Copyright Issues in Deep Learning

Compute Heavy Industries Incorporated
Douglas Gastonguay-Goddard
September 5th, 2023

Copyright in deep learning has three aspects. Datasets, model weights, and model outputs. Industry is currently taking advantage of the ambiguity and applying a laissez faire mentality to developing datasets and models. This may be the most beneficial approach for society, allowing anyone to innovate on anyone else's released work, but may hinder commercialization in the space. This document discusses some of the current issues in copyright in these three areas and proposes some alternative viewpoints.

Datasets

Copyright on datasets is the most clearly defined. In certain cases, copyrighted material can be compiled into collections, used, and redistributed under fair use. Fair use is decided by a multi factor test on a case-by-case basis in court. The fair use test considers things like the commerciality of the use, the amount of the work used, and how the use affects the market for the original.

Being a well defined area, most of the questions are around specific uses and will be decided by the courts. It is interesting to look at the cases like GitHub's Copilot and the image generation models through this lens.

Models

Deep learning models consist of weights and an architecture. The architecture is the portion that is defined in source code. For a specific architecture, this is a fixed set of operations that can be implemented in different languages (e.g.. C++, Python) and frameworks (e.g. TensorFlow, Pytorch). Source code copyright is extremely well defined so we will not address it here. The same architecture can be used for multiple purposes. For example, the same object detection architecture could have weights for detecting animals or weights for detecting vehicles.

The model weights are the learned parameters to an architecture. They are the (generally) floating point values that populate the model's neurons through which inputs pass. The neurons are generally grouped into matrices, so these computations are done massively in parallel via matrix math.

The copyrightability of model weights is currently undefined. These weights are derived through an automated process applied over a dataset (i.e. training). They are just a large collection of numbers. Different numbers can produce very similar results. For this reason, people have

referred to them as facts about a dataset, a distillation of facts in the dataset, or a generalization of those facts. The current consensus online (meaningless) is that these cannot be copyrighted.

Model Weights - Parallels to Object Code

Computer software has a similar issue as model weights. When the software industry was in its infancy, only source code was copyrightable. This meant that the object code, which is also the product of a mechanical process (compilation), was not initially considered copyrightable (by some parties).

This was argued in *Williams Electronics, Inc v. Artic International, Inc* and the court ruled that object code is copyrightable. This was further affirmed in *Apple Computer v. Franklin Computer Corp*. These cases considered the definition of copyrightable material in the 1976 Copyright Act -

> *Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.*

The courts also considered the definition of a literary work, which includes works expressed in numbers -

> *"Literary works" are works, other than audiovisual works, expressed in words, numbers, or other verbal or numerical symbols or indicia, regardless of the nature of the material objects, such as books, periodicals, manuscripts, phonorecords, film, tapes, disks, or cards, in which they are embodied.*

As well as a House Report on the 1976 Copyright Act which clarified that literary works do not need to have literary merit -

> *The term "literary works" does not connote any criterion of literary merit or qualitative value: it includes catalogs, directories, and similar factual, reference, or instructional works and compilations of data. It also includes computer data bases and computer programs to the extent that they incorporate authorship in the programmer's expression of original ideas, as distinguished from the ideas themselves.*

Considering these cases, there is a strong argument that model weights could also be a copyrightable form of expression. If datasets (or the compiled facts that they represent) are considered a form of expression, that could extend to their representation, extension, or generalization as model weights.

Model Weight Perturbation

Further attention should be paid to model weights. These weights are (generally) initialized randomly and then further refined during training. That is, without consideration, two training runs over the same dataset will produce different model weights. Additionally, model weights can be fine tuned or adjusted to not match an original distribution. Should model weights be copyrightable, this aspect should be considered. You do not need an exact copy of the weights to convey the functionality as derived from the contents of the dataset.

Copyrightability of Model Outputs

Further is the question of model outputs. Non-humans cannot hold copyright. However, with generative models there is a strong case that the person prompting the model is creating an original work of authorship by using a tool and could hold copyright over the output.

There is a wide variety of models though, some that produce facts (like the location of objects in an image or the classification of text), some produce art, and others produce text or source code. Due to this, general statements cannot be made about the outputs of models.

The general consensus online (again, meaningless) is that these outputs may be copyrightable, possibly by the user of the model, and almost certainly not by the developer of the model. Clarification in this space would be helpful.

Open Source Clauses Inviable

In open source software source code is distributed publicly under specific licenses. Some of these licenses impart no restrictions on the use of the source code or the compiled outputs. Others, however, impart clauses around the use and extension of the code. The most common is the requirement to release the source code of any derivative works or any works that include or use the original.

Due to the disjoint nature of datasets, model weights, and model outputs, there is no great way to impart such restrictions in any of those three areas.

If a dataset is released with a restriction on its use (non-commercial, non-military, etc.), that dataset can be used to train a model. The model could then be used to produce a new dataset. That new dataset would not hold any restrictions. This works similarly for model weights, they could be used to generate new outputs which are compiled into a new dataset, then that dataset is used to train new model weights.

Currently, it does not appear possible to impart restrictions on a copyrighted work that would extend to a dataset, which would in turn extend to a model trained on that dataset, which would further extend to the outputs of that model. The lack of circularity prevents any “open source” style licenses from being effective in deep learning.

This does not prevent people from releasing datasets, model weights, or model outputs openly. It just means that restrictions on their use are likely ineffective.

Conclusion

Due to ambiguity and the ease of launderability (datasets -> models -> outputs -> datasets -> …) an extremely open ecosystem has developed. This is greatly beneficial to scientific and societal progress.

However, on the commercial side it may limit investment and openness. A company might hold off on releasing a dataset because it would allow anyone to copy their product without restriction. Further, rather than running models locally on device, it would push businesses to a cloud-hosted system where the user would not be able to access a model's weights.

Being able to impart clauses on copyrighted material, datasets, model weights, and model outputs would enable an open source style licensing scheme to emerge. However, such a licensing scheme could create an untenable burden of compliance and also hinder technological and economic progress.

This is an extremely delicate issue and we must strike a balance that incentivizes innovation, openness, and commercialization.