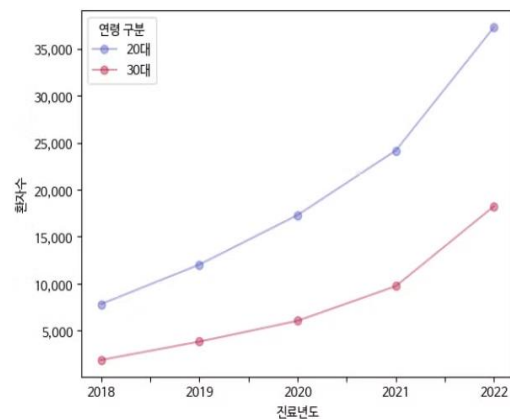


1. 분석 주제

성인 ADHD의 진단 접근성을 높이고 간편한 온라인 진단과 맞춤형 행동 치료법을 추천하는 시스템을 개발하여 극대화된 치료 효과를 기대할 수 있으며 지속적인 모니터링과 피드백을 통해 사용자가 자신의 상태를 관리할 수 있도록 지원하고자 함

2. 프로젝트 배경

구분	2018년	2019년	2020년	2021년	2022년
총 계	59,275	71,362	78,958	99,488	139,696
0~9세	19,509	23,138	23,049	28,966	37,609
10~19세	32,058	35,067	36,148	41,542	53,652
20~29세	7,610	11,308	16,123	22,087	33,672
30~39세	2,325	4,047	5,997	9,159	16,376
40~49세	1,047	1,512	1,967	2,829	4,867
50~59세	251	383	488	761	1,203
60~69세	54	69	97	150	229
70~79세	29	35	34	33	50
80세 이상	15	10	9	23	25



[국민건강보험공단이 제공한 ADHD 진료 현황]

[20~30 대 ADHD 환자의 급격한 증가 추이]

- 국민건강보험공단이 제공한 ADHD 진료 현황을 보면, ADHD 증상으로 진료 받은 환자는 2018년에서 2022년 사이에 2.4배가량 크게 증가
- 특히 2, 30대 ADHD 환자가 급격하게 증가
- 하지만 성인 ADHD는 대부분 성격 문제로 치부되기도 하고 유아 ADHD보다 검사 및 치료법에 대한 접근성이 떨어지는 문제점이 있음

3. 프로젝트 진행 과정



- 데이터 수집 및 전처리
 - ASRS 국제표준 ADHD 검사지를 사용하여 설문조사 진행
 - ADHD 치료법 관련 논문 및 의학 컬럼 수집
 - 설문조사 결과 데이터 수치화 진행 및 문항별 가중치 부여하여 ADHD위험군 판단
- 로지스틱 회귀분석

- ADHD 검사 결과는 유병률과 관련된 데이터로, 정규성을 띄지 않아 정규성의 가정이 필요하지 않으며 이항분류에 효과적인 로지스틱 회귀 분석을 활용함

- **요인분석**

- DSM-5에서는 18가지 문항을 부주의, 과잉행동, 충동성 증상으로 분류하고 있음
- 실제로 행동 추천 시스템을 구축하기 위해, 요인 분석을 진행함

- **텍스트 마이닝**

- 증상별 행동 추천을 위해 성인 ADHD 치료법과 관련된 데이터를 텍스트마이닝 후, 단어를 빈도에 따라 정렬
- 빈도가 높은 단어를 포함하고 있는 논문 및 의학 컬럼들의 문장들을 추출한 뒤 해당 문장의 내용들을 DSM-5 기준에 맞게 분류하였음
- 한글 텍스트마이닝은 의미 유사도를 분석하는 것이 아니라, 텍스트 길이만을 기반으로 단어를 추출해내는 한계점이 있기 때문에 중요 단어를 직접 선택
- 분류된 단어들로 다시 한 번 TF가 높은 즉, 단어의 가중치가 높은 상위 3개 문장을 추출하였고 해당 문장들을 종합하여 치료법을 작성

- **웹페이지 구현**

- 웹 페이지를 통해 사용자들에게 검사 결과 및 증상별 행동 추천 결과를 제공함

- **Oracle DB 반영**

- Oracle DB에 자동으로 적재될 수 있도록 연동

4. 분석 과정

1) 데이터 수집 및 전처리

- 구글폼을 이용하여 성인 ADHD 설문조사 진행

성인 ADHD 설문조사

※본 설문조사는 연구 목적으로만 이용됩니다. ♥

성인 ADHD 예방 추천 시스템 개발을 위한 설문조사 입니다. 5월 7일 오후 12시까지 부탁드립니다.

질문에 답하실 때에는 지난 6개월 동안 귀하가 어떻게 느끼고 행동하였는지를 가장 잘 설명하는 칸에 표시 해주
세요!

- 전혀 그렇지 않다(0 점), 드물게 그렇다(1 점), 가끔 그렇다(2 점), 보통 그렇다(3 점), 매우 자주 그렇다 (4 점)으로 가중치를 부여하여 데이터를 수치화 함

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	ADHD 위험군
0	0	0	0	1	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0
1	1	0	0	2	2	0	1	3	0	0	2	1	0	0	2	2	0	1	0
2	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0
3	2	2	0	2	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0
4	0	0	0	3	0	4	0	0	2	2	1	0	2	1	3	1	0	0	0
...
147	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
148	3	1	0	2	1	0	2	1	0	1	1	0	1	0	2	1	0	1	0
149	3	4	3	3	4	4	3	1	3	3	4	2	1	3	4	3	3	3	1
150	0	0	2	3	1	0	2	1	0	1	3	0	2	0	1	1	0	0	0
151	2	2	3	4	4	0	2	3	2	3	3	0	2	1	1	2	2	1	1

152 rows × 19 columns

[설문 데이터 수치화]

- Shapiro-Wilk 정규성 검정 및 등분산성 검정 진행

```
Q1|ADHD 위험군 => statistic = 0.86793 pvalue = 2.421e-10
Q2|ADHD 위험군 => statistic = 0.83992 pvalue = 1.364e-11
Q3|ADHD 위험군 => statistic = 0.83092 pvalue = 5.820e-12
Q4|ADHD 위험군 => statistic = 0.91514 pvalue = 8.915e-08
Q5|ADHD 위험군 => statistic = 0.88848 pvalue = 2.604e-09
Q6|ADHD 위험군 => statistic = 0.60791 pvalue = 1.768e-18
Q7|ADHD 위험군 => statistic = 0.87723 pvalue = 6.873e-10
Q8|ADHD 위험군 => statistic = 0.87405 pvalue = 4.784e-10
Q9|ADHD 위험군 => statistic = 0.81902 pvalue = 1.976e-12
Q10|ADHD 위험군 => statistic = 0.85333 pvalue = 5.165e-11
Q11|ADHD 위험군 => statistic = 0.8972 pvalue = 7.762e-09
Q12|ADHD 위험군 => statistic = 0.57475 pvalue = 3.435e-19
Q13|ADHD 위험군 => statistic = 0.82422 pvalue = 3.147e-12
Q14|ADHD 위험군 => statistic = 0.72838 pvalue = 1.921e-15
Q15|ADHD 위험군 => statistic = 0.73942 pvalue = 4.054e-15
Q16|ADHD 위험군 => statistic = 0.8346 pvalue = 8.214e-12
Q17|ADHD 위험군 => statistic = 0.54243 pvalue = 7.611e-20
Q18|ADHD 위험군 => statistic = 0.57663 pvalue = 3.760e-19
...
```

[Shapiro-Wilk 정규성 검정]

```
Q1|ADHD 위험군 => Statistic: 1.188e+02, P-value: 1.485e-23
Q2|ADHD 위험군 => Statistic: 9.028e+01, P-value: 6.660e-19
Q3|ADHD 위험군 => Statistic: 8.242e+01, P-value: 1.465e-17
Q4|ADHD 위험군 => Statistic: 1.127e+02, P-value: 1.384e-22
Q5|ADHD 위험군 => Statistic: 2.646e+02, P-value: 3.605e-43
Q6|ADHD 위험군 => Statistic: 2.086e+01, P-value: 7.219e-06
Q7|ADHD 위험군 => Statistic: 1.014e+02, P-value: 9.334e-21
Q8|ADHD 위험군 => Statistic: 1.575e+02, P-value: 2.359e-29
Q9|ADHD 위험군 => Statistic: 1.485e+02, P-value: 4.749e-28
Q10|ADHD 위험군 => Statistic: 9.942e+01, P-value: 1.985e-20
Q11|ADHD 위험군 => Statistic: 1.073e+02, P-value: 1.015e-21
Q12|ADHD 위험군 => Statistic: 1.135e+01, P-value: 8.527e-04
Q13|ADHD 위험군 => Statistic: 1.183e+02, P-value: 1.776e-23
Q14|ADHD 위험군 => Statistic: 4.432e+01, P-value: 1.313e-10
Q15|ADHD 위험군 => Statistic: 4.839e+01, P-value: 2.174e-11
Q16|ADHD 위험군 => Statistic: 1.217e+02, P-value: 5.289e-24
Q17|ADHD 위험군 => Statistic: 1.082e+01, P-value: 1.121e-03
Q18|ADHD 위험군 => Statistic: 1.177e+01, P-value: 6.858e-04
...
```

[등분산성 검정 진행(Levene)]

- Shapiro-Wilk 정규성 검정에서 p-value 가 유의수준(일반적으로 0.05)보다 작아, 귀무가설을 기각하고 "데이터가 정규 분포를 따르지 않는다"고 결론
- Levene 검정에서 p-value 가 유의수준보다 작으면 (보통 0.05 이하), 귀무 가설을 기각하고 대립 가설을 채택함
- 이 경우, 등분산성 가정이 기각되었으므로 등분산성을 가정할 수 없다는 것을 의미 따라서 등분산성 검정이 기각되었을 때 그룹 간의 분산이 서로 다르다고 해석할 수 있음
- 두 그룹 간의 잔차가 등분산성을 만족하지 않는다고 해석할 수 있음
- Levene 검정 결과, p-value 가 매우 낮으므로 등분산성을 충족하지 않고 이분산성이 있다고 판단
- ADHD 검사 결과는 유병률과 관련된 데이터로, 정규성 및 등분산성 검정을 통해 데이터가 이를 따르지 않음을 확인함

- 스피어만 상관관계 분석

```

Q1: correlation = 0.39643, p_value= 4.285e-07
Q2: correlation = 0.37247, p_value= 2.297e-06
Q3: correlation = 0.46155, p_value= 2.168e-09
Q4: correlation = 0.37618, p_value= 1.786e-06
Q5: correlation = 0.29469, p_value= 2.283e-04
Q6: correlation = 0.35882, p_value= 5.641e-06
Q7: correlation = 0.28547, p_value= 3.639e-04
Q8: correlation = 0.31311, p_value= 8.579e-05
Q9: correlation = 0.55483, p_value= 1.198e-13
Q10: correlation = 0.35552, p_value= 6.966e-06
Q11: correlation = 0.39851, p_value= 3.679e-07
Q12: correlation = 0.49604, p_value= 8.124e-11
Q13: correlation = 0.47203, p_value= 8.299e-10
Q14: correlation = 0.39403, p_value= 5.099e-07
Q15: correlation = 0.35102, p_value= 9.255e-06
Q16: correlation = 0.44821, p_value= 7.021e-09
Q17: correlation = 0.44182, p_value= 1.212e-08
Q18: correlation = 0.46347, p_value= 1.822e-09
...

```

- ADHD 위험군으로 판단한 결과와 설문 문항들의 상관관계를 분석하여 모든 변수들이 ADHD 위험군과 양의 상관관계를 가지며, 통계적으로 유의하다는 것을 확인함
- 모든 p-value 가 0.05 보다 작으므로, 귀무가설을 기각하고 대립가설을 채택함
- 즉, 모든 상관 관계는 통계적으로 유의미하다고 볼 수 있음
- 또한, 모든 상관 관계 계수가 양수인 것으로 보아, 모든 변수 간의 관계가 양의 상관 관계를 갖는다고 말할 수 있음

2) 로지스틱 회귀 분석

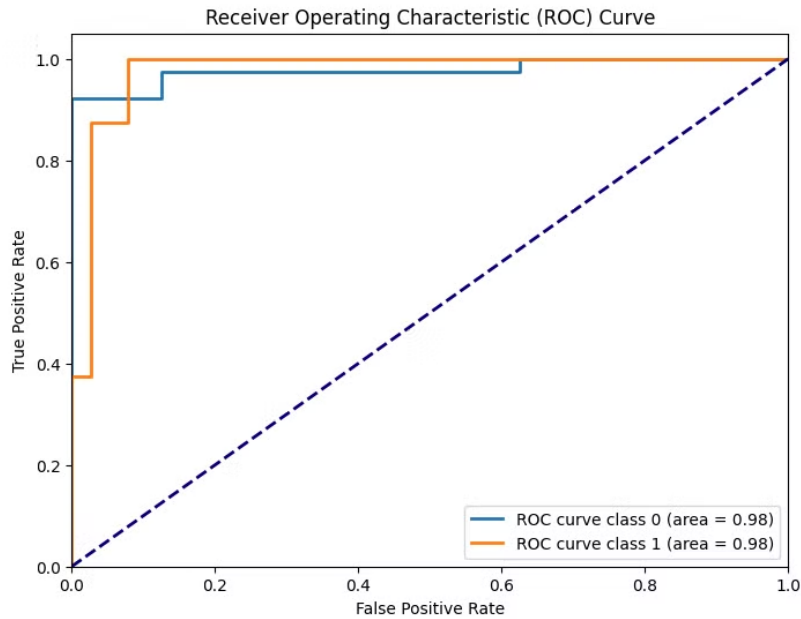
- ADHD 검사 결과는 유병률과 관련된 데이터로, 정규성을 띄지 않아 정규성의 가정이 필요하지 않으며 이항분류에 효과적인 로지스틱 회귀 분석을 활용함

	precision	recall	f1-score	support
0	0.97	0.95	0.96	38
1	0.78	0.88	0.82	8
accuracy			0.93	46
macro avg	0.88	0.91	0.89	46
weighted avg	0.94	0.93	0.94	46

[모델 성능 확인]

- 0 은 ADHD 위험군이 아님을 나타내고 1 은 ADHD 위험군임을 나타냄

- 정확도는 모델이 예측한 것이 맞춘 비율로 전체 샘플 중 약 93 프로가 올바르게 예측됐음을 알 수 있음
- 정밀도와 재현율에 대해 항목 간에 차이가 나는 이유는 ADHD 위험군임을 나타내는 1 이라는 결과에 대해서 보수적으로 도출했기 때문
- 쉽게 말하자면, ADHD 가 확실히 아닌 사람을 분류하는 것에 집중을 했기 때문에 두 항목 간에 차이가 발생함

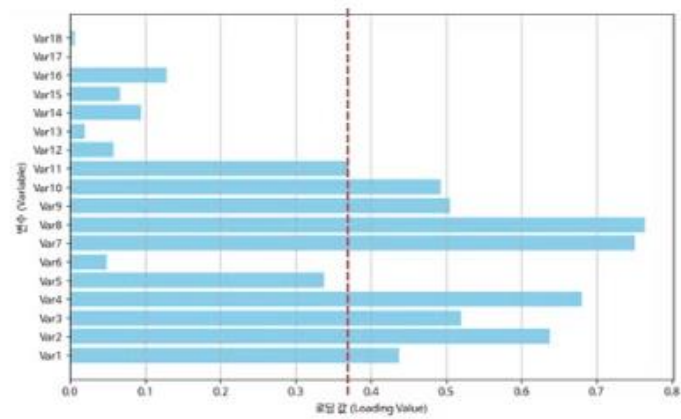


[모델 ROC Curve]

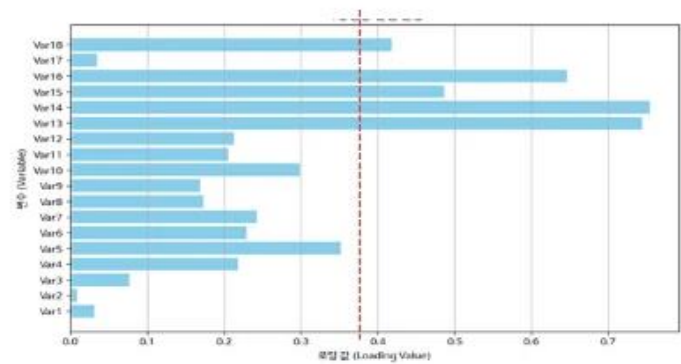
- ROC Curve 결과에서도 해당 모형이 매우 높은 성능을 보여주는 것을 확인

3) 요인 분석

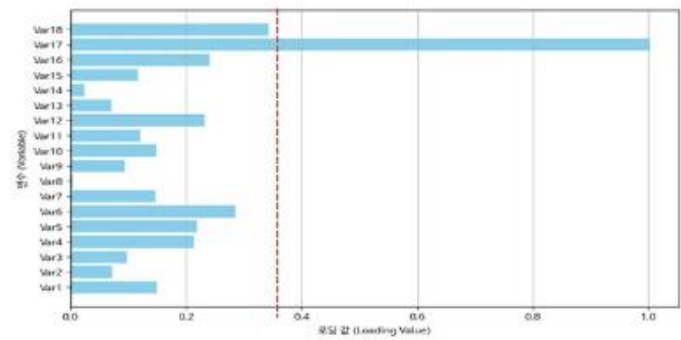
- 추천 시스템을 구축하기 위해서는 ADHD 위험군을 몇가지 분류로 나누는 것이 필요, 유사한 증상을 갖는 문항들을 분류하기 위해 요인 분석을 진행



[부주의 관련 문항]



[과잉행동 관련 문항]



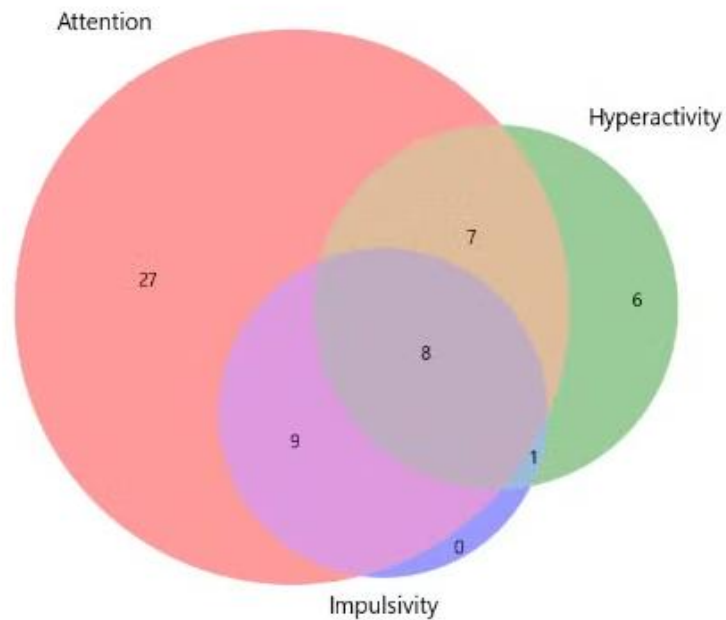
[충동성 관련 문항]

- DSM-5에서는 18 가지 문항을 부주의, 과잉행동, 충동성 증상으로 분류하고 있음
- 실제로 행동 추천 시스템을 구축하기 위해, 요인 분석을 진행하여 문항들이 요인 적재값 0.37 을 기준으로 부주의, 과잉행동, 충동성 증상으로 분류되는 것을 확인함

4) 텍스트 마이닝

- 왜 부주의에 관한 문항이 다른 두 요인보다 많을까?

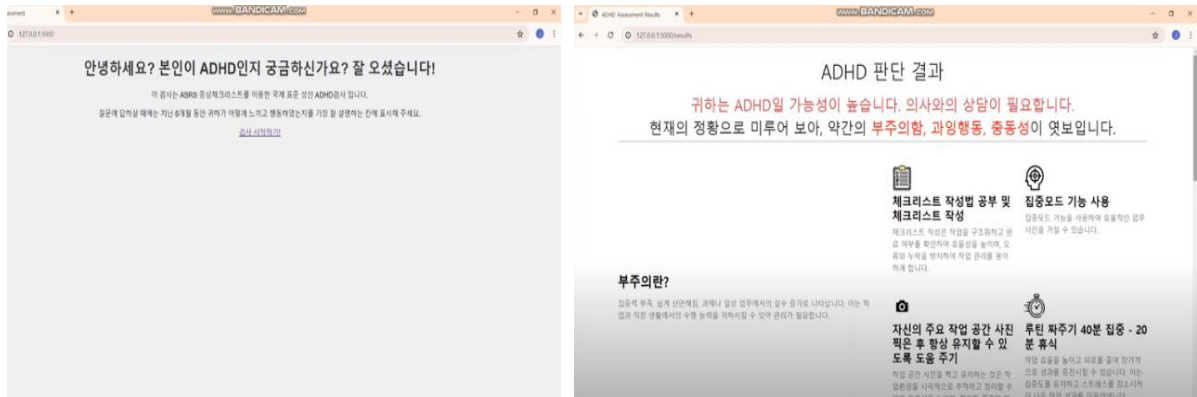
- 실제로 성인 ADHD 의 경우에는 아동 ADHD 와 달리, 성인이 되면서 과잉 행동 또는 충동성의 증상은 사라지고 부주의한 증상만 남는 경우가 많다고 함
- DSM-5 분류 기준에서도 부주의와 관련된 문항의 비율이 높음



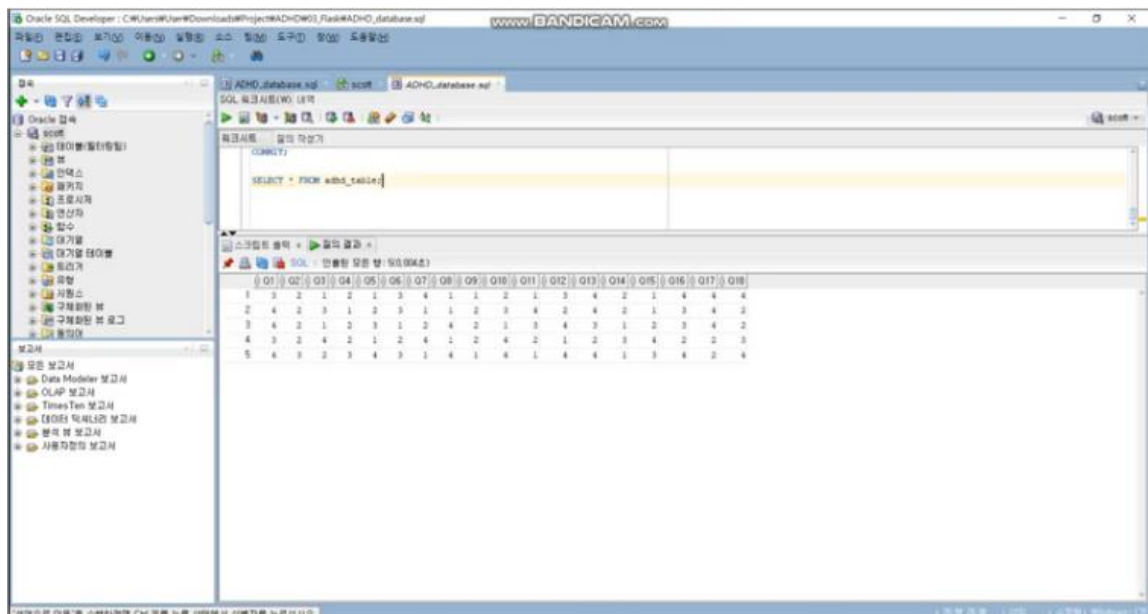
[실제 설문 결과의 문항 분류 편중 확인 시각화]

- 증상별 행동 추천을 위해 성인 ADHD 치료법과 관련된 데이터를 텍스트 마이닝한 다음, 단어를 빈도에 따라 정렬
- 빈도가 높은 단어를 포함하고 있는 논문 및 의학 컬럼들의 문장들을 추출하고, 해당 문장의 내용들을 DSM-5 기준에 맞게 분류하였음
- 한글 텍스트마이닝은 의미 유사도를 분석하는 것이 아니라, 텍스트 길이만을 기반으로 단어를 추출해내는 한계점이 있기 때문에 중요 단어를 직접 선택
- 분류된 단어들로 다시 한 번 TF 가 높은 즉, 단어의 가중치가 높은 상위 3 개 문장을 추출하였고 해당 문장들을 종합하여 치료법을 작성

5) 웹페이지 구현



6) Oracle DB 반영



- 웹 페이지를 통해 얻은 사용자들의 검사지 데이터가 Oracle DB 에 자동으로 적재될 수 있도록 연동

5. 분석 결과

- 웹 페이지를 통해 성인 ADHD 에 대한 간편한 검사과정을 제공하고 사용자들에게 검사 결과 및 증상별 행동 추천 결과를 제공함

6. 프로젝트를 통해 얻은 인사이트

1) 한계점



분석 가능 데이터 부족



치료법 선택



자기보고식 설문지의 한계점

- 개인정보 상의 이유로 환자 데이터를 구할 수 없어 설문지로 데이터를 대체했고, 이에 따라 분석 가능한 데이터가 부족했던 아쉬움이 있음
- 논문이나 의학 칼럼이 아닌 실제 환자 처방전으로 텍스트마이닝을 진행을 했다면 보다 전문적이고 현실적인 치료법을 제공할 수 있지 않았을까라는 아쉬움이 있음
- 자기보고식 설문지의 한계점으로, 응답자가 자신의 증상을 과장하거나 감추는 등의 의도적인 왜곡이 발생할 수도 있기 때문에 신뢰할 수 있는 정보를 제공하기에 제약을 가할 수 있음

2) 보완점

- 실제 ADHD 를 발생시키는 다양한 요인들, 예를 들어, 직종이나 수면 시간 등의 데이터들을 추가한다면 좀 더 세분화적으로 분석과 결과를 도출할 수 있을 것으로 보임
- 또한 실제로 성인 ADHD 환자들에게 약물 이외에 어떤 치료를 진행하는지에 대해서 알 수 있다면 좀 더 현실적이고 효과적인 치료법을 제공할 수 있을 것으로 보임

3) 서비스 발전 방향



알람 서비스



집중력 향상 게임 서비스



다이어리 기능 서비스

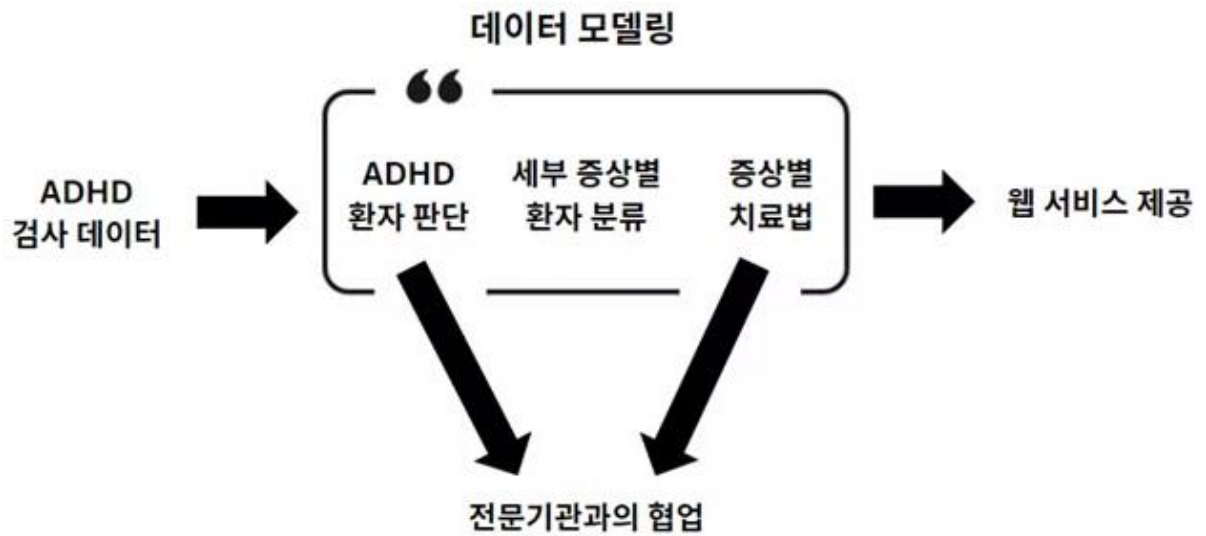


마음챙김명상 서비스



음악 추천 서비스

- 부주의 증상과 관련해서는 일정 알람 서비스나 집중력 향상 게임 서비스, 다이어리 기능 서비스 등을 추가
- 과잉 행동 및 충동성은 명상이나 클래식 음악 같은 음악 추천 서비스를 제공하는 방향으로 시스템 발전



[성인 ADHD 증상별 행동추천 시스템 개발 모식도]

7. 활용 데이터

- 구글폼을 통한 설문지
- ASRS 국제표준 ADHD 검사지
- 성인 ADHD 의 인지행동 치료 (방수영, 2016)
- 주의력결핍 과잉행동장애의 진단과 치료 (이정섭, 옥선명, 2008)
- ADHD 성향 대학생을 위한 인지행동치료 프로그램의 개발 및 효과 검증 (안재순, 2016)
- 주의력결핍 과잉행동장애(ADHD) 관련 예술치료 연구 동향 (김다윤, 신지현, 2023)
- 통제촉진매체를 활용한 집단미술치료가 주의력결핍과잉행동장애(ADHD)성향 성인의 문제행동감소에 미치는 영향 (홍아람, 이모영, 2016)
- 성인 ADHD 대학생을 위한 인지행동치료(CBT)와 마음챙김기반 스트레스 감소 개입(MBSR)의 효과 비교 (신주영, 2018)
- 성인 ADHD 성향 대학생을 위한 변증법적행동치료 프로그램 개발 및 효과 분석 : 주의집중 곤란과 정서조절 곤란을 중심으로 (조안남, 2023)
- 주의집중에 영향을 미치는 음악활동 효과 메타분석 (박충수, 2015)