

제19회 한국능률협회컨설팅 경영혁신 대학(원)생 공모전

대학 경영위기 조기 대응을 위한 셀프진단에 대한 연구: 머신러닝을 활용한 한계대학 예측 모델 탐색

팀 김나박이

강남대학교 교육학과

김승준
이민재
박지은
이동건

I. 서론

1. 저출산 및 학령인구 감소
2. 등록금 수입 및 국고보조금 의존율
3. 대학구조조정 정책 : 경영위기대학

II. 선행연구

1. 선행연구 검토 : 한계대학 예측 요인

III. 연구 방법

1. 한계대학 예측 시스템 구축 : 오렌지3 프로그램
2. 예측 모형 성능 평가 : 예측 모형
3. 예측 모형 성능 평가 : 성능평가 지표
4. 한계대학 예측을 위한 자료수집

IV. 연구결과

1. 한계대학 예측 모형 성능 평가
2. 한계대학 예측 실시
3. 2023년도 한계대학 예측 실시

V. 결론

1. 시사점 및 기대효과

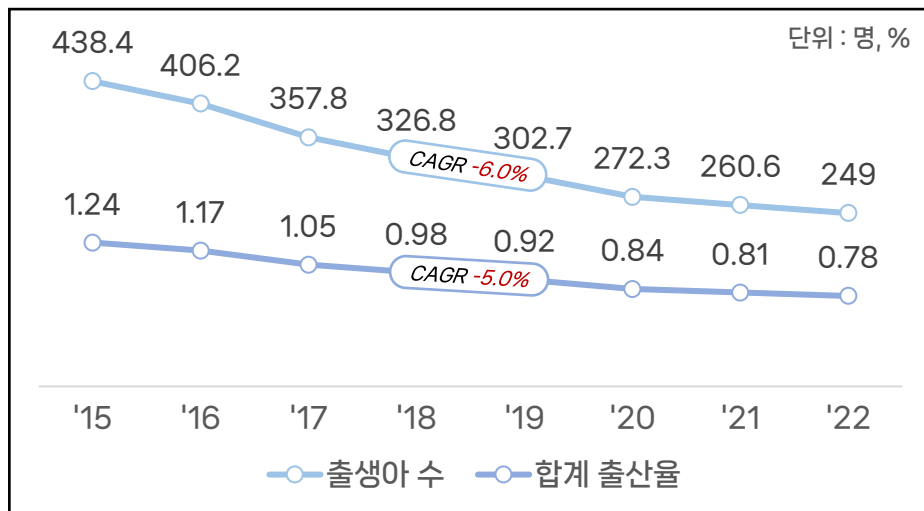
참고문헌

대학 경영위기 조기 대응을 위한 셀프진단에 대한 연구: 머신러닝을 활용한 한계대학 예측 모델 탐색

I. 서론

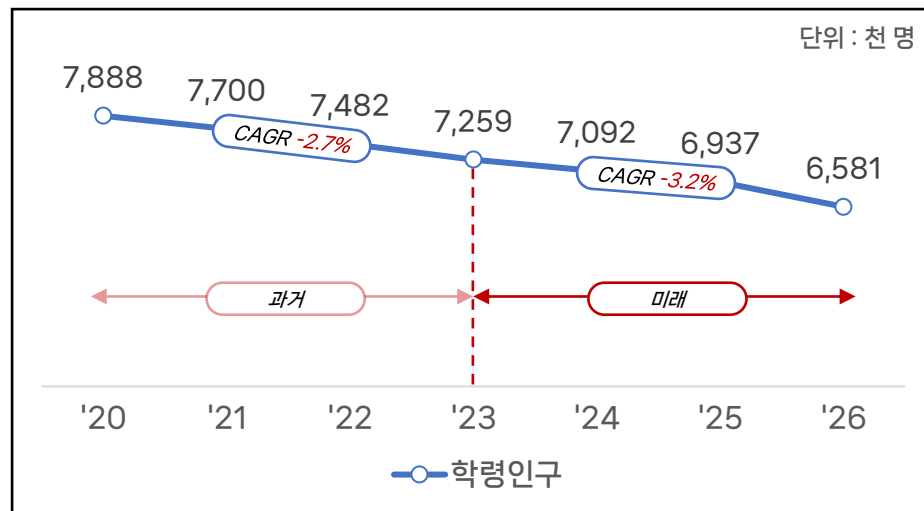
저출산 시대의 도래로 인한 학령인구의 지속적인 감소

연도별 대한민국 저출산 동향



- 출생아 수는 '15년 이후로 **지속적인 감소세**를 보이며 연평균 성장률은 **-6.0%**로 나타나고 있음
- 합계 출산율의 연평균 성장률은 **-5.0%**로 '22년 기준 **OECD 가입 38개국 중 38위**에 위치하고 있음

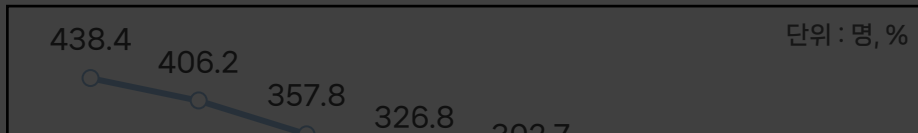
연도별 대한민국 학령인구 추이



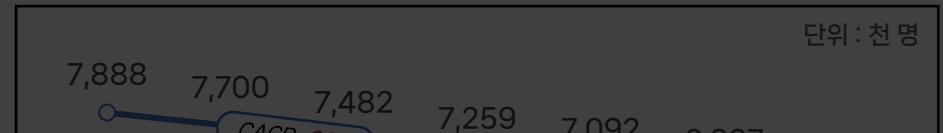
- '23년 기준 **과거 3년 간** 연평균 성장률은 **-2.7%**이고, **미래 3년 간** 연평균 성장률은 **-3.2%**로 예측됨
- '26년 학령인구 추이는 **6,581 천 명**으로 현재 대비 **약 9.3%** 수준의 감소가 예측됨

저출산 시대의 도래로 인한 학령인구의 지속적인 감소

연도별 대한민국 저출산 동향



연도별 대한민국 학령인구 추이



Q. 학령인구 감소로 인한 재정난을 겪는 대학들은 어떻게 극복하는가?

출생아 수 합계 출산율

출처: 통계청

- 출생아 수는 '15년 이후로 **지속적인 감소세**를 보이며 연평균 성장률은 **-6.0%**로 나타나고 있음
- 합계 출산율의 연평균 성장률은 **-5.0%**로 '22년 기준 **OECD 가입 38개국 중 38위**에 위치하고 있음

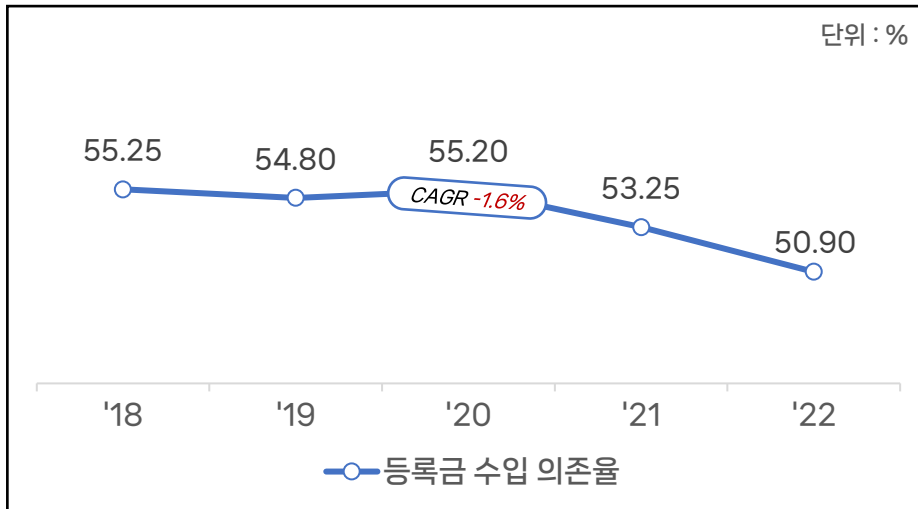
학령인구

출처: 통계청

- '23년 기준 **과거 3년 간** 연평균 성장률은 **-2.7%**이고, **미래 3년 간** 연평균 성장률은 **-3.2%**로 예측됨
- '26년 학령인구 추이는 **6,581 천 명**으로 현재 대비 **약 9.3%** 수준의 감소가 예측됨

등록금 수입 의존율 완화와 국고보조금 의존율 가속화 간의 연관성

연도별 대학 등록금 수입 의존율

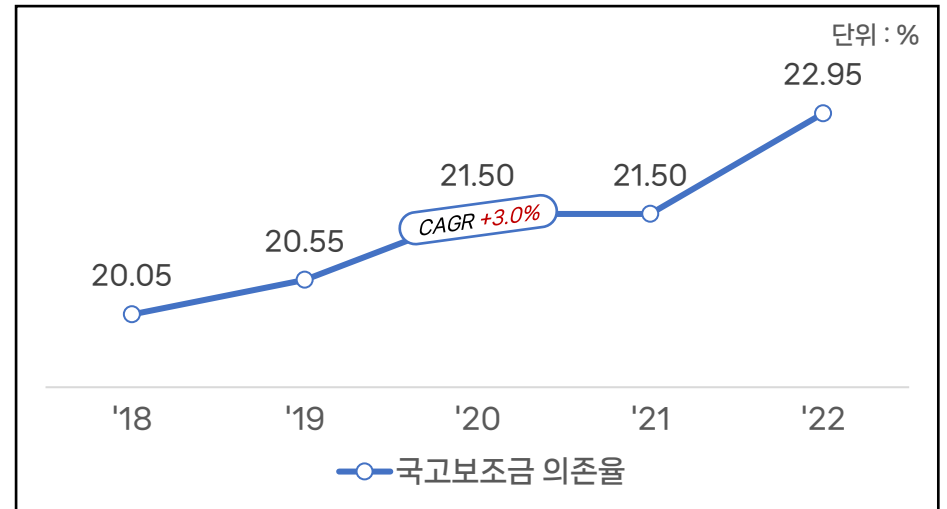


출처 : 대학재정알리미

산식 : 등록금 수입 / 자금 수입 총액 * 100

- '20년 이후, 등록금 수입 의존율은 **급격하게 감소하였음**
- '18년~'22년에 대한 연평균 성장률은 **-1.6%** 수준임

연도별 대학 국고보조금 의존율



출처 : 대학재정알리미

산식 : 국고보조금 수입 / 자금 수입 총액 * 100

- '16년부터 시행된 대학재정지원 정책으로 대학 국고보조금 의존율의 연평균 성장률은 **+3.0%** 수준임
- 대학의 **등록금 수입에 대한 의존이 국고보조금으로 전가**되며 국고보조금 의존율이 증가하고 있는 추세임

등록금 수입 의존율 완화와 국고보조금 의존율 가속화 간의 연관성

연도별 대학 등록금 수입 의존율

단위 : %

연도별 대학 국고보조금 의존율

단위 : %
22.95

A. 대부분의 대학이 등록금 수입 감소 대안으로 국고보조금에 의존하고 있음

—○— 등록금 수입 의존율

출처 : 대학재정알리미

산식 : 등록금 수입 / 자금 수입 총액 * 100

- '20년 이후, 등록금 수입 의존율은 **급격하게 감소하였음**
- '18년~'22년에 대한 연평균 성장률은 **-1.6%** 수준임

—○— 국고보조금 의존율

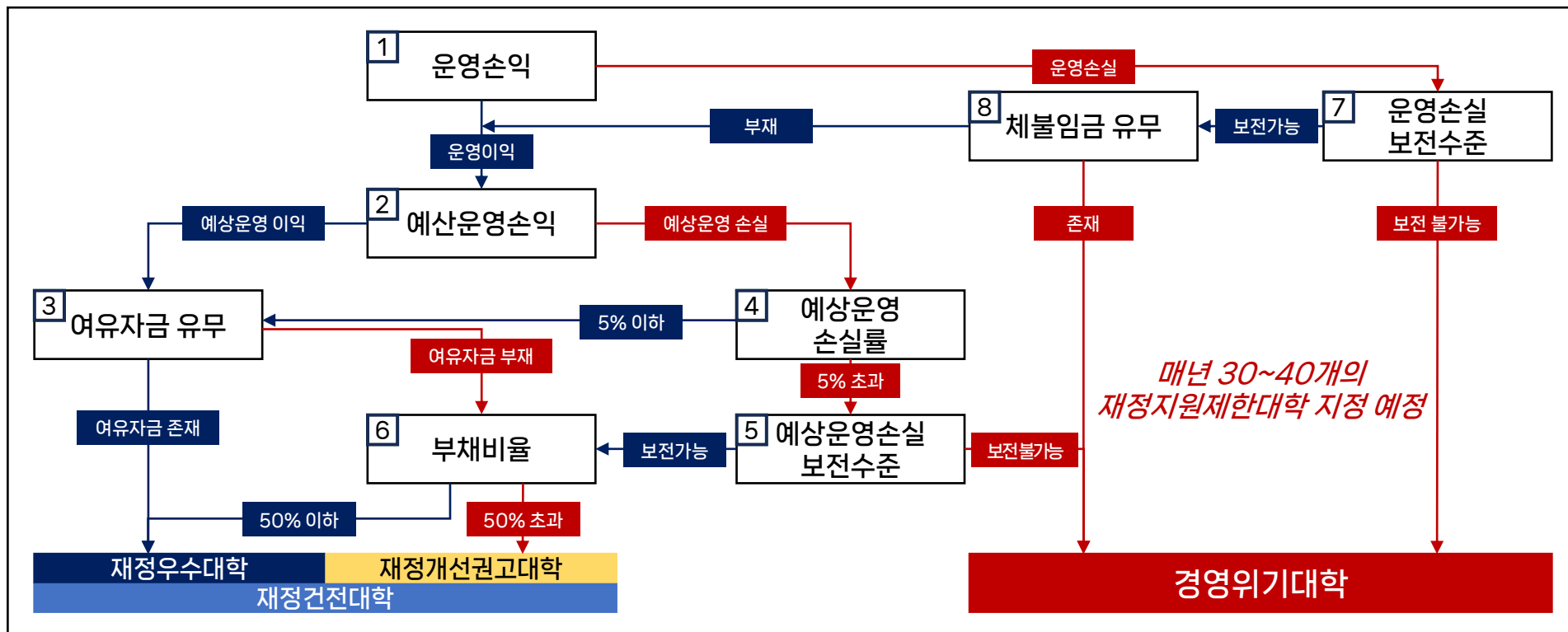
출처 : 대학재정알리미

산식 : 국고보조금 수입 / 자금 수입 총액 * 100

- '16년부터 시행된 대학재정지원 정책으로 대학 국고보조금 의존율의 연평균 성장률은 **+3.0%** 수준임
- 대학의 **등록금 수입 의존에 대한 부담이 국고보조금으로 전가**되며 국고보조금 의존율이 증가하고 있는 추세임

엄격한 평가 체계 구축을 통한 사후관리 초점 재정진단

한국사학진흥재단의 경영위기대학 재정진단 프로세스



한국사학진흥재단은 8개 주요 지표 중심 재정진단 프로세스를 구축하고 **절대평가 방식**으로 경영위기대학 지정 여부 결정함

대학 경영위기 조기 대응을 위한 셀프진단에 대한 연구: 머신러닝을 활용한 한계대학 예측 모델 탐색

II. 선행연구

김경민 외(2023)의 한계대학 예측 요인

일반대학에 대한 한계대학 예측 요인

정(+) 요인	부(-) 요인
중도탈락율	학생 1인당 교육비
지역_경남	교육비 환원율
	재학생 총원률
	유지 총원률
	신입생 총원률
	취업률
	운영경비부담율
	전임교원 1인당 연구비

출처 : 김경민 외(2023)

전문대학에 대한 한계대학 예측 요인

정(+) 요인	부(-) 요인
이월금비율	학생 1인당 교육비
교원확보율	재학생 총원률
지역_강원	대학규모
중도탈락율	교육비환원율
부채비율	유지 총원률
	운영경비부담율
	신입생 총원률
	장학금 지급율
	전임교원 1인당 연구비

김경민 외(2023)가 머신러닝을 활용하여 추출한 한계대학 예측 요인으로 일반대학 10개와 전문대학 14개 지표를 활용함

대학 경영위기 조기 대응을 위한 셀프진단에 대한 연구: 머신러닝을 활용한 한계대학 예측 모델 탐색

III. 연구방법

'오렌지3' 프로그램을 활용한 한계대학 예측 시스템 구축

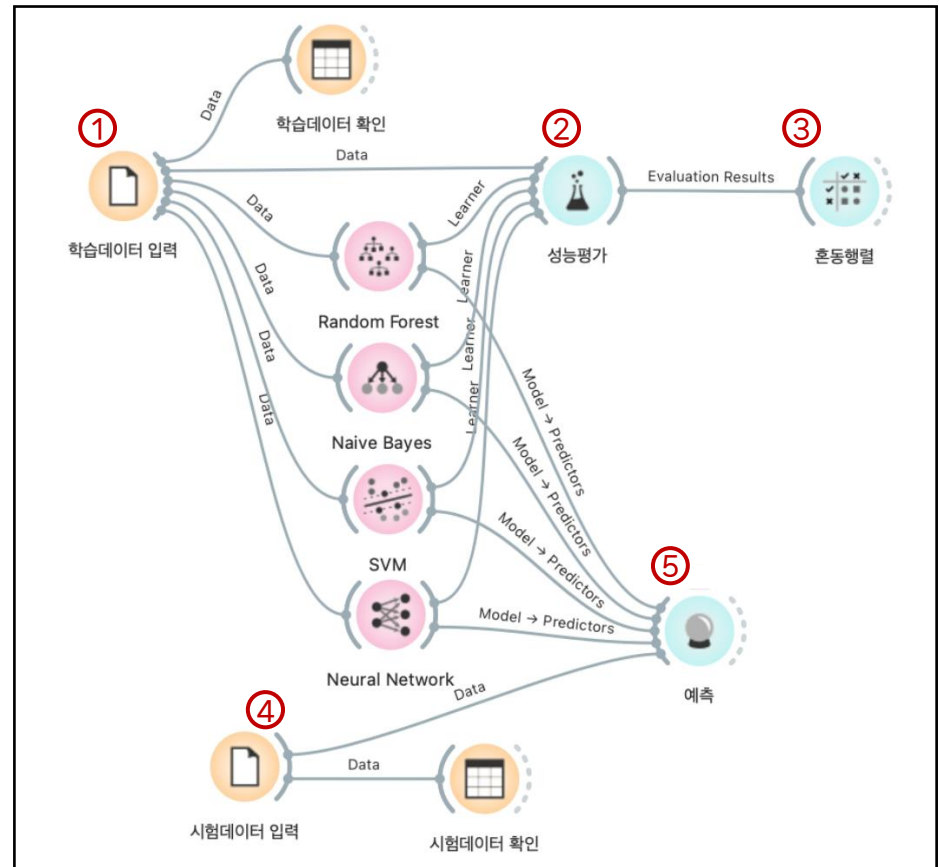
'오렌지3' 머신러닝 프로그램이란?



출처 : '오렌지3' 공식홈페이지

- 슬로베니아 루블라냐 대학에서 개발한 **코딩 없이** 인공지능 기법을 활용하여 데이터 분석을 할 수 있는 프로그램으로 무료로 전 세계에 공개한 **오픈소스 소프트웨어**
- 노코딩을 기반으로 **데이터를 분석하고 시각화**하여 머신러닝에 적용할 수 있는 프로그램

'오렌지3' 한계대학 예측 시스템 구조



지도학습 내 분류 모델에 대한 대표적인 네 가지 알고리즘

지도학습 내 분류 모델의 머신러닝 알고리즘 종류

랜덤포레스트

- 랜덤포레스트는 입력변수로부터 **랜덤하게** 학습하여 **다수의 의사 결정 나무**를 구현함
- 상호 경쟁을 통해 가장 신뢰도 높고 안정적인 예측모형을 선택하는 방식

서포트 벡터 머신

- 서포트 벡터 머신은 **패턴 인식** 분야에서 제안된 알고리즘으로 주어진 샘플 그룹에 대해 **그룹 분류 규칙**을 찾아내는 기법

나이브베이지

- 나이브베이지는 **다양한 요소**들이 **독립적**임을 가정하며, 요소들 간 주어진 특징에 대한 **조건부 확률**을 계산하는 기법

뉴럴 네트워크

- 뉴럴 네트워크는 인간의 신경 회로를 모방하는 네트워크로 **이미 알고 있는** 입력 자료와 출력 자료의 관계를 **학습**시켜 임의의 입력에 대한 출력 결과를 추정하는 기법

1종 오류와 2종 오류를 최소화할 수 있는 F1(조화평균)지표 적극 활용

성능평가지표

지표	내용
AUC	예측이 성공할수록 넓이는 1에 가까워지는 ROC 그래프의 넓이
CA	전체 예측 데이터 중 실제 값과 일치하는 데이터의 비율
F1	1종 오류와 2종 오류의 발생률을 낮추는 최적의 모형을 평가하는 지표
Prec	참으로 예측한 데이터 중 실제로 참인 데이터의 비율
Recall	실제로 참인 데이터 중 참으로 예측한 데이터의 비율
MCC	불균형한 데이터의 모델 성능 평가에 용이한 지표

F1(조화평균)지표

구분	Predicted(예측값)		
	True/False	0(False)	1(True)
	Actual (실제값)	0(False)	1(True)
Actual (실제값)	0(False)	TN(True Negative)	FP(True Positive)
	1(True)	FN(False Negative)	TP(True Positive)

[혼동행렬 매트릭스]

TN: 실제 값이 거짓인 데이터를 거짓으로 예측한 경우	FP: 실제 값이 거짓인 데이터를 참으로 잘못 예측한 경우 (2종 오류)
FN: 실제 값이 참인 데이터를 거짓으로 잘못 예측한 경우 (1종 오류)	TP: 실제 값이 참인데 데이터가 참으로 예측한 경우

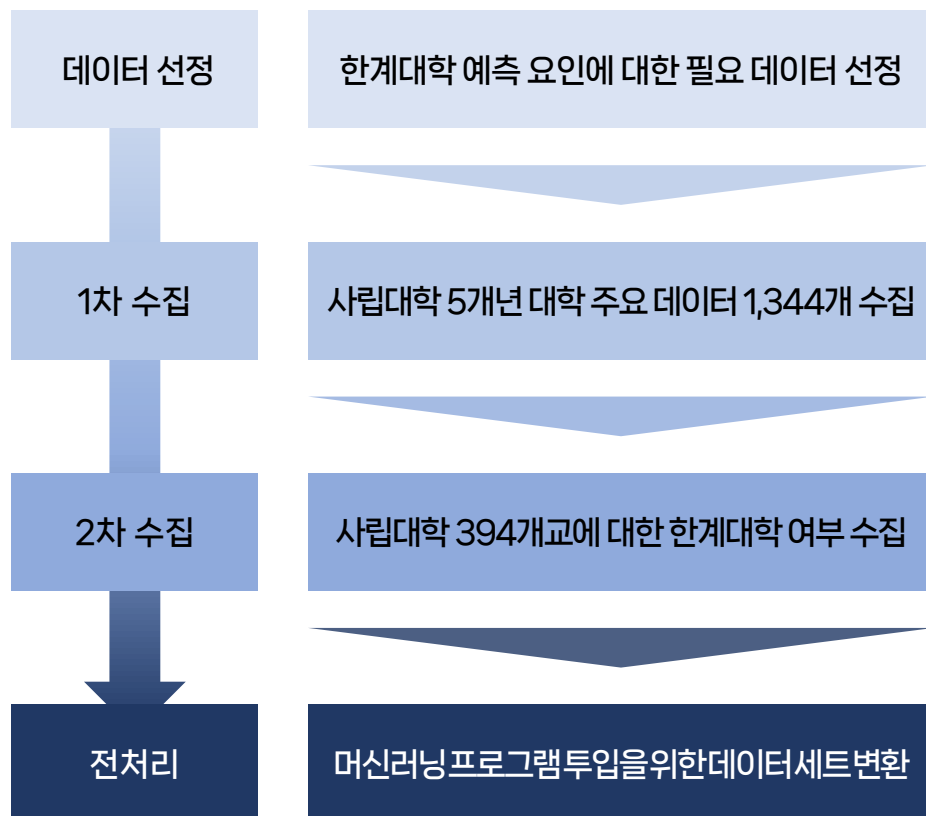
[혼동행렬에 대한 해설]

1종 오류와 2종 오류로 인한 머신러닝 예측력 저하

1종 오류와 2종 오류에 대한 발생률을 낮추는 최적의 지점을 평가하는 F1지표를 핵심 성능 평가 지표로 선정

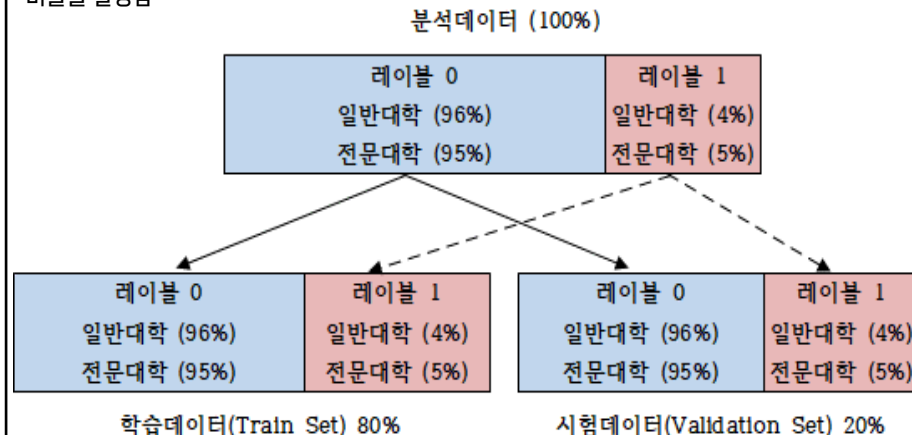
대학알리미 내 정보공시 자료 기반 1,344개 데이터 추출

대학알리미 내 데이터 추출 프로세스



학습데이터와 시험데이터의 구분 도식화

*김경민 외(2023)에서 학습데이터와 시험데이터의 비율을 8:2로 설정하여 본 연구도 동일한 비율을 설정함



- 1,344개에 대한 분석데이터를 8:2 비율로 학습데이터와 시험데이터로 구분함
- 일반대학 중 정상대학은 96%이고 한계대학은 4%이며 전문대학 중 정상대학은 95%이고 한계대학은 5%임

*레이블 0은 정상대학을 의미하며 레이블 1은 한계대학을 의미한다.

대학 경영위기 조기 대응을 위한 셀프진단에 대한 연구: 머신러닝을 활용한 한계대학 예측 모델 탐색

IV. 연구결과

4가지 예측 알고리즘에 대한 6가지 지표를 통한 성능평가 실시

일반대학 : 랜덤포레스트

모형	랜덤 포레스트	서포트 벡터 머신	나이브 베이즈	뉴럴 네트워크
AUC	0.886	0.824	0.910	0.922
CA	0.974	0.969	0.833	0.972
F1	0.970	0.961	0.881	0.967
Prec	0.971	0.966	0.964	0.970
Recall	0.974	0.969	0.833	0.972
MCC	0.595	0.480	0.374	0.561

일반대학의 경우, 랜덤포레스트의 F1지표가 0.970로 가장 높게 나타나고 있음

전문대학 : 뉴럴 네트워크

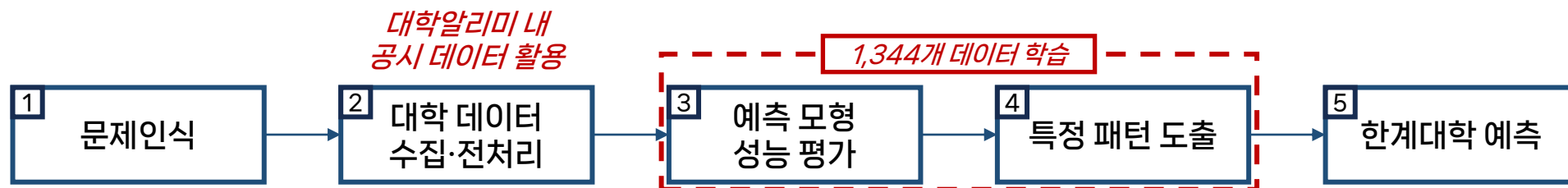
모형	랜덤 포레스트	서포트 벡터 머신	나이브 베이즈	뉴럴 네트워크
AUC	0.954	0.982	0.880	0.978
CA	0.968	0.966	0.900	0.974
F1	0.963	0.959	0.920	0.972
Prec	0.965	0.964	0.952	0.973
Recall	0.968	0.966	0.900	0.974
MCC	0.589	0.554	0.432	0.961

전문대학의 경우, 뉴럴 네트워크의 F1지표가 0.972로 가장 높게 나타나고 있음

*수치가 1에 가까울수록 모델 성능이 뛰어난을 의미함

대학 유형 별 한계대학 예측 알고리즘에 대한 성능평가 결과 일반대학은 랜덤포레스트, 전문대학은 뉴럴 네트워크가 선정되었음

일반대학 148개교와 전문대학 121개교에 대한 높은 예측 성공률



[머신러닝을 활용한 한계대학 예측 절차]

대학 유형	실시 대학 수	실제 한계대학 수	한계대학 예측 수	1종·2종 오류	예측 성공률
일반대학	148	6	6	0	99.99%
전문대학	121	6	5	1	99.17%

[머신러닝을 통한 한계대학 예측 결과]

대학 유형 별 한계대학 예측 알고리즘에 대한 성능평가 결과 일반대학은 **랜덤포레스트**, 전문대학은 **뉴럴 네트워크**가 선정되었음

임의의 23학년도 일반대학에 대한 한계대학 예측 결과

구분	A	B	C	D	E	F	G	H	I	J
중도탈락율	5.1	18.5	12.2	12.9	4.2	11.4	13	18.9	5	25.1
지역_경남	0	0	0	0	0	0	0	0	0	0
학생1인당교육비(천원)	12,501	10,716	28,700	12,291	12,608	9,842	5,724	13,872	12,358	9,693
교육비환원율	171.53	165.79	469.91	156.41	157.59	159.35	77.67	259.23	0.00	0.00
재학생총원율	98.00	18.00	47.50	37.00	97.20	71.70	75.30	29.40	92.80	36.40
신입생총원율	99.90	30.70	30.00	23.30	99.90	46.80	98.90	16.20	99.80	26.10
유지총원율	99.86	21.13	37.36	33.07	98.71	63.13	85.02	26.57	96.58	32.82
취업률	60.8	66.7	66.7	51.3	67.4	61.3	61.7	66.8	65.7	67.5
운영경비부담율	45.90	19.80	125.00	41.10	0.00	44.70	125.00	0.00	26.10	0.00
전임교원1인당연구비(천원)	22,650	0	0	0	143,049	2,200	1,050	5,031	77,212	0
한계대학	0	1	1	1	0	1	1	1	0	1

무작위로 선정한 2023학년도 일반대학 10개교의 한계대학 예측 성공률은 **약 90.00%**임

임의의 23학년도 전문대학에 대한 한계대학 예측 결과

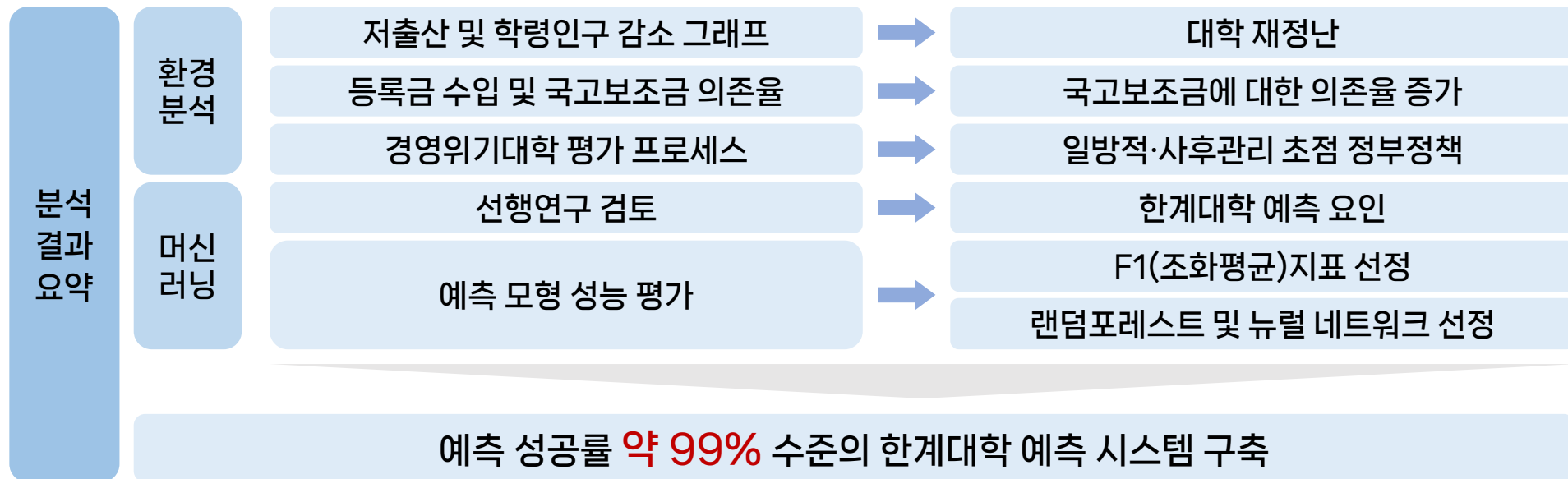
구분	a	b	c	d	e	f	g	h	i	j
중도탈락율	3.4	74.4	14.1	15.6	9.3	7.2	5.2	6.7	10.6	5.1
지역_강원	1	1	0	0	0	0	0	0	0	0
학생1인당교육비(천원)	10,151	9,233	8,514	13,251	11,360	10,501	12,059	5,543	7,971	11,080
교육비환원율	175.26	181.30	209.53	248.98	183.27	168.36	208.49	99.21	122.44	190.45
교원확보율	70.43	63.64	79.31	80.56	103.15	60.33	80.65	100	128.57	59.91
이월금비율	75.5	93.9	71	63.4	54.8	96.1	78	43.7	41.2	97.9
부채비율	79.8	100	73	57.8	38.6	100	82.6	47.4	29.4	100
대학 규모	4.93	23.51	21.16	12.58	39.25	1.90	11.66	52.44	3.66	2.88
재학생총원율	0.00	0.00	1.08	0.00	0.04	0.10	0.15	0.00	4.19	0.09
신입생총원율	47.58	62.87	39.18	38.88	35.88	59.23	50.54	37.82	22.75	59.96
유지총원율	10.42	10.30	10.07	10.38	10.88	11.08	10.53	10.11	10.15	10.57
장학금지급율	56.46	74.31	84.32	66.16	46.50	57.94	58.85	39.99	18.09	58.66
운영경비부담율	0.00	125.00	0.00	41.20	125.00	0.00	19.80	86.20	68.30	125.00
전임교원1인당연구비(천원)	333	0	0	4,186	2,082	3,027	0	0	0	2,212
한계대학	0	1	1	0	1	0	0	1	1	0

무작위로 선정한 2023학년도 전문대학 10개교의 한계대학 예측 성공률은 **약 80.00%**임

대학 경영위기 조기 대응을 위한 셀프진단에 대한 연구: 머신러닝을 활용한 한계대학 예측 모델 탐색

V. 결론

시사점 및 기대효과



시사점

평가자 중심의 일방적인 평가에서 평가 대상자가 **대응가능한 방향**으로 변화 시사

한계대학 예측 시스템을 통해 대학 평가에 대응하여 **각자의 생존 전략 수립**하도록 함

평가 대상자 중심의 대학 경영 및 생존 전략에 초점이 맞추어진 **후속연구 기대**

기대효과

머신러닝을 활용한 셀프진단을 통한 **솔루션 제공**

더 나아가, **국가 주도** 대학 구제사업 확대

- 김경민, 이종현. (2023). 머신러닝을 활용한 한계대학 예측 요인 탐색 연구 : 랜덤 포레스트 모형을 중심으로. 교육재정경제연구, 제32권 제1호, 6~24.
- 김민규 외 2명. (2022). 대학의 폐교지정이 주변 지역 공동주택 가격에 미치는 영향 - 동부산대학교를 중심으로. 대한토목학회논문집, 제42권 제2호, 279.
- 김병주. (2020). 대학구조조정 정책의 분석과 개선과제, 국회입법조사처, 1.
- 김한수. (2022). 학령인구 감소가 대학 재정에 미치는 영향:경영성과 판단을 위한 재무지표와 모형을 중심으로. 상업교육연구, 8.
- 문남철. (2021). 대학폐교가 지역 경제 및 지역 사회에 미친 영향에 관한 연구 : 남원 서남대학교를 사례로. 국토지리학회지, 제55권 제3호, 264.
- 박귀만, 배영철. (2019). 다양한 종류의 예측에서 머신러닝 성능 비교. 한국전자통신학회 논문지, 제14권 제1호, 169~178.
- 서영인 외 6명. (2020). 학령인구 감소에 따른 한계대학 대응 방안 연구, 한국교육개발원.
- 연덕원. (2022). 대학 구조조정 현재와 미래(정원 정책을 중심으로). 대학교육연구소, 2~3.
- 장제국. (2023). 학생 미충원에 따른 사립대학 재정손실 분석. 한국대학교육협의회 고등교육 포커스, 제8호, 1~4.
- 홍기혜. (2020). 랜덤포레스트 머신러닝 알고리즘 기반 남·여 청소년의 자살생각 예측 및 분석. 한국사회복지학, 제72권 제3호, 157~180.

- 공정뉴스. (2023). 경영위기 사립대 19개... 국고 끊긴 2025년 폐교
(<http://www.fairn.co.kr/news/articleView.html?idxno=67933>)
- 대한민국 정책 브리핑. (2023). '지역 혁신 중심 대학지원체계(RISE)' 구축 배경과 기대
(<https://www.korea.kr/news/contributePolicyView.do?newsId=148911817>)
- "오렌지" 설명 : 스마투스. (2021). 오렌지 소프트웨어 강좌. 제1강
(<http://www.sbr.ai/news/articleView.html?idxno=904>)
- 통계청. (2023). 2022년 출생통계(확정), 국가승인통계 제10103호 출생통계
(https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=1428)

Thank You !

Q&A