On the Dangers of Stochastic Parrots: Can Language Models Be too Big?

Emily M. Bender*
ebender@uw.edu
University of Washington
Seattle, WA, USA

Timnit Gebru*
timnit@blackinai.org
Black in AI
Palo Alto, CA, USA

Angelina McMillan-Major
aymm@uw.edu
University of Washington
Seattle, WA, USA

Shmargaret Shmitchell
shmargaret.shmitchell@gmail.com
The Aether

2021 ACM FAccT

HUMANE LAB

발제자 : 김태균

2025.01.03

Background

- One of the biggest trends in NLP has been the increasing size of LMs
- Is the ever-increasing size of LMs as the primary driver of increased performance of language technology?
- Any risks?

Trend of Language Modeling

- N-gram models
- Word embeddings
  - using pretrained representations of the distribution of words
- Transformer models
  - larger architectures and larger quantities of data
- Multilingual modelss
- Reducing the size of models
  - knowledge distillation
  - quantization
  - parameter sharing etc.

Risks Associated with LLMs

- **Environmental and Financial cost**
  - CO2 emissions
  - energy-performance trade-offs
    - e.g. increase in 0.1 BLEU score using neural architecture search for English to German translation results in an increase of $150,000 compute cost

- Most language technology is built to server the need of those who already have the most privilege in society
  - **disproportionately affect people** who are in marginalized positions

Unfathomable Training Data

- The size of data available on the web has enabled DL models to achieve high accuracy on specific benchmarks in NLP and CV applications
- However, the training data also contains problematic characteristics
  1. Size doesn't guarantee diversity
  2. Static data/changing social views
  3. Encoding bias

Unfathomable Training Data

1. Size doesn't guarantee diversity
    - Internet access is uneven, further reinforcing inequality
2. Static data/changing social views
    - misrepresent social movements and align with existing power structures
3. Encoding bias
    - encode stereotypes or negative sentiment towards certain groups

Down the Garden Path

- LLM are not performing NLU
  - just focused on measuring whether it performs well on existing and new benchmarks

Stochastic Parrots

LLM is a system for sequences of linguistic forms according to probabilistic information without any reference to meaning : **Stochastic Parrots**

- without any person or entity being accountable for it

Paths Forward

- In order to **mitigate the risks** associated with increasingly LLMs ...
  - Consider the financial and environmental costs of model development
  - Ensure datasets and models are curated to understand and address biases
  - Provide documentation on model suitability and benchmark evaluations for various conditions

Conclusion

- Identified a wide variety of cost and risks associated with LLM
- Carefully consider these risks while pursuing this research direction

Open Question

Are there other methods to predict and prevent the potential risks of LLMs in advance?