

# Big Data and Security

**Jeffrey Borowitz, PhD**

*Lecturer*

Sam Nunn School of International Affairs

Neural Nets vs Humans on Reading  
Comprehension

# Does the Success on Images Extend to Text?

- Yes and no.
- Neural networks for machine translation are great, close to humans in many language pairs
  - See Module for readings
  - Rather than just text in language pairs, they depend on some paired text and more unpaired text
  - While it's likely possible to do better English-Spanish translation than a machine, Hausa-Thai is likely better by machine
- But translating language (and high end language tasks generally) are inherently more nuanced than image recognition.
  - Great literature is often translated multiple times.
  - Google Translate feels totally functional, but definitely awkward

# Text Versions of ImageNet

- Stanford Question Answering Dataset (SQuAD 2.0)
  - 100,000+ reading comprehension questions, including those unanswerable given text (like SAT reading comprehension)
- General Language Understanding Evaluation (GLUE) with many tasks
  - Co-reference
  - Text summarization
  - Sentence similarity
  - Sentiment
  - Entailment

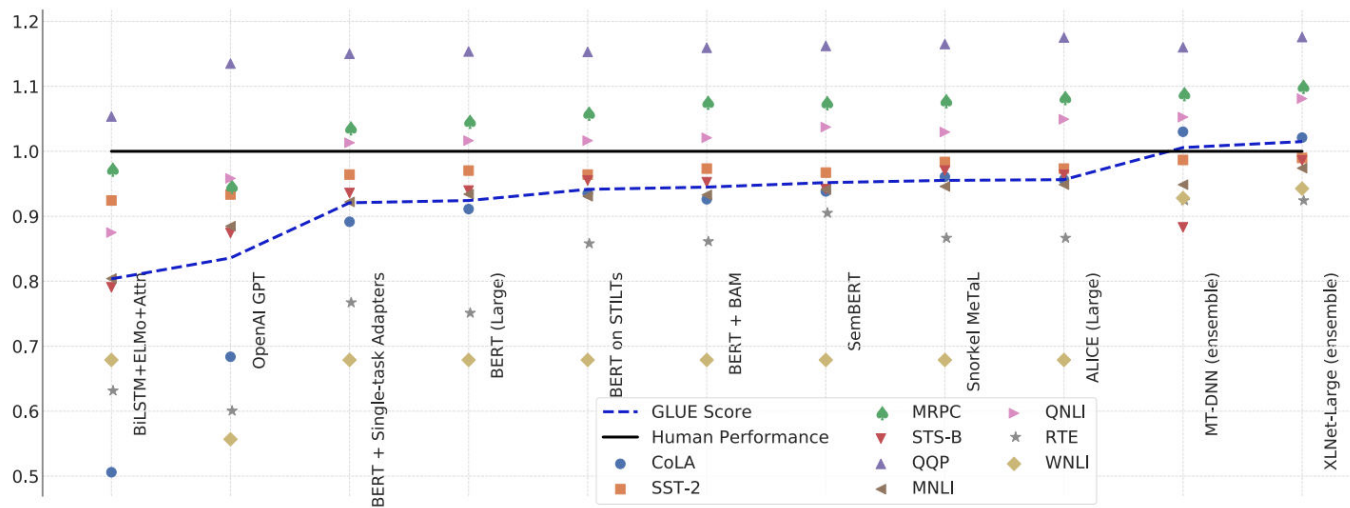
# Solution Evolution: Attention Mechanisms

- Things like Google Translate are built on neural nets using autoencoder-type architectures
- Specifically, there's a long-short term memory type of “neuron” that allows “memory” of state from a lot of words ago
- It turns out that you can do just as well with “only” an attention mechanism with no recurrence, and avoid like 99% of computation costs

# Solution Evolution: BERT

- Based on the attention idea
- Trained on books and Wikipedia
- First, whole model is “pre-fit”
- Then, the model is fine-tuned to answer specific questions
- This model, plus its derivatives, beat people on SqUAD and (most of) GLUE

# Performance on GLUE



- Sentence entailment and coreference resolution are still easier for people
- Meaning similarity is easier for computers (Quora Question Pairs, Microsoft Research Paraphrase Corpus)

# SuperGLUE

- Make GLUE harder Task
- Difficulty Goal:
  - *Tasks should be beyond the scope of current state-of-the-art systems, but solvable by most college-educated English speakers. We exclude tasks that require domain-specific knowledge, e.g. medical notes or scientific papers.*
- Also, include a task called Winogender, to score for gender bias
  - *The surgeon couldn't operate on {his,her,their} patient, it was {his,her,their} son!*

# Text Conclusions

- Computers can also do better than people at specific language tasks
- What does it mean for a model to be better than humans at these tasks?
  - Some humans are better than others at this sort of reasoning...
- Like in images, broadening question space is harder for computers than people
- There will likely be some really nice applications of these models in industry in the next five years



# Lesson Summary

- Translating language is inherently more nuanced than image recognition
- BERT is the state of the art basis for neural nets based on attention
- SQuAD 2.0 and GLUE are evaluation tasks encompassing many parts of language comprehension
- Text based neural nets, similar to image based ones, can show cultural biases