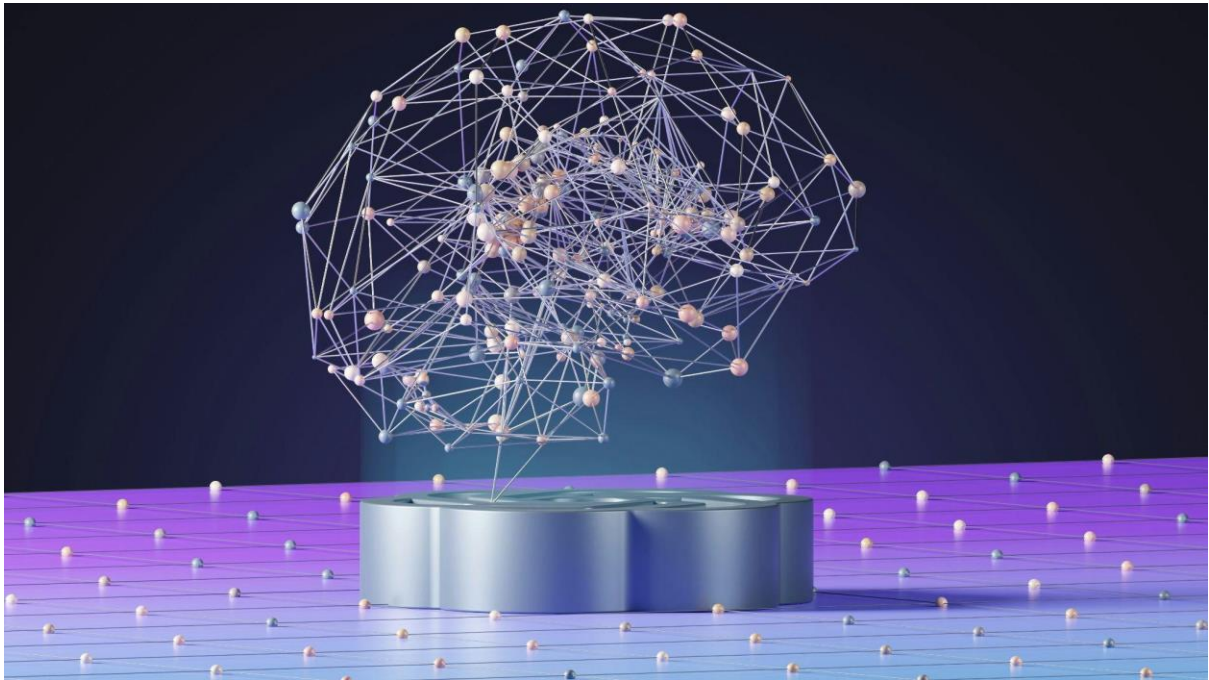




PRESS RELEASE

Golem.ai's training methods and sovereign approach: AI explained at last!



Paris, June 17th 2025,

Golem.ai is a leading explainable AI company whose approach differs from that of traditional machine learning methods. Unlike machine learning-based models, which require large amounts of data and complex training phases, Golem.ai uses neuro-symbolic technology. This approach guarantees technology that is understandable, auditable, and compliant with GDPR standards while avoiding biases and issues related to data storage. It is fully aligned with the concept of digital sovereignty, enabling companies to deploy controlled artificial intelligence independently of non-European infrastructure. In this context, exploring the various types of learning employed in AI can provide valuable insights into the specifics of the startup's approach.

Pre-training: an unsupervised or self-supervised stage

Both unsupervised and self-supervised learning rely on unlabelled data, i.e. raw information without human annotations. In unsupervised learning, an algorithm analyses this data to discover hidden structures, such as clusters or underlying

patterns. This approach is particularly useful for tasks such as customer segmentation or anomaly detection, as it can reveal unexpected trends or behaviours. Self-supervised learning takes this a step further by generating its own labels from the raw data. For instance, it can learn to predict a missing word in a sentence or fill in the gaps in a partially obscured image, much like solving an incomplete puzzle. These techniques are often employed in the pre-training of large language models (LLMs), such as GPT and Llama. These models are trained using millions of texts or images to develop a deep understanding of the relationships between words or visual elements. These capabilities open the door to a variety of applications, such as answering questions, translating text or predicting trends. Both unsupervised and self-supervised learning rely on unlabelled data. Golem.ai stands out by using technology that processes information flows directly without requiring extensive pre-training.

Fine-tuning and supervised and semi-supervised learning.

Fine-tuning is a key step in adapting a pre-trained artificial intelligence model to perform a specific task. Unlike pre-training, which uses large amounts of general data to provide the model with a knowledge base, fine-tuning focuses on more specific and specialised data. For instance, if a model has learned to understand language in general, fine-tuning can enable it to classify emails or detect emotions in texts. In supervised mode, this method uses annotated examples, i.e. data where the correct answers are already indicated. In semi-supervised mode, it combines this annotated data with a large amount of unlabelled information, reducing the effort involved in manual annotation while maintaining good performance. This approach is particularly useful in fields such as medicine and law, where producing annotated data is difficult and costly. Thanks to fine-tuning, models can be quickly and efficiently adapted to specific needs, saving time and resources, as they do not require complete training from scratch. Fine-tuning is a vital step in adapting a pre-trained model. By contrast, Golem.ai does not require a training phase or manual adjustment, offering fast and efficient implementation.

Reinforcement learning:

Reinforcement learning is similar to a game in that a model learns through trial and error. Each time a decision is made, a reward is given for correct action or a penalty for incorrect action. Unlike other learning methods, reinforcement learning does not require pre-annotated data; it learns by interacting directly with its environment. In video games such as AlphaGo Zero, for example, the model plays against itself repeatedly, refining its strategies with each game until it becomes unbeatable. In some cases, such as with large language models (LLMs), human feedback is incorporated into the process: evaluators judge the responses generated by the model, providing information such as 'this response is better than that one', which helps the model produce more useful and reliable results. This method is used in various fields, including robotics (to navigate autonomously), energy management (to optimise consumption) and recommendation systems (e.g. film or product suggestions). Sophisticated mathematical calculations enable

the model to strike the right balance between immediate objectives and long-term results.

The Golem.ai approach:

Golem.ai uses technology based on a symbolic understanding of natural language. This supervised method does not require a training phase or significant data storage, eliminating algorithmic bias and guaranteeing total transparency in the AI's decision-making process. By providing robust, economical and explainable artificial intelligence solutions, the start-up addresses current ethical and accountability issues, while offering businesses high-performance, customised solutions. This also forms part of a strategy of technological autonomy, offering businesses fully controllable AI solutions that are independent of non-European suppliers.

About Golem.ai:

Founded in 2016, Golem.ai is a French start-up specialising in artificial intelligence for human language analysis. Its solution helps companies manage incoming messages and attachments more effectively, offering full explainability for each decision. Today, Golem.ai helps to improve business productivity while ensuring data confidentiality. The start-up has won a dozen awards. Since 2019, it has been part of the French Assurtech and French IoT La Poste programmes, and won the '10,000 Startups to Change the World' award. Golem.ai employs around 30 people.

For more information: <https://golem.ai/fr/>

Press contacts

Agence Volga

Dorian Canit - +33 (0)7 77 78 79 23 - dorian@volgarp.com

Charlotte Mouret - +33 (0)6 58 46 57 99 - charlotte@volgarp.com