
기부에 영향을 미치는 주요 예측변인 탐색:

머신러닝을 이용하여

이영주 (연세대학교 교육연구소 연구원)



Table of *contents*

01

연구 필요성

02

연구 목적

03

연구 방법

04

결과 및 해석

연구 배경



- 최근 사회적 책임·공동체 의식 증대
- 기부에 대한 사회적 논의 증가
- 그러나 실제 기부율은 정체·감소 추세
(2022 Giving Korea 조사에서 기부율은 지속 하락 중)



- 기존 연구: 회귀분석 중심 → 복합적 상호작용 반영 한계
(선형성 가정, 변수 간 상호작용 반영 어려움, 다중공선성 문제)
 - 기부 결정 요인 연구 결과: 불일치·혼재
 - 머신러닝: 다양한 변수·비선형적 관계 반영 가능
-

선행 연구



신뢰

비영리조직 투명성
조직 신뢰도, 사회적 신뢰도
재단에 대한 인식
(여러 연구 공통)



사회참여·연대

이타행동
일반적 신뢰, 호혜성
사회적 규범
사회참여 태도



인구학적 요인

연령, 성별, 학력, 소득, 종교
자원봉사 경험
(대부분 연구에서 반복 확인)



기부행동

과거 기부 경험
기부 지속기간
기부액, 기부 분야
기부 이유/기피 이유
기부단체 선택기준

연구 문제

- ① 기부 예측에 적합한 머신러닝 모형은 무엇인가?
- ② 기부 행동에 영향을 주는 핵심 변수는 무엇인가?

(변수 중요도를 통해 정책적 시사점이 가능한 핵심 변수를 도출한다)

연구 자료

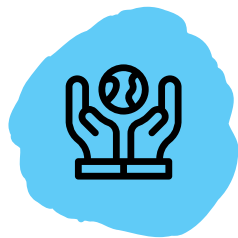
기빙코리아 2022년 자료
2022 6/22~7/6 온라인

<표 2> 연구대상의 일반적인 특성과 기술통계

		변인	n(%)			변인	n(%)
성별		남자	1,076(51.3)	지역		서울, 인천, 경기	1,069(50.9)
		여자	1,019(48.5)			강원, 제주, 세종	191(9.0)
		기타	4(0.2)			대전, 세종, 충북, 충남	227(10.8)
만나이		18세 이상~40세 미만	746(35.5)	자녀		광주, 전북, 전남	210(10.0)
		41세 이상~60세 미만	853(40.6)			대구, 부산, 울산, 경남	402(19.1)
		61세 이상~80세 미만	500(23.8)			예	1,266(60.3)
혼인 상태		미혼	711(33.9)	여부		아니요	833(39.7)
		기혼	1,266(60.3)			기독교	497(23.7)
		별거/이혼/가별	122(5.8)			천주교	258(12.3)
		상용근로자	1,105(52.6)	종교		불교	294(14.0)
		임시근로자	141(6.7)			원불교, 기타	16(0.8)
		일용근로자	35(1.7)			없음	1,034(49.3)
고용 상태		자영업자	173(8.2)	교육 수준		초등학교졸업이하	10(0.5)
		학생	156(7.4)			중학교졸업이하	24(1.1)
		주부	263(12.5)			고등학교졸업이하	368(17.5)
		실업/미취업	100(4.8)			대학교졸업이하	1,348(64.2)
		퇴직	126(6.0)			대학교졸업이상	349(16.6)

Note: 고용 상태에서 근로자별 구체적인 기간은 상용근로자는 계약기간 1년 이상, 임시근로자는 계약기간 1개월 이상 1년 미만, 일용근로자는 계약기간 1개월 미만이다.

연구 방법



연구자료

- 데이터: 2022 기빙코리아 조사
- 표본: 성인 2,501명 중 기부 경험 無 제외
→ 2,099명
- 평균 연령: 46.9세, 성별: 남 51%, 여 49%

표본특성

- 혼인 상태: 기혼 60%
 - 종교: 없음 49%, 종교 있음 51%
 - 교육 수준: 대학 이상 80%
 - 고용 상태: 상용근로자 53%
-

변수 구성

<표 3> 본 연구에서 모델별로 사용된 20개 변수

구분		설명변수
설명 변수	기부 관련 태도 및 사회 참여(8)	도움 제공 빈도, 사회 안녕 관련 동의도, 주변 사람과의 관계 동의도, 사회공동체 관련 동의도, 온라인 매체 활용도, 사회 참여 활동, 비영리 단체에 대한 동의도, 일반 사회 신뢰도(믿음에 대한 동의도)
	인구학적 특성(12)	성별, 연령, 거주 지역, 가구원 수, 가구 소득, 가구 소득변화, 혼인 상태, 자녀 유무, 종교, 정치적 성향, 고용 상태, 교육 수준
결과변수		기부 여부(예, 아니오)

분석 방법



분석 모형

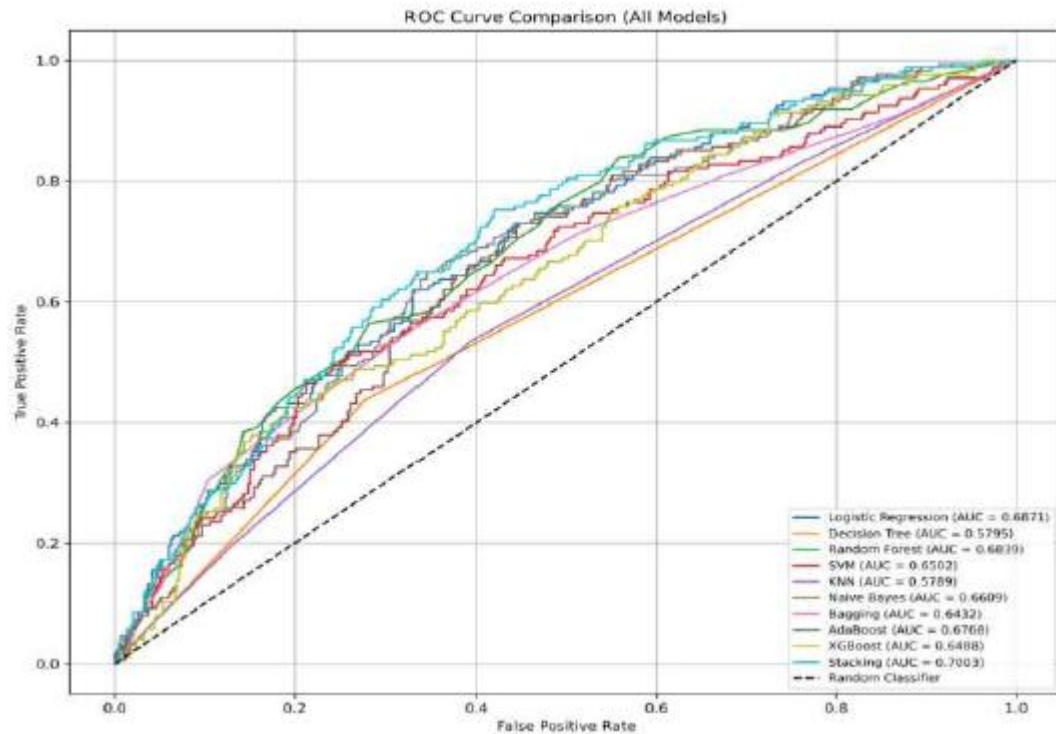
- 8개 머신러닝 기법 적용
Logistic, Decision Tree, Stacking, Random Forest, KNN, Naïve Bayes, Bagging, XGBoost
- 평가 지표: Accuracy, Precision, Recall, F1, AUC, Log-loss

결과 ① 모형 성능 비교

- 상위 모형: Logistic Regression, Random Forest, Stacking
- Accuracy: Logistic 0.726
- AUC: Random Forest 0.696
- Log-loss: Stacking 0.553

② ROC Curve

- ROC 곡선 비교
- Logistic / Random Forest / Stacking
→ 안정적 성능



<그림 1> 머신러닝 모형별 ROC 곡선

- 전체적으로 Logistic, RF, Stacking의 ROC 곡선이 직선보다 위에 있으며, 이는 기부 여부를 무작위보다 유의미하게 예측함을 의미함.
- 특히 Random Forest가 전 범위에서 가장 안정적 면적을 보여 AUC 최상

<표 6> 상위 3개 모형의 설명변수

순위	평균 순위	Logistic Regression	Random Forest	Stacking	설명변수
1	1	1	1	1	비영리 단체에 대한 동의도
2	3.67	2	7	2	사회 안녕 관련 동의도
3	6	6	8	4	일반 사회 신뢰도(믿음에 대한 동의도)
4	6	3	9	6	사회 참여 활동
5	6.33	9	5	5	기부 정보의 신뢰도 및 진실성에 대한 동의도
6	8.33	19	3	3	가구 소득
7	9.33	5	6	17	주변 사람과의 관계 동의도
8	10.33	8	12	11	고용 상태
9	11	7	16	10	교육 수준
10	11	17	2	14	연령
11	11	10	15	8	종교

결과③ 변수 중요도

- 상위 변수

1. 비영리단체 동의도 → 기부자의 윤리성·투명성에 대한 확신이 기부 결정을 좌우함
2. 사회 안녕 동의도 → 사회가 잘 되어야 한다는 믿음이 '공익 활동 지향성'으로 연결
3. 일반 사회 신뢰도 → 타인·사회에 대한 긍정적 신뢰가 기부 참여로 이어짐
4. 사회참여 활동 → 참여 경험이 기부의 자연스러운 확장 행동
5. 기부 정보 신뢰도

- 추가 변수

- 가구소득, 관계 인식 · 고용 상태, 교육 수준 · 연령, 종교 → 정책 타깃 세분화 가능
-

논의 및 시사점

- 기부 핵심 요인: 신뢰, 사회안녕, 사회참여
- 인구학적 요인: 타겟 마케팅 전략 활용 가능

- 실무적 시사점:

- 투명한 정보 공개를 통한 기부자 신뢰 확보→ 재단의 회계 투명성·프로젝트 성과 공개 확대
 - 사회적 연대·공익 메시지 강조→ 연대·공동체 중심의 캠페인 설계
 - 데이터 기반 CRM·예측 시스템 활용→ 과거 참여도 높은 집단 중심의 기부 재참여 유도
-

결론 및 한계

- 결론:

- 머신러닝 모형 비교 → 기부 예측 가능성 확인
- 주요 예측변수 도출 → 신뢰·사회적 가치·경제 요인의 중요성 재확인

- 한계:

- 하위집단 분석 미흡
 - 심리·정서 요인 일부 제외 (기부 동기, 감정적 반응 등)
 - 머신러닝 모델의 설명가능성(interpretability) 제한
 - 단면(1년) 데이터 기반 → 인과적 해석 불가- 향후 연구: 종단분석, 인과적 모형 확대
-

향후 연구

1) 종단 자료 활용→ 기부 지속성·변화 궤적 예측

2) 사회·심리 변인 추가→ 예측력 및 설명력 증대

(Giving Korea 조사에는 신뢰·사회참여·인구학 정보는 포함되어 있으나,

정서·동기·사회관계·효능감·미디어 영향과 같은 핵심 심리·사회적 변인은 부재함.

따라서 향후 연구에서는 공감·도덕적 의무감·기부 효능감·지인 권유·미디어 노출 등

새로운 변수를 추가하는 것이 예측력·설명력을 크게 향상시킬 수 있음)

3) 딥러닝 기반 텍스트 분석 결합→ 기부 후기·SNS 의견 데이터를 통한 정성·정량 융합 분석

