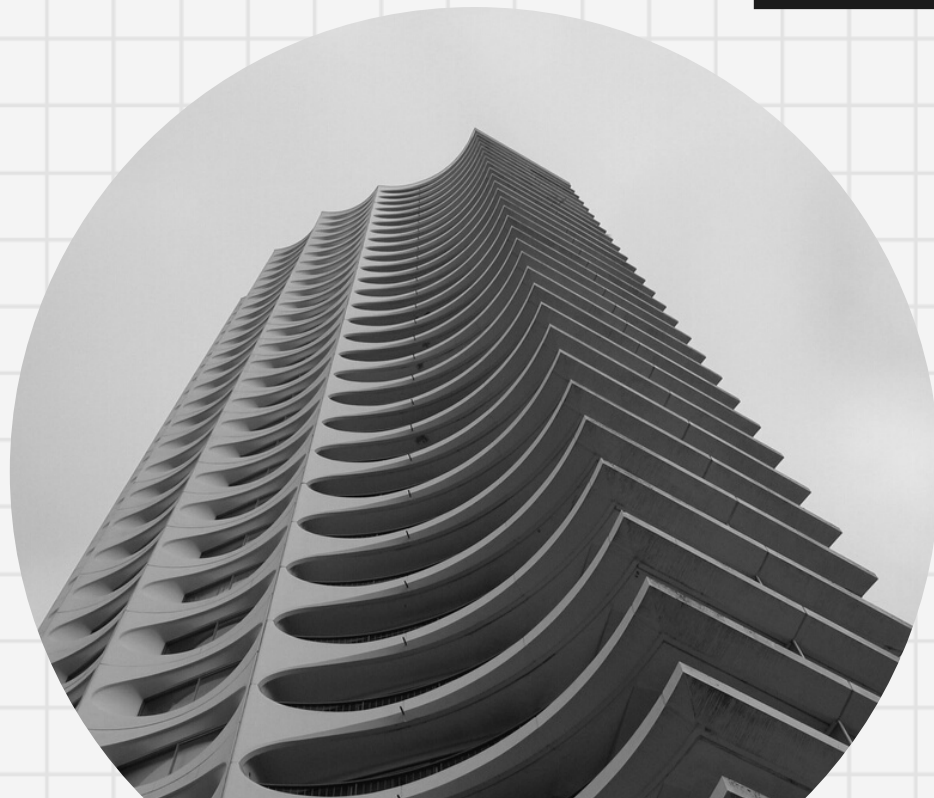


FITING

기업 부실 등급화 프로젝트 보고서

발표자: 이지현



1. 프로젝트 개요



WeakTing
기업 대출
부실평가 모델

배경 [기업 여신의 새 바람]

ESG 등 전통적인 평가 모형에 사용되지 않았던 새로운 Feature들을 추가한 대안 평가

목표

기존 부실 평가를 기반으로 하되,
"머신러닝을 활용한", 더 개선된,
부실기업(Kosdaq 제조업) 예측 모델 개발
-> NEW 부실 등급 제도

Business 활용

- 기업 추천
- 은행 대출심사
- 기업 분석 리포트

...

3. 프로젝트 수행절차 & 방법

데이터 수집 & 전처리

재무/비재무 데이터 수집
크롤링
데이터 전처리

데이터 모델링 & 시각화

적절한 모델 선정
모델 시각화

빅데이터 분석기획

기획안 작성
필요한 피처 선정

데이터 분석 & 시각화

Lasso - 피처 선정
다중공선성 해결
EDA -시각화
(Histogram, Scatter plot, ...)

데이터 모델 평가 및 검증

F1 Score,
Roc Curve - Auc score

3-1. 데이터 수집 & 전처리

FiTingDB



FiTing DB

A new tool for teams & individuals that blends everyday work apps into one.

Feature Data Set

표 표

목적	이름	파일	움김 여부	수정사항 특징	
기업 부실 이분류 학습	DB-0518-1800(태환)	DB-0518... DB-0518...	<input type="checkbox"/>		2022년 5월 19일 오후 1:07
기업 부실 이분류 학습	DB-0523-1400(민정)	DB-0523...	<input type="checkbox"/>	공공조달 미포함	2022년 5월 23일 오후 2:52
기업 부실 이분류 학습	DB-0524-1548(태환) (2)	DB-0524...	<input type="checkbox"/>	근속연수, 기부금, 등등 비율 환산 미포함	2022년 5월 26일 오후 3:04
기업 부실 이분류 학습	DB-0526(승연)	DB-0526...	<input type="checkbox"/>	'시총','유동비율', '당좌비율', '부채비율', '차입금의존도', '이자보상배율', '매출총이익률', '영업이익률', '당기순이익률', 'ROA', 'ROE', '총자산회전율', '현금순환주기', '매출액증가율', '영업이익증가율', '유형자산증가율', '총자산증가율', '재고자산회전율', '매출채권회전율', '매입채무회전율', '복리후생비/영업이익', '경상개발비/영업이익', '기부금/영업이익', '교육훈련비/영업이익', '자산/기말인원', '평균근속연수/기업존속연수', '비정규직비율', '여성고용률', '공공조달은1'	2022년 5월 26일 오후 4:04
기업 부실 이분류 학습	DB-0526-1813(승연)	DB-0526...	<input type="checkbox"/>	이자보상배율 역수로 변경	2022년 5월 26일 오후 6:13
기업 부실 이분류 학습	DB-0527-1427(승연)	DB-0527...	<input type="checkbox"/>	-(평균근속연수/평균근속연수.max())/((기업존속연수/기업존속연수.max())+(상장년수/상장년수.max())) -기업존속연수/기업존속연수.max() -상장년수/상장년수.max() -시가총액/영업이익	2022년 5월 27일 오후 3:26
기업 부실 이분류 학습	DB - 0528-1504(태환)	DB - 052...	<input type="checkbox"/>	시총/자산으로 수정	2022년 5월 28일 오후 8:23
기업 부실 이분류 학습	DB_0530_0100(태환)	DB_0530...	<input type="checkbox"/>	시총/자본으로 수정	2022년 5월 30일 오후 2:23
기업 부실 이분류 학습	DB_0531_1740(태환)	DB_0531...	<input type="checkbox"/>	ROE가 원래 당기순이익/자기자본 지금은 분모가 평균자본으로 들어가 있었다. 그랬더니 ROE보다 ROA가 더 중요해졌다. 따라서 ROA채택	2022년 5월 31일 오후 5:46
부실 등급 다분류 학습	prediction_with_predict_proba(지현)_0531		<input type="checkbox"/>	xgboost predict_proba	2022년 6월 1일 오후 7:38
부실 등급 다분류 학습	prediction_with_predict_proba(지현)_0531_2112	predicti...	<input type="checkbox"/>	catboost predict_proba	2022년 6월 1일 오후 12:38
부실 이분류 및 등급 검증	2022_부실폐지기업_catboost_predict_proba_06_02		<input type="checkbox"/>	2017~2021년 학습 세트에 대한 2022년 실제 부실 상장폐지 기업 7개에 대한 predict_proba 데이터	2022년 6월 2일 오후 3:42

3-1. 데이터 수집 & 전처리

TS2000 - 약 80여개의 초기 Features

회사명	거래소코드	회계년도	자산	부채	자본	매출액	매출원가	매출총이익	영업이익	당기순이익	유동자산	당좌자산	재고자산	유동부채	매출채권	매입채무
삼천당제약(주)	250	2017	235143417	36386893	198756523	161111597	75822851	85288747	26511831	23028009	131206711	109359039	21847673	23699293	31488048	4434709
삼천당제약(주)	250	2018	261158440	55394974	205763466	159960312	76762385	83197927	15500129	14592937	142271889	119842031	22429858	38419138	31897860	3393207
삼천당제약(주)	250	2019	290150966	66954254	223196712	186602477	75946716	110655762	25210117	19599720	162492554	139457883	23034671	45205983	31129224	3617239
삼천당제약(주)	250	2020	317078866	59382900	257695966	166862334	72862414	93999919	5522905	7494986	166527651	141191159	25336491	23658314	26655005	3177946
삼천당제약(주)	250	2021	343947668	93647149	250300519	167274904	84196690	83078214	-15293087	-10360886	168919766	141012020	27907746	54019011	31862868	4758228
(주)중앙에너지비스	440	2017	60493562	15299436	45194126	70818896	58635726	12183170	642132	1597262	5099750	3511535	1588215	15145366	2909282	730596
(주)중앙에너지비스	440	2018	58743738	13480020	45263718	76133494	63437337	12696157	1424706	889990	4174206	3240003	934203	13391590	2649735	297183
(주)중앙에너지비스	440	2019	61575390	16451398	45123992	65702344	53921816	11780528	638178	615333	4675590	2717425	1958165	14349625	2186140	194714
(주)중앙에너지비스	440	2020	71325527	9718861	61606666	53564699	42974238	10590461	128643	19200785	22613735	21229485	1384249	6816291	1567595	13338
(주)중앙에너지비스	440	2021	64464478	4806806	59657672	60712578	50307026	10405552	37956	251457	15692197	14011004	1681194	3482532	2008892	184798
교육훈련비/영업이익	경상개발비/영업이익	복리후생비/영업이익	기부금/영업이익	비정규직비율	여성고용률	공공조달은1	new_시총	new_평균근속연수	new_기업존속연수	new_상장년수	자산/기말인원	산업	제품			
0.058284846	0.098181563	0.099725704	0.003666288	0.030927835	0.242268041	0	1.718358698	0.179204821	0.948717949	0.53125	606039.7345	의약품 제조업	하메론, 아토로우			
0.100554518	0.57798848	0.152962211	0.019856351	0.051948052	0.249350649	0	4.08245942	0.182581265	0.961538462	0.5625	678333.6104	의약품 제조업	하메론, 아토로우			
0.062923151	0.710975717	0.104337358	0.008497184	0.019900498	0.248756219	0	3.525889754	0.180222612	0.974358974	0.59375	721768.5721	의약품 제조업	하메론, 아토로우			
0.285207694	3.835047679	0.460913414	0.005667307	0.012406948	0.255583127	0	7.253741799	0.186083499	0.987179487	0.625	786796.1935	의약품 제조업	하메론, 아토로우			
-0.089758464	-2.117913015	-0.173183021	-0.009105814	0.010204082	0.24744898	0	3.943239926	0.189007383	1	0.65625	877417.5204	의약품 제조업	하메론, 아토로우			
0.03621841	0	1.301838251	0.087324725	0.35	0.13	0	0.807427536	0.23045157	0.91025641	0.75	604935.62	기타 전문 도매업	석유류도소매			
0.012167423	0	0.533270724	0.007824772	0.387755102	0.142857143	0	0.921748651	0.224492549	0.923076923	0.78125	599425.898	기타 전문 도매업	석유류도소매			
0.02948237	0	1.192786339	0.019997556	0.317073171	0.12195122	0	1.116423425	0.032327741	0.935897436	0.8125	750919.3902	기타 전문 도매업	석유류도소매			
0.088182023	0	5.803720373	0.118638402	0.320987654	0.135802469	0	0.935990008	0.033958524	0.948717949	0.84375	880562.0617	기타 전문 도매업	석유류도소매			
0.278849194	0	19.43666351	0.44367162	0.320512821	0.166666667	0	1.831885093	0.030776235	0.961538462	0.875	826467.6667	기타 전문 도매업	석유류도소매			
0.00015255	0	0.043132476	0	0.181818182	0.181818182	0	4.44543458	0.150239368	0.525641026	0.71875	2920188.182	가전제품 및 정보	부동산			
9.89657E-05	0	0.021267078	0	0.166666667	0.25	0	3.681042956	0.134977287	0.538461538	0.75	2654723	가전제품 및 정보	부동산			
0.000116098	0	0.02271706	0	0.181818182	0.181818182	0	3.35892072	0.189244163	0.551282051	0.78125	2923788.364	가전제품 및 정보	부동산			
5.54119E-05	0	0.016410678	0	0.111111111	0.222222222	0	3.441584245	0.192660829	0.564102564	0.8125	3599093.222	가전제품 및 정보	부동산			
8.52619E-05	0	0.019038844	0	0.142857143	0.142857143	0	3.554449111	0.2417715	0.576923077	0.84375	4692788.571	가전제품 및 정보	부동산			

3-1. 데이터 수집 & 전처리

TS2000

지분법손실	복리후생비	교육훈련비	경상개발비	기부금	시총	기말인원	평균근속연수	유동비율	당좌비율	부채비율	차입금의존도	이자보상배율(역수)	매출총이익률	영업이익률	당기순이익률
0	2643911	1545238	2602973	97200	341535000	388	6.1	553.6313298	461.4443098	18.30726984	2.269434147	0.007454747	52.93768331	16.45556961	14.29320386
0	2370934	1558608	8958896	307776	840021000	385	6.4	370.3151513	311.9331595	26.92167617	3.595187274	0.01496394	52.01160585	9.689984226	9.122848548
0	2630357	1586300	17923781	214215	786967000	402	6.5	359.4492216	308.4943048	29.99786753	3.409181481	0.032400881	59.30026427	13.51006557	10.50346186
0	2545581	1575175	21180604	31300	1869260000	403	6.9	703.8863843	596.7929879	23.04378331	6.941807342	0.115593696	56.33381528	3.309857214	4.491718305
0	2648503	1372684	32389428	139256	986995000	392	7.2	312.7042922	261.0414693	37.41388527	8.036118157	-0.124014007	49.66567729	-9.142487387	-6.193927333
0	835952	23257	0	56074	36490981.8	100	8.8	33.67201559	23.18554071	33.85270909	16.07167883	0.538685192	17.2032758	0.906724104	2.25541782
0	759754	17335	0	11148	41721771	98	8.8	31.17035393	24.1943115	29.78107101	17.47773184	0.226811005	16.67617803	1.871326173	1.16898615
0	761210	18815	0	12762	50377481.7	82	1.3	32.58336019	18.93725446	36.45820609	17.9334731	0.647034213	17.93014873	0.971316944	0.936546495
0	746608	11344	0	15262	57663223.8	81	1.4	331.760117	311.4521519	15.7756646	0	1.426288255	19.77134418	0.240163769	35.84596826
ROA	ROE	재고자산회전율	매출채권회전율	총자산회전율	매입채무회전율	현금순환주기	매출액증가율	영업이익증가율	유형자산증가율	총자산증가율	교육훈련비/영업이익				
9.793176136	11.58603937	3.697011372	5.055307532	0.731898145	18.35406423	-323.2834841	9.322329735	8.842057148	-4.751757847	14.64065477	0.058284846				
5.587771546	7.092093307	3.467329061	5.04718847	0.64460896	19.61247029	-383.3249653	-0.714588535	-41.53504901	10.06202724	11.06347068	0.100554518				
6.755007667	8.781365919	3.340921711	5.921342545	0.676942838	21.66672876	-363.2509388	16.65548452	62.64456251	-5.586416605	11.10150834	0.062923151				
2.363760819	2.90846074	3.012638563	5.775359017	0.549585429	21.44530693	-473.2088344	-10.57871434	-78.09250548	10.36253405	9.280651507	0.285207694				
-3.012343727	-4.139378552	3.162659275	5.717053455	0.506106473	21.21845867	-534.5150151	0.247251725	-376.9029523	17.12616199	8.473854577	-0.089758464				
2.640383451	3.53422478	17.75737879	14.28314631	1.02614738	26.13046846	-305.3490051	-22.71069538	-69.51161668	-13.1741247	-21.97916813	0.03621841				
1.515038079	1.966232646	50.29882993	27.39099161	1.277008017	123.445482	-261.6087271	7.504491457	121.8712041	-1.671715596	-2.892578883	0.012167423				
0.999316448	1.363649298	37.28558468	27.17288764	1.092134644	219.2402718	-306.7259804	-13.70113133	-55.20633731	4.628227565	4.820346979	0.02948237				
26.91993429	31.16673283	25.71449138	28.53941421	0.806084716	413.1105493	-419.9891606	-18.47368642	-79.84214435	-3.827232301	15.83447056	0.088182023				
0.390070637	0.421499853	32.8220267	33.9509569	0.894212766	507.8029838	-381.0169759	13.34438377	-70.49509107	-0.182951719	-9.61934568	0.278849194				
1.272502052	2.901027561	2.868857972	2.191374339	0.12366358	3.431440163	-2621.357933	4.532955825	-30.12884172	-8.318922745	-1.226304643	0.00015255				
0.537717746	1.202450711	4.604328421	2.568329382	0.147007945	426.0789775	-2230.490995	17.65586545	40.98524734	-61.89114371	-0.826204538	9.89657E-05				

3-1. 데이터 수집 & 전처리

TS2000

기업존속연수	상장년수	지역	Z-Score	Z-Score_ad	K1_Score	K1_S
74	17	경기도	63.90158057	건전	17.22136806	건전
75	18	경기도	101.7384407	건전	14.40066256	건전
76	19	경기도	79.25456993	건전	13.80365617	건전
77	20	경기도	209.1892709	건전	14.59111989	건전
78	21	경기도	70.51387034	건전	11.01480535	건전
71	24	서울특별시	17.6189413	건전	10.4997435	건전
72	25	서울특별시	22.39096555	건전	11.40731305	건전
73	26	서울특별시	21.90903542	건전	9.936566064	건전
74	27	서울특별시	41.89549433	건전	18.22124173	건전
75	28	서울특별시	151.8951162	건전	28.03889918	건전
41	23	대구광역시	23.11812606	건전	-2.42768191	부실

단기차입금	장기차입금	사채	유형자산	이익잉여금	이자비용	영업활동으로 인한 현금흐름	투자활동으로 인한 현금흐름	재무활동으로 인한 현금흐름	이자수익	법인세비용	지분법이익
2336425	3000000	0	55174955	111080921	197639	28631055	-11900641	-3305826	1414387	3593123	0
2211871	0	7177264	60726674	120001443	231943	25034032	-50272218	14808451	1649749	-131529	0
2439279	0	7452494	57334229	128792637	816830	30603122	-19339984	-2415022	2032619	-282442	0
2011004	20000000	0	63275508	126987000	638413	3249733	-19206331	14661975	1630997	9085845	0
1579730	0	26060311	74112174	112514106	1896557	-7706965	-30050293	27691109	1435326	-9188213	0
9699831	22500	0	45744683	21685552	345907	1838540	3886279	-5709338	1438	439097	0
10262573	4500	0	44979962	21722567	323139	2102430	-96689	-2038171	689	257131	0
10860606	182000	0	47061737	21573653	412923	983261	55535	-950910	2388	168531	0
0	0	0	45260575	39992857	183482	2492070	17945828	-14416302	39625	5542455	0
0	0	0	45177770	38085628	85353	-4369800	1094121	-2806640	277459	47357	0
13000000	3000000	0	12520011	-2992334	575227	1085400	347554	-1282770	1353	15639	0
9800000	3000000	0	4771233	-2836485	574195	1044626	-31757	-1013144	5515	1762	0
8000000	1600000	0	4700565	-2626747	586958	1037999	-38098	-880493	1349	0	0
14000000	1600000	0	4713294	-2268538	442998	1332653	2684	-502350	1564	0	0
14000000	0	0	4711649	-1680243	419969	398771	117555	-472966	5519	-293960	0
10000000	746170	0	49137924	119700588	821980	17463947	-8652785	-7182474	365282	4339847	440610

3-1. 데이터 수집 & 전처리

수집 전제

TS2000

1. 6개년(16-21) 데이터가 모두 존재할 것
2. 코스닥&제조업인 기업
3. 재고자산이 0이 아닌 기업
4. 매출원가가 0이 아닌 기업
5. 매입채무가 0이 아닌 기업



비제조업일 가능성이 높기 때문

3-1. 데이터 수집 & 전처리

Y: 부실 여부(1/0)

$$K1 \text{ score} = -17.862 + 1.472X1 + 3.041X2 + 14.839X3 + 1.516X4$$

X1 : log (총자산) ← 기업규모

X2 : log (매출액/총자산), 즉 loge (총자산회전율) ← 회전율

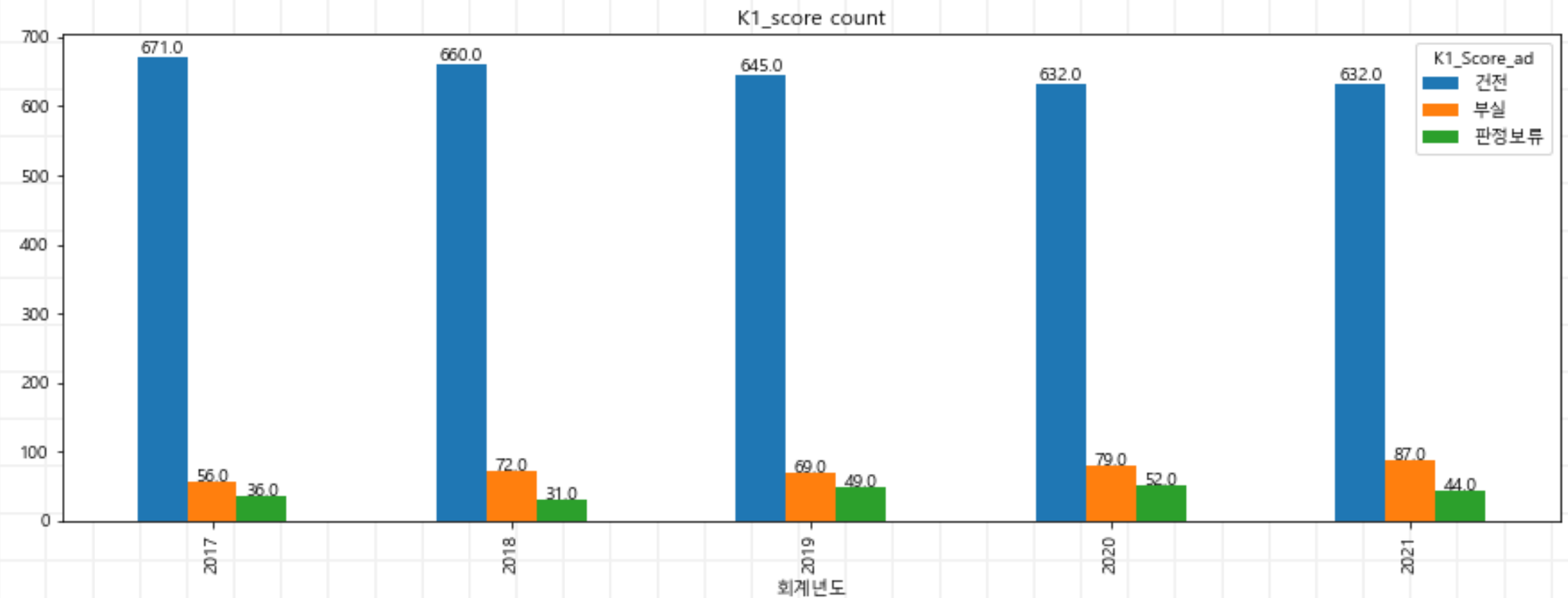
X3 : 이익잉여금/총자산 ← 누적 수익성

X4 : 자기자본/총부채 (즉, 부채비율의 역수) ← 재무구조

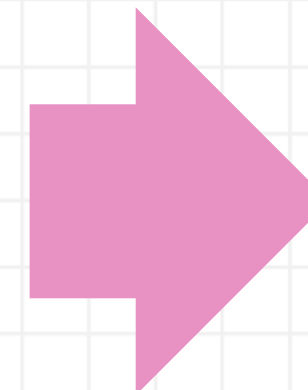
K1의 범위	판정내용
$0.75 < K1$	부실가능성은 없음.
$-2.0 < K1 < 0.75$	판정유보
$K1 < -2.0$	부실가능성이 심각

3-1. 데이터 수집 & 전처리

Y: 부실 여부(1/0)



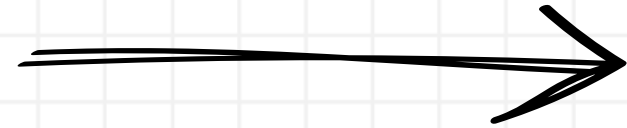
	2017	2018	2019	2020	2021
건전	671	660	645	632	632
부실	56	72	69	79	87
판정보류	36	31	49	52	44



보수적인 평가를 위해
판정보류를
'부실 가능성 있음'에
포함하기로 결정

3-1. 데이터 수집 & 전처리

X Features

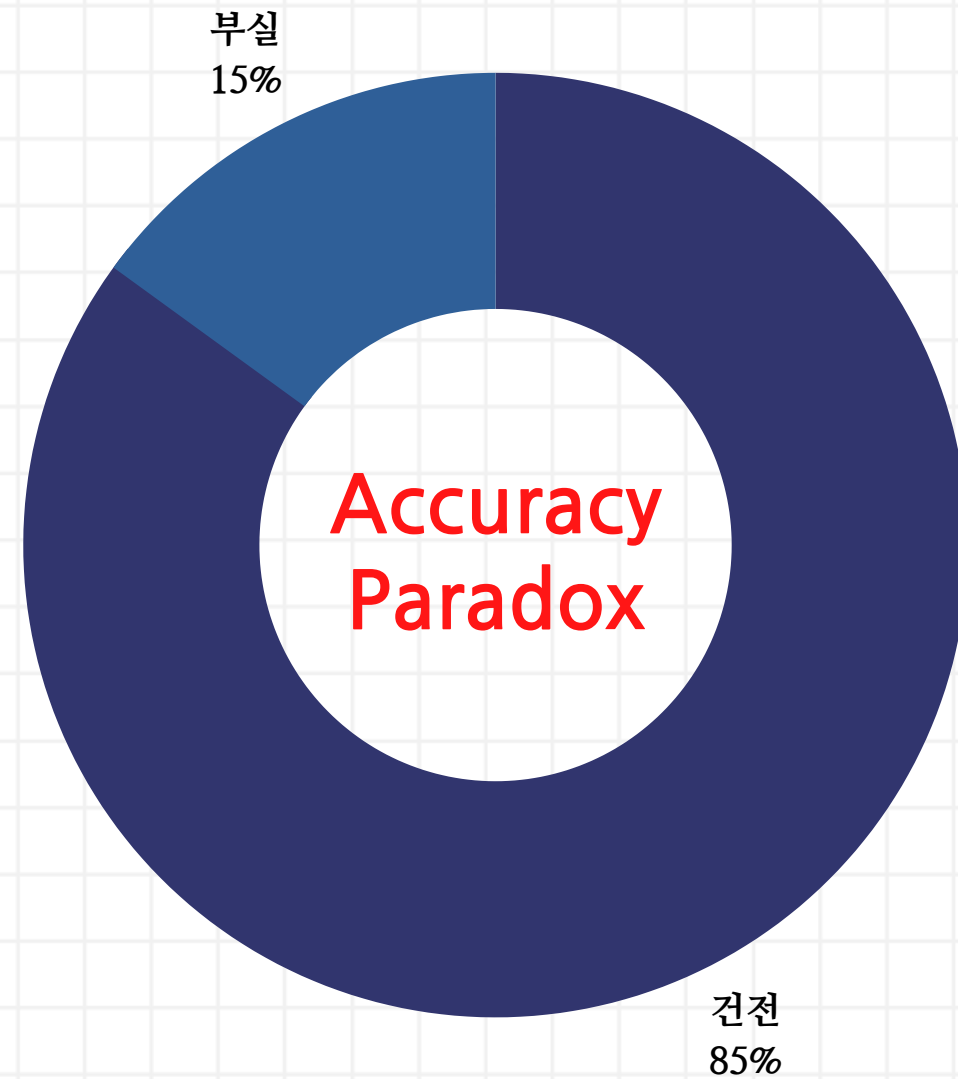


Standard Scaler 처리

- new_복리후생비 -> 복리후생비 / 영업이익
- new_기부금 -> 기부금 / 영업이익
- 기업존속연수 -> 회계년도 - 설립일
- new_시총 -> 시총 / 자본
- new_평균근속연수 ->
$$\frac{(\text{평균근속연수} / \max(\text{평균근속연수}))}{((\text{기업존속연수} / \max(\text{기업존속연수})) + (\text{상장년수} / \max(\text{상장년수}))}$$
- new_기업존속연수 = 기업존속연수 / max(기업존속연수)
- new_상장년수 = 상장년수 / max(상장년수)
- 자산 -> 자산 / 기말인원
- 교육훈련비 -> 교육훈련비 / 영업이익
- 경상개발비 -> 경상개발비 / 영업이익

3-1. 데이터 수집 & 전처리

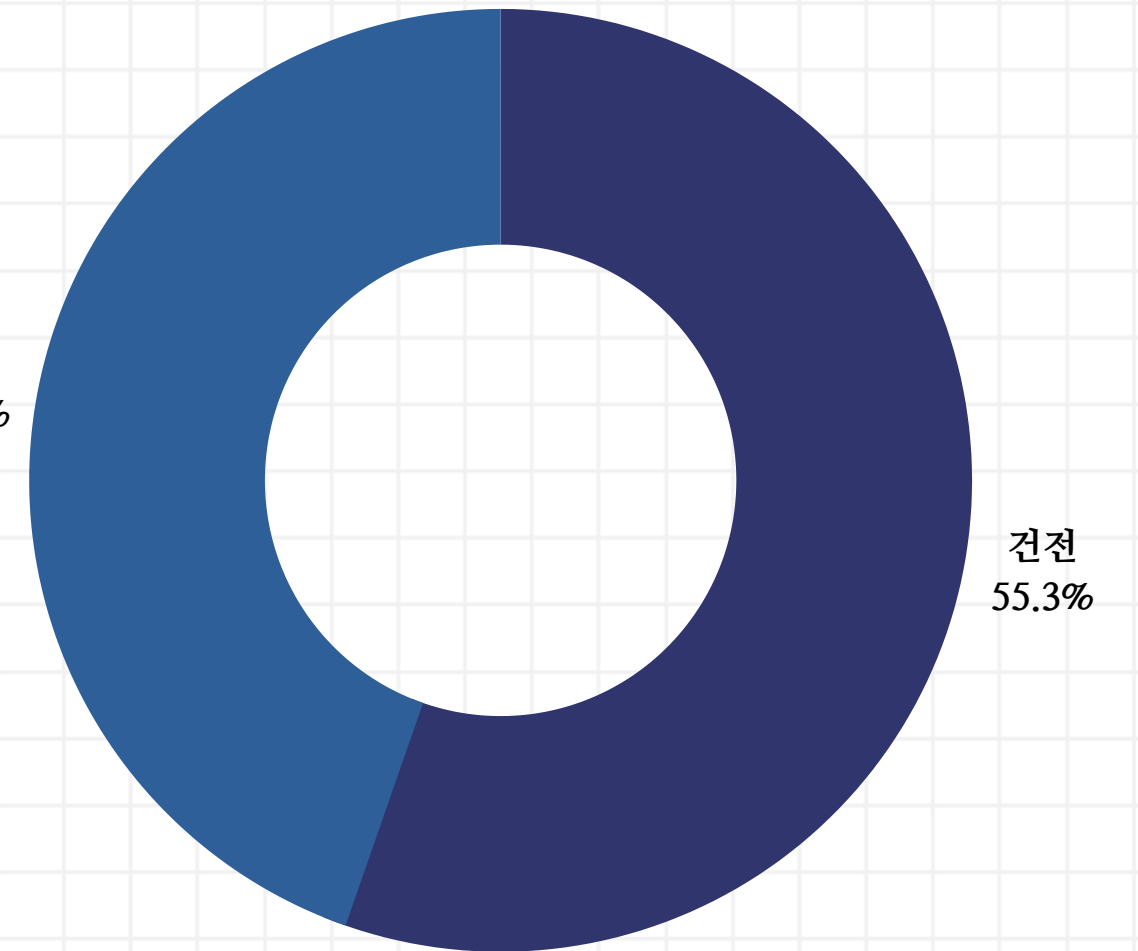
Under Sampling



Imbalanced data

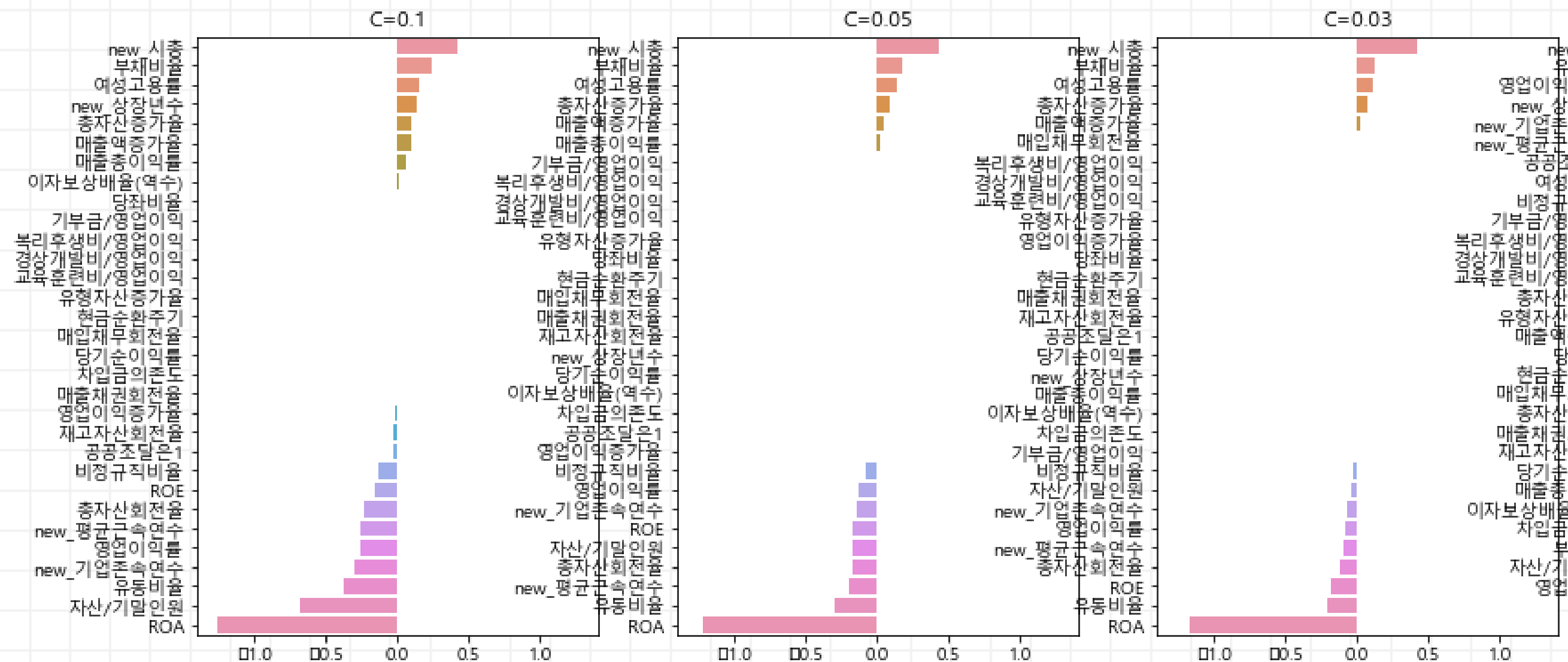


부실
44.7%



Balanced data

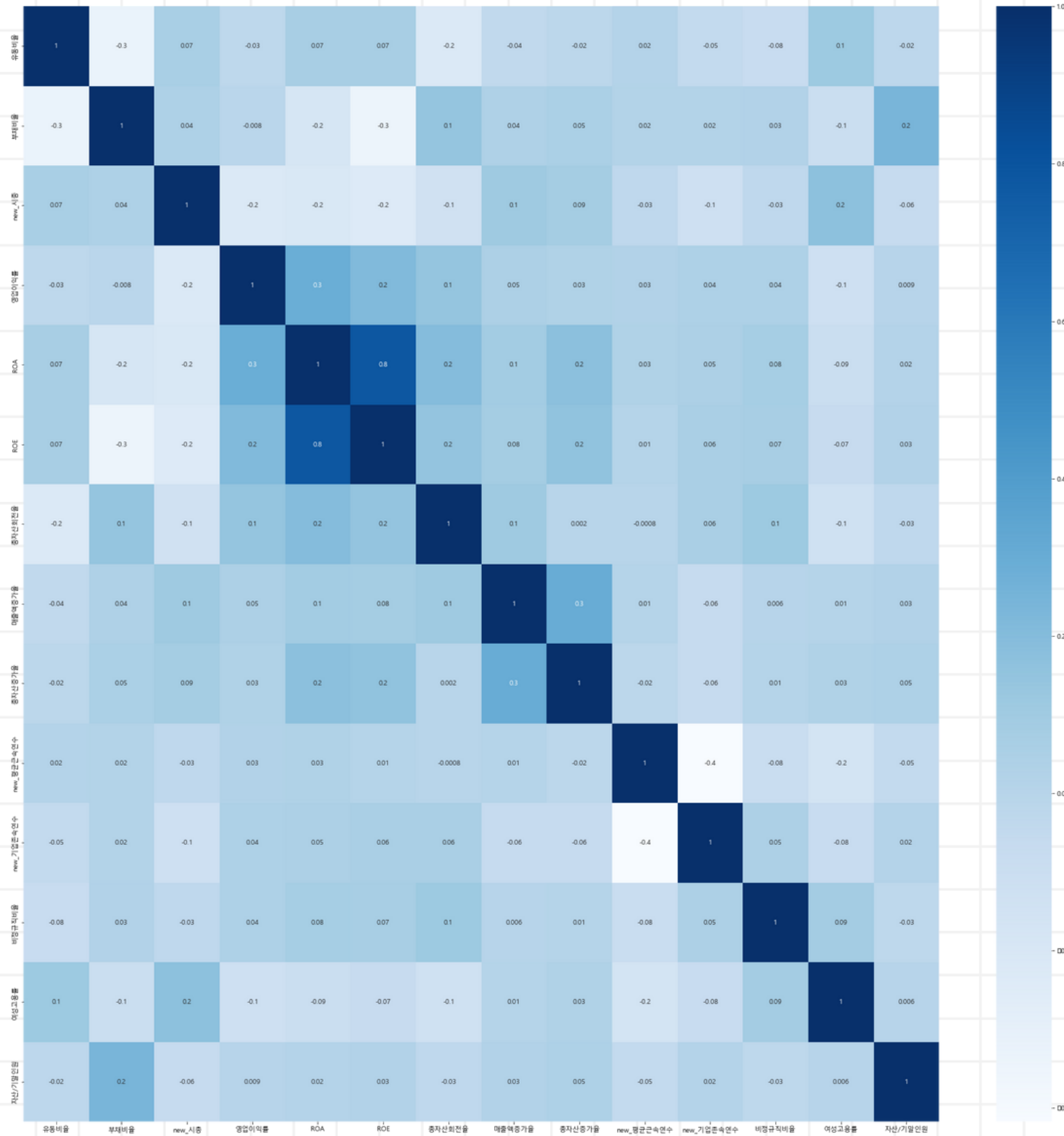
3-2. Feature Selection -Lasso



C = 0.03

new_시총	0.416428
부채비율	0.123789
여성고용률	0.115295
총자산증가율	0.067970
매출액증가율	0.021619
비정규직비율	-0.033397
자산/기말인원	-0.039737
new_기업존속연수	-0.075034
영업이익률	-0.077519
new_평균근속연수	-0.093421
총자산회전율	-0.123895
ROE	-0.182513
유동비율	-0.216182
ROA	-1.173990

3-2. 다중공선성 -Heatmap



ROE & ROA: 상관관계 높음

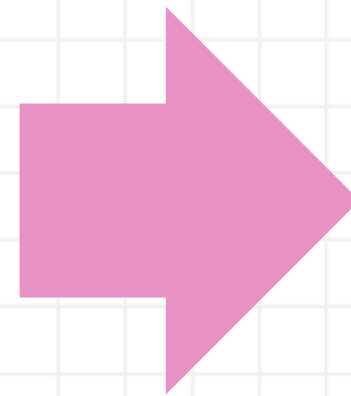
ROE = 당기순이익/자본
ROA = 당기순이익/자산

-> ROE 제거

3-2. 다중공선성 -Heatmap

* ROE 제거

	VIF Factor	features
0	1.044331	비정규직비율
1	1.109056	자산/기말인원
2	1.122338	여성고용률
3	1.137532	매출액증가율
4	1.143419	영업이익률
5	1.163315	총자산증가율
6	1.170764	총자산회전율
7	1.172344	new_시총
8	1.188396	유동비율
9	1.273630	new_기업존속연수
10	1.285346	new_평균근속연수
11	1.468526	부채비율
12	2.884574	ROA
13	3.074571	ROE

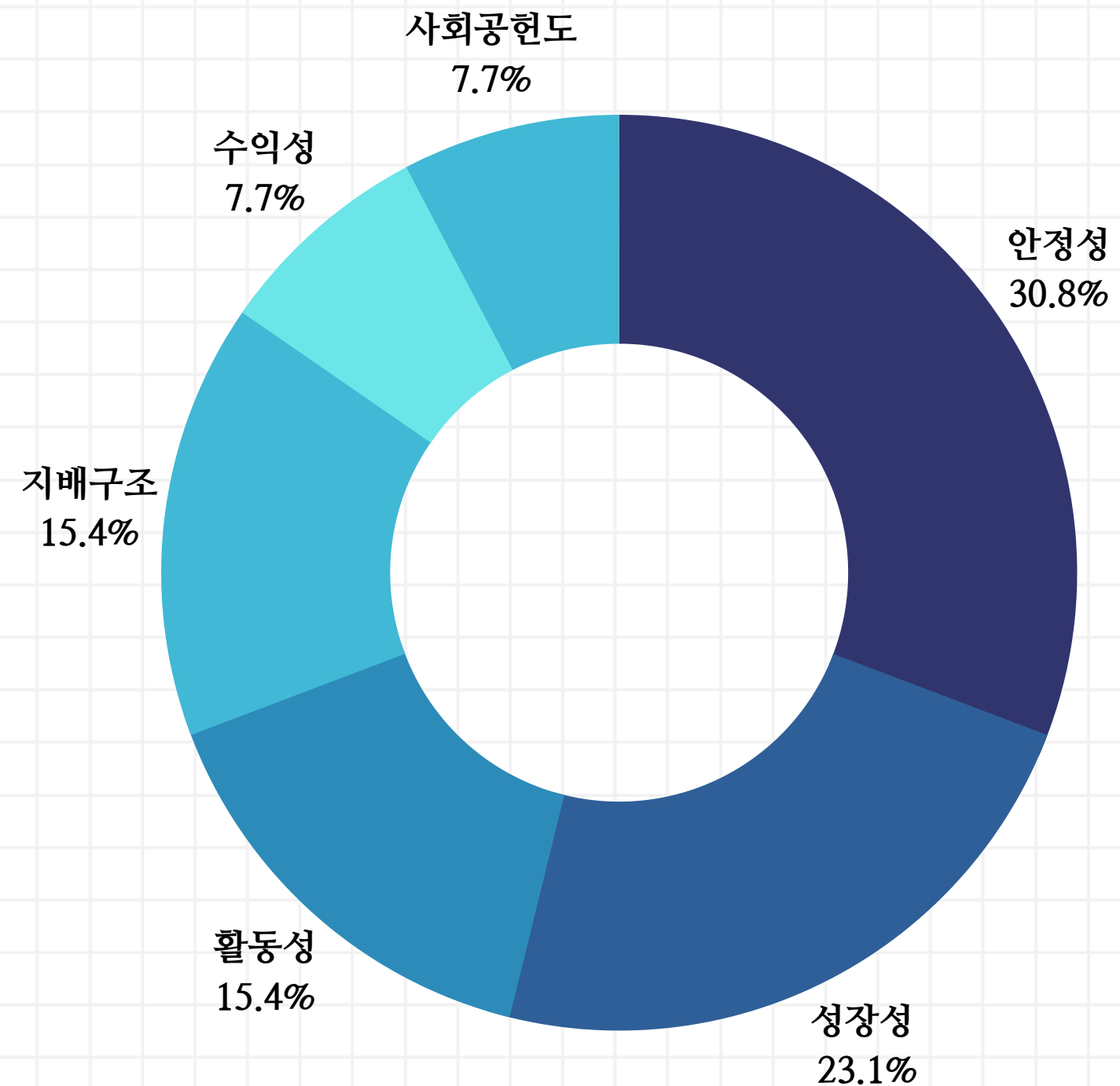


	VIF Factor	features
0	1.043772	비정규직비율
1	1.094175	자산/기말인원
2	1.122171	여성고용률
3	1.136847	매출액증가율
4	1.141735	영업이익률
5	1.156192	new_시총
6	1.157630	총자산증가율
7	1.169295	총자산회전율
8	1.181702	유동비율
9	1.272470	new_기업존속연수
10	1.285194	new_평균근속연수
11	1.286272	부채비율
12	1.288233	ROA

3-2. Feature Selection -Lasso

C=0.03

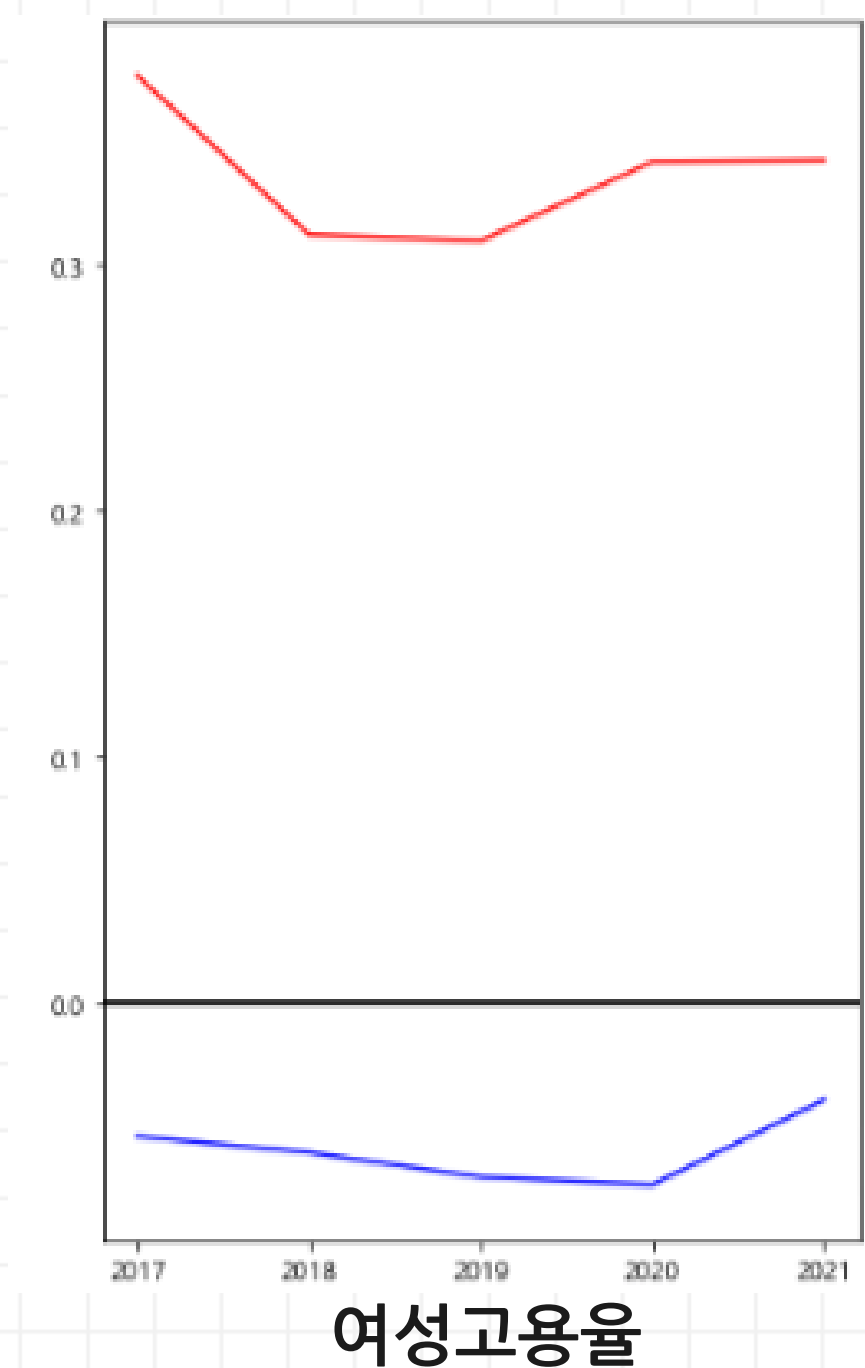
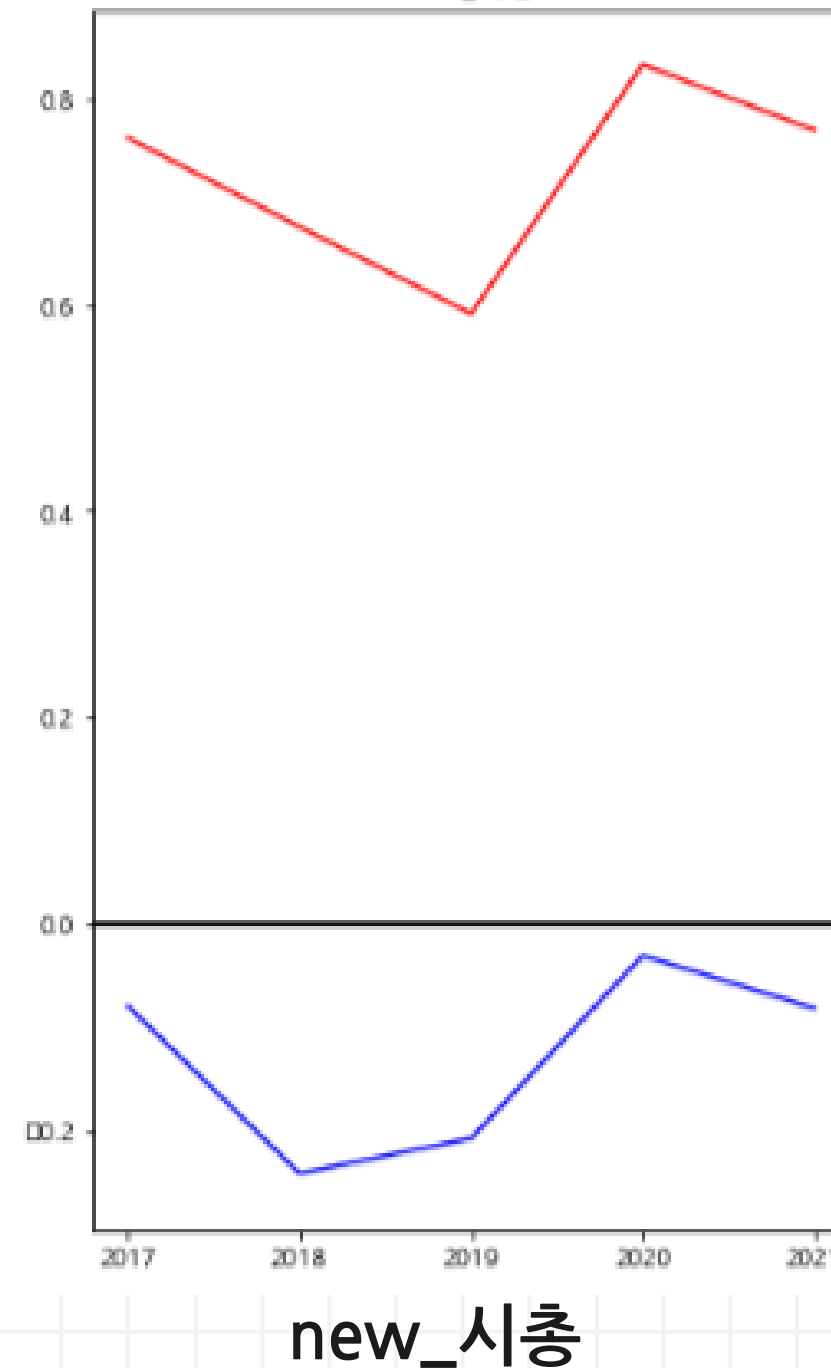
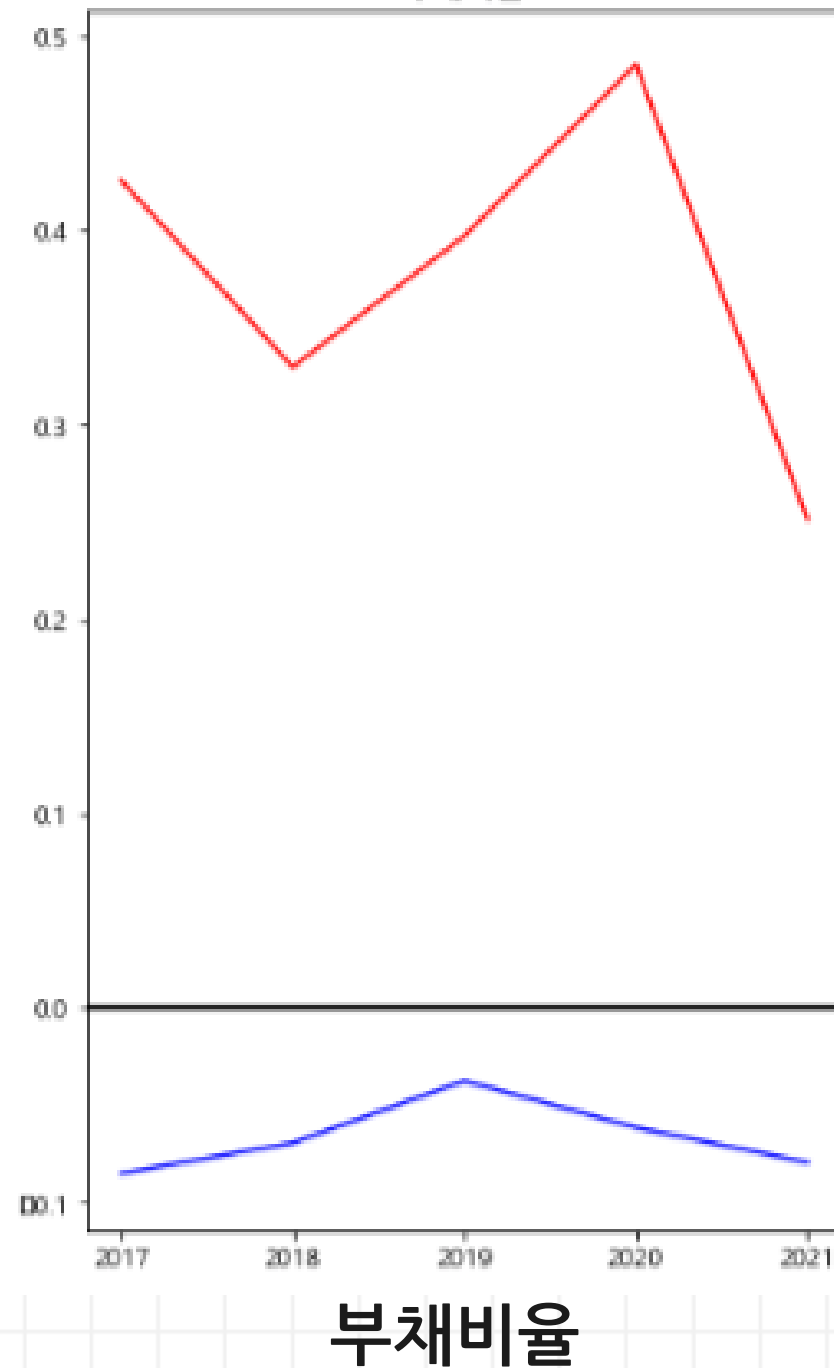
new_시총	0.422701
부채비율	0.167312
여성고용률	0.117814
총자산증가율	0.064107
매출액증가율	0.021515
비정규직비율	-0.034285
자산/기말인원	-0.056414
영업이익률	-0.074527
new_기업존속연수	-0.077219
new_평균근속연수	-0.091843
총자산회전율	-0.126583
유동비율	-0.211195
ROA	-1.300430



3-3. Feature Visualizing

부실 여부와 양의 상관관계(5개)

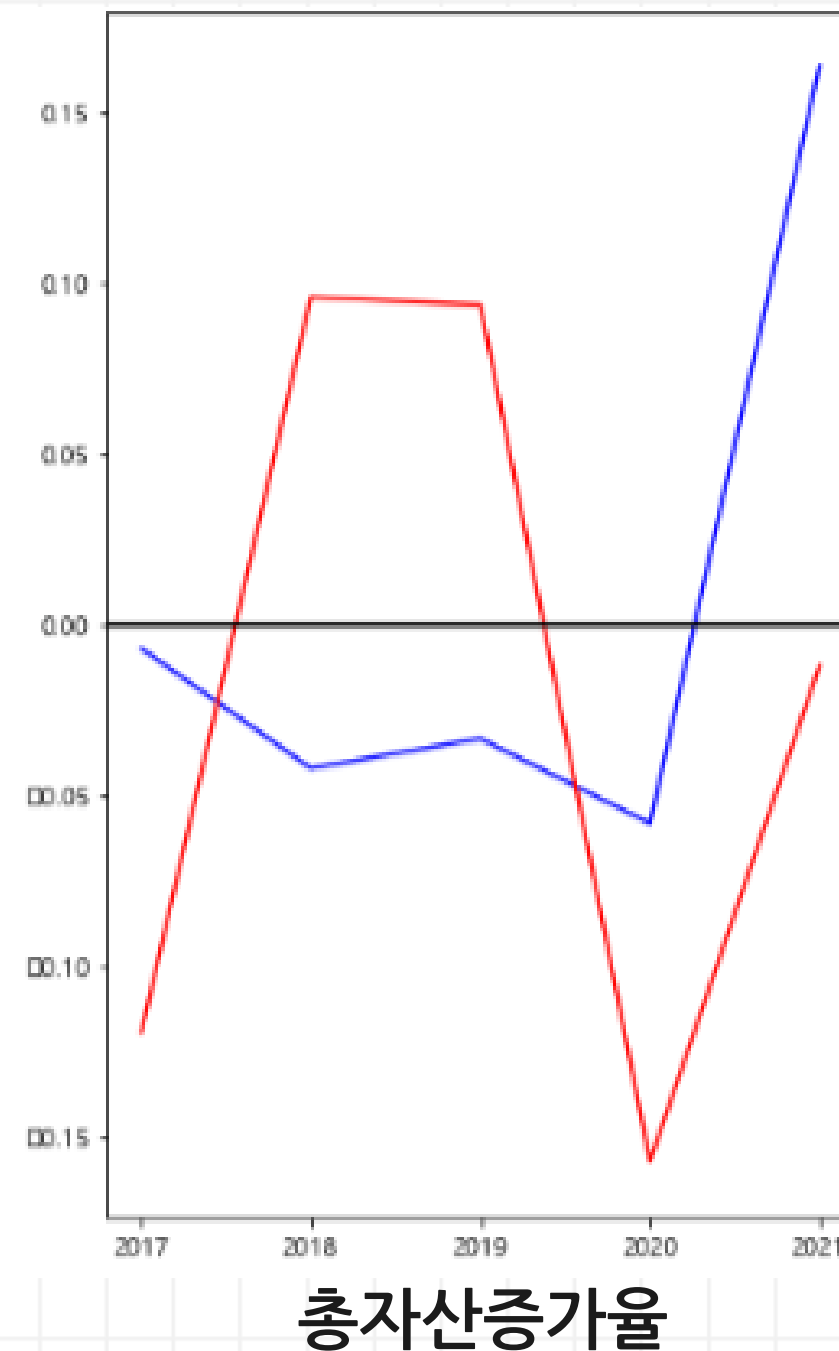
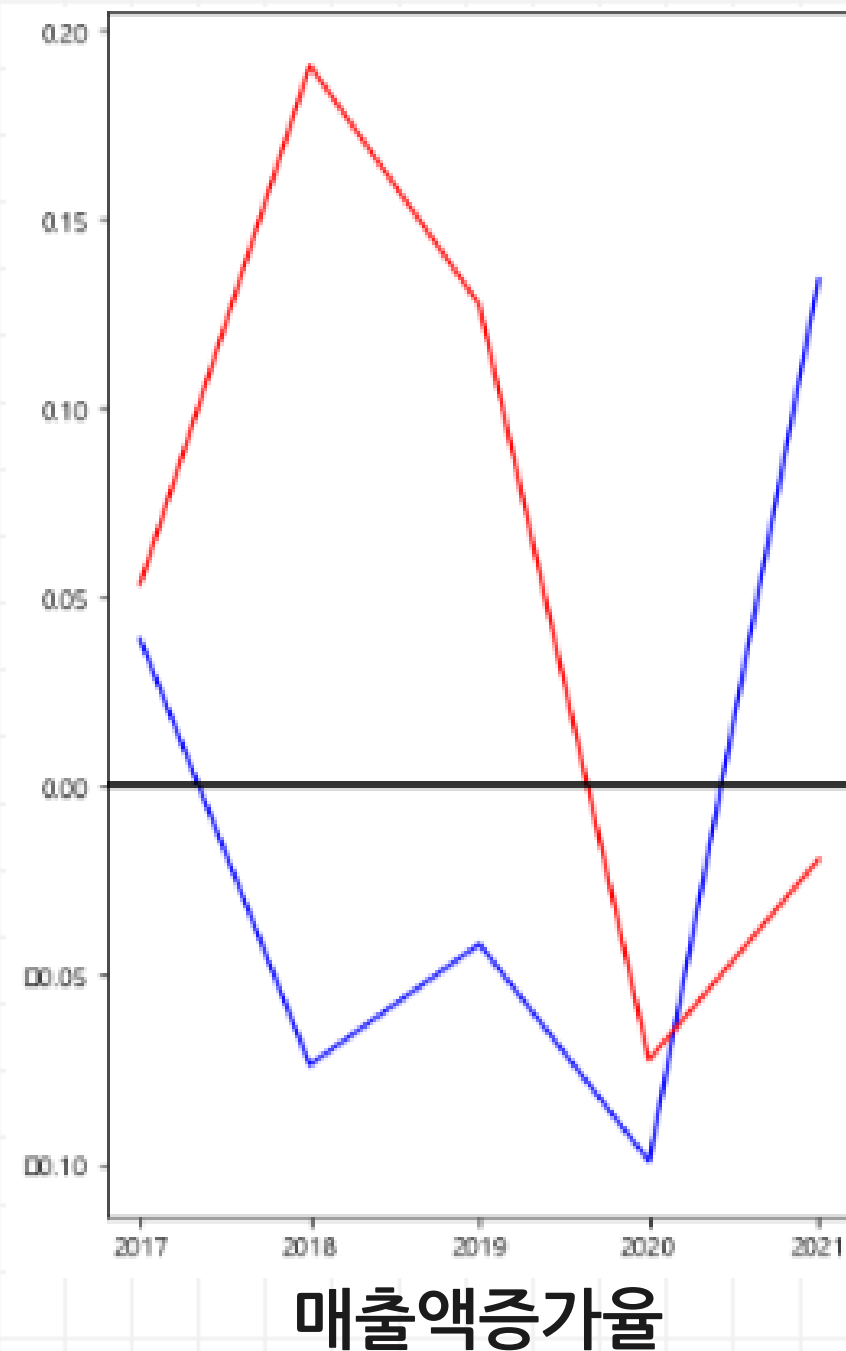
빨강: 부실
파랑: 건전



3-3. Feature Visualizing

부실 여부와 양의 상관관계

빨강: 부실
파랑: 건전



3-3. Feature Visualizing

	회사명	회계년도	차입금의존도	부채비율	매출액증가율	총자산증가율	부채
585	서울전자통신(주)	2017	29.731873	0.094892	-0.129552	-0.400914	37751998
586	서울전자통신(주)	2018	20.600580	0.449984	-0.297195	-0.110230	46830310
587	서울전자통신(주)	2019	14.305609	-0.422075	-0.522540	1.760695	43897434
588	서울전자통신(주)	2020	15.219142	-0.437042	-0.311634	0.174839	50177448
589	서울전자통신(주)	2021	22.359794	-0.214276	0.062921	-0.499021	61748671
610	(주)광무	2017	2.921994	0.220804	-0.461111	-0.348921	15398329
611	(주)광무	2018	19.836979	0.716197	-0.066660	-0.388349	17702532
612	(주)광무	2019	9.183240	0.826137	0.718411	3.330873	40295715
613	(주)광무	2020	15.361190	-0.301477	-0.653522	-1.500334	14403170
614	(주)광무	2021	4.784979	0.129626	-0.558893	4.750928	53084091
1035	(주)오스코텍	2017	12.549174	-0.315099	-0.022679	-0.354339	14891654
1036	(주)오스코텍	2018	5.341594	-0.527960	5.651559	1.865972	18589269
1037	(주)오스코텍	2019	3.334806	-0.555142	-1.356342	-0.243308	18027159
1038	(주)오스코텍	2020	1.648748	-0.552341	13.605154	-0.168876	19146658
1039	(주)오스코텍	2021	13.403267	-0.325189	-1.577744	-1.412369	17778799
1180	(주)상지카일룸	2017	4.799445	-0.383186	1.790443	2.548512	13658940
1181	(주)상지카일룸	2018	39.191759	0.924856	1.833799	6.138099	83106261
1182	(주)상지카일룸	2019	16.860190	1.475272	3.492412	0.294416	109458452
1183	(주)상지카일룸	2020	29.556558	1.081210	-1.394865	0.267394	124126763
1184	(주)상지카일룸	2021	33.858414	0.425608	-0.143455	-1.154928	77338625

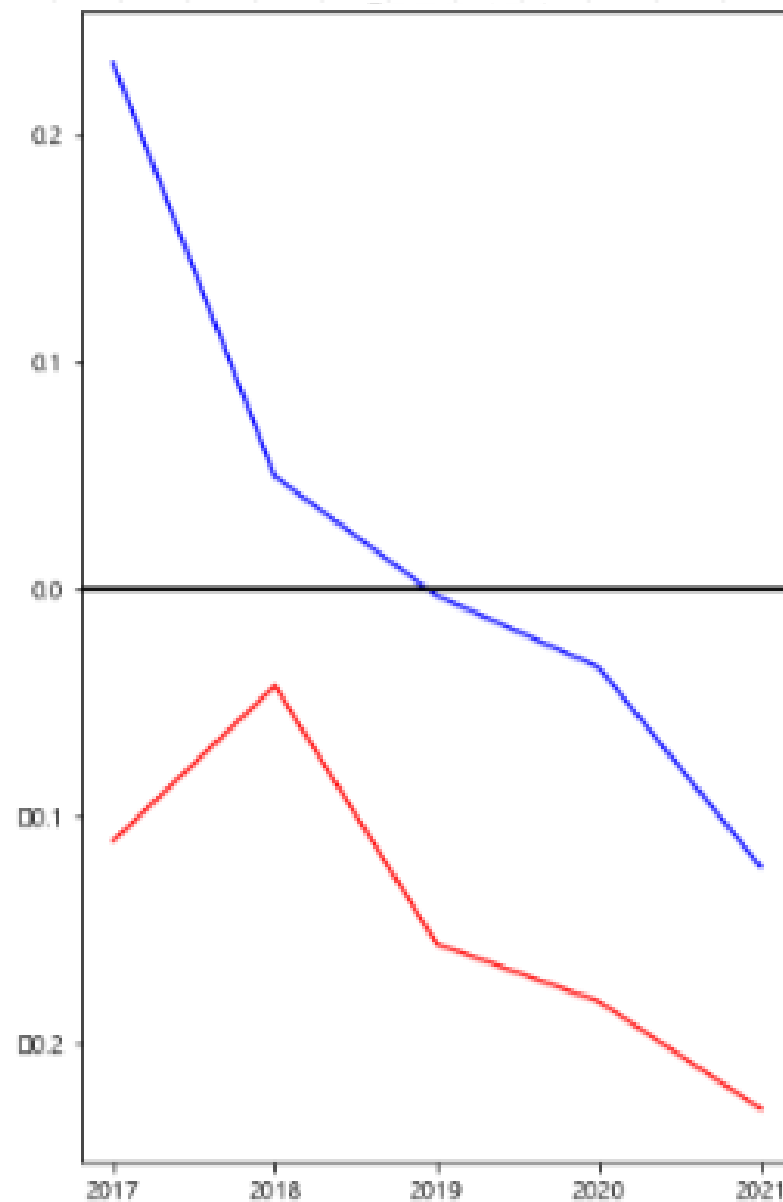
- 총자산증가율이 증가했을때
부실비율이 높아지는 것 확인
⇒ 급격한 부채 증가로 인한 자산
증가로 봄

- 이는 부채 창출로 인한

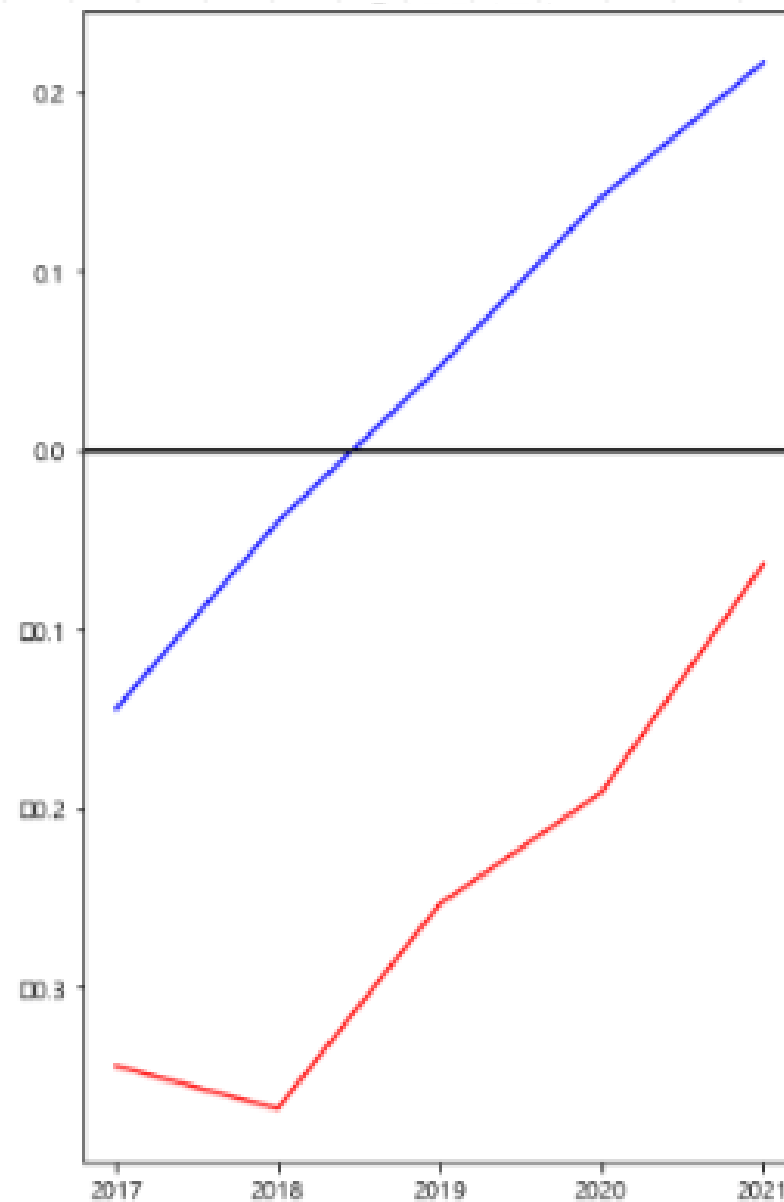
3-3. Feature Visualizing

부실 여부와 음의 상관관계(8개)

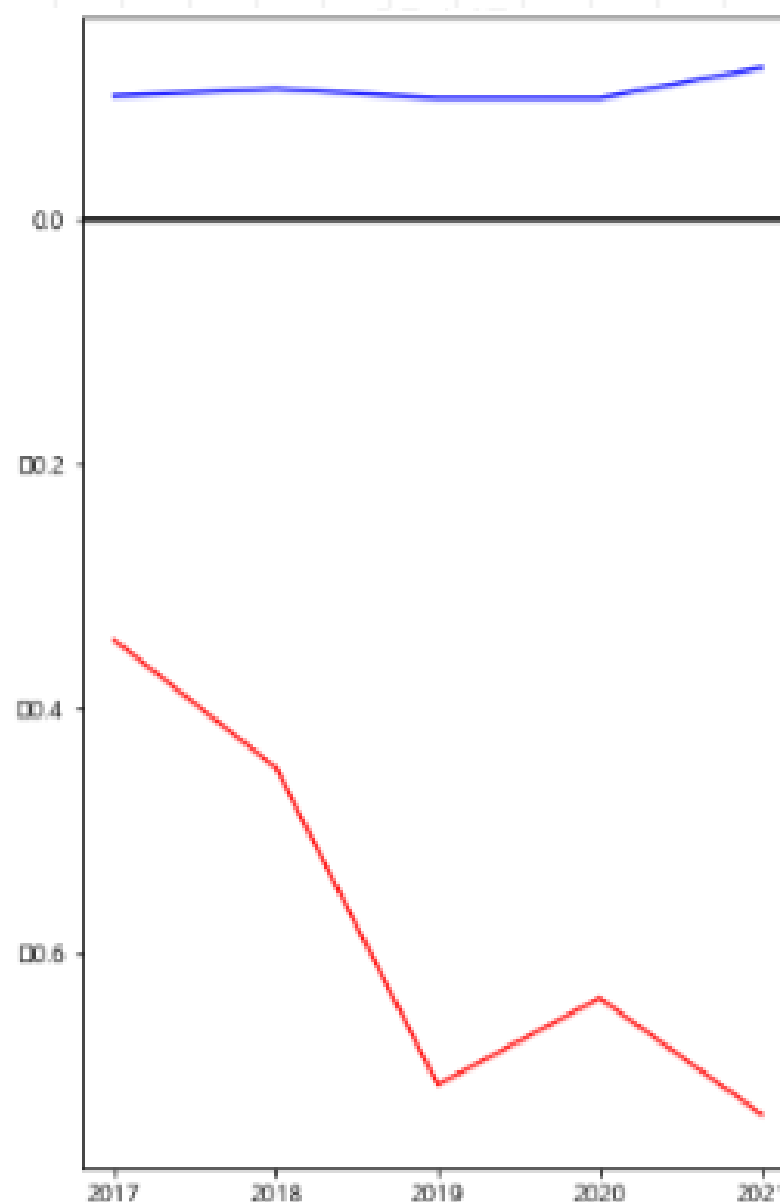
파랑: 건전
빨강: 부실



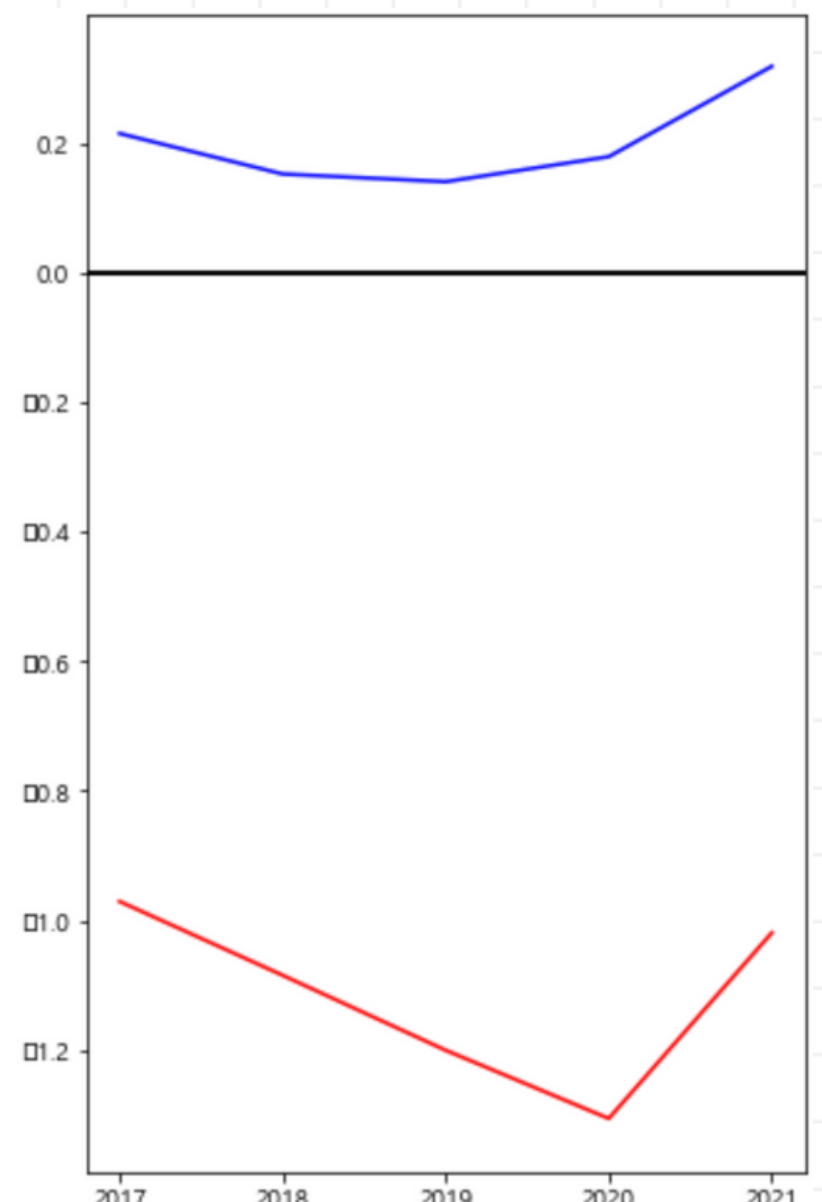
new_평균근속연수



new_기업존속연수



영업이익률

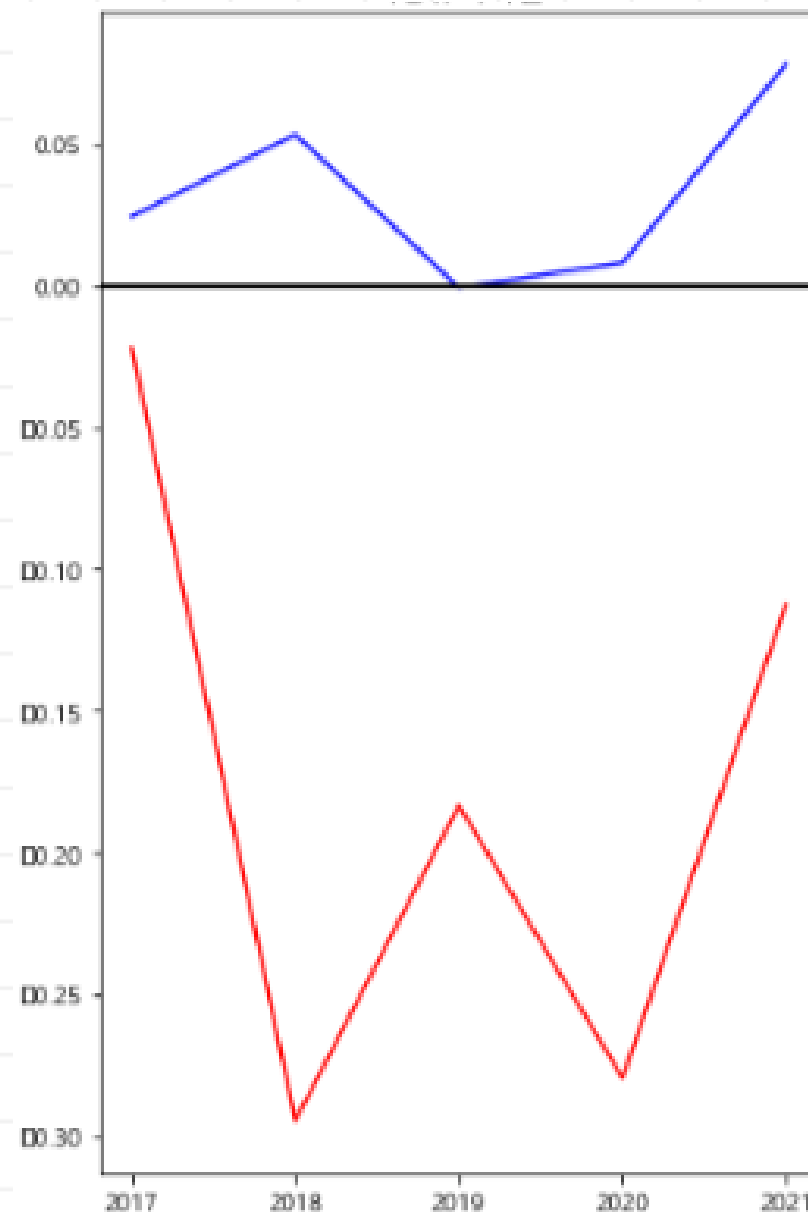


ROA

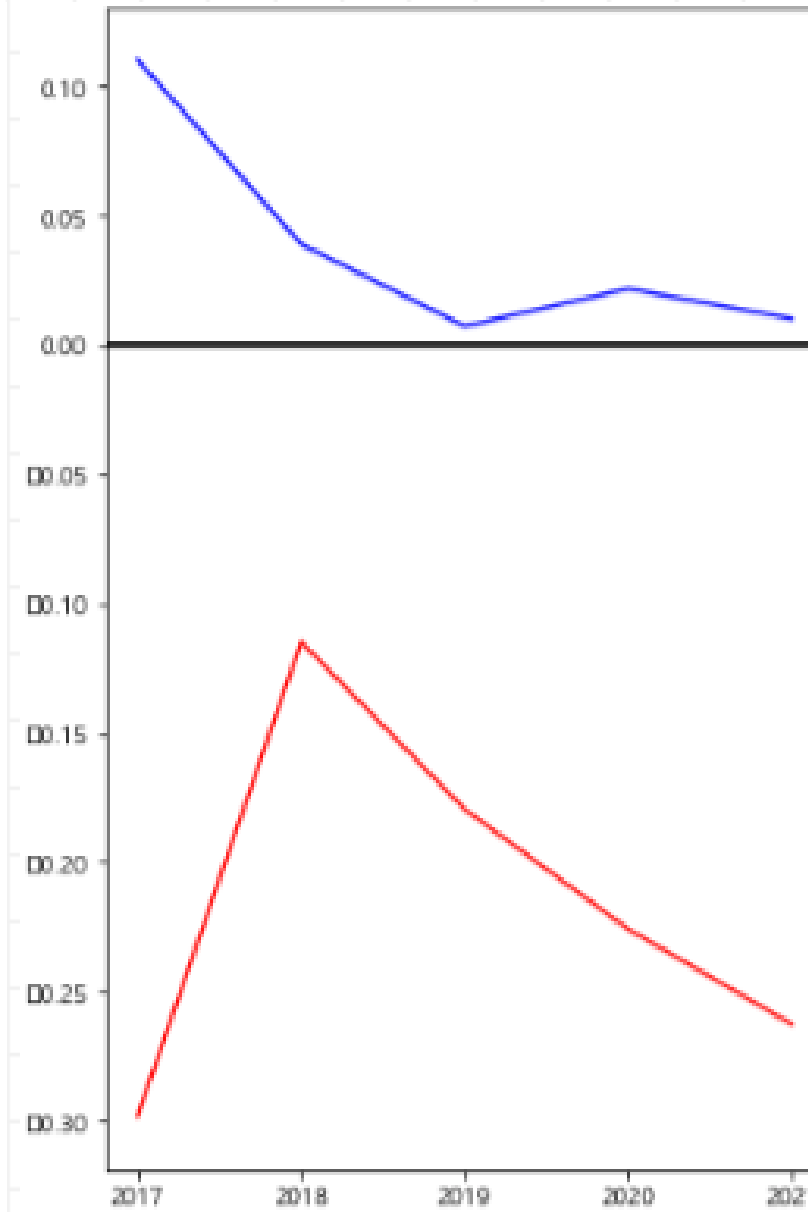
3-3. Feature Visualizing

부실 여부와 음의 상관관계

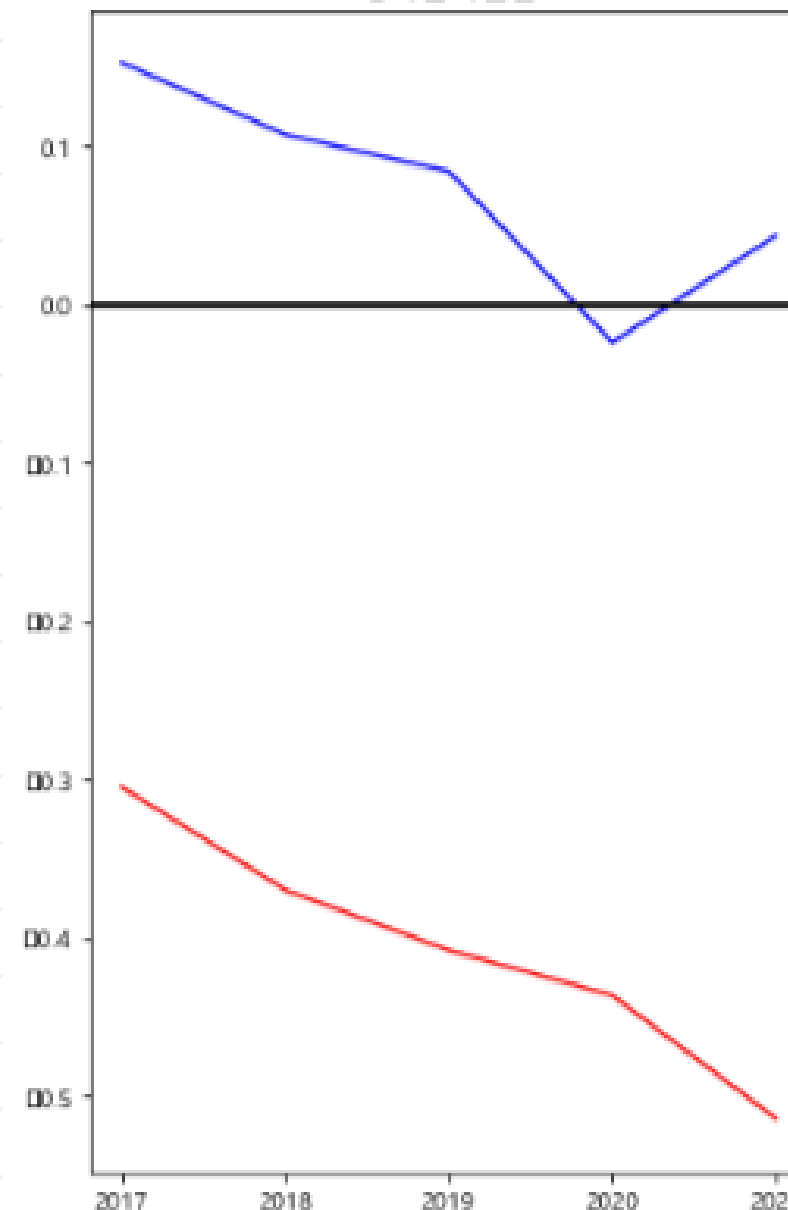
파랑: 건전
빨강: 부실



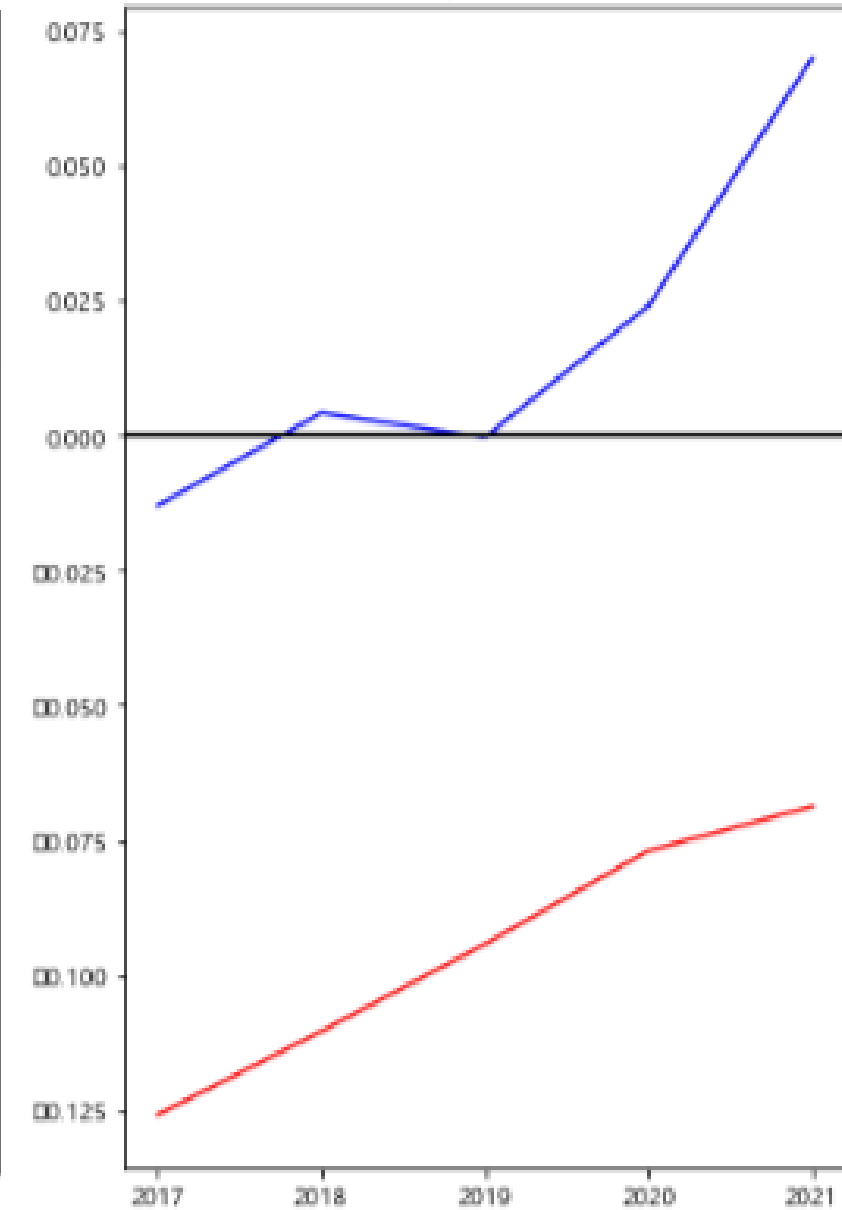
비정규직비율



유동비율



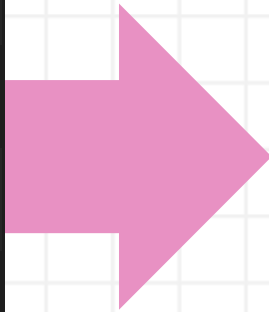
총자산회전율



자산/기말인원

4. 프로젝트 결과

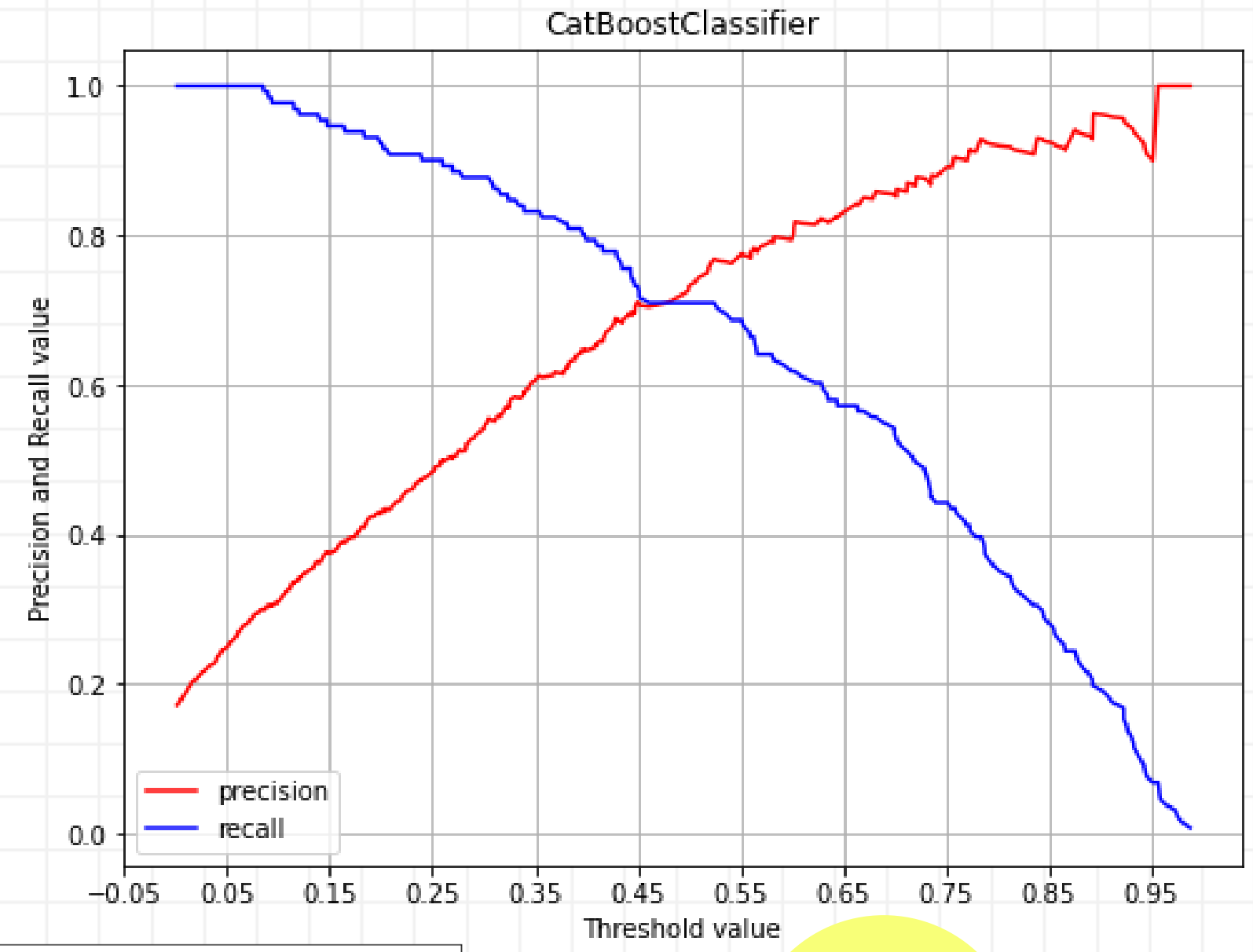
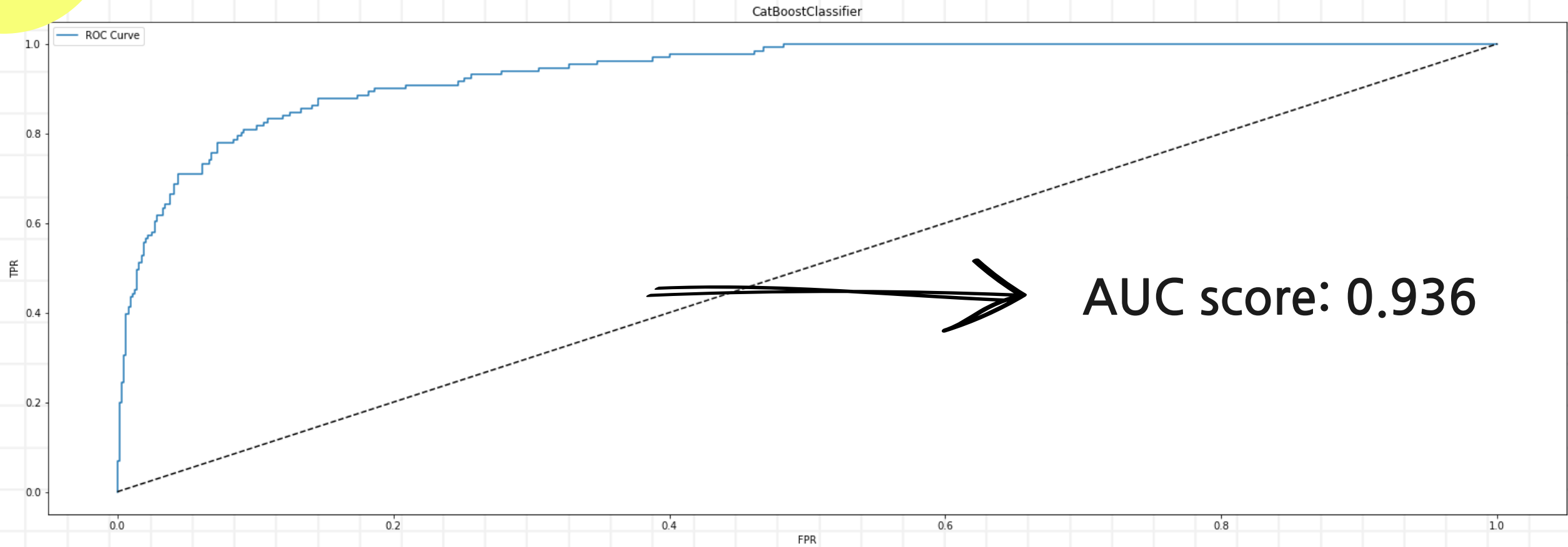
	threshold	Accuracy	Precision	Recall	F1 score	Auc_Score
models						
DecisionTreeClassifier	0.500000	0.813000	0.472477	0.786260	0.590258	0.875266
RandomForestClassifier	0.500000	0.898000	0.702290	0.702290	0.702290	0.926587
LogisticRegression	0.500000	0.889000	0.709091	0.595420	0.647303	0.889096
AdaBoostClassifier	0.500000	0.882000	0.651852	0.671756	0.661654	0.899398
GradientBoostingClassifier	0.500000	0.883000	0.647887	0.702290	0.673993	0.917595
XGBClassifier	0.500000	0.885000	0.650350	0.709924	0.678832	0.914352
LGBMClassifier	0.500000	0.889000	0.671642	0.687023	0.679245	0.921514
CatBoostClassifier	0.500000	0.907000	0.738095	0.709924	0.723735	0.936286
KNeighborsClassifier	0.500000	0.857000	0.575342	0.641221	0.606498	0.866273
GaussianProcessClassifier	0.500000	0.893000	0.709402	0.633588	0.669355	0.918748
GaussianNB	0.500000	0.843000	0.582090	0.297710	0.393939	0.804806
QuadraticDiscriminantAnalysis	0.500000	0.843000	0.582090	0.297710	0.393939	0.804806



임계값이 0.5일때
F1 score가 0.72로
가장 높은
"CatBoostClassifier"
선택

4. 프로젝트 결과

ROC Curve

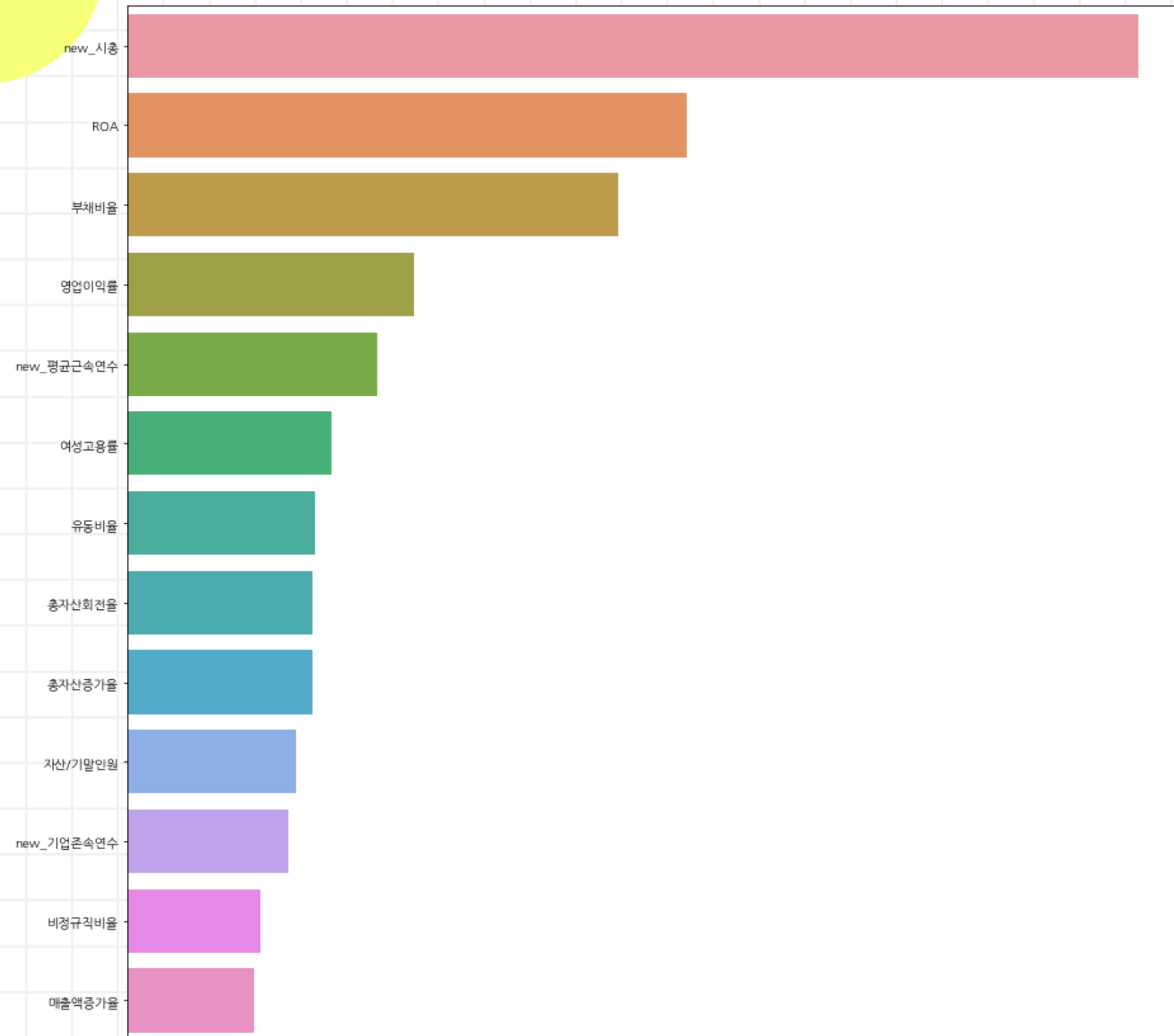


F1 Score

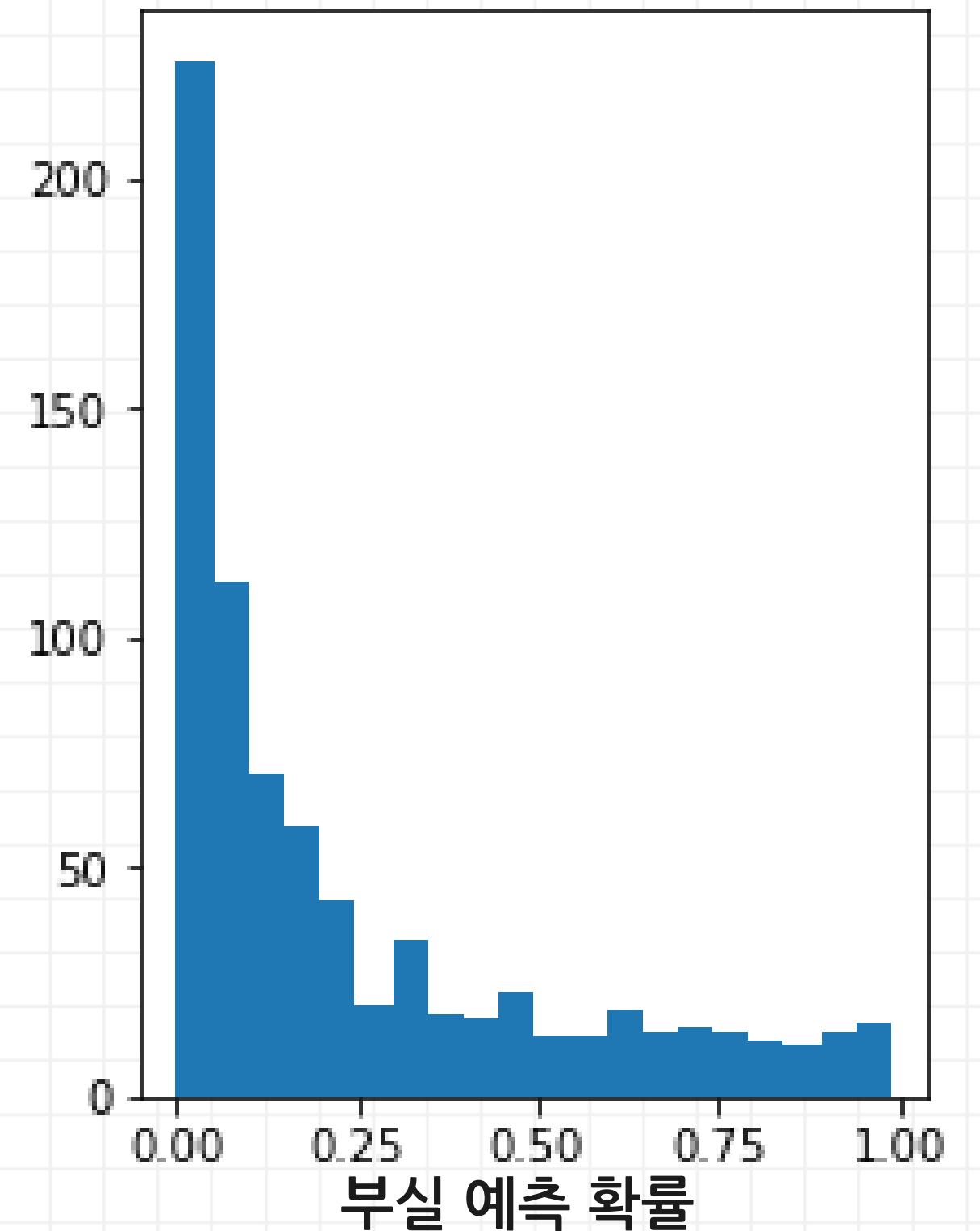
AUC score: 0.936

4. 프로젝트 결과

CatBClassifier - Feature Importance

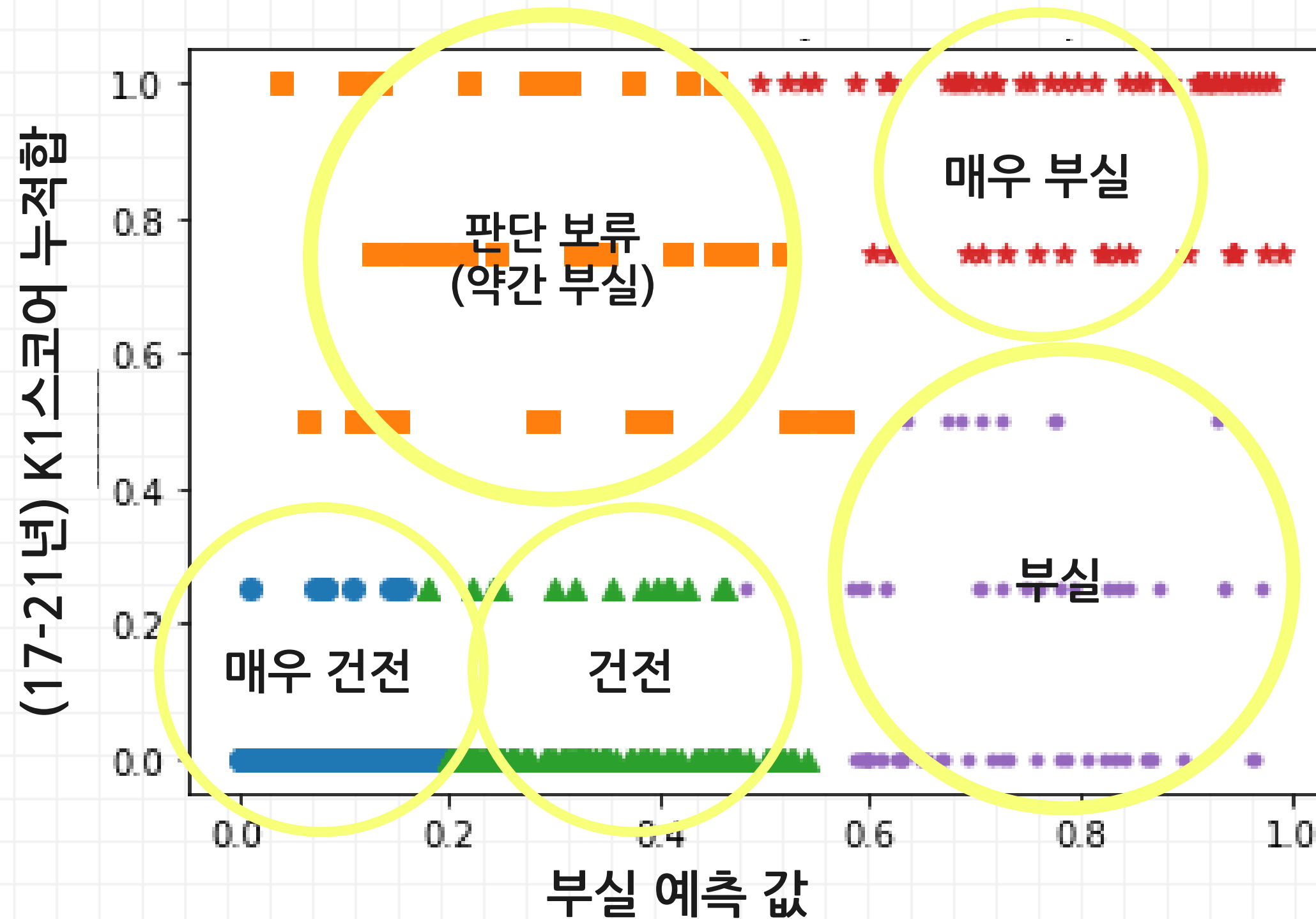


Test Data(21년도) - CatB에 모델링한
PredictProba Histogram



4. 프로젝트 결과

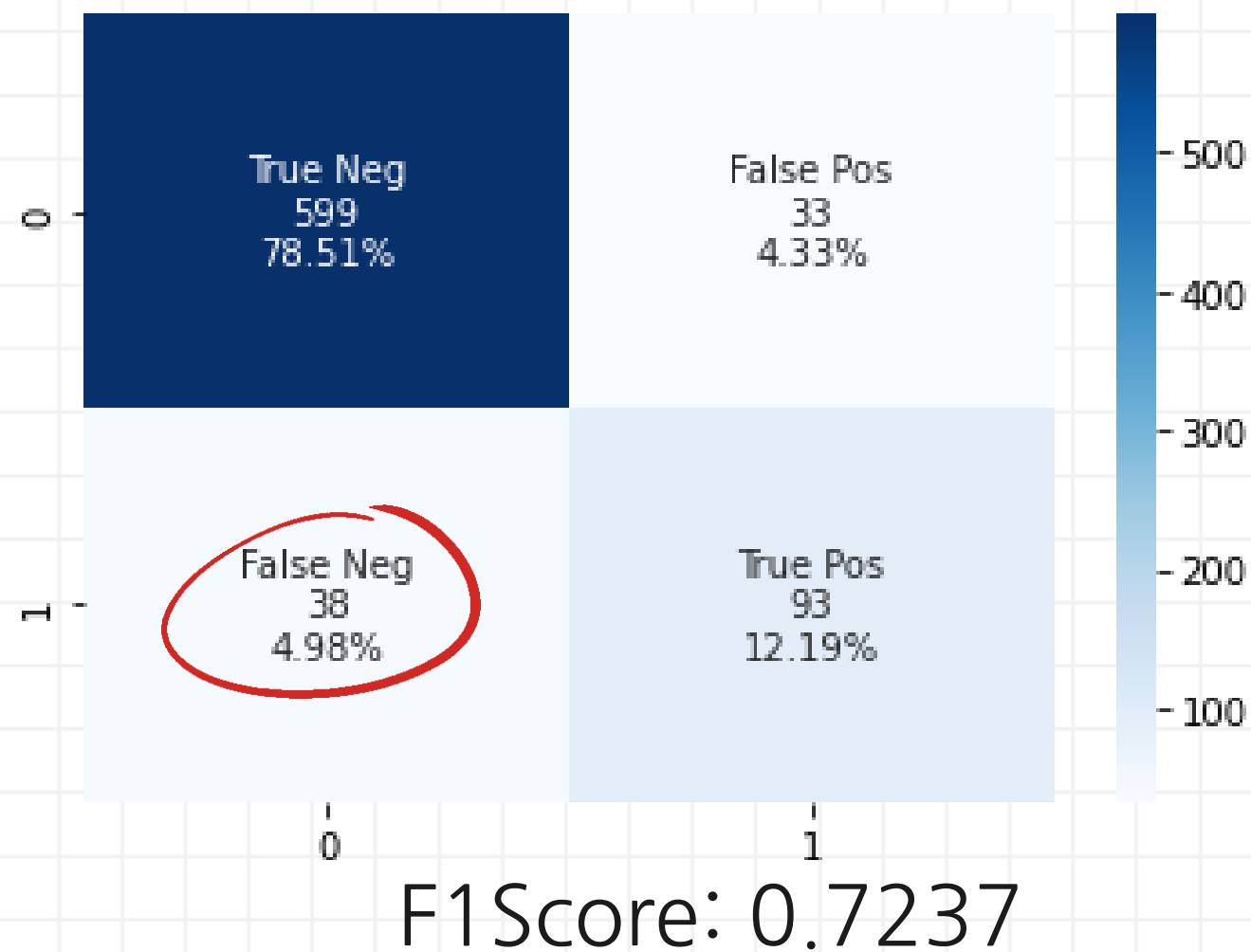
K-Means 알고리즘으로 등급화



4. 프로젝트 결과

등급화 - F1 Score 개선

K1-Score 예측 확률으로 도출한
confusion matrix



WeakTing등급으로 도출한
confusion matrix

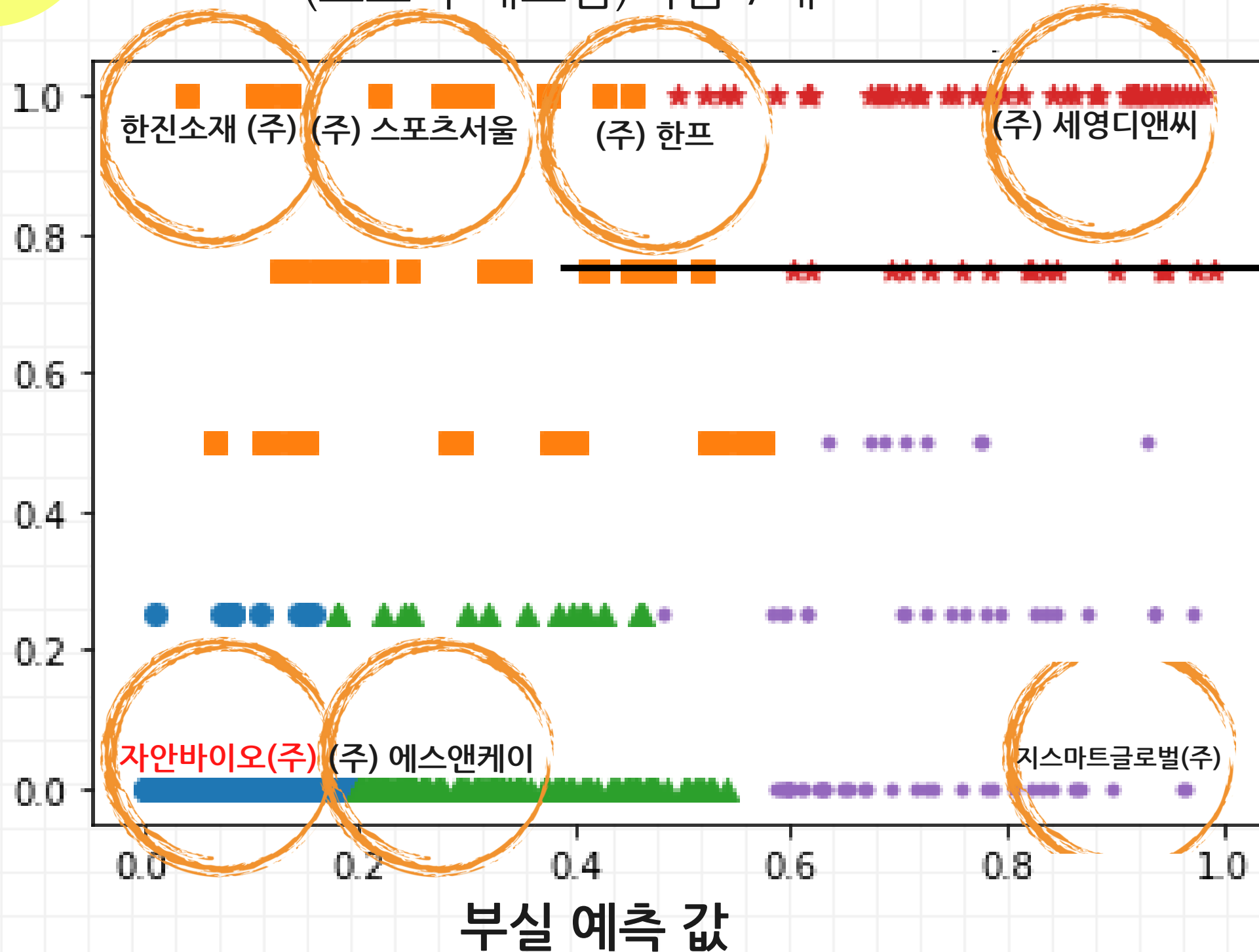


실제 부실하지만
건전하다고 판단한
FN 개수 감소

4. 프로젝트 결과

추가 검증- 22년에 상장폐지된
(코스닥 제조업)기업 7개

(17-21년) K1스코어 누적합



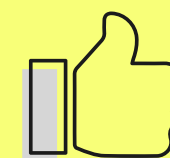
회사명	회계년도	Positive(1)	y_pred	K1_score_ad
(주)스포츠서울	2017	0.620841	1	1
(주)스포츠서울	2018	0.828641	1	1
(주)스포츠서울	2019	0.971738	1	1
(주)스포츠서울	2020	0.131930	0	1
(주)스포츠서울	2021	0.297513	0	1
(주)세영디앤씨	2017	0.947532	1	1
(주)세영디앤씨	2018	0.772481	1	1
(주)세영디앤씨	2019	0.527575	1	1
(주)세영디앤씨	2020	0.051616	0	1
(주)세영디앤씨	2021	0.934496	1	1
현진소재(주)	2017	0.625296	1	1
현진소재(주)	2018	0.385946	0	1
현진소재(주)	2019	0.074066	0	1
현진소재(주)	2020	0.045366	0	1
현진소재(주)	2021	0.017136	0	1
(주)한프	2017	0.829824	1	1
(주)한프	2018	0.961500	1	1
(주)한프	2019	0.510316	1	1
(주)한프	2020	0.973974	1	1
(주)한프	2021	0.437978	0	1
지스마트글로벌(주)	2017	0.284653	0	0
지스마트글로벌(주)	2018	0.400423	0	0
지스마트글로벌(주)	2021	0.903842	1	1
자안바이오(주)	2017	0.198922	0	0
자안바이오(주)	2018	0.325675	0	0
자안바이오(주)	2019	0.115120	0	0
자안바이오(주)	2021	0.164780	0	1
(주)에스앤케이	2019	0.007691	0	0
(주)에스앤케이	2020	0.017798	0	0
(주)에스앤케이	2021	0.270079	0	0

4. 프로젝트 결과

활용 - 기업 추천

기준

1. 산출한 등급: 완전건전, 완전부실 (우리가 판단한 부실등급)
2. 베타 1미만 ,1이상 (베타1미만은 경기를 잘안탐, 1이상이면 경기를 탄다.)
3. 확률 낮은순, 높은순 (우리가 만든 부도확률)
4. 평점 3.2이상, 평점 2.5미만 (리뷰가 최소 50개이상인 기업들만 선정)



한국알콜산업(주)
대양전기공업(주)
(주)슈피겐코리아
(주)다우데이타
(주)원익머트리얼즈



(주)엔케이맥스
(주)나이벡
(주)펍트론
원풍물산(주)
휴림로봇(주)

한국알콜 저 PBR 종목, 투자 매력도 높아지는 2022년 실적 개선도 기대

이광희 기자 | 승인 2022.01.17 19:32 | 댓글 0



PBR 0.67배, PER도 4.98배로 업종 PER 73.17배 대비 저평가 상태

인기뉴스

5. 자체 평가 의견

한계

1. 거시경제지표 반영하지 못함
2. 동종업계 비교 반영하지 못함
3. 코스닥, 제조업에만 적용되는 한계
4. ESG 중 E(environment)반영 미비

의의

1. 새로운 피처를 반영했다는 점
2. k1 score의 3단계를 5단계로 세분화했다는 점
3. 시계열이 짧음에도 불구하고 일정 수준 이상의 F1 score를 도출했다는 점
4. 상대성을 최대한 반영하려고 노력했다는 점(비율, 전년대비 증가율)
5. 전년대비 증감율을 반영함으로써 데이터의 끈김을 방지한 점

5. 자체 평가 의견



추후 연구 방향

- 최소 10개년 데이터로 다시 연구할 것
- 적시성 해결을 위해 실시간 재무제표 업로드 및 열람이 가능한 사이트의 필요성
- 제조업 이외의 다른 업종도 반영한 모델을 개발할 것