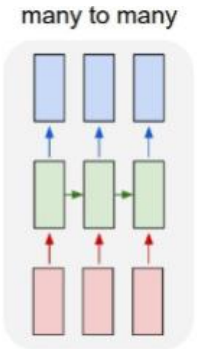
# Chocolate Wafers and Hamburger Casserole

Character and Word-Level Recurrent Neural Network Language Models

# Introduction

- Build language models that can generate output that imitates input text
- Use deep learning models instead of traditional probabilistic methods
  - Recurrent neural network/long short-term memory (RNN/LSTM)
  - Character-level and word-level models
- Compare models using different text inputs
  - Different structures, styles, and lengths
  - Different amounts of data
- Explore models' ability to learn
  - Words, grammar, formatting, context
- Hard to train but fun generate
- "Sometimes the ratio of how simple your model is to the quality of the results you get out of it blows past your expectations..." - Andrej Karpathy



# Methodology: Data

| Category | Type | Documents | Maximum Length | Character Time/Epoch | Word Time/Epoch |
|---|---|---|---|---|---|
| Short | Trump tweets | 4575 | 334 | 300s | 40s |
| Short | Motivational quotes | 4136 | 421 | 180s | 25s |
| Short | Jokes | 8007 | 900 | 900s | 150s |
| Medium | Song lyrics | 100, 500, 1000, 2500 | 3437, 4015, 3991, 4018 | 45s*, 230s, 450s, 1150s | 5s, 30s, 70s, 170s |
| Medium | Food recipes | 2221 | 1497 | 750s | 80s |
| Long | Abstracts | 1281 | 2000 | 530s | 50s |
| Long | Conference Papers | 50, 100 | 40000 | 660s, 1310s | 95s, 185s |

# Methodology: Models

Base Model

- 4 layers with 128 LSTM nodes each
- Bidirectional
- Embeddings dimension 100
- Batch size 512
- Training size 0.8
- Drop out rate 0.25
- Attention layer
- Momentum-based optimizer
- Linearly decaying learning rate

Character-Level Parameters

- Max input length 40
- 10 epochs

Word-Level Parameters

- Max input length 10
- Max vocabulary 10,000
- 50 epochs

*package developed and architecture suggested by **https://github.com/minimaxir/textgenrnn**

# Results

https://github.com/ll3091/ANLY-580-01-NLP-Project/tree/master/GeneratedText

| Character | | Word | |
|---|---|---|---|
| Albert Einstein: A man comes from failure.<br><br>Search One-Liners Like Persistics: A man walks into a bar. The bartender asked, "What do you want?"<br><br>The second guy says, "Yes, as well down the alphabet with a shower.<br><br>The doctor said, "It's just that makes you a ferry son?"<br><br>The bartender says, "I don't know what the hell are you eating a sign of the same time?" | -Stormy There by David Allan Coe-<br>I'm a little more<br><br>Oh when I was on the night<br>When the world is gone<br><br>It's how I feel like I don't know much<br>The way I know<br>When you walk into the sky<br>I know the truth will be<br>The story of the darkness<br>I'm a long way to be there<br>Oh baby I know what I want to<br>I don't want to say goodbye | chinese proverb : the best time to plant a tree was 20 years ago . the second best time is now .<br><br>michael jackson : what do you call a smart blonde ?<br><br>a golden retriever .<br><br>skunk : a married couple are driving down the road , when a car pulls over his hands and says , " let me guess what you ' ve been doing . "<br><br>the man replies , " i can ' t believe that my wife could be dead ! " | - i want to disappear by marilyn manson -<br>i ' m not talking<br>that ' s what i got to say<br>think i ' d die before i ' ll make it known<br><br>you ' ve been out there foolin , but i ' m not thinkin bout you<br>i ' m into your dream , i ' m a god ,<br>i ' m a woman , a<br>i ' m a liar , you ' re the only one<br><br>you ' re all i need , oh yes you are<br>you ' re all i need<br>oh , oh , oh , oh , oh , oh , oh , oh |
| Title: Chicken and No Guinne<br>Categories: Poultry, Main dish<br>Yield: 6 servings<br><br>2 Eggs -- cheeries<br>1/4 c Chopped onion 1/2 c Chopped onion<br>1 c Raisins 1 c Chopped pecans<br>1 ea Egg 1 ts Salt<br><br>Bacon skillet over medium heat. Add onion and butter in a saucepan. Add the sugar and salt to the mixture. Add the sugar and bring to a boil. Remove from heat. Add all other ingredients and mix well. Add the flour and salt. Pour into lightly oiled 8-inch square pan. Bake at 350 degrees for 20 minutes. Serves 4 | | title : chicken casserole<br>categories : poultry , poultry , main dish<br>yield : 4 servings<br><br>4 lb spareribs cut to serve 1 / 2 c chili sauce<br>1 c brown sugar 1 ts honey<br>1 ts cinnamon 2 x eggs<br>1 / 2 ts cloves 1 / 2 c raisins ( optional )<br>1 / 2 c milk<br><br>sprinkle chocolate wafers over bottom of a 7 " springform pan . add water , green pepper , and hamburger . heat to boiling ; reduce heat . simmer uncovered , stirring occasionally unti potatoes are tender , 30 to the meat cooked all . | |
| Stochastic Gradient Descent (DANN) and SVM. We propose a novel approach to the generative model that relates the correlation between a set of arms in terms of the contextual learning and samples. We focus on the predictions of the standard error covers a simple and provide results that are possible with a regression and demonstrate its complex data compared to the expressive posterior distribution is a principled data set is a generalization of the data is a class of structured performance in the sparse variants. | | markov chain monte carlo for matrix factorization<br>a process of truncated gradient descent ( em ) is developed for linear value function $ a $ ( a ) $ lower bound in the regret of the method will be almost positive and is defined by the best tree of the markov chain , the resulting polytope . we develop a new algorithm for using the approximate size of the standard generative models , and can be applied to the multi - task aimed at which it is not the optimal . | |

# Discussion

- Character-level models
  - Generally able to learn words within 2 epochs
  - Learns format and structure of texts
  - Starts with capital letters, ends with punctuation
  - Generate short grammatically correct phrases
  - Usually doesn’t make sense all together
- Word-level models
  - Quickly overfit/memorize training data
  - Some motivational quotes generated were input text
  - Repeated punctuation, new lines, and/or phrases
  - Doesn’t have to learn the words and tends to make more sense than character-level models
- Passable as input text when more figurative/symbolic than literal
- Hard to train, not guaranteed good outputs, but interesting to look at