Session on Analytics, Machine Learning, and Data Mining

Mary Brady and Paul Cohen, moderators

Report compiled by Paul Cohen

The session was guided by nine questions, though we strayed:

1. What do we do well already, what are incremental next steps?
2. What are new research issues in the near term?
3. What are really hard problems, where would new funding do most good?
4. How much should be driven by application domains?
5. How good is Watson today in your domain?
6. What else should Watson be able to do?
7. Are these things hard to do?
8. What do you anticipate will be the interaction between domain experts and technology experts?
9. Where are the datasets we need, can they be created or curated?

Because the workshop was about Predictive Analytics, we added some questions:

10. How do you make predictions?
11. What is the role of Big Data in making predictions, from your perspective?
12. How do you manage / control / coordinate predictive analytics?

The morning session got off to a slow start because many attendees didn't know much about how Watson worked, and so interpreted "Beyond Watson" as "any future intelligent system"; and because a significant number of attendees were preoccupied with current practice (e.g., "Never mind Beyond Watson, we're just trying to get beyond Snowden."); and because many were reluctant to describe their jobs. Issues such as interoperability, data standards, security, privacy and access dominated the conversation (as presaged by Dr. Seigel's talk), and there was a general sense of the futility of discussing future intelligent systems while the current ones are still so dumb.

There was a spirited discussion of Enterprise Architectures, data and metadata standards, and other essentially top-down design approaches. Spirited because some participants insisted that nothing else would work, while others insisted that no such standards would ever be sufficient and they would probably make things worse, not better.

The discussion became a bit more prospective when we asked about use cases. In the morning session this tactic turned up three issues: 1) How hard NLP is, 2)

whether we're aiming too low by focusing on predictive analytics, 3) the worrying reliance on supervised ML. In a bit more detail:

- The shallow language processing of Watson probably isn't sufficient for some intelligence analysis tasks. We deferred a discussion of deep semantics because another working group had that responsibility. However, we did note that Watson's approach is "transactional," meaning it answers one question and moves onto another, without any context or persistent structures of any kind; whereas language processing for intelligence analysis involves building a case or telling a story through repeated interactions with analysts and data, all contextualized by the case or story and the analysts's goals. Where Watson was a pure question-answering system, the language processing we want is more like conversation or storytelling.

- Prediction certainly is valuable and it is a step up from forensics and classification, but other modes of inference are at least as valuable. Chief among these is explanation, but there's also simulation, causal modeling, judgments of similarity, and various kinds of inference on social networks. Explanation is important because analysis involves explanation of observations. Again, the transactional nature of Watson seemed inadequate if the task is to accumulate evidence for multiple competing explanations.

- There was a tendency to view Machine Learning as the solution to whatever ails us. One participant insisted that the best use of graduate students would be to annotate cases from which algorithms could learn. One of the moderators was reminded of 1) the free-lunchism that machine learning has encouraged in AI since its earliest days, 2) the parallels between graduate students annotating for machine learning algorithms (today) and graduate students manually entering knowledge (in the expert systems era). Neither is sustainable -- there's no research in either, students get bored, no-one is willing to pay enough to do the job properly, etc. -- so something other than supervised machine learning, or some way to pay for supervision, will be required if learning is to be our salvation. There was also some discussion of whether to trust the results of machine learning.

The afternoon session focused entirely on use cases. The participants were asked to talk around their respective tables and come up with cases, then discuss them in the whole group, then revise.

For the most part, the use cases had nothing to do with Watson, either as it is today or as it might be in the future. We heard about the familiar quartet -- known knowns, known unknowns, unknown knowns and unknown unknowns -- as if its invocation in this forum would somehow transformed it into a computational problem.   We heard about the importance of sensitivity analysis, decision support and management, confirmation bias, and image analysis. We heard about

"empowering" doctors with technology that would actually preempt doctors, by performing diagnosis in advance of the doctor and contacting patients when new results become available. We heard about personalized medicine, detecting disruptive technologies before they are generally recognized, detecting drug interactions, and monitoring a security environment.

The latter two use cases hinted at Watson-like abilities: Because Watson relies on many sources to answer a question, and is good at consolidating information from multiple sources, it seemed a good starting point for applications that require wide surveillance and integration, such as detecting drug interactions or side effects by monitoring physicians' reports. Some participants thought that detecting interactions, generally, would be a good task for Watson. Similarly, a security environment in which dozens or hundreds of people work is potentially very insecure if a "big picture" can be pieced together from the transactions of individuals. Something like Watson, which might monitor many transactions and ask whether they collectively spill the beans, even if none does individually, would be useful.