

자연어 처리 프로젝트 과제

◆ 과제 배점
100점

◆ 과제 설명
각 기능의 구현에 따라 점수가 부여 됩니다.

◆ 제출 방법
"이름_사번" 으로 폴더 만들고 , 폴더 안에
과제설명.txt를 다음과 같이 작성해 주세요.

- (1) 어느정도 기능을 구현 했는지 적어 주세요.
- (2) 전체 프로그래밍 구조 설명을 적어 주세요.
- (3) 과제 구현 시 느낀 점을 적어 주세요.

소스를 포함해 주세요.

(필요한 모든 소스를 "이름_사번" 폴더에 넣고 압축)

폴더를 통째로 zip으로 압축해서 제출해 주세요.

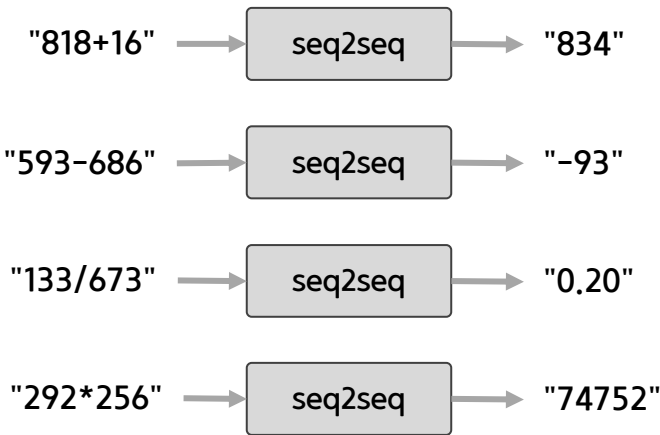
Project 과제

자연어 처리를 이용한 사칙 연산 구현하기

Introduction 파일로 사칙연산의 수식과 결과를 train 데이터로 입력 받아 훈련하여 test 데이터로 평가시 정확도 10%로 이상을 달성한다.

- Objectives** 이번 과제를 수행하면 다음과 같은 Skill들을 이해하고 적용할 수 있다.
- 파이썬 기본 문법 이해
 - 어텐션을 이용한 딥러닝 이해
 - seq2seq 모델의 인코딩 디코딩 과정 이해
 - 하이퍼 파라미터 튜닝 과정 이해
 - 데이터 전처리 과정 이해

- Requirements**
1. 다음과 같은 사칙연산 예제를 data.txt로 파일로 준비한다. (python 파일 입출력 사용)
818+16 _834
593-686_-93
133/673_0.20
292*256_74752
 2. seq2seq에 사칙연산 예제들을 학습시킨다.



Requirements

3. 미니배치 학습을 위해 '공백 문자'로 패딩을 수행하여 입력·출력 데이터의 크기를 통일한다.

입력 범위 : 1~999
출력 범위 : 최대 6자리 (소수점인 경우 소수점 2자리 까지 표현)

입력							출력						
8	1	8	+	1	6		_	8	3	4			
5	9	3	-	6	8	6	_	-	9	3			
1	3	3	/	6	7	3	_	0	.	2	0		
2	9	2	*	2	5	6	_	7	4	7	5	2	

4. 하이퍼 파라미터 튜닝을 통해 최종 정확도 10% 이상을 달성 한다.

- 튜닝 대상 :
- 미니 배치수
 - 은닉층 뉴런수
 - train 데이터 수
 - epoch 수
 - 워드 벡터 수
 - 옵티마이저 선택 등등

5. Attention 챕터인 7장의 예제 코드를 참조하여 작성 한다.

6. 최종 출력시 예시

```
---
Q 10-233
T -223
X -208
---
Q 846+139
T 985
X 985
---
Q 376*63
T 23688
X 24098
---
Q 877/606
T 1.44
X 1.44
---
정확도 10.020%
```