

Machine Learning, Big Data, Deep Learning, Data Mining, Statistics, Decision & Risk Analysis, Probability, Fuzzy Logic FAQ

Last Updated 3 December 2012, 7:24 AM EST.

What's the difference between machine learning, deep learning, big data, statistics, decision & risk analysis, probability, fuzzy logic, and all the rest?

- None, except for terminology, specific goals, and culture. They are all branches of probability, which is to say the understanding and sometime quantification of uncertainty. Probability itself is an extension of logic.

So what's the difference between probability and logic?

- Not much, except probability deals with uncertainty and logic with certainty. Machine learning, statistics, and all the rest are matters of uncertainty. A statement of logic is a list of premises and a conclusion, and either the conclusion follows validly and the conclusion is true, else it is false. A statement of probability is also a list of premises and a conclusion, though usually the conclusion does not follow with certainty.

In mathematics there are many “logic” theories that have more than one truth value, and not just one universal “logic.” What's up with that?

- The study of “logics” is just one more branch of math. Plus, these special many-valued “truth” logics are all evaluated with the standard, Aristotelian two-value logic, sometimes called “meta-logic”, where there is only truth and falsity, right and wrong, yes and no. There is only one logic at base.

Is probability a branch of philosophy, specifically epistemology?

- Of course probability is part of epistemology, as evidenced by the enormous number of books and papers written by philosophers on the subject, and over the period of centuries, most or all of which remain hidden from mathematical practitioners. See *inter alia* Howson & Urbach, or Adams, or Swinburne, Carnap, Hempel, Stove, that guy who just wrote a book on objective Bayes whose name escapes me, and on and on for a long stretch. Look to this space for a bibliography.

Probability can also be pure mathematical manipulation: theorems, proofs, lemmas, papers, tenure, grants. Equations galore! But the very instant you apply that math to propositions (e.g. “More get better with drug A”) you have entered the realm of philosophy, from which there is no escape. Same applies for *applied math*: it's pure mathematics *until* it's applied to any external proposition (“How much weight will this bridge hold?”)

Isn't fuzzy logic different than probability?

- No. It sometimes has, like mathematics, many-valued “truths” (but so can probability models), but the theory itself is also evaluated with standard logic like probability. Fuzzy logic in practical applications makes statements of uncertainty or of things which are not certain, and that makes it probability. Fuzzy logic is one of the many rediscoveries of probability, but the best in the sense of possessing a cuddly slogan. Doesn't *fuzzy* logic sound cute? Meow.

What is a model?

- A list of premises said to support some conclusion. Premises are usually propositions like “I observed $x_1 = 12$ ” or “My uncertainty in the outcome is quantified by this probability distribution”, but they can be as simple as “I have a six-sided object, just one side of which is labeled 6, which when tossed will show only one side.” The conclusions (like premises) are always up to us to plug in: the conclusion arises from our desires and wants. Thus I might choose, with that last proposition in mind, “A 6 shows.” We now have a complete probability model, from which we can *deduce* the conclusion has probability $1/6$. Working probability models, such as those described below, are brocaded with more and fancier premises and complex conclusions, but the philosophy is identical.

Physical models, that is, models of physical systems, are squarely within this definition. There is nothing in the framework of a model which insists outcomes must be uncertain, so even so simple a (deterministic) equation $y = a + b*x$ (where a and b are known with certainty) is a model. If the parameters (a and b) are not known with certainty, the model switches from deterministic to probabilistic.

Surely exploratory data analysis (EDA) isn't a model?

- Yes it is, and don't call me Shirley. Once a picture, plot, figure, table, or summary is printed and then it is acted on it in the sense of explaining the uncertainty of some proposition, you have a premises (the pictures, assumptions) probative toward some conclusion. The model is not a formal mathematical one, but a model it still is.

What is reification?

- This is when ugliness of reality is eschewed in favor of a beautiful model. The model, created by great credentialed brains, is a jewel, an object of adoration so lovely that flaws noted by outsiders are seen as gratuitous insults. The model is such an intellectual achievement that reality, which comes free, is felt an intrusion; the third wheel in the torrid love affair between modeler and model. See, e.g., *climate models*, *econometrics*.

What's the difference between probability and decision analysis?

- A bet, which if made on an uncertain outcome, becomes a decision. The probability, given standard evidence of throwing a 6 with a die is $1/6$, but if you bet a six will show you have made a decision. The amount of money wagered depends on a host of factors, such as your total fortune, the level of your sanity, whether it is your money or a taxpayer's, and so forth. Decision analysis is thus the marriage of psychology with probability.

Probability models (in all their varied forms) sometimes become decisions when instead of telling us the uncertainty of some outcome, the model insists (based on non-deducible evidence) that the outcome will *be* some thing or that it will take a specific value or state. See *machine learning* below.

Is all probability quantifiable?

- We had a saying in the Air Force which began, “Not only no...” This answer applies here with full force. The mad rush to quantify that which is unquantifiable is the primary cause of the fell plague of over-certainty which inflicts mankind.

Example? Premise: “Some X are F & Y is X”. Conclusion: “Y is F”. Only an academic could quantify that conclusion with respect to that (and no other) premise.

What is a statistical model?

- Same as a regular model, but with the goal of telling us not about the conclusion or outcome, but about the premises. In a statistical model, some premises will say something like, “I quantify the uncertainty in the outcome with this distribution, which itself has parameters a, b, c, ...” The conclusion(s) ignore the outcome *per se* and say things instead like, “The parameter a will take these values...” This is well and good when done in a Bayesian fashion (see *Bayesian* and *frequentism* below), but becomes a spectacular failure when the user forgets he was talking about the parameters and assumes the results speak of the actual outcome.

This all-too-common blunder is the second great cause of over-certainty. It occurs nearly always when using *statistical models*, but only rarely when using *machine learning* or *deep learning* models, whose practitioners usually have the outcomes fixed firmly in mind.

What is a neural network?

- In statistics they are called non-linear regressions. These are models which take inputs or “x” values, have multitudinous parameters associated with these x values, all provided as functions of the uncertainty of some outcome or “y” values. Just like any other statistical model. But neural nets sound slick and mysterious. One doesn’t “fit” the parameters of a neural network, as one does in a non-linear regression, one lets the network “learn”, a process which when contemplated puts one in mind of Skynet.

What is machine learning?

- Statistical modeling, albeit with some “hard code” written into the models more blatantly. A hard code is a rule such as “If $x_{17} < 32$ then $y = \text{'artichoke'}$.” Notice there is no uncertainty in that rule: it’s strictly if-then. These hard codes are married to typical uncertainty apparatuses, with the exception that the goal is to make direct statements about the outcome. Machine learning is therefore modeling with uncertainty with a direct view to making decisions.

This is the right approach for many applications, except when the tolerance for uncertainty of the user does not match that of the modeler

What is big data?

- Whatever the labeler wants it to be; data that is not small; a faddish buzz word; a recognition that it’s difficult to store and access massive databases; a false (but with occasional, and temporary, bright truths) hope that if characteristics down to the microsecond are known and stored we can predict everything about that most unpredictable species, human beings. See [this Guardian article](#). See also *false hope* (itself contained in the *hubris* entry in any encyclopedia).

Big data is a legitimate computer science topic, where timely access to tidbits buried under mountains of facts is a major concern. It is also of interest to programmers who must take and use these data in the models spoken of above, all in finite time. But more data rather than less does not imply a new or different philosophy of modeling or uncertainty.

What is data mining?

- Another name for modeling, but with attempts at automating the modeling process, such that fooling yourself happens faster and with more reliability than when it was done by hand. Data mining can be useful however as the first step in a machine learning process, because if the user has *big data* going through by hand is not possible.

What is “deep learning”?

- The opposite of shallow learning? It is nothing more than the fitting or estimating of the parameters of complex models, which are (to repeat) long lists of human-chosen premises married to human-chosen conclusions. It is also a brilliant marketing term, one of many which flow from the fervid, and very practically minded, field of computer science.

The models are usually a mixture of *neural networks* and hard codes, and the concentration is on the outcomes, so these practices are sound in nature. The dangers are when practitioners either engage in *reification* (man is a loving creature) or when they start believing their own press. as in “If the *New York Times* thinks I’m a genius. I must be.”

The latter is all to apt to happen (and has, many times in the past) because it is to be noted that “deep learning” applications are also simple in the sense that (e.g.) when a human being mouths the sounds *flee* it’s a 50-50 bet the model predicts *free* (perhaps, too, the locutor is saying *free* with an accent; as in what do you call a Japanese lady with one leg shorter than the other? *Irene*.). Accomplishments in this field are thus over celebrated. In contrast, no model, “deep learning” or otherwise, is going to predict skillfully where and when the next wars will occur for the next fifty years. See *artificial intelligence*, or *AI*.

What is artificial intelligence?

- Another name for probability models (but with much hard coding and few statements on uncertainty). Also, See *neural nets* or entries under *New and Improved!*.

What is Bayesianism?

- Another name for probability theory, with hat tip to the God-fearing Reverend Thomas Bayes who earned naming rights with his Eighteenth century mathematical work.

What is frequentism?

- A walking-dead philosophy of probability which, via self-inflicted wounds, handed in its dinner pail about eighty years ago; but the inertia of its followers, who have not yet all fallen, ensures it will be with us for a short while longer.

What is a p-value?

- Something best discussed with your urologist? Also an unfortunate frequentist measure, which contributes mightily to the great blunder mentioned above.

Where can I learn more about these fascinating subjects?

- Glad you asked: [click here for the world’s wonders to be described](#) (scroll down to statistics).

This is only a draft, folks, with the intention of being a permanently linked resource. I’m bound to have forgotten much. There is even the distinct possibility of a typo. Remind me of my blindnesses below. Still to come: within-page HTML anchors.



wmbriggs.com



<http://wmbriggs.com/post/6465/>



<http://goo.gl/ffgS>



