

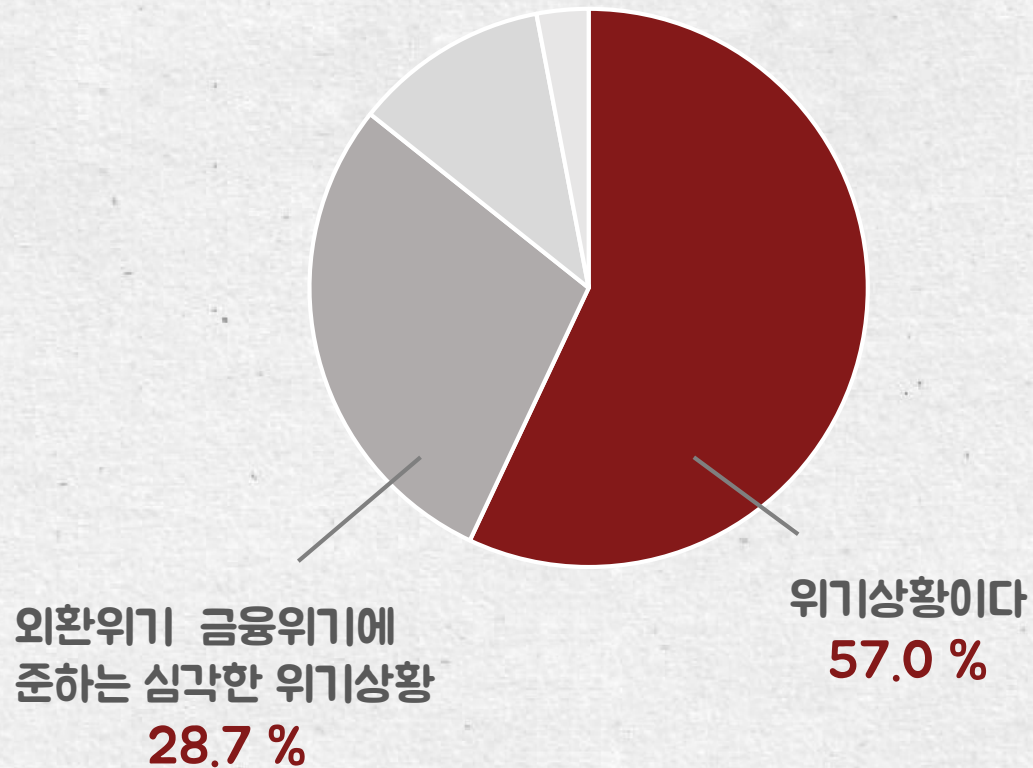
A low-angle, upward-looking perspective of several tall, modern skyscrapers with glass facades, creating a sense of height and urban density. The buildings are dark and silhouetted against a lighter sky.

중소기업 폐업 분석과 예측

비주얼라이제이션

중소기업 현재 어떠한가?

현재 경제상황에 대한 중소기업 인식
중소기업 86 % '경제위기 우려'



경제위기 감내 가능 시간 12개월 미만

12개월 이내 47.1 %

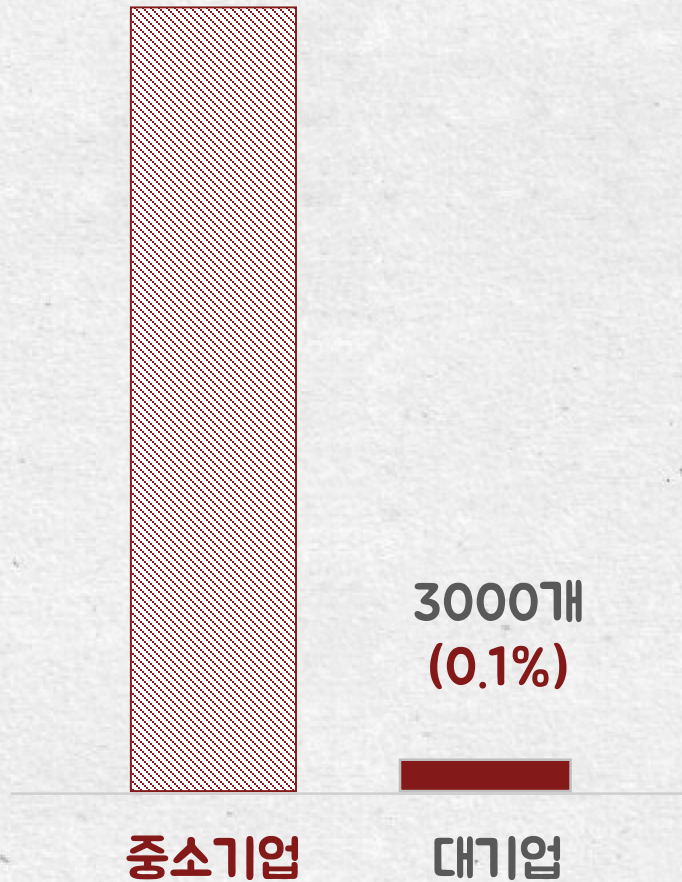
6개월 이내 19.5 %

9개월 이내 5.1%

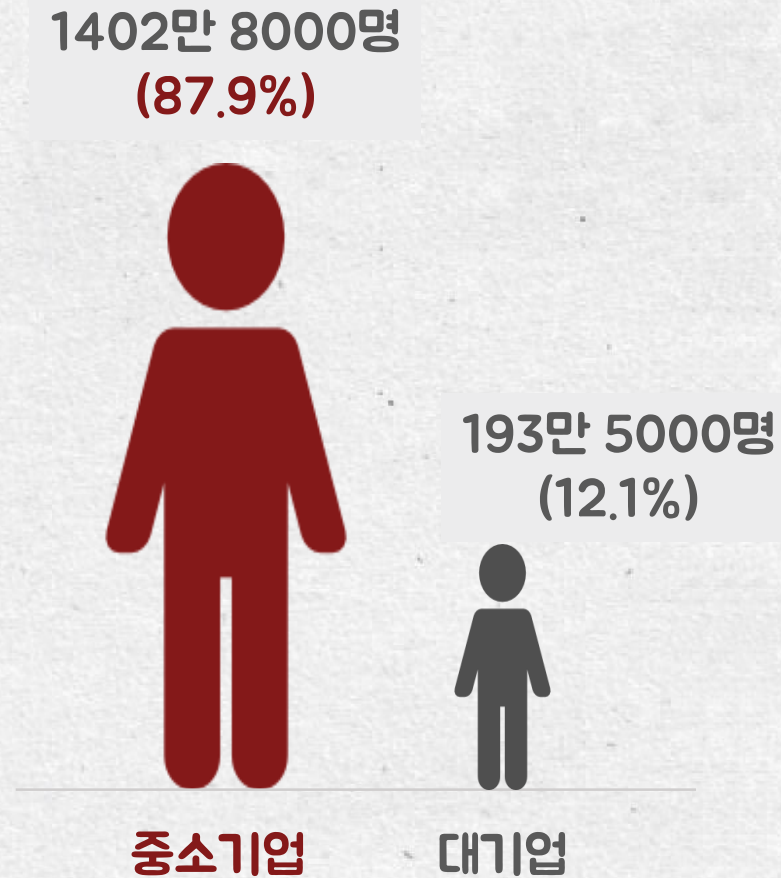
3개월 이내 4.7%

한국 경제의 핵심 축 역할 하는 중소기업

〈 사업체 〉



〈 종사자 〉



데이터 설명

2010

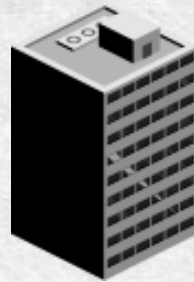
2011

2012

2013

2014

2015



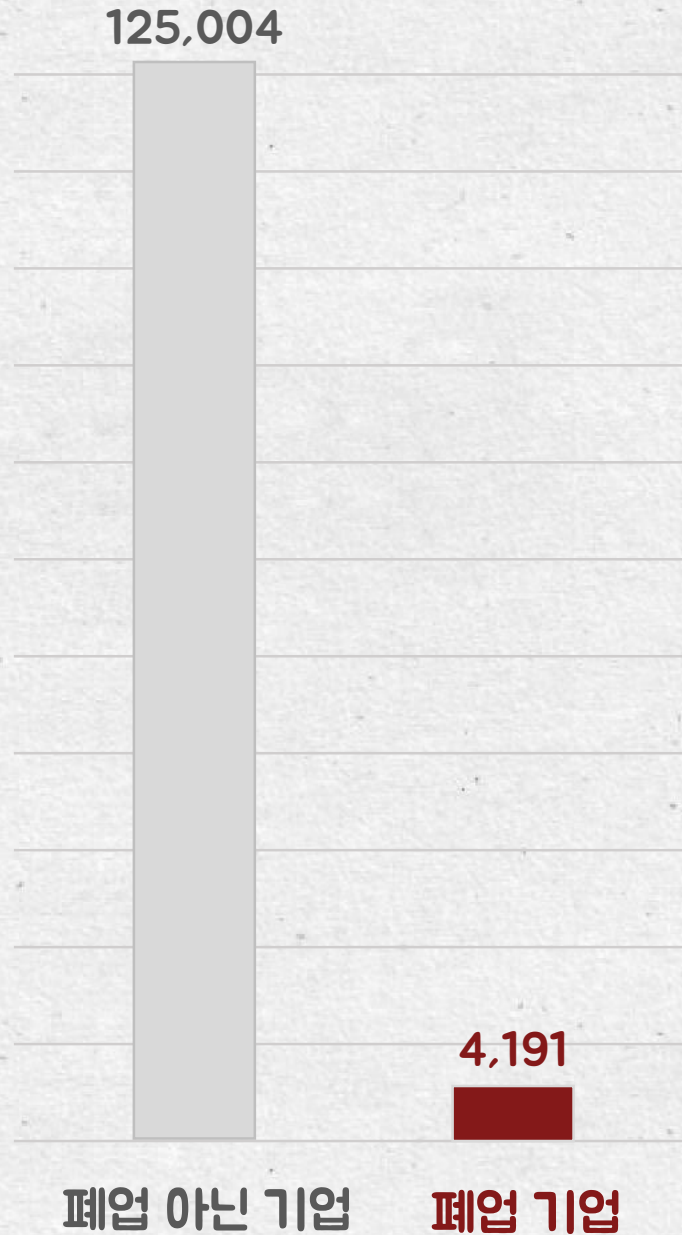
중소기업
약 13 만개

연도에 상관없이 모두 고유한 기업으로 봄

사용한 변수

업체코드	순이익증가율
기업공개	영업이익이자보상비율
유동비율	총자본순이익율
당좌비율	재고자산회전율
부채비율	차입금매출액비율
매출액순이익율	차입금의존도
매출액영업이익율	차입금자기자본
매출액증가율	총자본투자효율
매출채권회전율	종업원수
설비투자효율	폐업여부 (0:폐업x , 1: 폐업)

데이터 전처리



결측치 **중앙값** 처리

표본 129,195개 중 폐업 기업 4,191개

이상치 제거 시 폐업 기업의 개수가 작아짐

데이터 웨손 최소화시키는 중앙값으로 처리

중앙값 -> 결측치 처리, 이상치 포함

데이터 분할



약 13만개 데이터
예측용(train) : 검증용(valid) = 7 : 3

2만개 데이터
평가용(test) 으로 뒀음

의사결정나무를 통해 교호작용 파악

'변수들 간 교호작용은 없음'



로지스틱 회귀분석

변수선택법 - 전진선택법

가장 중요한 변수들을 골라 차례대로 모형에 포함시켜 나가는 과정을 반복하여 회귀방정식을 적합시키는 방법

INWON_x 변수의 상관성이 가장 눈에 띄게 강하여 전진선택법을 채택

로지스틱 회귀분석

변수선택법 - 전진선택법

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Exp(Est)
Intercept	1	-3.2327	0.0223	21074.56	<.0001		0.039
INWON_x	1	-0.0119	0.00110	117.82	<.0001	-0.4114	0.988
x1	1	4.875E-7	1.934E-7	6.35	0.0117	0.0111	1.000
x4	1	-1.08E-6	5.304E-7	4.14	0.0419	-0.00778	1.000

INWON_x(종업원수), x1(부채비율), x4(매출액영업이익율)

유의수준 0.05 보다 작아 유의함

3개의 변수는 폐업의 여부에 영향을 줌

로지스틱 회귀분석

변수선택법 - 전진선택법

Analysis of Maximum Likelihood Estimates							
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Exp(Est)
Intercept	1	-3.2327	0.0223	21074.56	<.0001		0.039
INWON_x	1	-0.0119	0.00110	117.82	<.0001	-0.4114	0.988
x1	1	4.875E-7	1.934E-7	6.35	0.0117	0.0111	1.000
x4	1	-1.08E-6	5.304E-7	4.14	0.0419	-0.00778	1.000

하지만, 오즈비를 보면 **INWON_x (종업원수)** 외에는 의미를 찾기 어려움

종업원 수가 작을수록 기업 폐업에 영향을 준다는 의미를 갖음

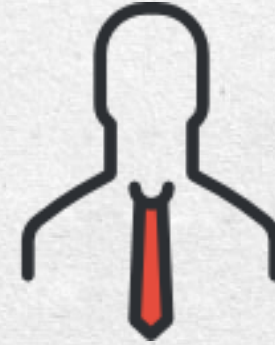
폐업 기업 예측



부채비율



매출액영업이익율



종업원수

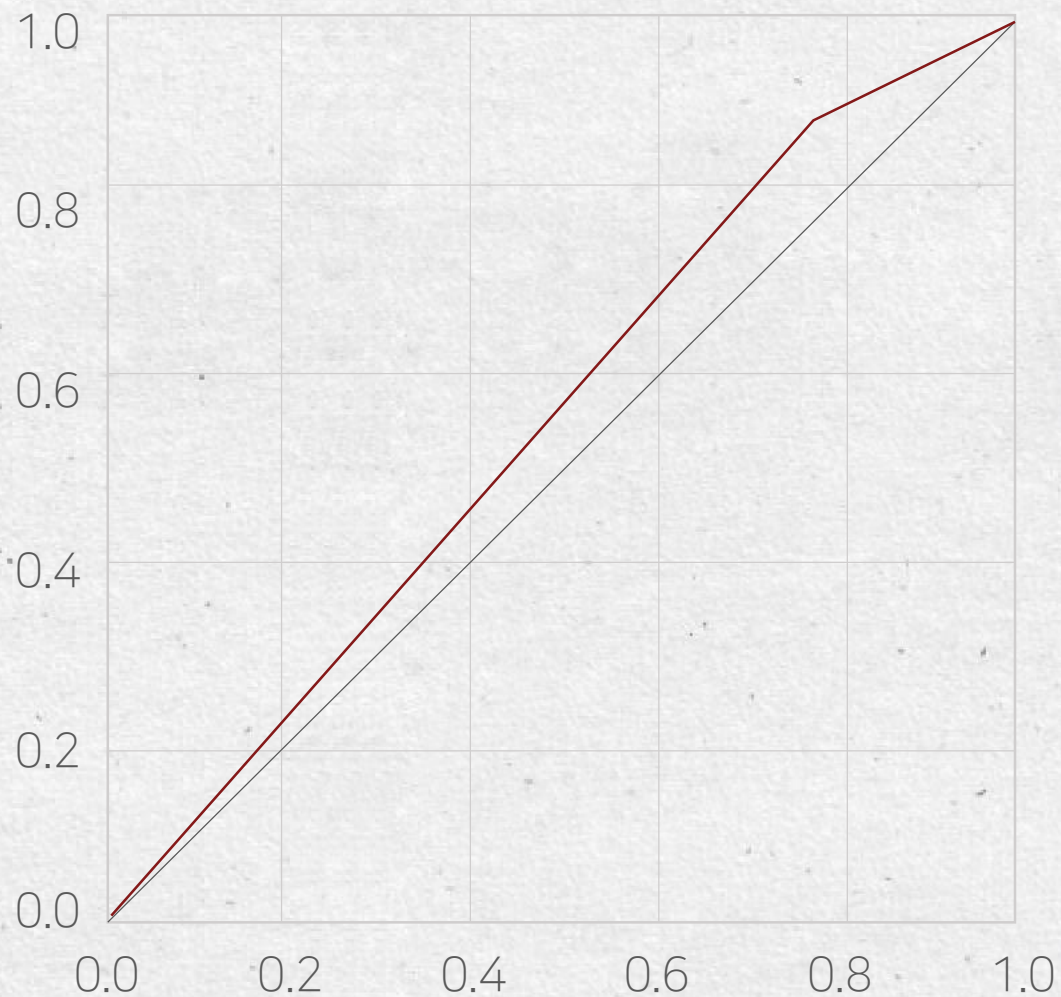
앞서 폐업하는 기업에 영향을 주는 변수는 실질적으로 **종업원수** 밖에 없었음
하지만 폐업 기업 예측에 중요한 변수로 **부채비율**, **매출액영업이익율**도 포함함

2933 : 87501 (training)

1258 : 37502 (test)

폐업한 기업과 아닌 기업의 **비율 차이가 큼**

하지만 오버샘플링을 하지 않았음
모형 적합이 가시적으로 나아지려면
경진대회 이외 데이터와 변수를 고려해야한다고 판단



ROC 그래프

ROC 인덱스가 0.609로 낮게 나옴

모형 적합이 낮게 나온 것

모형 예측면에서는 설명력이 부족

— 회귀-전진
— 기준선

폐업 기업 예측 결과

제시된 20,403 데이터 예측
예측 결과 **11개의 폐업 기업 예측**

모형 적합이 낮아 올바른 예측이 되었다고
판단하기 어려움

폐업 기업의 데이터 양이 적고
모형을 만들기 어려운 데이터라고 판단됨