Incompleteness and Machine Learning

Long ago, Betrand Russell and Alfred Whitehead together set out to distill all mathematics down to a logical system which was expressive and self consistent. They were chiefly interested in this pursuit because of a general belief that all mathematical proofs were really logical proofs in disguise. They published the Principia Mathematica, which did a pretty good job of all this, with a relatively small group of required axioms.

Then, Kurt Godel came with his two important incompleteness theorems:

1) No consistent system of axioms can prove all truths.

2) No system can demonstrate its own consistency.

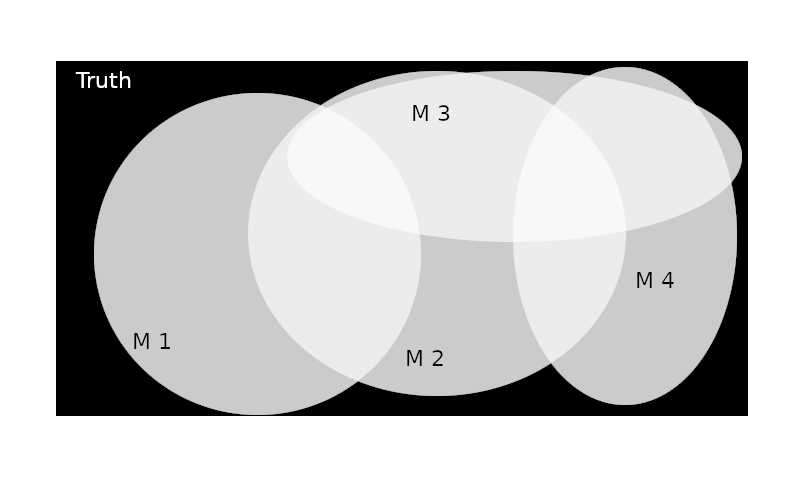
His proofs involved codifying operations as numbers, and showing that it is possible to represent any formal expression or system of expressions as individual numbers. He called his system Typographical Number Theory, and it is considered to be generally equivalent to the system in Principia Mathematica. He then showed that it is possible to create an expression which produces itself, and to show that it is undecidable whether or not it is possible to make a Godel number that represents its own contradiction in this special case.

It may be strange to take these concepts and place them directly onto the context of machine learning. However, effectively, when we train models, we are making a self-consistent system of coefficients or weights that is intended to represent the mechanism that produced some data. In essence, we have a very small and weakly expressive system of equations, and it is meant to describe reality accurately.

Taking the two theorems into account, one way we may interpret them in this context as follows:

1) No model can be identical to the mechanism that created the data.

2) It’s okay for your model to be inconsistent with itself, as this is inevitable anyway.



Each circle represents accurately predicted values. The black space in the rectangle represents ground truth not expressed by any of the 4 models.

The logical extension of these theories doesn’t finish with set or number theory. It suggests that truth itself always will hang outside of even the best model available. Equally that the best model will always have significant error, it is necessary to entertain mutually exclusive systems of axioms all at the same time to reduce this error.

Although there is no published material I am able to find on this relationship, I feel it flows logically that in a scenario in which reduction of error is more pressing than computing restraints, a multilevel multivariate regression model is prudent.

The effect on our discipline is that one can use all manners of model selection – stepwise, experimental, astrological – and find a very good fit. Counter-intuitively, lesser models with higher error will probably account for some of the truth space that wasn’t captured by the best model. At this point, we are discussing the creation of a secondary model which considers the outcomes of generalized models in relation to their effective spaces. The unfortunate translation is the suggestion that combining mutually exclusive models to produce a master construct (a tensor?) is the only way of reducing error… But, it will still exhibit inescapable error.