

Evaluating Text Generated by Probabilistic Language Models

Emily Sheetz

Overview

- Project Description
- Probabilistic Language Models
- Methods for Evaluation
- Tools and Project Timeline
- Socio-Technical Context

Project Description

- Quantitatively evaluate generated text
- Probabilistic language models
- Generate horror text at word level
- Baseline measurements from text written by human authors
- Compare different language models
- Compare models to human authors

Probabilistic Language Models

- Natural language processing
- Train from datasets of text samples
- Represent probabilities of word sequences in language
- Generate text by predicting next word in sequence of words

N-Gram Models

- Predict n^{th} word in sequence given previous $n-1$ words
- Estimating the probability:

$$P(w_n | w_1 \dots w_{n-1})$$

- Works on word level of meaning
- Probability estimators

Variations on N-Gram Models

- Maximum Likelihood Estimation
- Expected Likelihood Estimation
- Held Out Estimation

Markov Models

- Consider order of words or categories of words
- Work with syntax or sentence level of meaning
- Visible Markov Model (VMM): know the sequence of previous states, such as n-gram model
- Hidden Markov Model (HMM): know the probabilistic function of state sequence, which allows for higher abstraction

Methods for Evaluation

- Perplexity
- Context free grammar
- Probabilistic context free grammar

Tools

- Project Gutenberg
- Python
- Natural Language Toolkit

Project Timeline

- Checkpoint 1: Script for text preprocessing, perplexity
- Checkpoint 2: Context free grammar
- Checkpoint 3: Probabilistic context free grammar, baseline evaluation on existing horror text
- Checkpoint 4: Build n-gram model(s)
- Checkpoint 5: Build Markov model
- Checkpoint 6: Generate text and evaluate quality

Socio-Technical Context

- Evaluating other approaches to text generation
- Humans tell stories to communicate
- Artificial intelligence given human-like intelligence through storytelling
- Horror elicits strong emotions
- Emotions of horror fiction depend on person
- Language models participating in human tasks

Concluding Remarks

- Evaluating text generated by probabilistic language models
- N-gram models and Markov models trained to generate horror text at word level
- Perplexity, context free grammar, probabilistic context free grammar
- Compare models to human authors
- Language models participating in human task of storytelling

Select Bibliography

- Horror Writers Association. “What is Horror Fiction?” *Horror Writers Association*, <http://horror.org/horror-is.htm>. 2009. Accessed Nov. 2017.
- Liddy, Elizabeth D. “Natural Language Processing.” *Encyclopedia of Library and Information Science*. 2001.
- Manning, Christopher D., and Hinrich Schütze. *Foundations of Statistical Natural Language Processing*. Vol. 999. Cambridge: MIT Press. 1999.
- NLTK Project. “Natural Language Toolkit.” *NLTK Project*, <http://www.nltk.org/>. 2017. Accessed Oct. 2017.
- Press, Ofir, et al. “Language Generation with Recurrent Generative Adversarial Networks without Pre-training.” *arXiv preprint arXiv:1706.01399*. 2017.
- Project Gutenberg. “Free ebooks – Project Gutenberg.” *Project Gutenberg*, www.gutenberg.org/. 2017. Accessed Oct. 2017.
- Rajeswar, Sai, et al. “Adversarial Generation of Natural Language.” *arXiv preprint arXiv:1705.10929*. 2017.
- Riedl, Mark. “Why Artificial Intelligence Should Read and Write Stories.” *Huffington Post*, www.huffingtonpost.com/mark-riedl/why-artificial-intelligen_b_8287478.html. 2015. Accessed Oct. 2017.
- Theis, Lucas, Aäron van den Oord, and Matthias Bethge. “A Note on the Evaluation of Generative Models.” *International Conference on Learning Representations*. 2016.