NeurIPS WORKSHOP ON INTEGRATION OF DEEP LEARNING THEORIES, 2018
PANEL DISCUSSION

PANELISTS

- Moderator: Ankit Patel, Baylor College of Medicine
- Anima Anandkumar, Caltech/NVIDIA
- Tom Goldstein, University of Maryland
- Sanjeev Arora, Princeton University
- Stephane Mallat, College de France
- Richard Baraniuk, Rice University

List of Questions from the Audience

1. Sanjeev Arora's "Agenda for the theory of deep learning" - *What's missing?*
    a. Optimization
    b. Over-parameterization/generalization
    c. New models
2. How to incorporate data models into theory of deep learning?
    a. Sanjeev Arora talk about how understanding gradient descent and loss function is not enough info
3. When you think about machine learning, in what sense do you imagine the machine is learning? (Q from Boris Hanin)
4. The "rethinking style" experiments have resulted in many (including speakers here) that generalization bounds from statistical learning theory are overly pessimistic to be effective. While empirically over-parametrized models appear to have given remarkable results, recent studies show the brittleness of models and ease to find adversarial distributions. If we consider adversarial distributions, then the bounds provided from STL are completely sensible and unsurprising! **So why do we think that its possible to solve adversarial robustness when STL says its impossible? Is the current efforts of theory not considering adversarial distribution not ultimately useful since the theory can easily be broken with such distributions?**
5. Ideas like group theory are very powerful for describing invariants in data. However, are they too rigid for use in real datasets acquired in the wild?
6. What are some useful insights learned from statistical learning theory so far? Will statistical learning theory ever give tight bounds for practical deep learning? Kenji Kawaguchi's work argues that SLT is based on worst-case analysis. Can something like Kawaguchi's analytical learning theory be more useful in this sense (bounds are tight in this case, by definition)? (question particularly aimed at Arora).
7. Why has it taken so long for progress to be made in deep learning theory?
8. besides all variants of neural nets, trees, kernels etc, what other structural models do you think worth exploring? how much does the similarity between biological structures and nn helps the deep nn to be the leading contender in ML?

9. There's been some suggestion that there might be parallels to historical developments in physics and science in general. In what ways do you expect the theory of deep learning/machine learning to be different from that for physics?