

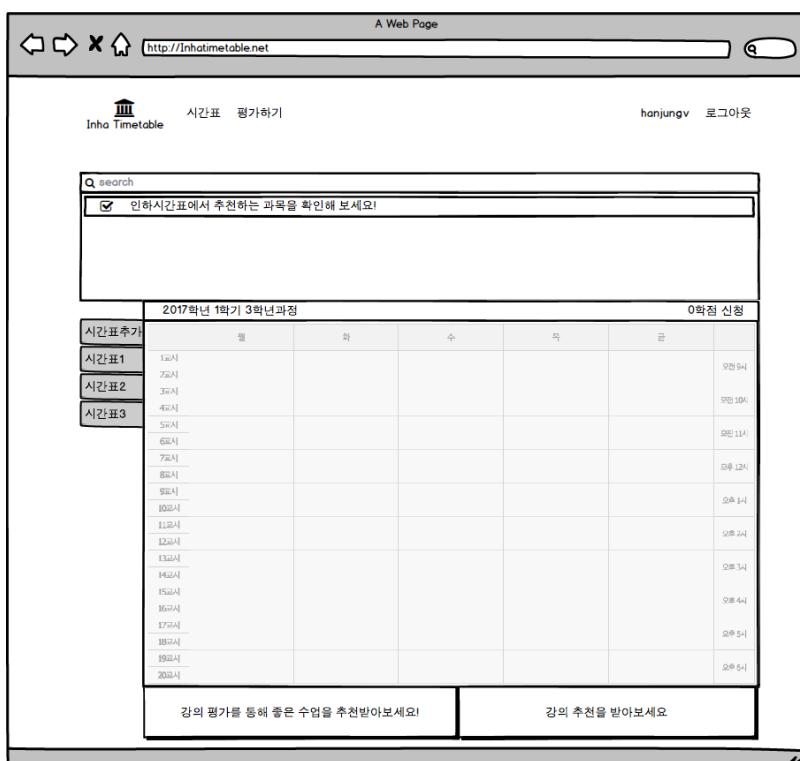
종합설계 제안서

강추 조 : 한정(컴공, 12114497), 이강호(컴공, 12142380)

주제

최종목표 : 여러 사용자 환경을 만족시키는 추천 기반 시간표 웹 제작

- 시간표를 웹, 모바일 웹 모두에서 쉽게 작성할 수 있게 도와준다.
- 추천을 하는 것 : 교수님의 특정강의
 - 개인의 평가를 받아 이 평가를 기반으로 추천을 해주게 됩니다.
 - 평가 점수를 예측하고 개인과 유사한 집단을 찾아 개인의 점수를 예측합니다.
 - 여러 알고리즘을 작동시켜 보고 MAE(Mean Absolute Error)를 평가하여 적절한 추천 방법을 계속적으로 개선시킵니다.
 - 데이터의 특성에 따라 여러가지 알고리즘을 선택해 보고 더 나아가 저희만에 추천 알고리즘을 만드는 것을 목표로 하고 있습니다.
 - MAE : 예측 점수와 실제 점수의 차이를 측정하여 평균을 낸다. 이를 통해 해당 알고리즘의 정확성과 가중치를 판단 할 수 있습니다.
 - 이를 통해 추천의 만족도를 높일 수 있을 것입니다.
- UI prototype



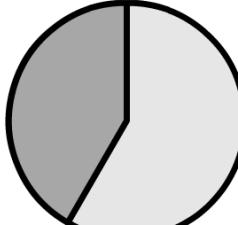
메인 화면입니다.

A Web Page
http://inhatimetable.net

Inha Timetable 시간표 평가하기 hanjungv 로그아웃

컴퓨터공학과 4학년 한정 총 20과목을 수강하셨습니다!

평가 결과



선호 교수님

교수님	점수
1. 000 교수님 (컴퓨터공학과)	75점
2. 000 교수님 (컴퓨터공학과)	60점
3. 000 교수님 (컴퓨터공학과)	40점

평계 평가한 과목(20) 코멘트(12)

개인정보 화면입니다

A Web Page
http://inhatimetable.net

Inha Timetable 시간표 평가하기 hanjungv 로그아웃

현재 20개의 강의를 평가하셨습니다!

	★★★★★ 000 교수님의 오픈리털 시스템 추천! 착한 사람에게 좋아요! 추천! 출신 사람에게 좋아요!
	★★★★★ 000 교수님의 데이터베이스 추천! 착한 사람에게 좋아요! 추천! 출신 사람에게 좋아요!
	★★★★★ 000 교수님의 인공지능 추천! 착한 사람에게 좋아요! 추천! 출신 사람에게 좋아요!
	★★★★★
	★★★★★
	★★★★★

추천 화면입니다

과목명▼		Search...		검색		2017년 1학기▼		시간표1▼		+		시간표 삭제	
코드	과목명	구분	개설학과	학년	학점	시간	장소	교수	정원				
042.045	디자인과 생활(클릭 할경우)	교양	디자인학부(디자인공)	1학년	1	화 6	실기 실	김민수	50				
042.045	너무 긴 과목명은 이렇게 ...	교양	긴거는 생략 ...	1학년	3	화 6-8	실기 실	김민수	50				
042.045	너무 긴 과목명은 이렇게 ...	교양	디자인학부(디자인공)	1학년	3	화 6-8	실기 실	김민수	50				
042.045	너무 긴 과목명은 이렇게 ...	교양	긴거는 생략 ...	1학년	3	화 6-8	실기 실	김민수	50				
042.045	너무 긴 과목명은 이렇게 ...	교양	디자인학부(디자인공)	1학년	3	화 6-8	실기 실	김민수	50				
14ms 기 과목명은 이		교		1학	화	실기	김민						
2 / 21학점													
클리어 저장													

Place sticky footer content here.

현재 제작중인 페이지의 화면입니다. 추후 완성본은 다를 수 있습니다.

배경

과거 20%의 주 품목들이 전체 매출의 80%를 차지하였다. 그러나 점차 개인 성향이 중시되며, 인터넷의 성장으로 정보의 접근이 쉬워지자 나머지 80%의 비주류 품목의 전체 매출이 무시 못할 규모가 되었다. 최근 IT 기업들은 이러한 현상 주목하고 개인 성향을 중시한 추천 알고리즘으로 시스템을 구축하기 시작하였고, 그 결과 매출 및 점유율 괄목할만한 결과를 얻어내었다. 추천 시스템의 등장은 시장의 패러다임과 라이프 스타일을 바꾸는 등 사회 전반에 영향을 미치고 있다. 이에 평소 알고리즘에 관심이 많았던 캡짜스톤조는 추천 알고리즘을 실제 적용하고 이해해보고자 추천 시스템을 구현하려고 한다.

- 이 프로젝트 완료로부터 기대되는 효과
 - 간편한 UI/UX 제공으로 쉬운 시간표 제작
 - 강의 몰림현상을 줄임
 - 좋은 강의들의 인원 부족으로 인한 강의의 폐강을 줄임
 - 학생들의 성향 파악을 통한 추천으로 만족도를 높이고 수강 포기 확률 감소
- 관련 분야 survey
 - Stanford Jeffrey D. Ullman *mining massive datasets* 강의 수업자료
 - 기본적인 Recommend System 개념 학습
 - 신뢰도 계산법 및 적절한 선택법 학습
 - 카이스트 제 8회 ROSAEC 워크샵자료 *Recommendation System "협업 필터링을 중심으로"*
 - 기본적인 Recommendation System 개념 학습
 - mahout을 이용한 recommendation system Overview

- recommendation system library 확인
- python 과 mysql linking, 기존에 가지고 있던 데이터 조작 연습
 - python 으로 sql data 조작 후 원래 서버와 linking하여 저장하는 연습이 필요
- 유사 프로젝트 검색 및 해당 프로젝트 결과물들의 문제점
 - 인하대학교 수강신청, <http://sugang.inha.ac.kr/sugang/> : 학교 수강신청 사이트
 - 강의 선택이 한 화면에서 이뤄지지 않음
 - 모바일에 최적화 되어있지 않음

수강신청 장바구니 | 인쇄

학부(과)/전공 컴퓨터공학과 / 컴퓨터공학 조회
기타 E-Learning 조회
강의시간 선택 선택 선택
* 최소 2시간 이상 선택 전체 전선 전필 교선 교필 일선 교직 조회

2017 학년도 1학기 컴퓨터공학과 / 컴퓨터공학 강의시간표 및 강의계획서

학수번호	분반그룹	과목명	학년	학점	과목구분	시간 및 강의실	담당교수	평가방식	비고	시간표
GEB1111-030		인하 새내기 세미나	1	1.0	교양필수	목13,14,15(60주년-201)	원주연	Pass/Fail	컴퓨터/정보통신,	<button>추가</button>
GEB1111-031		인하 새내기 세미나	1	1.0	교양필수	목13,14,15(60주년-206)	성동기	Pass/Fail	컴퓨터/정보통신,	<button>추가</button>
GEB1111-032		인하 새내기 세미나	1	1.0	교양필수	목13,14,15(60주년-207)	이경자	Pass/Fail	컴퓨터/정보통신,	<button>추가</button>
GEB1111-039		인하 새내기 세미나	1	1.0	교양필수	수17,18,19(본관소강당)	안영미	Pass/Fail	외국인전용,	<button>추가</button>
GEB1112-001		크로스오버 1: 인간의 탐색	1	2.0	교양필수	월4,5,6,7(60주년-101)	정연재	Pass/Fail	자연계열 학생 수강,	<button>추가</button>
GEB1112-002		크로스오버 1: 인간의 탐색	1	2.0	교양필수	월9,10,11,12(60주년-101)	정연재	Pass/Fail	자연계열 학생 수강,	<button>추가</button>
GEB1112-003		크로스오버 1: 인간의 탐색	1	2.0	교양필수	월4,5,6,7(60주년-108)	도윤정	Pass/Fail	자연계열 학생 수강,	<button>추가</button>
GEB1112-004		크로스오버 1: 인간의 탐색	1	2.0	교양필수	월9,10,11,12(60주년-108)	도윤정	Pass/Fail	자연계열 학생 수강,	<button>추가</button>
GEB1112-005		크로스오버 1: 인간의 탐색	1	2.0	교양필수	월4,5,6,7(60주년-201)	정영진	Pass/Fail	자연계열 학생 수강,	<button>추가</button>
GEB1112-006		크로스오버 1: 인간의 탐색	1	2.0	교양필수	월9,10,11,12(60주년-201)	성희활	Pass/Fail	자연계열 학생 수강,	<button>추가</button>
GEB1112-007		크로스오버 1: 인간의 탐색	1	2.0	교양필수	월4,5,6,7(60주년-206)	정태욱	Pass/Fail	자연계열 학생 수강,	<button>추가</button>
GEB1112-008		크로스오버 1: 인간의 탐색	1	2.0	교양필수	월9,10,11,12(60주년-206)	임학성	Pass/Fail	자연계열 학생 수강,	<button>추가</button>
GEB1114-001		크로스오버 3: 사회의 탐색	전체	2.0	교양필수	수4,5,6,7(60주년-101)	장세진,김영순,윤여준,안형태	Pass/Fail		<button>추가</button>
GEB1114-002		크로스오버 3: 사회의 탐색	전체	2.0	교양필수	수9,10,11,12(60주년-101)	장세진,장연화,윤여준,안형태	Pass/Fail		<button>추가</button>
GEB1114-003		크로스오버 3: 사회의 탐색	전체	2.0	교양필수	목4,5,6,7(60주년-101)	김성탁,이영운,고재민,정진영	Pass/Fail		<button>추가</button>
GEB1114-004		크로스오버 3: 사회의 탐색	전체	2.0	교양필수	목9,10,11,12(60주년-101)	이영운,고재민,최준혁,정진영	Pass/Fail		<button>추가</button>

시간표등록

저장 | 인쇄

학수번호	분반그룹	과목명	학년/영역	학점	과목구분	시간 및 강의실	담당교수	평가방식	비고	NO
삭제	CSE4205-004	컴퓨터공학 종합설계	4/	3.0	전공필수	목19,20,21,22,23,24(화-220)	이문규	상대평가		1
삭제	CSE3206-001	오플레이팅시스템	3/	3.0	전공선택	화4,5,6(화-001) / 목7,8,9(화-120)	송민석	상대평가		3

* 수강을 희망하는 과목(순위) 순으로 입력하기 바랍니다.
* 경대/IT공대 수강생들은 버고린에 "선수과목 미이수"로 표시된 과목은 지정된 선수과목을 먼저 이수해야 하는 과목으로서, 본 수강신청화면에서는 해당 과목이 표시되지 않습니다. (공학인증 제외학과는 선택사항)

2017 학년도 1학기 수강회장

학부	컴퓨터공학과	학번	12114497		
전공	컴퓨터공학	성명	한정		
교시/요일	월	화	수	목	금
1교시 [09:00 ~ 09:30]					
2교시 [09:30 ~ 10:00]					
3교시 [10:00 ~ 10:30]					
4교시 [10:30 ~ 11:00]	오플레이팅시스템 CSE3206-001 송민석				
5교시	오플레이팅시스템				

■ 컴퓨터에서 본 화면입니다.

The screenshot shows a table of assignments with the following columns: 학번 (Student ID), 본번그룹 (Group), 과목명 (Subject), 학년 (Year), 학점 (Grade), 과목구분 (Subject Type), 시간 및 강의실 (Time and Classroom), 담당교수 (Instructor), 평가방식 (Evaluation Method), 비고 (Notes), and 시간표 (Schedule). The data includes:

학번	본번그룹	과목명	학년	학점	과목구분	시간 및 강의실	담당교수	평가방식	비고	시간표
CHM1001-001		서비스플랫폼과 인터랙션	1	3.0	교양필수	설비(점검이)	김민정(으)씨 영	상대평가	추가	
CEM2001-001		비즈니스 모델 개발 기 조	2	3.0	교양필수	설비(점검이)	박승숙	상대평가	추가	
CHM1021-013		화학1	1	3.0	교양필수	설비(점검이)	최병식	상대평가	추가	
CHM1021-013		일반화학	전체	3.0	교양필수	설비(점검이)	이정환	상대평가	추가	
GEC1111-007		생활한문	1	1.0	교양필수	설비(점검이)	윤민현	상대평가	추가	
GEC1111-008		생활한문	1	1.0	교양필수	설비(점검이)	김경	상대평가	추가	
GEC1111-009		생활한문	1	1.0	교양필수	설비(점검이)	박숙숙	상대평가	추가	
GEB1131-010		생활한문	1	1.0	교양필수	설비(점검이)	류준동	상대평가	추가	
GEC1111-011		생활한문	1	1.0	교양필수	설비(점검이)	이우희	상대평가	미래융합대학	추가
GEC1011-001		당신에게 개인화 서비스(으)	전체	3.0	교양선택	설비(점검이)	조별조	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC1014-002		상상력개척형 예술(으)	전체	3.0	교양선택	설비(점검이)	주승준	상대평가	미래융합대학 설. 1학기	추가
GEC1014-001		고스케어리어학	전체	3.0	교양선택	설비(점검이)	조병준	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC1015-001		영화로보는문화세계	전체	3.0	교양선택	설1(점검이)	김장원	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC2004-001		국제관찰학개론	전체	3.0	교양선택	설기(점검이)	주승호	점대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC2004-002		국제관찰학개론	전체	3.0	교양선택	설기(점검이)	주승호	점대평가	EnginA국제학부 부 학연단 수강 신청(5월 1일 까지)	추가
GE2010-001		동북아의 친밀한족	전체	3.0	교양선택	설기(점검이)	김경희	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC3001-001		한반도이야기	전체	3.0	교양선택	설1(점검이)	박재선	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC3002-001		환경과 인간	전체	3.0	교양선택	설기(점검이)	조강현	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC3003-001		생활 속 미어모아이기	전체	3.0	교양선택	설기(점검이)	김유경	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC3005-001		생활과학학	전체	3.0	교양선택	설기(점검이)	이우주	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC3008-001		지속가능성 위한 환경 의미(으)	전체	3.0	교양선택	설기(점검이)	이익모	상대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가
GEC3009-001		현대경영과학의 이해	전체	3.0	교양선택	설기(점검이)	소자원	점대평가	기말고사 6월 1주 수강신청(5월 31일 까지)	추가

The screenshot shows a weekly schedule table with the following columns: 학수번호 (Course Number), 분반그룹 (Group), 과목명 (Subject), 학년(학기) (Year), 학점 (Grade), 과목구분 (Subject Type), 시간 (Time), and 평가방식 (Evaluation Method). The data includes:

학수번호	분반그룹	과목명	학년(학기)	학점	과목구분	시간	평가방식	제작	인쇄
CSE4205-004		컴퓨터공학 통합설계	4/	3.0	전공필수	목 19:30-21:22(3) 24(4-22)	이종규 상대평가	1	
CSE3208-002		오피레이팅시스템	3/	3.0	전공선택	화 13:14(13:14-08-1) 1/1(13:14)(13:14- 12)	송은석 상대평가	2	
CSE3206-001		오피레이팅시스템	3/	3.0	전공선택	화 5:50(화)(001)/ 5:7:00(화)(12)	송은석 상대평가	3	

■ 모바일에서 본 화면입니다.

- 에브리타임, <http://everytime.kr/> : 시간표 제작 및 익명 커뮤니티

■ 추천 시스템 부재

■ 실제 시간표가 수강신청 사이트에 등재 된 후 약 1주일 후 반영

로그인

에브리타임 회원가입

페이스북으로 시작하기 아이디/비밀번호 찾기

우리 학교 커뮤니티 둘러보기

찾으시는 캠퍼스를 검색하세요.

경희대	53,728명
연세대 신촌캠퍸	49,815명
성균관대	42,195명
중앙대	36,365명
이화여대	33,525명
한양대 서울캠퍸	29,418명
홍익대 서울캠퍸	27,250명
고려대 안암캠퍸	25,117명
단국대	23,813명
숙명여대	21,970명

**대학 생활을 더 편하고 즐겁게,
에브리타임**

GET IT ON Google Play Download on the App Store

**350만 대학생을 위한
국내 1위 대학생 서비스 에브리타임!**

2017년 1학기

시간표 1
8학점 (03/23 16:15 변경)

기본시간표

시간표 2

+ 새 시간표 만들기

월	화	수	목	금	
1교시					오전 9시
2교시					
3교시					
4교시					오전 10시
5교시					
6교시					오전 11시
7교시	오피데이팅시스템 이상진 학-001	오피데이팅시스템 이상진 학-222			
8교시					오후 12시
9교시					
10교시					오후 1시
11교시					
12교시					오후 2시
13교시					
14교시					오후 3시
15교시					
16교시					오후 4시
17교시					
18교시					오후 5시
19교시					
20교시					
21교시					
22교시					
23교시					
24교시					
25교시					
26교시					

교수평론 3: 사랑의 탐색
김상학/사정윤, 고예나, 정진경 60주년 101

상업정보교재구매지도법
윤기중 9회334

■ 컴퓨터에서 본 화면입니다.

시간표 1

학기변경

시간표 1 시간표 2

	월	화	수	목	금	
1교시						오전 9시
2교시						
3교시						오전 10시
4교시				크로스오버 3 : 사생의 탐색 김성태, 이명 윤, 고민정 진영 60주년-101		오전 11시
5교시						
6교시						
7교시	오피레이팅 시스템 이성진 하-001	오피레이팅 시스템 이성진 하-222				오후 12시
8교시						
9교시						오후 1시
10교시						
11교시						오후 2시
12교시						
13교시						오후 3시
14교시						
15교시						오후 4시
16교시						
17교시						오후 5시
18교시						

■ 모바일에서 본 화면입니다.

- **SNUTT, <http://snutt.kr/>** : 서울대학교 시간표 제작

- UI/UX 디자인 부자
- 서울대 자대에만 사용되는 시스템
- 추천 시스템 부자

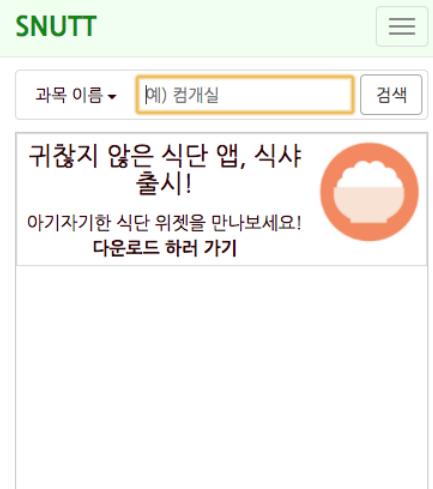
SNUTT 2017-1▼ 강의 찾기 내 강의 / 0학점 저장하기

과목 이름 ▾ 예) 경계설 검색

코드	번호	이름	구분	개설학과	학년	학점	시간	장소	교수	정원
귀찮지 않은 식단 앱, 식사 출시! 아기자기한 식단 위젯을 만나보세요! 다운로드 하러 가기										

	월요일	화요일	수요일	목요일	금요일	토요일
0교시 08:00~09:00						
1교시 09:00~10:00						
2교시 10:00~11:00						
3교시 11:00~12:00						
4교시 12:00~13:00						
5교시 13:00~14:00						
6교시 14:00~15:00						
7교시 15:00~16:00						
8교시 16:00~17:00						
9교시 17:00~18:00						
10교시 18:00~19:00						
11교시						

■ 컴퓨터에서 본 화면입니다.



	월	화	수	목	금	토
0교시 08:00-09:00						
1교시 09:00-10:00						
2교시 10:00-11:00						
3교시 11:00-12:00						
4교시 12:00-13:00						

- 모바일에서 본 화면입니다.

- 주관식 강의백서, <https://goo.gl/PKoyQc> : 인하대학교 주관식 강의백서
 - 불편한 UI/UX
 - 어려운 검색, 강의에 대한 자료의 낮은 가독성
 - 모바일 최적화 부재
 - 단순한 엑셀자료, 데이터 부족

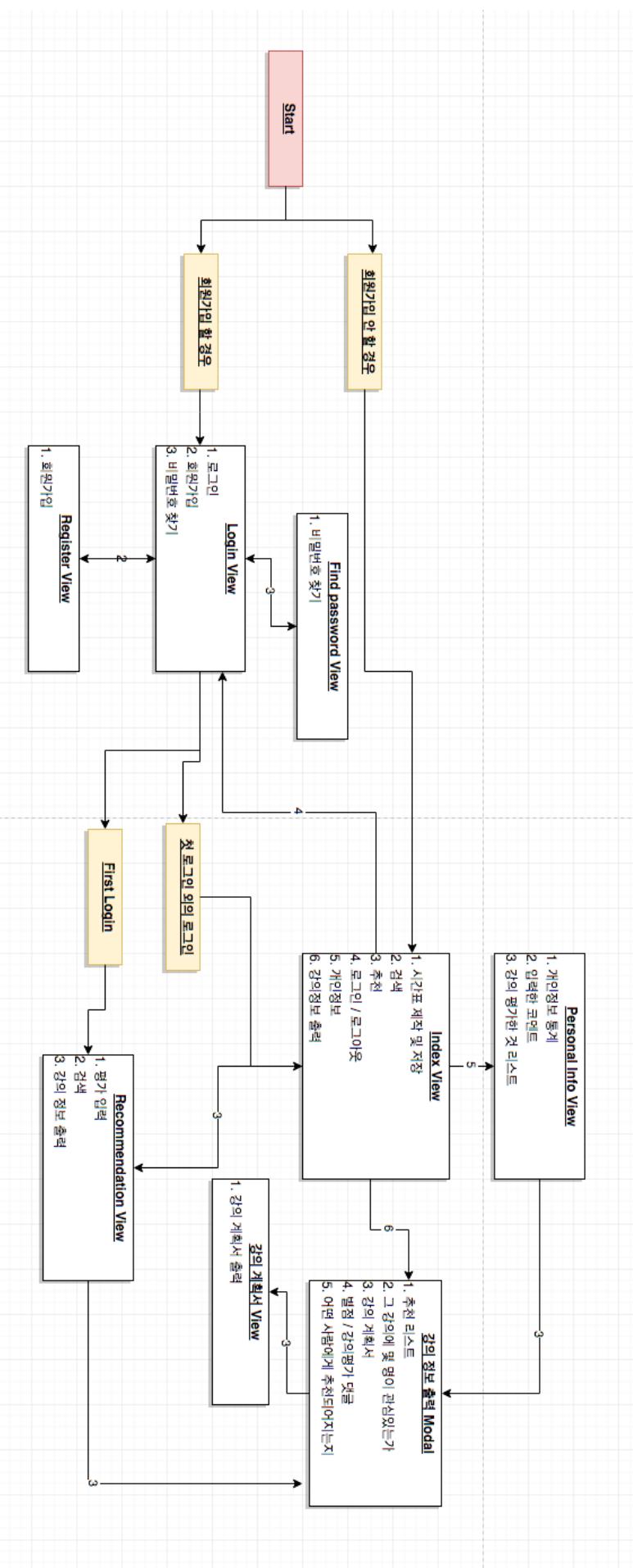
번호	강의명	교수명	결과확인
1	3D프린팅의 이해	이재원	(확인)
2	경영학의 이해	윤여준	(확인)
3	경영PC활용A	구대환	(확인)
4	경영PC활용B	구대환	(확인)
5	경영학의 이해	매초	(확인)
6	공학과 여성	이경자	(확인)
7	광고와 생활	임수현	(확인)
8	교양가창	조병욱	(확인)
9	국제개발협력의 이해	김용호	(확인)
10	기업가 정신과 창업	소국천	(확인)
11	기업수요지향기술 종합설계	최종식	(확인)
12	기초수학A	이영진	(확인)
13	기초수학B	이영진	(확인)
14	기초일본어	최태화	(확인)
15	논리학의 이해	우환식	(확인)
16	동화와 마법의 상상력	김미령	(확인)
17	리더십 개발을 위한 멘토링	이영규	(확인)
18	민법의 이해	최홍섭	(확인)
19	범죄와 사회	강은영	(확인)
20	범죄와 사회	심영주	(확인)
21	법학의 이해	심영주	(확인)
22	비즈니스커뮤니케이션	김진상	(확인)
24	생활 중국어 회화	한서영	(확인)
25	생활속의 경제원리A	김종철	(확인)
26	생활속의 경제원리B	김종철	(확인)
27	생활속의 경제원리C	김종철	(확인)
28	생활속의 경제원리D	김종철	(확인)
29	서양 음악사A	반율리	(확인)

- 기존 프로젝트와 제안 내용과의 차이 및 기존 문제 해결 방안
 - 추천 시스템

- 사용이 간편한 UI/UX
- 기술적인 어려움
 - 특정 OS 기반이 아닌 브라우저 기반으로 어려운 디바이스 최적화
 - 다양한 브라우저 호환 필요
 - 특정 기간에 급격히 늘어나는 요청에 대한 대응 필요
 - 잦은 연산에 의해 로드가 커져 시스템을 분리 필요
 - 강의에 해당하는 데이터를 정보 통신처에서 직접 얻지 못해 크롤링을 하여 사용함
 - 크롤링이 적절하게 모두 이뤄지는지에 대해 검토 필요
 - 데이터가 모두 공통적인 속성을 갖고 있지 않아 적절한 평가와 추천을 위해 데이터가공이 필요함
 - 악의적 데이터 입력과 추천 몰림 현상으로 적절한 추천이 이뤄지지 않을 수도 있음
 - 이 악의적 데이터를 어떻게 관리할지에 대한 정책이 필요함
 - 새로 추가되는 강의를 추천할 때 어떠한 방식으로 할 지에 대한 방식또한 고민이 필요
 - 새로 추가되는 강의는 평가를 할 수 없음

개발 내용

- 개발 플랫폼 : **Web Application**
 1. 높은 호환성(OS에 국한되지 않음)
 2. 자바스크립트 기반의 높은 확장성
 3. 상대적으로 많은 라이브러리와 자료 존재
 4. 서버 언어로 `Node.js` 를 사용하여 클라이언트와 동시 개발 용이
- 상세 요구사항
 - 기능적 요구사항 (functional requirements)
 - 과목 recommendation systems
 - 시간표 제작
 - 모바일 환경을 고려한 반응형 웹
 - 기존보다 쉽고 간편한 검색
 - 비기능적 요구사항 (non-functional requirements)
 - 해당 과목에 관심이 있는 인원 명시를 통한 경쟁 가능성 제시로 수강 인원 분산 유도
 - 추천 시스템 사용 유도



플로우차트입니다. 가로로 출력했습니다.

개발 방법

- 사용하게 될 open sources
 - React
 - facebook에서 만든 유저인터페이스 라이브러리
 - 페이스북, 인스타그램, 넷플릭스에서 널리 쓰임
 - UI를 component화 해서 손쉬운 재사용이 쉽게 만드는게 특징
 - VIRTUAL DOM이라는 개념을 사용해서 상태의 변함에 따라서 선택적으로 UI를 렌더링
 - Babel
 - ES6 와 ES7 모든 웹 브라우저에서 호환 가능하도록 변환해주는 transcompiler
 - Redux
 - React로 이루어진 앱에서 효율적인 데이터 흐름을 위해 사용하는 라이브러리
 - Node.js
 - Javascript 엔진으로 만들어진 언어로써, Non-blocking I/O와 단일 스레드 이벤트 루프를 통한 높은 처리 성능이 특징인 언어
 - express
 - Node.js에서 가장 많이 쓰이는 웹 프레임워크
 - passport
 - Node.js에서 가장 많이 쓰이는 유저 인증을 위한 미들웨어
 - SASS
 - CSS를 확장하는 pre-processor
 - webpack
 - Javascript 모듈화를 지원해주는 모듈화 엔진
 - mysql
 - 가장 널리 쓰이고 있는 오픈소스 RDBMS
 - redis
 - 메모리 기반의 key/value NoSQL DBMS
- 추천 : 특정 교수님의 강의를 추천해줍니다.
 - 사용하는 방식 : Collaborative filtering
 - 유사도 측정
 1. Cosine-similarity : 가장 많이 사용하는 방식. sparse한 데이터에 적합합니다.
 2. Manhattan Distance / Euclidean Distance : Missing value가 적을 때 효과를 발휘한다. dense한 데이터에 많이 사용합니다.
 3. Pearson Correlation Coefficient : grade-inflation이 있는 곳에 적합합니다.
 - 저희의 점수 입력방식은 학생들이 모든 데이터를 입력할 것이라 보장이 없기에 sparse한 데이터 구조를 띄게 될 것입니다.
 - 그래서 시작을 Cosine-similarity로 하게 되었습니다.
 - 이후 MAE가 낮게 나와 신뢰도가 낮게 판단이 된다면 다른 방식을 찾아 반영할 예정입니다.

- 실제 구현하는 방식

- 먼저 개인의 강의 성향을 파악하기 위해 수강했던 과목에 점수를 입력하게 됩니다.(1점 ~ 10점)
- 입력받은 점수를 기반으로 각 사용자들과 Similarity를 구하게 됩니다. 이 방식으로 Cosine-similarity를 사용합니다.
- 유사도를 구하여 가장 유사도가 높은 K명의 집단을 구합니다.
 - 여기서 저희는 계산횟수와 신뢰성을 높이기 위해 과내에서 검색을 하게 됩니다.
- K명의 유사집단 내에서 제가 듣지 않은 과목을 찾고 그 과목에 해당하는 예상 평점을 유사도를 이용하여 계산합니다.
- 이후 이 예상 평점이 높은 과목을 추천하게 됩니다.
 - 만약 그 과목을 수강하고 이후에 그 과목에 대한 평점을 입력 했을 때 MAE를 계산하게 됩니다.
 - 주기적으로 유저의 MAE를 확인하여 제대로 추천이 되고 있는지 판단하여 알고리즘을 개선합니다.

- 가용 가능한 data sources

- sugang.inha.ac.kr에서 수강가능한 시간표
- 예전 총학생회에서 조사한 주관식 교양 백서
- 인하대학교 정보통신처에서 학생들의 수강신청 기록들
 - 위 사항은 정보통신처에서 연락을 주지 않아 불가능할 것으로 예측됩니다.
- '인하대학교 시간표'에서 데이터 요청
 - 위 데이터는 입력 받아 사용할 알고리즘(cosine-similarity)를 파이썬과 몽고디비를 이용하여 구현해보는데 사용했습니다.
 - 결과

<u>id</u>	<u>group</u>	<u>title</u>	<u>grade</u>	<u>credit</u>	<u>class_type</u>	<u>time</u>	<u>instructor</u>
2917		의사소통 영어: 중급	1	3	교양필수	월7,8,9,목1,2,3(60주년-701)	잉글스
3373		데이터통신	3	3	전공선택	월13,14,15,수13,14,15(하-222)	노영태
4062		오퍼레이팅시스템	3	3	전공선택	화10,11,12(하-120) /목4,5,6(하-001)	송민석
4434		인터넷 프로그래밍	2	3	전공선택	수4,5,6,7(하-001) /목16,17,18,19(하-324)	이영석
4614		데이터베이스	3	3	전공선택	수10,11,12(하-120) /금13,14,15(하-220)	최범기
4769		문제해결기법	3	3	전공선택	수19,20,21,22,23,24(하-121)	이문규

- 위 과목을 들을 예정이던 학생과 비슷한 성향을 가진 학생을 찾아 유사도를 판단하고 그 유사도를 가진 학생들이 내가 선택하지 않은 어떤 과목을 선택했는지에 대해 검토해 보았습니다.

```
/Library/Frameworks/Python.framework/Versions/3.7/lib/python3.7/site-packages/itertools.py:9: DeprecationWarning: itervalues() is deprecated. Use dict.values() instead.
  return iter(values, None)
it works
('3779', 2.714389224171833)
('4615', 2.0514159629793616)
('3371', 2.0514159629793616)
('5083', 2.0477225575051663)
('2617', 0.7302967433402214)
('3062', 0.6666666666666667)
('3780', 0.5477225575051661)
('4202', 0.5477225575051661)
('2888', 0.5000000000000001)
('3068', 0.5000000000000001)
('4462', 0.5000000000000001)
('4199', 0.3651483716701107)
('5070', 0.3651483716701107)
('3199', 0.3651483716701107)
('4063', 0.3651483716701107)
('5350', 0.3651483716701107)
('3782', 0.3651483716701107)
('2743', 0.3651483716701107)
```

- 유사도 측정 후 가중치를 뒤 과목에 합을 구한 결과입니다. 이 결과를 변환해 보면

id	title	class_type	time	instructor
2617	의사소통 영어: 중급	교양필수	월1,2,3,수4,5,6(60주년-805)	헴브리
3062	의사소통 영어	교양필수	월10,11,12,목4,5,6(60주년-706)	코너
3371	데이터베이스	전공선택	월13,14,15(화-120) /수16,17,18(화-001)	이주홍
3779	오퍼레이팅시스템	전공선택	화4,5,6(화-001) /목10,11,12(화-220)	송민석
4615	데이터통신	전공선택	수10,11,12,목10,11,12(화-001)	타메르
5083	지식재산 이론과 실제	교양선택	금1,2,3,4,5,6(60주년-511)	최규승

- 이러한 과목을 다른 학생들은 선택했던것을 알 수 있습니다.
- 크롤링한 데이터를 바로 사용할 수 있는 상태가 아니기 때문에 가공이 필요

```
월1,2,3,수4,5,6(2복695)
월1,2,3(2복463A) /수9,10,11(2복463B)
월1,2,3,수1,2,3(2복467)
월1,2,3,수1,2,3(2복467)
월1,2,3,화4,5,6(2복671)
월1,2,3,4,5,6(2복474)
월1,2,3,금9,10,11(2동252)
월1,2,3,목4,5,6(2동254)
월1,2,3,수4,5,6(5동505)
월1,2,3,4(5서455) /금9,10,11,12,13,14,15,...
월1,2,3(2남225) /금4,5,6(2남227)
월1,2,3,화9,10,11(2남227)
월1,2,3,4,5,6(2남409)
```

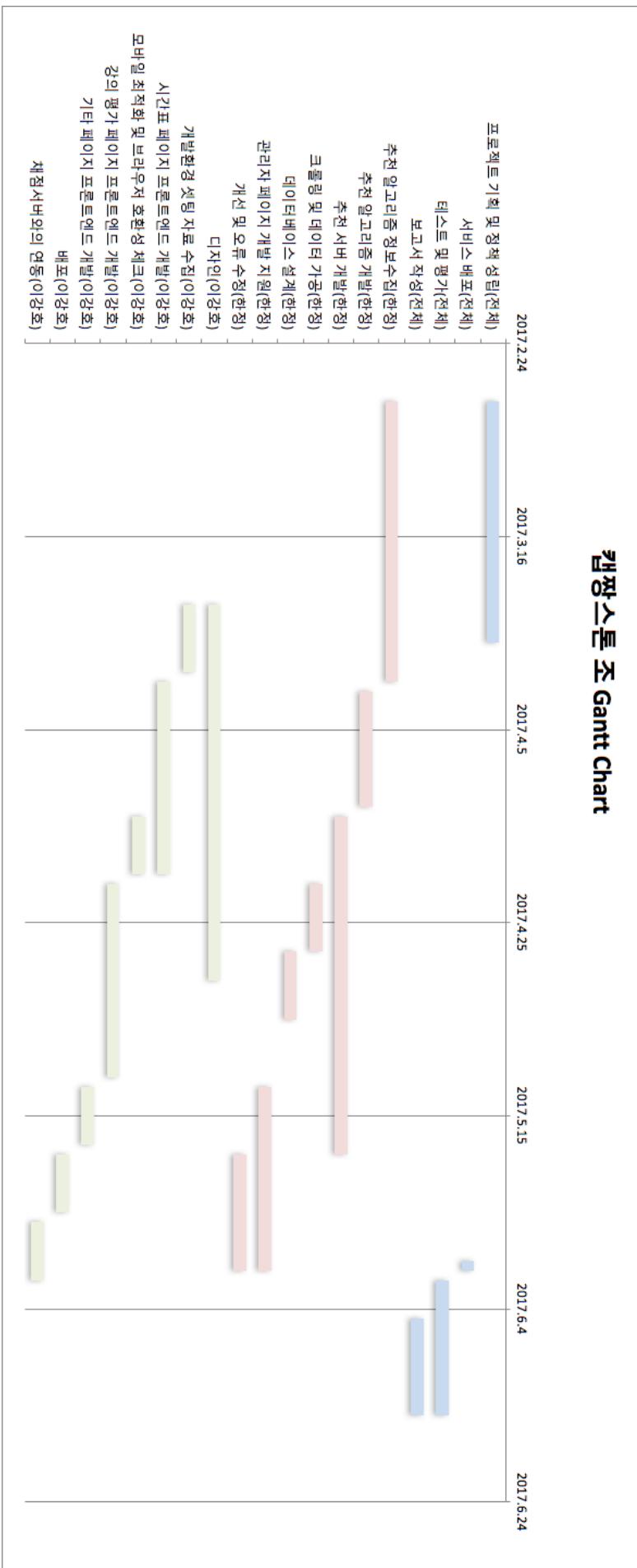
- 크롤링한 데이터의 시간 부분을 예로 들면 각 시간이 균일하게 나타나지 않는 것을 볼 수 있습니다.
- 이것을 어떻게 균일하게 만들고 자동화 시킬지에대한 고민 또한 필요합니다.
- 윤리, 정책 등 비기술적 고려사항
 - 평가가 이뤄지는 시기(6월)가 수강신청이 실제로 이뤄지지 않는 시기
 - 학생들의 지속적인 유입 방안 필요
 - 정보 수집 범위
 - 회원가입시 정보 수집에 대한 동의
 - 개인정보 없이 직접적인 강의 평가를 받아 정보 누적

- 학생들의 악의적으로 입력하는 데이터의 필터링 문제

계획 및 일정 (Plan and schedule)

- 프로젝트 규모 산정
 - 예상 소요기간
 - 기획 및 디자인 : 2017.03.02 ~ 2017.03.27
 - 개발 기간 : 2017.03.27 ~ 2017.05.31
 - 성과 측정 기간 : 2017.06.01 ~ 2017.06.15
 - 프로젝트 규모 : 두개의 서버(추천서버, 웹서버)와 8개(예정)의 뷰
- 요소 작업들에 대한 roadmap 구성
 - Gantt chart를 이용한 각 구성 작업 당 소요 기간

캡짱스톤 조 Gantt Chart



간트차트입니다.

- 평가기준

- 알고리즘 개선에 따라 강의 추천의 정확도가 올라가는 것을 확인
 - MAE(Mean Absolute Error) : 예측 점수와 실제 점수의 차이를 측정하여 평균을 낸다. 이를 통해 해당 알고리즘의 정확성과 가중치를 판단 할 수 있습니다.

$$MAE = \frac{\sum_{i=1}^n |p_i - a_i|}{n}$$

- MAE값이 낮을 수록 점수 예측이 잘 된것이고 신뢰성있게 추천이 되었음을 의미합니다.
- 추천 시스템 연산의 속도
 - 사용자의 요청에 따라 추천을 해줘야하는데 이때 많이 일어나게 되는 연산을 어떻게 나누게 될 지에 대해, 데이터 관리를 어떻게 할지에 대해 평가가 이뤄져야 합니다.
- 사용자의 웹/모바일 환경 접근성 향상 및 기존 사이트보다 편리한 UI/UX 제공 여부

참고문헌

- Stanford Jeffrey D. Ullman *mining massive datasets* 강의 수업자료 (<http://infolab.stanford.edu/~ullman/mmds/ch9.pdf>)
- 카이스트 제 8회 ROSAEC 워크샵자료 *Recommendation System* "협업 필터링을 중심으로"(<http://rosaec.snu.ac.kr/meet/file/20120728b.pdf>)
- 미네소타 대학 Technical Report, "Application of Dimensionality Reduction in Recommender System -- A Case Study"
- mahout을 이용한 recommendation system Overview(<https://mahout.apache.org/users/recommender/recommender-documentation.html>)
- mysql-python connector overview(<https://dev.mysql.com/doc/connector-python/en/connector-python-connectargs.html>)