Published on *The National Law Review* https://natlawreview.com

Patent Protection on AI Inventions

Article By:

Weiguo (Will) Chen

Yunlai Zha

In recent years, AI patent activity has exponentially increased. The figure below shows the volume of public AI patent applications categorized by AI component in the U.S. from 1990-2018. The eight AI components in FIG. 1 are defined in an article published in 2020 by the USPTO. Most of the AI components have experienced explosive growth in the past decade, especially in the areas of planning/control and knowledge processing (e.g., using big data in automated systems).

AI technology is complex and includes different parts across different fields. Inventors and patent attorneys often face the challenge of effectively protecting new AI technology development. The rule of thumb is to focus the patent protection on what the inventors improve over the conventional technology. However, inventors often need to improve various aspects of an existing AI system to make it fit and work for their applications. In the following sections, we will discuss an illustrative list of subject areas that may offer patentable AI inventions.

Figure 1. AI patent activities by year

(1) Training phase

The training phase of an AI system includes most of the exciting technical aspects of machine learning algorithms exploring the latent patterns embedded in the training data. A typical training process includes preparing training data, transforming the training data to facilitate the training process, feeding the training data to a machine learning model, fitting (training) the machine learning model based on the training data, testing the trained machine learning model, and so on. Different AI models or machine learning models may have different training processes, such as supervised training based on labeled training data, unsupervised training that infers a function to describe a hidden structure from unlabeled training data, semi-supervised training based on partially-labeled training data, reinforcement learning (RL), etc. Common areas in the training phase that may yield patent-protectable ideas include:

- Training data preparation: collecting meaningful training data, balancing positive/negative samples in the training data, labeling the training data, standardization of the training data, encoding or embedding of the training data, synthetic training data generation.
- Novel machine learning architectures: new neural network architecture, hybrid model (e.g., a group of homogeneous neural networks working collectively, or a neural network trained based on training data from a general domain and subsequently transformed by training based on training data from specific domains), hierarchical model (e.g., federated learning).
- Loss function: a new loss function that improves training efficiency.
- Sparsification/Pruning of neural networks: reducing the number of active neurons in neural networks, reducing the number of channels/layers in neural networks.
- Output post-processing: converting predictions to probabilities when definitiveness is harmful.

(2) Application (Inferencing) phase

The application phase of an AI system includes applying the trained models to make predictions, inferences, classifications, etc. This phase generally covers the real application of the AI system. It can provide easier infringement detectability and thus valuable patent protection for the AI system. In this digital era, AI systems can be applied to almost every aspect of our life. For example, an AI patent can claim or describe how the AI system helps the user to make better decisions or perform previously impossible tasks. These applications may be deemed as practical applications that are powerful in overcoming potential “abstract idea” rejections during the prosecution of the AI patent.

On the other hand, simply claiming an AI system as a magical black box that generates accurate predictions based on input data will likely trigger rejections during prosecution, such as patentable subject matter rejections (e.g., a simple application of the black box may be categorized as human activities). There are various ways to reduce the chances of getting such rejections. For example, adding a brief description of the training process or the machine learning model structure helps overcome U.S.C. §101 rejections.

(3) Between software and hardware

Another flavor of AI patents is related to accelerators, hardware pieces with built-in software logic accelerating training and/or inferencing process. These AI patents may be claimed from either a software perspective or hardware perspective. Some examples include specially designed hardware to improve training efficiency by working with GPU/TPU/NPU/xPU (e.g., by reducing data migrations among different components/units), memory layout changes to improve the computational efficiency of computing-intensive steps, arrangement of processing units for easy data sharing, and efficient parallel training (e.g., segmenting tensors to evenly distribute workloads to processors), an architecture that fully exploits the sparsity of tensors to improve computation efficiency.

(4) Robustness, safety, reliability, and data privacy of AI models

The state-of-art AI systems are far from perfection. Robustness, safety, reliability, data privacy, are just some of the most noticeable pain points in training and deploying AI systems. For example, an AI model trained from a first domain may have near-perfect accuracy for inferencing in the first domain, but generate disastrous inferences when being deployed in a second domain, even though the domains share some similarities. Therefore, how to train an AI model efficiently and adaptively so that it is robust when being deployed in all domains of interest is both challenging and intriguing.

As another example, AI systems trained based on training data may be easily fooled by adversarial attacks. For instance, a second deep neural network may be designed to compete against the first one to identify its weaknesses. The safety and reliability of such AI systems will be critical in the coming years and may be important patentable subject matters.

As yet another example, training data in many cases may include sensitive data (e.g., customer data), directly using such training data may result in serious data privacy breaches. This problem becomes more alarming when a plurality of entities collectively train a model using their own training data. Accordingly, researchers and engineers have been exploring differential privacy protection and federated learning to address these issues.

National Law Review, Volumess XI, Number 242