

Machine learning prediction of suicidal ideation, planning, and attempt among Korean adults: A population-based study

Tae-Young Pak, Ph.D.

SKKU Consumer Science

제 10회 수선포럼

October 6, 2022

Suicide in South Korea

[표 2] 2011년, 2018년, 2019년 사망원인 순위 추이

[단위: 명, %, 인구 10만 명당 명]

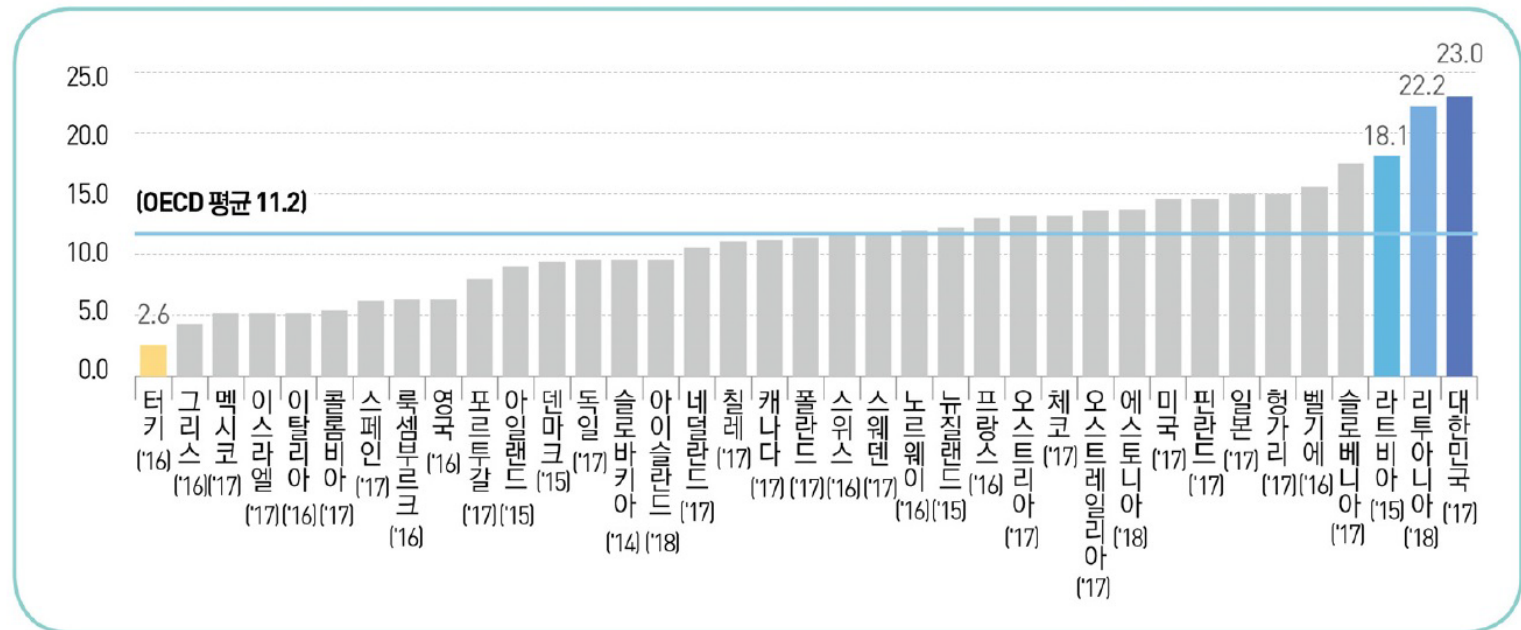
순위	2011년		2018년		2019년					
	사망원인	사망률	사망원인	사망률	사망원인	사망자수	구성비	사망률	'11 순위 대비	'18 순위 대비
1	악성신생물	142.8	악성신생물	154.3	악성신생물	81,203	27.5	158.2	—	—
2	뇌혈관 질환	50.7	심장 질환	62.4	심장 질환	31,030	10.5	60.4	↑+1	—
3	심장 질환	49.8	폐렴	45.4	폐렴	23,168	7.9	45.1	↑+3	—
4	고의적 자해(자살)	31.7	뇌혈관 질환	44.7	뇌혈관 질환	21,586	7.3	42	↓-2	—
5	당뇨병	21.5	고의적 자해(자살)	26.6	고의적 자해(자살)	13,799	4.7	26.9	↓-1	—
6	폐렴	17.2	당뇨병	17.1	당뇨병	8,102	2.7	15.8	↓-1	—
7	만성 하기도 질환	13.9	간 질환	13.4	알츠하이머병	6,744	2.3	13.1	↑+5	↑+2
8	간 질환	13.5	만성 하기도 질환	12.9	간 질환	6,496	2.2	12.7	—	↓-1
9	운수 사고	12.6	알츠하이머병	12.0	만성 하기도 질환	6,176	2.1	12.0	↓-2	↓-1
10	고혈압성 질환	10.1	고혈압성 질환	11.8	고혈압성 질환	5,631	1.9	11	—	—

출처: 2021 자살예방백서, 보건복지부

Suicide in South Korea

[그림 62] OECD 회원국 자살률

[단위: OECD 표준인구 10만 명당 명]

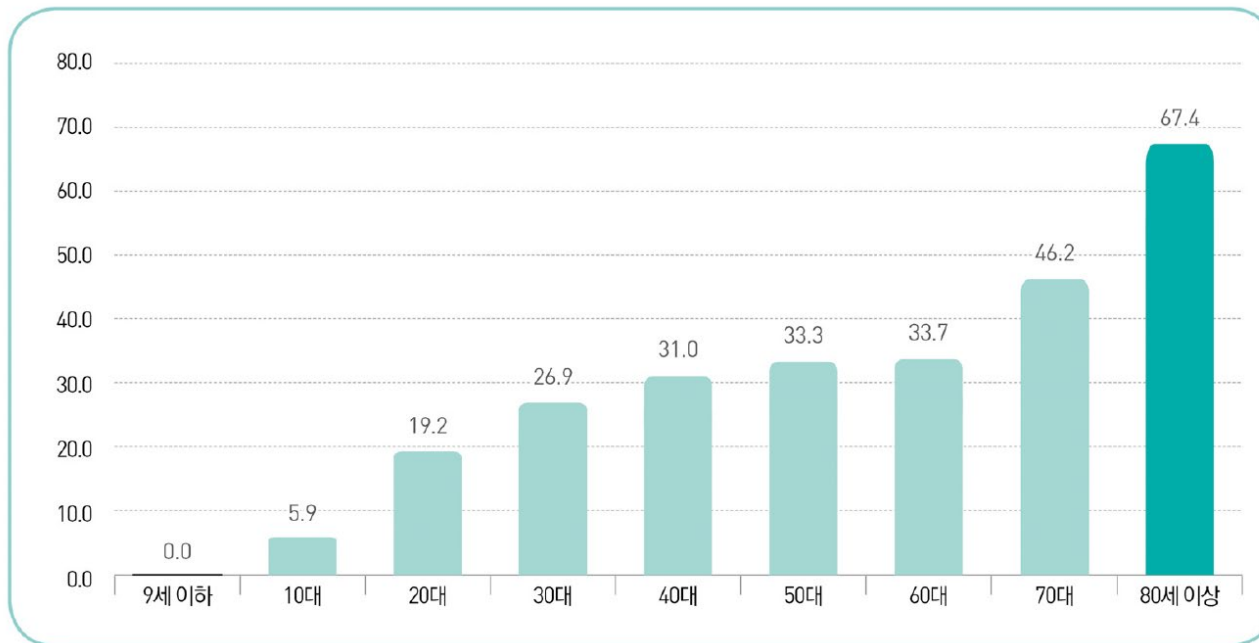


출처: 2021 자살예방백서, 보건복지부

Suicide in South Korea

[그림 6] 2019년 연령대별 자살률

[단위: 인구 10만 명당 명]

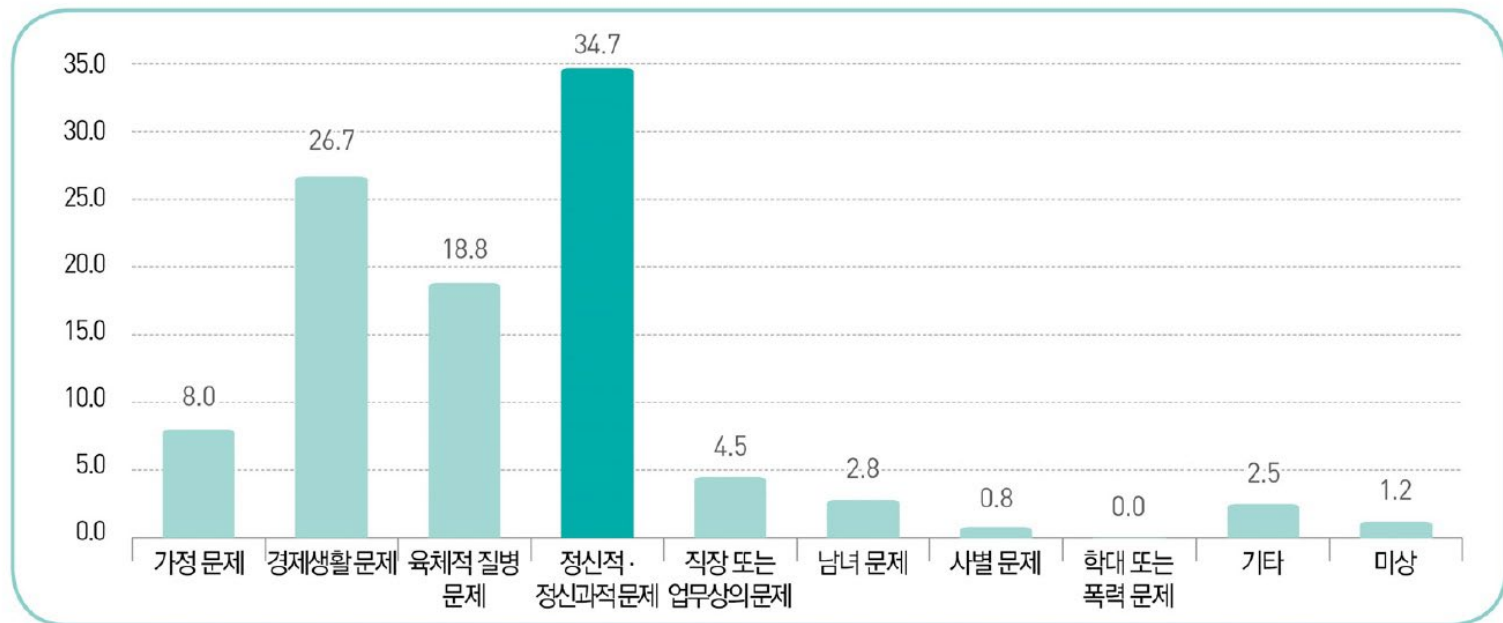


출처: 2021 자살예방백서, 보건복지부

Suicide in South Korea

[그림 45] 2019년 동기별 자살 비율

[단위: %]



출처: 2021 자살예방백서, 보건복지부

Suicide in South Korea

[그림 2] 1988~2019년 자살률 추이

[단위: 인구 10만 명당 명]



출처: 2021 자살예방백서, 보건복지부

Suicide in South Korea

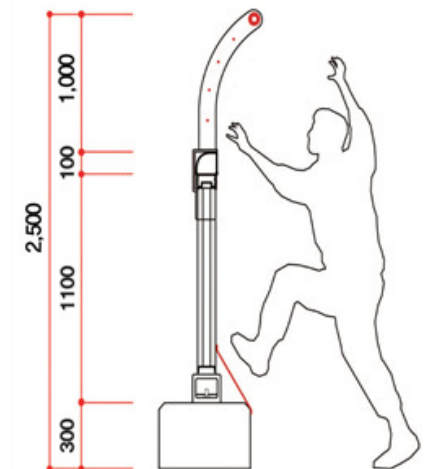
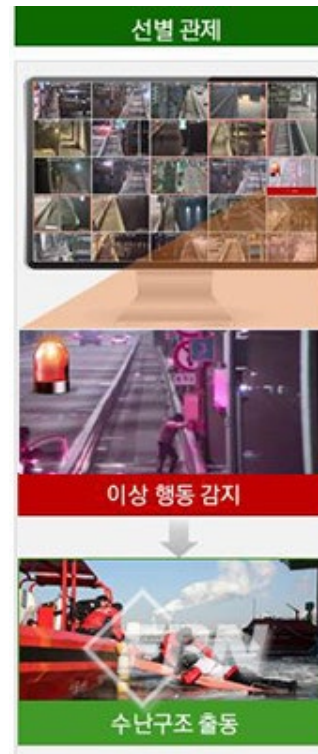
[그림 3] 1988~2019년 연령집단에 따른 자살률 추이

[단위: 인구 10만 명당 명]

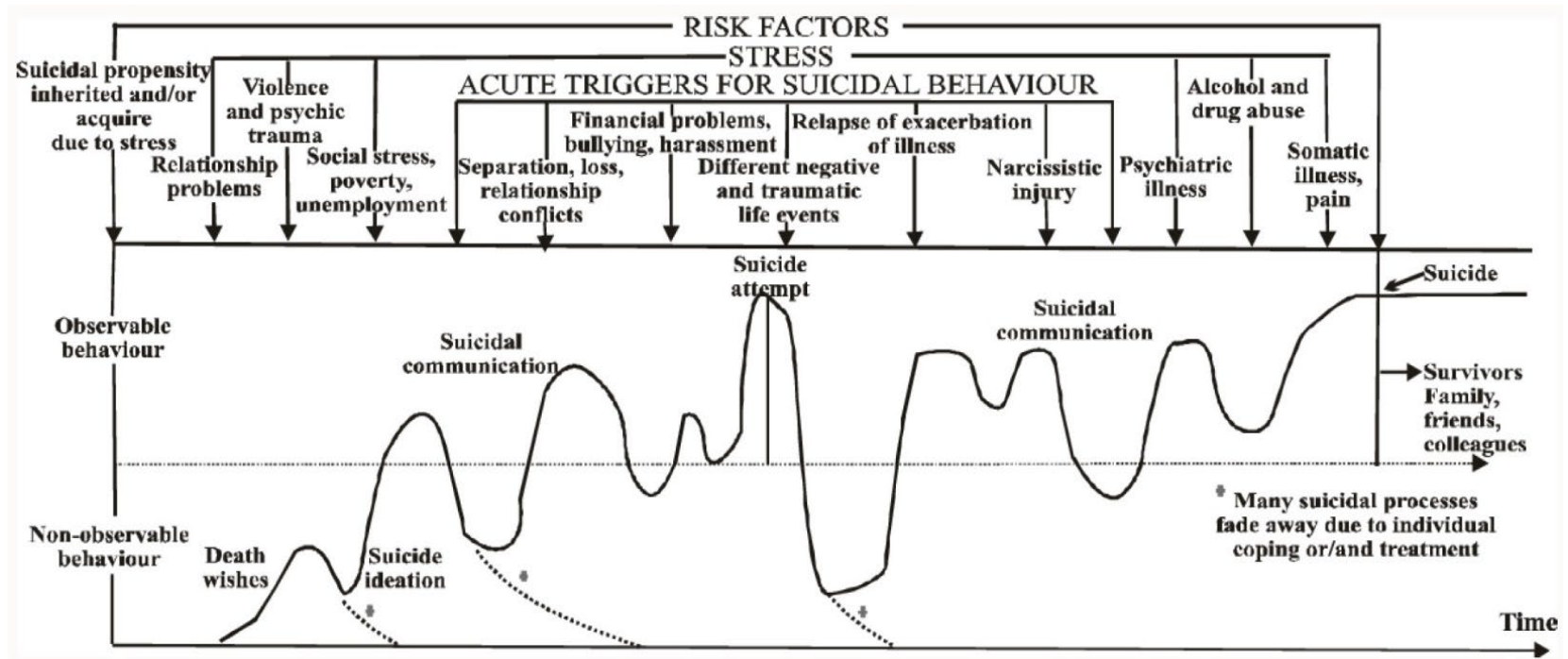


출처: 2021 자살예방백서, 보건복지부

Suicide Prevention in South Korea



Suicidal Process from Suicidal Ideation to Suicide



출처: Wasserman, D. (2001). A Stress-Vulnerability Model and the Development of the Suicidal Process. In D. Wasserman (Ed.), Suicide (pp. 13-27). London: CRC Press.

Suicide Prevention in South Korea

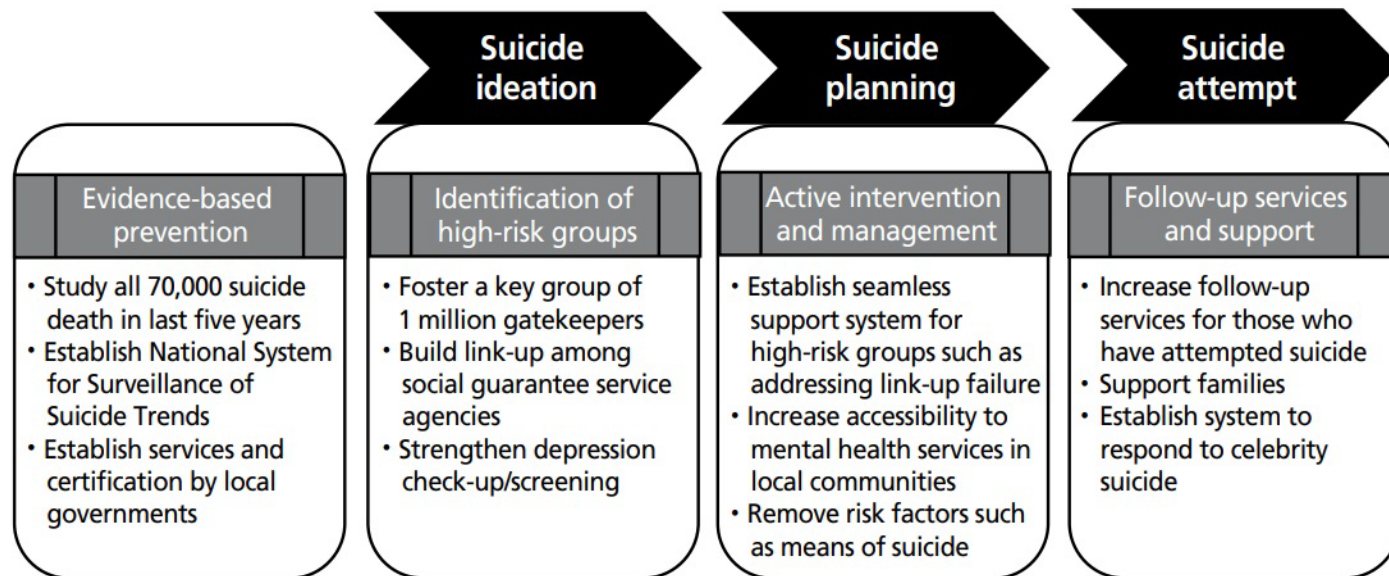


Figure 1. Basic paradigm of National Action Plan for suicide prevention.

출처: 한창수 (2019), 자살예방 어떻게 접근할 것인가?, 대한의사협회지, 62(2): 79-84.

Challenges in Suicide Prevention

- 자살, 정신질환 -> Taboo
- 자살자 중 정신질환을 앓았던 사람의 비율 87% (Bae et al., 2022)
- 자살자 중 사망 전까지 정신질환 치료를 위해 병원에 방문한 비율 15% (Bae et al., 2022)
- But, 자살자 네 명 중 세명은 정신적 괴로움으로 인한 신체 증상(e.g., 어지럼증, 두통, 소화불량, 요통, 답답함, 피로 등)을 치료하기 위해 병원(주로 1차 의료기관)을 방문
- 자살자의 1/3은 사회복지 서비스를 이용하고 있는 취약계층
-> 1차 의료기관, 사회복지센터, 상담센터 등에서 활용 가능한 자살예방 시스템?

Suicide Prediction System

- 자살 위험도 평가/예측을 위한 선행연구
 - 타당도, 신뢰도를 기준으로 구성한 자살 위험 자가진단표
 - 데이터 사이언스, 빅데이터 기반 자살 위험도 평가 시스템

Suicide Prediction System

• 타당도, 신뢰도를 기준으로 구성한 자살 위험 자가진단표

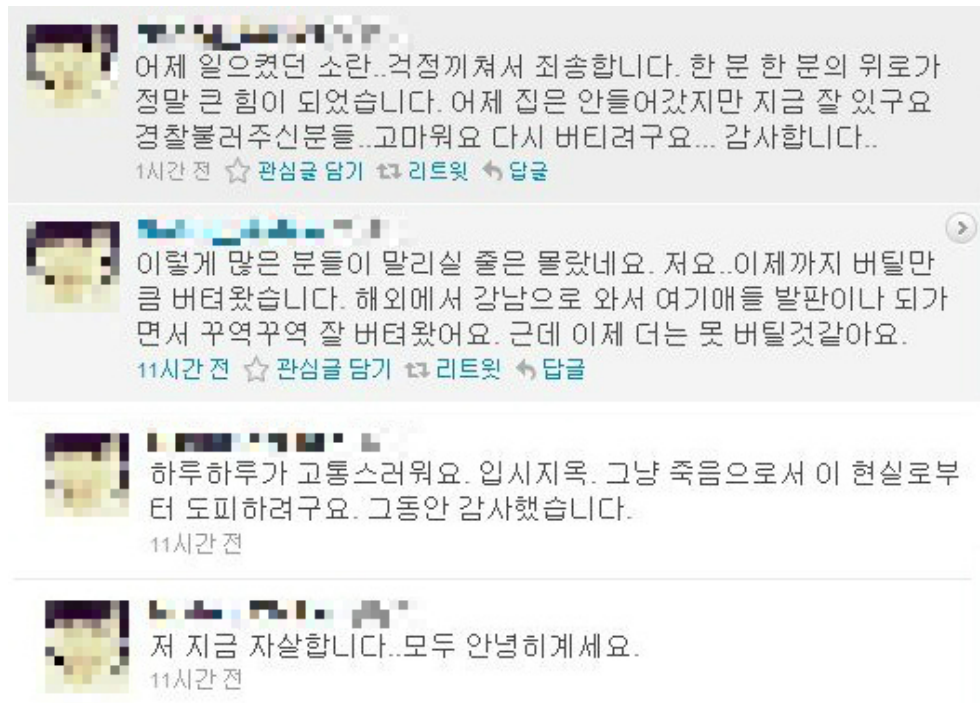
▶ 내원한 환자 중 자살 위험에 처한 환자를 감지하기 위해 자살 위험성 선별 도구를 사용하여 환자를 평가하세요.

자살 위험성 선별 도구	
1. 지난 몇 주 동안 죽고 싶다고 생각한 적이 있습니까?	☐ 예 ☐ 아니요
2. 지난 몇 주 동안 자신이 죽으면 본인 또는 본인의 가족이 더 나아진다고 생각한 적이 있습니까?	☐ 예 ☐ 아니요
3. 지난 몇 주 동안 자살에 관해 생각한 적이 있습니까?	☐ 예 ☐ 아니요
4. 자살을 시도했던 적이 있습니까? '예' 라면, 어떻게 시도했습니까? _____ 언제 시도했습니까? _____	☐ 예 ☐ 아니요
※ 환자가 1~4번까지 모든 질문에 '아니요'라고 응답한 경우 5번 질문은 하지 않아도 되며, 검진을 종료합니다. ※ 이 때 검진 결과는 추가 중재가 필요하지 않습니다. – 음성 (참고 : 임상적 판단이 항상 선행되어야 함) ※ 환자가 1~4번까지의 질문 중 '예'라고 응답한 질문이 있거나 응답을 거부했을 경우, 양성으로 판단됩니다. 급성여부를 평가하기 위해 5번 질문을 합니다.	
5. 본인은 지금 자살하고 싶은 생각이 있습니까? '예' 라면, 구체적으로 설명해주시시오. _____	☐ 예 ☐ 아니요
※ 5번 질문에 '예'로 응답한 환자의 검진 결과는 급성 양성이며, 3단계*로 넘어가세요. * 3단계 종합진단평가는 [별첨] 자살 위험성 선별 검사 사용 가이드 p7~8 참고 – 즉각적인 위험 확인 · 환자는 종합진단평가가 필요합니다. · 환자를 주시해야 합니다. 환자의 주변에서 위험한 물건을 모두 치우세요. 환자의 치료를 담당하고 있는 의사에게 환자상태를 알려주세요. ※ 5번 질문에 '아니요'로 응답한 환자의 검진 결과는 비급성 양성이며, 2단계로 넘어가세요. – 잠재적 위험 확인 · 환자에게 정밀 정신건강평가가 필요한지 여부를 결정하기 위해 간편진단평가가 필요합니다. · 환자의 치료를 담당하고 있는 의사에게 환자상태를 알려주세요.	
자살 예방 정보 • 자살예방상담전화 : 1393 • 자살예방핫라인 : 1577-0199 • 희망의 전화 : 129 • 생명의 전화 : 1588-9191	

자살관련 면담기록지	
날짜: 20 년 월 일 학년 ()반 ()번 이름: _____ 상담교사 : _____ (인)	
면담 내용	① 죽고 싶거나, 죽는 게 더 낫겠다고 생각한 적이 있니? ☐ 예 ☐ 아니요 ② 자살을 할까 생각하고 있니? ☐ 예 ☐ 아니요 ③ 최근(6개월)에 어떻게 죽어야겠다고 계획을 세운 적이 있니? ☐ 예 ☐ 아니요 □ 언제: _____ □ 장소: _____ □ 방법: _____ ④ 최근에 이 계획을 실행해야겠다는 생각을 한 적이 있니? ☐ 예 ☐ 아니요 ⑤ 혹시 죽어야겠다는 생각에 어떤 행동을 해본 적이 있니? ☐ 예 ☐ 아니요 □ 언제: _____ □ 장소: _____ □ 방법: _____ ⑥ 무엇이 너를 그렇게 힘들게 만들고 죽고 싶게 하는지 그 이유를 물어봐도 될까?
자살위험성 단계	□ 위험성 없음 혹은 낮은 위험 : 모두 '아니오' 인 경우, 혹은 ①번만 '예'이고 나머지는 모두 '아니오' 인 경우 : 수동적 자살사고는 있지만, 적극적인 자살 사고, 계획, 의도, 시도 경험은 없음에 해당 □ 중간 위험 : ①, ②번이 모두 '예'이나 나머지는 모두 '아니오' 인 경우 : 수동적/적극적 자살 사고는 있으나 계획, 의도, 시도 경험은 없음에 해당 □ 높은 위험 : ⑤번이 '예'이면서, ③번과 ④번 모두 '아니오'인 경우(①②번의 응답은 상관없음) : 과거 자살시도라는 위험요인이 있으므로, 관리 요망 □ 응급 상황 : ③번이나 ④번 중 하나라도 '예'인 경우(①②번의 응답은 상관없음) : 최근 6개월간 구체적 자살계획과 실행 의도가 있으므로, 응급대처 요망
조치	□ 위험성 없음 혹은 낮은 위험 -> 코넬 상담 인력의 지속적 관찰과 상담 □ 중간위험 -> WEE클래스 우선 연계, 학부요 연락 자살예방센터 및 Wee센터 정신건강지원센터 등 교육장에서 자살예방 전문기관으로 □ 높은위험 -> 학부요에게 연락하고 자살예방센터 또는 전문병원 진료 권유 □ 응급상황 -> 학생을 혼자 두지 않고, 안전조치 부모에게 즉시 알리고 자살예방센터 또는 전문병원 진료 적극 권유

Suicide Prediction System

- 데이터 사이언스, 빅데이터 기반 자살 위험도 평가 시스템
 - 소셜 빅데이터, 비정형 데이터 활용
 - 머신러닝/AI 기반 예측 시스템



어제 일으켰던 소란..걱정끼쳐서 죄송합니다. 한 분 한 분의 위로가 정말 큰 힘이 되었습니다. 어제 집은 안들어갔지만 지금 잘 있구요 경찰불러주신분들..고마워요 다시 버티려구요... 감사합니다..
1시간 전 ☆ 관심글 담기 리트윗 답글

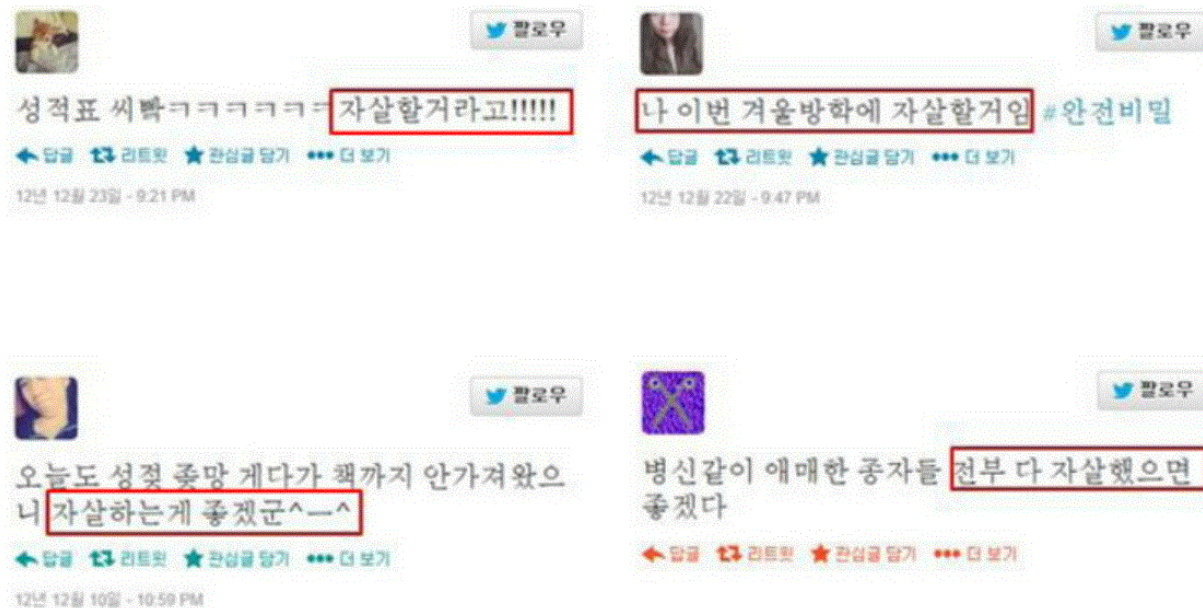
이렇게 많은 분들이 말리실 줄은 몰랐네요. 저요..이제까지 버틸만큼 버텼습니다. 해외에서 강남으로 와서 여기애들 발판이나 되가면서 꾸역꾸역 잘 버텼어요. 근데 이제 더는 못 버틸것 같아요..
11시간 전 ☆ 관심글 담기 리트윗 답글

하루하루가 고통스러워요. 입시지옥. 그냥 죽음으로서 이 현실로부터 도피하려고요. 그동안 감사했습니다..
11시간 전

저 지금 자살합니다..모두 안녕히계세요..
11시간 전

Suicide Prediction System

- 데이터 사이언스, 빅데이터 기반 자살 위험도 평가 시스템
 - 소셜 빅데이터, 비정형 데이터 활용
 - 머신러닝/AI 기반 예측 시스템

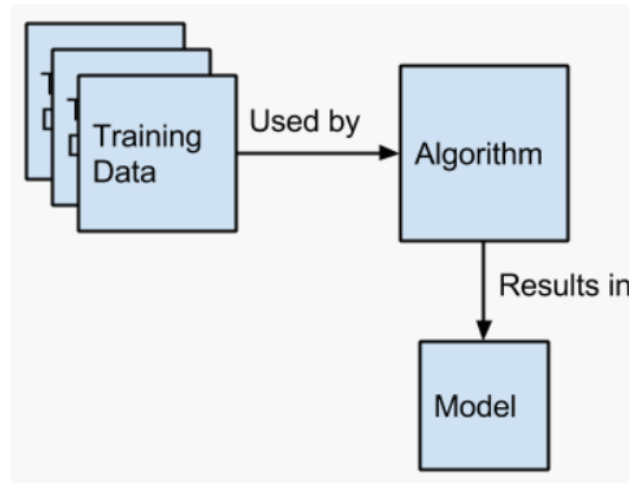


출처: 송태민 (2015), 소셜 빅데이터를 활용한 청소년 자살담론 분석.

Suicide Prediction System

- 데이터 사이언스, 빅데이터 기반 자살 위험도 평가 시스템
 - 소셜 빅데이터, 비정형 데이터 활용
 - 머신러닝/AI 기반 예측 시스템

Model training:



Prediction:



Suicide Prediction System

- 선행 연구의 한계점
 - 타당도, 신뢰도를 기준으로 구성한 자살 위험 자가진단표
 - Social desirability bias
 - 자살 위험을 self-reporting 하는 부분에서 생기는 reporting error
 - 데이터 사이언스, 빅데이터 기반 자살 위험도 평가 시스템
 - 자살 고위험군 선별에 적용 가능한 모형
 - 소셜 빅데이터에서의 측정오차 이슈
 - 실제 자살위험 예측에 활용 불가능

Research Objectives

- 머신러닝을 활용하여 자살위험도를 평가하지만
- 소셜 빅데이터보다는 자기보고형 데이터를 활용하고
- 응답자가 조사에 부담을 느끼지 않으며
- 정신건강에 대한 전문성이 없는 비의료인도 쉽게 활용할 수 있고
- 실시간으로 자살생각, 자살행동 위험을 정량화 해줄 수 있는
- 핸드폰 앱 or 인터넷 웹페이지 기반 자살 위험 스크리닝 시스템의 구축

-> 1차 의료기관, 사회복지센터, 상담센터 등에서의 활용 가능성을 우선적으로 고려

Research Objectives

The screenshot shows a web browser window with the address bar displaying <http://ami.ajou.ac.kr/bcr/>. The page title is "BREAST CANCER RECURRENCE PREDICTION". Below the title, there are navigation links for "Home" and "About". The main heading is "BREAST CANCER RECURRENCE PREDICTION BASED ON SVM (BCRSVM)".

The form contains the following input fields and radio buttons:

- Histologic grade: ☐ Grade 1 ☐ Grade 2 ☒ Grade 3
- Local invasion of tumor: ☐ No ☒ Yes
- Number of tumor:
- Tumor size (cm):
- Lymphovascular invasion: ☐ No ☒ Yse
- Estrogen receptor: ☒ Negative ☐ Positive
- Number of matastatic lymph node:

A "Calculate" button is located below the input fields. The result displayed is: "Recurrence probability at 5 years after beast cancer : 76.2 %".

출처: Kim, W., Kim, K. S., Lee, J. E., Noh, D. Y., Kim, S. W., Jung, Y. S., ... & Park, R. W. (2012). Development of novel breast cancer recurrence prediction model using support vector machine. *Journal of Breast Cancer*, 15(2), 230-238.

Data Description

- 한국복지패널, 2012-2019
 - 연도별 조사
 - 높은 표본유지율
 - 전국의 모수치와 유사한 표본
 - 높은 저소득층 가구 비율
 - 자살생각, 계획, 시도에 대한 질문 포함
- 만 19-64세의 응답자 11,114명
- Number of observations: 60,568 (5.45 obs. per respondent)

Measures of Suicide

- Suicidal ideation:

Have you seriously considered committing suicide in the past year?

(1646 observations w/ experiences of suicidal ideation)

- Suicide planning or attempt:

Have you made a concrete plan to commit suicide in the past year?

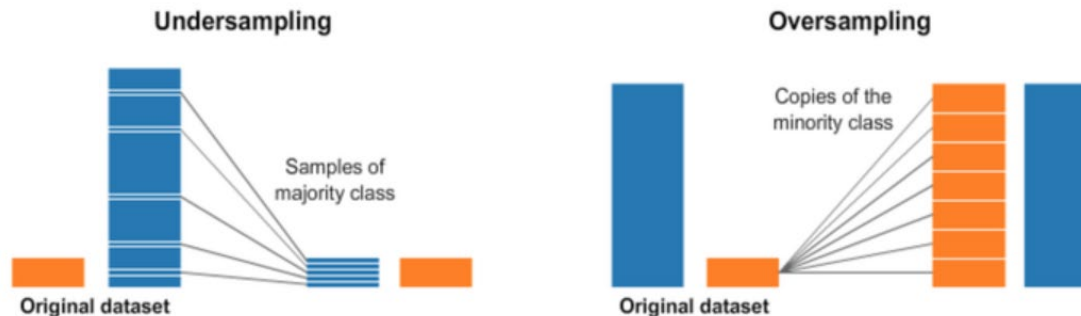
Have you made an attempt to commit suicide in the past year?

(244 observations w/ experiences of suicide planning or attempt)

– Response options: yes; no; refused to answer

Class Imbalance Problem

- 불균형 데이터의 문제:
 - 대부분의 응답자들은 $y=0$ 의 경우
 - 불균형 데이터를 그대로 사용해 모델을 학습하게 되면 $y=0$ 인 관측치들이 cost function 갱신과정의 대부분을 차지하게 되어 학습을 압도하게 됨
- Solutions:
 - Data resampling (undersampling or oversampling)



- Classification by an optimal cutoff point

Class Imbalance Problem

- Undersampling:
 - Data A (3292 observations):
 - 1646 observations w/ experiences of suicidal ideation +
 - 1646 observations randomly selected from the remaining sample
 - Data B (488 observations):
 - 244 observations w/ experiences of suicide planning or attempt +
 - 244 observations randomly selected from the remaining sample

Descriptive Statistics

Table 1. Descriptive statistics

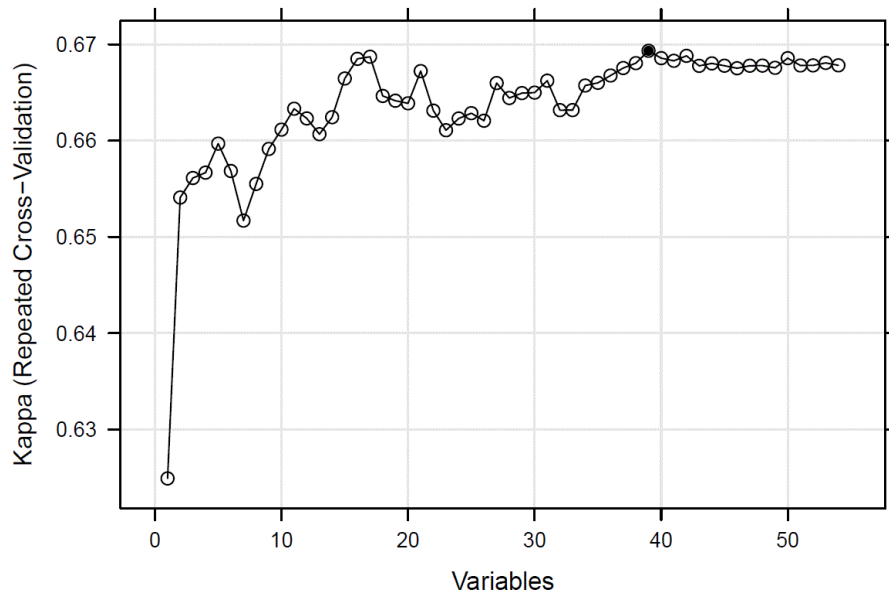
Outcome:	Full sample (N=60,568)		Suicidal ideation (N=3292)		Suicide planning or attempt (N=488)	
	Mean	SD	Mean	SD	Mean	SD
Age (19-64)	43.9	12.2	45.7	12.2	45.9	12.4
Female (0,1)	0.54		0.57		0.60	
Education background (0,1)	0.45		0.34		0.31	
Marital status (0,1)	0.67		0.58		0.57	
No. of household members	3.28	1.21	3.01	1.29	2.98	1.26
Employment status (0,1)	0.71		0.62		0.56	
Region of residence (0,1)	0.40		0.41		0.43	
Religion (0,1)	0.46		0.45		0.46	
Household income [†]	5962.5	5900.4	4803.3	3776.4	4598.4	3603.6
Household consumption [†]	423.3	249.5	360.2	238.0	344.3	230.2
Household net worth	13625.9	36446.7	10024.8	31557.9	9378.9	32144.1
Social welfare receipt (0,1)	0.06		0.15		0.19	
No. of outpatient visits	10.2	19.4	16.7	30.0	18.6	32.6
Poor self-rated health (0,1)	0.25		0.40		0.45	
Disability (0,1)	0.06		0.12		0.14	
Any chronic disease (0,1)	0.37		0.49		0.55	
Smoking (0,1)	0.22		0.26		0.27	
Drinking (0,1)	0.59		0.55		0.51	
CESD score (0-33)	2.68	4.03	6.68	7.10	8.31	8.42
Self-esteem score (0-30)	21.3	3.81	18.8	5.15	17.8	5.86
Experience of physical abuse from spouse (0,1)	0.67		0.60		0.58	

Notes: N, number of observations; SD, standard deviation. [†] in the 2019 South Korean won.

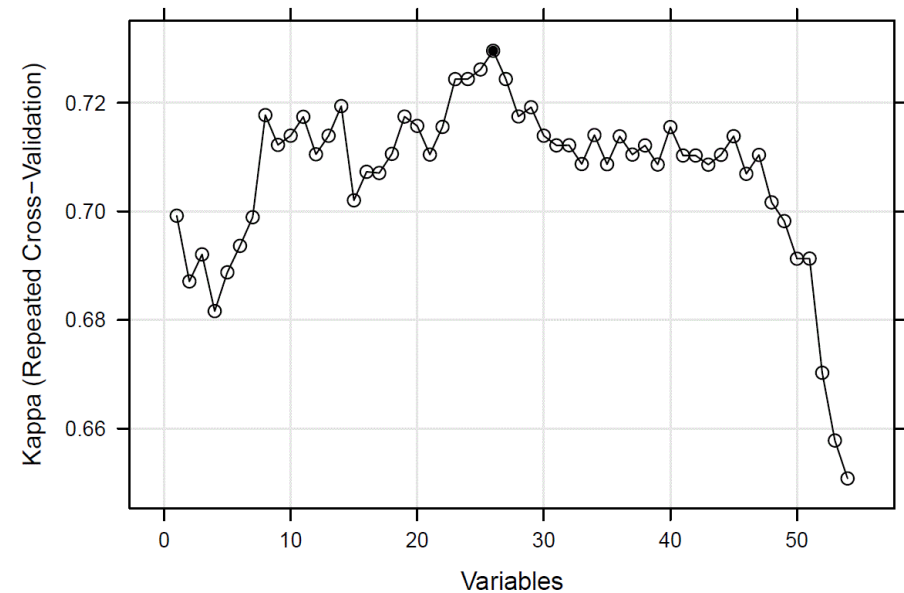
Predictor Selection

- Structured literature search through the PubMed database
- 응답자 배경 및 개인사, 가구여건, 사회복지 가입 및 수급, 소득/소비, 경제활동상태, 정신/신체건강, 유년기 특성 관련 55개 변수 선택
<https://drive.google.com/file/d/1Eoc7Cpb9X2IKAdU2oghM5km0pyZpE-tC/view?usp=sharing>
- Recursive feature elimination:
 - 로짓모형을 활용하여 recursive feature elimination을 수행 (10-fold cross validation with three repeats)
 - 각 모형의 classification performance는 kappa 지표로 평가

Recursive Feature Elimination



39 predictors for models predicting suicidal ideation

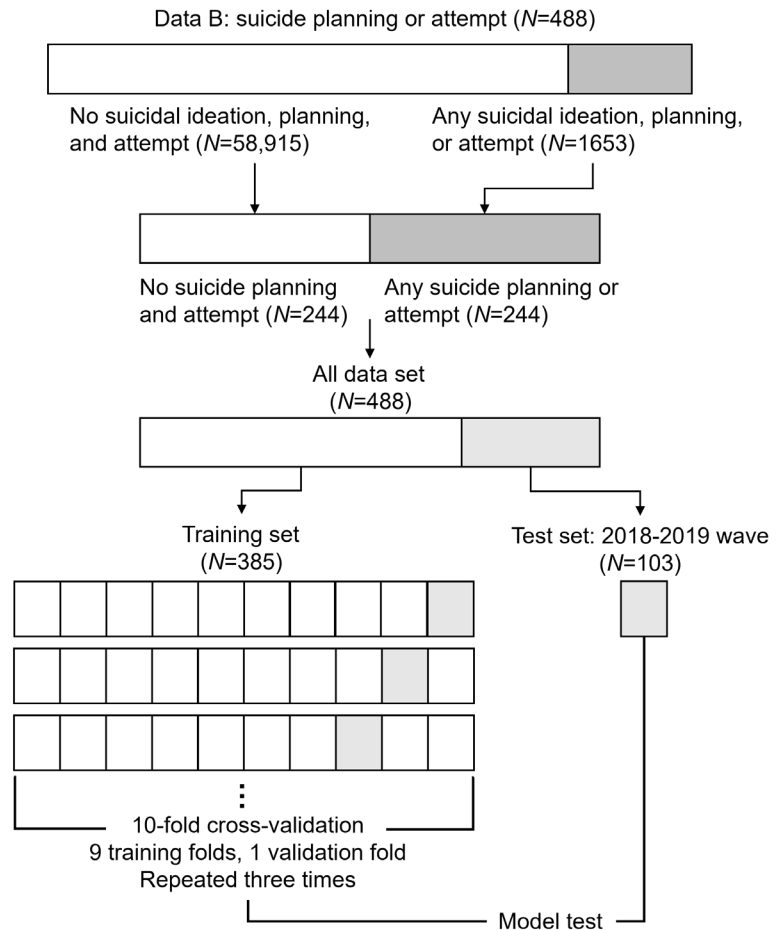
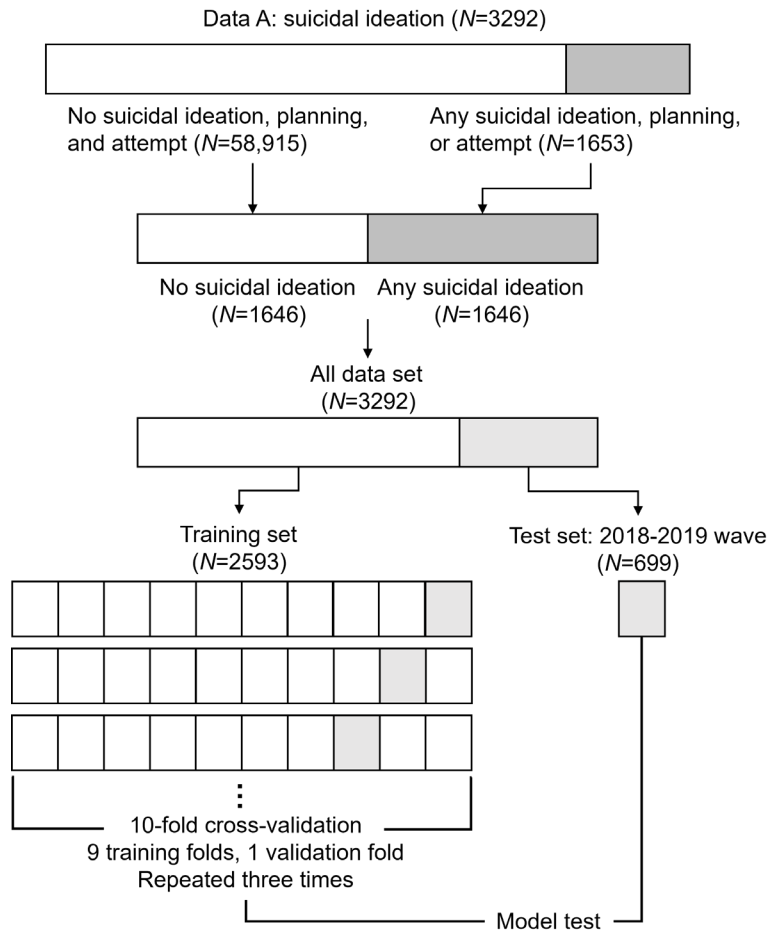


26 predictors for models predicting suicide planning or attempt

Algorithm Development and Validation

- Four ML algorithms:
 - Logistic regression
 - Support vector machine (SVM)
 - Random forest (RF)
 - Extreme gradient boosting (XGBoost)
- Training set: 2012-2017 waves, Test set: 2018, 2019 waves
- Hyperparameter tuning using a grid search on 10 randomly selected training and validation sets (10-fold cross-validation with three repeats)
- 분류 정확도 평가 지표: area under the curve (AUC), sensitivity, specificity, positive predictive value, negative predictive value, accuracy

Algorithm Development and Validation



Hyperparameter Tuning

Table 2. Tuned hyperparameters

Panel A: suicidal ideation ($N=3292$)	
SVM	kernel=linear, cost=1.5
RF	split rule=Gini, minimal node size=2, number of randomly selected predictors=6
XGBoost	number of boosting iterations=120, max tree depth=3, shrinkage=0.04, minimum loss reduction=2, subsample ratio of columns=0.55, minimum sum of instance weight=5, and subsample percentage=1
Panel B: suicide planning or attempt ($N=488$)	
SVM	kernel=linear, cost=0.5
RF	split rule=Gini, minimal node size=2, number of randomly selected predictors=7
XGBoost	number of boosting iterations=60, max tree depth=5, shrinkage=0.04, minimum loss reduction=3, subsample ratio of columns=0.5, minimum sum of instance weight=6, and subsample percentage=1

Notes: SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

Results

Table 3. Algorithm performance

	Logistic	SVM	RF	XGBoost
Panel A: suicidal ideation ($N=3292$)				
Area under the curve	0.837	0.844	0.851	0.861
Sensitivity	0.808	0.811	0.850	0.853
Specificity	0.867	0.877	0.852	0.869
Positive predictive value	0.808	0.820	0.799	0.819
Negative predictive value	0.867	0.870	0.891	0.895
Accuracy	0.843	0.850	0.851	0.863
Panel B: suicide planning or attempt ($N=488$)				
Area under the curve	0.872	0.872	0.857	0.880
Sensitivity	0.861	0.861	0.814	0.861
Specificity	0.883	0.883	0.900	0.900
Positive predictive value	0.841	0.841	0.854	0.861
Negative predictive value	0.898	0.898	0.871	0.900
Accuracy	0.874	0.874	0.864	0.884

Notes: SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

Results

Table 4. Top five most important predictors

Algorithm:	Logistic	SVM	RF	XGBoost
Panel A: suicidal ideation ($N=3292$)				
Feature 1	CESD score	CESD score	CESD score	CESD score
Feature 2	Self-esteem	Self-esteem	Self-esteem	Self-esteem
Feature 3	Satisfaction with family relation	Income	Income	Income
Feature 4	Smoking	Life satisfaction	Consumption	Satisfaction with family relation
Feature 5	Unpaid rent	Consumption	Net worth	Life satisfaction
Panel B: suicide planning or attempt ($N=488$)				
Feature 1	CESD score	CESD score	CESD score	CESD score
Feature 2	Mother's education background	Self-esteem	Self-esteem	Self-esteem
Feature 3	Smoking	Income	Income	Life satisfaction
Feature 4	Religion	Life satisfaction	Consumption	Income
Feature 5	Age	Consumption	Life satisfaction	Smoking

Notes: SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

Results

Table 5. Algorithm performance for the reduced model

	Logistic	SVM	RF	XGBoost
Panel A: suicidal ideation ($N=3292$)				
Area under the curve	0.835	0.838	0.845	0.843
Sensitivity	0.801	0.804	0.839	0.829
Specificity	0.869	0.872	0.850	0.857
Positive predictive value	0.809	0.813	0.795	0.801
Negative predictive value	0.863	0.865	0.884	0.878
Accuracy	0.841	0.844	0.846	0.846
Panel B: suicide planning or attempt ($N=488$)				
Area under the curve	0.869	0.877	0.865	0.869
Sensitivity	0.837	0.837	0.814	0.837
Specificity	0.900	0.917	0.917	0.900
Positive predictive value	0.857	0.878	0.875	0.857
Negative predictive value	0.885	0.887	0.873	0.885
Accuracy	0.874	0.884	0.874	0.874

Notes: SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

Results

Table A3. Algorithm performance, train and test set randomly drawn from the full sample

	Logistic	SVM	RF	XGBoost
Panel A: suicidal ideation ($N=3292$)				
Area under the curve	0.829	0.837	0.832	0.830
Sensitivity	0.797	0.806	0.819	0.815
Specificity	0.860	0.867	0.844	0.846
Positive predictive value	0.835	0.844	0.824	0.825
Negative predictive value	0.826	0.834	0.839	0.837
Accuracy	0.830	0.838	0.832	0.831
Panel B: suicide planning or attempt ($N=488$)				
Area under the curve	0.852	0.843	0.869	0.869
Sensitivity	0.789	0.788	0.824	0.824
Specificity	0.915	0.898	0.915	0.915
Positive predictive value	0.931	0.918	0.933	0.933
Negative predictive value	0.750	0.747	0.783	0.783
Accuracy	0.840	0.833	0.861	0.861

Notes: SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

Table A4. Algorithm performance, using one observation for each participant

	Logistic	SVM	RF	XGBoost
Panel A: suicidal ideation ($N=2515$)				
Area under the curve	0.827	0.836	0.847	0.838
Sensitivity	0.814	0.822	0.814	0.838
Specificity	0.861	0.871	0.881	0.878
Positive predictive value	0.803	0.817	0.810	0.804
Negative predictive value	0.861	0.867	0.883	0.874
Accuracy	0.839	0.849	0.855	0.847
Panel B: suicide planning or attempt ($N=438$)				
Area under the curve	0.842	0.842	0.858	0.842
Sensitivity	0.800	0.800	0.800	0.800
Specificity	0.883	0.883	0.917	0.883
Positive predictive value	0.821	0.821	0.865	0.821
Negative predictive value	0.869	0.869	0.873	0.869
Accuracy	0.850	0.850	0.870	0.850

Notes: SVM, support vector machine; RF, random forest; XGBoost, extreme gradient boosting.

Conclusion

- ML, 자기보고형 데이터를 활용하여 85-90%의 분류 정확도를 갖는 자살 예측 시스템 구축 가능
- Best performing model -> XGBoost
- 모형의 해석 가능성을 고려한다면 Logit 모형을 선택
- Depression, self-esteem만을 예측 변수로 활용하는 reduced model도 거의 비슷한 수준의 분류 정확도를 보임
- Best subset of predictor를 찾는 방식으로 분류 정확도의 손실 없이 단순한 형태의 ML 모형을 구축할 수 있음
- 모형 추정 결과를 활용하여 앱 기반 or 웹페이지 기반 실시간 자살예측 시스템 구축 가능

Things to Note

- 상대적으로 social desirability bias가 적은 변수를 활용
- 자기보고형 데이터를 input으로 활용하여 실제 현장에서도 활용 가능
- 비의료인도 쉽게 administration할 수 있는 단순한 시스템
- Suicidal chain의 끝단이 아닌 early stage symptom을 예측하는 방식으로 조기 선별, 개입을 가능하게 해주는 방식
- 자살 스크리닝 시스템 자체를 구축하는 것이 목적이 아닌 reduced model을 만드는 과정을 보여줌으로써 다방면에 적용 가능하다는 점
- Nothing new.... 선행연구에서 보고된 자살 위험인자 대부분이 자살 위험을 예측하는 ML모형에서 그대로 활용

Limitations

- 다양한 특성을 갖는 개인/가구로 구성된 샘플
- 자살 위험이 가장 높은 노인 인구를 고려하지 못함
- 소수의 suicide ideator, planner, and attempter
- 모집단에 대한 대표성 부족
- Still.... social desirability bias

Future Research

- 국민건강영양조사를 활용한 replication study
- 빈곤층, 차상위층, 특정 취약계층에 적용 가능한 자살위험 예측 모형의 개발
 - 65세 이상 노인
 - 중고등학생
 - 이주민, 이주노동자
 - 자영업자, 임시근로자/일용근로자 등
- 유족의 정신건강, 자살위험을 예측해줄 수 있는 경보시스템의 개발
- 자살 위험 요인을 포괄하는 전국규모의 조사 수행