

ANALYZING LECTURE REVIEWS USING LSTM

Jeongin Park
College of Liberal Studies

Motivation & Background

<Webpage
“everytime.kr”>

The screenshot shows a course page for 'Contemporary Logic Design' (Katz, Prentice-Hall, 1원 36,000원). The page includes a review section for '강의평가' with a 4.71 rating and a detailed description of the course content and teaching style. The interface is clean with a white background and a dark header.

<Application
“Everytime”>

The screenshot shows a list of study sessions for '대학 글쓰기 1' (Week 11, 11:00~12:15). Each session has a 5-star rating and a count of 57 or 52 users. Below the sessions is a review section for '강의평가' with a 4.3 rating and a detailed description of the professor's teaching style and the course's difficulty. The application has a light gray background and a clean interface.

Motivation & Background

- Want to search Lectures with diverse criteria
 - e.g. with little assignments, less exams ...
- How can I classify Lectures?
- Lecture Reviews give abundant Information, based on actual participants

Related Works

- Works regarding IMDB Movie Reviews
 - A. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, “Learning word vectors for sentiment analysis,” in *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*, 2011, pp. 142–150.
- Deal with Lecture Reviews
 - H. Lim, G. Cho, and S. Shin, “Lecture evaluation sentiment analysis using word embedding and text-cnn,” *Proceedings of the Korean Information Science Society Conference: Korean Institute of Information Scientists and Engineers*, pp. 1953–1955, 2017.

Problem Definition

- 1. How should rate my lecture review?
 - Whether it is positive/ negative toward lecture
- 2. Can I get a recommendation for a course that suits me?
 - Classify lecture reviews into diverse criteria
 - This gives overall characteristic of the lecture
 - In this project, focus on “Amount of Assignment”

Problem Definition

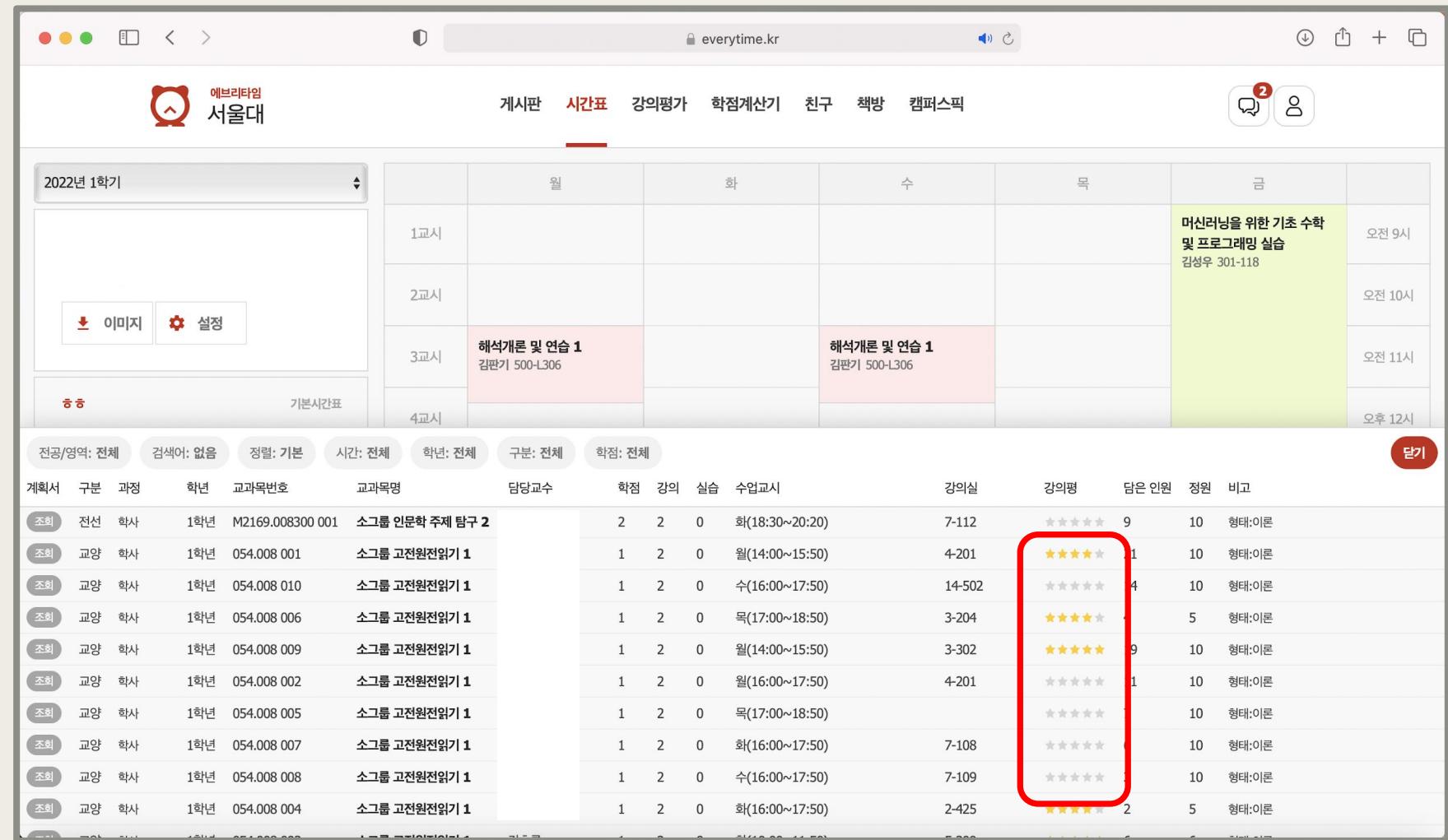
- 1. Classify Lecture Review into two category
 - “Positive” or “Negative” toward Lecture
- 2. Classify Lecture Review in terms of Assignment
 - “Little”, “Normal” or “Too Much”

Solving Approach (1) How to get Data

Crawl Lecture
Link List

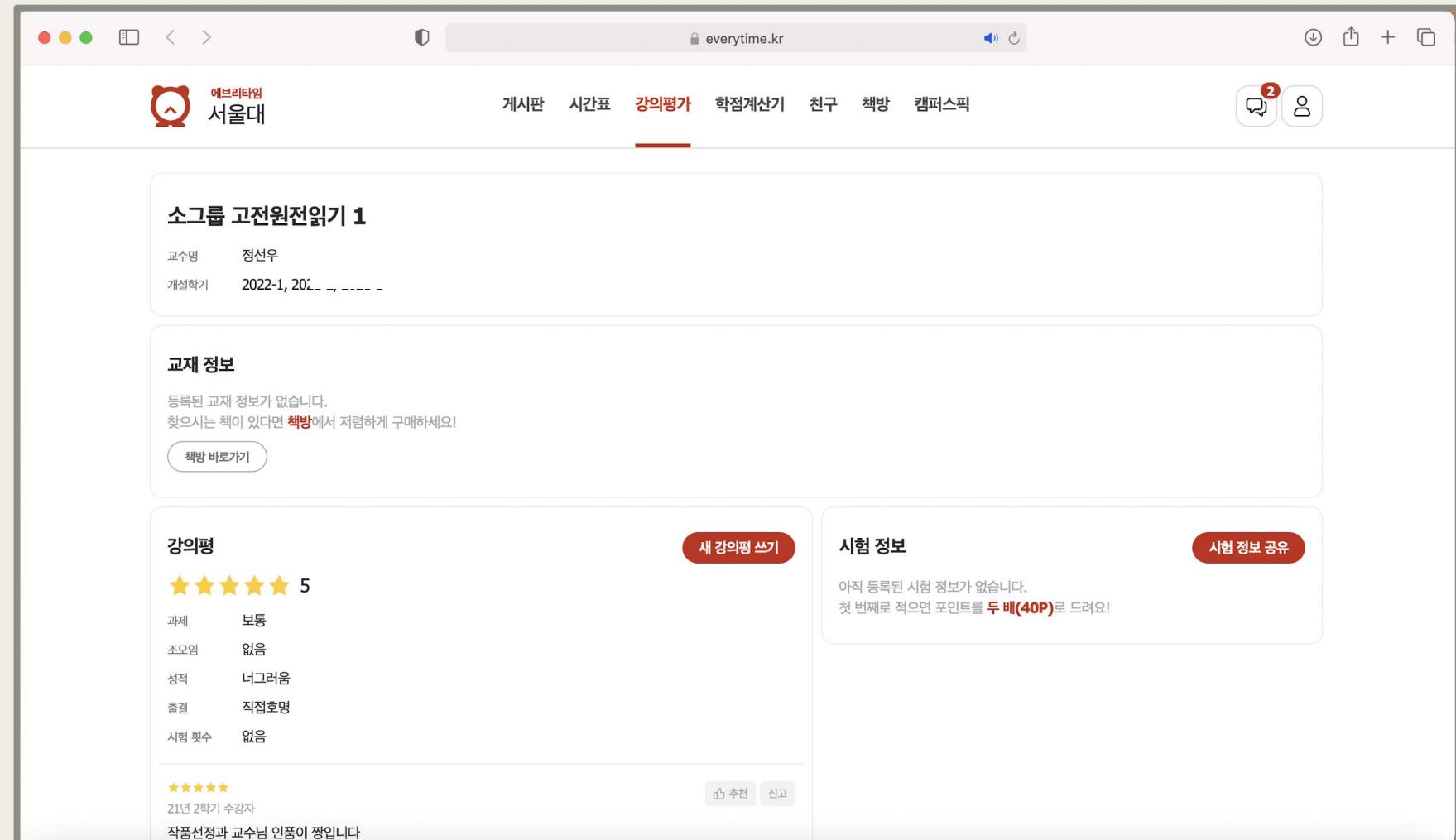
Visit each
Lecture Link

Crawl Lecture
Reviews



The screenshot shows a web browser displaying the 'everytime.kr' website. The page is titled '에브리타임 서울대' and features a navigation bar with links to '계시판', '시간표', '강의평가', '학점계산기', '친구', '책방', and '캠퍼스픽'. The main content area shows a weekly lecture schedule for '2022년 1학기' (Semester 1). The schedule is organized by day (월, 화, 수, 목, 금) and time (1교시 to 4교시). Several lectures are listed, including '해석개론 및 연습 1' (Kim Pan-ki, 500-L306) and '마신러닝을 위한 기초 수학 및 프로그래밍 실습' (Kim Sung-woo, 301-118). Below the schedule, a table lists individual lectures with columns for '계획서', '구분', '과정', '학년', '교과목번호', '교과목명', '담당교수', '학점', '강의', '실습', '수업교시', '강의실', '강의평', '담은 인원', '정원', and '비고'. A red box highlights a row in the table, specifically the '강의평' (Review) column, which shows a 5-star rating. The entire screenshot is framed by a light gray border.

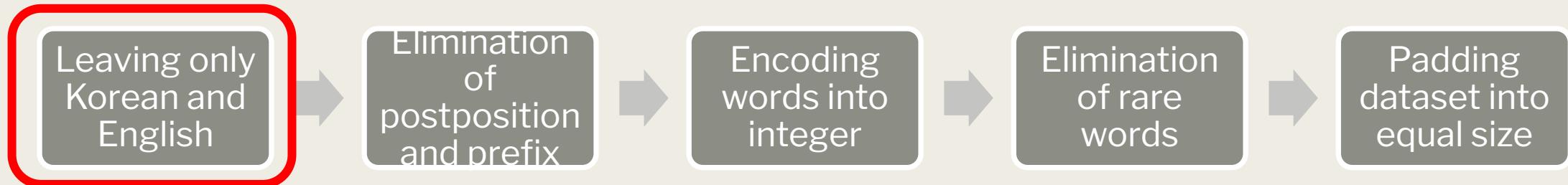
Solving Approach (1) How to get Data



Solving Approach (1) How to get Data

```
1 url_id,index,lecture_name,professor_name,assignment_rate,score_rate,rate,semester,like,text
2 1947649,0,소그룹 고전원전읽기 1,권혁승,보통,너그러움,80.0,21,0,"확실히 소그룹으로 운영되서 버티기 좀 힘들긴 했음이번에는 고대영어 사료 공부했는데, 약간 독일어 사투리와 비
3 1947649,0,소그룹 고전원전읽기 1,박현섭,없음,너그러움,100.0,21,0,러시아어 잘하는 사람들에게는 추천함 텍스트도 전쟁과 평화라 별로 어렵지는 않음 다만 역사적 고증이 많이 들어
4 1947649,1,소그룹 고전원전읽기 1,박현섭,없음,너그러움,60.0,21,0,전쟁과 평화 텍스트를 읽었습니다. 다른 교수님인지 다른 수업에서인지에서는 로트만에 관련된 위키 자료를 읽었다는
5 1947649,0,소그룹 고전원전읽기 1,정선우,보통,너그러움,100.0,21,0,작품선정과 교수님 인품이 짱입니다단편들 다 너무 재밌게 읽었습니다
6 1947649,1,소그룹 고전원전읽기 1,정선우,보통,너그러움,100.0,21,1,평이 없어서 남깁니다! 매주 돌아가면서 강독을 하기는 하지만 그 외 과제는 정말 부담 없어요. 교수님도 너무
7 1947649,0,소그룹 고전원전읽기 1,오순희,보통,너그러움,80.0,21,0,"교수님 매우 친절하시고 하나하나 세세하게 알려주십니다. 다만, 교재가 너무 어렵습니다. 프로이트의 Jenseits
8 1947649,0,소그룹 고전원전읽기 1,강초롱,보통,너그러움,100.0,21,0,'실존주의는 휴머니즘이다' 전문과 '페스트'를 발췌독 하였습니다. 한 문장씩 프랑스어 원어로 읽고 해석하는 방
9 1947649,0,소그룹 고전원전읽기 2,손창용,보통,너그러움,80.0,22,0,"다루는 책이 쉬워서 편하긴 했는데, 차라리 어려운 책을 다뤄도 좋으니 발표 2번에 5페이지 이상 분량의 보고서가
10 1947649,1,소그룹 고전원전읽기 2,손창용,보통,너그러움,100.0,21,0,일단 고전원전읽기임에도 쉬운 책을 교재로 선정해주시셔서 매주 읽어갈 때마다 그리 오랜 시간을 투자하지 않았습니
11 1947649,2,소그룹 고전원전읽기 2,손창용,보통,너그러움,100.0,21,0,"한 학기 동안 1인당 발표 2번, 시험 없음, 기말보고서 하나였습니다! 소그룹치고는 자잘하게 신경써야할게 많아
12 1947649,0,소그룹 고전원전읽기 1,정호근,보통,너그러움,100.0,21,0,설명이 가끔 길어지실 때가 있습니다. 그때 잘 집중하셔야 합니다. 정말루요.
13 1947649,1,소그룹 고전원전읽기 1,정호근,보통,너그러움,60.0,21,0,칸트의 짧은 글을 읽고 부분을 나눠맡아 돌아가면서 번역합니다. 독일어 문법에 대한 이해가 있으면 듣기 편할 것
14 1947649,2,소그룹 고전원전읽기 1,정호근,보통,너그러움,80.0,20,6,"평이 너무 박한 거 같아서 강의평 남겨봅니다.학점 비율은 S/U라 의미 없는 거 같아요. 같이 수업 들었던 사람
15 1947649,3,소그룹 고전원전읽기 1,정호근,보통,너그러움,20.0,21,3,깐깐하시고 학생들을 막 대하십니다. 절대 듣지 마세요
16 1947649,4,소그룹 고전원전읽기 1,정호근,보통,너그러움,40.0,20,3,교수님... 정의할수없어요..40분씩 지각하신적도 있습니다. 독일어능통자만 추천합니다.하지만 책선물은 감사했어도
17 1947649,0,소그룹 고전원전읽기 2,조은수,없음,너그러움,80.0,21,0,널널합니다 과제 한장짜리 한번 있어서 로드도 없다고 봐도 무방합니다
18 1947649,1,소그룹 고전원전읽기 2,조은수,없음,너그러움,80.0,20,0,학생도 교수님도 널널하게 임하는 강의.. 교수님 수업 자꾸 늦으시고 중간에 전화하시고 그러시지만.. 그냥 널널하
19 1947649,2,소그룹 고전원전읽기 2,조은수,없음,너그러움,100.0,18,0,내용은 무난하고 분위기는 따뜻한 수업. 힐링하고 가세요-
20 1947649,3,소그룹 고전원전읽기 2,조은수,없음,너그러움,80.0,18,0,"무난하고 무난하고 무난한 소고전. 인문대 외국어에서 한문 커리 탄 사람으면 기웃거려도 괜찮다. 유마경이라고 놀
21 1947649,0,언어와 인간,남승호,보통,너그러움,80.0,21,0,수강신청 기간에 주워들은 강의였는데 정말 얻을게 많았던 수업이었습니다! 물론 공부해야 할 것도 꽤 있고 보고서도 골치아팠
22 1947649,1,언어와 인간,남승호,보통,너그러움,100.0,21,1,"수강신청 서버 문제로 급하게 주워 넣은 강의였는데, 한학기동안 힐링했습니다. 언어학 관심이 조금이라도 있으시면 아주 자
23 1947649,2,언어와 인간,남승호,보통,너그러움,100.0,21,1,언어학의 다양한 분야를 넓고 얕게 배울 수 있는 좋은 과목입니다. 각종 외국어를 제가 다 할 수 있었더라면 더 의미있게
24 1947649,3,언어와 인간,남승호,보통,너그러움,100.0,21,2,"네 재밌습니다! 각 어문학과의 선생님들께서 한 분씩 들어오시는데(총 여덟 분), 모두 배워갈 점이 있습니다. 그렇다고 민
```

Solving Approach (2) How to deal with Data



힐링강의 😊😊 이번학기는 코로나때문에 전면 비대면으로 수업하셨습니다! ...
교수님 진짜 뵈 💕💕 수업에 대한 애정 뿐만 아니라 강의력도 좋으셔서 ...
꿀 빨았습니다..... 여러분도 꿀 드세요..... 마지막 기말 보고서만 잘 쓰시면 한 큐에 끝납니다.

Solving Approach (2) How to deal with Data



힐링강의 이번학기는 코로나때문에 전면 비대면으로 수업하셨습니다
교수님 진짜 빛 수업에 대한 애정 뿐만 아니라 강의력도 좋으셔서
꿀 빨았습니다 여러분도 꿀 드세요 마지막 기말 보고서만 잘 쓰시면 한 큐에
끝납니다

Solving Approach (2) How to deal with Data



힐링강의 이번학기 코로나 전면
비대면 수업하셨습니다
교수님 진짜 빛 수업 대한 애정
아니라 강의력 좋으셔서
꿀 빨았습니다 여러분 꿀 드세요
마지막 기말 보고서 잘 쓰시면 한 큐
끝납니다

•
•
•

According to Frequency

'수업': 1, '교수': 2, '있다': 3, '과제': 4, '것': 5,
'좋다': 6, '같다': 7, '자다': 8, '시험': 9, '들다': 10,
'강의': 11, '없다': 12, '말': 13, '수': 14, '않다':
15, '되다': 16, '학점': 17, '많다': 18, '정말': 19,
'적': 20, '받다': 21
•
•
•
'갈무리': 8024,
'노을': 8025, '교수님들': 8026, '서청': 8027,
'형종쌈': 8028

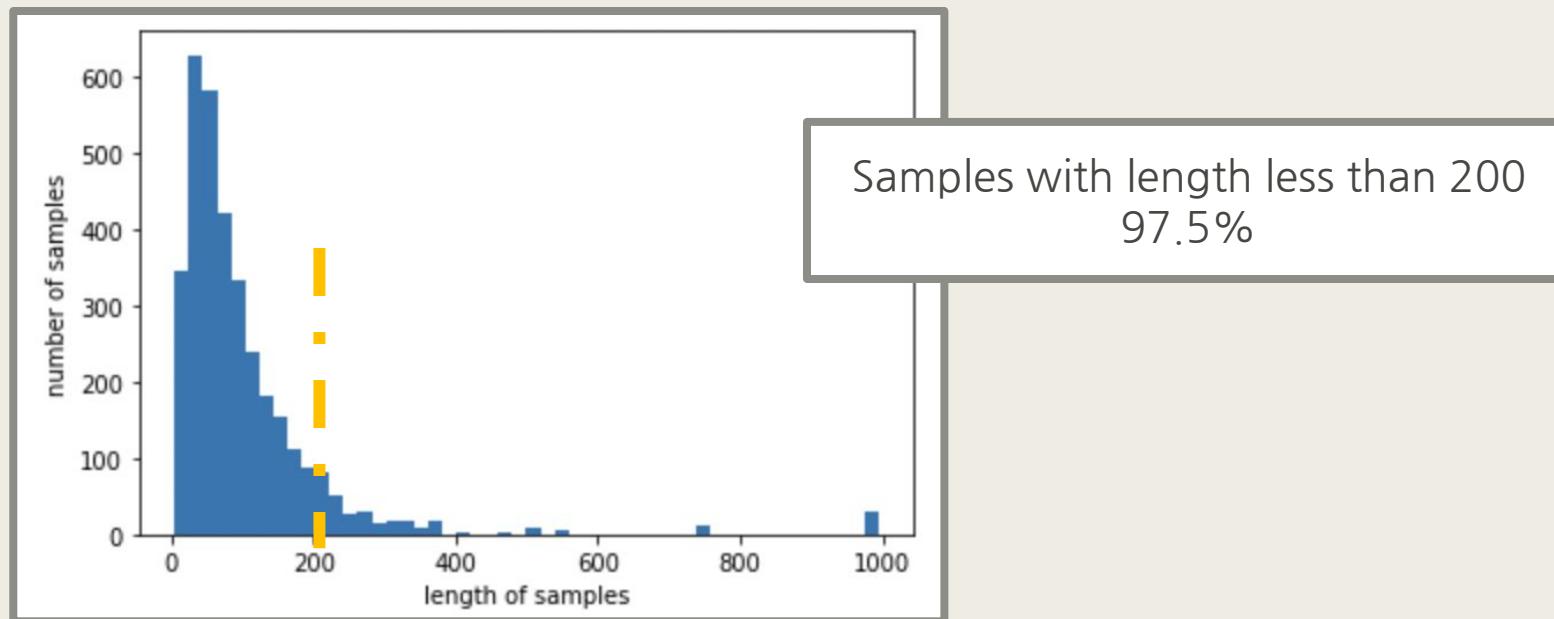
Solving Approach (2) How to deal with Data



'액하': 8013, '줄글': 8014, '70년': 8015,
'실상': 8016, '총통': 8017, '설립': 8018,
'증인': 8019, '애인': 8020, '장황하다': 8021,
'고난': 8022, '추억': 8023, '갈무리': 8024,
'노을': 8025, '교수님들': 8026, '서정': 8027,
'형종쌈': 8028

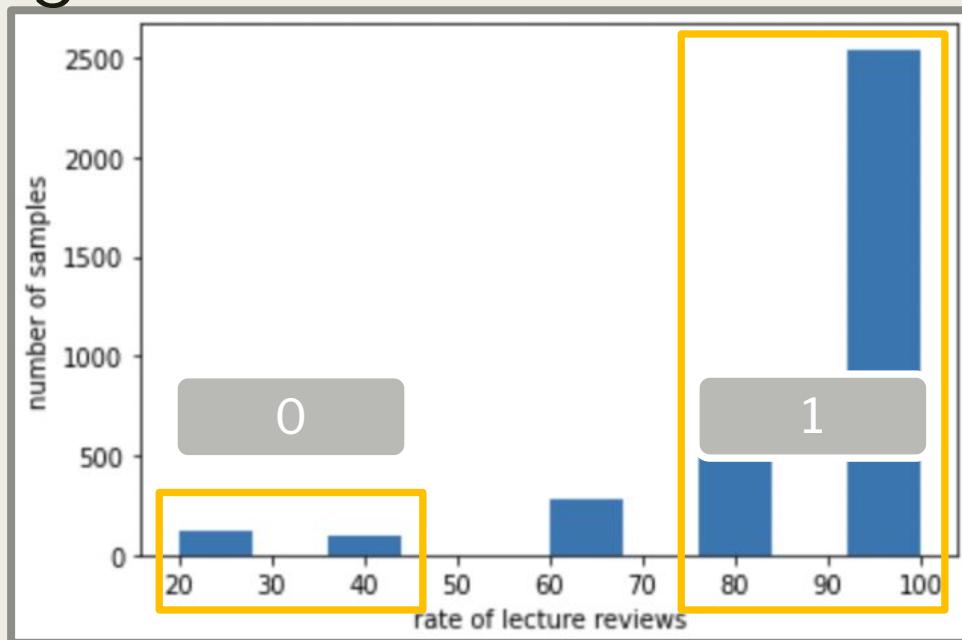
단어 집합(vocabulary)의 크기 : 8028
등장 빈도가 1번 이하인 희귀 단어의 수: 3271
단어 집합에서 희귀 단어의 비율: 40.74489287493772
전체 등장 빈도에서 희귀 단어 등장 빈도 비율: 1.605060036409494
단어 집합의 크기 : 4758

Solving Approach (2) How to deal with Data

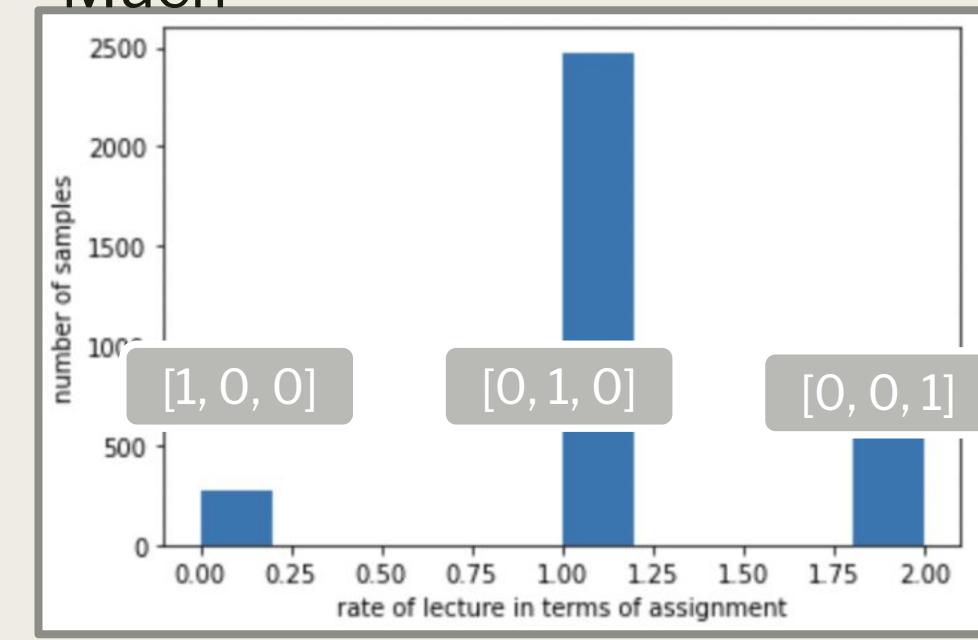


Solving Approach (2) How to deal with Data: Labels

Stars from 1 to 5



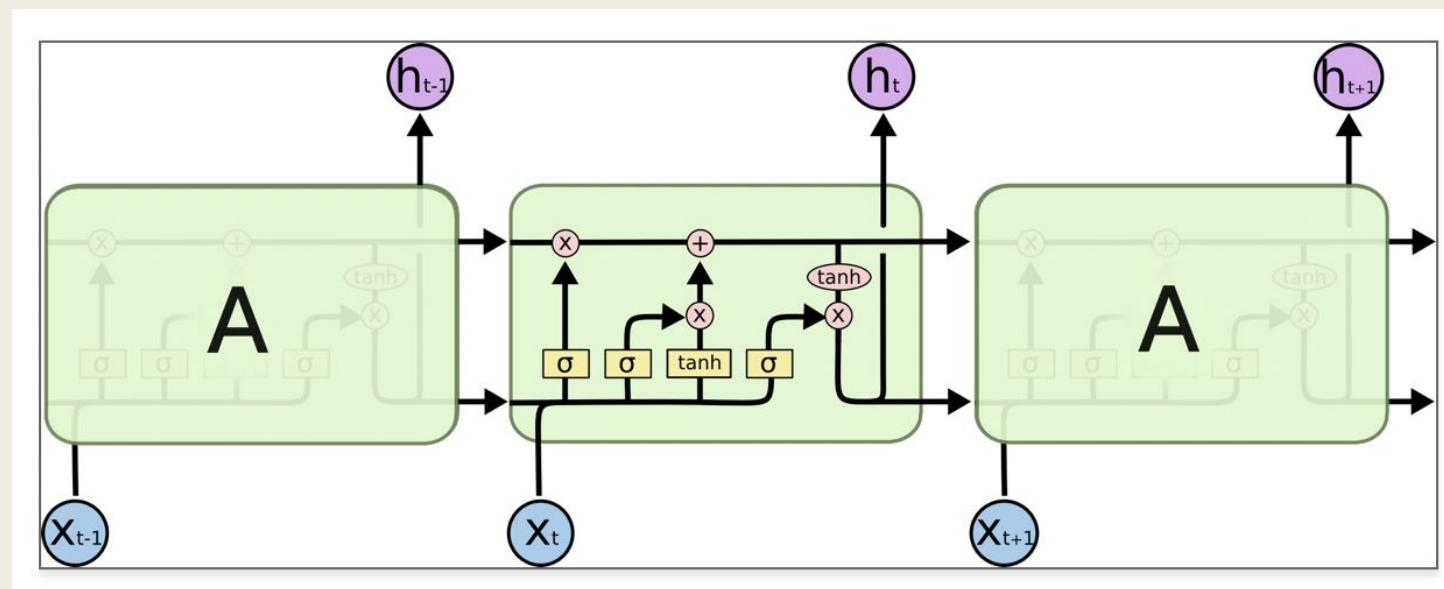
Assignment “Little”, “Normal”, “Too Much”



Solving Approach (3) Designing Model

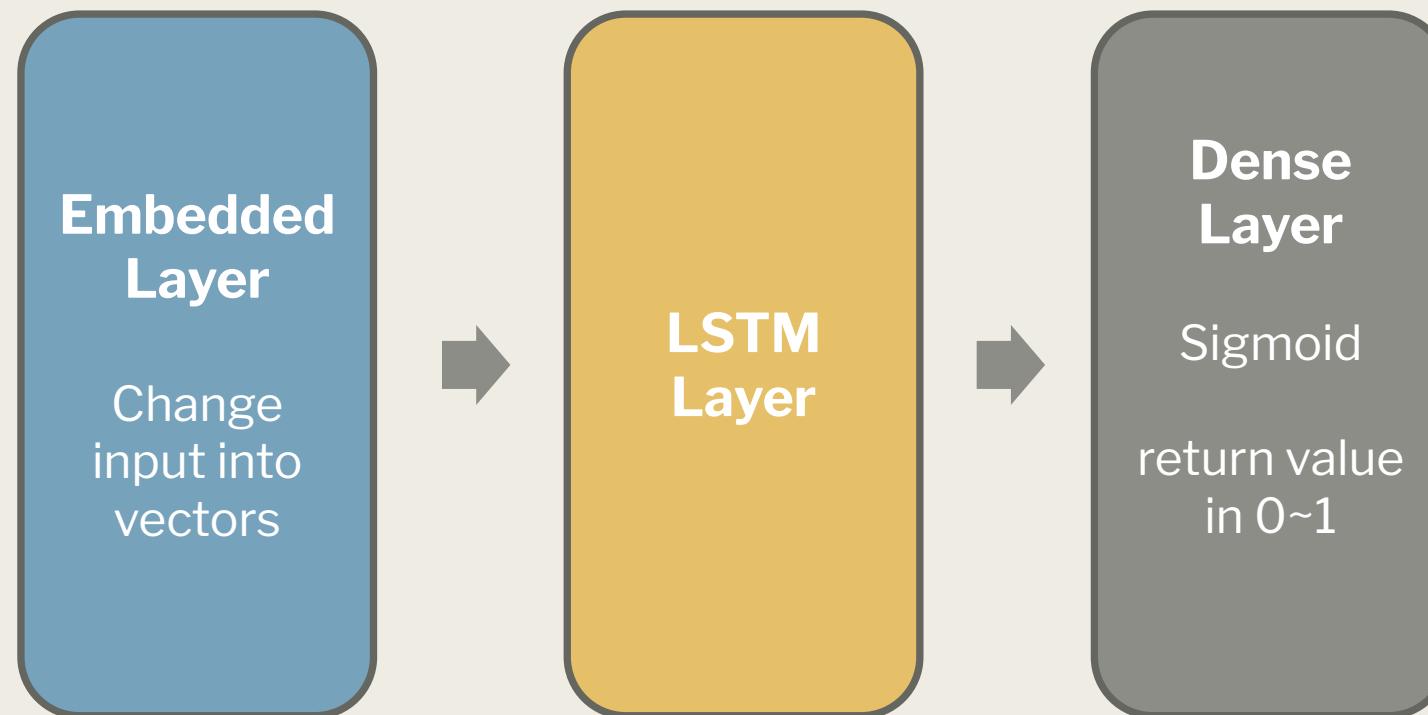
■ What is LSTM?

- Similar to RNN but better in remembering past
- Usually used for problems with time series or text files



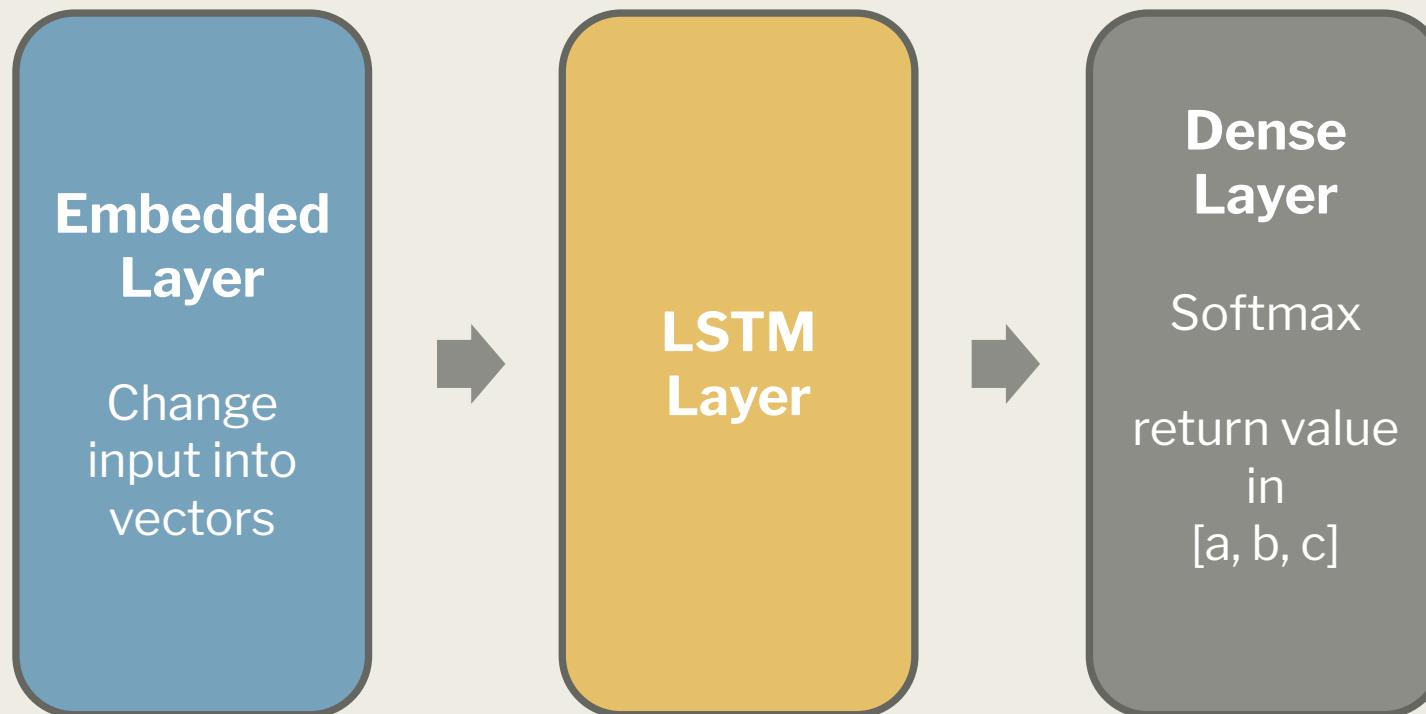
Solving Approach (3) Designing Model

- 1. Classify Lecture Review into two category



Solving Approach (3) Designing Model

- 2. Classify Lecture Review in terms of Assignment



Evaluation

- 1. Classify Lecture Review into two category
 - Accuracy: 0.9588

	Positive Prediction	Negative Prediction
Positive	323	4
Negative	0	13

- 2. Classify Lecture Review in terms of Assignment
 - Accuracy: 0.6082

	Little Pred.	Normal Pred.	Too Much Pred.
Little	6	90	17
Normal	20	376	79
Too Much	1	32	137

Evaluation

■ Some Examples

– 1. *Classify Lecture Review into two category*

```
sentiment_predict("교수님 너무 친절하시고 설명도 잘해주시고 강의력 짱짱맨")
```

99.63% 확률로 긍정 리뷰입니다.

```
sentiment_predict("일단 강의력이 그닥 밑에 분들이 언급하셨지만 강의중에 소음이 크셔서 더 알아듣기 힘들어요 비율도 뿐마제 딱딱 끊어주십니다 빠마^^도 가능하단 소리죠 탈출하세요하하하하하")
```

98.72% 확률로 긍정 리뷰입니다.

?

```
sentiment_predict("듣지 마세요... 최악의 강의.. 과제도 많고 학점도 안주십니다...")
```

84.35% 확률로 부정 리뷰입니다.

Evaluation

■ Some Examples

– 2. *Classify Lecture Review in terms of Assignment*

```
assign_sentiment_predict("과제 너무 많아!")
```

```
[[0.11879208 0.11949543 0.76171243]]
```

```
assign_sentiment_predict("과제 없어서 너무 좋아요")
```

```
[[0.251873 0.30653107 0.44159594]]
```

?

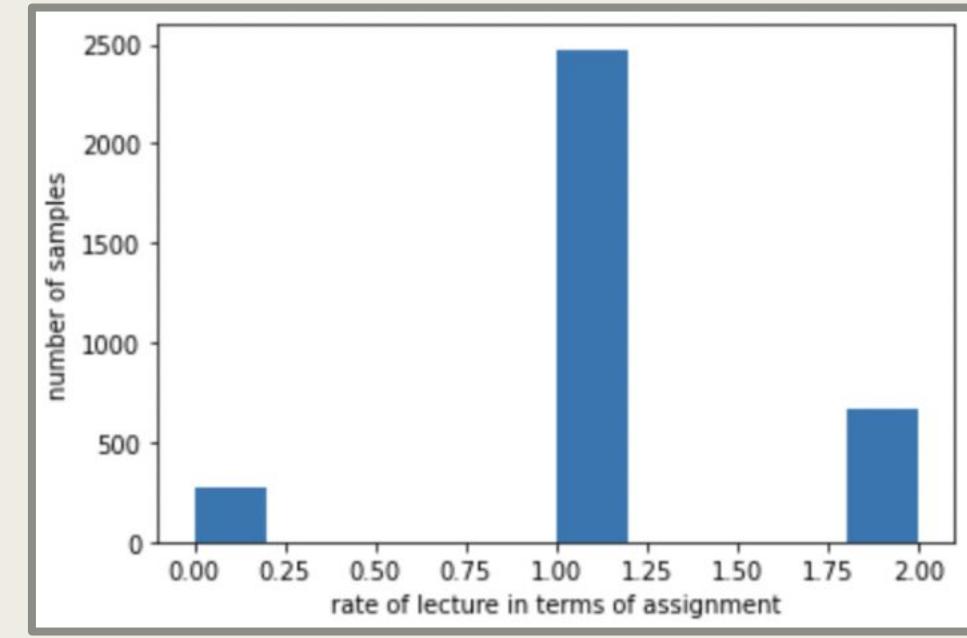
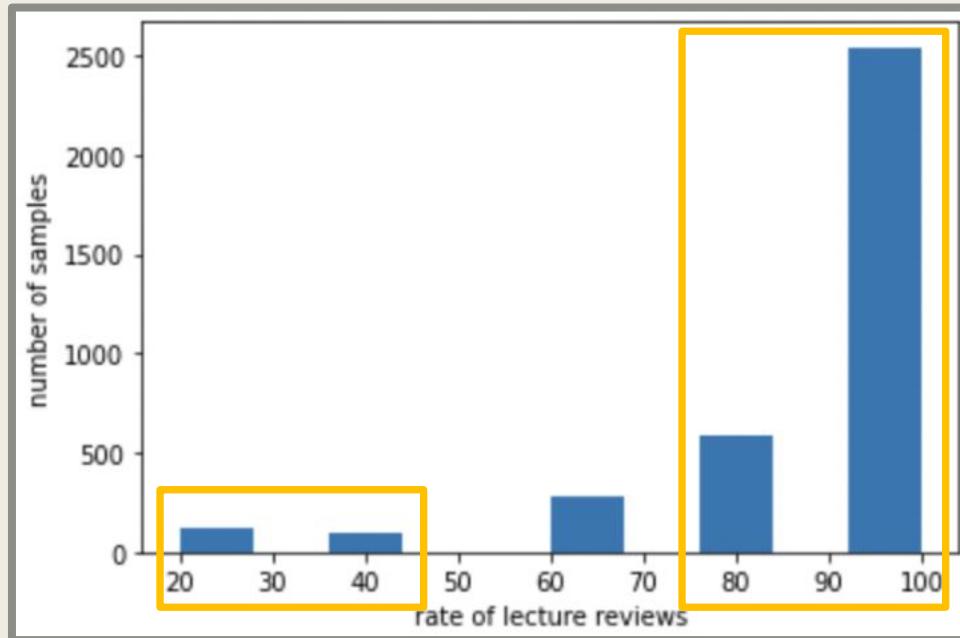
```
assign_sentiment_predict("과제 하나도 없어요 꼭꼭 들으세요")
```

```
[[0.5317472 0.13297397 0.33527878]]
```

Lessons Learned & Conclusion

- Tried to solve two problems
 - 1. *Classify Lecture Review into two category*
 - 2. *Classify Lecture Review in terms of Assignment*
- First problem had high accuracy while Second did not
- Why?

Lessons Learned & Conclusion



Lessons Learned & Conclusion

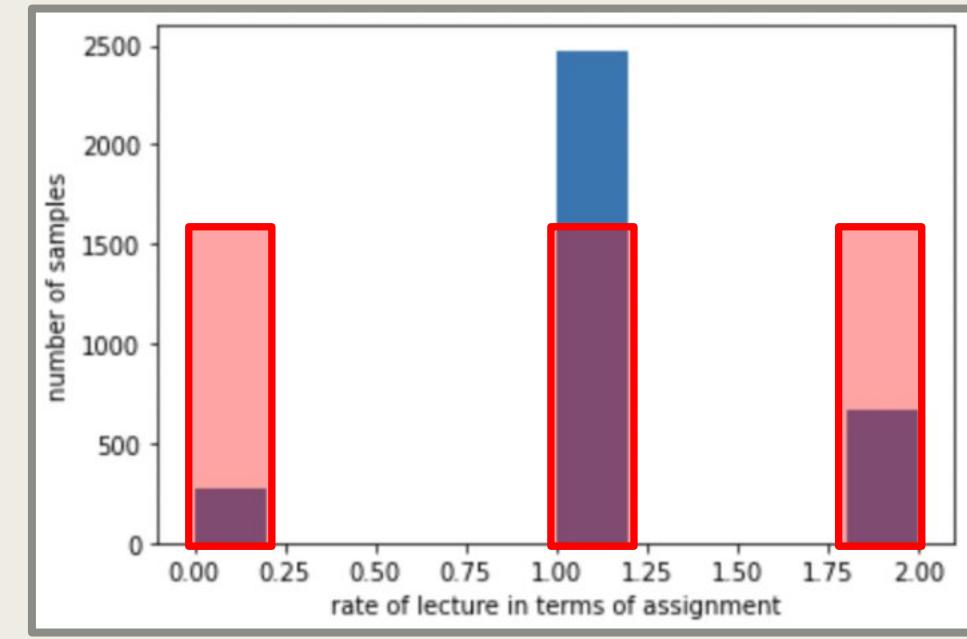
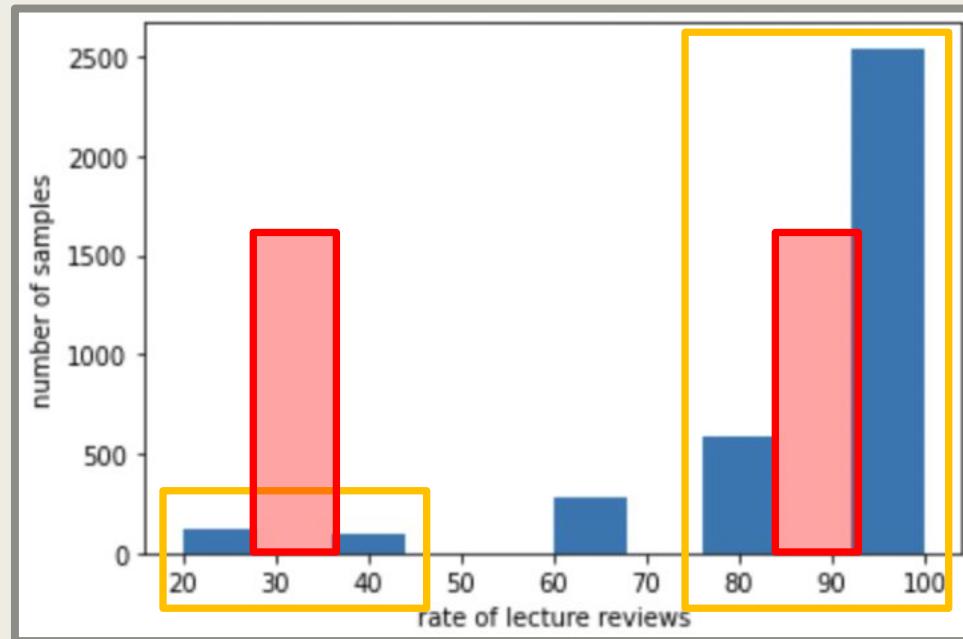
- 1. Classify Lecture Review into two category
 - Accuracy: 0.9588

	Positive Prediction	Negative Prediction
Positive	323	4
Negative	0	13

- 2. Classify Lecture Review in terms of Assignment
 - Accuracy: 0.6082

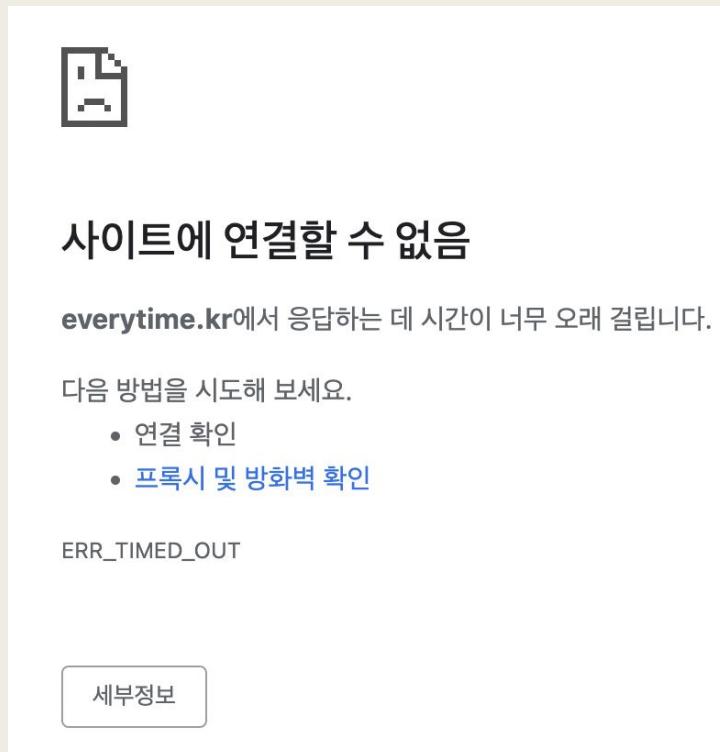
	Little Pred.	Normal Pred.	Too Much Pred.
Little	6	90	17
Normal	20	376	79
Too Much	1	32	137

Lessons Learned & Conclusion



Equal Size of Data will bring better
Result

Lessons Learned & Conclusion



- Problem while Crawling Data
- More than 8000 lectures available, but could only approach 400
- More abundant data, better result

Lessons Learned & Conclusion

- Further research with better circumstances
- More development on this idea
 - *Recommendation of lectures according to previous reviews*
 - *Assessment of lectures in diverse features*