Machine Learning and Patents – A Guide for Patent Attorneys in Chemistry and Life Sciences

Introduction

Machine learning and artificial intelligence have gained an explosion of attention in recent years as a tool that can be deployed against a wide variety of problems. Indeed, the applications of machine learning are no longer limited to the field of computer science. In chemistry and life sciences machine learning is being actively used in techniques such as drug discovery, immunotherapy and molecule behaviour prediction. This cross-over in technologies means that it is essential for IP advisors to up-skill to ensure they are able to provide the best advice to their clients or employers. For many chemists this has come as a bit of a shock. Years of training in advanced chemical disciplines are of limited use when confronted with the question of how best to protect an invention that uses machine learning to predict the quantity of a target organism in an assay.

This paper is designed as a guide for those chemist patent attorneys. Our hope is to help you understand what you need to know, and when to know that you might need assistance from a friendly patent attorney specialising in computer implemented inventions.

What is Machine Learning or Artificial Intelligence?

Artificial Intelligence or AI is a bit of buzz-word that is often used carelessly or to cover anything that involves the use of a computer. In this paper we will predominantly refer to machine learning, but it is helpful to start with a brief guide to the terminology:

Artificial Intelligence

The modern definition of AI is "the study and design of intelligent agents" where an intelligent agent is a system that perceives its environment and takes actions which maximize its chances of success.

Artificial General Intelligence

Artificial general intelligence is a machine that could successfully perform any intellectual task that a human being can. In effect, true intelligence and not data analysis. Most academics believe we are decades from this level of technology.

Machine Learning

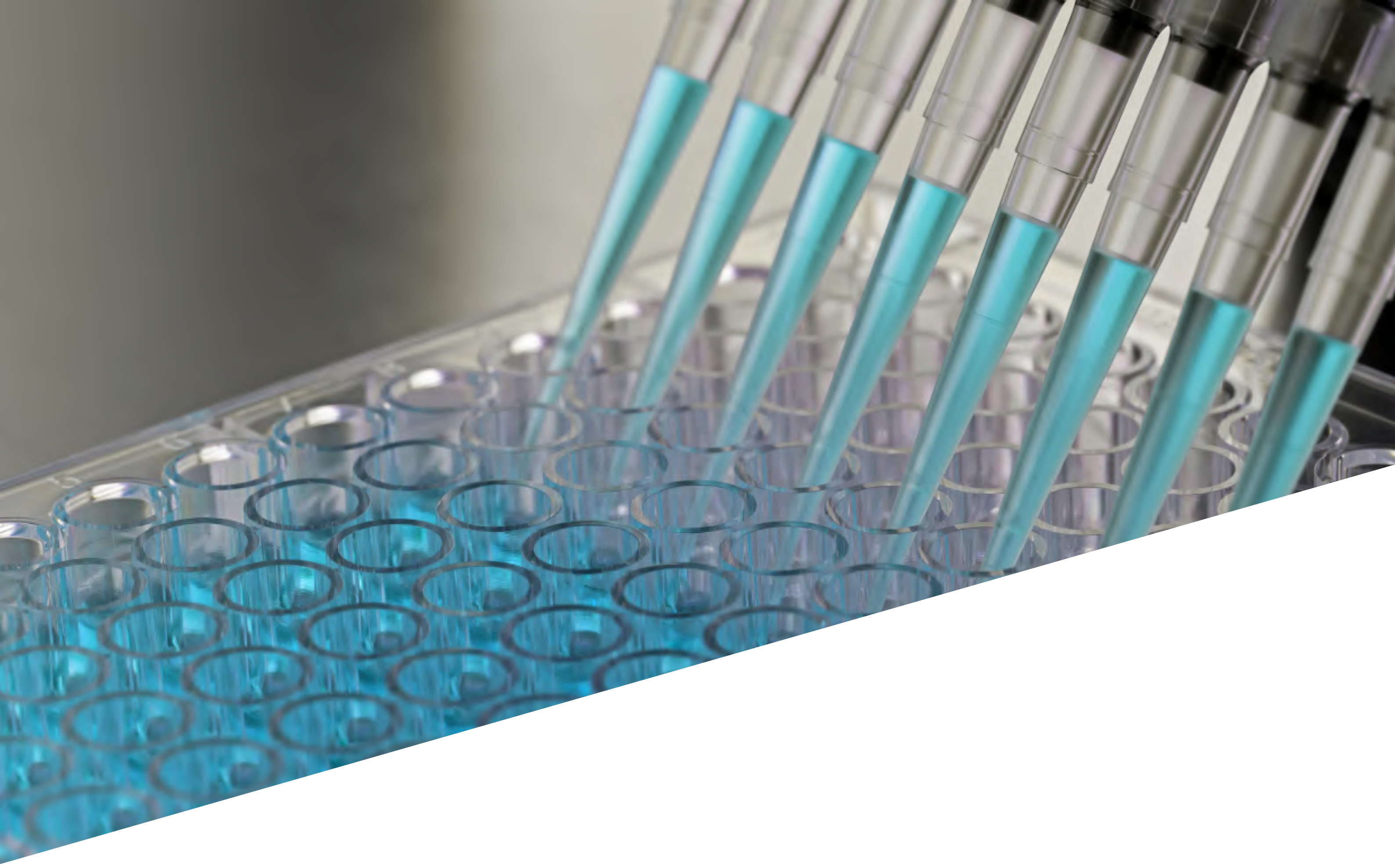
Machine learning (ML) is a sub-field of AI. Machine learning is the programming of a digital computer to behave in a way which, if done by human beings or animals, would be described as involving the process of learning. Machine learning uses statistical techniques to give computer systems the ability to "learn" (e.g., progressively improve performance on a specific taskform data, without being explicitly programmed.

Typically the process of machine learning is building a mathematical model from a set of input training data and then applying that model to a set of test data to provide a prediction or output.

Deep Learning

Deep learning is a subset of machine learning and is based on learning data representations, as opposed to task-specific algorithms. With deep learning it is generally not possible to identify "how it did it". Since there is generally no standard definition of the term we encourage readers to avoid its use where possible.

"AI/ML is no longer confined to the field of computer science – all modern patent attorneys need a basic understanding of the issues"



Examples of Machine Learning in chemistry

ML in chemistry and life sciences has seen unrelenting growth in recent years as diverse applications have been explored and exploited. This in turn has led to an influx of inventions and patent applications in the field.

One such invention, described in 3M applications WO2020/234718 and WO2020/170051, involves a ML model for detecting inhibition of a biological assay, for example in food or water, and quantifying the target organisms using nucleic acid amplification assays. Using ML in this way can reduce the cost and time needed for pathogen detection in food and protect against the threat of foodborne bacterial diseases to public health.

Another innovative use of ML, described in WO2020/243643 (an application by Harvard College) relates to a system in mass spectrometry for detecting and identifying regulatory molecules, such as signalling proteins and membrane receptors, when present in low amounts. This system can also allow high-throughput proteomics studies to be carried out in tandem, greatly increasing the efficiency of complex sample detection and identification.

In the MedTech industry, inventive uses for ML have been found in the operating theatre. WO2021/033061, which is also in the name of 3M, describes continuous video-based product authentication which can be used to automatically provide a medical professional with instructional material that is specific to the product. In addition, the ML model may be trained to recognise the changes a product undergoes when applied to a patient and assess compliance with best practices.

Patentability of Machine Learning in Europe

Machine learning is a special category of computer implemented invention (CII). However, there is nothing particularly special about ML when it comes to assessing the requirements for patentability. At the EPO, the standard method for assessing patentability for CII inventions is to analyse inventive step using the *Comvik* approach (as decided by the EPO Boards of Appeal in T 641/00).

Inventive step at the EPO is assessed using the "problem-and-solution" approach. This involves determining the differences between the invention and the closest prior art, and then determining the effect of those differences. Patent examiners will then establish an "objective problem", which is a hypothetical problem-to-be-solved based on the closest prior art. The question then is whether the solution to this problem offered by the invention is obvious. The problem-and-solution approach is modified slightly by *Comvik* for computer implemented inventions, such that the EPO insists that there must be ***technical*** differences between the invention and the prior art, and that those technical differences must provide a ***technical solution*** to a ***technical problem***. This means that a computer implemented invention needs to provide some kind of tangible, real-world, technical improvement to the operation of a machine if it is to be successful in a patent application.

Exactly the same principles apply when considering the patentability of a machine learning technique. The decisive question for machine learning is whether the technique provides a technical solution to a technical problem, based on the closest known prior art.

The application of this test depends a little on the nature of the invention because, broadly speaking, there are three types of invention that can make use of AI/ML:

AI invention type 1 — 'Core AI'

These are inventions that relate to underlying ML algorithms, absent any particular application. These are mathematical techniques, operating on a computer.

Inventions of this type will be very difficult to protect at the EPO. The EPO is likely to view this category of invention as excluded from patentability on the grounds that they relate to no more than a mathematical method that does not provide a technical effect.

> "At the EPO there is nothing particularly special about ML when it comes to assessing the requirements for patentability"

Fortunately, these inventions will be rarely encountered by most chemist patent practitioners.

AI invention type 2 — Generating a training set or training a model

A ML algorithm is typically trained using a training set, which is a database of underlying data. For example, a ML algorithm that uses optical character recognition may make use of a training set, which is a large number of characters that have been previously sorted by human users.

A patent application could conceivably be based on the training set itself, or, indeed, on the method for training the algorithm. In practice these inventions will be difficult to protect as it will usually be difficult to demonstrate a clear link between a training set and a technical result.

The method for training an algorithm is a process that is undertaken entirely within a computer, without any interaction with the outside world. Following the guidance from the Enlarged Board of Appeal in G1/19, it will be challenging to protect inventions of this nature. As stated in G1/19, no category of inventions is a priori excluded from patent protection, but it will be difficult to establish a technical result when the process is undertaken entirely within a computer, unless it somehow makes the computer work more effectively as a machine.

AI invention type 3 — AI as a tool

This type of invention is the use of ML in an applied field, defined by way of technical effects. This is the most likely avenue for success for patent applicants. Thus, a ML technique that automatically controls the focus of a microscope may yield a technical result. In another example, a ML technique that controls the relative concentration of chemicals in an industrial process is likely to produce a technical advantage. These are likely to be the most common areas where patent attorneys will see ML in action. In these cases, the decisive question will be whether the ML technique involves a technical solution to a technical problem.

EPO Guidelines

In November 2018 the EPO updated their examination guidelines to specifically include sections dealing with ML.

These sections explain that the computational models and algorithms used in AI and ML are considered to be mathematical methods; therefore the general principles regarding examination of these methods apply.

The EPO Guidelines give some examples of AI and ML techniques that find applications in various fields of technology, such as:

Use of a neural network in a heart monitoring apparatus for the purpose of identifying irregular heartbeats. Such a technique would make a technical contribution.

The classification of digital images, videos, audio or speech signals based on low-level features (e.g. edges or pixel attributes for images). These are typical technical applications of classification algorithms.

Classifying text documents solely in respect of their textual content. This is unlikely to be regarded as a technical purpose, rather the purpose is linguistic (T 1358/09), which is non-technical.

Classifying abstract data records or even "telecommunication network data records" without any indication of a technical use of the resulting classification. This is also not a technical purpose (T 1784/06), according to the EPO.

These sections in the examination guidelines are useful in providing some clear examples of the kind of ML techniques that should be considered patentable.

Drafting Considerations

There are lots of things to think about when considering a patent application that involves machine learning. The following is probably the key list of key considerations before a decision is made to proceed:

Is it technical?

What is the purpose of the invention? Does it provide some kind of advantage for a product or process outside of a computer, or does it somehow make the computer better as a technical tool? If the answer to any of these questions is 'yes' then the invention might stand a chance of success as a patent application in Europe. If the answer is 'no' then things are not looking promising in Europe. However, it may still be worth considering patent applications in other jurisdictions, such as the US, where the requirements are more relaxed.

Is the invention clearly defined?

One of the challenges identified by the EPO is potential problems with clarity due to the prevalence of buzz-words. Poorly drafted patent applications are likely to fail if they merely amount to a sprinkling of different terms without any technical detail. As an example, identifying the scope of a claim reciting "applying deep learning" will be difficult: is a neural network implicitly essential to this claim or is the use of classifiers excluded?

Location of infringement

A machine learning process can be categorised into discrete sub-processes. It is common for those sub-processes to be carried out in a distributed manner and potentially across borders. For example, data might be gathered by a smartphone in the UK and sent to a cloud server to apply a model to the data in Ireland using a model trained in the US. Thus multiple steps of a claim would be carried out in different jurisdictions on a claim to the whole process. Wherever possible claims should be drafted to cover steps carried out by discrete entities as well as methods and computer program products. The EPO has also indicated that it may object to the clarity of claims that do not explicitly identify where a particular operation is performed within a distributed system, increasing the importance of considering infringement location up-front when drafting a specification.

Enablement/plausibility

It may be important to detail aspects of the algorithm that some applicants may prefer to keep secret. In some circumstances it may be enough to simply say "inputting the data to an artificial neural network" as long as the skilled person would understand how to put the wider invention into practice. However, it is important to consider the level of detail disclosed carefully as more detail is likely to be required depending on the interplay between the ML algorithm and the application in question.

Detectability/Enforceability

Patent practitioners should consider carefully how possible it will be to detect the use of an invention by third-parties when developing their filing strategies. For example, if the invention is in the way the model is trained and only a model or coefficients are shared with third parties, will it be possible to reverse engineer how the model was trained?

The value of patents should also be questioned for inventions that can be performed anywhere in the world. For example, a method of training a model is something that can often be performed entirely within a computer, which could be located anywhere. An infringer could easily avoid patents by training their model in a jurisdiction where no patent exists. In a case like this it would probably be better to focus protection on the deployment of the trained model, linked to technical advantages, because such a patent is likely to be infringed in the territory in which the technical advantages are experienced.

Patent practitioners should also think about how the model is trained when preparing an application. Hypothetically, if an invention is trained once on a training dataset and then a set of coefficients or a model is applied repeatedly, then there is limited value in obtaining patent protection for the training process. In this case, the real value in the invention is likely to lie in the trained model.

Physical hardware

Practice in computer implemented inventions has changed dramatically over the years when it comes to the amount of hardware that must be described and/or claimed in an application. This is true of machine learning applications, particularly if one considers the distributed nature of many implementations. It is important to balance the need to include hardware into a claim and the likely implementation. Lengthy recitations of conventional computer hardware in the description of an application will not assist in obtaining an inventive step at the UK IPO or EPO. If relying upon a technical effect tied to particular hardware, it is recommended to ensure that the details of this hardware that give rise to the technical effect are clearly described in the context of their interaction with the ML algorithm.

Conclusion

Providing effective advice regarding machine learning is rapidly becoming an essential part of a patent attorney's skillset and we can see no evidence of that trend abating. Many chemists will be relieved not to be involved in drafting these patent applications, but they may still be called upon to provide some initial advice, and hopefully some of those individuals will start to feel more confident in spotting cases that are likely to be patentable. In those cases, it might be best to speak to a colleague who has a specialism in computer implemented inventions, but it is quite possible that a team-based approach will be needed to draft a good patent application and to provide the best advice to the applicant.

Gill Jennings & Every LLP

The Broadgate Tower
20 Primrose Street
London EC2A 2ES

T +44 (0)20 7655 8500
F +44 (0)20 7655 8501

gje.com