

Predviđanje ranog stadija Alzheimerove bolesti

Sara Rivić Carević i Tea Madžarac

A dark blue diagonal gradient bar that starts from the bottom left corner and extends towards the top right corner, covering the lower half of the slide.

MOTIVACIJA

Kavitha C, Mani V, Srividhya SR, Khalaf
OI and Tavera Romero CA (2022)
Early-Stage Alzheimer's Disease
Prediction Using Machine Learning
Models. Front. Public Health 10:853294.
doi: 10.3389/fpubh.2022.853294

- vodeći uzrok demencije među odraslim osobama
- težak za predvidjeti u ranijim stadijima
- tretman u kasnijim stadijima ima puno slabiju učinkovitost nego u ranim
- korištenje strojnog učenja za raniju predikciju

Korišteni skup podataka

Longitudinal MRI Data in Nondemented and Demented Older Adults

oasis_longitudinal.csv

- 150 subjekata u dobi od 60 do 98 godina
- za svaki MRI se bilježilo:
 - Group: Class
 - Visit: Visit Order
 - MR Delay: MR Delay Time (Contrast)
 - M/F: Gender
 - Hand: Dominant hand
 - Age: Subject's age
 - EDUC: Years of education
 - SES: Socioeconomic Status
 - MMSE: Mini Mental State Examination
 - CDR: Clinical Dementia Rating
 - eTIV: Estimated total intracranial volume
 - nWBV: Normalize Whole Brain Volume
 - ASF: Atlas Scaling Factor

OAS2 Study Data															
Subject Information				Visit Data				Demographics				Cognitive & Clinical Measures			
Study ID	Subject ID	MRI ID	Group	Visit	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
0	OAS2_0001	OAS2_0001_MR1	Nondemented	1	0	M	R	87	14	2.0	27.0	0.0	1987	0.696	0.883
	OAS2_0001	OAS2_0001_MR2	Nondemented	2	457	M	R	88	14	2.0	30.0	0.0	2004	0.681	0.876
2	OAS2_0002	OAS2_0002_MR1	Demented	1	0	M	R	75	12	NaN	23.0	0.5	1678	0.736	1.046
3	OAS2_0002	OAS2_0002_MR2	Demented	2	560	M	R	76	12	NaN	28.0	0.5	1738	0.713	1.010
4	OAS2_0002	OAS2_0002_MR3	Demented	3	1895	M	R	80	12	NaN	22.0	0.5	1698	0.701	1.034

Priprema podataka

- uklanjanje monotonih značajki
- pronalazak null vrijednosti
- detekcija stršućih vrijednosti
- uklanjanje „converted” skupine
- transformacija kategoričkih u numeričke
- standardizacija

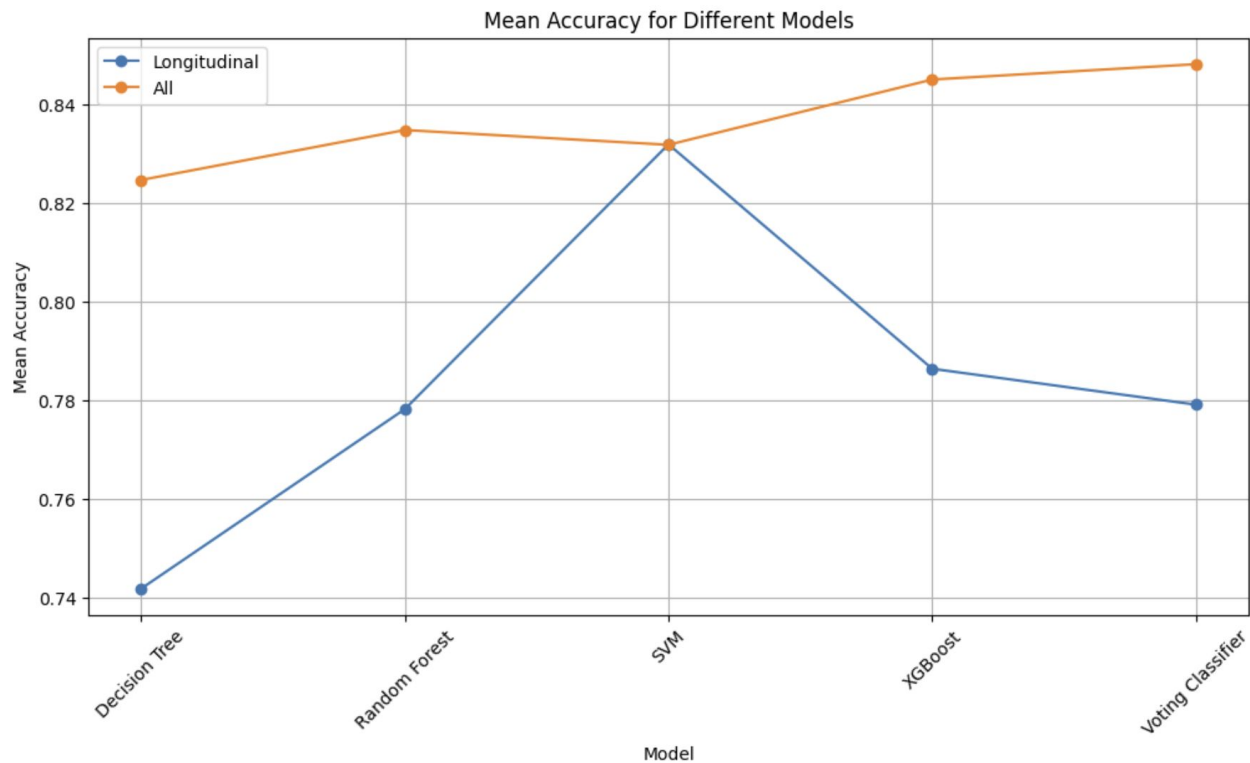
Replikacija rezultata

- nasumična podjela podataka na skupove za treniranje i testiranje u omjeru 80:20
- procjenjujemo različite metrike performansi poput točnosti, preciznosti, odziva i F1 ocjene
- kako bismo odredili najbolje parametre za svaki model, provodimo 10-erostruku unakrsnu validaciju za sljedeće algoritme:
 - Decision Tree
 - SVM
 - Random Forests
 - XGBoost
 - Voting

Poboljšanje rezultata

1. Pokušaj korištenja tri nova modela:
 - Logistic Regression
 - KNN Classification
 - SGDClassifier
2. Proširenje skupa podataka:
 - integracija cross-sectional podataka s prethodno korištenim longitudinalnim skupom

Rezultati



Decision Tree
Mean Accuracy: 0.7418181818181818
Test Accuracy: 0.8571428571428571
Test Precision: 0.8571428571428571
Test Recall: 0.8571428571428571
Test F1 Score: 0.8571428571428571

Random Forest
Mean Accuracy: 0.7781818181818181
Test Accuracy: 0.8571428571428571
Test Precision: 0.8571428571428571
Test Recall: 0.8571428571428571
Test F1 Score: 0.8571428571428571

SVM
Mean Accuracy: 0.8318181818181818
Test Accuracy: 0.8928571428571429
Test Precision: 0.8948717948717949
Test Recall: 0.8928571428571429
Test F1 Score: 0.8927203065134101

XGBoost
Mean Accuracy: 0.7863636363636364
Test Accuracy: 0.8928571428571429
Test Precision: 0.8948717948717949
Test Recall: 0.8928571428571429
Test F1 Score: 0.8927203065134101

Voting Classifier
Mean Accuracy: 0.7790909090909091
Test Accuracy: 0.9285714285714286
Test Precision: 0.9375
Test Recall: 0.9285714285714286
Test F1 Score: 0.9282051282051282

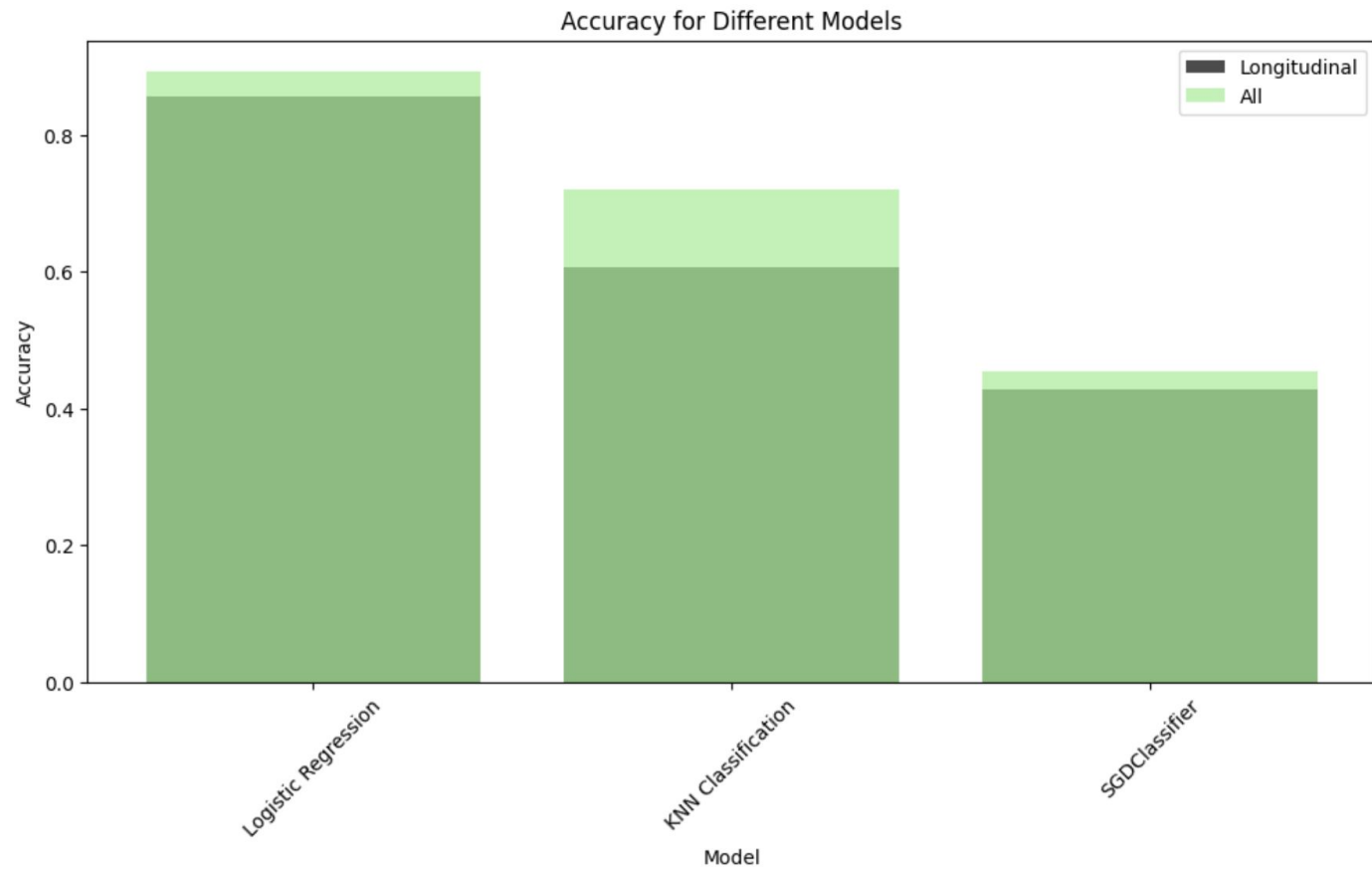
Decision Tree
Mean Accuracy: 0.8245977011494254
Test Accuracy: 0.8533333333333334
Test Precision: 0.853110465116279
Test Recall: 0.8533333333333334
Test F1 Score: 0.8530678733031675

Random Forest
Mean Accuracy: 0.8347126436781609
Test Accuracy: 0.8933333333333333
Test Precision: 0.8966116941529234
Test Recall: 0.8933333333333333
Test F1 Score: 0.8923167155425218

SVM
Mean Accuracy: 0.8317241379310344
Test Accuracy: 0.88
Test Precision: 0.8916483516483518
Test Recall: 0.88
Test F1 Score: 0.8774967405475881

XGBoost
Mean Accuracy: 0.8449425287356321
Test Accuracy: 0.92
Test Precision: 0.9207038123167155
Test Recall: 0.92
Test F1 Score: 0.9196802325581395

Voting Classifier
Mean Accuracy: 0.8480459770114942
Test Accuracy: 0.88
Test Precision: 0.8813333333333333
Test Recall: 0.88
Test F1 Score: 0.8792118226600985



Hvala na pažnji !