캡스톤 디자인- Proposal

‘채점해드림’

팀 이름 : 문차장

팀원 : 20193494문서형, 20190101장세환, 20193038차선우

# 목차

* 지난 발표 내용
* 발표 피드백 내용
* 수정된 프로젝트 소개
* 프로젝트 세부사항
* 사용자 시나리오
* 구현 계획 및 고려사항
* 기술 스택 및 프로젝트 스케줄

# 지난 발표 내용

처음 프로젝트를 준비하면서 저희가 찾은 문제 상황은 학원에서 많은 학생들의 시험지를 채점하면서 낭비되는 시간과 인력이었습니다. 실제로 학원 아르바이트를 경험하면서 가장 번거로웠던 업무가 쪽지 시험 채점이었습니다. 수십 명이나 되는 학생들의 단답형, 문장형 답안을 정답과 일일이 확인하고, 점수를 합산하는 과정이 고되었던 경험이 있습니다. 시간이 많이 소요될 뿐더러 단순 작업의 반복에 집중력이 떨어지면서 채점 오류도 간간히 발생했습니다.

보통 채점을 하는 방식은 다음과 같습니다.

1. 답안지를 학생들에게서 걷는다.

2. 각 답안지의 답을 답지와 비교한다.

3. 정답과 오답을 표시한다.

4. 각 문제의 점수를 확인해 총 점수를 계산한다.

5. 학생 수만큼 이를 반복한다.

6. 결과를 학생들에게 공지한다.

7. 점수에 대한 이의 제기를 받는다.

이 과정들을 자동으로 수행해줌과 동시에 그 결과 통계까지 제시하여 교사의 번거로움을 제거하고 학생들의 학습 효율을 높이고자 ‘채점해드림’서비스를 구상하게 되었습니다.

‘채점해드림’의 핵심 서비스는 다음과 같습니다.

1. 출제자가 시험문제를 출제하고, 시험에 대한 답을 ‘채점해드림’에 입력한다.
2. 선생님은 ‘채점해드림’에 방을 만들고, 학생들은 방에 접속한다.
3. 학생들은 수기로 작성한 답안을 사진을 찍어 방에 업로드한다.
4. ‘채점해드림’이 출제자가 입력한 답안과 비교하여 객관식, 단답형, 문장형 문제까지 채점해서 결과를 공지한다.
5. 해당 시험 뿐만 아니라 이전 기록들 까지 각종 통계를 출제자, 학생 모두에게 제공한다.

# 발표 피드백 내용

위 주제에 대해 저희가 받았던 피드백 내용은 다음과 같습니다.

1. 기존 한글 수기 텍스트 인식 기술의 정확도에 대한 조사가 부족해 보인다.
2. 한글 OCR의 정확도를 개선하는 방법, 알고리즘을 제시하지 않는다면 기술적 난이도는 낮고, 제시한다고 하면 기술적 난이도가 매우 높은데 이에 대해 조사가 부실한 것 같다.
3. 서비스의 타겟층이 너무 광범위하다. 축소된 타겟층에게 필요한 핵심적 서비스를 제공하는 것도 주제 선정의 방법 중 하나다.

# 프로젝트 소개

채점이라는 서비스 개념을 유지한 채 타겟 층을 더욱 섬세히 정하는 방향으로 토의를 해서 새로 기획해보았습니다. 채점 서비스가 가장 필요한 타겟층은

1. 페이지 안에 문제의 집적도가 높을수록 수요가 높을 것이다.

2. 채점에 여러가지 해석이 필요하지 않을수록 수요가 높을 것이다.

3. 주기적으로 시험을 보는 과목에서 수요가 높을 것이다.

라는 점에 착안하여 새로 기획한 채점해드림은 영어학원 교사들의 영어 단어 시험 업무를 도와주는 오프라인 영어 단어 시험 출제, 채점, 관리 서비스입니다.

주요기능은 다음과 같습니다.

1. 출제자는 간편하게 영어 단어 시험지를 생성, 출력할 수 있습니다.
2. 생성된 시험지에 수기로 작성된 답안을 스캔 혹은 업로드하면 자동으로 채점합니다.
3. 각 피시험자에 대한 데이터를 수집하여 개인에게 맞춤 단어장/시험을 제공합니다.
4. 피시험자들의 데이터를 종합하여 출제자에게 학습에 도움이 되는 정보를 제공합니다.

# 프로젝트 동기

이전 기획과 마찬가지로 학원 아르바이트 중 반복되는 채점으로 인한 피로, 스트레스가 이 프로젝트를 생각하게 된 이유입니다. 특히 영어 단어 시험의 경우 위에서 문제로 삼았던 세가지 기준(

1. 페이지 안에 문제의 집적도가 높을수록 수요가 높을 것이다.

2. 채점에 여러가지 해석이 필요하지 않을수록 수요가 높을 것이다.

3. 주기적으로 시험을 보는 과목에서 수요가 높을 것이다.)

에 부합하기에텍스트, 스크린샷, 번호, 폰트이(가) 표시된 사진

자동 생성된 설명 저희는 모든 시험에 대한 솔루션을 제공하려는 기존 기획에서 영어 단어 시험을 대상으로 좁혔습니다. 영어 단어 시험을 제공하는 기존 서비스(Word TEST, 클래스카드 등)을 분석해본 결과 답안의 형태는 OMR, 객관식, 모바일기기나 컴퓨터에서 작성된 텍스트였습니다.

위는 구인 사이트의 영어학원 채점 아르바이트 구인 글들입니다. 2023.09.22 기준, “알바몬” 어플에는 <채점 조교>를 구하는 모집글이 서울 기준 35건, 경기도까지 기준을 넓히면 102건이 올라와 있습니다. 기존 서비스들의 해결 시도에도 불구하고 채점에 대한 수요가 여전히 많습니다.

또한 프로젝트 진행을 위해 현재 학원에서 조교를 하고 있는 3명(앞으로 계속 더 많은 인원에게 실시할 예정입니다.)에게 인터뷰를 진행했습니다.

이XX (용산구 J영어학원 조교)

* 우리 학원은 학생들을 관리하는 학원이다 보니, 학생들마다 스케줄이 다르고, 진도가 다르다. 맞춤형 시험지를 빠르게 만들어주는 기능은 유용하다. 또한 채점하는 시간을 줄이면 다른 학생들의 질문을 더 많이 받아줄 수 있고 효율적일 것 같다.

장XX(대치동 O영어학원 조교)

* 학원 조교의 인력을 많이 줄일 수 있을 것 같다. 오-영어학원의 경우 한 수업 당 150~200명의 학생들이 있다. 수업 시작 전 단어시험을 보고, 다음 수업 쉬는 시간 전까지 조교 4명이서 한시간동안 시험지만 채점하고 있어야 한다. 이렇게 시간이 오래 걸리고 단순한 업무를 대체할 수 있으면 대체하고 싶다.

남XX(송파구 2영어학원 조교)

* 다른 것보다도 재시험지를 자동으로 만들어 주는 기능은 정말 편리할 것 같다. 학생들에게 틀린 단어들에 대해 재시험을 시킬 때, 할 수 없이 같은 시험지로 계속 보게 되는데, 순서를 외워서 답을 적는 등 효율이 떨어진다. 맞춤형으로 자동으로 시험지를 만들어 주는 기능은 매우 도움이 될 것 같다.

저희는 영어학원 선생님들, 조교님들의 고충을 해결하는 서비스를 제공하고자 합니다.

# 프로젝트 세부사항

채점해드림 프로젝트의 목표는 오프라인 영어 단어 시험 출제와 채점, 통계 제공입니다.

1. 시험 출제
2. 선생님이 입력한 단어장에 기반한 기본 시험 출제
3. 학생의 통계(이전 시험의 점수, 오답 문제)에 기반한 맞춤 시험 출제

두가지 기능을 제공할 예정입니다.

맞춤 시험의 경우 각 학생의 시험코드와 함께 인쇄되어 각 시험에 맞는 답지로 채점이 될 예정입니다.

1. 시험 채점

기존 서비스와 달리 수기로 작성된 답안지의 채점을 제공합니다.

답안지 입력은 복합기로 미리 스캔한 답안지를 받는 방식, 사진을 찍어 업로드 하는 방식 두 가지 방식을 제공할 예정입니다.

1. 통계제공

정기적으로 보는 단어시험의 특성을 반영해 해당 분반의 통계와 각 학생의 데이터를 제공할 예정입니다. 이 통계는 위에 기술한 1-b)의 맞춤 출제에도 이용될 예정입니다.

# 사용자 시나리오

1. 영어 학원 선생님은 매일 학생들의 영단어 암기를 확인하기 위해 영단어 시험을 출제해야 한다.

2. 선생님은 채점해드림에 로그인한다.

3. 선생님은 해당 분반의 정보와 학생 명단을 입력하여 새 클래스를 생성한다.

4. 선생님은 채점해드림에 그날에 테스트해야 할 영단어를 입력한다.

5. 채점해드림이 해당하는 영단어 시험지를 만들어주고 선생님은 출력하여 사용한다.

6. 학생들은 시험지에 답안을 쓴다.

7. 시험이 종료되면 선생님이 시험지를 걷는다.

8. 선생님은 학생들의 시험지를 스캔 혹은 촬영하여 채점해드림에 업로드한다.

9. 채점해드림이 시험지에 학생들의 답안을 추출하여 채점을 수행한다.

10. 채점 결과를 통해 학생이 틀린 문제를 따로 분류해 주고, 학생들의 성적 통계를 계산한다. (학생들의 평균 점수, 편차, 높은 오답률을 보인 단어 등 )

11. 이후 틀린 문제 데이터에 따라 채점해드림이 영단어 시험 맞춤 출제를 지원한다.

위 과정을 수행하기 위해 필요한 기능을 정리해 보았습니다.

* 단어장 입력기능
* 시험지 생성 및 내보내기 기능
* 단어장에 기반한 일반 출제
* 학생 통계에 기반한 맞춤 출제
* 답안지 이미지 인식 기능
* 채점 기능
  + - (단어일 경우) 유의어 판별 기능
    - (문장일 경우) 의미 유사도 판별 기능
* 시험 결과 통계 제공 기능

# 구현 계획 및 고려사항

이전 기획과 마찬가지로 사진 업로드가 필요하므로 모바일, 컴퓨터 모두 사용가능한 웹으로 개발 예정입니다.

영어 단어 시험은 크게 두가지로 진행됩니다.

1. 한국어 뜻이 제공되고, 영어 철자를 맞추어야 하는 문제

영어 OCR은 현재 많은 연구가 이루어졌으며, 현재 99.5%이상의 높은 정확도를 보이고 있습니다.

1. 영어 단어가 제공되고, 한국어 뜻을 맞추어야 하는 문제

한국어 뜻을 맞추는 문제는 다음과 같은 순서로 해결할 예정입니다.

1. Wordnet을 이용하여 정답의 유의어를 찾습니다.
2. OCR을 이용하여 한국어 단어를 읽습니다.
3. 해당 단어를 영어로 번역한 후, 정답 단어와 유의어와 비교합니다.
4. 일치하는 단어가 있는 경우 정답을, 아닐 경우 오답을 반환합니다.

한글 OCR의 경우 다른 언어들에 비해 정확도가 떨어지는 편입니다. 이는 한 음운이 한 글자로 되어있어 구조가 단순하고 알파벳 26자 만을 인식하면 되는 영어와 비교하면, 초성-중성-종성으로 이루어진 한글은 조합의 가짓수가 11,172나 되기 때문입니다. 이중 자주 사용되는 글자만 추려도 2350개(EUC-KR)로 매우 많은 양입니다. 현재 상용화 된 프로그램인 ABBYY사의 OCR의 한글 인식률은 98%로 소개되고 있으며, 2018년 논문에 따르면 98.56%까지 증가하였다고 발표되었습니다. Github의 다른 프로젝트들은 95% 이상의 정확도를 가지고 있습니다.

한국어에는 각 글자마다 비슷하게 생긴 글자들이 있습니다.(ex. 관, 판) 이러한 글자들로 인해 정확도가 떨어지게 됩니다. 한국어 뜻을 맞추는 문제의 핵심은 한국어의 맞춤법 또는 한국어의 철자가 아닙니다. 이 점을 이용하여 채점하는 것에 있어서는 위의 연구 이상의 정확도를

방법은 다음과 같습니다.

1. 한국어 글자들을 모양을 기준으로 벡터 공간에 나열합니다.
2. 위의 4단계를 거쳤을 때 오답이 나왔다면, 벡터 좌표 상 가까운 글자로 변환 후 위의 3,4번 단계를 다시 진행합니다.
3. 벡터 좌표상 거리가 일정 거리 이하인 문자들에 대해서 반복합니다.

영단어의 뜻을 학생이 수기로 작성하는 경우 정답을 여러가지 방식으로 쓸 수 있다는 문제점이 있었습니다. 예를 들어 take an exam은 시험을 치다, 시험을 치르다, 시험을 보다 등 여러가지 방식으로 쓸 수 있습니다. 이를 해결하는 방법을 생각해보았습니다.

1. 한글 뜻을 번역기능을 사용하여 영단어로 바꿉니다.

2. 피시험자의 답변과 출제자의 답변을 영어로 비교합니다. 이 때 각 단어별로 동의어들도 확인하여 정확도를 높입니다. Python 모듈인 wordnet을 이용하여 영단어끼리의 동의어를 비교할 예정입니다.

3. 답이 단어가 아닌 구일 경우 유사도 비교 방법(ex. 맨해튼 유사도, 코사인 유사도) 등을 이용하여 최종적인 답안을 비교하여 일정 수치 이상일 경우 정답으로 처리합니다.

# 기술 스택 및 프로젝트 스케줄

**텍스트, 폰트, 스크린샷, 디자인이(가) 표시된 사진

자동 생성된 설명**

**팀 스케줄**

**텍스트, 도표, 라인, 번호이(가) 표시된 사진

자동 생성된 설명**

**프론트엔드(문서형)**

**텍스트, 스크린샷, 라인, 도표이(가) 표시된 사진

자동 생성된 설명**

**백엔드(차선우)**

**텍스트, 번호, 라인, 스크린샷이(가) 표시된 사진

자동 생성된 설명**

**텍스트 인식 및 채점(장세환)**

**텍스트, 스크린샷, 라인, 번호이(가) 표시된 사진

자동 생성된 설명**

# 그 외 고려해본 프로젝트 아이디어

구체화는 하지 못했지만 이런 아이디어도 있었는데 피드백 해 주실 수 있으면 좋겠습니다.

답지 채점 및 오답노트 관리 어플

문제집을 푸는 학생을 위한 어플리케이션입니다.

1. 학생은 채점을 원하는 문제집을 메인 페이지에서 선택합니다.
   1. 해당 문제집이 존재하지 않을 경우 학생은 답지 pdf파일을 올립니다.

-문제집 출판사에서 답지 pdf를 배포하고 있기 때문에 저작권 문제는 없을 것이라고 생각합니다.

2. 학생은 채점을 원하는 문제집 페이지를 사진 찍습니다.

3. 어플리케이션은 해당 페이지를 채점합니다.

a. 수식, 문장 등 채점이 기술적으로 힘든 문제의 경우 답을 화면에 표시해 학생이 직접 채점합니다.

4. 채점이 끝나면 채점결과를 표시합니다. 표시될 정보는 다음과 같습니다.

a. 각 문제의 정/오답 여부

b. 문제를 누를 시 답지의 풀이가 제시

5. 학생은 자신이 원하는 문제를 오답노트에 저장해두고 원할 때 확인할 수 있습니다.

고등학생 때 학교에서 문제집을 풀며 자꾸 덮이는 답지, 좁은 책상 공간, 다수 페이지의 채점을 위해 문제집을 넘겼다가 돌아왔다가 하는 경험이 매우 귀찮았던 경험이 있습니다. 채점하기 귀찮아서 조금 쉴까 생각했던 경우도 자주 있습니다. 이런 귀찮음을 해소하는 어플리케이션이 될 수 있을 것이라 생각합니다.

여러가지 형태의 답지를 분석하는 기술, 문제집 사진에서 답을 뽑아내는 ocr, 이미지 인식 기술이 필요할 것 같습니다.