

**Project 6**

---

# **GLM/GAM/Cox-PHM을 이용한 기업 부도 예측**

강아미 고유정 이혜린 윤보인 홍지원

# Contents

00 데이터 전처리

01 Logistic GLM/GAM 이용한 부도 예측

02 생존 분석 기법을 이용한 부도 예측 방법

03 Lift Chart 이용한 예측력 비교

04 Data Mining 기법

05 Shiny Application

# 01 Logistic GLM/GAM 이용한 부도 예측

# 00 데이터 전처리

## 1) 변수 변환

로그 변환 (12개)		제곱근/세제곱근 변환 (13개)	
고정부채비율	X13	총자본투자효율	X1
고정비율	X14	경영자본순이익율	X4
부채비율	X15	자기자본순이익율	X7
유동비율	X19	자본금순이익율	X8
고정자산/차입금비율	X23	총자본순이익율	X10
유동부채구성비율	X30	총자산사업이익율	X12
...		...	

## 2) 변수 제거 및 결측치 처리

- 변수 간의 상관관계를 살펴보고 0.95 이상의 상관성을 가진 변수 중 하나를 선택
- 9999.99, -9999.99의 결측치를 갖는 값은 knn을 통해 대체

# 01 Logistic GLM/GAM 이용한 부도 예측

Logistic Regression을 이용한 방법: 각 기업의 향후 12개월 이내 부도확률 예측 모형

Model	Formula	AIC
Link = probit	<code>glm(formula = delta ~ x2 + x4 + x5 + x10 + x16 + x20 + x27 + x29 + x30 + x31 + x40 + x42 + x44 + x1 + x2:x29 + x5:x16 + x5:x27 + x5:x29 + x5:x31 + x10:x29 + x10:x42 + x16:x20 + x16:x40 + x20:x30 + x20:x42 + x27:x31 + x27:x40 + x29:x31 + x30:x40 + x40:x42 + x40:x44 + x42:x44 + x20:x1 + x40:x1 + x44:x1, family = binomial(link = probit), data = train)</code>	854.08

## 선택된 주요 정량변수와 부도확률의 연관성

Variable	변수 효과		Variable	변수 효과	
x2 매출채권증가율	-0.13765*	-	x27 총CF/차입금비율	-0.37524**	-
x4 경영자본순이익율	0.84674*	+	x29 순운전자본/총자본비율	0.35158***	+
x5 금융비용/총부채비율	0.44136***	+	x30 유동부채구성비율	-0.16545*	-
x10 총자본순이익률	-1.36900**	-	x31 현금비율	-0.63595**	-

- 경영자본순이익율, 금융비용/총부채비율, 순운전자본/총자본비율: 부도확률과 양의 상관관계
- 매출채권증가율, 총자본순이익율, 총CF/차입금비율, 유동부채구성비율, 현금비율: 부도확률과 음의 상관관계

# 01 Logistic GLM/GAM 이용한 부도 예측

GAM을 이용한 방법: 각 기업의 향후 12개월 이내 부도확률 예측 모형

Model	Formula	AIC	p
Link=probit	gam(delta ~ x1 + s(x2) + s(x4) + s(x5) + s(x10) + x13 + x22 + s(x23) + s(x25) + s(x28) + s(x30) + x31 + x33 + s(x39) + s(x40) + s(x43), family = binomial(link=probit) , data = train)	825.82	44.61

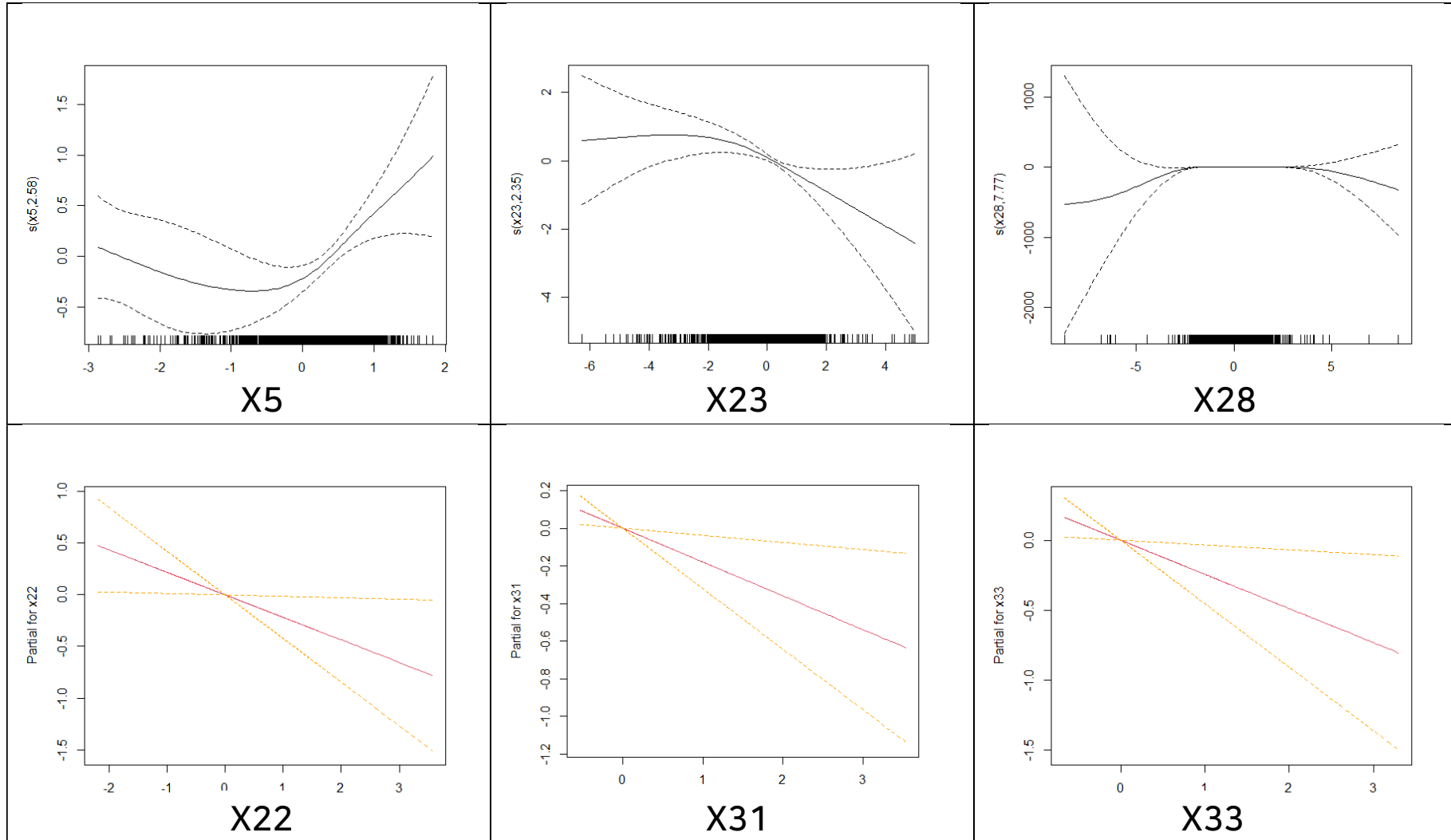
선택된 주요 정량변수와 부도확률의 연관성

Variable	변수 효과(edf)		Variable	변수 효과	
x5 금융비용/총부채비율	2.549**	+	x22 차입금의존도	-0.21836*	-
x23 고정자산/차입금비율	2.446**	+	x31 현금비율	-0.17947*	-
x28 CF/차입금비율	7.759***	+	x33 고정자산회전율	-0.24386*	-

- 금융비용/총부채비율, 고정자산/차입금비율, CF/차입금 비율: 부도 확률과 양의 상관관계
- 차입금의존도, 현금비율, 고정자산회전율: 부도 확률과 음의 상관관계

# 01 Logistic GLM/GAM 이용한 부도 예측

주요 정량변수에 대한 최적 함수의 그래프와 의미



## 02 생존 분석 기법을 이용한 부도 예측 방법



## 02 생존 분석 기법을 이용한 부도 예측 방법

공변량 (covariates) 이  $(x_1, \dots, x_p)$ 인 기업의 t시점의 순간 부도율에 대한 최적 Cox PHM (Proportional Hazard Model)

Model	Formula	AIC
Cox PHM	<code>coxph(formula = Surv((exp(x40) + y)/365, delta) ~ x2 + x4 + x5 + x10 + x16 + x20 + x27 + x30 + x31 + x42 + x44 + x22 + x6 + x7 + x14 + x19 + x23 + x33 + x41 + x43 + x1 + x8 + x13 + x15 + x28 + x25 + x34 + x36 + x37 + x39 + x29 + x12 + x5:x16 + x5:x27 + x10:x30 + x10:x31 + x42:x44 + x23:x1 + x4:x23 + x14:x23 + x23:x15 + x8:x28 + x14:x28 + x15:x28 + x19:x28 + x28:x25 + x30:x28 + x33:x28 + x33:x13 + x13:x36 + x5:x14 + x5:x23 + x5:x36 + x5:x37 + x5:x39 + x5:x41 + x5:x43 + x2:x29 + x7:x29 + x27:x1 + x27:x6 + x10:x27 + x27:x12 + x27:x25 + x27:x30 + x27:x34 + x27:x43 + x27:x44, data = train)</code>	1599.82

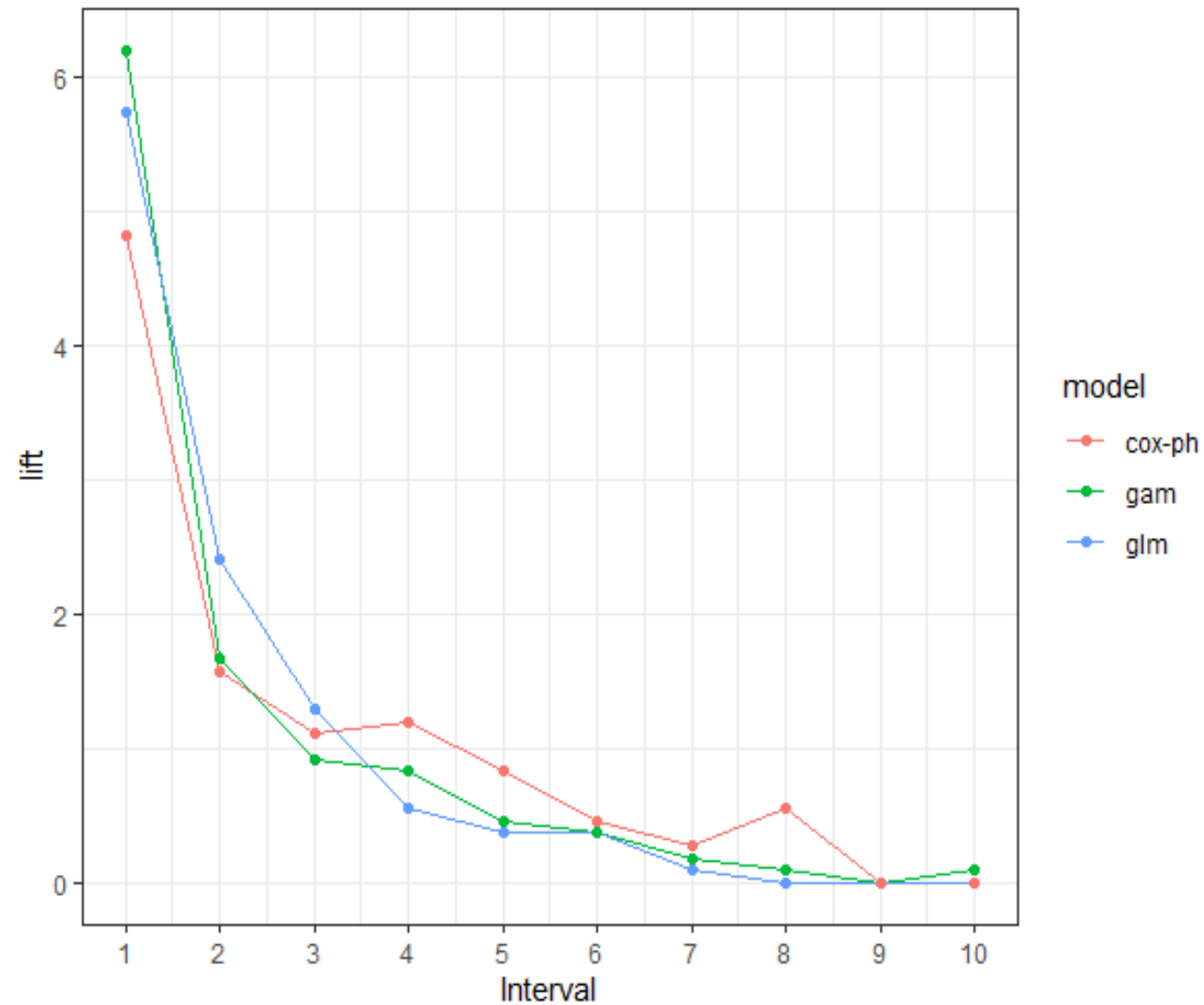
### 선택된 주요 정량변수와 부도확률의 연관성

Variable	변수 효과(edf)		Variable	변수 효과	
x2 매출채권증가율	-0.36334*	-	x27 총CF차입금비율	-1.08098**	-
x8 자본금순이익율	-0.70068**	-	x31 현금비율	-0.91961***	-
x23 고정자산/차입금비율	-1.34722**	-	x33 고정자산회전율	-0.99603**	-

- 매출채권증가율, 자본금순이익율, 고정자산/차입금비율, 총CF/차입금비율, 현금비율, 고정자산회전율이 낮을수록 부도 확률이 높음

## 03 Lift Chart를 이용한 예측력 비교

### 03 Lift Chart를 이용한 예측력 비교



### 모형별 AIC

GLM	GAM	Cox PHM
854.0817	852.82	1599.815

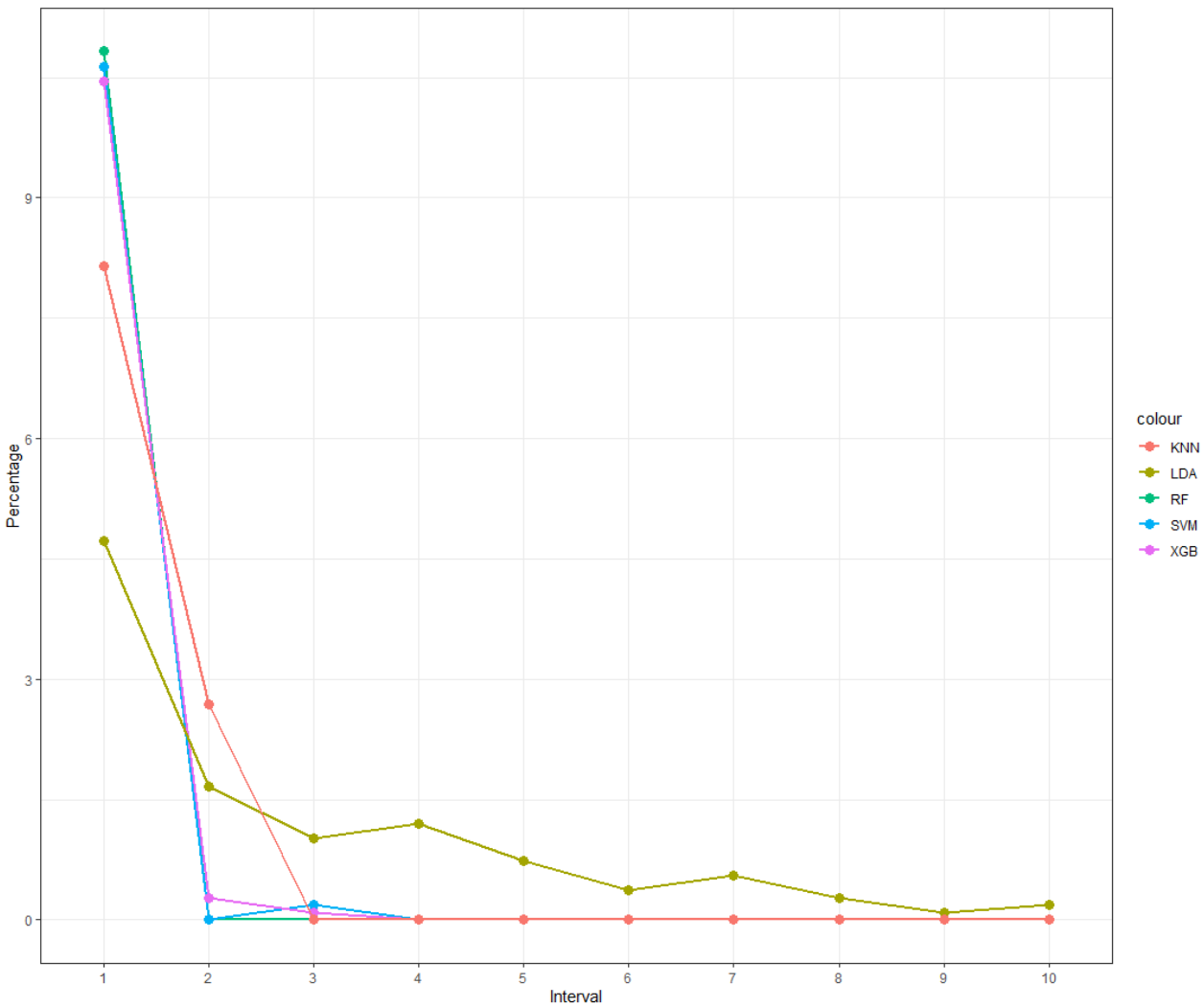
### Lift Chart

Interval	GLM	GAM	Cox PHM
10%	0.1956	0.2114	0.164
20%	0.082	0.0568	0.0536
30%	0.044	0.0315	0.0379
40%	0.018	0.0284	0.041
50%	0.012	0.0158	0.0284
60%	0.01	0.0126	0.0158
70%	0.003	0.0063	0.0095
80%	0	0.0032	0.0189
90%	0	0	0
100%	0	0.0032	0

## 04 Data Mining 기법을 이용한 예측력 비교

# 04 Data Mining 기법을 이용한 예측력 비교

a) 위에서 사용한 방법 외에 아래의 다양한 Data Mining 기법을 이용한 분석을 추가하여 각 방법들의 장단점을 서로 비교 검토해 보시오.



## 모형별 정확도

KNN	LDA	RF	SVM	XGB
91.86%	89.52%	93.42%	93.29%	93.16%

## Lift Chart

Interval	KNN	LDA	RF	SVM	XGB
10%	8.14	4.72	10.83	10.64	10.46
20%	2.68	1.67	0	0	0.28
30%	0	1.02	0	0.19	0.09
40%	0	1.20	0	0	0
50%	0	0.74	0	0	0
60%	0	0.37	0	0	0
70%	0	0.56	0	0	0
80%	0	0.28	0	0	0
90%	0	0.09	0	0	0
100%	0	0.19	0	0	0

## **05 Shiny Application**