



## [기후금융] 재무정보를 이용한 상장폐지 - 시차변수 추가



### 필수 과제

- 주가 데이터 제외해서 모델 돌려보기 → O
- 시차 변수 만들기 (Xt-1를 row로 붙여서) (상장폐지연도 Xt row는 제거하고 학습, Xt예측) → O
- 부도 예측, 부도확률예측
- auto ml



### 자율 과제

#### • 결측치 채우기

- Nan값이 하나라도 있는 결측치 row를 모두 drop하면, row개수: 111903 → 36084
- 모든 columns에서 Nan값인 row는 drop하되,  
일부 columns에서 Nan값인 row는 중앙값이나 평균값으로 fillna하는 것이 필요한가
- 데이터 개수가 크게 줄어들긴 하지만, 36084개로도 성능이 충분함

⇒ XGB에서 중요도가 높은 변수 ∩ 다른 변수들 간에 유의미한 상관관계가 있는 변수

•



### 고민해보기

- {상장유지를 상장폐지로 예측}인 예측오류가 예측정답보다 압도적으로 많음
  - 상장폐지 사유 중 '상장폐지 조건을 2년 계속 미충족할 경우'가 많으므로 이 부분을 이용한 파생변수 적절한 파생변수를 추가할 것
    - 상장연도~폐지연도 변화율 기울기 (1차함수 또는 지수함수 기울기값)
- ## ? . 변화율의 변화율 처리
  - 변화율의 변화율
- 지수함수 파라미터 기울기만

## ▼ 1. 전처리 & 변수 추가

### ▼ (1) 결측치 제거

- 재무정보 `[ 'cash/assets': 'net_income_grow' ]` 이 모두 Nan값인 row만 drop
  - 주가정보는 제외
  - 열 하나라도 Nan값이 아니라면 drop하지 않음
  - 상장 전이거나 상장폐지 이후 연도의 row 삭제 목적(기업이 존재하지 않을 때의 row)
- Shape (55876, 38)

```

code 0
cor_name 0
Accounting_Closing 245
year 0
cash/assets 1138
liabilities/assets 1137
liabilities-cash/assets 1137
sales/cur_liabilities 1138
cur_assets/cur_liabilities 1137
retained_earnings/cur_liabilities 1138
non_cur_lia/non_cur_lia+worth 1138
interest/cash_flow 1145
interest/ebitda 1155
inven/sales 1290
sales/net_income 1145
gross_margin/sales 1323
sales/retained_earnings 1147
change_in_roa 10427
change_in_AR 4021
change_in_TP 5440
assets 1137
sales 1132
roa 5819
sales_grow 6002
net_income 1132
net_income_grow 5855
closing_price 12609
revised_price 12609
market_capital 12609
return 15106
코드대분류 437
최초상장일자 61
상장일자 61
unlist_date 44822
unlist_reason 0
unlist_year 44822
unlist_reason_ctg 0
fnc_rsn_unlist 0
dtype: int64

```

## ▼ (2) 상장폐지 날짜에 대한 class형태 파생변수

	재무정보	연도	상장폐지연도
	cor_name	year	unlist_year
44788	대우중공업	2002	2001
44789	대우중공업	2003	2001
44790	대우중공업	2004	2001
44803	동아건설	2007	2001
44804	동아건설	2008	2001
...	...	...	...
56615	지디	2021	2018
56698	파티게임즈	2021	2020
56813	쿠첸	2020	2019
56814	쿠첸	2021	2019
56964	웨이포트	2018	2017

- 'year' 는 해당 연도의 재무정보
- 'unlist\_year' 는 제공되는 상장폐지 연도 정보
- 이미 상장폐지 연도가 지났음에도, 이후 재무정보 존재
- 즉, 마지막 결산재무정보연도와 상장폐지연도가 다름

## ▼ 상장폐지날짜 정보 기준 (panel\_a)

- 'year'=='unlist\_year' (회계결산월) ≥ (상장폐지 월)는 drop
- 상장폐지 판단에 반영되지 않은 정보이므로 제외

연도는 같아도  
(회계결산 월) ≠ (상장폐지 월)

재무정보 연도      상장폐지연도

year	Accounting_Closing	unlist_date	unlist_year
44722	2001	12.0	2001-05-23
44762	2001	6.0	2001-11-10
44807	2015	12.0	2015-04-15
44848	2015	12.0	2015-02-24
44865	1999	12.0	1999-05-25
...	...	...	...
55608	2021	12.0	2021-06-10
55621	2020	12.0	2020-05-14
55696	2019	12.0	2019-09-16
55711	2021	12.0	2021-06-11
55836	2017	12.0	2017-07-25

230 rows × 4 columns

- 달이 같아도 월말에 결산하므로 상장폐지 결정에 영향 X
- 예를들어,  
2022년 6월에 상장폐지가 결정되었다면, 2022년 12월 회계결산정보는 이 결정에 영향을 주지 못 함

연도는 같아도  
(회계결산 월) ≠ (상장폐지 월)

재무정보 연도      상장폐지연도

year	unlist_year	Accounting_Closing	unlist_mon	unlist_date
44722	2001	2001	12.0	5.0
44807	2015	2015	12.0	4.0
44848	2015	2015	12.0	2.0
44865	1999	1999	12.0	5.0
44937	2000	2000	12.0	6.0
...	...	...	...	...
55608	2021	2021	12.0	6.0
55621	2020	2020	12.0	5.0
55696	2019	2019	12.0	9.0
55711	2021	2021	12.0	6.0
55836	2017	2017	12.0	7.0

208 rows × 5 columns

- (회계결산월) > (상장폐지 월)
- 'Accounting\_Closing' > 'unlist\_mon'

- 최종 종속변수 'fnc\_rsn\_unlst\_year'

재무정보 연도      재무부실사유 유무      상장폐지연도 유무      재무부실사유 ∩ 상장폐지연도 유무      상장폐지연도

cor_name	year	fnc_rsn_unlst	unlist_last_year	fnc_rsn_unlst_year	unlist_date
54702	중국고섬	2008	1	0	2013-10-04
54703	중국고섬	2009	1	0	2013-10-04
54704	중국고섬	2010	1	0	2013-10-04
54705	중국고섬	2011	1	0	2013-10-04
54706	중국고섬	2012	1	1	2013-10-04
54707	SNK	2017	0	0	2022-05-18
54708	SNK	2018	0	0	2022-05-18
54709	SNK	2019	0	0	2022-05-18
54710	SNK	2020	0	0	2022-05-18
54711	SNK	2021	0	1	2022-05-18

종속변수

▼ 폐지 결정 후 마지막 재무정보 기준 (panel\_b)

- 상장폐지변수를 무시하고, 단순히 마지막 연도 재무정보 기준으로 라벨링 == '1'
- 최종 종속변수 'fnc\_rsn\_unlst\_year'

		재무정보 연도	재무부실사유 유무	마지막 연도 유무	재무부실사유 ∩ 마지막 연도 유무	상장폐지연도
	cor_name	year	fnc_rsn_unlst	unlst_last_year	fnc_rsn_unlst_year	unlst_date
55866	중국고섬	2008	1	0	0	2013-10-04
55867	중국고섬	2009	1	0	0	2013-10-04
55868	중국고섬	2010	1	0	0	2013-10-04
55869	중국고섬	2011	1	0	0	2013-10-04
55870	중국고섬	2012	1	1	1	2013-10-04
55871	SNK	2017	0	0	0	2022-05-18
55872	SNK	2018	0	0	0	2022-05-18
55873	SNK	2019	0	0	0	2022-05-18
55874	SNK	2020	0	0	0	2022-05-18
55875	SNK	2021	0	1	0	2022-05-18

▼ (3) 다음 해에 상장폐지가 되었는지 유무에 대한 변수 [종속변수]

- 일반적인 경우 → shift() 필요 X

		재무정보 연도	상장폐지 날짜	재무부실사유 ∩ 상장폐지연도 유무	상장폐지연도
cor_name	Accounting_Closing	year	unlst_date	fnc_rsn_unlst_year	unlst_reason
54690	중국고섬	12.0	2008	2013-10-04 2008년 재무정보: 2009년에 폐지 하지 않음 → 0	0 감사의견 의견거절
54691	중국고섬	12.0	2009	2013-10-04 2009년 재무정보: 2010년에 폐지 하지 않음 → 0	0 감사의견 의견거절
54692	중국고섬	12.0	2010	2013-10-04	0 감사의견 의견거절
54693	중국고섬	12.0	2011	2013-10-04 2011년 재무정보: 2012년에 폐지 하지 않음 → 0	0 감사의견 의견거절
54694	중국고섬	12.0	2012	2013-10-04 2012년 재무정보: 2013년에 폐지함 → 1	1 감사의견 의견거절
54695	SNK	7.0	2017	2022-05-18	0 상장폐지 신청('22.04.14)
54696	SNK	7.0	2018	2022-05-18	0 상장폐지 신청('22.04.14)
54697	SNK	7.0	2019	2022-05-18	0 상장폐지 신청('22.04.14)
54698	SNK	7.0	2020	2022-05-18	0 상장폐지 신청('22.04.14)
54699	SNK	7.0	2021	2022-05-18	0 상장폐지 신청('22.04.14)

즉, 이 변수는 다음해  
상장폐지 여부에 대한 변수

- 폐지된 당해 연도 재무정보가 존재하고, 상장폐지 결정에 반영된 경우
  - 상장폐지 연도 이전의 데이터로 폐지 예측을 하려고 하는 것이므로, 아래 경우 연도는 drop한다.

		재무정보 회계결산 월		재무정보 연도	상장폐지 날짜	재무부실사유 n	상장폐지연도 유무
cor_name	Accounting_Closing	year	unlist_date	fnc_rsn_unlst_year			
44743	해태제과	6.0	2001 → 2001-11-10	1			
47084	삼애인더스	6.0	2002 → 2002-10-22	1			
47955	인터피온반도체	6.0	2005 → 2005-12-28	1			
48302	웨스테이트	6.0	2012 → 2012-10-12	0			
50085	소예	6.0	2009 2009-11-03	1			
50428	플러스프로핏	6.0	2009 2009-11-20	1			
50569	한신코퍼	6.0	2004 2004-10-21	1			
51039	피케이엘	10.0	2005 2005-11-09	0			
51654	알에스넷	6.0	2012 2012-12-05	0			
51685	덴소풍성	3.0	2004 2004-08-14	1			

즉, 이 변수는 당해 연도 상장폐지 여부에 대한 변수

## ▼ 2. 데이터 변환

<결과비교>	1차	2차(n=임시1)	3차(n='all')	4차(n='all')
정확도	0.992	0.9191	0.9064	0.9031
상장폐지를 상장유지로 예측 (오답률 %)	12 / 10838 (0.1%)	22 / 509 (4%)	19 / 506 (3%)	23 / 506 (4%)
상장유지를 상장폐지로 예측 (오답률 %)	75 / 102 (70%)	27 / 97 (27%)	38 / 103 (36%)	36 / 103 (34%)

### ▼ (1) 원데이터 - 직전 연도 재무정보로 다음 해 상장폐지 예측

#### • 학습

- 독립변수 ['cash/assets': 'return'], ['code\_label'], ['year']
- 종속변수 ['fnc\_rsn\_unlst\_year']

			재무정보 연도	상장폐지 날짜	재무부실사유 유무	상장폐지연도
	cor_name	Accounting_Closing	year	unlist_date	fnc_rsn_unlst_year	unlist_reason
54690	중국고섬	12.0	2008	2013-10-04	2008년 재무정보: 2009년에 폐지 하지 않음 → 0	0
54691	중국고섬	12.0	2009	2013-10-04	2009년 재무정보: 2010년에 폐지 하지 않음 → 0	0
54692	중국고섬	12.0	2010	2013-10-04	2010년 재무정보: 2011년에 폐지 하지 않음 → 0	0
54693	중국고섬	12.0	2011	2013-10-04	2011년 재무정보: 2012년에 폐지 하지 않음 → 0	0
54694	중국고섬	12.0	2012	2013-10-04	2012년 재무정보: 2013년에 폐지함 → 1	1
54695	SNK	7.0	2017	2022-05-18		0
54696	SNK	7.0	2018	2022-05-18		0
54697	SNK	7.0	2019	2022-05-18		0
54698	SNK	7.0	2020	2022-05-18		0
54699	SNK	7.0	2021	2022-05-18		0

즉, 이 변수는 다음해  
상장폐지 여부에 대한 변수

#### • 결과

y_real	y_pred	
0	0	10826
1	1	12
1	0	75
	1	27

- 정확도: 0.9920
- 상장폐지를 상장유지로 예측: 12개
- 상장유지를 상장폐지로 예측: 75개

### ▼ (2) 원데이터 - 특정 기간 연도의 재무정보를 하나의 row로 나열

#### • 학습

- $X_t$  시점 재무정보 +  $X_{t-1}$  시점 재무정보 + ... +  $X_{t-n}$  시점 재무정보
- 독립변수 ['cash/assets', 'return'], ['code\_label'], ['year']
- 종속변수 ['fnc\_rsn\_unlst\_year']

$X_t$ = 마지막연도		$X_t$ 시점 재무정보		$X_{t-1}$ 시점 재무정보		기업코드	종속변수
		cash/assets_0	liabilities/assets_0	cash/assets_1	liabilities/assets_1	code_label	fnc_rsn_unlst_year
0		0.077683	0.270434	0.091490	0.285315	5930	0
1		0.074875	0.620626	0.053984	0.632119	373220	0
2		0.041813	0.270672	0.052476	0.354774	660	0
3		0.004491	0.284147	0.005949	0.373765	207940	0
4		0.071791	0.379642	0.090027	0.411737	6400	0
...		...	...	...	...	...	...

#### • 결과

y_real	y_pred	
0	0	487
1	1	22
1	0	27
	1	70

- 정확도: 0.9191
- 상장폐지를 상장유지로 예측: 22개
- 상장유지를 상장폐지로 예측: 27개

### ▼ (3) 데이터 변환 - 변화율, 증가율

- 학습

- $t$ 시점과  $t - n$  시점 변화율:  $(X_t - X_{t-1}) / (X_{t-1})$
- 독립변수 ['cash/assets': 'return'], ['code\_label'], ['year']
- 종속변수 ['fnc\_rsn\_unlst\_year']

$X_t$ 과 $X_{t-1}$ 시점 재무정보 변화율			기업코드	종속변수
	cash/assets_growth	liabilities/assets_growth	code_label	fnc_rsn_unlst_year
0	2.486240	-0.645929	5930	0
1	-0.279015	0.018519	373220	0
2	1.530986	-0.582872	660	0
3	-0.721334	0.627304	207940	0
4	0.556421	-0.334759	6400	0
...	...	...	...	...
3040	0.637586	0.454579	900180	1
3041	1.546060	-0.188975	950010	1
3042	-0.752784	2.667578	950030	1
3043	-0.910588	-0.076373	950070	1
3044	0.135113	-0.397054	950180	0

- 결과

y_real	y_pred	
0	0	487
	1	19
1	0	38
	1	65

- 정확도: 0.9064
- 상장폐지를 상장유지로 예측: 19개
- 상장유지를 상장폐지로 예측: 38개

#### ▼ (4) 데이터 변환: 최소제곱법을 바탕으로 한 지수함수 파라미터

- 학습

- $t$ 시점과  $t - n$  시점까지 최소자승법 지수함수 파라미터:  $a * \exp(bx)$
- 독립변수 ['cash/assets': 'return'], ['code\_label'], ['year']
- 종속변수 ['fnc\_rsn\_unlst\_year']

	$X_t$ 과 $X_{t-n}$ 시점까지 최소자승법 지수함수 기울기값		$X_t$ 과 $X_{t-n}$ 시점까지 최소자승법 지수함수 절편값		기업코드	종속변수
	cash/assets_a	cash/assets_b	liabilities/assets_a	liabilities/assets_b	code_label	fnc_rsn_unlst_year
0	0.026277	0.047779	0.910205	-0.044553	5930	0
1	0.074875	-0.327137	0.620626	0.018350	373220	0
2	NaN	NaN	1.034392	-0.043382	660	0
3	0.109713	-0.308367	0.323608	0.013360	207940	0
4	0.034340	0.040577	0.612500	-0.028288	6400	0
...	...	...	...	...	...	...
3040	0.106058	0.056724	0.274297	0.036408	900180	1
3041	0.150862	0.044876	0.389416	-0.056806	950010	1
3042	0.191154	-0.297272	0.304774	0.179470	950030	1
3043	0.233210	-0.530861	0.328443	-0.009892	950070	1
3044	0.595607	0.038371	0.147613	-0.194057	950180	0

#### • 결과

y_real	y_pred	
0	0	483
1	1	23
1	0	36
1	1	67

- 정확도: 0.9031
- 상장폐지를 상장유지로 예측: 23개
- 상장유지를 상장폐지로 예측: 36개

### ▼ 3. 데이터 범위

- $X_t \dots X_{t-n}$  에서 최적 예측성능을 내는  $n$  을 찾는 것이 목표

<n에 따른 결과비교>	1차	2차(n=1)	3차(n=5)	4차(n=3)
정확도	0.9392	0.9191	0.9097	0.9146
상장폐지를 상장유지로 예측 (오답률 %)	27 / 506 (5%)	(4.3%)	(4.2%)	(4.5%)
상장유지를 상장폐지로 예측 (오답률 %)	10 / 103 (9%)	(27.8%)	(33.0%)	(30.1%)

#### ▼ (1) 원데이터 사용해 최적범위 탐색: 마지막 회계결산 재무정보로만 예측

##### • 학습

- 독립변수 ['cash/assets': 'return'], ['code\_label'], ['year']
- 종속변수 ['fnc\_rsn\_unlst\_year']



	데이터 원형		재무정보 연도	기업코드	종속변수
	cash/assets	liabilities/assets	year	code_label	fnc_rsn_unlst_year
0	0.091490	0.285315	2021	5930	0
1	0.053984	0.632119	2021	373220	0
2	0.052476	0.354774	2021	660	0
3	0.005949	0.373765	2021	207940	0
4	0.090027	0.411737	2021	6400	0
...	...	...	...	...	...
3040	0.170397	0.315057	2017	900180	1
3041	0.296593	0.284505	2014	950010	1
3042	0.073408	0.707520	2010	950030	1
3043	0.008260	0.425390	2012	950070	1
3044	0.694562	0.101859	2021	950180	0

마지막 회계연도 재무정보만 iloc한  
다음해 상장폐지 여부에 대한 변수

#### • 결과

y_real	y_pred	
0	0	479
	1	27
1	0	10
	1	93

- 정확도: 0.9392
- 상장폐지를 상장유지로 예측: 27개
- 상장유지를 상장폐지로 예측: 10개

#### ▼ (2) 원데이터 사용해 최적범위 탐색: 마지막 연도 + (1년전 | 2년전 | .. | n년도까지)

{  $X_t$  시점 재무정보 +  $X_{t-1}$  시점 재무정보 + ... +  $X_{t-n}$  시점 } 에서 최적의 예측성능을 내는  $n$ 을 찾는 실험

- '최대 길이 연도까지(최장영업기간)'는  $n=32$ 으로 총 92개의 기업이 장기적으로 사업한 기업 개수가 매우 적고, 92개 기업 모두 상장폐지되지 않았으므로 결과가 무의미하게 나타남 → 실험X
- '최소 길이연도까지(최단영업기간)'는  $n=1$ 이므로 결과가 무의미함 → 실험X

#### • 학습

- 독립변수 ['cash/assets': 'return'], ['code\_label'], ['year']
- 종속변수 ['fnc\_rsn\_unlst\_year']

$x_t$ =마지막연도		$x_t$ 시점 재무정보		$x_{t-1}$ 시점 재무정보		기업코드	종속변수
		cash/assets_0	liabilities/assets_0	cash/assets_1	liabilities/assets_1	code_label	fnc_rsn_unlst_year
0		0.077683	0.270434	0.091490	0.285315	5930	0
1		0.074875	0.620626	0.053984	0.632119	373220	0
2		0.041813	0.270672	0.052476	0.354774	660	0
3		0.004491	0.284147	0.005949	0.373765	207940	0
4		0.071791	0.379642	0.090027	0.411737	6400	0
...		...	...	...	...	...	...

- o  $n$ 에 따른 loop

												Unlist
		Year	1990	1991	...	2003	2004					1
		기업 A	{assets : ..									
Year	1990	1991	...	2003	2004	...	2015	2016	...	2020	2021	0
기업 B	{assets : ..											
		Year	2003	2004	...	2015	2016					1
		기업 C	{assets : ..									

#### • 결과

- o 기간별 정확도, 오답률

	term	score	real0_pred1(%)	real1_pred0(%)
0	1년전까지	0.9191	4.322	27.835
1	2년전까지	0.9104	5.389	26.471
2	3년전까지	0.9027	6.275	26.471
3	4년전까지	0.9043	5.179	36.145
4	5년전까지	0.9011	5.819	28.431
5	6년전까지	0.9016	4.646	34.021

#### ▼ (3) 변화율 데이터 사용해 최적범위 탐색: 마지막 연도 + (1년전 | 2년전 | .. | $n$ 년도까지)

{  $(X_t - X_{t-1})/(X_{t-1}) \dots (X_t - X_{t-2})/(X_{t-2}) \dots (X_t - X_{t-n})/(X_{t-n})$  } 에서 최적의 예측성능을 내는  $n$ 을 찾는 실험

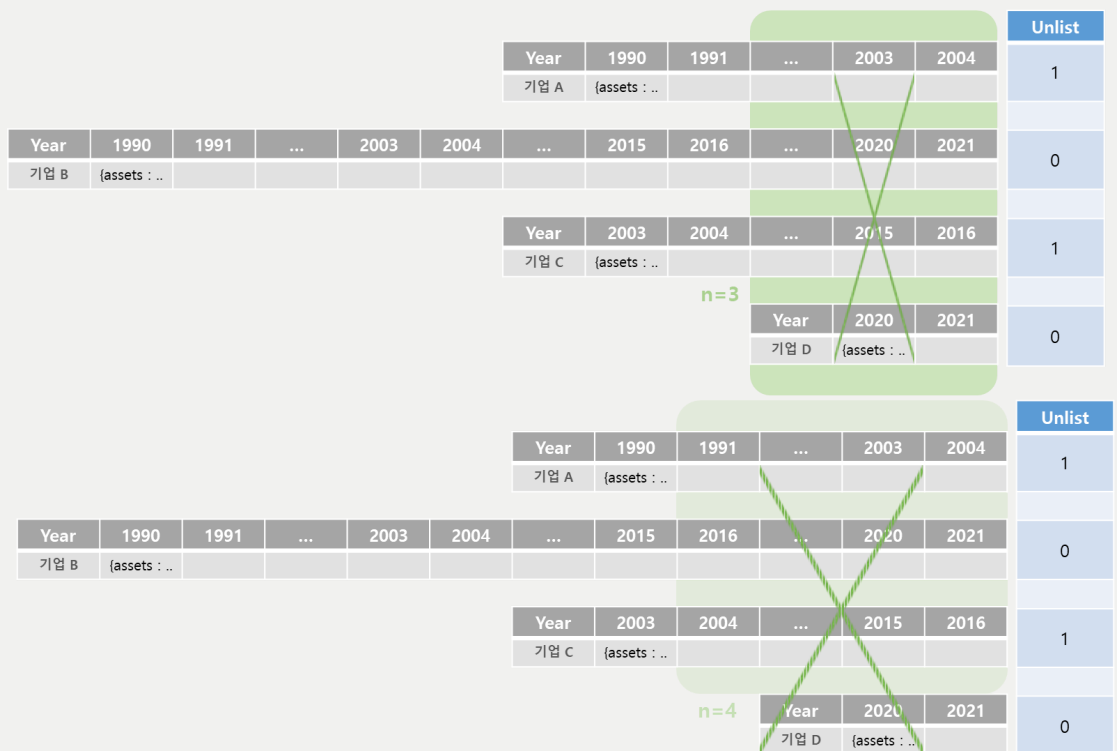
- o  $t$ 시점과  $t - n$  시점 변화율:  $(X_t - X_{t-1})/(X_{t-1})$

#### • 학습

- o 독립변수 ['cash/assets': 'return'], ['code\_label'], ['year']
- o 종속변수 ['fnc\_rsn\_unlst\_year']

$X_t$ 과 $X_{t-1}$ 시점 재무정보 변화율			기업코드	종속변수
	cash/assets_growth	liabilities/assets_growth	code_label	fnc_rsn_unlst_year
0	2.486240	-0.645929	5930	0
1	-0.279015	0.018519	373220	0
2	1.530986	-0.582872	660	0
3	-0.721334	0.627304	207940	0
4	0.556421	-0.334759	6400	0
...	...	...	...	...
3040	0.637586	0.454579	900180	1
3041	1.546060	-0.188975	950010	1
3042	-0.752784	2.667578	950030	1
3043	-0.910588	-0.076373	950070	1
3044	0.135113	-0.397054	950180	0

- o n에 따른 loop



#### • 결과

- o 기간별 정확도, 오답률

	term	score	real0_predl(%)	real1_pred0(%)
0	1년전까지	0.9048	3.953	36.893
1	2년전까지	0.9064	3.953	35.922
2	3년전까지	0.9064	5.138	30.097
3	4년전까지	0.9080	4.941	30.097
4	5년전까지	0.9097	4.150	33.010
5	6년전까지	0.9064	4.545	33.010
6	7년전까지	0.9031	4.941	33.010

▼ (4) 지수함수 기울기 데이터 사용해 최적범위 탐색: 마지막 연도 + (1년전 | 2년전 | .. | n년도까지)

$\{ a * \exp(bX_1) \dots a * \exp(bX_2) \dots a * \exp(bX_n) \}$  에서 최적의 예측성능을 내는  $n$ 을 찾는 실험

- $t$ 시점과  $t - n$  시점까지 최소자승법 지수함수 파라미터:  $a * \exp(bX)$

• 학습

- 독립변수 `['cash/assets': 'return'], ['code_label'], ['year']`
- 종속변수 `['fnc_rsn_unlst_year']`

	$X_t$ 과 $X_{t-n}$ 시점까지 최소자승법 지수함수 기울기값			$X_t$ 과 $X_{t-n}$ 시점까지 최소자승법 지수함수 절편값		기업코드	종속변수
	cash/assets_a	cash/assets_b	liabilities/assets_a	liabilities/assets_b		code_label	fnc_rsn_unlst_year
0	0.026277	0.047779	0.910205	-0.044553		5930	0
1	0.074875	-0.327137	0.620626	0.018350		373220	0
2	NaN	NaN	1.034392	-0.043382		660	0
3	0.109713	-0.308367	0.323608	0.013360		207940	0
4	0.034340	0.040577	0.612500	-0.028288		6400	0
...	...	...	...	...		...	...
3040	0.106058	0.056724	0.274297	0.036408		900180	1
3041	0.150862	0.044876	0.389416	-0.056806		950010	1
3042	0.191154	-0.297272	0.304774	0.179470		950030	1
3043	0.233210	-0.530861	0.328443	-0.009892		950070	1
3044	0.595607	0.038371	0.147613	-0.194057		950180	0

- $n$ 에 따른 loop

• 결과

- 기간별 정확도, 오답률

	term	score	real0_predl(%)	real1_pred0(%)
0	1년전까지	0.8309	0.000	100.000
1	2년전까지	0.9113	4.545	30.097
2	3년전까지	0.9146	3.953	31.068
3	4년전까지	0.9064	3.953	35.922
4	5년전까지	0.9015	4.743	34.951
5	6년전까지	0.9031	4.348	35.922
6	7년전까지	0.9064	5.336	29.126

## ▼ 4. 추가 데이터 제외

<주식정보 제외 결과비교>	1차	2차(n=임시1)	3차(n='all')	4차(n='all')
정확도	0.991	0.911	0.933	0.924
상장폐지를 상장유지로 예측 (오답률 %)	2 / 10838 (0.019%)	23 / 509 (4%)	20 / 506 (3%)	23 / 506 (4%)
상장유지를 상장폐지로 예측	98 / 102 (96%)	31 / 97 (31%)	21 / 103 (20%)	23 / 103 (22%)

<n에 따른 결과비교>	1차	2차(n=1)	3차(n=3, 6)	4차(n=3)
정확도	0.938	0.9109	0.9310	0.9442
상장폐지를 상장유지로 예측 (오답률 %)	24 / 506 (5%)	(4.5%)	(4.3%)	(5.9%)
상장유지를 상장폐지로 예측 (오답률 %)	14 / 103 (9%)	(32.0%)	(19.4%)	<b>(3.9%)</b>

	term	score	real0_predl(x)	real1_pred0(x)
0	1년전까지	0.9109	4.519	31.959
1	2년전까지	0.9071	5.389	28.431
2	3년전까지	0.9094	5.466	26.471
3	4년전까지	0.9162	4.781	30.120
4	5년전까지	0.8940	4.741	37.255
5	6년전까지	0.8998	4.646	35.052

	term	score	real0_predl(x)	real1_pred0(x)
0	1년전까지	0.9294	4.348	20.388
1	2년전까지	0.9245	4.545	22.330
2	3년전까지	0.9310	4.348	19.417
3	4년전까지	0.9195	4.150	27.184
4	5년전까지	0.9261	4.545	21.359
5	6년전까지	0.9310	4.348	19.417
6	7년전까지	0.9261	4.150	23.301

	term	score	real0_predl(x)	real1_pred0(x)
0	1년전까지	0.8309	0.000	100.000
1	2년전까지	0.9392	6.126	5.825
2	3년전까지	0.9442	5.929	3.883
3	4년전까지	0.9360	5.138	12.621
4	5년전까지	0.9360	6.126	7.767
5	6년전까지	0.9360	6.522	5.825
6	7년전까지	0.9392	6.522	3.883

## ▼ 2. 직전 연도 재무정보로 다음 해 상장폐지 예측 (연도 별 교차검증 O)

- 학습
  - Index ['year']
  - 독립변수 ['cash/assets': 'return'], ['code\_label']
  - 종속변수 ['fnc\_rsn\_unlst\_year']

		재무정보 연도	상장폐지 날짜	재무부실사유 n 유무	상장폐지연도	
cor_name	Accounting_Closing	year	unlist_date	fnc_rsn_unlist_year	unlist_reason	
54690	중국고섬	12.0	2008	2013-10-04 2008년 재무정보: 2009년에 폐지 하지 않음 → 0	0	감사의견 의견거절
54691	중국고섬	12.0	2009	2013-10-04 2009년 재무정보: 2010년에 폐지 하지 않음 → 0	0	감사의견 의견거절
54692	중국고섬	12.0	2010	2013-10-04	0	감사의견 의견거절
54693	중국고섬	12.0	2011	2013-10-04 2011년 재무정보: 2012년에 폐지 하지 않음 → 0	0	감사의견 의견거절
54694	중국고섬	12.0	2012	2013-10-04 2012년 재무정보: 2013년에 폐지함 → 1	1	감사의견 의견거절
54695	SNK	7.0	2017	2022-05-18	0	상장폐지 신청('22.04.14)
54696	SNK	7.0	2018	2022-05-18	0	상장폐지 신청('22.04.14)
54697	SNK	7.0	2019	2022-05-18	0	상장폐지 신청('22.04.14)
54698	SNK	7.0	2020	2022-05-18	0	상장폐지 신청('22.04.14)
54699	SNK	7.0	2021	2022-05-18	0	상장폐지 신청('22.04.14)

즉, 이 변수는 다음해  
상장폐지 여부에 대한 변수

• 결과

y	pred_y	
0	0	19880
1	1	11
1	0	112
1	1	10

<교차검증 결과비교>	1차	2차	3차	4차	5차
정확도	0.994				
상장폐지를 상장유지로 예측	11 / 19891				
상장유지를 상장폐지로 예측	112 / 122				

<마지막 연도 결과비교>	1차	2차	3차	4차	5차
정확도	0.939				
상장폐지를 상장유지로 예측	27 / 506				
상장유지를 상장폐지로 예측	27 / 102				