

# Text Generation with RNNs

James Otto

December 8, 2021

# Introduction

- Recurrent Neural Networks (RNNs) are networks that make one step ahead predictions based on sequential inputs
- Information from previous inputs is preserved in the network's "memory" which is maintained via transferring information in hidden layers between sequential predictions
- RNNs can be used to predict arbitrarily long continuations of sequences by treating predictions as new data

# Introduction

- We are going to consider the use of RNNs in language modeling
- If we treat text as sequences of characters, we can generate text by making predictions with RNNs

# The Goal

- It is a common tactic to train a network on sequences of length  $k$  and pad inputs with whitespace to be of length  $k$  for predictions on "seeds" of arbitrary lengths
  - ▶ We believe this may have adverse effects on a network's predictions, padded seeds do not resemble training observations
- We will train several networks, of lengths 1, 5, 10, and 30 and "bootstrap" prediction process
  - ▶ This avoids the necessity of padding!
- While we are considering the specific application of character generation, the idea of bootstrapping via several models trained on sequences of increasing lengths generalizes to other applications of RNNs as well

# The Data

- Our data comes from Project Gutenberg, a digital library which provides public domain books in plain text
- We will use several author's works:
  - ▶ Jane Austen
  - ▶ Merry Shelley
  - ▶ Miguel de Cervantes

# The Data

```
## [1] "Letter 1"
## [2] ""
## [3] "_To Mrs. Saville, England..."
## [4] ""
## [5] ""
## [6] "St. Petersburg, Dec. 11th, 17--."
## [7] ""
## [8] ""
## [9] "You will rejoice to hear that no disaster has accompanied the"
## [10] "commencement of an enterprise which you have regarded with such evil"
## [11] "forebodings. I arrived here yesterday, and my first task is to assure"
## [12] "my dear sister of my welfare and increasing confidence in the success"
## [13] "of my undertaking."
## [14] ""
## [15] "I am already far north of London, and as I walk in the streets of"
## [16] "Petersburgh, I feel a cold northern breeze play upon my cheeks, which"
## [17] "braces my nerves and fills me with delight. Do you understand this"
## [18] "feeling? This breeze, which has travelled from the regions towards"
## [19] "which I am advancing, gives me a foretaste of those icy climes."
## [20] "Inspired by this wind of promise, my daydreams become more fervent"
## [21] "and vivid. I try in vain to be persuaded that the pole is the seat of"
## [22] "frost and desolation; it ever presents itself to my imagination as the"
## [23] "region of beauty and delight. There, Margaret, the sun is for ever"
## [24] "visible, its broad disk just skirting the horizon and diffusing a"
## [25] "perpetual splendour. There--for with your leave, my sister, I will put"
## [26] "some trust in preceding navigators--there snow and frost are banished;"
## [27] "and, sailing over a calm sea, we may be wafted to a land surpassing in"
## [28] "wonders and in beauty every region hitherto discovered on the habitable"
## [29] "globe. Its productions and features may be without example, as the"
## [30] "phenomena of the heavenly bodies undoubtedly are in those undiscovered"
## [31] "solitudes. What may not be expected in a country of eternal light? I"
```

# The Data

```
## # A tibble: 10 x 11
##   X1    X2    X3    X4    X5    X6    X7    X8    X9    X10   Y
##   <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
## 1 "Y"    "o"    "u"    " "    "w"    "i"    "l"    "l"    " "    "r"    "e"
## 2 "o"    "u"    " "    "w"    "i"    "l"    "l"    " "    "r"    "e"    "j"
## 3 "u"    " "    "w"    "i"    "l"    "l"    " "    "r"    "e"    "j"    "o"
## 4 " "    "w"    "i"    "l"    "l"    " "    "r"    "e"    "j"    "o"    "i"
## 5 "w"    "i"    "l"    "l"    " "    "r"    "e"    "j"    "o"    "i"    "c"
## 6 "i"    "l"    "l"    " "    "r"    "e"    "j"    "o"    "i"    "c"    "e"
## 7 "l"    "l"    " "    "r"    "e"    "j"    "o"    "i"    "c"    "e"    " "
## 8 "l"    " "    "r"    "e"    "j"    "o"    "i"    "c"    "e"    " "    "t"
## 9 " "    "r"    "e"    "j"    "o"    "i"    "c"    "e"    " "    "t"    "o"
## 10 "r"    "e"    "j"    "o"    "i"    "c"    "e"    " "    "t"    "o"    " "
```

```
## # A tibble: 10 x 11
##   X1    X2    X3    X4    X5    X6    X7    X8    X9    X10   Y
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 47    5    14    1    18    7    12    12    1    9    2
## 2 5      14    1    18    7    12    12    1    9    2    32
## 3 14     1    18    7    12    12    1    9    2    32    5
## 4 1      18    7    12    12    1    9    2    32    5    7
## 5 18     7    12    12    1    9    2    32    5    7    15
## 6 7      12    12    1    9    2    32    5    7    15    2
## 7 12     12    1    9    2    32    5    7    15    2    1
## 8 12     1    9    2    32    5    7    15    2    1    3
## 9 1      9    2    32    5    7    15    2    1    3    5
## 10 9      2    32    5    7    15    2    1    3    5    1
```

## Preliminary Results (Simple RNN)

**You will** revenge nearly by this desorated; ce. headd never attempted to resolving not by my musulle and repoused at a convent, or carinisy is us. Weres; a lift of infunciultasing, hapten, and my enulance for the pleasant forteneane air, deageranest, Felix apprayed to an boing earliking place in any other remains and perceived of your patt began. "A more that half kiederer, whose species of the were in my death. Etrebzed this, he spitits and this were, erroneo for several hours not; and on the observe himself. You paused injonttly which is true the greatest events with. My rever by the hour of a soul this opened before her whose tast of some country; a seriem, and they will not resore me, and then to have changed that to my lotter began to engaged. The spprention of my misfortuness are cluelt of its melancholy hour, dress had placed remember me to laugh in my mind whch I had been the difficulty "Masine still sumper I will not rather has begine to any rosts of your mind which had started in preserve in wreck. I inhabited, except the wood; you will not lean hampines and roney from the innocent straw, she bediced her what materials of Agatha and their bither, and condemned by hall first accordingly occupied your atmost agitated in my concertion of daultle purssity and mead and happy and horror and armight of blessed here so subject, but these troubled possessing quantiby and hill with my unusa; they land; but none now with gustious occupations,