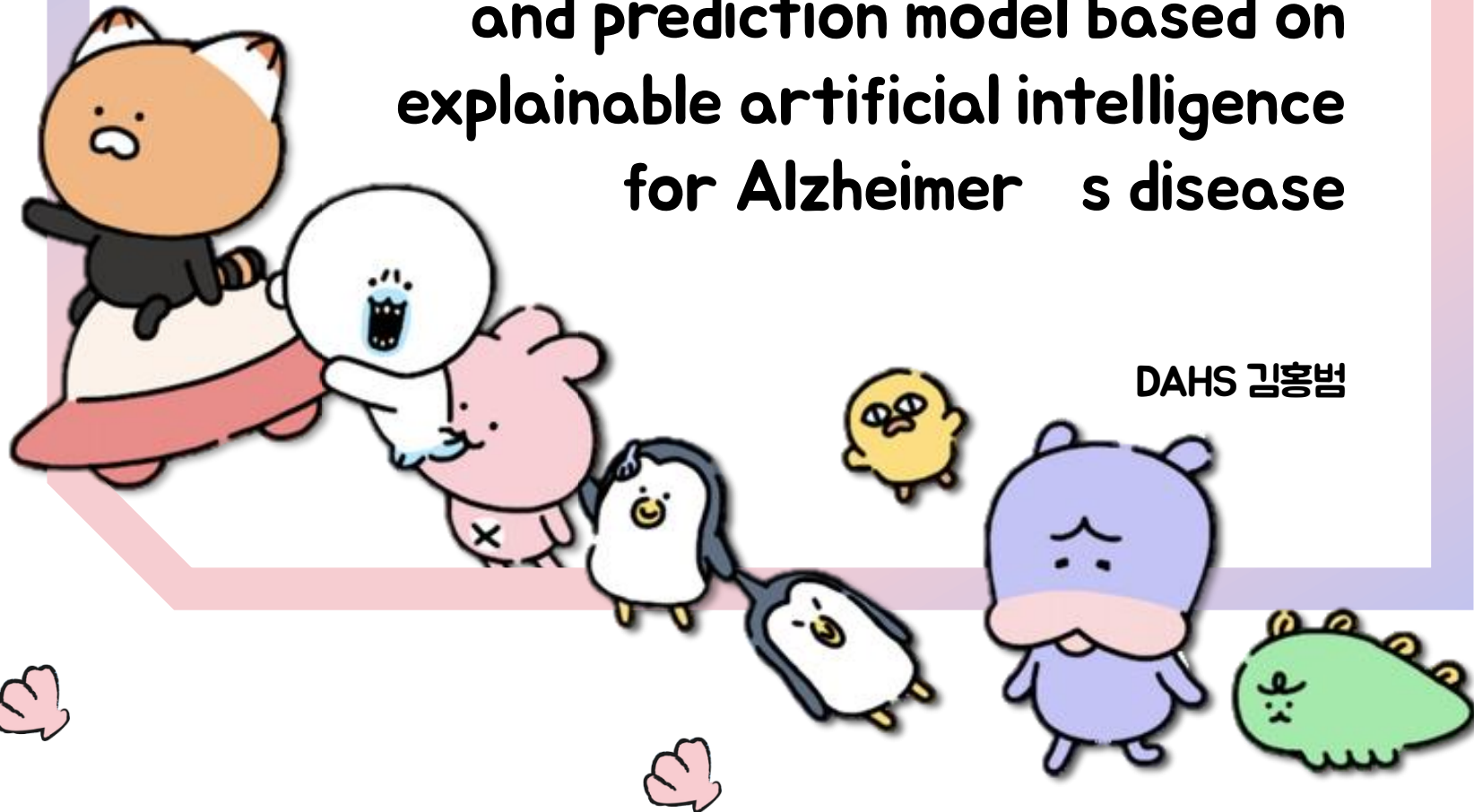


**A multilayer multimodal detection  
and prediction model based on  
explainable artificial intelligence  
for Alzheimer's disease**

DAHS 김홍범





# Contents



Introduction



Preprocessing



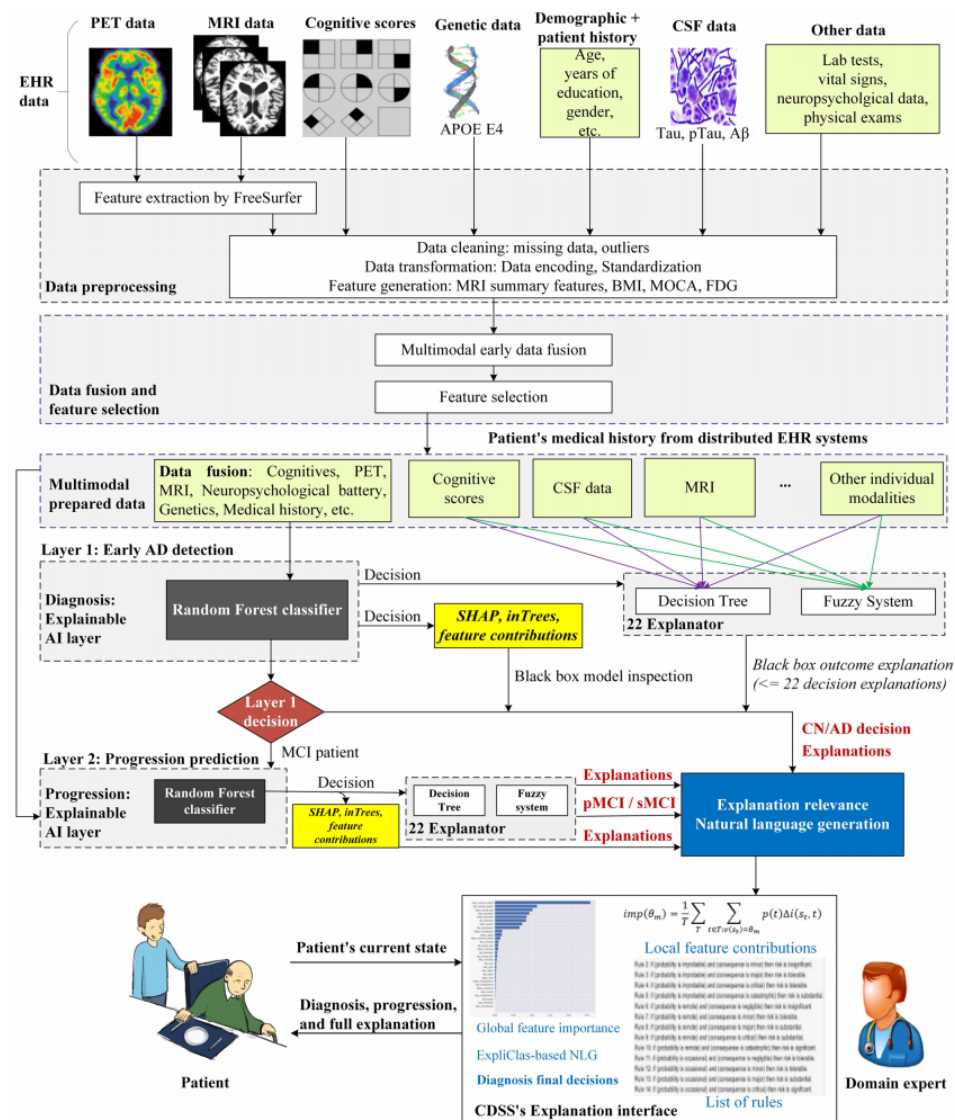
Modeling



Interpretability



Results





# Introduction

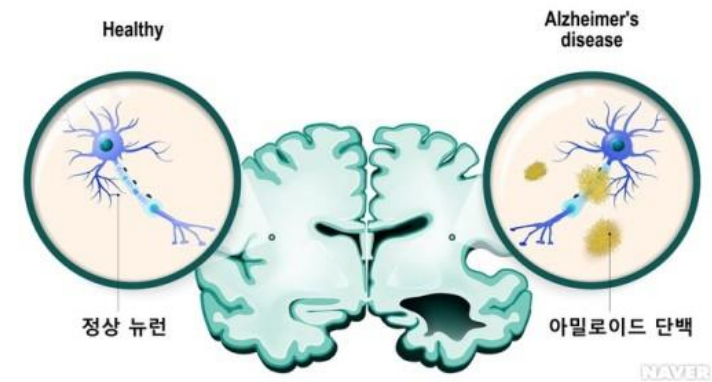
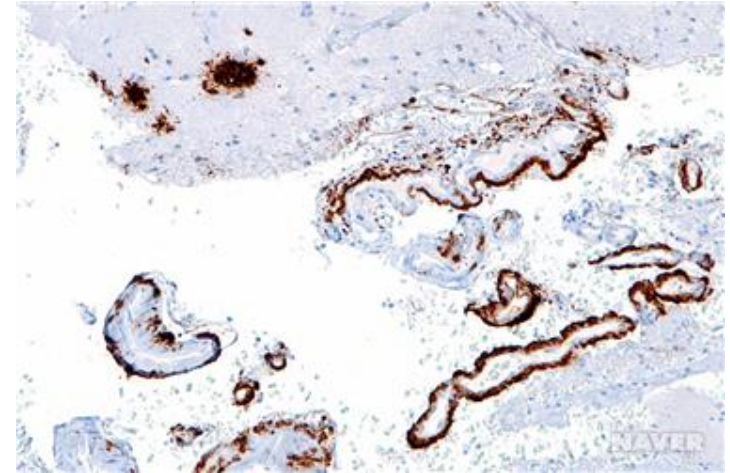


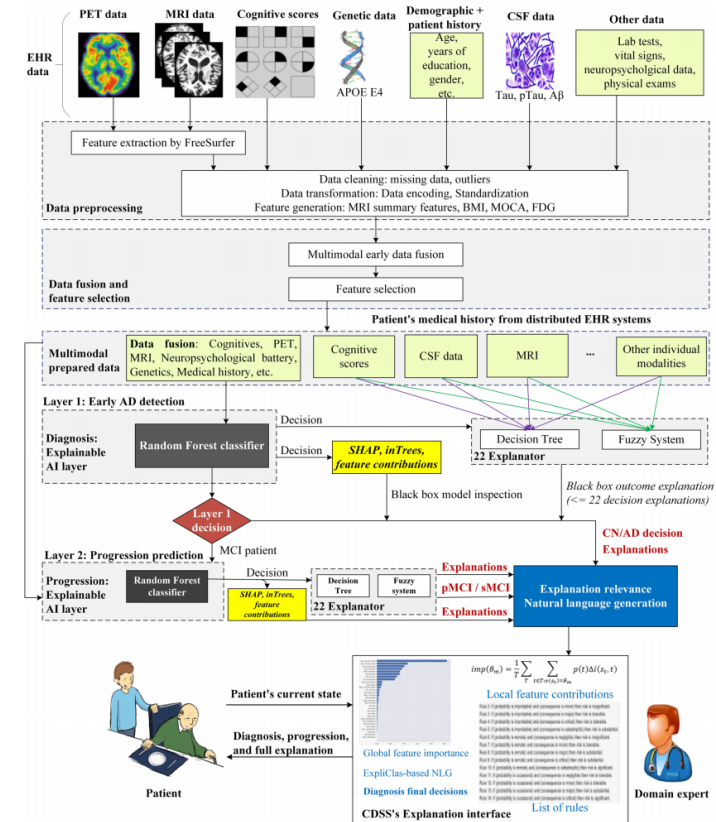
## Alzheimer's disease(AD) 연구의 어려움

- 대부분의 연구가 neuroimaging 데이터셋에만 의존
- Diagnosis와 progression이 독립적인 문제로서 연구
- 모델 결과 해석을 등한시하고 모델 성능 최적화에만 연구가 집중

## Alzheimer's disease(AD)란?

- 만성 신경 퇴행성 질환으로서, 아밀로이드반의 비정상적인 축적과 뇌의 신경섬유매듭이 엉키는 특징을 가지고 있음
- AD 환자는 행동변화와 함께 기억력, 사고력, 언어 능력이 감소
- 현재까지 치료법이 발견되지 않았기 때문에 조기 발견이 매우 중요함

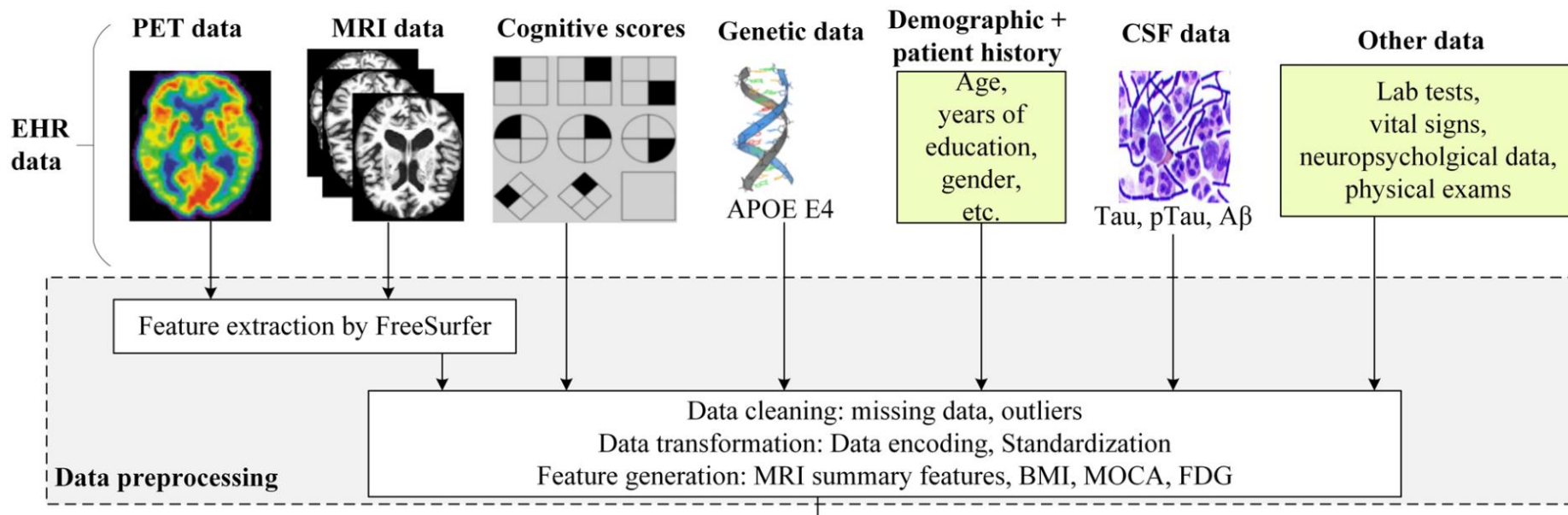
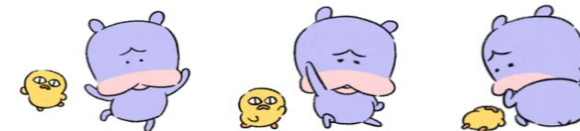








# Preprocessing



## Material and methods

- DATASET : 1048 Patients
- 294명 Cognitively normal
- 254명 sMCI (MCI판정만 받음)
- 232명 pMCI (3년이내 AD로 판정)
- 268명 - Alzheimer's disease(AD)

PET(Positron emission tomography) data : 양전자 방사 단층 촬영 데이터

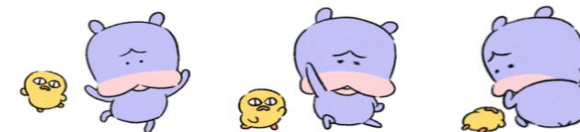
CSF(cerebrospinal fluid) data : 뇌척수(중추신경) 데이터

FreeSurfer : 뇌영상 이미지에서 자동으로 중요 feature 추출하는 소프트웨어





# Preprocessing



	CN (n = 294)	sMCI (n = 254)	pMCI (n = 232)	AD (n = 268)	Combined (n = 1048)
Gender (M/F)	140 ± 154	144/110	136/96	151 ± 117	571/477
Age (years)	74.120 ± 5.890	72.202 ± 07.553	73.771 ± 7.0840	75.241 ± 7.610	73.864 ± 07.107
Education (years)	16.405 ± 2.733	15.9530 ± 2.867	15.784 ± 2.7830	15.175 ± 2.923	15.844 ± 02.858
FAQ	0.1940 ± 0.720	01.539 ± 02.817	5.7110 ± 4.8680	13.146 ± 6.814	05.053 ± 06.754
MMSE	29.085 ± 1.143	27.941 ± 01.722	26.7590 ± 1.736	23.235 ± 2.015	26.797 ± 2.7960
MOCA	25.569 ± 1.866	23.493 ± 02.452	20.947 ± 01.908	17.553 ± 3.377	21.993 ± 3.9450
FDG	6.5690 ± 0.477	06.3820 ± 0.599	05.800 ± 00.462	5.4060 ± 0.614	6.0560 ± 0.7180
APOE4	0.2520 ± 0.472	0.4610 ± 0.6380	0.8660 ± 00.686	0.8880 ± 0.710	0.6010 ± 0.6850
CSF PTAU pg/mL	19.423 ± 6.820	25.640 ± 11.703	35.2240 ± 13.20	35.717 ± 13.11	28.5940 ± 13.29
CSF TAU pg/mL	215.07 ± 67.28	270.861 ± 106.4	352.86 ± 116.27	361.2 ± 121.41	296.46 ± 120.58
ADAS-Cog 11	05.617 ± 2.784	8.6260 ± 3.5200	13.412 ± 4.3850	19.318 ± 6.569	11.576 ± 06.970
ADAS-Cog 13	08.600 ± 4.108	13.791 ± 05.303	21.580 ± 05.841	29.706 ± 7.835	18.129 ± 10.085
RAVLT immediate	045.595 ± 9.64	37.705 ± 10.308	27.444 ± 06.510	22.466 ± 7.069	33.750 ± 12.585
RAVLT learn	06.139 ± 2.143	04.799 ± 02.403	02.853 ± 02.219	01.799 ± 1.810	3.9770 ± 02.752
RAVLT forgetting	03.582 ± 2.810	04.343 ± 02.497	05.039 ± 02.193	04.381 ± 1.783	4.2930 ± 02.420
RAVLT % forget	32.612 ± 27.53	50.000 ± 30.027	78.188 ± 27.892	88.562 ± 21.22	61.223 ± 35.098
CDR-SB	0.0390 ± 0.141	1.1970 ± 0.6390	02.004 ± 0.9980	04.347 ± 1.591	01.856 ± 01.896
GDTOTAL	0.7890 ± 1.056	1.7090 ± 01.462	01.668 ± 01.423	01.634 ± 1.454	01.423 ± 01.404
HCI	8.9500 ± 3.330	11.066 ± 04.080	15.560 ± 04.770	21.158 ± 7.384	14.048 ± 06.996
Hippo. vol. (cm <sup>3</sup> ) (/1000)	7.4520 ± 0.920	07.106 ± 01.074	06.083 ± 01.038	05.713 ± 0.995	06.621 ± 01.240

**Table 11.** Descriptive statistics from the dataset used. AD, Alzheimer's disease; MCI, mild cognitive impairment; pMCI-sMCI, progressive MCI – stable MCI; CN, cognitive normal; CDR, clinical dementia rating; ADAS-Cog, Alzheimer's Disease Assessment Scale–Cognitive Subscale test; RAVLT, Rey Auditory Verbal Learning Test; FAQ, Functional Assessment Questionnaire; MMSE, Mini–Mental State Examination; FDG, sum of mean glucose metabolism uptake in regions of angular, temporal, and posterior cingulate; TAU, CSF level of TAU; Aβ42, CSF level of amyloid β1–42 peptide; HCI, hypometabolic convergence index; AV45, Average AV45 SUVR of frontal, anterior cingulate, precuneus, and parietal cortex relative to the cerebellum; Hippo, Hippocampus; GDTOTAL, Geriatric Depression Scale. \*Data are mean ± standard deviation.





# Preprocessing

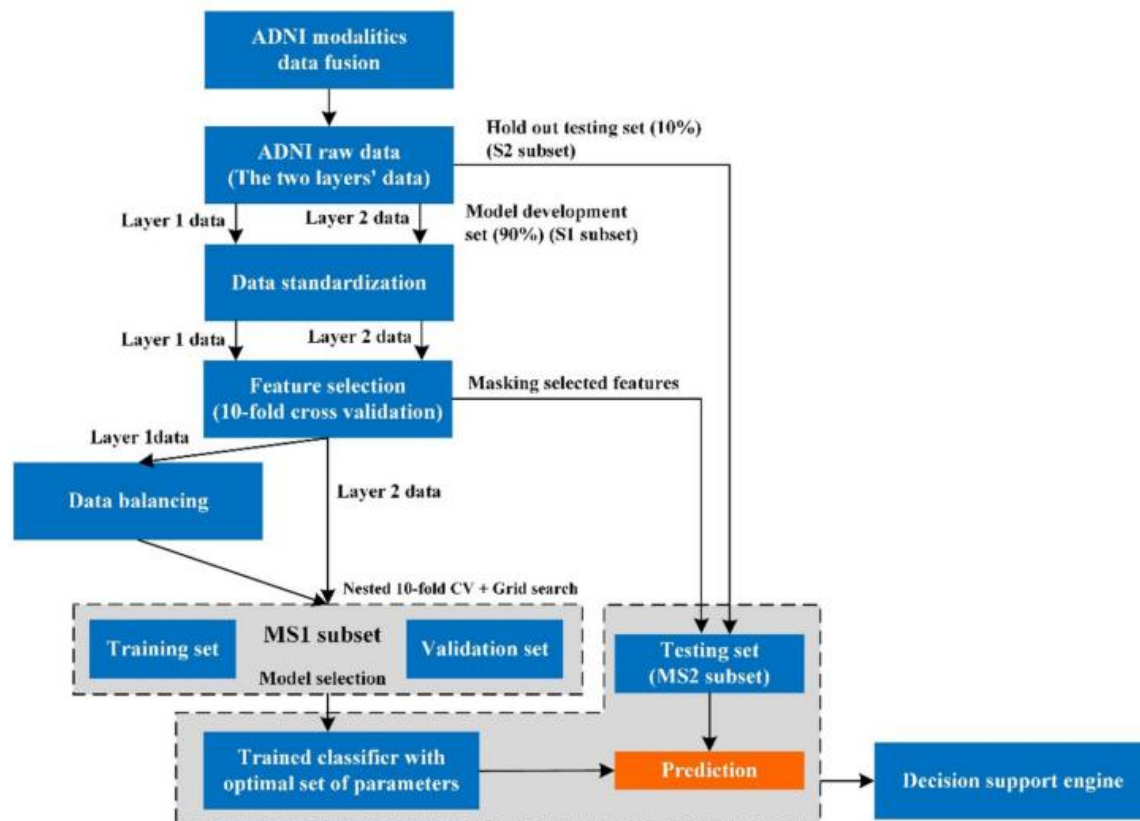
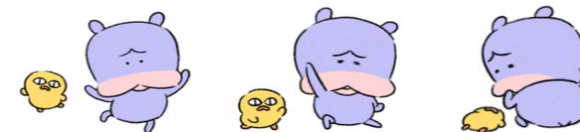
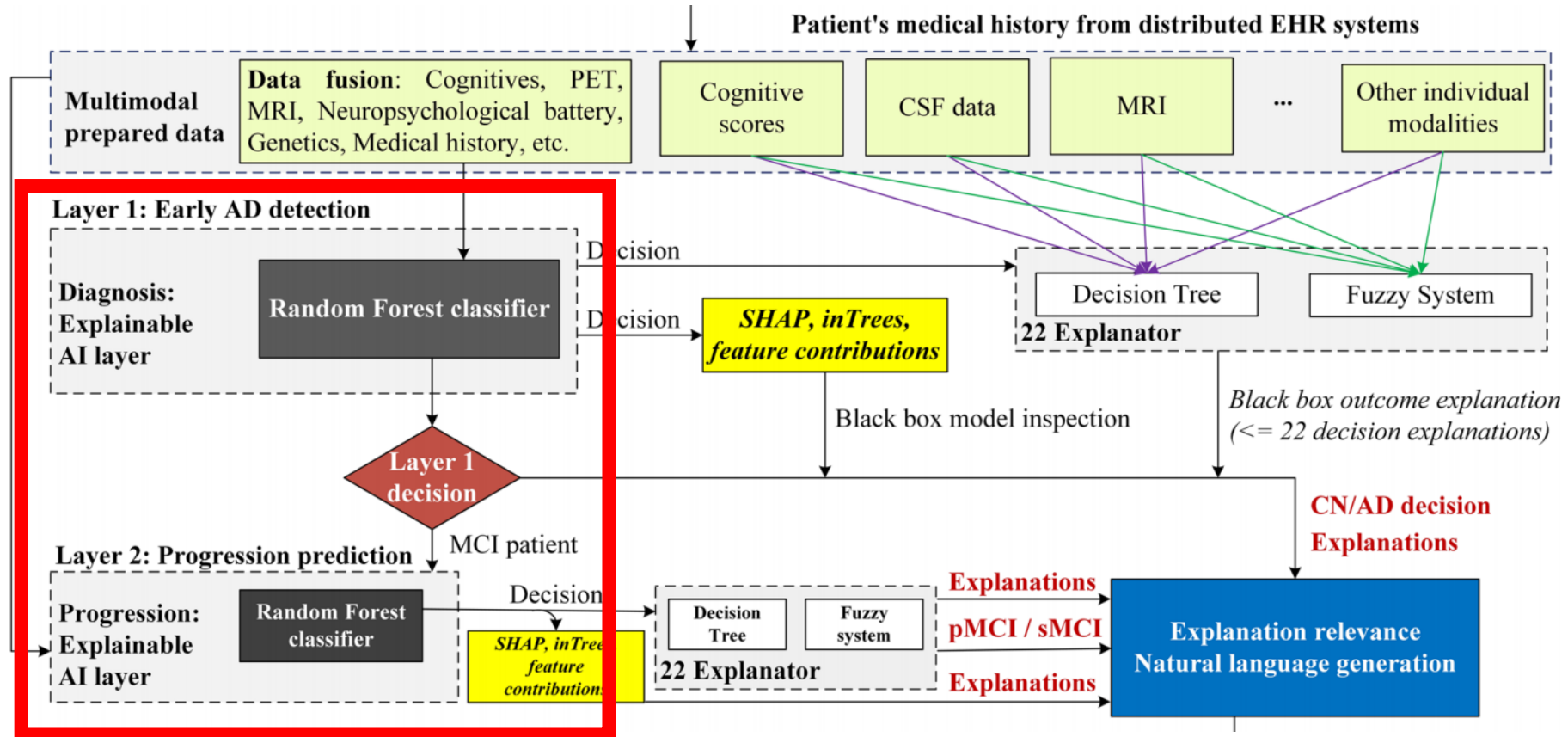
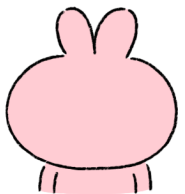


Figure 5. Development process for the oracle model in each layer.

1. train\_test\_split : S1(90%), S2(10%)
2. Standardization : 표준 정규분포로 정규화 진행
3. S1(train set)을 이용한 Feature selection 진행과 CNN을 이용한 결측치 처리 진행
4. 첫번째 모델의 imbalance 존재 (28.05% CN, 46.37% MCI, and 25.58% AD) -> oversampling 실시 (SMOTE)
5. Nested k-fold stratified cross-validation을 이용한 하이퍼파라미터 튜닝 진행









# Modeling



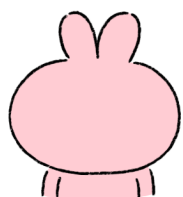
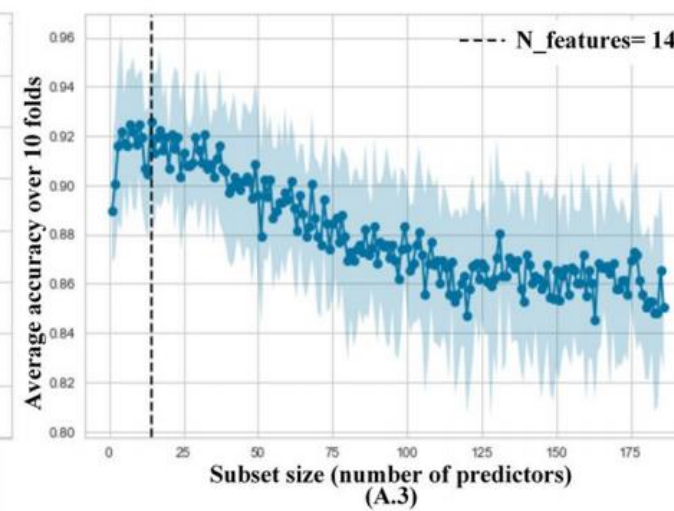
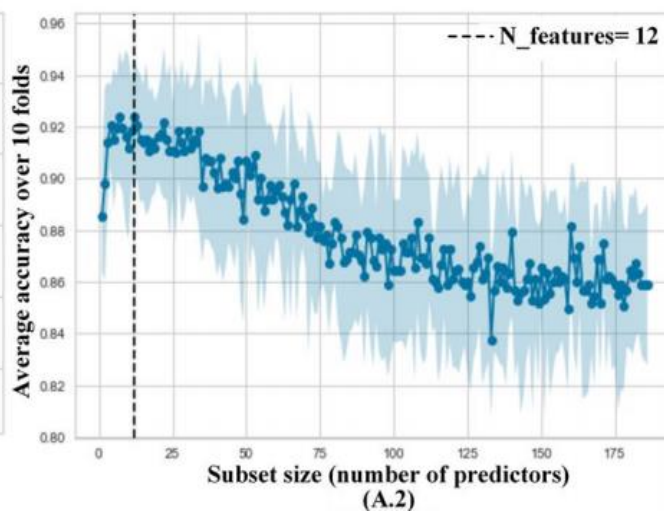
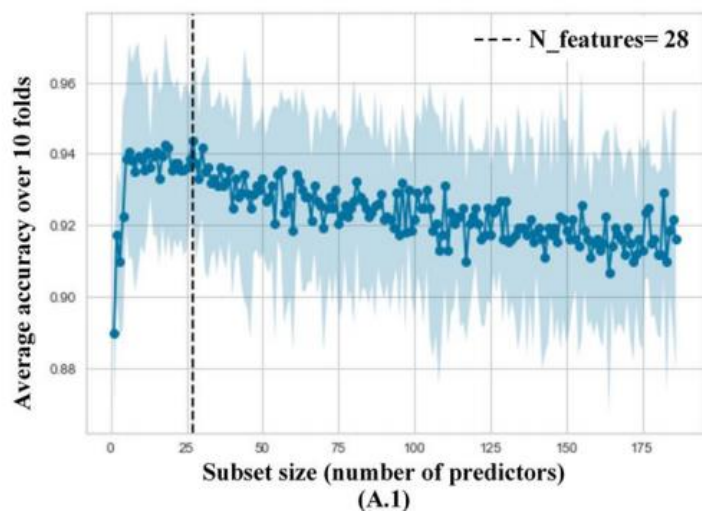
## Layer1 : multilable classification

Label : CN(정상), MCI(Progression), AD(알츠하이머)

A1 : RF-RFE 94.4%(28 features)

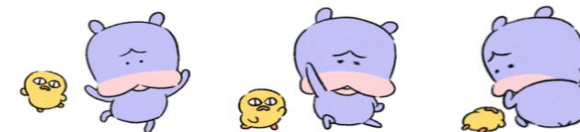
A2 : SVM-RFE 92.4%(12 features)

A3 : GB-RFE 92.6%(14 features)





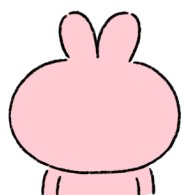
# Modeling



Modalities	Precision (%)			Recall (%)			MCA (%)	MCF (%)
	CN	MCI	AD	CN	MCI	AD		
All modalities	98.83 ± 1.33	90.91 ± 2.55	92.41 ± 1.99	96.21 ± 1.11	95.45 ± 2.00	86.61 ± 2.10	93.42 ± 2.73	93.39 ± 2.19
CS	99.60 ± 0.17	89.76 ± 1.99	88.51 ± 1.04	93.94 ± 0.18	93.64 ± 1.08	87.03 ± 2.09	92.00 ± 2.26	92.08 ± 2.00
NB	84.86 ± 1.32	74.11 ± 2.06	77.93 ± 2.00	80.68 ± 1.09	80.68 ± 1.92	69.46 ± 2.01	77.83 ± 2.33	77.94 ± 2.10
MRI	45.19 ± 2.46	47.16 ± 4.34	48.99 ± 3.01	46.21 ± 2.41	50.91 ± 3.99	40.59 ± 3.31	46.99 ± 4.01	46.50 ± 3.91
PET	61.59 ± 1.89	65.06 ± 2.22	70.35 ± 3.03	70.45 ± 1.40	61.36 ± 3.50	66.53 ± 2.87	65.23 ± 2.98	65.89 ± 2.11
Genetics	61.28 ± 2.60	58.82 ± 3.33	54.90 ± 2.42	68.94 ± 2.00	59.09 ± 3.91	46.86 ± 2.00	58.75 ± 3.11	58.31 ± 2.89
MH	46.60 ± 1.42	50.08 ± 2.91	29.29 ± 1.67	36.36 ± 1.79	67.95 ± 3.82	17.15 ± 1.77	46.22 ± 2.99	41.22 ± 3.00
CS + NB	99.22 ± 1.09	90.65 ± 2.06	90.79 ± 0.90	95.83 ± 1.11	94.77 ± 2.31	86.61 ± 2.90	93.00 ± 2.61	92.97 ± 2.25
CS + MRI	99.20 ± 2.11	89.42 ± 3.62	89.57 ± 2.39	93.94 ± 2.70	94.09 ± 3.59	86.19 ± 3.20	92.05 ± 3.99	92.06 ± 4.01
CS + PET	99.20 ± 2.01	89.25 ± 2.99	89.87 ± 2.51	94.32 ± 2.22	94.32 ± 3.87	85.36 ± 2.91	92.05 ± 3.33	92.05 ± 3.41
CS + Genetics	99.20 ± 1.99	89.01 ± 2.53	89.08 ± 2.33	93.94 ± 1.83	93.86 ± 3.01	85.36 ± 2.43	91.73 ± 2.91	91.74 ± 3.08
CS + MH	98.43 ± 1.55	89.15 ± 4.01	89.04 ± 1.92	94.70 ± 1.77	93.41 ± 4.31	84.94 ± 3.91	91.62 ± 4.00	91.61 ± 3.81
CS + NB + MRI	99.22 ± 2.05	90.87 ± 3.91	91.19 ± 2.71	96.21 ± 2.09	95.00 ± 3.61	86.61 ± 2.99	93.21 ± 3.40	93.18 ± 3.47
CS + NB + PET	99.22 ± 1.99	91.45 ± 2.06	90.83 ± 2.90	96.97 ± 1.78	94.77 ± 3.01	87.03 ± 2.72	93.42 ± 2.97	93.38 ± 2.61
<b>CS + NB + Genetics*</b>	<b>99.22 ± 1.01</b>	<b>91.72 ± 2.01</b>	<b>92.54 ± 0.91</b>	<b>96.21 ± 1.00</b>	<b>95.68 ± 2.00</b>	<b>88.28 ± 1.81</b>	<b>93.95 ± 2.30</b>	<b>93.94 ± 2.07</b>
CS + NB + MH	99.22 ± 2.01	91.50 ± 3.87	92.07 ± 2.33	96.59 ± 2.33	95.45 ± 4.10	87.45 ± 3.83	93.74 ± 4.00	93.71 ± 3.61
CS + NB + Genetics + MRI	99.22 ± 2.30	91.50 ± 4.10	92.11 ± 2.97	96.21 ± 2.22	95.45 ± 4.20	87.87 ± 4.86	93.74 ± 4.01	93.72 ± 4.44
CS + NB + Genetics + PET	99.61 ± 2.03	90.46 ± 2.06	90.27 ± 3.04	96.59 ± 1.83	94.77 ± 3.31	85.36 ± 2.77	92.89 ± 2.99	92.84 ± 3.20
CS + NB + Genetics + MH	99.22 ± 2.22	91.34 ± 4.39	92.83 ± 2.31	96.97 ± 2.40	95.91 ± 4.90	86.61 ± 3.09	93.85 ± 4.30	93.81 ± 4.06
CS + NB + Genetics + MH + MRI	99.22 ± 2.30	90.93 ± 4.55	92.41 ± 3.60	96.21 ± 2.32	95.68 ± 4.24	86.61 ± 3.95	93.53 ± 4.32	93.51 ± 3.97
CS + NB + Genetics + MH + PET	98.84 ± 1.89	91.29 ± 3.41	92.04 ± 2.87	96.59 ± 2.00	95.23 ± 3.05	87.03 ± 2.78	93.53 ± 3.11	93.50 ± 3.01

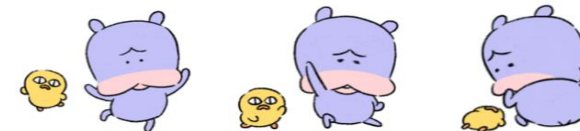
**Table 2.** Random Forest performance validation for detecting AD patients based on *tenfold cross-validation*. MCA: multiclass classification accuracy, MCF: multiclass F1 score; Asterisk (\*): is the subset of features with the best predictive performance; italic text is the best of single and pairs of modalities.

- 가장 성능이 좋은 RF-RFE 모델을 선정
- 6가지 selected modalities를 이용하여 최고의 성능을 가지는 데이터셋 선정
- CS : cognitive scores ( 8 features)
- NB : neuropsychological battery (6 features)
- MH : medical history (lab test and demographics)





# Modeling



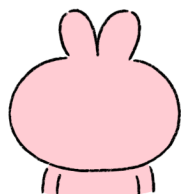
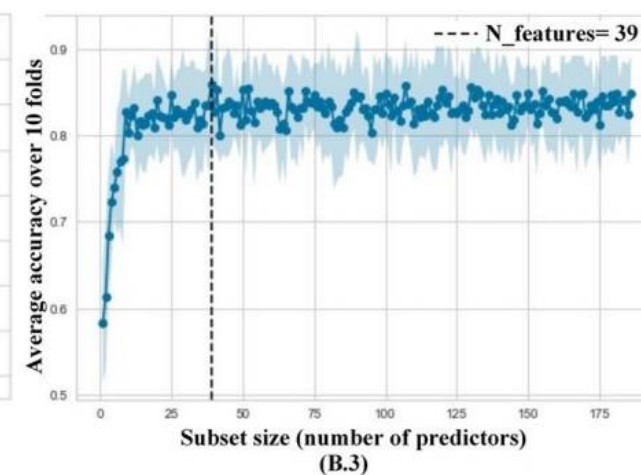
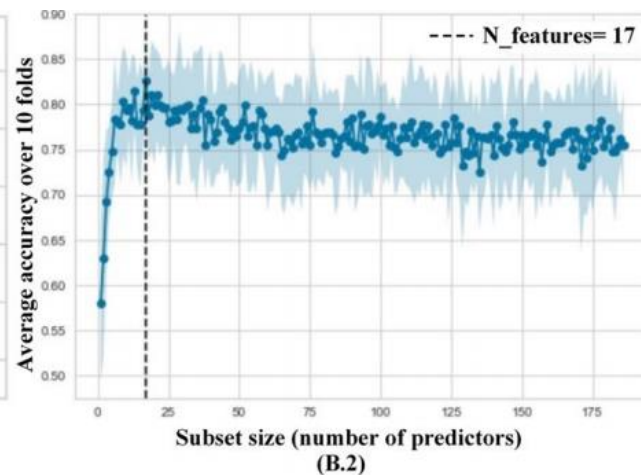
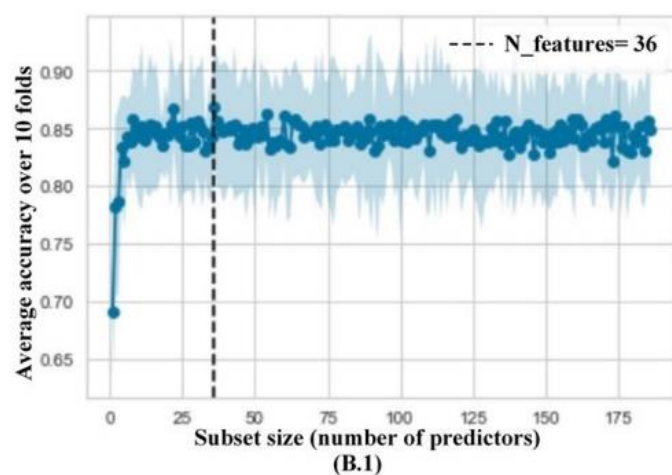
## Layer2 : Binary classification

Lable : pMCI, Smci

B1 : RF-RFE 86.8%(36 features)

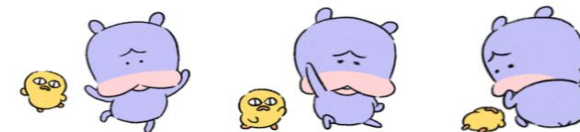
B2 : SVM-RFE 82.6%(17 features)

B3 : GB-RFE 86.0%(39 features)





# Modeling



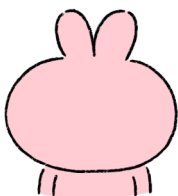
## Layer 2 experiments

Modalities used	Precision (%)	Recall (%)	Accuracy (%)	F1-score (%)	AUC
All	87.12 ± 1.52	81.31 ± 2.00	84.18 ± 1.77	83.21 ± 2.43	84.17 ± 1.99
CFA	82.14 ± 1.40	84.19 ± 1.90	82.16 ± 1.60	83.15 ± 1.50	82.15 ± 1.50
MRI	75.23 ± 1.88	72.25 ± 1.59	71.18 ± 2.01	72.17 ± 1.92	71.18 ± 1.89
PET	68.22 ± 2.22	68.53 ± 1.98	68.25 ± 1.99	66.39 ± 1.99	68.24 ± 2.01
Genetics	73.11 ± 1.73	68.36 ± 1.72	70.14 ± 1.79	69.24 ± 1.89	70.13 ± 1.80
MH	58.16 ± 4.60	52.22 ± 4.90	55.15 ± 3.73	54.19 ± 3.76	55.15 ± 3.76
CFA + MRI	83.11 ± 2.31	84.22 ± 2.20	83.14 ± 2.51	83.15 ± 2.26	83.14 ± 2.22
CFA + PET	85.17 ± 1.70	84.29 ± 2.90	84.19 ± 1.90	84.21 ± 2.10	84.19 ± 1.90
CFA + Genetics	82.14 ± 1.91	81.27 ± 1.98	81.16 ± 1.93	81.17 ± 1.96	81.16 ± 1.95
CFA + MH	84.16 ± 3.77	82.23 ± 4.60	82.17 ± 3.80	83.18 ± 3.33	82.17 ± 3.80
CFA + PET + MRI	86.09 ± 2.10	84.23 ± 2.30	84.11 ± 2.15	85.14 ± 2.22	85.11 ± 2.15
CFA + PET + Genetics	90.11 ± 1.50	83.21 ± 2.21	86.08 ± 1.04	85.11 ± 2.00	86.08 ± 1.05
CFA + PET + MH	86.09 ± 3.45	84.17 ± 4.20	84.08 ± 3.71	85.09 ± 4.01	84.08 ± 3.72
<b>CFA + PET + Genetics + MRI*</b>	<b>88.07 ± 0.70</b>	<b>86.13 ± 1.30</b>	<b>87.08 ± 0.80</b>	<b>87.09 ± 0.90</b>	<b>87.08 ± 0.80</b>
CFA + PET + Genetics + MH	86.09 ± 3.51	86.13 ± 4.70	86.08 ± 3.32	86.08 ± 3.99	86.08 ± 3.36

**Table 4.** Random Forest performance validation for predicting whether MCI subjects will progress to AD or not (*tenfold cross-validation*; Second Layer). BA: Balanced accuracy. Asterisk (\*): is the subset of features with the best predictive performance. Performance: Mean ± standard deviation.

Modalities used	Precision (%)	Recall (%)	Accuracy (%)	F1-score (%)	AUC
All	87.50	87.50	87.76	87.75	0.953
CFA	91.30	87.50	89.80	89.81	0.926
MRI	56.67	70.83	59.18	59.66	0.691
PET	72.73	66.67	71.43	71.44	0.812
Genetics	76.00	79.17	77.55	77.58	0.787
MH	57.14	50.00	57.14	57.07	0.562
CFA + MRI	90.91	83.33	87.76	87.86	0.903
CFA + PET	95.45	87.50	91.84	91.95	0.955
CFA + genetics	84.00	87.50	85.71	85.75	0.926
CFA + MH	91.30	87.50	89.80	89.81	0.918
CFA + PET + MRI	88.00	91.67	89.80	89.83	0.949
CFA + PET + genetics	91.67	91.67	91.85	91.83	0.956
CFA + PET + MH	87.50	87.50	87.76	87.75	0.943
<b>CFA + PET + genetics + MRI*</b>	<b>91.70</b>	<b>91.70</b>	<b>91.86</b>	<b>91.84</b>	<b>0.963</b>
CFA + PET + genetics + MH	87.50	87.50	87.76	87.75	0.948

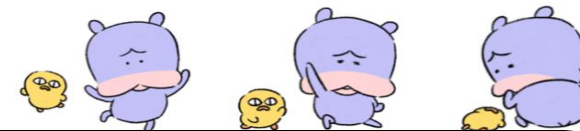
**Table 5.** Random Forest performance measures for AD progression prediction of MCI subjects based on CFA, MRI, PET, genetics, and MH modalities (*MS2 test dataset*; Second Layer). Asterisk (\*): is the subset of features with the best predictive performance.







# Modeling



## Different classifiers experiments

Classifier	Precision (%)			Recall (%)			MCA (%)	MCF (%)
	CN	MCI	AD	CN	MCI	AD		
SVM	100.0	91.43	83.87	96.67	89.13	89.66	91.43	91.74
KNN	65.71	59.57	73.91	76.67	60.87	58.62	64.76	65.89
DT	100.0	86.54	96.30	86.67	97.83	89.66	92.38	92.81
NB	90.32	92.68	84.85	93.33	82.61	96.55	89.52	90.05
<b>RF (our model)*</b>	<b>100.0</b>	<b>86.79</b>	<b>100.0</b>	<b>86.67</b>	<b>100.0</b>	<b>89.66</b>	<b>93.33</b>	<b>93.82</b>

**Table 6.** Comparison of different classifiers (*MS2 test dataset*; First Layer). MCA: multiclass classification accuracy, MCF: multiclass F1 score; Asterisk (\*): is the model with the best predictive performance.

Modalities used	Precision	Recall	Accuracy	F1-score	AUC
SVM	81.82	75.00	79.59	79.64	0.867
KNN	83.33	83.33	83.67	83.66	0.869
DT	85.71	75.00	81.63	81.82	0.850
NB	84.00	84.00	83.67	83.67	0.907
<b>RF (our model)*</b>	<b>87.50</b>	<b>87.50</b>	<b>87.76</b>	<b>87.75</b>	<b>0.953</b>

**Table 7.** Comparison of different classifiers (*MS2 test dataset*; Second Layer). Asterisk (\*): is the model with the best predictive performance.





# Explainability



## RF/DT Explainer

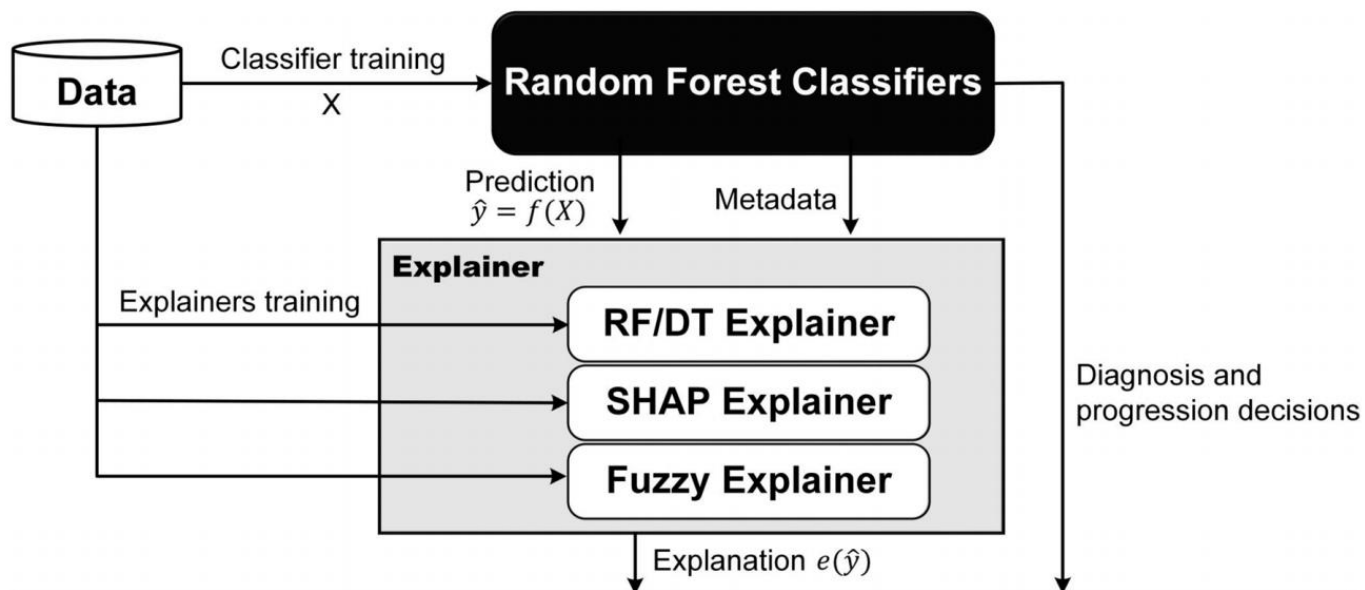
- Global feature importance 즉 전체적인 feature가 얼마나 예측에 영향을 미치는지 알 수 있는 지표

## SHAP Explainer

- 해당 환자가 AD로 분류됨에 있어 feature가 얼마나 영향을 미치는지 나타내는 지표

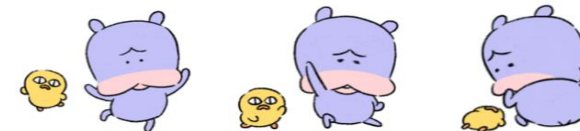
## Fuzzy Explainer

- 해당 feature가 각각 얼마만큼의 영향을 분류됨에 있어 미치는지 나타내는 지표





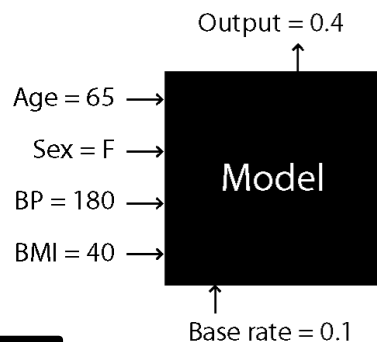
# Explainability



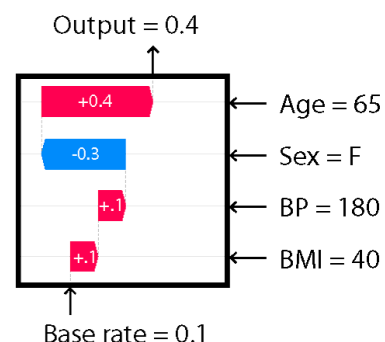
## SHAP VALUE



SHAP



Explanation



SHAP VALUE : 특정 변수가 예측에 미치는 영향을 알 수 있는 지표

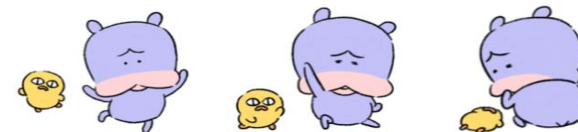
장점 : 변수 간의 의존성을 고려할 수 있고 음의 영향력을 계산할 수 있음  
(Feature importance는 불가능)

또한 각 환자별 예측값에서 특정 변수들의 영향도를 확인 가능함





# Explainability



## SHAP VALUE

- **Shapley Value**
  - \*Additive Feature Attribution 성질을 만족시키는 식 : **Shapley Value**
  - $$\phi_i(f, x) = \sum_{z' \subseteq x'} \frac{|z'|!(M-|z'|-1)!}{M!} [f_x(z') - f_x(z' \setminus i)]$$
  - $\phi_i(f, x)$ : Shapley value
  - Cooperative game theory result에 기반
- **Key Idea (of Shapley value):**
  - 해석을 위한 각 feature importance 필요
  - Feature들을 하나씩 조합
  - 조합한 각 feature에 따른 평균적인 변화 관찰



### ✓ Game Theory

개인 또는 기업이 어떠한 행위를 했을 때, 그 결과가 게임에서와 같이 자신뿐만 아니라 다른 참가자의 행동에 의해서도 결정되는 상황에서, 자신의 최대 이익에 부합하는 행동을 추구한다는 수학적 이론

### ✓ 협력적 게임 이론 (Cooperative Game Theory)

비협력적으로 게임을 했을 때 각 개인이 취하는 이득보다, 협력적으로 게임을 했을 때의 각 개인이 취하는 이득이 더 크다면, 긍정적인 협동이 가장 최선의 선택지

대표적인 예시 : 죄수의 딜레마

그림 01 게임이론 사례 1 - 죄수의 딜레마

[A와 B가 경찰에게 받은 제안]

		[은행강도 B]	
		자백+공범지목	무죄주장
[은행강도 A]	자백+공범지목	10년형 / 10년형	25년형 / 1년형
	무죄주장	1년형 / 25년형	3년형 / 3년형





# Explainability



문제: ('XAI, 인공지능을 해부하다'에서 발췌)

'당신은 운송 수단 스타트업에서 근무 중인 데이터 과학자다. 당신이 운영하는 서비스는 최근 많은 사람의 호응을 받으며 승승장구 하고 있다. 어느 날 기획자들로부터 합승 제도를 도입하는게 어떠냐는 제안을 받았다.

문제는 합승 요금을 분배할 때 발생했다. 중간에 탑승한 사용자에게 합승 요금을 어떻게 분배해야 하는지 논쟁이 발생한 것이다. 베타 기간에는 동승자들이 알아서 청구 금액을 분배하기로 했다. 그러나 이 방식은 합리적이지 않았고 곧 고객들의 불만이 폭주했다. 이제 당신은 운송 비용을 어떻게 나눠야 더 합리적인지 구분해 달라는 요청을 받았다. 어떻게 하면 사용자들의 합승 요금 기여도를 합리적으로 계산할 수 있을까?'

굉장히 예민





# Explainability



사용자	요금 (원)
A	10,000
B	30,000
C	5,000
A, B	50,000
A, C	4,000
B, C	35,000
A, B, C	100,000

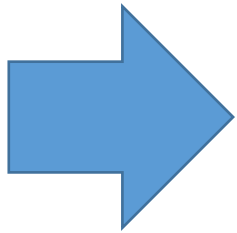
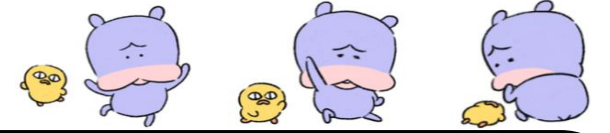


순서	A 기여도	B 기여도	C 기여도
A, B, C	$V(A) = 10,000$	$V(A,B)-V(A)=40,000$	$V(A,B,C)-V(A,B)=50,000$
A, C, B	$V(A) = 10,000$	$V(A,B,C)-V(A,C)=60,000$	$V(A,C)-V(A)=30,000$
B, A, C	$V(A,B)-V(B)=20,000$	$V(B)=30,000$	$V(A,B,C)-V(A,B)=50,000$
B, C, A	$V(A,B,C)-V(B,C)=65,000$	$V(B)=30,000$	$V(B,C)-V(B)=5,000$
C, A, B	$V(A,C)-V(C)=35,000$	$V(A,B,C)-V(A,C)=60,000$	$V(C)=5,000$
C, B, A	$V(A,B,C)-V(B,C)=65,000$	$V(B,C)-V(C)=30,000$	$V(C)=5,000$





# Explainability



사용자	Shapley Equation	Shapley Value
A	$\frac{10+10+20+65+35+65}{6} \times 1,000$	34.170
B	$\frac{40+60+30+30+60+30}{6} \times 1,000$	41.700
C	$\frac{50+30+50+5+5+5}{6} \times 1,000$	24.170

- A 기여도: 34.17%
- B 기여도: 41.17%
- C 기여도: 24.17%

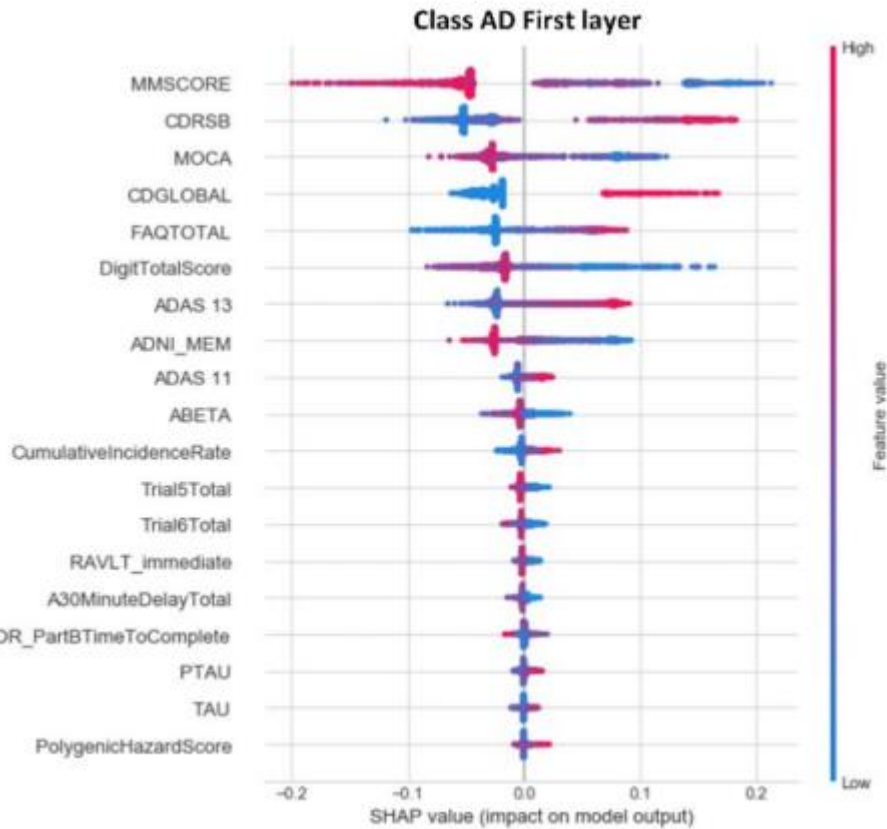
## 결과해석

- B는 C에 비해서 많은 개인 운송 비용(6배) 부담하지만, 합승시에는 상대적으로 적은 금액을 부담하고 있음
- 사용자 A는 B와 비교했을 때 개인 운송 요금이 3배나 차이가 나지만, 합승시 훨씬 많은 요금 부담(서비스 이용에 적극적)





# Explainability



## SHAP summary plots

- 전체 Feature들에 대해 Shapley Value가 어떠한 분포를 미치는지 시각화

Ex) MMSE CDRSM MOCA

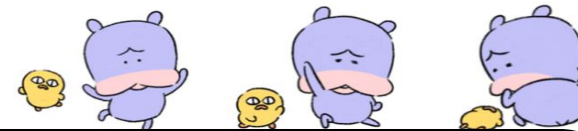
- MMSE는 낮을수록 CDRSM은 높을수록 MOCA는 낮을수록 AD CLASS로 분류될 확률이 높다.







# Explainability



## FURIA(Fuzzy Unordered Rule Induction Algorithm) & DT Explainer

- RIPPER 알고리즘(IF-THEN : 분류 알고리즘)의 변형으로서 결과값(IF)의 구조를 Fuzzy하게 나타냄
- 이를 통해 정량화된 수치가 아니라 그보다 넓은 범위(Fuzzy)의 결과를 해석 가능

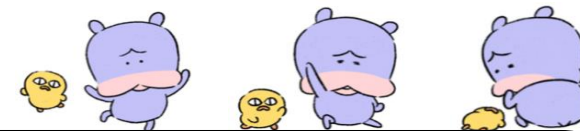
### ExpliClas

- FURIA 알고리즘의 하나로서, 이를 이용하면 (textual + graphical)한 데이터에 대해 언어학적인 용어로서 정의 되어 설명되어진다.ex) low, medium, high

## Fuzzy natural language explanation (local explanation)

```
IF Hippocampus is HIGH and Ventricles and ICV are LOW  
THEN DX is CN  
IF ICV is HIGH, Ventricles is LOW, and Hippocampus is MEDIUM  
THEN DX is MCI  
IF Entorhinal and Hippocampus are LOW  
THEN DX is AD  
IF ABETA is LOW and PolygenicHazardScore and TAU are MEDIUM  
THEN DX is AD
```



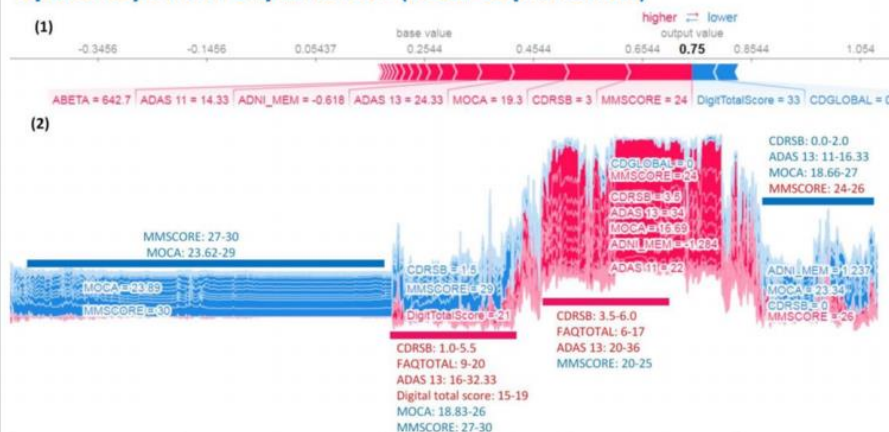


## RF interpretability

### Explanation Interface

## Why These decisions?

### Specific patient explanation (local explanation)



### Fuzzy natural language explanation (local explanation)

```

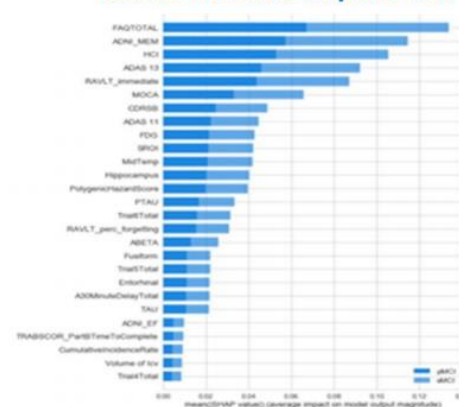
IF Hippocampus is HIGH and Ventricles and ICV are LOW
THEN DX is CN
IF ICV is HIGH, Ventricles is LOW, and Hippocampus is MEDIUM
THEN DX is MCI
IF Entorhinal and Hippocampus are LOW
THEN DX is AD
IF ABETA is LOW and PolygenicHazardScore and TAU are MEDIUM
THEN DX is AD

```

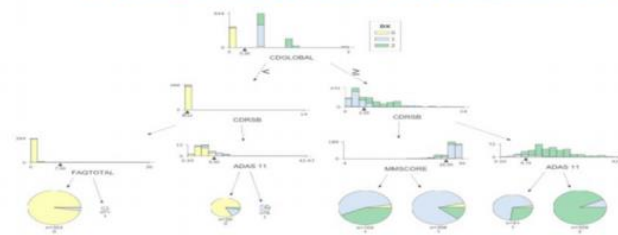
**Diagnosis Decision: CN, MCI, AD**

### Progression Decision: sMCI, pMCI

## Global feature importance



### Visualization: Decision tree explanation





# Conclusion



## Results

- 다중분류를 수행함에 있어 multimodal이 단일 modal보다 뛰어난 성능을 발휘함
- SHAP VALUE와 DT & FURIA, Feattrue importance를 통해 Explanation capability를 달성

## 한계점

- 제안된 ML 모델이 정확하고 설명가능하지만 실제 임상 시나리오에서 적용하기에는 힘들다.

## Discussion

- Layer1, layer2 두가지 모델을 사용함으로써 predict, progression 모두 예측할수 있다고 저자가 주장한다. 또한 각각의 모델에서 사용하는 데이터셋이 다른것도 납득이 가지만, Layer1의 MCI Patient 예측결과를 가져와 Layer 2에서 사용한다는 점이 정말 효율적일까? 의문이 들었습니다.



감사합니다.

질문은 언제든지 환영입니다:3

