Specializing NLP Models through Reinforcement Learning

Bringing together two research lines in NLP:

Reality Check Questions:

- not all error are equal
- challenge sets
- enriched datasets (meta-analysis)

Reinforcement Learning:

- successfully applied in NLP tasks
- interactive with the users
- adaptive learning

Idea: use reinforcement learning to obtain models targeted for specific phenomena*: e.g. avoiding societal bias, strong on an important class

*(rather than simply increasing a benchmark score)