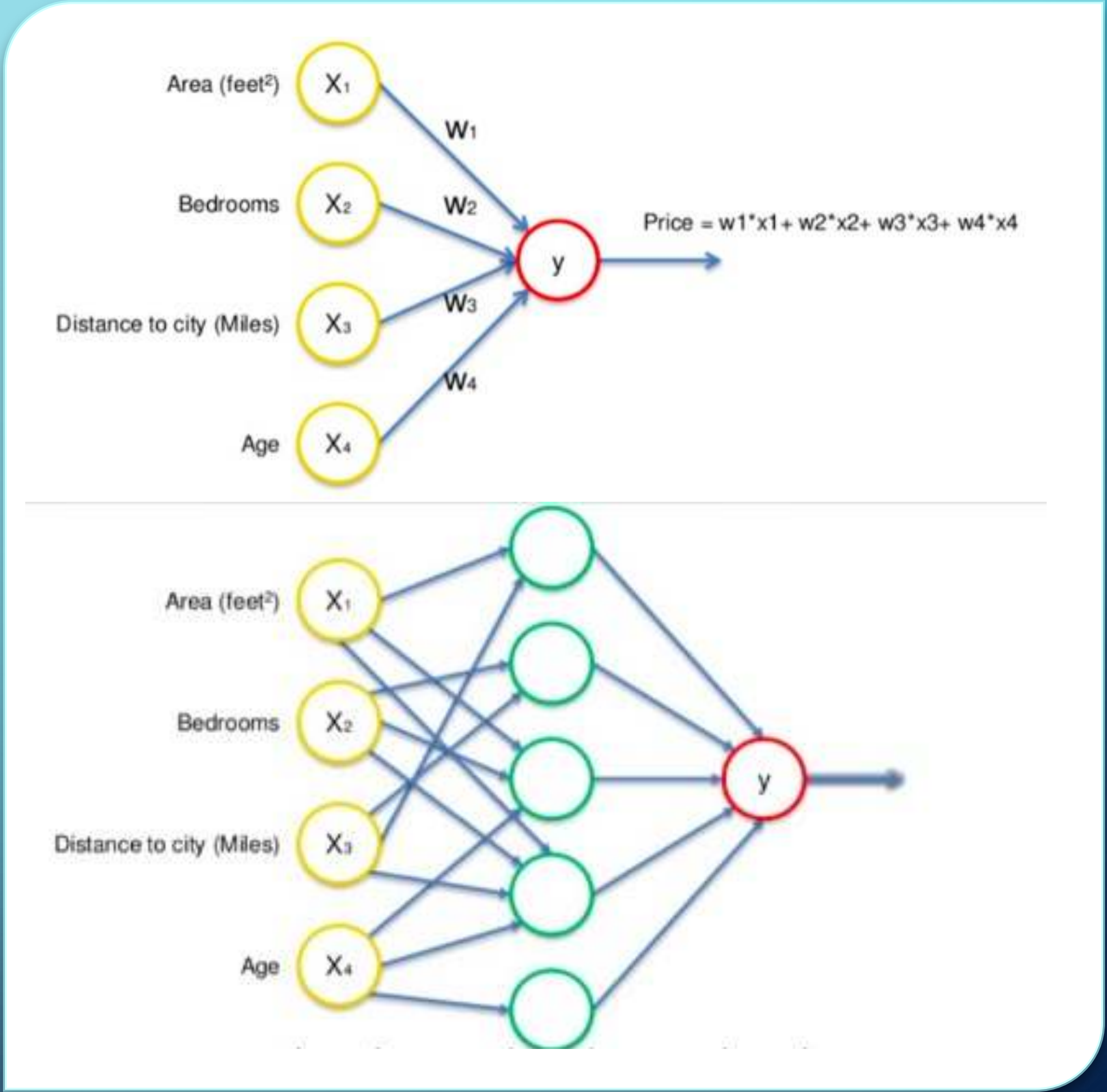


PRINCIPALES MÉTHODES SUPERVISÉES

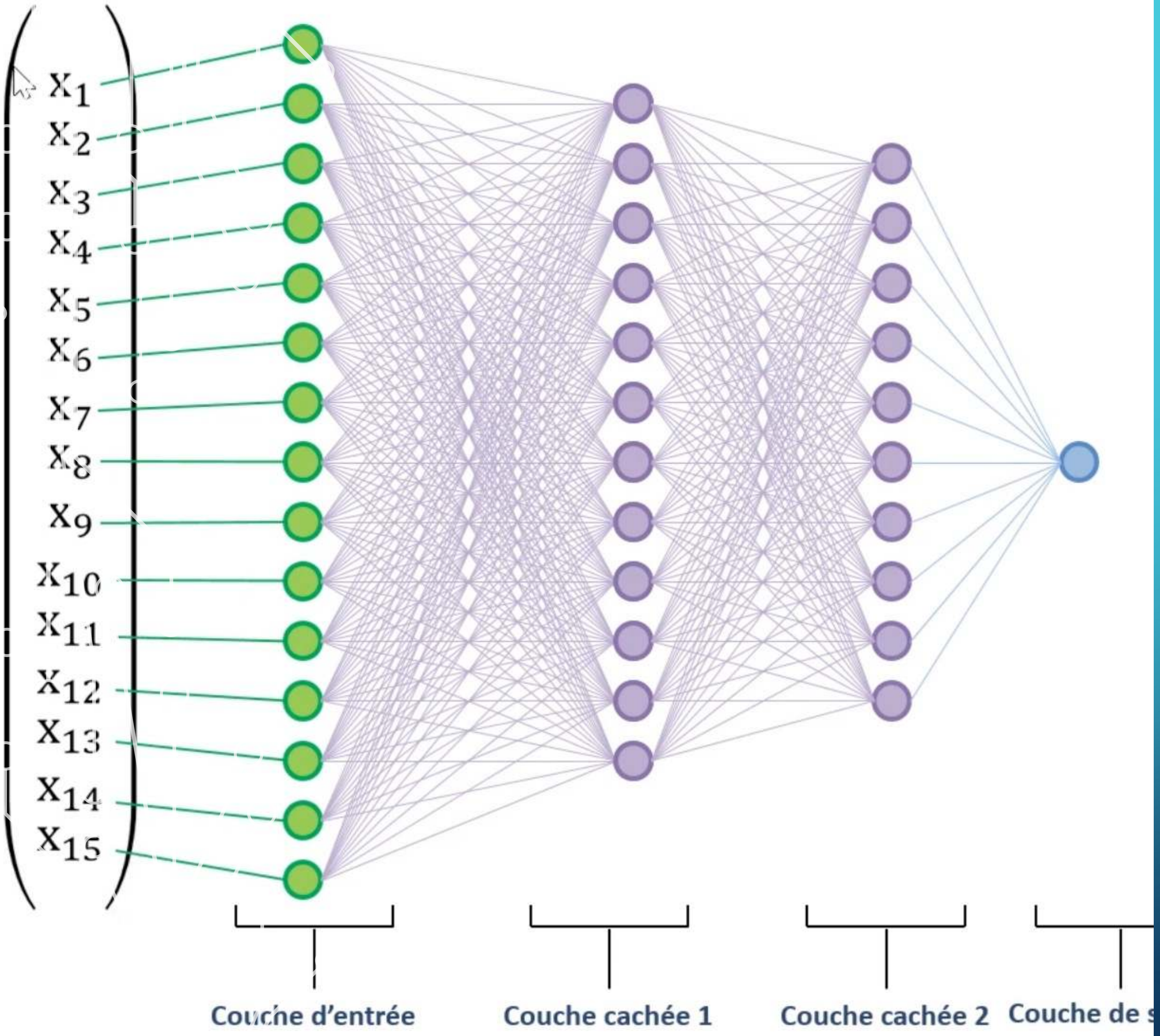
- A : Régression linéaire
- B : Régression logique
- C : KNN (K - plus proches voisins)
- … et d'autres : random forest, reinforcement learning…

RÉSEAU DE NEURONES

EXEMPLE : PRIX D'UNE MAISON



APPRENTISSAGE PROFOND (*DEEP LEARNING*)



UN EXEMPLE POUR TOUT RÉSUMER : LA VOITURE AUTONOME

EN 2017 : L'ARRIVÉE DES TRANSFORMERS

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature.

1 Introduction

Recurrent neural networks, long short-term memory [12] and gated recurrent [7] neural networks in particular, have been firmly established as state of the art approaches in sequence modeling and transduction problems such as language modeling and machine translation [29, 2, 5]. Numerous efforts have since continued to push the boundaries of recurrent language models and encoder-decoder architectures [31, 21, 13].

*Equal contribution. Listing order is random. Jakob proposed replacing RNNs with self-attention and started the effort to evaluate this idea. Ashish, with Illia, designed and implemented the first Transformer models and has been crucially involved in every aspect of this work. Noam proposed scaled dot-product attention, multi-head attention and the parameter-free position representation and became the other person involved in nearly every detail. Niki designed, implemented, tuned and evaluated countless model variants in our original codebase and tensor2tensor. Llion also experimented with novel model variants, was responsible for our initial codebase, and efficient inference and visualizations. Lukasz and Aidan spent countless long days designing various parts of and implementing tensor2tensor, replacing our earlier codebase, greatly improving results and massively accelerating our research.

†Work performed while at Google Brain.

‡Work performed while at Google Research.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova
Google AI Language
{jacobdevlin,mingweichang,kentonl,kristout}@google.com

Abstract

We introduce a new language representation model called **BERT**, which stands for **B**idirectional **E**ncoder **R**epresentations from **T**ransformers. Unlike recent language representation models (Peters et al., 2018a; Radford et al., 2018), BERT is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers. As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task-specific architecture modifications.

BERT is conceptually simple and empirically powerful. It obtains new state-of-the-art results on eleven natural language processing tasks, including pushing the GLUE score to 80.5% (7.7% point absolute improvement), MultiNLI accuracy to 86.7% (4.6% absolute improvement), SQuAD v1.1 question answering Test F1 to 93.2 (1.5 point absolute improvement) and SQuAD v2.0 Test F1 to 83.1 (5.1 point absolute improvement).

1 Introduction

Language model pre-training has been shown to be effective for improving many natural language processing tasks (Dai and Le, 2015; Peters et al., 2018a; Radford et al., 2018; Howard and Ruder, 2018). These include sentence-level tasks such as natural language inference (Bowman et al., 2015; Williams et al., 2018) and paraphrasing (Dolan and Brockett, 2005), which aim to predict the relationships between sentences by analyzing them holistically, as well as token-level tasks such as named entity recognition and question answering, where models are required to produce fine-grained output at the token level (Tjong Kim Sang and De Meulder, 2003; Rajpurkar et al., 2016).

There are two existing strategies for applying pre-trained language representations to downstream tasks: *feature-based* and *fine-tuning*. The feature-based approach, such as ELMo (Peters et al., 2018a), uses task-specific architectures that include the pre-trained representations as additional features. The fine-tuning approach, such as the Generative Pre-trained Transformer (OpenAI GPT) (Radford et al., 2018), introduces minimal task-specific parameters, and is trained on the downstream tasks by simply fine-tuning *all* pre-trained parameters. The two approaches share the same objective function during pre-training, where they use unidirectional language models to learn general language representations.

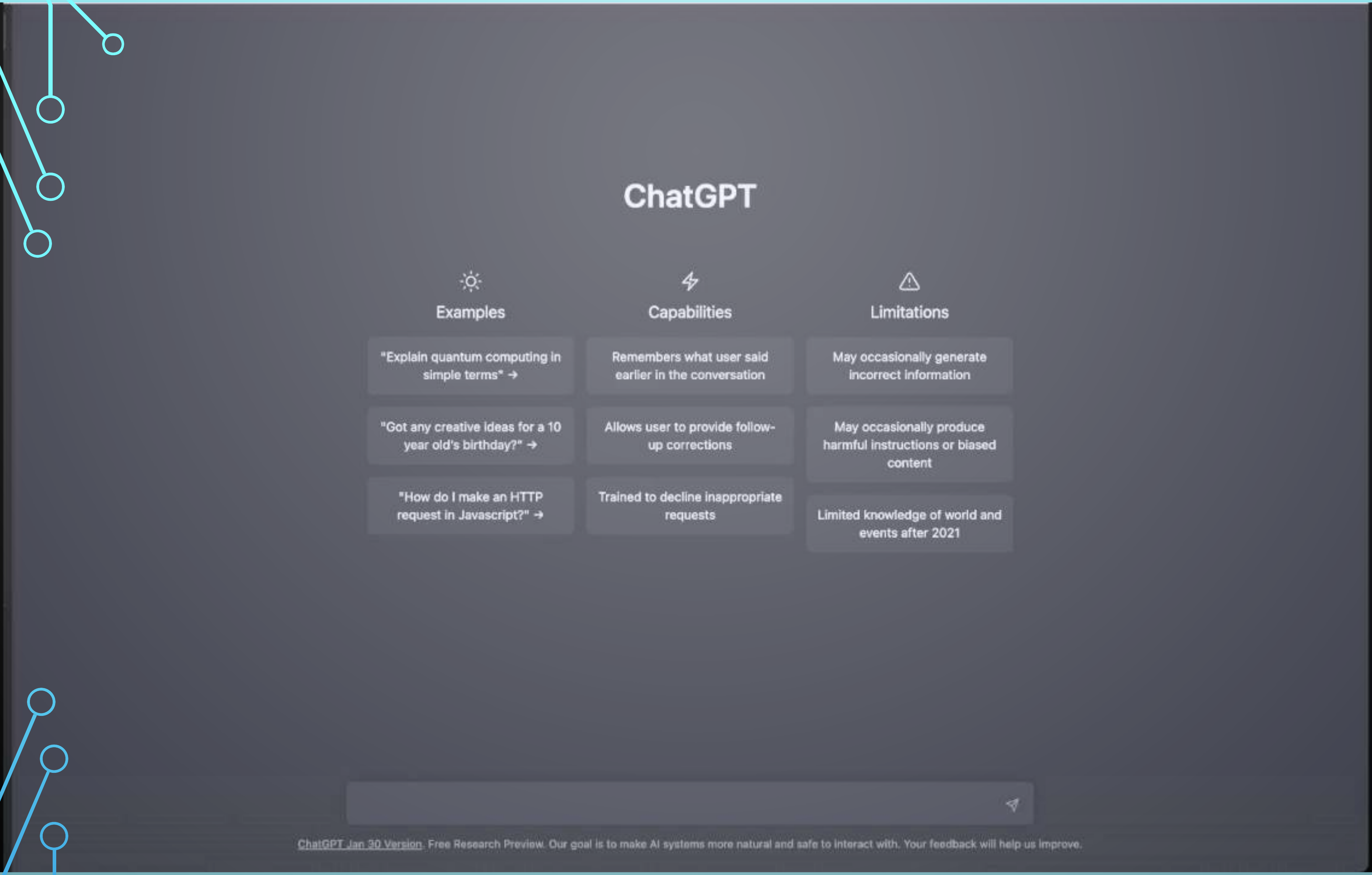
We argue that current techniques restrict the power of the pre-trained representations, especially for the fine-tuning approaches. The major limitation is that standard language models are unidirectional, and this limits the choice of architectures that can be used during pre-training. For example, in OpenAI GPT, the authors use a left-to-right architecture, where every token can only attend to previous tokens in the self-attention layers of the Transformer (Vaswani et al., 2017). Such restrictions are sub-optimal for sentence-level tasks, and could be very harmful when applying fine-tuning based approaches to token-level tasks such as question answering, where it is crucial to incorporate context from both directions.

In this paper, we improve the fine-tuning based approaches by proposing BERT: **B**idirectional **E**ncoder **R**epresentations from **T**ransformers. BERT alleviates the previously mentioned unidirectionality constraint by using a "masked language model" (MLM) pre-training objective, inspired by the Cloze task (Taylor, 1953). The masked language model randomly masks some of the tokens from the input, and the objective is to predict the original vocabulary id of the masked

Proceedings of NAACL-HLT 2019, pages 4171–4186
Minneapolis, Minnesota, June 2 - June 7, 2019. ©2019 Association for Computational Linguistics

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez (2017). Attention is all you need, Advances in Neural Information Processing Systems , page 5998--6008. (2017)

(Devlin et al., 2018) Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.



Démo 1

FIN 2022 : CHAT-GPT

QUELQUES CAS CONCRETS :

1. Établir une demande de remise de majoration de TVA (génération de courriers)
2. Établir une relance de facture (fichier externe ; génération de courriers)
3. Définir une stratégie (réflexion ; génération de documents complexes)
4. Préparer une campagne marketing (génération d'images)
5. Établir un prévisionnel (analyse de données)
6. Concevoir un assistant d'aide à la vente (GPTs)

AUTRES POSSIBILITÉS DES MÉTHODES GÉNÉRATIVES

Midjourney

OpenAI
Sora

D’AUTRES MANIÈRES D’INTERAGIR

La gouvernance des plateformes et des modèles

Les questions éthiques et juridiques

Notre capacité à identifier les résultats des algorithmes de ML

ENJEUX SOCIÉTAUX

CONCLUSION

« … il est important de noter que l'impact de l'IA sur les métiers varie d'un secteur à l'autre et dépend de nombreux facteurs, notamment la maturité de la technologie, la réglementation, l'acceptation sociale et la capacité d'adaptation des travailleurs et des entreprises. Certaines professions pourraient être automatisées de manière significative, tandis que d'autres pourraient être renforcées par l'IA en améliorant les capacités humaines. Il est essentiel que les individus et les organisations s'adaptent et développent des compétences pour tirer parti de ces changements technologiques. »

Chat-GPT 4
Réponse à la question de l'impact de l'IA sur les métiers par Vincent Jolivet.

L'INTELLIGENCE ARTIFIICELLE : VERS UNE NOUVELLE ÈRE DES POSSIBLES

D'UN PROCESSUS « *MAGIQUE* » À UN MODÈLE STATISTIQUE ET PROBABILISTE, POUVANT ÊTRE SOURCE DE CONNAISSANCE, DE CRÉATIVITÉ ET DE CROISSANCE !

Partie 1 : Dîtes nous, comment ça marche…

Partie 2 : Montrez-nous des exemples !

Partie 3 : Et après ?

8 novembre 2024

Vincent JOLIVET

LA CRÉATIVITÉ DANS LES RÉSULTATS

AlphaGo - The Movie | Lee Sedol - Full award-winning documentary (2017)