

2022년 데이터 분석 청년 인재양성 데이터 분석 전문교육과정

데이터로 같이, 가치 있게(With Value)!

데이턴십 해커톤 제 5회

경기도 원격화상 투약기 설치 최적 입지 선정

분석 결과보고서

참여조: 5조

참여자: 김인제(조장), 김진석

강민주, 김민주, 김정경, 김재연

씨에스리 컨소시엄

CSLEE **kpc** 한국생산성본부

Copyright © CSLEE Consortium

CSLEE Consortium의 사전 승인 없이 본 내용의 전부 또는 일부에 대한 복사, 배포, 사용을 금합니다.

목 차

1. 분석 개요	4
1.1. 분석 배경 및 개요	4
1.2. 분석 목적 및 방향	7
1.3. 분석 결과 활용 방안	14
2. 분석 데이터	14
2.1. 분석 데이터 목록	14
2.2. 데이터 상세 설명	15
2.3. 데이터 정제 방안	23
3. 분석 프로세스	27
3.1. 분석 프로세스	28
3.2. 분석 내용 및 방법	29
4. 분석결과	32
4.1. 현황 분석	32
4.2. 분류 분석	33
4.3. 최종 모델	37
4.4. 입지 순위 결과	38
5. 활용 방안	48
5.1. 문제점 개선 방안	48
5.2. 업무 활용 방안	48
5.3. 한계점	48

6. 참고자료(Reference)	50
7. 부록	53

1. 분석 개요

1.1 분석 배경 및 개요

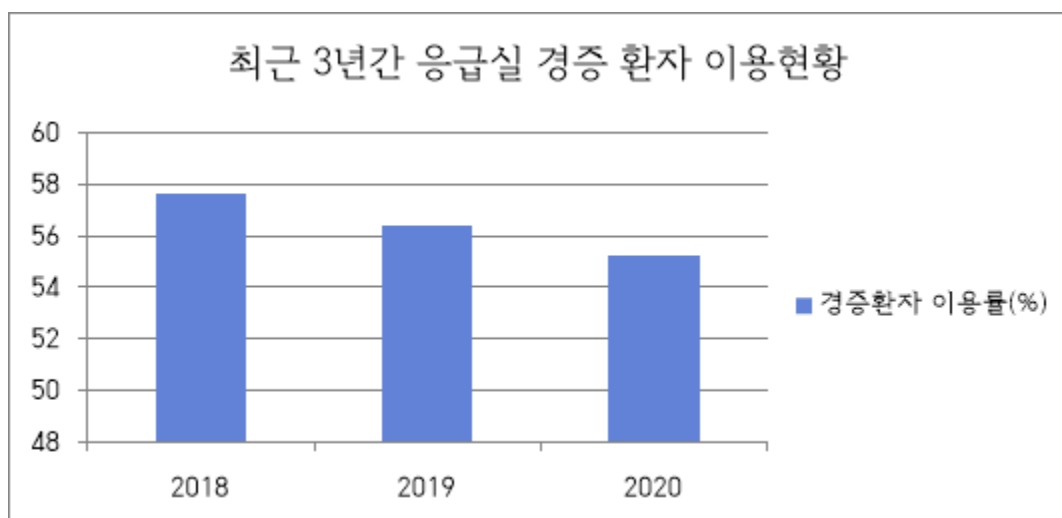
- 심야시간대 환자들이 이용할 수 있는 의료서비스 폭이 좁고 심야의료시설 이용 부담이 큼

☞ 응급실

- 경증 환자의 비용부담이 큼

* 전국 20곳에 위치한 심야의료시설 이용 시, 권역응급의료센터는 54,830원, 전문 및 응급의료센터는 47,520원, 그 밖의 응급실은 18,280원을 별도로 부담해야 함. 비응급환자가 응급실에서 단순 치료를 받거나 약 처방을 받으면 진찰료와는 별도로 이 비용을 전부 부담해야 함 ¹⁾

* 보건복지부와 국립중앙의료원의 ‘최근 3년간 응급실 경증 환자 이용현황’에 따르면 2018년 57.6%, 2019년 56.4%, 2020년(1~6월) 55.2%로 경증 환자 응급실 이용이 다소 줄어드는 경향이 있었으나 여전히 절반이 넘음 ²⁾



1) <https://www.mk.co.kr/news/it/view/2015/12/1229007>

2) https://m.health.chosun.com/svc/news_view.html?contid=2020103002051

☞ 24시 편의점

- 의약품의 종류 부족

* 경제정의실천시민연합에서 1745명 대상 설문 결과, 편의점 안전상비의약품목을 확대해야 한다는 응답 비율이 86.8%로 높음 ³⁾

- 부작용과 오남용에 대한 소비자의 우려

* 광주시 공공심야약국 위원회에서 281명 대상 설문 결과, 편의점에서 약구매 시 부작용과 오남용에 대한 우려로 심야약국을 주로 이용한다는 응답 비율이 22.1%로 나타남

■ 정부는 공공심야약국을 위 대안으로 시행

☞ 공공심야약국

- 정의

「약사법」 제20조에 따라 개설 등록된 약국으로서 심야시간대와 공휴일에 의약품 및 의약외품 구매의 편의를 제공하는 약국을 뜻함

- 운영체제

22시부터 새벽 1시까지 운영 중 (27개 약국 365일 운영 : 8개 약국 요일별 운영)

■ 공공심야약국의 한계점

- 부족한 수

3) <http://www.dailypharm.com/Users/News/NewsView.html?ID=242643>

* 현재 운영 중인 공공심야약국은 2022년 기준 총 163개이며 응급실 개수 (510개)의 약 31.9%로 절반을 못 미치는 수치임

- 짧은 운영시간

* 현재 공공심야약국의 운영종료시간은 새벽 1시까지이지만, 자정부터 새벽 3시까지 심야의료시설을 이용하는 환자들은 전체 9%로 적지 않은 수치임

0-3시	3-6시	6-9시	9-12시	12-15시	15-18시	18-21시	21-24시	계
423,250	271,010	368,421	763,145	722,491	738,992	801,060	721,292	4,809,661

[표 1] 전국 응급실 내원시간 현황

- 약사들의 참여도 저조 및 반발

* 공공심야약국의 운영비용과 정부 지원금이 적어 약사들의 참여도가 저조함

* 2022년 7월 약사회는 기획재정부에게 예산안을 약 24억원 지원금을 요청했지만, 난항을 겪어 운영에 문제가 생김

■ 정부는 2022년 연말에 원격화상 투약기 시범사업을 본격 게시할 것으로 계획 진행 중

"화상투약기 시범사업, 1단계는 비수도권 배분"

김정주 기자 2022-07-23 06:00:52

가 가

- ❖ 복지부, 실증특례 연말에 시작...한약사 참여는 불가능
- ❖ 복지부 관계자, 전문가자협의회와 질의답변
- ❖ 편의점 안전상비약 품목 확대는 당장 논의할 사항 아냐

[데일리팜=김정주 기자] 정부는 화상투약기 실증특례 시범사업을 연말에 본격 개시할 것으로 전망했다. 기계 생산과 인증 등 절차가 남아있기 때문에 도입을 서두르기 어렵다는 게 이유다. 다만 애초에 한약사는 논의 대상이 아니어서, 여러 방법을 적용하더라도 참여하는 것은 사실상 불가할 것으로 보인다.

편의점 안전상비약 품목 확대의 경우 업계 일각에서 요구는 있지만 당장 논의할 사항은 아니라는 게 정부의 판단이다.

하태길 보건복지부 약무정책과장과 양대형·여정현 사무관은 전문가자협의회의 현안 질의에 이 같이 답했다. 다음은 현안에 대한 일문일답이다.

[그림 1] 원격화상 투약기 시범사업 운영 예정 (출처 : 데일리팜)

1.2. 분석 목적 및 방향

■ 목적

경기도 지역 심야시간대, 주말이나 공휴일에 의약품을 구입함에 있어 소비자들의 불편한 문제를 해소하고자 원격으로 의사와 약사의 통제를 받을 수 있는 원격화상 투약기의 최적 입지를 선정하여 의료서비스의 접근성을 높이고 경증환자의 심야의료시설 이용 부담을 줄이고자 함

■ 원격화상 투약기

- 정의

약사와 화상통화를 통해 약사가 원격으로 상담한 약을 내주는 기계로 약국 개설자는 등록된 약국에 '일반의약품 원격화상 투약기'를 설치하고 고용한 약사가 시스템을 통해 일반의약품을 판매하도록 함

- 관련 법 및 제도

* 약사법 제21조 제1항 약사 또는 한약사는 하나의 약국만을 개설할 수 있다.

- * 약사법 제21조 제2항 약국 개설자는 자신이 그 약국을 관리하여야 한다. 다만, 약국 개설자 자신이 그 약국을 관리할 수 없는 경우에는 대신할 약사 또는 한약사를 지정하여 약국을 관리하게 하여야 한다.
- * 약사법 제23조 제1항 약사 및 한약사가 아니면 의약품을 조제할 수 없으며, 약사 및 한약사는 각각 면허 범위에서 의약품을 조제하여야 한다.
- * 약사법 제23조 제2항 약사 또는 한약사가 의약품을 조제할 때에는 약국 또는 의료기관의 조제실(제92조 제1항 제2호 후단에 따라 한국 희귀·필수의약품센터에 설치된 조제실을 포함한다)에서 하여야 한다.
- * 약사법 제23조 제3항 의사 또는 치과의사는 전문의약품과 일반의약품을 처방할 수 있고, 약사는 의사 또는 치과의사의 처방전에 따라 전문의약품과 일반의약품을 조제하여야 한다.
- * 약사법 제24조 제4항 약사는 의약품을 조제하면 환자 또는 환자 보호자에게 필요한 복약지도를 구두 또는 복약 지도서(복약지도에 관한 내용을 환자가 읽기 쉽고 이해하기 쉬운 용어로 설명한 서면 또는 전자문서를 말한다)로 하여야 한다.
- * 약사법 제50조 제1항 약국 개설자 및 의약품 판매업자는 그 약국 또는 점포 이외의 장소에서 의약품을 판매하여서는 아니 된다.
- * 약사법 제50조 제2항 약국 개설자는 의사 또는 치과의사의 처방전에 따라 조제하는 경우 외에는 전문의약품을 판매하여서는 아니 된다.
- * 약사법 제50조 제3항 약국 개설자는 의사 또는 치과의사의 처방전이 없이 일반의약품을 판매할 수 있다.
- * 약사법 제50조 제4항 약국 개설자는 일반의약품을 판매할 때에 필요하다고 판단되면 복약지도를 할 수 있다.
- * 정보통신기술 규제 샌드박스 심의위원회에서 규제 특례 과제로 승인받았으며 현재 판매 가능한 일반의약품은 해열, 진통, 소염제, 제산제 등 11개 약효군

■ 원격화상 투약기 도입 시, 장점

- 불필요한 병원 응급실 방문 대체를 통해 2조원 이상의 막대한 재정 절감 등 경제적 효과 기대 가능

- 공공심야약국과 비교 시, 비용 대비 효과가 더 높음

* 원격화상 투약기 한 대의 설치비용이 2000만원인데 이를 1000대를 설치하면 약 200억원 정도가 듦. 원격화상 투약기 1000대를 10시부터 8시까지 운영한다고 할 때, 이를 공공심야약국으로 운영하면 약 500~600억의 재정이 들어감 4)

* 원격화상 투약기 도입 시 1명의 관리 약사가 20~30대의 기계를 운영하므로 약사 인력이 부족한 곳에 대체 가능하며 인건비 절감 효과가 있음

- 비대면 화상 진료를 통해 오남용 문제 방지

* 편의점에서 판매하는 상비약은 의약품 오남용 방지를 위해 1회 1일분 또는 1일 1개만 판매하도록 규정하고 있지만 각각 다른 편의점에서 쉽게 여러 상비약품을 구매할 수 있는 것과 같은 편의점의 허술한 관리시스템을 엿볼 수 있음 5)



[그림 2] 편의점 오남용 발생 사례 (출처 : 충청투데이)

* 실제 2011~2015년 5년간 식품의약품안전처에 보고된 의약품 부작용 보고 4만여건 중 1023건이 상비약에 대한 부작용이었음. 편의점에서 쉽게 판매하는 A사 해열진통제의 경우 650건 이상의 부작용(두드러기, 구토, 어지럼증 등)이 보고됨 (출처 : 충청투데이)

4) <http://www.dailypharm.com/Users/News/NewsView.html?ID=289450>

5) <https://www.cctoday.co.kr/news/articleView.html?idxno=1036388>

- * 24시간 편의점은 안전상비의약품을 아르바이트생이 관리하는 반면, 원격화상 투약기는 약사가 약을 직접 관리하므로 안전성 문제 해결 가능

■ 원격화상 투약기 운영 방향

- 의약품 종류별 정의

* 일반의약품

질병의 치료를 위하여 의사 또는 치과의사의 전문적 지식이 필요하지 않은 의약품

→ 주요 품목 > 아스피린프로텍트정, 마이폴텍장용정, 돌코락스에스장용정, 비오플250산, 지르텍정, 비판텐연고 등

* 안전상비의약품

일반의약품 중 가벼운 증상에 시급하게 사용하며 환자 스스로 판단하여 사용할 수 있는 의약품으로 편의점과 같이 24시간 운영되는 곳에서 판매되는 의약품

→ 주요 품목 > 해열, 진통, 소염제(5종), 소화제(4종), 감기약(2종), 파스(2종)

* 전문의약품

의약품의 한 분류로, 일반의약품과는 달리 반드시 의사나 한의사, 치과의사의 처방이 있어야 함. 병원에서 의사의 처방 정보가 적힌 처방전을 발급받아 약국에 가서 그 처방전을 제출하여야 약사에게 구매가 가능한 약. 즉, 처방전 없이는 절대로 구매할 수 없음

→ 주요 품목 > 주사제, 혈압약, 당뇨약, 디스토마 치료제[6], 통풍약, 편강탕, 향생제, 니트로글리세린, 사후피임약, 전문한의약품, 치료용 마약, 한외마약, 향정신성의약품

- 투약기 별 판매용도 구분

현재 시범 운영되고 있는 일반의약품 한정으로 판매되고 있는 원격화상 투약기의 이용 범위가 좁다고 판단하였음. 따라서 소비자들의 편의성 증대를 위해 판매 범위를 더 넓히고자 하는 취지로 일반의약품 전용 원격화상 투약기와 전문의약품 전용 원격화상 투약기를 분리하여 설치하고자 함

* 현재 미국, 독일, 스웨덴 등 선진 각국에서는 최첨단의 원격적인 의약품 판매시스템을 통해 일반의약품은 당연히 가능하고 심지어 전문의 약품에 대해서도 판매할 수 있도록 하는 문제가 논의되고 있음 ⁶⁾

- 일반의약품 전용 투약기

의약품 중 일반의약품으로만 구성된 원격화상 투약기

→ 안전상비약품을 포함한 모든 일반의약품 구입 전 약사와의 화상상담 필수

- 전문의약품 전용 투약기

의약품 중 전문의약품으로만 구성된 원격화상 투약기

→ 전문의약품은 의사 처방이 필수적이므로 전문의약품 원격화상 투약기의 경우 의사와의 비대면 진료 필요함

→ 의사가 환자의 상태 파악 위해 기계에 체온계, 심박 측정기, 자동혈압계를 설치하도록 함

→ 대면 진료 후, 실물 처방전을 약국 영업시간 내에 제출하지 못하였을 경우를 대비한 실물 처방전 투입구 설치

- 공통 운영 사항

* 화상 투약기 설치 위치는 약국 앞으로 한정하여 현행법에 위반되지 않도록 함

6) <http://www.kyeongin.com/main/view.php?key=20220210010001851>

- * 원격화상 투약기 운영시간은 병원 및 약국 운영시간 종료 시점부터 시작 시간 전까지로 한정함 (병원 및 약국 영업방해 방지)
- * 병원이나 약국이 운영하지 않는 공휴일에는 24시간 운영
- * 비대면 진료 전, 환자 본인 확인 절차 필수
- * 원격화상 투약기 설치를 원하는 약국의 경우, 반드시 설치 전 약사면허증과 비대면 진료를 하는 의사의 의사면허증 확인 필요

■ 경기도 선정 이유

- 원격화상 투약기의 운영과 관리를 원활하게 하기 위해서는 수도권 대상으로 운영하는 것이 적합하다고 판단
- * 7월 10일 복지부와 과기부에서는 약국에 설치, 운영하는 것과 관리의 문제가 중요하기 때문에 수도권에 우선적으로 배치하는 것이 적합하다고 판단하였음
- 수도권 중 인구수 대비 공공심야약국의 수가 서울에 비해 적고 의약품 접근이 취약해 경기도로 선정하였음
- * 의료 취약지대와 시간을 없애기 위한 대안책 중에 하나인 공공심야약국이 경기도 내 의료취약지로 꼽히는 포천, 동두천 등을 포함한 시군 21곳에는 단 한 곳도 없는 것으로 조사되었음 ⁷⁾

	경기도	서울
공공심야약국	39	35
총인구수 (만 명)	1358.6	949.5
비율	0.028	0.036

[표 2] 총인구수 대비 공공심야약국 수

- * 의료기관 출입이 쉽지 않은 코로나 특수상황과 1천380만에 달하는 도내 인구 등 심야 약국 방문 수요 등을 고려했을 때 턱없이 부족한 수준. 2019년 16곳, 2020년 18곳, 2021년 16곳

7) <http://www.joongboo.com/news/articleView.html?idxno=363522418>

재정자립도 낮은 지역 예산 부담 올 첫 복지부 예산은 전국 17억뿐

도는 이 같은 문제를 알고 있지만, 예산 확대에는 "검토가 필요하다"며 소극적인 입장이다.

올해부터 시행된 '경기도 공공심야약국 운영 지원 조례' 제정 당시 비용 추계안을 보면, 인건비 확대 없이 5년 동안 매년 2개소씩 공공심야약국만 늘린다고 해도 43억7천만원(도비 30%, 시·군비 70%)의 예산이 필요하다. 여기에 수당 등 지원을 확대하면 예산이 더 필요한 상황인데, 재정자립도가 낮은 시·군은 예산 부담이 클 수밖에 없다.

지방정부 예산만으로는 공공심야약국 확대가 쉽지 않은 상황에 보건복지부가 올해 처음으로 공공심야약국 운영비 지원 예산 17억원을 마련했다.

하지만 200여개가 넘는 전국 기초단체 중 공공심야약국이 없는 60곳만 지원할 예정이라서 실효성이 클지는 의문이다.

[그림 3] 공공심야약국 지역 예산 부담

■ 공공심야약국 확충에 난항을 겪고 있음

* 경기도 내 약사 인력 부족으로 심야약국 지정 및 운영이 어려움

행정구역	주민등록세대수	약사수	비율
경기도	5,871,380	9,521	0.001621595
서울특별시	4,451,444	9,843	0.002211193
부산광역시	1,551,767	2,964	0.001910081

[표 3] 인구수 상위 3지역의 약사 비율

경기도, 약사 인력난 속 '공공심야약국' 확충 난항

✎ 이지는 | ✎ 입력 2022.01.18 18:51 | ✎ 수정 2022.01.19 17:21 | ✎ 댓글 0



경기도가 '공공심야약국' 확충에 난항을 겪고 있다. 약사 인력난이 그 이유다.

18일 경기도에 따르면 경기도 공공심야약국은 도민에게 의약품 구매 편의를 제공하기 위해 1년 365일 심야 시간인 오후 10시부터 다음날 새벽 1시까지 운영되는 약국이다. 이를 위해선 별도의 '경기도지사 지정'을 받아야 한다. 현재 도내 공공심야약국은 11개 시·군에서 18곳이 운영되고 있으며, 연내 3곳이 추가될 계획이다.

올해 공공심야약국은 지난해보다 5곳이 늘어 모두 21곳으로 확대됐지만, 의료기관 출입이 쉽지 않은 코로나 특수 상황과 1천380만에 달하는 도내 인구 등 심야 약국 방문 수요 등을 고려했을 때 턱없이 부족한 수준이라는 의견이 나온다. 앞서 도내 심야약국은 ▶지난해 16곳 ▶2020년 18곳 ▶2019년 16곳이 운영된 바 있다.

공공심야약국 확대를 추진하고 있는 도와 달리, 약사들은 인력 부족 등의 이유로 심야약국 지정 및 운영이 힘들다는 입장이다.

[그림 4] 경기도의 약사 인력 부족 현황

1.3. 분석 결과 활용 방안

- (편의성 증가) 병원, 약국이 운영하지 않는 심야시간대나 주말, 공휴일에 국민들이 의약품을 쉽게 구입 가능
- (안정성 증가) 원격으로 의료 전문가와 복약 상담을 통해 의약품을 처방받음으로써 편의점 부작용과 오남용을 방지
- (의료비 절감) 경증 환자의 경우 심야에 응급실까지 가지 않아도 되므로 의료비 지출 비용 감소
- (정책 활용) 자동판매기 도입 시 원격화상 투약기 최적 입지 분석을 바탕으로 자동판매기 위치 설정에 도움이 될 수 있음

2. 분석 데이터

2.1. 분석 데이터 목록

‘원격화상 투약기 입지 분석’을 위한 모델을 구축하기 위해서 데이터를 보유하고 있는 공공 포털 기관과 민간기관, 지자체에 데이터를 수집했음. 모델의 정확성을 위해서 최대한 생성 주기가 최신인 데이터를 사용하였고 입지 분석에 타당성을 높이기 위해 다양한 변수들 데이터를 수집하였음.

대상 구분	수집 데이터	구분	중요도	생성 주기	지역 속성	수집 방법
인구데이터	총인구	정형/외부	필수	월간	지점	다운로드
	유아인구					
위치데이터	서울시	정형/외부	필수	수시	지점	-
	행정구역 경기도 행정구역					
적합 입지 데이터	공공심야약국 소유 현황	정형/외부	필수	수시	지점	-
	응급실 소유 현황			수시		
	약국 소유 현황			수시		
수요 밀도 데이터	안전상비의약품 판매업소 현황	정형/외부	필수	월간	지점	-
	어린이집 위치 정보			매년		
	소방서 소유 정보			수시		
	경찰서 소유 정보			수시		
경제 활동 데이터	버스 정류장 정보	정형/외부	필수	수시	지점	-
	지하철 정류장 정보					

[표 4] 수집 데이터 목록

2.2. 데이터 상세 설명

활용(정제) 데이터는 수집된 데이터에서 정제 과정을 거친 데이터임. 입지 분석의 최적화를 위해 행정구역과 인구 데이터를 제외한 모든 데이터의 위치 정보가 필수적으로 요구됨.

수집 데이터	정제 데이터	형식	비고
응급실 데이터	경기도 응급실 위치	CSV	위경도 필수 준비
	서울시 응급실 위치		
약국 데이터	경기도 약국 위치	CSV	위경도 필수 준비
	서울시 약국 위치		
공공심야약국 데이터	경기도 공공심야약국 위치	CSV	위경도 필수 준비
	서울시 공공심야약국 위치		
시군구별 인구 데이터	경기도 시군구별 총인구	CSV	-
	서울시 시군구별 총인구		
시군구 행정구역	경기도 격자별 유아인구수	SHP	-
	서울시 격자별 유아인구수		
	경기도 격자별 총인구수		
	서울시 격자별 총인구수		
	경기도 행정구역		
	서울시 행정구역		
어린이집 및 유치원 현황	경기도 어린이집 및 유치원 위치	CSV	위경도 필수 준비
	서울시 어린이집 및 유치원 위치		
소방, 경찰서 현황	경기도 소방서 위치	CSV	위경도 필수 준비
	서울시 소방서 위치		
	경기도 경찰서 위치		
	서울시 경찰서 위치		
버스 정류소 데이터	경기도 정류소 위치	CSV	위경도 필수 준비
	서울시 정류소 위치		
지하철 데이터	경기도 지하철 위치	CSV	위경도 필수 준비
	서울시 지하철 위치		
안전상비의약품 판매업소	경기도 판매업소 위치	CSV	위경도 필수 준비
	서울시 판매업소 위치		

[표 5] 활용 데이터 목록

■ 응급실 위치 [.csv]

☞ 경기도 및 서울지역의 응급실 위치(주소), 병원명 등이 필수로 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
병원명	텍스트	not-null	-
주소	텍스트	not-null	-
진료시간	텍스트	not-null	-
업무구분	텍스트	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 6] 경기도, 서울 응급실 데이터

■ 약국 위치 [.csv]

☞ 경기도 및 서울지역의 약국 위치(주소), 약국명 등이 필수로 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
약국명	텍스트	not-null	-
주소	텍스트	not-null	-
운영시간	텍스트	not-null	-
대표전화	텍스트		-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 7] 경기도, 서울 약국 데이터

■ 공공심야약국 위치 [.csv]

☞ 경기도 및 서울지역의 공공심야약국의 위치(주소), 약국명 등이 필수로 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
약국명	텍스트	not-null	-
주소	텍스트	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 8] 경기도, 서울 공공심야약국 데이터

■ 시군구별 총인구 [.csv]

☞ 해당 구역별 거주하고 있는 총인구 데이터

컬럼명	데이터 타입	공백(null)	비고
지역명	텍스트	not-null	-
총인구수	숫자	not-null	-
남자인구수	숫자	not-null	-
외국인인구수	숫자	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 9] 경기도 시군구별 총 인구 데이터

■ 시군구별 유아인구 [.csv]

☞ 해당 구역별 거주하고 있는 유아인구 데이터

컬럼명	데이터 타입	공백(null)	비고
지역명	텍스트	not-null	-
유아인구수	숫자	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 10] 경기도 시군구별 유아인구 데이터

■ 격자 [.shp]

☞ 해당 지자체의 표준격차(1km * 1km) 데이터. 경기도 시군구 행정구역 데

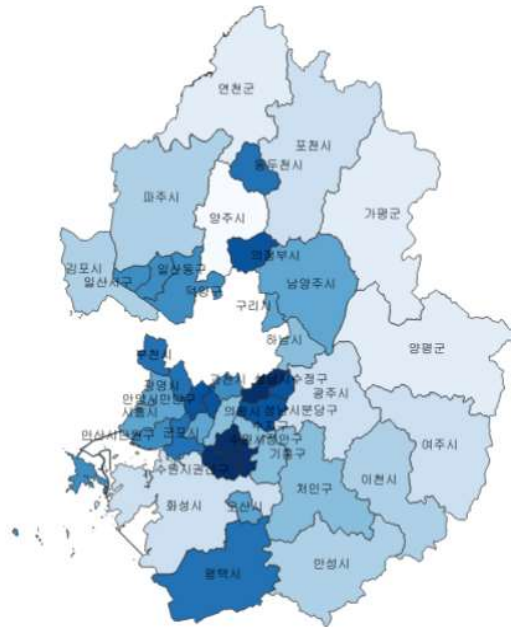
이터에서 격자만 남긴 데이터



[그림 5] 경기도 격자

■ 행정구역 [.shp]

- ☞ 국가교통부에서 제공하는 데이터로 해당 지자체 내의 정보만 추출한 데이터



[그림 6] 경기도 시군구 행정구역

■ 어린이집 및 유치원 위치 [.csv]

- ☞ 경기도 및 서울지역의 어린이집 및 유치원 위치(주소), 시 설명 등이 필수

로 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
시설명	텍스트	not-null	-
주소	텍스트	not-null	-
우편번호	숫자	not-null	-
공사립구분	텍스트	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 11] 경기도, 서울 어린이집 및 유치원 데이터

■ 소방서 위치 [.csv]

☞ 경기도 및 서울지역의 소방서 위치(주소), 기관명 등이 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
시설구분	텍스트	not-null	-
기관명	텍스트	not-null	-
주소	텍스트	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 12] 경기도, 서울 소방서 데이터

■ 경찰서 위치 [.csv]

☞ 경기도 및 서울지역의 경찰서 위치(주소), 기관명 등이 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
시설구분	텍스트	not-null	-
기관명	텍스트	not-null	-
주소	텍스트	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 13] 경기도, 서울 경찰서 데이터

■ 버스 정류소 위치 [.csv]

☞ 경기도 및 서울지역의 버스 정류소 위치(주소), 정류장 명 등이 필수로 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
정류장명	텍스트	not-null	-
정류장 id	숫자	not-null	-
위치	텍스트	not-null	-
관할관청	텍스트	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 14] 경기도, 서울 버스 정류장 데이터

■ 지하철 위치 [.csv]

☞ 경기도 및 서울지역의 지하철 위치(주소), 정거장 명 등이 필수로 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
정거장명	텍스트	not-null	-
위치	텍스트	not-null	-
행정동명	텍스트	not-null	-
법정동명	텍스트	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 15] 경기도, 서울 지하철 정거장 데이터

■ 판매업소 위치 [.csv]

☞ 경기도 및 서울지역의 안전상비의약품 판매업소로 지정된 곳의 위치(주소), 업소명 등이 필수로 포함된 데이터

컬럼명	데이터 타입	공백(null)	비고
업소명	텍스트	not-null	-
영업상태	텍스트	not-null	-
대표전화	텍스트	not-null	-
위치	텍스트	not-null	-
위도	숫자	not-null	-
경도	숫자	not-null	-

[표 16] 경기도, 서울 안전상비의약품 판매업소 데이터

2.3. 데이터 정제 방안

2.3.1. 데이터 정제 방향

- ‘원격화상 투약기 입지 분석’을 하기 위해 격자별 점수를 매겨 최적 입지 순위를 정함. 격자별 입지 분석을 하기 위해서는 격자와 인구를 제외한 모든 데이터에 위치(경도, 위도)에 대한 정보가 필수적임. 다음과 같은 정제 방향을 제시하려 함.

- ▶ 분석 목적에 적합한 데이터인가
- ▶ 약국, 소방서 등 현재 존재하는 데이터인가
- ▶ 각 데이터의 위치(위도, 경도) 정보가 유용한가
- ▶ 각 데이터의 이상치가 존재하는가

2.3.2. 데이터 정제 방법

■ 경기도 격자별 지하철 역 개수를 세기 위한 작업

1) raw 데이터

(1) 이름 : 경기도 지하철 위치 정보

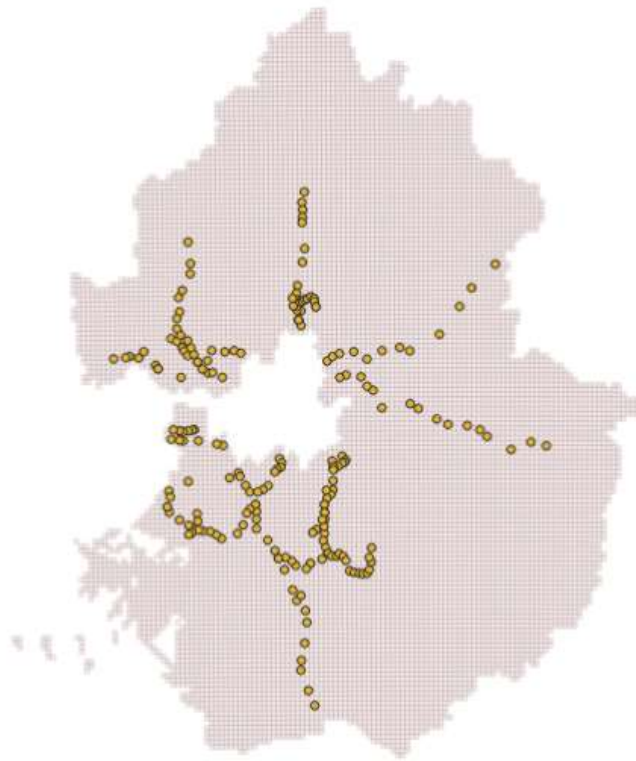
2) 정제 과정

(1) Geocoding Tool64를 통해 데이터의 도로명주소를 바탕으로 위도, 경도를 구함



[그림 7] Geocoding Tool

(2) Q-GIS를 이용하여 경기도 격자별 지도 위에 지하철 좌표 삽입



[그림 8] 서울시 격자별 지하철 좌표

(3) ‘폴리곤에 포함하는 포인트 개수 계산’ 도구를 이용하여 서울시 각 행정동 위에 존재하는 지하철 개수 추출



[그림 9] ‘폴리곤에 포함하는 포인트 개수 계산’ 도구 창

(4) 지하철 개수를 추출할 새로운 열(지하철)이 추가된 ‘경기도 격자’ 생성



경기도 격자 — 총 객체 수: 10907, 필터

	gid	지하철
1	다사3545	3.000000000000...
2	다사6067	3.000000000000...
3	다사3058	3.000000000000...
4	다사6938	3.000000000000...
5	다사4928	2.000000000000...

[그림 10] ‘경기도 격자’ 레이어 속성 테이블

■ 컬럼 병합

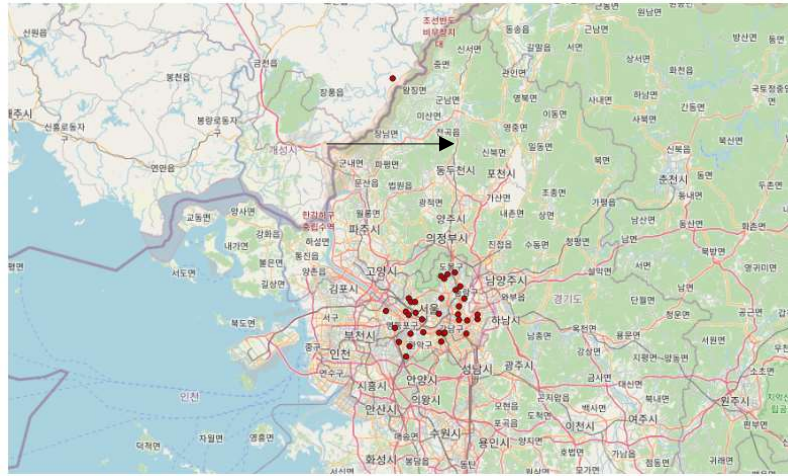
- 격자별 단위로 병합

gid	심_응	약국	소방서	경찰서	안전상비	버스정류소	지하철	어린이집	유아인구수	총인구수
다사6734	0	27	0	1	24	34	1	9	559	18910
다사6424	0	14	0	0	6	27	0	10	401	10680
다사5720	0	11	0	1	11	19	0	7	180	10137
다사5230	0	9	0	2	6	22	0	7	333	10575
다사5130	0	9	0	0	18	23	0	11	407	8364
다사7459	0	9	0	1	4	17	1	8	134	8924
다사3645	0	11	0	0	20	26	0	15	653	18610
다사6530	0	14	0	0	27	18	0	10	773	27656
라사1221	0	13	0	0	9	16	0	8	153	5704
다사6919	0	18	0	1	15	19	1	5	347	11182
다사5319	0	14	0	0	25	27	0	18	676	24082

[그림 11] 격자별 단위 병합 결과

- Q-GIS를 통해 수집 데이터의 위치, 경도 등을 폴리곤 포인트 계산 도구를 통하여 격자 별 변수 개수를 구하고 이상값을 찾음

유치원 데이터 이상치



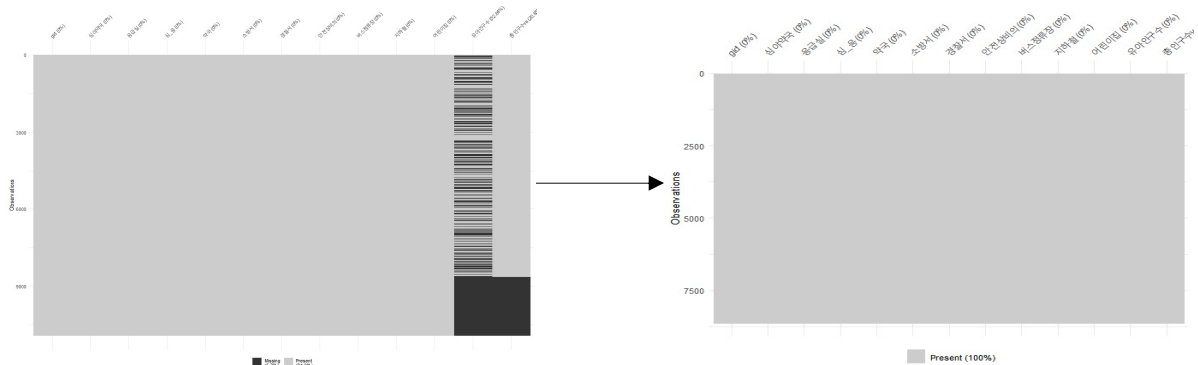
[그림 12] 경기도 유치원 위치 데이터 이상치

■ R 언어를 사용하여 결측치와 이상치를 찾고 정제하는 과정을 거침

- 결측치 처리

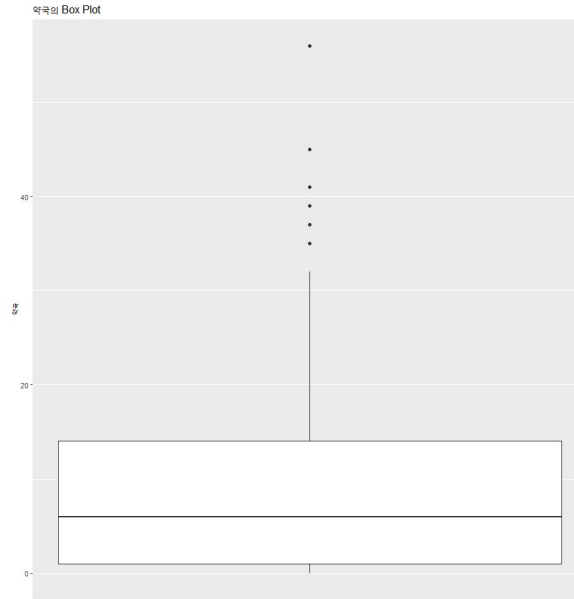
격자별 총인구수 데이터 : 사람이 살기 힘든 곳이 포함된 격자에 NA가 존재하여 입지선정에 적합하지 않다고 판단해 최종 데이터에서 제외함

격자별 유아인구수 데이터 : 유아인구가 살지 않는 격자에 NA가 존재하여 0으로 대체함



- 이상치 처리

각 변수별 이상치가 존재하였으나 검증해보았을 때, 유효하다고 판단해 제거하지 않음



[그림 13] 약국 데이터 boxplot

- 변수 간 상관관계 분석

공공심야약국과 응급실 유무가 약국, 안전상비의약품, 버스 정류소, 총인구수와 어느 정도 상관관계를 갖고 있다는 것을 통해 추후 분류분석에서 중요한 변수가 될 가능성이 있음. 유아인구수와 총인구수의 상관관계가 매우 강한 상관관계를 띄고 있어 다중공선성 문제를 일으킬 가능성이 있다고 판단해 이를 검정해보았지만, 문제가 되지 않아 변수를 제거하지 않고 분석을 진행함

	심_응	약국	소방서	경찰서	안전상비	버스정류소	지하철	어린이집	유아인구수	총인구수
심_응	1.000	0.393	0.061	0.138	0.376	0.297	0.188	0.188	0.177	0.263
약국	0.393	1.000	0.049	0.438	0.837	0.665	0.527	0.548	0.559	0.654
소방서	0.061	0.049	1.000	0.081	0.057	0.097	0.029	0.063	0.059	0.070
경찰서	0.138	0.438	0.081	1.000	0.417	0.436	0.271	0.300	0.286	0.350
안전상비	0.376	0.837	0.057	0.417	1.000	0.663	0.483	0.528	0.552	0.690
버스정류소	0.297	0.665	0.097	0.436	0.663	1.000	0.383	0.636	0.666	0.737
지하철	0.188	0.527	0.029	0.271	0.483	0.383	1.000	0.268	0.298	0.338
어린이집	0.188	0.548	0.063	0.300	0.528	0.636	0.268	1.000	0.843	0.821
유아인구수	0.177	0.559	0.059	0.286	0.552	0.666	0.298	0.843	1.000	0.943
총인구수	0.263	0.654	0.070	0.350	0.690	0.737	0.338	0.821	0.943	1.000

[그림 14] 상관관계 heatmap

3. 분석 프로세스

3.1. 분석 프로세스



■ 데이터 수집 및 처리 : 목적에 맞는 분석을 수행하기 위해 관련된 정보를 조사하고 분석에 필요한 데이터를 수집하고 분석 가능한 형태로 변경 및 정확성을 높이기 위한 전처리

* 원격화상 투약기 입지선정 모델을 위해 필요한 데이터 수집, 수집된 데이터로부터 분석 요소 추출, 데이터 전처리

■ 표준모델 분석: 데이터의 패턴을 파악하고, 현황분석 결과를 기반으로 분류분석 기능을 모델링 하는 단계

* 약국, 총인구수 등 각 데이터별 변수 중요도를 통해 입지분석

■ 시각화 도출: 목적에 맞는 전략을 제공하기 위해 도출해낸 결과를 시각화 하여 분석하는 단계

* 원격화상 투약기 입지순위 시각화

3.2. 분석 내용 및 방법

여러 가지 분류 모델의 성능을 비교하여 화상 투약기 최적 입지를 알려주는 모형을 만들고자 함. 각 분석에서 train set과 test set의 비율은 7:3을 사용하였고 이상치를 제거하지 않은 데이터를 이용하여 분석을 진행함

3.2.1. 평가지표

		예측 클래스	
		Negative	Positive
실제 클래스	Negative	TN (True Negative)	FP (False Positive)
	Positive	FN (False Negative)	TP (True Positive)

[표 17] 혼동행렬

■ 정확도(Accuracy)

▶ 전체 예측에서 옳은 예측의 비율

$$\Rightarrow \frac{TP + TN}{TP + TN + FP + FN}$$

■ 정밀도(Precision)

▶ Positive로 예측된 것 중 실제로 Positive인 경우의 비율

$$\Rightarrow \frac{TP}{TP + FP}$$

■ 재현율(Recall)

▶ Positive 검출율

$$\Rightarrow \frac{TP}{TP + FN}$$

■ F1 score

▶ 정밀도와 재현율의 조화평균

$$\Rightarrow 2 * \frac{Precision * Recall}{Precision + Recall}$$

3.2.2. train-test split

데이터가 Imbalanced data임을 고려해 아래 제시된 방법을 사용하여 train, test data를 생성하여 분석을 진행함

■ 층화 무작위 추출 적용

- ▶ 모집단을 일정한 기준에 따라 2개 이상의 동질적인 층으로 구분하고 각 층별로 단순무작위추출방법을 적용하는 방법

■ SMOTE (Synthetic Minority Oversampling Technique)

- ▶ 소수 클래스에 속하는 데이터 주변에 원본 데이터와 동일하지 않으면서 소수 클래스에 해당하는 가상의 데이터를 생성하여 데이터를 증식시키는 방법

3.2.3. 분류 모델

■ Logistic Regression

- ▶ 일반적인 회귀분석의 목표와 동일하게 종속 변수와 독립 변수 간의 관계를 구체적인 함수로 나타내어 향후 예측 모델에 사용함. 로지스틱 회귀는 알고리즘의 이름은 회귀이지만 실제로는 분류 작업에 사용함
- ▶ 단계 선택법 : 전진 선택법에서 후진 소거법을 추가한 방법임. 전진 선택법에 의해 변수를 추가하면서 새롭게 추가된 변수에 기반하여, 기존 변수의 중요도가 약화되면 해당 변수를 제거한 후 단계별로 추가 또는 제거되는 변수의 여부를 검토하여 더 이상 변경사항이 없을 때까지 진행함

■ LGBM Classifier

- ▶ 다른 알고리즘이 트리를 수평으로 확장하는 반면 이는 트리를 수직으로 확장함. 그렇게 확장하면, 빠른 속도로 학습을 하면서 손실을 최소화할 수 있기에 높은 정확도를 나타내지만 과적합 문제가 발생할 가능성이 높

음. 적은 양의 데이터 학습에서는 과적합되기 쉽기 때문에 일반적으로 10000개 이하의 데이터셋에서 사용되는 것은 추천되지 않음

■ XGB Classifier

- ▶ Gradient Boosting의 단점을 보완하기 위해 나온 알고리즘으로 약한 분류기를 세트로 묶어 정확도를 예측하는 기법. 여러 개의 Decision Tree를 조합해서 사용하는 Ensemble 알고리즘

■ K Neighbors Classifier

- ▶ 지도학습 알고리즘 중 하나로 새로운 데이터를 입력받았을 때, 해당 데이터와 가장 가까이 있는 k개의 데이터를 확인하여 새로운 데이터의 특성을 파악하는 방법

■ Gaussian Naive Bayes

- ▶ 클러스터링 방법 중 하나로 데이터의 군집을 가우시안 모델로 표현하는 기법. 가우시안 모델의 평균과 분산(μ 와 σ)으로부터 군집의 특성을 알 수 있음. 다시 말해서 가우시안 나이브 베이즈는 표본 평균과 표본 분산을 가진 정규분포 하에서 베이즈 정리를 사용한 것임. 정규분포를 가정한 표본들을 대상으로 조건부 독립을 나타내 항상 같은 분모를 갖는 조건에서 분자의 값이 가장 큰 경우, 즉 확률이 가장 높은 경우를 선택하는 것임

■ Decision Tree Classifier

- ▶ 설명변수들의 규칙, 관계, 패턴 등으로 관심 대상인 목표변수를 분류하는 나무 구조의 모델을 만들고, 설명변수의 값을 생성된 모델에 입력하여 목표변수를 분류/예측하는 지도학습 기법. 데이터에 있는 규칙을 학습을 통해 자동으로 찾아내 분류 규칙을 트리 형태로 만드는 것임. 결정 나무는 목표변수에 영향을 주는 설명변수를 탐색하고 해당 설명변수의 최적 분리기준을 제시함

■ Random Forest Classifier

- ▶ 분류, 회귀분석 등에 사용되는 앙상블 학습 방법의 일종으로 훈련 과정에서 구성한 다수의 결정 트리로부터 분류 또는 평균 예측치(회귀분석)를

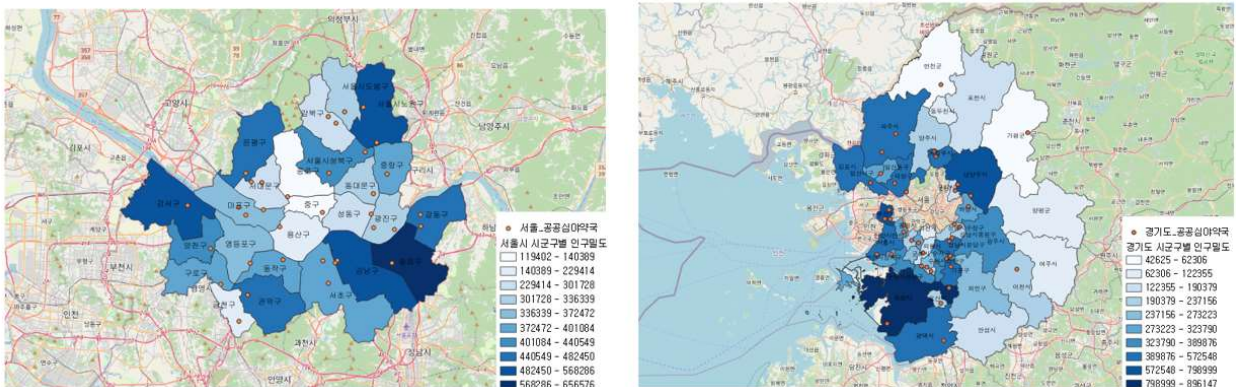
출력함으로써 동작함. 좋은 성능을 얻기 위해 다수의 학습 알고리즘을 사용하는 걸 앙상블 학습법이라고 일컬음

4. 분석결과

4.1. 현황 분석

■ 서울시 공공심야약국 균등한 배치

- ▶ 서울시 공공심야약국은 각 시군구마다 한 개씩 골고루 배치가 되어있지만, 경기도는 지역 차가 큼. 공공심야약국이 없는 경기 지역은 동두천시, 양주시, 포천시, 양평군, 여주시, 처인군, 안성시, 과천시, 광주시 등등이 있음



[그림 15] 서울, 경기 공공심야약국 위치

■ 서울 공공심야약국 입지 재선정

- ▶ 2022년 1월까지 서울 공공심야약국은 40개를 운영했지만, 5개 운영 지역을 축소 하여 현재는 35개가 운영되고 있음. 축소된 5개 운영 약국은 입지상 가시성이 좋지 않은 곳에 위치되어 이용 건수가 저조하고 판매 건수의 불충족으로 인해 운영을 축소했음. 현재는 이용자 수가 많고, 판매 건수가 높은 약국만 공공심야약국을 운영하고 있음

☞ 따라서 서울시 공공심야약국 입지를 통해서 경기도 원격화상 투약기 최적 입지 선정을 진행하기로 했음

4.2. 분류 분석

4.2.1. 로지스틱 회귀분석 (모든 변수)

■ 층화 무작위 추출 적용

Logit Regression Results

Dep. Variable:	심_응	No. Observations:	422
Model:	Logit	Df Residuals:	412
Method:	MLE	Df Model:	9
Date:	Sun, 14 Aug 2022	Pseudo R-squ.:	0.2040
Time:	12:15:38	Log-Likelihood:	-142.95
converged:	True	LL-Null:	-179.59
Covariance Type:	nonrobust	LLR p-value:	3.453e-12

	coef	std err	z	P> z	[0.025	0.975]
const	-3.3985	0.385	-8.839	0.000	-4.152	-2.645
약국	0.1007	0.030	3.306	0.001	0.041	0.160
소방서	1.0596	0.571	1.856	0.063	-0.059	2.178
경찰서	-0.1349	0.242	-0.559	0.576	-0.608	0.338
안전상비	0.0042	0.024	0.175	0.861	-0.043	0.051
버스정류소	0.0195	0.015	1.268	0.205	-0.011	0.050
지하철	-0.1328	0.172	-0.771	0.441	-0.470	0.205
어린이집	0.0587	0.042	1.382	0.167	-0.025	0.142
유아인구수	-0.0022	0.001	-2.338	0.019	-0.004	-0.000
총인구수	4.784e-05	3.53e-05	1.356	0.175	-2.13e-05	0.000

■ SMOTE 적용

Logit Regression Results

Dep. Variable:	심_응	No. Observations:	716			
Model:	Logit	Df Residuals:	706			
Method:	MLE	Df Model:	9			
Date:	Sun, 14 Aug 2022	Pseudo R-squ.:	0.3035			
Time:	02:45:15	Log-Likelihood:	-345.67			
converged:	True	LL-Null:	-496.29			
Covariance Type:	nonrobust	LLR p-value:	1.414e-59			
	coef	std err	z	P> z	[0.025	0.975]
const	-1.8643	0.212	-8.784	0.000	-2.280	-1.448
약국	0.1773	0.024	7.365	0.000	0.130	0.225
소방서	0.0764	0.492	0.155	0.877	-0.889	1.041
경찰서	-0.6412	0.176	-3.635	0.000	-0.987	-0.295
안전상비	0.0264	0.018	1.488	0.137	-0.008	0.061
버스정류소	0.0320	0.011	2.853	0.004	0.010	0.054
지하철	-0.6545	0.126	-5.199	0.000	-0.901	-0.408
어린이집	0.0606	0.033	1.814	0.070	-0.005	0.126
유아인구수	-0.0023	0.001	-3.364	0.001	-0.004	-0.001
총인구수	9.793e-06	2.67e-05	0.367	0.714	-4.25e-05	6.21e-05

분류방법	추출방법	정확도	정밀도	재현율	F1 score
로지스틱 회귀 (모든 변수)	층화 무작위	0.84	0.33	0.11	0.17
	SMOTE	0.74	0.31	0.63	0.41

4.2.2. 로지스틱 회귀분석 (일부 변수)

■ 층화 무작위 추출 적용

Logit Regression Results

Dep. Variable:	심_응	No. Observations:	422
Model:	Logit	Df Residuals:	417
Method:	MLE	Df Model:	4
Date:	Sun, 14 Aug 2022	Pseudo R-squ.:	0.1927
Time:	12:43:30	Log-Likelihood:	-144.99
converged:	True	LL-Null:	-179.59
Covariance Type:	nonrobust	LLR p-value:	3.335e-14

	coef	std err	z	P> z	[0.025	0.975]
const	-3.2023	0.333	-9.613	0.000	-3.855	-2.549
약국	0.1006	0.020	5.076	0.000	0.062	0.139
소방서	1.2118	0.564	2.147	0.032	0.106	2.318
유아인구수	-0.0018	0.001	-2.326	0.020	-0.003	-0.000
총인구수	7.186e-05	2.88e-05	2.495	0.013	1.54e-05	0.000

■ SMOTE 적용

Logit Regression Results

Dep. Variable:	심_응	No. Observations:	716
Model:	Logit	Df Residuals:	708
Method:	MLE	Df Model:	7
Date:	Sun, 14 Aug 2022	Pseudo R-squ.:	0.3033
Time:	02:49:34	Log-Likelihood:	-345.74
converged:	True	LL-Null:	-496.29
Covariance Type:	nonrobust	LLR p-value:	3.528e-61

	coef	std err	z	P> z	[0.025	0.975]
const	-1.8569	0.211	-8.807	0.000	-2.270	-1.444
약국	0.1776	0.024	7.371	0.000	0.130	0.225
경찰서	-0.6414	0.176	-3.635	0.000	-0.987	-0.296
안전상비	0.0294	0.015	1.907	0.056	-0.001	0.060
버스정류소	0.0324	0.011	2.907	0.004	0.011	0.054
지하철	-0.6598	0.124	-5.306	0.000	-0.904	-0.416
어린이집	0.0637	0.032	1.962	0.050	6e-05	0.127
유아인구수	-0.0021	0.001	-4.158	0.000	-0.003	-0.001

분류 방법	추출 방법	정확도	정밀도	재현율	F1 score
로지스틱 회귀 (일부 변수)	층화 무작위	0.85	0.50	0.11	0.18
	SMOTE	0.79	0.37	0.63	0.47

4.2.3. LGBM Classifier

분류방법	추출방법	정확도	정밀도	재현율	F1 score
LGBM Classifier	층화 무작위	0.86	0.58	0.26	0.36
	SMOTE	0.81	0.36	0.33	0.35

4.2.4. XGB Classifier

분류방법	추출방법	정확도	정밀도	재현율	F1 score
XGB Classifier	층화 무작위	0.86	0.57	0.30	0.39
	SMOTE	0.82	0.39	0.41	0.40

4.2.5. K Neighbors Classifier

분류방법	추출방법	정확도	정밀도	재현율	F1 score
K Neighbors Classifier	층화 무작위	0.79	0.26	0.22	0.24
	SMOTE	0.64	0.20	0.48	0.28

4.2.6. Gaussian Naive Bayes

분류방법	추출방법	정확도	정밀도	재현율	F1 score
Gaussian Naive Bayes	층화 무작위	0.76	0.34	0.63	0.44
	SMOTE	0.66	0.28	0.78	0.41

4.2.7. Decision Tree Classifier

분류방법	추출방법	정확도	정밀도	재현율	F1 score
Decision Tree Classifier	층화 무작위	0.80	0.29	0.22	0.25
	SMOTE	0.79	0.34	0.48	0.40

4.2.8. Random Forest Classifier

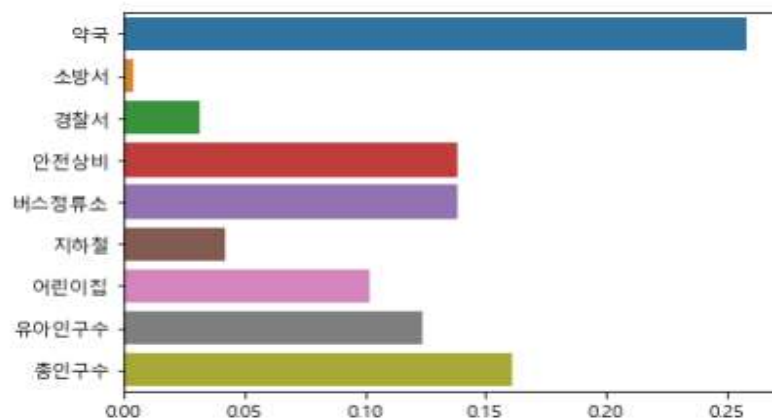
분류방법	추출방법	정확도	정밀도	재현율	F1 score
Random Forest Classifier	층화 무작위	0.85	0.50	0.15	0.23
	SMOTE	0.84	0.45	0.52	0.48

4.3. 최종 모델

정확도와 F1 score를 모두 고려해 봤을 때, SMOTE 방법을 사용해 생성된 train data를 통해 훈련된 Random Forest Classifier로 test data를 분류했을 때 가장 좋은 성능을 보여줌. 변수의 중요도를 보았을 때, 약국과 총인구수가 공공심야약국과 응급실의 여부에 영향을 가장 많이 준다는 사실을 알 수 있음.

```
Feature importances:
[0.259 0.004 0.031 0.138 0.139 0.042 0.102 0.124 0.161]
약국 : 0.259
소방서 : 0.004
경찰서 : 0.031
안전상비 : 0.138
버스정류소 : 0.139
지하철 : 0.042
어린이집 : 0.102
유아인구수 : 0.124
총인구수 : 0.161
```

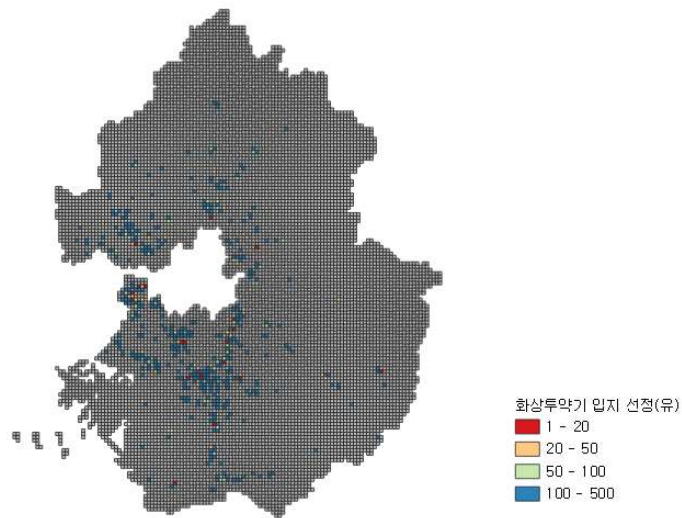
<AxesSubplot :>



4.4. 입지 순위 결과

4.4.1. 화상 투약기 입지선정 순위

■ (1위~500위) 예측 결과 시각화



■ (1위~50위) 예측 결과

순 위	약 국	소 방 서	경 찰 서	안전 상비	버스 정류장	지 하 철	어린 이집	유아 인구수	총 인구수	최종 점수	심야약국과 응급실 여부	장소 설명
1	27	0	1	24	34	1	9	559	18910	0.93	X	진흥더블파크아파트
2	14	0	0	6	27	1	10	401	10680	0.89	X	수지구청역롯데캐슬하이브엘아파트
3	20	0	1	20	24	1	6	128	7690	0.89	O (1)	광장약국, 빌라
4	11	0	1	11	19	0	7	180	10137	0.88	X	빌라
5	20	0	1	53	32	0	12	682	26324	0.87	O (2)	뿌리약국, 빌라
6	9	0	2	6	22	0	7	333	10575	0.87	X	샘마을대우한양아파트
7	9	0	0	18	23	0	11	407	8364	0.85	X	평촌어바인퍼스트아파트
8	23	0	1	29	28	0	24	821	32421	0.84	O (1)	단골약국, 빌라
9	19	0	0	17	20	0	8	237	9819	0.82	O (1)	수원센트럴아이파크자이아파트, 수원제일큰약국
10	9	0	1	4	17	1	8	134	8924	0.81	X	서울아파트
11	11	0	0	20	26	0	15	653	18610	0.80	X	중동대림아파트
12	20	1	0	44	30	2	16	570	30465	0.79	O (2)	성모월병원, 빌라
13	14	0	0	27	18	0	10	773	27656	0.79	X	양지마을1단지한양아파트
14	13	0	0	9	16	0	8	153	5704	0.78	X	현대아파트
15	18	0	1	15	19	0	5	347	11182	0.78	X	성산마을아파트

순 위	약 국	소 방 서	경 찰 서	안전 상비	버스 정류장	지 하 철	어린 이집	유아 인구수	총 인구수	최종 점수	심야약국과 응급실 여부	장소 설명
16	14	0	0	25	27	0	18	676	24082	0.78	X	삼성아파트
17	19	0	0	29	26	1	2	705	20555	0.78	O (1)	백석역, 일산요진와이시티아파트
18	5	0	0	9	24	0	8	296	12470	0.78	X	신안파크아파트
19	4	0	0	9	15	0	10	338	11369	0.77	X	호원우성아파트
20	10	0	0	12	21	0	5	249	5223	0.77	O (1)	성심중앙병원, 빌라
21	16	0	0	19	26	0	18	513	19350	0.77	O (1)	풍림아파트
22	9	0	1	19	12	0	10	268	11007	0.76	O (1)	두산아파트
23	19	0	1	27	35	0	13	372	17906	0.75	O (2)	경기도의료원의정부병원, 빌라
24	23	0	0	17	46	1	20	717	22627	0.74	O (1)	역곡e편한세상아파트, 새현대약국
25	14	0	0	19	28	0	19	807	17101	0.73	X	평촌자이아이파크아파트
26	16	0	1	23	36	1	13	524	20549	0.72	X	금정역, 매화주공아파트
27	17	0	0	19	20	0	16	823	20024	0.72	O (1)	백현마을휴먼시아아파트
28	15	0	2	13	14	1	6	175	7655	0.71	X	양평역, 양평현대아파트
29	20	0	1	23	39	0	31	907	34666	0.71	O (2)	은행현대아파트
30	12	0	0	29	27	0	43	869	27924	0.71	X	한양아파트
31	5	0	0	6	21	0	17	275	11649	0.71	X	한산그린아파트
32	17	0	1	20	12	0	7	291	9775	0.71	X	송탄역서희스타필드아파 트
33	19	0	0	22	20	0	12	241	11609	0.71	O (1)	장안아파트
34	25	0	1	22	38	0	7	136	7678	0.70	O (1)	광주역금호리첸시아아파 트
35	10	0	0	6	13	0	3	294	4937	0.70	X	신원마을LH3단지아파트
36	21	0	1	22	26	0	17	844	27928	0.69	X	한일아파트
37	14	0	0	23	26	0	31	1216	23306	0.69	X	산들마을길훈신안현대아 파트
38	8	0	1	11	15	0	11	283	10118	0.69	X	마송현대아파트
39	17	0	0	24	61	0	13	620	23208	0.69	X	원당e편한세상아파트
40	23	0	0	20	22	0	1	363	7064	0.69	O (1)	이천코아루휴티스아파트
41	30	0	0	19	26	0	2	89	5492	0.68	O (1)	안성성모병원, 빌라
42	19	0	1	25	35	0	13	271	9769	0.68	O (1)	금강아파트
43	6	0	1	6	17	0	11	262	9912	0.68	X	흰돌마을아파트
44	19	0	0	21	18	0	15	658	21606	0.68	X	빌라
45	9	0	0	16	16	0	14	388	16119	0.68	X	권선데시앙아파트
46	11	0	1	7	13	0	8	274	4638	0.68	X	화성남양시티프라디움아 파트
47	16	1	1	29	40	0	24	1294	30704	0.67	X	꿈마을아파트
48	12	0	0	17	28	0	21	799	19381	0.67	X	대한대우아파트
49	25	0	1	29	36	0	27	869	22696	0.67	X	한라마을주공아파트
50	11	0	1	11	11	1	6	233	5437	0.65	X	은동마을주공아파트

■ 1위 비타민약국 (성남시) : 버스터미널, 지하철, 아파트단지 최종점수 0.93

순 위	약 국	소 방 서	경 찰 서	안전 상비	버스 정류장	지 하 철	어린 이집	유아 인구수	총 인구수	최종 점수	심야약국과 응급실 여부	장소 설명
1	27	0	1	24	34	1	9	559	18910	0.93	X	진흥더블파크아파트



[그림 16] 비타민 약국

■ 2위 삼성메디컬약국 (용인시) : 지하철, 학교, 아파트단지 최종점수 0.89

순위	약국	소방서	경찰서	안전상비	버스정류장	지하철	어린이집	유아인구수	총인구수	최종점수	심야약국과응급실여부	장소설명
2	14	0	0	6	27	1	10	401	10680	0.89	X	수지구청역롯데캐슬하이브엘아파트



[그림 17] 삼성메디칼 약국

■ 3위 광장약국 (용인시) : 지하철, 학교 최종점수 0.89 (공공심야약국 존재)

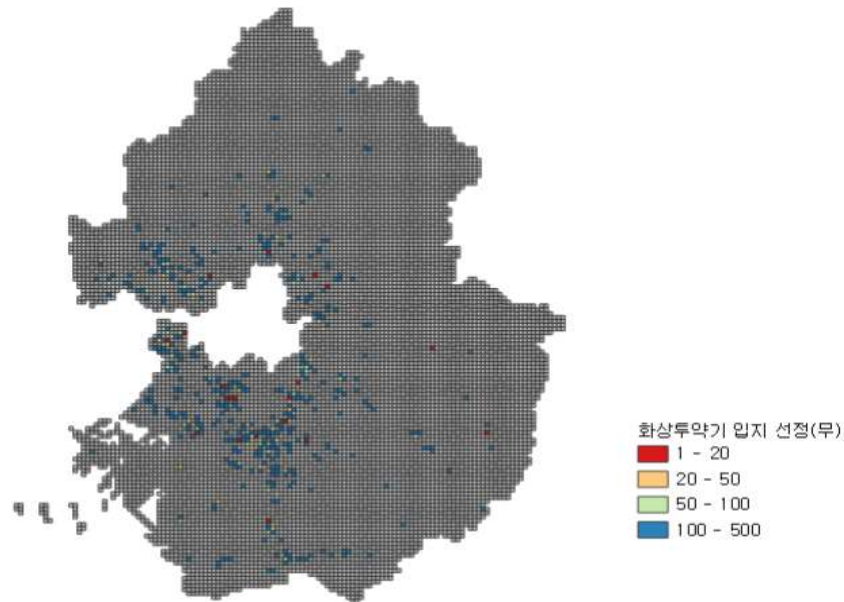
순 위	약 국	소 방 서	경 찰 서	안전 상비	버스 정류장	지 하 철	어린 이집	유아 인구수	총 인구수	최종 점수	심야약국과 응급실 여부	장소 설명
3	20	0	1	20	24	1	6	128	7690	0.89	O (1)	광장약국, 빌라



[그림 18] 광장 약국

4.4.2. [설치완료지역 외] 화상 투약기 입지선정

■ (1위~500위)



■ (1위~50위) 예측 결과

순 위	약 국	소 방 서	경 찰 서	안전 상비	버스 정류장	지 하 철	어린 이집	유아 인구수	총 인구수	최종 점수	장소 설명
1	27	0	1	24	34	1	9	559	18910	0.93	진흥더블파크아파트
2	14	0	0	6	27	1	10	401	10680	0.89	수지구청역롯데캐슬하이브엘아파트
3	11	0	1	11	19	0	7	180	10137	0.88	빌라
4	9	0	2	6	22	0	7	333	10575	0.87	샘마을대우한양아파트
5	9	0	0	18	23	0	11	407	8364	0.85	평촌어바인퍼스트아파트
6	9	0	1	4	17	1	8	134	8924	0.81	서울아파트
7	11	0	0	20	26	0	15	653	18610	0.8	중동대림아파트
8	14	0	0	27	18	0	10	773	27656	0.79	양지마을1단지한양아파트
9	13	0	0	9	16	0	8	153	5704	0.78	현대아파트
10	18	0	1	15	19	1	5	347	11182	0.78	성산마을아파트
11	14	0	0	25	27	0	18	676	24082	0.78	삼성아파트
12	5	0	0	9	24	0	8	296	12470	0.78	신안파크아파트
13	4	0	0	9	15	0	10	338	11369	0.77	호원우성아파트
14	14	0	0	19	28	0	19	807	17101	0.73	평촌자이아이파크아파트
15	16	0	1	23	36	0	13	524	20549	0.72	금정역, 매화주공아파트

순 위	약 국	소 방 서	경 찰 서	안전 상비	버스 정류장	지 하 철	어린 이집	유아 인구수	총 인구수	최종 점수	장소 설명
16	15	0	2	13	14	1	6	175	7655	0.71	양평역, 양평현대아파트
17	12	0	0	29	27	0	43	869	27924	0.71	한양아파트
18	5	0	0	6	21	0	17	275	11649	0.71	한산그린아파트
19	17	0	1	20	12	0	7	291	9775	0.71	송탄역서희스타필드아파트
20	10	0	0	6	13	0	3	294	4937	0.7	신원마을LH3단지아파트
21	21	0	1	22	26	0	17	844	27928	0.69	한일아파트
22	14	0	0	23	26	0	31	1216	23306	0.69	산들마을길혼신안현대아파트
23	8	0	1	11	15	0	11	283	10118	0.69	마송현대아파트
24	17	0	0	24	61	0	13	620	23208	0.69	원당e편한세상아파트
25	6	0	1	6	17	0	11	262	9912	0.68	흰돌마을아파트
26	19	0	0	21	18	0	15	658	21606	0.68	빌라
27	9	0	0	16	16	0	14	388	16119	0.68	권선데시앙아파트
28	11	0	1	7	13	0	8	274	4638	0.68	화성남양시티프라디움아파트
29	16	1	1	29	40	0	24	1294	30704	0.67	꿈마을아파트
30	12	0	0	17	28	0	21	799	19381	0.67	대한대우아파트
31	25	0	1	29	36	0	27	869	22696	0.67	한라마을주공아파트
32	11	0	1	11	11	1	6	233	5437	0.65	덕정역, 은동마을주공아파트
33	16	0	1	15	32	1	5	195	7837	0.65	동두천중앙역, 빌라
34	13	0	0	26	30	0	30	1098	27127	0.64	느치미마을주공프란체아파트
35	20	0	0	15	34	0	21	864	23403	0.64	한솔리치빌아파트
36	10	0	0	20	23	1	13	205	8757	0.64	한대앞역, 늘푸른아파트
37	11	1	0	22	31	0	17	890	17433	0.64	향남시범한우물휴먼시아아파트
38	8	0	0	16	21	0	11	322	11374	0.64	원일산호아파트
39	26	0	0	21	30	1	6	619	16296	0.64	중앙역, 안산센트럴푸르지오아파트
40	10	0	0	16	24	0	16	346	15966	0.63	현대아파트
41	19	0	0	16	24	0	15	452	15286	0.63	미금역, 까치마을롯데선경아파트
42	9	0	0	5	14	0	11	640	16293	0.63	성동마을LG빌리지아파트
43	9	1	0	19	37	0	17	835	19110	0.62	공작부영아파트
44	3	0	0	9	16	0	7	327	10481	0.62	부영아파트
45	22	1	0	44	22	1	1	166	12866	0.62	정발산역, 양지마을건영1단지
46	15	0	0	19	20	0	14	329	19452	0.62	빌라
47	19	0	1	27	37	2	16	986	25922	0.62	정자역, 상록마을우성아파트
48	14	0	1	24	27	0	36	923	30997	0.61	미리내마을은하수타운아파트
49	14	0	1	14	24	0	7	527	12776	0.61	죽현마을아이파크아파트
50	13	0	0	7	31	1	9	246	6441	0.61	문산역, 두산위브아파트

■ 1위 비타민약국 (성남시) : 버스터미널, 지하철, 아파트단지 최종점수 0.93

순위	약국	소방서	경찰서	안전상비	버스정류장	지하철	어린이집	유아인구수	총인구수	최종점수	장소 설명
1	27	0	1	24	34	1	9	559	18910	0.93	진흥더블파크아파트



[그림 19] 비타민 약국

■ 2위 삼성메디컬약국 (용인시) : 지하철, 학교, 아파트단지 최종점수 0.89

순 위	약 국	소 방 서	경 찰 서	안전 상비	버스 정류장	지 하 철	어린 이집	유아 인구수	총 인구수	최종 점수	장소 설명
2	14	0	0	6	27	1	10	401	10680	0.89	수지구청역롯데캐슬하이브엘아파트



[그림 20] 삼성메디컬 약국

■ 3위 신통약국 (수원시) : 파출소, 학교, 주택 입지 최종점수 0.88

순위	약국	소방서	경찰서	안전상비	버스정류장	지하철	어린이집	유아인구수	총인구수	최종점수	장소 설명
3	11	0	1	11	19	0	7	180	10137	0.88	빌라



[그림 21] 신통 약국

5. 활용 방안

5.1. 문제점 개선 방안

- (응급실 본연 기능 상실 방지) 심야시간대 경증 환자의 응급실 이용률을 줄임으로써 중증 환자의 치료에 집중
- (예산 효율화) 공공심야약국의 운영예산을 전환해 원격화상 투약기의 신속한 설치 및 효과적인 예산 활용
- (불편 해소) 심야시간대 경증 환자의 비용 부담을 감소하고 의료시설의 이른 운영종료시간으로 인한 불편 해소
- (오남용 방지) 이용자가 의약품을 선택하지 못하도록 외부에 의약품 선택기능을 제거하여 의약품 오남용 방지

5.2. 업무 활용 방안

- (표준화 모델) 원격화상 투약기 시범운영 계획에 대하여 최적 입지 제안
- (정책적 효율성) 인구 특성, 교통접근성, 위치 적합성을 고려해 원격화상 투약기의 최적 후보지를 제안함으로써 입지 선정 시 현실성과 효율성 극대화
- (정책 발전) 현재 약사법의 한계가 존재하며 해외 사례를 모델로 발전 될 것으로 기대

5.3. 한계점

5.3.1 분석 한계점

- 데이터 변수의 다양성 부족으로 인해, 분석 모델의 정확성이 낮아짐
- ☞ 각 응급실별 이용량과 경증 환자의 이용량, 건물 데이터 등 필요함

5.3.2 원격화상 투약기 한계점과 개선방안

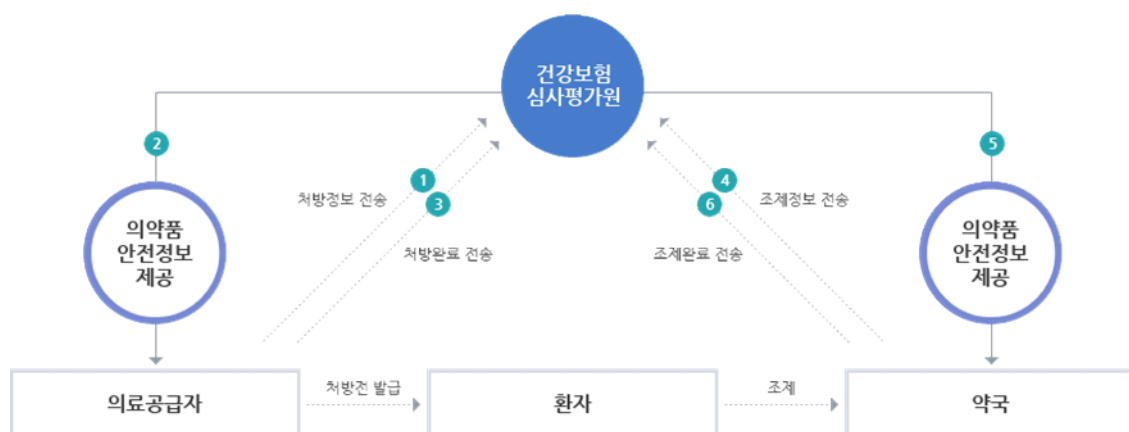
- 화상상담으로 진행되지만, 대면 상담을 하지 않는 점에서 안정성, 전문성 우려

▶ 개선방안

- 건강보험심사평가원의 DUR 서비스망을 활용한 표준적인 공적 전자처방전 전달 시스템 구축

☞ DUR 서비스 절차

의사가 처방단계에서 환자의 처방정보를 건강보험심사평가원으로 전송함. 건강보험심사평가원은 환자의 투약이력 및 DUR 기준과 비교해서 문제되는 의약품이 있으면 의사의 컴퓨터 화면에 0.5초 이내로 경고 메시지를 띄움. 의사는 처방을 변경하거나 임상적 필요에 의해 부득이하게 처방시에는 예외사유를 기재하여 처방을 완료하고 그 정보는 건강보험심사평가원에 전송. 약사도 동일한 과정을 거치게 되며, 경고 메시지가 있는 의약품에 대해 처방의사에게 변경여부를 물어 변경에 동의하는 경우 변경하여 조제할 수 있음, 조제 완료한 내역은 건강보험심사평가원에 전송함



[그림 22] DUR 서비스 절차

- 노년층의 이용 어려움

▶ 개선방안

- 기계 사용이 어려운 노년층들을 위해 원격화상 투약기 화면 상단에 ‘사용법’에 관한 동영상을 만들어 원격화상 투약기 활용 교육을 시행함.

* 키오스크 활용을 어려워 하는 노인들을 위해 창원시설공단 성산노인복지관에서 키오스크 활용 교육을 실행하였음.

■ 국내 법 조항으로 인한, 한정적 설치 장소

약국 앞에만 설치할 수 있는 장소적 한계로 인해, 다양한 지역에 의료 서비스를 제공하지 못함

▶ 개선방안

- 기존의 법을 개선하여 대학교, 공원 근처 등으로 장소를 확대하여 접근 용이성을 높임

6. 참고자료(Reference)

■ 학술 논문

윤희영, 「비대면 시대 원격의료의 정착을 위한 법제도 개선 방안」, 『법학논총』 vol.45. no.1, 단국대학교 법학연구소, 2021, pp. 449-498

유주선, 「원격 화상투약기의 전망과 법적 과제」, 『한국의료법학회지』 vol.22.no2, 한국의료법학회, 2014, pp. 93-111

고대균, 여정성, 「안전상비의약품 약국 외 판매에 관한 소비자 평가」 『소비자정책교육연구』, 제13권 제2호, 한국소비자정책교육학회, 2017, pp. 51-78

장후석, 「일반의약품 약국외 판매의 경제적 효과」, 『이슈리포트』, 2011권 15호, 현대경제연구원, 2011, pp.1-11

■ 기사

이금숙, "아직도... 응급실 환자 절반이 '경증'", 헬스조선, 2020.10.30.,

https://m.health.chosun.com/svc/news_view.html?contid=2020103002051

신현정, "공공심야약국 '지원 빈약'... 경기도내 21개 시군은 아예 없다", 경인일보, 2022.02.11.,

<http://www.kyeongin.com/main/view.php?key=20220210010001851>

황재용, "안전상비의약품, 구입 편한 대신 값은 두배?", 청년의사, 2012.11.18.,

<http://www.docdocdoc.co.kr/news/articleView.html?idxno=123939>

김지은, "코로나에 해외도 비대면진료가 대세...약국 지형도는?", 데일리팜, 2022.06.04.,

<http://www.dailypharm.com/Users/News/NewsView.html?ID=288727>

이상우, "달힌 문을 스스로 여는 '화상투약기'", 약국신문, 2021.08.11.,

<http://www.pharm21.com/news/articleView.html?idxno=113674>

이혜선, "약사회가 밝힌 화상투약기를 도입해선 안되는 이유들", 청년의사, 2016.05.27.,

<http://www.docdocdoc.co.kr/news/articleView.html?idxno=203586>

박민식, "전화기 너머 환자 잡아라"...코로나 팬데믹에서 국내 비대면 진료 플랫폼 20여곳으로 늘어", medigatenews, 2022.03.06.,

<https://medigatenews.com/news/578148095>

김명석, "한번에 2개 결제 불가능 한개씩 2번 결제는 가능" 편의점 상비약 오남용 우려",충청투데이, 2017.02.02.,

<https://www.cctoday.co.kr/news/articleView.html?idxno=1036388>

김민준, "허용 확대 요구 높아질 원격의료, 의료전달체계 개선 차원에서 이뤄져야", 메디컬투데이, 2022.05.04.,

<https://mdtoday.co.kr/news/view/1065587311478776>

강신국, "독일 약업계 신기술 경연장엔 '이런 신기한 것도'", 데일리팜, 2015.10.06.,

<http://www.dailypharm.com/Users/News/NewsView.html?ID=203572>

정홍준, "약사회 Vs 쓰리알코리아, 유튜브서 화상투약기 설전", 데일리팜,

2022.06.29.,<http://www.dailypharm.com/Users/News/NewsView.html?ID=289450>

남영진, "약사와 화상상담으로 약을 살 수 있다?", 시민일보, 2015.09.14.,

<http://www.siminilbo.co.kr/news/articleView.html?idxno=413151>

이지은, "경기도, 약사 인력난 속 '공공심야약국' 확충 난항", 중부일보, 2022.01.18.,

<http://www.joongboo.com/news/articleView.html?idxno=363522418>

이상배, "드론 택배·위치정보…朴대통령, 신산업 규제 깨부순다", 머니투데이, 2016.05.10.,

<https://news.mt.co.kr/mtview.php?no=2016051017047680405>

강혜경, "10년만에 빗장 풀린 화상투약기…무너진 대면투약 원칙", 데일리팜, 2022.06.21.,

<http://www.dailypharm.com/Users/News/NewsView.html?ID=289184>

"비응급환자 응급실 이용 땀 '의료비 폭탄'", 매일경제, 2015.12.31,

<https://www.mk.co.kr/news/it/view/2015/12/1229007>

정홍준, "환자 10명 중 9명 '공공심야약국' 필요성 공감", 데일리팜, 2019.12.31,

<http://www.dailypharm.com/Users/News/NewsView.html?ID=260513>

이혜경, "국민 86% 편의점약 확대 찬성…"지사제·제산제 원해"", 데일리팜, 2018.08.09.,

<http://www.dailypharm.com/Users/News/NewsView.html?ID=242643>

임도이, "응급실 찾은 경증환자 진료비 '깜놀'", 2015.05.01,

<http://www.dttoday.com/news/articleView.html?idxno=61935>

■ 데이터 활용 소스

KOSIS

https://kosis.kr/statHtml/statHtml.do?orgId=411&tblId=DT_41104_121&vw_cd=MT_ZTITLE&list_id=D1_41101_01&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE

공공데이터포털

<https://www.data.go.kr/data/15085014/fileData.do?recommendDataYn=Y>

■ 웹 페이지

<https://www.hira.or.kr/dummy.do?pgmid=HIRAA990001000330>

<https://www.pharmaself24.co.uk/>

[부록]

1. 사례조사 정리

해외 사례

- 해외의 경우 화상투약기보다 소비자들의 편의성을 높인 자동판매기 운영중 현재 미국, 독일, 스웨덴 등 선진 각국에서는 최첨단의 원격적인 의약품 판매시스템을 통해 일반의약품은 판매 가능하며, 전문의 약품에 대해서도 판매할 수 있도록 하는 문제가 논의되고 있음
- 미국 최대 약국 체인인 CVS 파머시는 ‘셀프 혁신’ 차원에서 일반의약품과 건강보조 제품 수십 가지가 들어가 있는 약 자판기를 도입. 주 별로 의사 처방약 공급·판매기, 낙태·피임약 판매기 등이 다양하게 허용돼 있어 공항이나 대학 캠퍼스 등 공공장소에서 쉽게 찾을 수 있음
- 스웨덴은 2011년 일반의약품에 한해 자판기 판매를 허용하고, 약사가 아니어도 기기를 운영할 수 있도록 함. 자판기가 음성 안내로 각 약품에 대해 설명해 주기 때문에 소비자들은 신용카드와 주민번호만 입력하면 약사와 영상 통화를 하지 않고도 약을 살 수 있음
- 독일·영국 등에선 터치스크린 방식 판매기가 존재함. 환자가 스크린 속 의약품을 클릭하면 약값과 복약 정보 등이 화면에 뜨고, 약사와 전화 상담 후 로봇이 조제해주는 약을 받아갈 수 있는 ‘24시간 의약품 조제 서비스’. 일반약이 아닌 의사 처방이 필요한 전문의약품도 정보통신(IT) 기기와 터치 스크린을 통해 복약지도 정보를 환자에게 제공함. 독일은 처방약 우편배달, 의약품 자판기 등이 합법화 됨

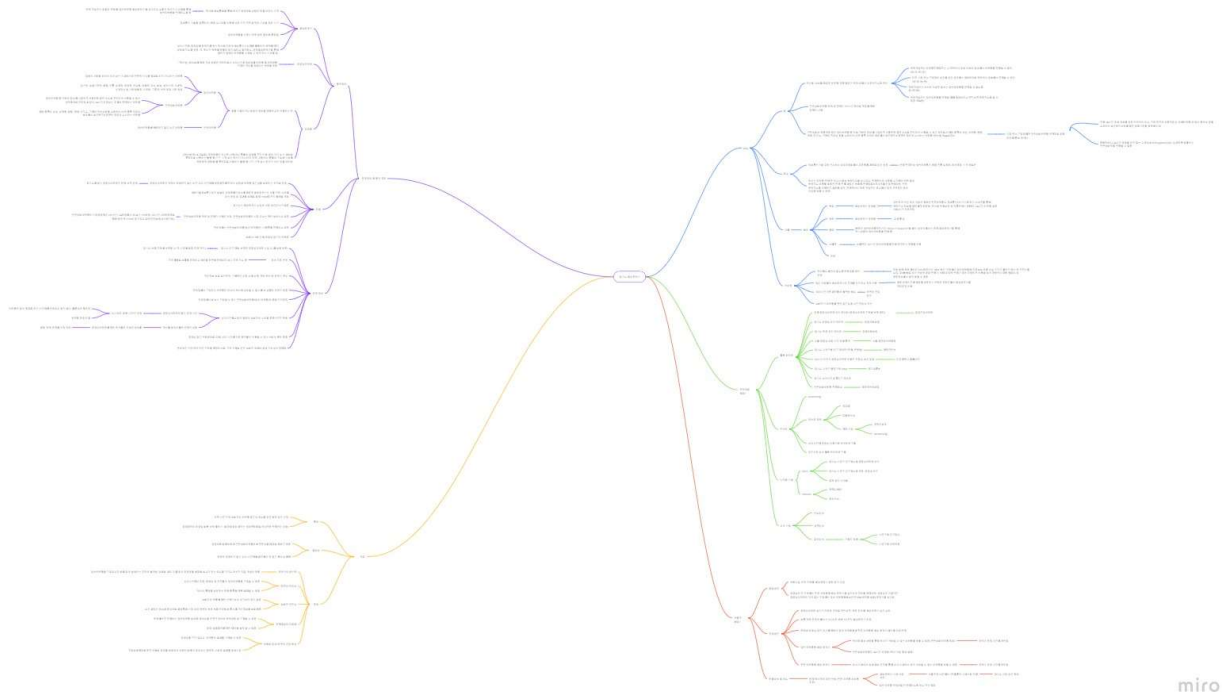
독일 약업계 신기술 경연장엔 '이런 신기한 것도'

해당 의약품을 터치하면 가격, 복약지도 정보 등이 화면에 표시돼 약사가 상담을 주도할 수 있다. 이를 멀티 터치 카운터 디스플레이라고 소개했다.



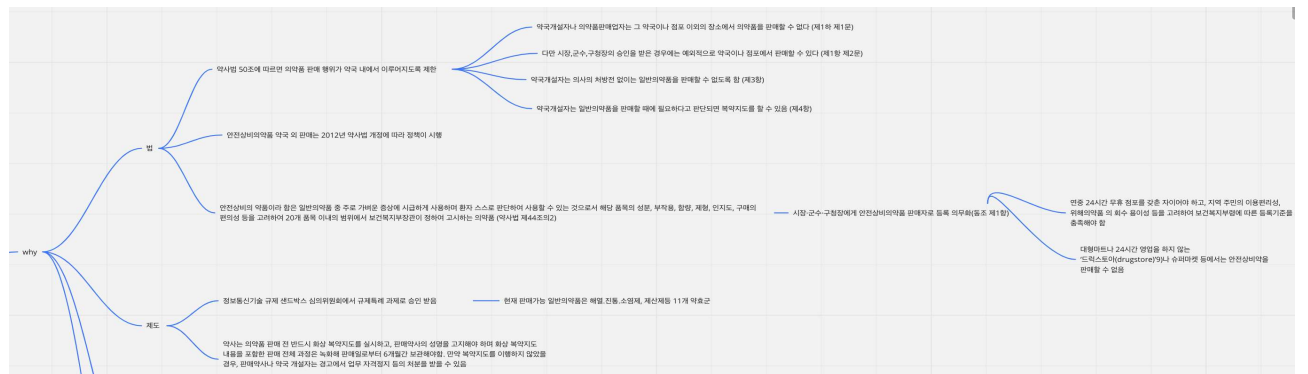
[그림 23] 독일에 있는 조제로봇 (출처 : 데일리팝)

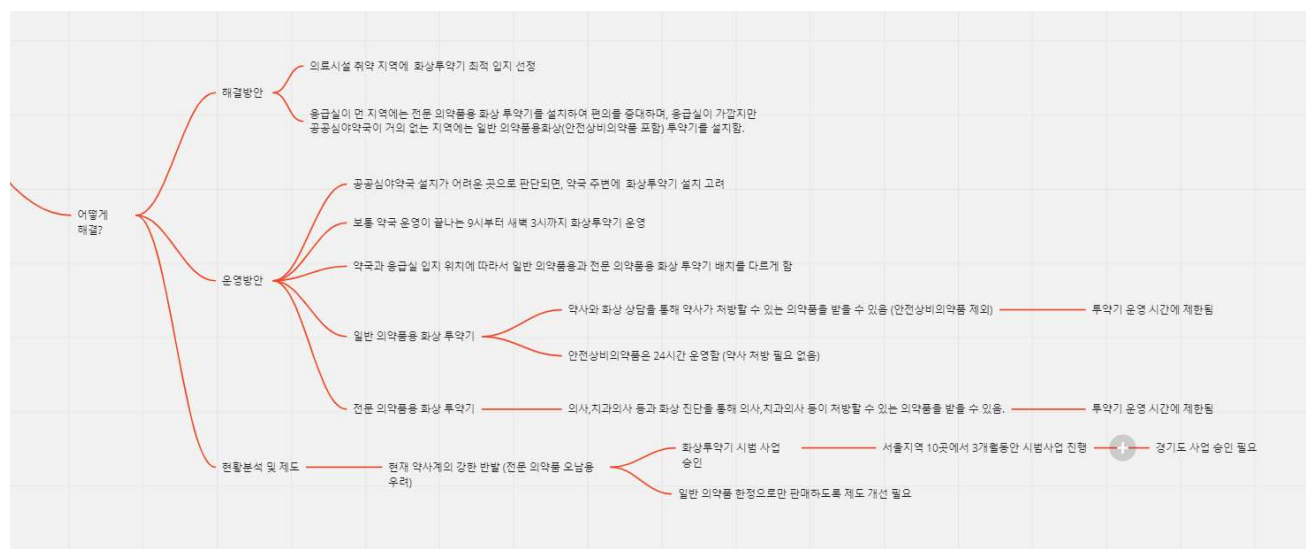
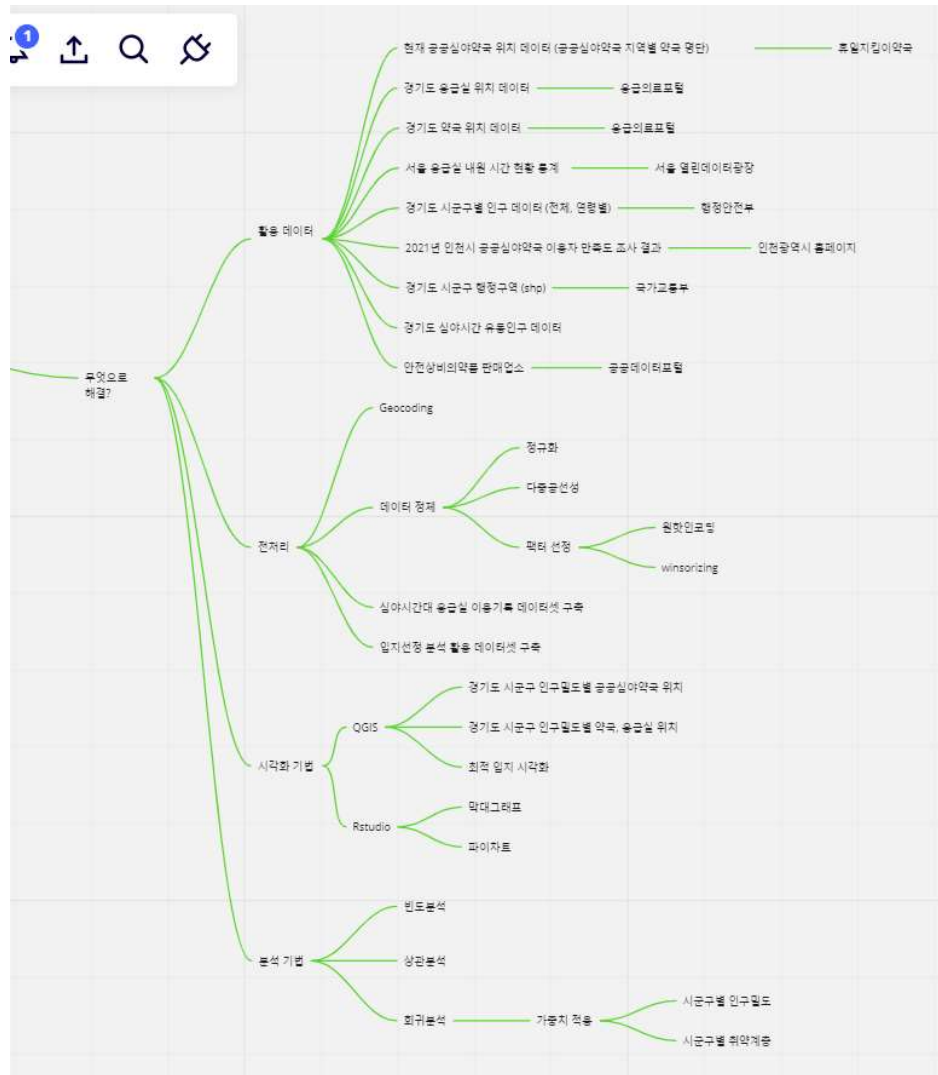
2. 주제설계를 위한 마인드맵



miro







3. 분석 상세코드

```
import numpy as np

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

%matplotlib inline

import warnings

warnings.filterwarnings('ignore')

Seoul = pd.read_csv('C:/Users/user/Desktop/서울/서울데이터_정제.csv', encoding='cp949')

Seoul

Seoul1=Seoul.drop(['심야약국', '응급실'],axis=1)

X = Seoul1.drop(['심_응'], axis=1)

y = Seoul1['심_응']

y.value_counts()

from sklearn.model_selection import train_test_split

X_train_s,X_test_s,y_train_s,y_test_s = train_test_split(X,y,test_size=0.3,random_state=2)

# random_state는 seed값

y_train_s.value_counts()
```

```
y_test_s.value_counts()
```

```
from sklearn.metrics import accuracy_score
```

```
from sklearn.metrics import precision_score
```

```
from sklearn.metrics import recall_score
```

```
from sklearn.metrics import f1_score
```

```
from sklearn.metrics import roc_auc_score
```

```
#모델링
```

```
def modeling(model,x_train,x_test,y_train,y_test):
```

```
    model.fit(x_train,y_train)
```

```
    pred = model.predict(x_test)
```

```
    metrics(y_test,pred)
```

```
#평가 지표
```

```
def metrics(y_test,pred):
```

```
    accuracy = accuracy_score(y_test,pred)
```

```
    precision = precision_score(y_test,pred)
```

```
    recall = recall_score(y_test,pred)
```

```
    f1 = f1_score(y_test,pred)
```

```
    print('정확도 : {0:.2f}, 정밀도 : {1:.2f}, 재현율 : {2:.2f}'.format(accuracy,precision,recall))
```

```
    print('f1-score : {0:.2f}'.format(f1))
```

```
# Logistic Regression : 모든 변수
```

```
from sklearn.linear_model import LogisticRegression

lr = LogisticRegression()

modeling(lr, X_train_s.iloc[:,1:], X_test_s.iloc[:,1:], y_train_s, y_test_s)

# 모두 0이라고 했을 때 : 로지스틱이 정확도가 더 떨어짐

pred = np.zeros(182)

metrics(y_test_s, pred)

# 층화추출

from sklearn.model_selection import train_test_split

X_train_c, X_test_c, y_train_c, y_test_c = train_test_split(X, y, test_size=0.3, random_state=20,
stratify=y)

import statsmodels.api as sm

X_train_cc = sm.add_constant(X_train_c.iloc[:,1:])

X_test_cc = sm.add_constant(X_test_c.iloc[:,1:])

# Logistic Regression : 모든 변수

lr = LogisticRegression()

modeling(lr, X_train_cc, X_test_cc, y_train_c, y_test_c)

model = sm.Logit(y_train_c, X_train_cc)

result = model.fit()

result.summary()
```

```
# Logistic Regression : 일부 변수
```

```
lr = LogisticRegression()
```

```
modeling(lr, X_train_cc.iloc[:,[0,1,2,8,9]], X_test_cc.iloc[:,[0,1,2,8,9]], y_train_c, y_test_c)
```

```
model = sm.Logit(y_train_c, X_train_cc.iloc[:,[0,1,2,8,9]])
```

```
result = model.fit()
```

```
result.summary()
```

```
from imblearn.over_sampling import SMOTE
```

```
smote = SMOTE(random_state=200)
```

```
X_train_over,y_train_over = smote.fit_resample(X_train_cc,y_train_c)
```

```
print('SMOTE 적용 전 학습용 피쳐/레이블 데이터 세트: ', X_train_cc.shape, y_train_c.shape)
```

```
print('SMOTE 적용 후 학습용 피쳐/레이블 데이터 세트: ', X_train_over.shape,  
y_train_over.shape)
```

```
print('SMOTE 적용 후 레이블 값 분포: \n', pd.Series(y_train_over).value_counts())
```

```
# Logistic Regression : 모든 변수
```

```
lr = LogisticRegression()
```

```
modeling(lr,X_train_over,X_test_cc,y_train_over,y_test_c)
```

```
model = sm.Logit(y_train_over, X_train_over)
```

```
result = model.fit()
```

```
result.summary()
```

```
model = sm.Logit(y_train_over, X_train_over.iloc[:,[0,1,3,4,5,6,7,8,9]])
```

```
result = model.fit()

result.summary()


model = sm.Logit(y_train_over, X_train_over.iloc[:,[0,1,3,4,5,6,7,8]])

result = model.fit()

result.summary()


# Logistic Regression : 일부 변수

lr = LogisticRegression()

modeling(lr,X_train_over.iloc[:,[0,1,3,4,5,6,7,8]],X_test_cc.iloc[:,[0,1,3,4,5,6,7,8]],y_train_over,
y_test_c)


# LGBM Classifier

from lightgbm import LGBMClassifier

lgb =
LGBMClassifier(n_estimators=1000,num_leaves=64,n_jobs=-1,boost_from_average=False)

modeling(lgb,X_train_cc.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_c,y_test_c)


# LGBM Classifier

from lightgbm import LGBMClassifier

lgb =
LGBMClassifier(n_estimators=1000,num_leaves=64,n_jobs=-1,boost_from_average=False)

modeling(lgb,X_train_over.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_over,y_test_c)


# XGB Classifier

import xgboost as Xgb
```

```
xgb = Xgb.XGBClassifier()

modeling(xgb,X_train_cc.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_c,y_test_c)


# XGB Classifier

import xgboost as Xgb

xgb = Xgb.XGBClassifier()

modeling(xgb,X_train_over.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_over,y_test_c)


# K Neighbors Classifier

from sklearn.neighbors import KNeighborsClassifier

knn = KNeighborsClassifier(n_neighbors = 3)

modeling(knn,X_train_cc.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_c,y_test_c)


# K Neighbors Classifier

from sklearn.neighbors import KNeighborsClassifier

knn = KNeighborsClassifier(n_neighbors = 3)

modeling(knn,X_train_over.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_over,y_test_c)


# Gaussian Naive Bayes

from sklearn.naive_bayes import GaussianNB

gaussian = GaussianNB()

modeling(gaussian,X_train_cc.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_c,y_test_c)


# Gaussian Naive Bayes

from sklearn.naive_bayes import GaussianNB
```

```
gaussian = GaussianNB()

modeling(gaussian,X_train_over.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_over,y_test_c)


# Decision Tree

from sklearn.tree import DecisionTreeClassifier

decision = DecisionTreeClassifier()

modeling(decision,X_train_cc.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_c,y_test_c)


# Decision Tree

from sklearn.tree import DecisionTreeClassifier

decision = DecisionTreeClassifier()

modeling(decision,X_train_over.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_over,y_test_c)


# Random Forest

from sklearn.ensemble import RandomForestClassifier

random_forest = RandomForestClassifier(n_estimators=100, random_state=300)

modeling(random_forest,X_train_cc.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_c,y_test_c)


# Random Forest

from sklearn.ensemble import RandomForestClassifier

random_forest = RandomForestClassifier(n_estimators=100, random_state=300)

modeling(random_forest,X_train_over.iloc[:,1:],X_test_cc.iloc[:,1:],y_train_over,y_test_c)


import seaborn as sns

import numpy as np
```

```
%matplotlib inline

# feature importance 추출

print("Feature importances:\n{0}".format(np.round(random_forest.feature_importances_,
3)))

# feature별 importance 매핑
for name, value in zip(X_train_over.keys()[1:], random_forest.feature_importances_):

    print('{0} : {1:.3f}'.format(name, value))

import matplotlib

matplotlib.rcParams['font.family'] = 'Malgun Gothic'

matplotlib.rcParams['axes.unicode_minus'] = False

# feature importance를 column 별로 시각화 하기

sns.barplot(x=random_forest.feature_importances_, y=X_train_over.keys()[1:])

pd.set_option('display.max_rows', 500)

Gyeonggi = pd.read_csv('C:/Users/user/Desktop/경기데이터_정제.csv', encoding='cp949')

Gyeonggi1 = Gyeonggi.drop(['심야약국', '응급실'],axis=1)

Gyeonggi1

classifier_rf = random_forest.fit(X_train_over.iloc[:,1:], y_train_over)
```



```
predictions = classifier_rf.predict_proba(Gyeonggi1.iloc[:,2:])

print(predictions)

pred_t = pd.DataFrame(predictions)

pred_t

final = pd.concat([Gyeonggi1, pred_t], axis=1)

final

final.to_csv('C:/Users/user/Desktop/경기_예측.csv', encoding='cp949', index = False)
```