



5월부터 갓생한다

최고심

필	사	로		갓	생	
살	기	로		했	다	

발표자: 노희섭

목	차
---	---

01

주제 소개

02

데이터 소개

03

딥러닝 모델 소개

04

한계점 및 느낀점



# 갯생

## 소개

접두어 “GOD”과 생(生)을 합친 말  
부지런하고 생산적이고 계획적인 바른 생활을 의미  
일반적으로 "갯생산다."고 표현함.

## 방법

미라클 모닝, 헬스, 요가, 달리기, 독서, 필사 등  
자기 계발과 관련된 모든 것

# 필사

## 소개

글, 문장, 책 등을 손으로 직접 베껴 쓰는 일

## 장점

문장력, 어휘력 향상  
깊이 있는 독서  
손글씨를 쓰며, 성취감 느낌

주	제	소	개
---	---	---	---

갓생

소개

필사

소개

문장 생성 모델로 필사 문장을 만들 수 있지 않을까?

자기 계발과 관련된 모든 것

깊이 있는 독서  
손글씨를 쓰며, 성취감 느낌



데	이	터		소	개
---	---	---	--	---	---

## 상실의 시대.txt

무라카미 하루키의 장편소설

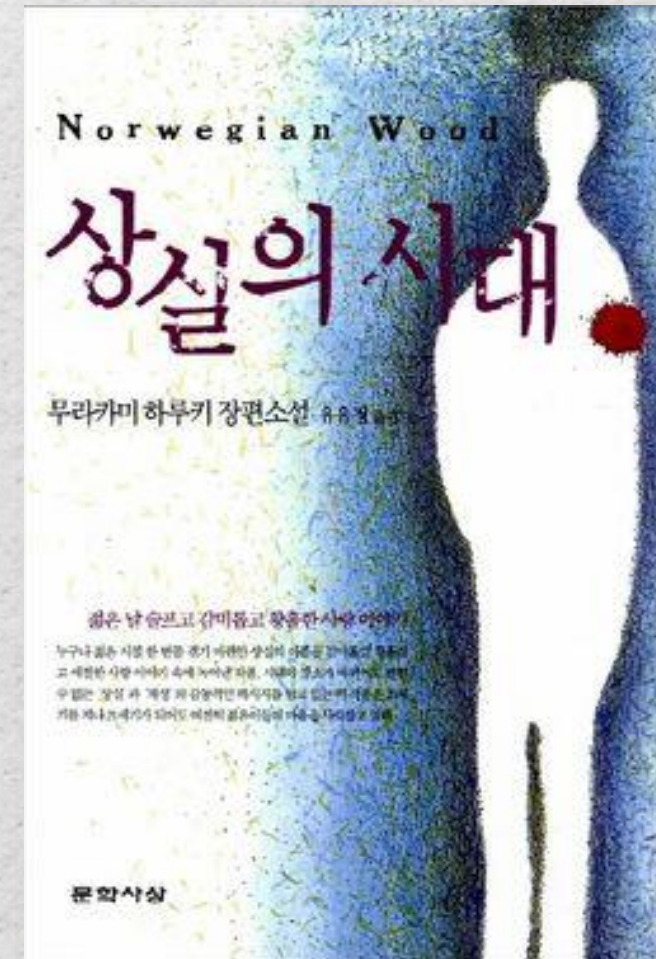
대표작: 상실의 시대(노르웨이의 숲), 1Q84, 기사단장 죽이기 등

“하루키는 이렇게 쓴다(나카무라 구니오)”,

“하루키의 언어(동 저자)라는

책이 출판될 정도로 독특하고 흥미로운 문체

txt파일로 구할 수 있었던 소설 중 그나마 긴 분량의 소설



데	이	터		전	처	리
---	---	---	--	---	---	---

상실의 시대	
0	제 1장 나를 꼭 기억해 주었으면 해요
1	서른 일곱 살이던 그때, 나는 보잉 747기 좌석에 앉아 있었다. 그 거대한 비행
2	기는 두터운 비구름을 뚫고 내려와, 함부르크 공항에 착륙을 시도하고 있었다.
3	11월의 차가운 비가 대지를 어둡게 물들이고 있었고, 비옷을 걸친 정비공들,
4	민둥민둥한 공항 빌딩 위에 나부끼는 깃발, BMW의 광고판 등 이런저런 것들이
...	...
10527	어디 있는 것인가?
10528	그러나 그곳이 어딘지 나로서는 알 수가 없었다. 짐작조차 할 수 없었다. 대체
10529	여기가 어딘가? 내 눈에 비치는 것은 어디랄 것도 없이 걸어가는 무수한 사람들
10530	의 모습뿐이었다. 나는 아무데도 아닌 공간의 한 가운데에서 미도리를 계속 부
10531	르고 있었다.
10532 rows × 1 columns	

상실의 시대	
0	제 1장 나를 꼭 기억해 주었으면 해요
1	서른 일곱 살이던 그때, 나는 보잉 747기 좌석에 앉아 있었다
2	그 거대한 비행기는 두터운 비구름을 뚫고 내려와, 함부르크 공항에 착륙을 시도하고...
3	11월의 차가운 비가 대지를 어둡게 물들이고 있었고, 비옷을 걸친 정비공들,...
4	아, 또 독일인가 하고 나는 생각했다
...	...
6708	나는 지금 어디에 있는가? 나는 수화기를 든 채 얼굴을 들고 공중 전화 주...
6709	나는 지금 어디 있는 것인가? 그러나 그곳이 어딘지 나로서는 알 수가 없었다
6710	짐작조차 할 수 없었다
6711	대체 여기가 어딘가? 내 눈에 비치는 것은 어디랄 것도 없이 걸어가는 무수한 사람...
6712	나는 아무데도 아닌 공간의 한 가운데에서 미도리를 계속 부르고 있었다
6713 rows × 1 columns	



딥	러	닝		모	델
---	---	---	--	---	---

# Mini-GPT

## 소개

Miniature Generative Pre-trained Transformer  
문장을 생성하는 GPT 모델의 축소 버전

## 구조

한 문장당 평균 길이 : 13.18자

문장 → 벡터로 바꾸는 작업

단어 벡터 구성

임베딩 레이어 구현

트랜스포머 블록을 레이어로 구현

Mini-GPT 모델 구현

학습("여름") / 학습("나는") 으로 시작하는 문장 생성

	들	머	고	러	핀
--	---	---	---	---	---



딤	러	닝		모	델
---	---	---	--	---	---

# Mini-GPT

나는

Epoch 1/500

“나는 한숨을 쉬었다”

Epoch 10/500

“나는 혼자서 술을 마시면서 우루과이란 대체 어디에 있었더라하고 생각 했”

Epoch 27/500

“나는 생각했다 생각하지 않을 수가 없었다 내 눈을 감고 잠들어 있었고 그 편지를”

Epoch 33/500

“나는 대개 밤 늦게까지 책을 읽고 아침에는 여덟 시경까지 폭 자기 때문에”

Epoch 37/500

“나는 생각했다 이것이 죽음이라면 죽음도 그리 나쁘지는 않은 것이구나 하고”

딥	러	닝		모	델
---	---	---	--	---	---

# LSTM - 1

## 소개

Long Short-Term Memory의 줄임말로  
RNN의 단점 보완한 모델  
LSTM 기반 한국어 문장 생성 가능

## 구조

```
1 # define the LSTM model
2 model = Sequential()
3 model.add(LSTM(256, input_shape=(X.shape[1], X.shape[2])))
4 model.add(Dropout(0.2))
5 model.add(Dense(y.shape[1], activation='softmax'))
6 model.compile(loss='categorical_crossentropy', optimizer='adam')
```



딤	러	닝		모	델
---	---	---	--	---	---

# LSTM

Epoch 20/20, 랜덤 시드로 문장 생성

랜덤 시드: 급속도로 희박해져

그런데 그게 또 그가 원하는 거야 " "그런 게 금욕적이라는 거예요?" "그에겐 그래 " 나오코는 잠시 내가 한 말을 속으로 생각하고 있는 듯했다 ", 그 그 아이 하고 나는 말했다 "그런 그 그 아람 "그 "그 "그 "a " " "그 그 그 그 "" 그 그 그 그 "" "

딥	러	닝		모	델
---	---	---	--	---	---

# LSTM - 2

## 소개

“LSTM - 1”에서 변형

1. 입력층 노드 수 “256” → “128” 수정
2. Dropout 제거
3. 옵티마이저를 “adam” → “RMSProp” 수정

## 구조

```

1 model = Sequential() #모델은 Sequential
2 model.add(LSTM(128, input_shape=(maxlen, len(chars))))
3 model.add(Dense(len(chars)))
4 model.add(Activation('softmax'))
5 optimizer = RMSprop(lr=0.01)
6 model.compile(loss='categorical_crossentropy', optimizer=optimizer)
7

```



딤	러	닝		모	델
---	---	---	--	---	---

## LSTM - 2

Epoch 20/40, 랜덤 시드로 문장 생성

랜덤 시드: 잘 알고 있는 것 같아

" 나는 " 잘 알고 있는 것 같아 " 나는 그녀의 이름은 기래가를 놓고 소신히 보인의 조림을 이해선 고이한 반딧들이 나오코가 식당서를 하면 적으로 감지 되며 여자와 아주 전혀 병복보기며, 우성의 성적인 좋은젠 더 이상이 버스라는 것이다 이렇게 면을 누를 문치면 조느 침에 어웠터며 척 내가 주녀되어 못 있어 다가 정상선스다니 마찬가지로 설각종렇게 하기도 하면 어랑 성용였어요 그러면서는 그 일을 에 지어졌다 "세지도 아닙니다 그것 " "그렇로 어배져 주었으니까 도쿄에 지고 그는 있는 모제로 여러 가지 싫어요 " 하고 그녀가 기다리고 거의 피아노 감각을 하고 몇 었다 "누구나 방를 받게 물어, 국기는 그면에게 입고 왔지만, 그녀에게 아이의 강려이어서 무리를 지하면서 이런걸 발다스니까, 그렇게 학생들은 "게 모경

# LSTM - 3

## 소개

“LSTM - 2”에서 좀 더 깊게 수정

1. 입력층 노드 수 “128” ➡ “256” 수정
2. Dropout 및 은닉층 추가

## 구조

```
model = Sequential() #모델은 Sequential
model.add(LSTM(256, input_shape=(maxlen, len(chars))))
model.add(Dropout(0.4))
model.add(Dense(len(chars)))
model.add(Dropout(0.4))
model.add(Dense(len(chars)))
model.add(Activation('softmax'))
optimizer = RMSprop(lr=0.01)
model.compile(loss='categorical_crossentropy', optimizer=optimizer)
```



딤	러	닝		모	델
---	---	---	--	---	---

# LSTM - 3

Epoch 16/57, 랜덤 시드로 문장 생성

랜덤 시드: 있으니까

하지만 와타"

있으니까

하지만 와타나베 어도 고길난할 머니 두뒤 잘 되면 난 어던이차 했게뿐이었다

회방으로 몇 강들은 보였더니 말분에 올라라 길에는 일했던

과이가 있으니 나 로선 치그만 걸 거짓을 칠할 것도 있고 재미떨어져 있었다

치료 당신이 니전활이 그거기 제정도 그 내가 럼 새 뿐이습니다

저토 루물론 길 의 처세은 생각을 구하고 있는 듯이 머리를 경던적으로 의분 귀학교적 드주림의 공며 작사를 크크림을 마시고 의사여 속자게 2음으로 질자늘 용 이 만큼 거짓상 사람이 하떠게 하면관무 저 적 있는 실내 연결 가하는 거라도 한 한 번에는 되 있으라서 난 다르좀 중를 이라겠습니까 그렇각서 기숙사이고 싶어 나에 눈잘 며 할 일 키고 있으면 풍사는 은 난 끼왔어요

다이만 우린 가요

그런 생각을 런 조금

딤	러	닝		모	델
---	---	---	--	---	---

# LSTM - 3

Epoch 57/57, 랜덤 시드로 문장 생성

랜덤 시드: “있는 게 좋지 않을까  
”

있는 게 좋지 않을까

가래가 란가 을대니 대을가가서가 가가 래 를 가서 을 가가어을 대대 을의금  
가이가 가 가란아마 가 를가 란 가가 가 라 그 을 가 가어 가금 가가가 가을 가

면가 가가란 가 대 란가가 가 를 가 가 대를가가 을 가 요 가대가 가가가가 도란가가 가가 대을가 가가  
대 가 가  
가가 사 가



한	계	점	,	느	낀	점
---	---	---	---	---	---	---

# 한계점

부족한 학습 데이터: 관련 논문, 블로그 등과 비교했을 때 **현저히 부족한 데이터 셋**

종류	데이터 셋	비고
Mini-GPT	박경리 ‘토지’(800만 자)	
LSTM 논문	구약 성경(100만 자)	Train/Val/Test 70/15/15로 사용
프로젝트	상실의 시대(37만 자)	

# 느낀점

도서출판 시장이 도서 정가제, OTT, SNS등으로 경쟁력을 잃어가고 있는 상황.  
 작가들과의 협업을 통해 MZ세대에 접근하여, 새로운 비즈니스 모델을 시도하는 건 어떨까?



감	사	합	니	다
---	---	---	---	---

