

# LSTM 작가 구현

Huere

# 목차

1. 개요
2. 기술 구현 소개
  - A. RNN?
  - B. LSTM?
  - C. Crawling?
3. 기술 구현
  - A. One-hot encoding
  - B. 머신러닝 모델
    - i. LSTM 오픈소스
    - ii. 기타 요소들
    - iii. 모델 테스트
  - C. Crawler
  - D. 결과물
4. 참고자료

## 1. 개요

LSTM은 이전에 주식 그래프 예측 보고서에서 한 번 다뤘던 적이 있는데, 그 때 이해했던 개념들을 실습하고 싶다는 생각으로 시작하게 되었다.

LAFE 프로젝트의 건강한 콘텐츠 공급에 쓰이는 기술도 비슷한 맥락이기 때문에 기획하게 되었다. 정직하고 중립적인 글쓰기는 머신러닝 모델이 맡게 될 것이기 때문이다. 대체물로 소설을 쓸 수 있는 모델을 제작하게 되었다. 약 100권의 소설의 구문을 구사할 수 있는 모델을 제작한다.

## 2. 기술 구현 개요

### A.RNN?

머신러닝 학습을 할 때는 다양한 방법을 사용한다. ANN, CNN, RNN, DNN 등이 있다. 각각 학습하는 방식에 따라 나뉘게 된다. ANN이 가장 많이 알려진 방법이다. Artificial Neural Network의 약자인 ANN은 인공 신경망을 통해 학습하는 기계학습 방식이다. ANN에서 파생된 RNN은 반복적이고 순차적인 데이

터 학습에 특화된 인공 신경망 학습 방식이다.

학습을 할 때에는 Input Layer(데이터 입력 노드<sup>1</sup>), Hidden Layer(실질적인 데이터 학습 노드), Output Layer(데이터 출력 노드)로 이루어진 인공 신경망을 사용해서 학습한다. 이 중 보편적인 ANN은 Hidden Layer를 인공뉴런처럼 여러 층으로 나누어 학습 노드의 결합 세기를 변화 시켜서 학습하는 방식이다. 예를 들면 Input은 16개의 노드로 데이터를 받는다면 Hidden Layer를 64 – 128 – 64개의 노드로 이루어진 3개의 층으로 형성해서 학습한 뒤 Output을 16개로 받는 것이다. (찾아본 자료에 따르면 Input과 Output의 노드의 개수는 똑같은 필요는 없는 것 같지만 나는 똑같이 맞추도록 배웠었다.) 이 경우 ANN의 인공 신경망의 개수와 한 번에 학습하는 데이터의 양 등 여러 요소를 조정해서 여러 개의 노드가 동시에 데이터를 받아 학습하게 된다.

그에 비해 RNN(Recurrent Neural Network)은 연산이 순차적으로 진행된다. 데이터가 노드 한 개를 거쳐 출력이 되면 같은 Hidden Layer의 다른 노드로

---

<sup>1</sup> 노드 : 학습하는 지점 (뉴런의 신경)

상속되면서 학습이 되는 방식이다. 이런 방식의 장점은 데이터의 길이에 상관없이 유연하게 대처할 수 있다는 것이다. 하지만 RNN 알고리즘에는 장기 의존성 이슈도 존재한다.

장기 의존성 이슈는 제공된 데이터와 학습할 데이터의 차이가 크지 않을 때에는 과거 데이터와 크게 다르지 않아 과거 데이터를 기반으로 학습을 할 수 있는 것이다. 하지만 데이터가 차이가 크다면 제대로 학습하기 힘들어진다. 이 문제는 관련 데이터와 그 데이터를 학습하는 노드의 거리가 멀 경우 데이터의 차이가 커 학습 능력이 저하되는 것이다.

또한 일부 노드에서 편미분 값(자료의 기울기)가 지나치게 높아지거나 낮아지는 등 기울기 소실 문제가 발생한다.

## B. LSTM?

기존의 RNN은 매우 간단한 구조인데, 활성화 함수<sup>2</sup>로 데이터의 깊이를 높인다. 그 후 각 노드를 거치며 학습이 되는 것이다. 하지만 앞서 언급했던 이슈

---

<sup>2</sup> 활성화 함수는 여러가지가 있는데, 대표적으로 시그모이드 함수, tanh 함수, ReLu 함수 등 다양하게 있다. 공통점은 이들 모두가 비선형 함수이다. 선형함수가 되면 이런 Layer를 나누는 것이 의미가 없어지기 때문이다.

가 있었다. 이 장기 의존성 이슈, 기울기 소실 이슈를 해결하기 위해 LSTM을 비롯한 다양한 RNN의 파생형이 제시되었다.

LSTM은 기본적으로 RNN의 알고리즘에서 망각 게이트(Forget Gate)를 추가적으로 가지는 알고리즘이다. 이를 이용하여 불필요한 데이터를 제거하여 정제된 데이터를 상속해서 이슈를 줄이는 것이다.

LSTM의 동작 방식은

- (1) 입력 게이트를 통해 데이터를 받는다.
- (2) 활성화 함수(tanh, 시그모이드)로 처리한 데이터 (시그모이드 함수는 0부터 1의 값, tanh 함수는 -1부터 1의 값을 가진다) 값을 갖는 데이터를 만든다.
- (3) 망각 게이트로 데이터를 받는다. 이 경우 이미 거쳐진 데이터가 아닌 독립적으로 시그모이드 활성화 함수로 처리하여 셀을 삭제할 것인지 기억할 것인지 결정한다. 값이 0에 가까울수록 많이 망각한 것이고 1에 가까울수록 온전하게 남긴 것이다.

(4) 입력 게이트의 두 값의 원소별 곱을 진행한다.  
이것이 이번 셀에서 기억할 내용이다.

(5) 망각 게이트의 값을 입력 게이트의 값을 연산한 값과 더한다.

이런 방식으로 여러 이슈들을 해결할 수 있게 설계된 것이다.

## C.Crawling?

크롤링(Crawling), 혹은 스크래핑(Scraping)은 웹 페이지를 그대로 가져와서 데이터를 추출해내는 행위이다. 크롤링하는 프로그램을 크롤러(Crawler)라고 부르며 검색엔진에서도 유사한 것을 필수적으로 사용한다. 대표적으로 네이버에 무언가를 검색하면 모종의 링크 아래에 미리보기로 제공하는 내용이 크롤링한 정보인 것이다.

## 3. 구현

### A.One-hot encoding

One-hot encoding은 선택해야 하는 선택지의 개수만큼의 차원을 가지면서, 각 선택지의 인덱스에 해당

하는 원소에는 1(True), 나머지 원소는 0(False)의 값을 가지도록 하는 표현 방법이다. 예를 들어 A, B, C라는 3개의 선택지가 있다면 각각 [1,0,0], [0,1,0], [0,0,1]의 벡터가 생성된다. 이를 One-hot vector(이하원핫벡터)라고 한다. 컴퓨터는 문자보다는 숫자를 처리하는 데에 특화되어 있기에 선택하는 방법이다.

한글은 영어와 달리 초;중;종성에 따라 매우 다른 글자가 만들어지기 때문에 1453개라는 벡터가 생성된다.

현재 구현한 모델에서는 한 번의 epoch<sup>3</sup>에 10글자로 이루어진 문장을 시드로 글짓기를 한다. 모델은 글자를 처리할 수 없으므로 벡터를 사용해서 배치를 정하는데, 미리 준비한 자료를 기준으로 정답을 정해 놓고 예측을 시작하는 것이다.

## B. 머신러닝 모델

### i. LSTM 오픈소스

이번 프로젝트는 머신러닝을 사실상 처음 만들어

---

<sup>3</sup> 학습 횟수 ; 전체 데이터를 학습한 횟수



보는 것이기 때문에 기존에 만들어져 있던 LSTM 모델을 개조해서 사용하는 것이 낫다고 판단했다. 공개된 코드를 받아와서 학습에 사용되는 글자 수와 학습하는 글자의 파라미터, 학습횟수를 조정해 (각각 40->10, 3->1, 60->200) 학습률을 올리는 동시에 한 번의 epoch에 걸리는 시간을 15분에서 30초로 줄이는 성과를 얻었다. 이 덕분에 총 15시간이 걸릴 예정이었던 모델 학습이 1시간 40분 정도로 단축되는 등의 이점이 있었다.

또한 이번 모델은 저번과 달리 PyTorch가 아닌 Keras 라이브러리를 사용함에 따라 새로운 문법과 개념들을 익히면서 모델을 개조했다.

이 과정에서 LSTM 모델을 구성하는 방법을 알게 되었다. Keras에는 Sequential() 이라는 함수가 있다. 이 함수는 모델을 간단하게 설계할 수 있게 해주는 함수인데, model.add()함수로 블록을 쌓듯이 모델의 계층을 쌓을 수 있게 해준다. 이 모델에서는 LSTM층 1개, Dense층<sup>4</sup> 1개로 이루어져 있다. LSTM의 출력은 1024개이고 Dense층은 1453

---

<sup>4</sup> 완전연결계층, Input과 Output을 모두 연결해주는 역할을 한다.

개를 출력한다. 이 두 층을 쌓아 올려 간단한 모델을 생성하고 국립국어원이 제공하는 말뭉치를 사용해 약 100권의 도서를 200번 학습시켰다.

이 과정에서도 조정이 많았는데, 20188개의 자료 모두를 학습시키니 자료 로드도 지나치게 오래걸리고 모종의 오류로 인해 학습 자료가 5468개까지 급락하는 것이 발견되어 여러 번의 테스트 끝에 100개가 가장 적합하다는 것이 발견되었다. 이 과정은 머신러닝 모델 개발에 흔히 있는 디버깅이며 간단한 모델이어서 이 정도로 끝난 것이지 복잡한 모델은 자료 선정부터 파라미터 조정까지 몇 년이 걸리기도 한다.

활성화함수는 Softmax를 사용했는데, 보통 LSTM에서는 tanh와 Sigmoid 함수를 상용한다. Softmax는 새로운 활성화 함수인데, 모든 합이 1이 되도록 만들어주는 활성화 함수이다. 계층이 2개일 때는 Softmax와 Sigmoid가 같은 식이 된다고 한다. 보통 멀티 클래스(계층이 여러 개)에서 하나의 클래스를 구분할 때 Softmax와 categorical\_crossentropy 조합을 많이 사용한다고 한다. 그래서 Sigmoid 대신 Softmax를 사용한 것

인가 추측한다.

## ii. 기타 요소들

국립국어원의 문어 말뭉치로 학습이 이루어졌다. 이 말뭉치는 책, 잡지, 보고서 등의 자료를 2017년부터 2019년까지 모아 구성되었다. 총 20188개의 자료가 포함되어 있으며 이번 모델에서는 상위 100개의 자료만 사용했다. 이 모델을 LAFE에 활용할 수 있게 제작하려면 신문 말뭉치나 문서 요약 말뭉치를 사용해 학습하면 될 것이다.

Json 파일을 다루는 방법을 알게 되었다. Json파일은 다량의 데이터를 제공하는 데에 특화된 파일로, 머신러닝에서는 매우 흔하게 쓰인다. 저번 Skeleton Detection의 모델 생성 취소 이유 중 하나가 Json파일과 영상파일의 매칭 실패이기도 한 만큼 이번에는 Json파일을 다룰 수 있게 배워야겠다고 생각했다. 파이썬의 Json 모듈을 이용하여 로드했다.

## iii. 모델 테스트

머신러닝은 epoch를 반복할수록 loss가 줄어들면

서 정확도가 높아진다. 다음은 epoch가 1일 때의 결과물이다.

----- Generating text after Epoch: 0

----- Generating with seed: "구, 당꼬 바디 란"

구, 당꼬 바디 란 하사는 어느 새로 서 하게 될 시간 막 상사를 가 어느 데 사내가 지금 시습니다. 그 사람들이 있다. 그녀는 하사의 안에 후 번 죽어서 있다. 그는 그는 전쟁을 나 난민들은 집에 명의 모양이다. "그 사내 있다가 내려 왔다. 자기 시간"

하사가 없다. "아니 그들 여인 자기 위로 터 오늘 동안 있습니다. 그들은 방마시 부르는 자기 남 사 한 형 안 되어 효진은 있다. 그는 자기 그 사람이 얼굴에 잡는 것이다. 그러나 그는 발태인 개를 바로 로 리리 내려주고 있다. "여, 모두 하구 따라게"

장 한 명의 손으로 무려를 달고 있다. "야: "어요? 그들은 만 되어 있었습니다." "제 나 국 사람 전에 거나 그 동렬의 그들은 차가 내리를 수으로 차려 이 있다. 그는 다시 가까이 않습니다. 한 국 사람은 사람이 보습니다." "나 앓는 한다. 그들은 "

난민들을 받고 있다. 그는 이 분명히 땀을 오시오?" "그 사람들 속에 집 안 하는 거제 니다. "아니 말을 돌아본다. 그는 성한 말이 그 사람들이 다시 걸어서 주어 있다. 효진이 동무 그 안 안 안 뒤으로 내려든다. "오늘어 바라본 장님은 수 있다. 그는"

대위는 "지금 어느 가로 당신 긴 복정하게 지금 새로 나 있습니다. 그녀는 다시 시작한 소대가 무금가 돌려 있다. 그는 바로이 담배 있다. "그 사람이 그러나 보고 있다. 그는 모두 없이 그 부리는 상사는 채 마을 하고 있다. 그가 어서 속에 도로 앞 "

가 부물 안 있다. 그는 지금 이 정도로 제가 지금 그들은 그는 시작한다. "아니 요?" "제 대가 소영을 걸어 있다. "자네 지금 그는 것이 없다. 서 명이 자기 이쪽에 서 올라고 고개를 보고 있다. 일는 마리 이리 그 사람들에게 마을 보고 있다. 그 사람가"

따라 지 앓는 어요?" "어니 요? 하게 사내 있는 ?" "남새 조금 좀 말 아 착선 것으로 지나 때문이 적인 후 모두 번 있습니다." "아니 국 오늘도 동무 그는 네가 아니 나 형 부리 아무 명의 안 사람은 '왔이 야 속 안으로 나와 듯 바라 고개를 잡고 있었습니다. 그"

다음은 epoch가 200일 때의 결과물이다.

Epoch 200/200↵

214301/214301 [=====] - 79s 369us/step - loss: 0.0358↵

↵

----- Generating text after Epoch: 199↵

----- Generating with seed: "은 연거푸 마른기침"↵

은 연거푸 마른기침을 토해내며 여맹원의 끈덕진 권유를 애원조로 거절하고 있다. "동무 말씀은 저도 충분하게 이해하겠습니다. 저도 실은 이번 해방 전쟁에 지원할 생각이 없지 않은 사람입니다. 허지만 저는 동무도 보다시피 폐결핵 제3기 환자입니다↵

. 천천히 걷는 것도 숨이 차서 이렇게 혈떡이는 형편인데 어떻게 제가 총을 들고 놈들과 싸우겠습니까?" "동무가 폐병인지 아닌지는 병원에서 의사들이 가려낼 거예요. 그런 건 조금두 염려하지 말고 동무 뜻대로 과감히 지원하세요. 동무가 만일 폐병 ↵

이라면 우리 인민 군대에서 동무 병까지 치료해줄 거예요." "박숙희 동무, 여기 정말 훌륭한 동무가 한 명 있소." 박숙희라고 불린 여기서 급번 자갈 이백철 이리의 나감들이 불편하하. 비비 가지 않고 조금런 그들을 일어보려 아바로이 대신 소식거나 ↵

낙해대장 대부분의 부산 외자와에 어한 이후이 전쟁 전에 사람들은 모두 이이 내 어루 가까이 해방시라 에는 가지 못할 것 같았습니다. 제가 27일날 잠깐 만났는데 우연히 말끝에는 별루 큰 불려은 없이 처음 죽거고 시체하기가 자는 에서워 서울 골짜↵

아래로는 요행하지 않는다. 그는 자신을 언제 어디서나 장사꾼이라고 떳떳하게 밝혔으며, 남들도 자기를 장사꾼으로 대하도록 끊임없이 일깨우며 자기를 굳게 지켜온 인물이다. 효진은 그러나 이런 한익에게 난리 이후 최근 며칠 새에 예전에는 미처 ↵

론이고 적어도 효진이 겪어본 사람들 중에는 가장 속이 깊고 언행이 신중하며 참을성이 많은 사람이다. 그는 자기의 말이나 역활이 꼭 필요하다고 생각될 경우에는, 전문 농사꾼들도 무색할 정도의 엄청난 농사 지식을 예사롭게 털어놓는다. 가마니 치 ↵

이런 식으로 정말 큰 개선이 이루어지는 것이 보이지만 아직 문법이나 맥락이 어색한 부분이 많다. 짧은 시간 내에 큰 진전이다.

또한 앞선 코드에서 모델을 저장하고 다시 사용하는 코드를 새로이 만들었는데, 기존 코드는 모델을 저장,

불러올 수 있는 기능이 없었는데, 새로이 개조한 코드는 모델을 저장하고 불러올 수 있는 기능을 추가했다. 또한 내가 만든 크롤러와 호환되면서 그 크롤러로부터 정보를 정상적으로 넘겨받을 수 있게 개조했다.

## C.Crawler

네이버는 공식적인 검색API를 제공한다. 자신의 네이버 계정 앞으로 애플리케이션을 등록하면 API 계정을 발급해준다. 또한 크롤링 코드를 제공하기 때문에 네이버 뉴스탭과 블로그탭에서 크롤링 대상 링크를 쉽게 얻을 수 있다. 두 개 모두 테스트해본 결과 정보의 양이나 크롤링 가능한 사이트<sup>5</sup>의 양이 뉴스탭이 월등히 좋았다. 그래서 사용자가 원하는 키워드를 입력하면 네이버의 뉴스탭에서 링크를 긁어오도록 설정했다. 또한 이 과정에서 세부적으로 검색 결과를 조정할 수 있도록 몇 개의 링크를 가져올 것인지, 유사도순으로 검색할 것인지 최신순으로 검색할 것인지도 설정할 수 있게 했다.

---

<sup>5</sup> 사용하는 크롤링 라이브러리로 내용을 긁어올 수 없는 사이트가 간혹 있다.

이렇게 수집된 링크들은 파이썬의 newspaper 라이브러리를 통해 내용이 수집된다. BeautifulSoup4라는 크롤러를 사용하는 선택지도 있었지만 이 경우 언론사마다 내용이 있는 html값<sup>6</sup>이 다르기 때문에 언론사 사이트를 전수조사를 필요로 하기 때문에 자동으로 본문을 수집해주는 newspaper 라이브러리를 사용했다. 수집된 링크들을 newspaper로 크롤링해 본문 내용을 수집한 후 gensim 라이브러리로 summarize<sup>7</sup>한다. 그러면 사용자가 설정한 수 만큼의 요약된 본문이 나오는데, 이 본문들을 합해서 큰 본문을 한 개 만든다. 그리고 이 본문을 또 한 번 summarize한 뒤 머신러닝 코드로 요약된 본문을 넘긴다. 이후 머신러닝 코드에서는 이 본문에서 랜덤한 지점의 10글자를 시드로 사용하게 된다.

## D. 결과물

테스트하는 키워드는 '바지락'으로 설정했고 포스트는 10개를 수집했으며 유사도 순으로 검색하게 했다. 본래의 모델의 레퍼런스대로 10글자를 시드로 제공

---

<sup>6</sup> 웹사이트의 내용은 html에 저장되는데 본문이 있는 html값은 웹사이트마다 다름

<sup>7</sup> Gensim이 제공하는 글을 요약해주는 함수

했다. Loss는 평균적으로 0.04대에서 결과물이 생성되었다. 1000자를 출력하게 했으며 모델은 개조된 코드에서 제작한 모델을 사용했다. 전체적으로 정말 사람이 쓴 것 처럼 부드러운 글이 생성되지는 않았지만, 대체로 문장의 구색이 맞춰져 있고, 자신의 데이터셋에 있는 자료가 아닌 랜덤한 자료를 제공해도 온전한 성능을 발휘할 수 있음을 알 수 있었다. 하드웨어 성능이나 코드가 전문적이지 않음에도 불구하고 준수한 성능을 보여주었다.

Presented word : "봄 기운을 담은 봄"↵

봄 기운을 담은 봄꽃 려나 지만 있다. 부근 남을 김하사 역시 앞에서 엄물난 호 나쳐나 문 내는 북한군이 남한선의 강장으로 작정오?" "대구가 여인있은 대원 두 명이 장바닥는 발 앞소리는 분이에 내려오시 기가의 군을 받는 것은 고향 남포 사람이면 누구나 다 아는 일이다. 그는 문득 방금 전에 만난 또 한 명의 고향 사람 최가를 머릿속에 떠올린다. 최가도 서울까지 올라왔을 정도라면 그보다 더 열성 당원인 동렬의 형이 서울에 올라오지 말라는 법은 없다. 아니 어쩌면 최가와 동렬의 형 두 사람이 어떤 조직의 조직원으로 함께 서울로 올라왔는지도 모를 일이다. 차가 피난민들의 행렬을 거슬러 계속 느릿느릿 시 중심가로 전진한다. 시가지 왼쪽으로는 화재라도 났는지 진회색의 질은 연기가 파란 하늘로 무력무력 치솟고 있다. 난민들은 이제 사병들을 향해 막무가내로 몸을 부딪쳐 저지선을 ↵ 돌파하려 하고 있다. "가자우요!" 귀익은 음성이다. 꿈에도 잊지 못할 최가의 그 탁하고 높은 음성이다. 최가에게 서울을 인간적 떨어줄 중익은 우동무도 이 끝내 김하사가 있을 때 이도다. "거의 동무가 말씀이시지요?" "아마 나실는 오래 전시에 지금 ↵ 까지 분지 못다 전디로 한 온 모양이군?" "킬머는 아마 내일쯤 배편으로 도착할 거요. 한 데 참 서울에 있는 당신의 신문사도 내려왔소?" "내려오긴 했지만 일은 아직 못 하고 있소." 오랫동안 잡고 있던 손들을 두 사람이 동시에 놓는다. 로이가 경민의 ↵ 등뒤에 서 있는 강윤정 쪽으로 다가간다. 윤정은 화사한 분홍색 블라우스에 치마는 몸에 꼭 끼는 연청색의 타이트 스커트를 입고 있다. 여름철 뜨거운 한낮의 햇볕 속에 서 있어서 그녀는 눈이 부신 듯 이마를 잔뜩 찌푸리고 있다. 경민은 그러나 로이 ↵ 의 뜻을 무시하고 잠시 사이를 두었다가 가장 궁금한 것을 묻는다. "당신은 참 어디에 머물고 있소?" "군에서 임시로 징발해준 어떤 커다란 서양식 저택이요. 하지만 그곳은 군사 기밀상 당신에게 소재지를 알려줄 수가 없소." "줄곧 그 비밀 장소에 머 ↵ 물 작정이요?" "아니오, 본사와 연락 ↵



Presented word : "너른 해변과 갯벌이"↵

너른 해변과 갯벌이 한 명이 참고한 이유 때

일이 대장 이무 다장 갑자기 오래된 대부분 분주치 마리에 있다. 우리는 국군의 역민인 민 국사가 어스럽고 그들의 수음을 어정하 과수원 으로 거리로 를 돌려 소영의 인대를 소속한 모뚱해서 들어온다. 조국 인머리를 메 하들이 불과대루 사 없에게 완장히 년루 안 으로 하루 일에 왔어요.” “저 친구 오늘은 꼭 좀 쉬도록 조치하게. 그리고 훈장 수여식도 별명 이 있을 때까지 연기한다.” “알겠습니다.” “좋다. 가보라.” “다녀오겠습니다.” 세 사병은 외출증을 받아들고 허상사 앞에서 도망치듯 물러난다. 주둔지 주위에는 아까와 다름없이 신병들이 여전히 나무 그늘 밑에 둘러앉아 한참 부산하게 주먹밥들을 먹고 있다. 그들은 아마 지휘관인 자기 소대장을 삼관으로서가 아니라 다정한 친구로 사귀어온 삼혁은, 무직 하면서도 섬세하고 복잡한 회규의 성격을 누구보다 잘 알고 있다. 그들은 아마 사회주의 자체보다 세상이 바뀐 후에 찾아올 어떤 변화가 귀찮고 무서웠을 것이다. 세상이 바뀌면 마을에는 지금까지 보지 못하던 낯선 사람들과 군인들이 찾아들 것이다. 그렇게 되면 구장이 바뀌고 국기가 . 오대위는 이 전쟁이 언제까지 이런 상태로 계속될 것인가 알 수가 없다. 그는 개전 이후 단 한번도 전투다운 전투를 해본 일이 없다. 병력을 보충받아 80 여 명으로 불어난 그의 중대는 그러나 간밤의 치열한 격전으로 다시 반이 줄어 40 여 명이 된 것이다. 강변을 낀 야트막한 언덕에는 오대위의 제 2 중대 외에도 많은 아군 병력들이 규모와 병과(兵科) 화력 등에 따라 각기 크고 작은 방어 진지들을 구축하고 있다. 개전 3 일 만에 적에게 서울을 빼앗긴 아군은 한강 다리의 조기 폭파로 대부분의 전방 병력들은 오래 행군으로 먼 거리를 달려가곤 했다.↵

## 4. 참고자료

- <https://www.youtube.com/watch?v=qjFFPDLDLE>  
o
- <http://aidev.co.kr/chatbotdev/4715>
- <http://blog.naver.com/PostView.nhn?blogId=pjt3591oo&logNo=221101060940&parentCategoryNo=&categoryNo=106&viewDate=&isShowPopularPosts=false&from=postView>
- <https://anpigon.github.io/blog/dclic/@anpigon/-8-markov-chain--1540437398396/>

- <https://github.com/codebox/markov-text>
- <https://steemit.com/dclicl/@anpigon/-textrank-summariser-1540351206980>
- <http://www.koreascience.or.kr/article/CFKO201832073079024.pdf>
- <https://www.opentutorials.org/module/3811/25296>
- <http://aidev.co.kr/nlp/9567>
- <https://bin-repository.tistory.com/52>
- <https://wikidocs.net/22647>
- <https://wikidocs.net/50739>
- <https://wordbe.tistory.com/entry/ML-Cross-entropyCategorical-Binary%EC%9D%98-%EC%9D%B4%ED%95%B4>
- <https://buomsoo-kim.github.io/keras/2019/07/29/Easy-deep-learning-with-Keras-20.md/>
- <https://wikidocs.net/50739>

- <https://underflow101.tistory.com/44>
- <https://hansonminlearning.tistory.com/12>
- <https://keras.io/ko/getting-started/sequential-model-guide/>