

# 강의평, 경쟁률과 강의실간 이동시간을 고려한 최적의 시간표 추천 서비스

[1차 발제]  
6조 '수신통'  
기다연, 임효진, 이가영

# 목차

## 1. 팀 소개

- '수신통' 이름의 근원
- 팀원/역할배분 소개



## 2. 프로젝트 소개

- 전체적 아이디어 개요
- 데이터 소개
- 추천 알고리즘 소개
- 시각화 알고리즘 소개

## 3. 주차별 활동 계획

1~7주차별 세부적인 활동 계획



## 4. 논의사항

현재까지 진행된 논의사항 요약



# 1. 팀 소개

## '수신통' ?

수강신청 도와주는 통계학과 친구들

저희는,

### 기다연

- 데이터 수집
- 데이터 모델링  
(Genetic Algorithm)
- 감정 사전 구축
- 선행연구 조사

### 임효진

- 데이터 전처리
- 데이터 모델링  
(Apriori)
- 워드클라우드 시각화
- 선행연구 조사

### 이가영

- 데이터 전처리
- 데이터 모델링  
(Genetic Algorithm)
- 베타 테스트
- 선행연구 조사

## 2. 프로젝트 소개



매학기 신입생부터 졸업을 앞둔 고려대학교 학생들이 공통적으로 갖는 고민이 있습니다.

**"이번 학기 시간표 어떻게 짜지?"**

### 서비스 소개

학생들이 중요시하는 요소나 전체적인 강의 특성 (강의평, 수강신청 경쟁률, 강의실간 이동 소요시간)을 기반으로 강의별 우선순위를 부여, 우선순위가 높은 강의로 구성된 시간표 후보 추천 서비스  
&  
연관분석을 통해 해당 강의와 연관성 높은 키워드로 구성된 워드클라우드 시각화 서비스

### 목표

- 하나의 시간표 후보를 완성하기까지의 여정 축소
- 강의 선택 결정 비용 감소
- 학생 개개인 취향별 적합한 시간표를 1대1로 맞춤형 추천
- 고려대학교 스마트 캠퍼스 구축 니즈 해소

### 기대효과

- 코로나19로 인해 비대면 강의를 주로 들어 캠퍼스가 익숙하지 않은 20,21학번에게 강의실간 이동거리를 고려한 시간표 추천
- 시선배와의 협업 가능성 (과목 추천 이후 해당 과목들로 구성된 시간표 추천)

# 데이터 소개

상대평가  
/ 절대평가

출석확인자율화

강의실간 이동  
소요시간  
(연강인 경우만)

수강신청 경쟁률

강의평  
(학점/성취감/  
학습량/난이도)

# 추천 알고리즘 소개

Genetic Algorithm (유전 알고리즘)

▷ 생명체가 환경에 적응해가면서 진화하는 모습을 모방해 최적의 해를 구하는 기법

▷ 구성 요소

- 염색체 (chromosome)
- 유전자 (gene)
- 자손 (offspring)
- 적합도 (fitness)

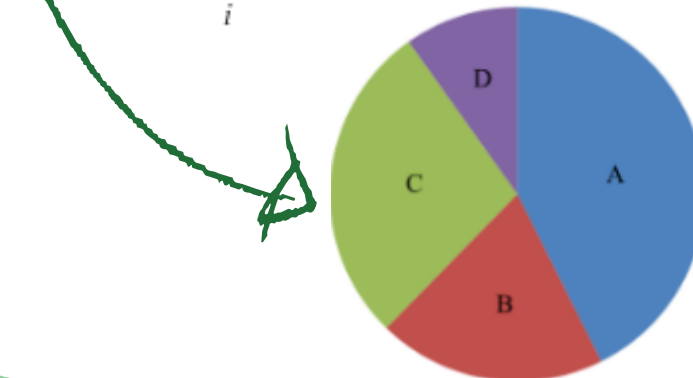
## ▷ 유전 알고리즘의 연산

① 초기 염색체 생성 연산

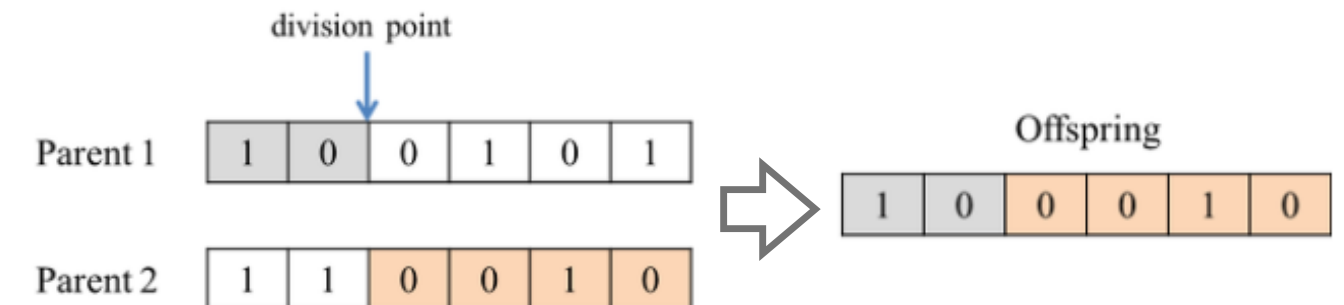
② 적합도 계산 연산

③ 적합도를 기준으로 염색체를 선택하는 연산

$$P(Ch_j) = \frac{f(Ch_j)}{\sum_i^N f(Ch_i)}$$

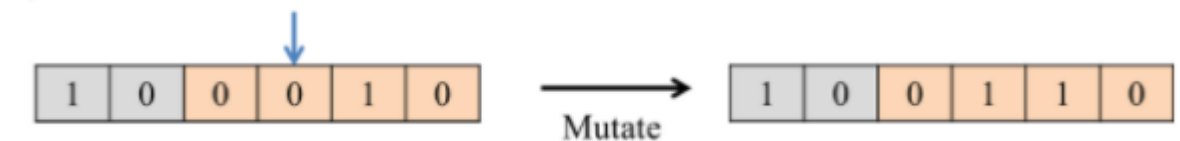


④ 선택된 염색체들로부터 자손을 생성하는 연산



⑤ 돌연변이 연산

a) reverse



b) exchange

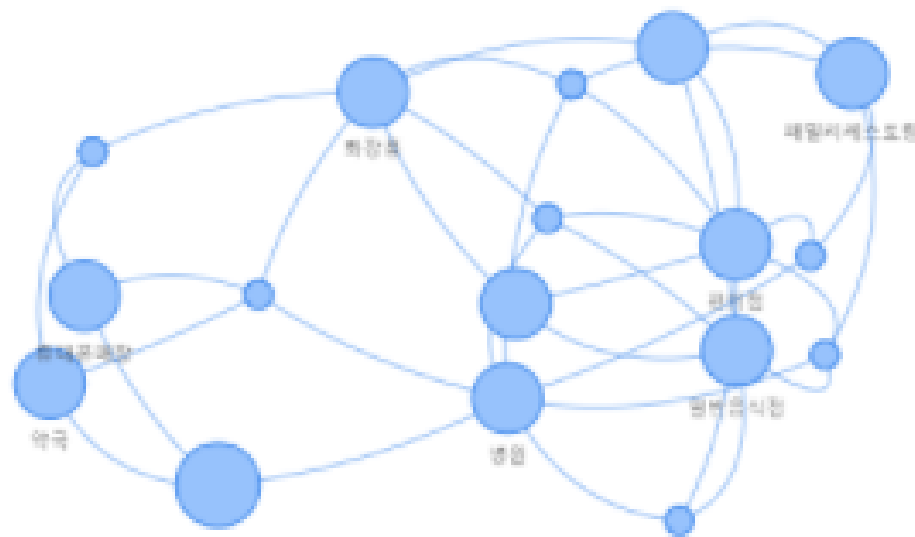


# 시각화 알고리즘 소개

Apriori Algorithm (연관규칙 알고리즘)

## 연관규칙

: 특정 사건이 발생하였을 때 함께 빈번하게 발생하는 또 다른 사건의 규칙



## 지지도 (Support)

: 전체 거래에서 X, Y를 동시에 포함한 거래 비율

$$P(X \cap Y) = (\text{X, Y를 동시에 포함하는 거래 수}) / (\text{전체 거래수})$$

## 신뢰도 (Confidence) ( $X \Rightarrow Y$ )

: X를 포함한 거래 중 Y를 동시에 포함하는 비율

$$P(Y|X) = P(X \cap Y) / P(X) \\ = (\text{X, Y를 동시에 포함하는 거래수}) / (\text{품목 X를 포함하는 거래수})$$

▷ 지지도와 신뢰도 계산 결과를 적용해 의미 있는 연관규칙을 선별한 후 시각화

- 원의 개수 : 연관규칙의 개수
- 원의 크기 : 지지도(support)에 비례해서 커짐

# 감정사전 소개

형태소 분석, 토큰화, 감성분석

## 1) 토큰화

: Khaiii 형태소 분석기를 이용해  
KLUE 강의평 토큰화

## 2) 감정사전 구축

: 과제량, 강의력, 출석, 시험 난이도  
등의 카테고리 별로 직접 감정사전  
구축

## 3) 단어 긍부정 평가

: 직접 구축한 감정사전을 바탕으로  
수강평에 등장하는 단어들을  
긍부정 평가



### 3. 주차별 활동 계획

#### 1~3주차

1. 데이터톤 측 제공 데이터, 강의실간 이동거리 데이터 수집
2. 데이터 전처리
3. KLUe 강의평 데이터 크롤링 작업
4. Genetic Algorithm/Apriori Algorithm 선행 연구 스터디
5. 2차 발제 준비

#### 4~5주차

1. Genetic Algorithm 기반 시간표 추천 알고리즘 개발
2. 강의평 데이터로 감정사전 구축

#### 6주차

Apriori Algorithm과 감정사전을 이용한 강의평 워드클라우드 시각화

#### 데이터톤 이후

1. 알고리즘 고도화로 성능 향상
2. '수신통' 홈페이지/앱 개발

#### 7주차

1. 서비스 베타 테스트
2. 최종 발표 준비

## 4. 논의사항



- ☑ 정기 회의 일자 조정 (매주 화/금요일 오후 10시)
- ☑ 팀 Notion 생성 (팀 내 원활한 커뮤니케이션을 위한 도구)
- ☑ 필요한 외부데이터 컨택요청/확보  
(KLUE에 협조요청, 카카오지도 API 확보, 강의실간 이동거리 데이터 직접 구축)
- ☑ Genetic Algorithm/Apriori Algorithm 선행 논문/연구 조사

# 팀 Notion 활용

## 데이터톤 회의록

☰ 리스트 보기 ▾

📄 [05.28] 온라인 설명회

📄 [06.25] O.T.

📄 [06.26] 1차 발제 회의

📄 [06.29] 화요일 회의

📄 [07.02] 금요일 회의

### [06.26] 1차 발제 회의

📅 회의 일자 2021년 6월 26일

📎 첨부 서류 비어 있음

+ 속성 추가

🗨 댓글 추가

조이름: 수신통 (수강신청 도와주는 통계학과 친구들)

#### Specifications (7주)

##### 1) 데이터 수집, 전처리

- 데이터톤측에서 주어지는 데이터: 서울캠 개설과목 데이터, 개설과목 강의계획안, 강의별 인원 데이터 (2021-1학기 기준)
- 수집해야 할 데이터:

지도API (카카오/네이버)

건물간 거리는 시간 + (층 수 소요시간) → 층마다 다르게 적용 (2층=1분, 3층=2분 등)

2021-1학기 수강희망등록 인원 데이터 문의해보기 → 강의별 경쟁률 계산하기 위함  
(데이터를 얻을 수 있으면 2021-1학기로, 데이터를 얻을 수 없으면 2019-2학기)

KLUE 크롤링 컨택해보기 [klueask@gmail.com](mailto:klueask@gmail.com) (어떤 특정한 패턴이 있는지/따로 데이터를 로깅하고 있으면 줄 수 있는지)

##### 2) 선행 연구/논문 스터디

(유전 알고리즘 / apriori 알고리즘 / Greedy Algorithm)

모든 회의록을 정리해서  
아이디어 발전 과정을 상세히  
기록하고 있습니다.

속성 필터 정렬 🔍 검색 ...

새로 만들기 ▾

2021년 5월 28일

2021년 6월 25일

2021년 6월 26일

2021년 6월 29일

2021년 7월 2일

Notion 워크스페이스를 형성해  
팀원끼리 원활한 커뮤니케이션과  
아이디어 공유를 하고 있습니다.



이번주에 진행한 선행 연구  
스터디는 논문별 중요 내용을  
요약하고 있습니다.

## 선행 연구 스터디

Genetic Algorithm / Apriori Algorithm

### 논문 목록

📄 유전알고리즘을 이용한 강의시간표 작성 시스템 설계

한국컴퓨터정보학회

📄 벌칙 함수에 기반한 유전 알고리즘을 사용한 강의 시간표의 자동 작성

한국정보과학회

📄 Solving Timetable Problem by Genetic Algorithm and Heuristic Search



# 감사합니다

질문이 있다면 말씀해주세요.