

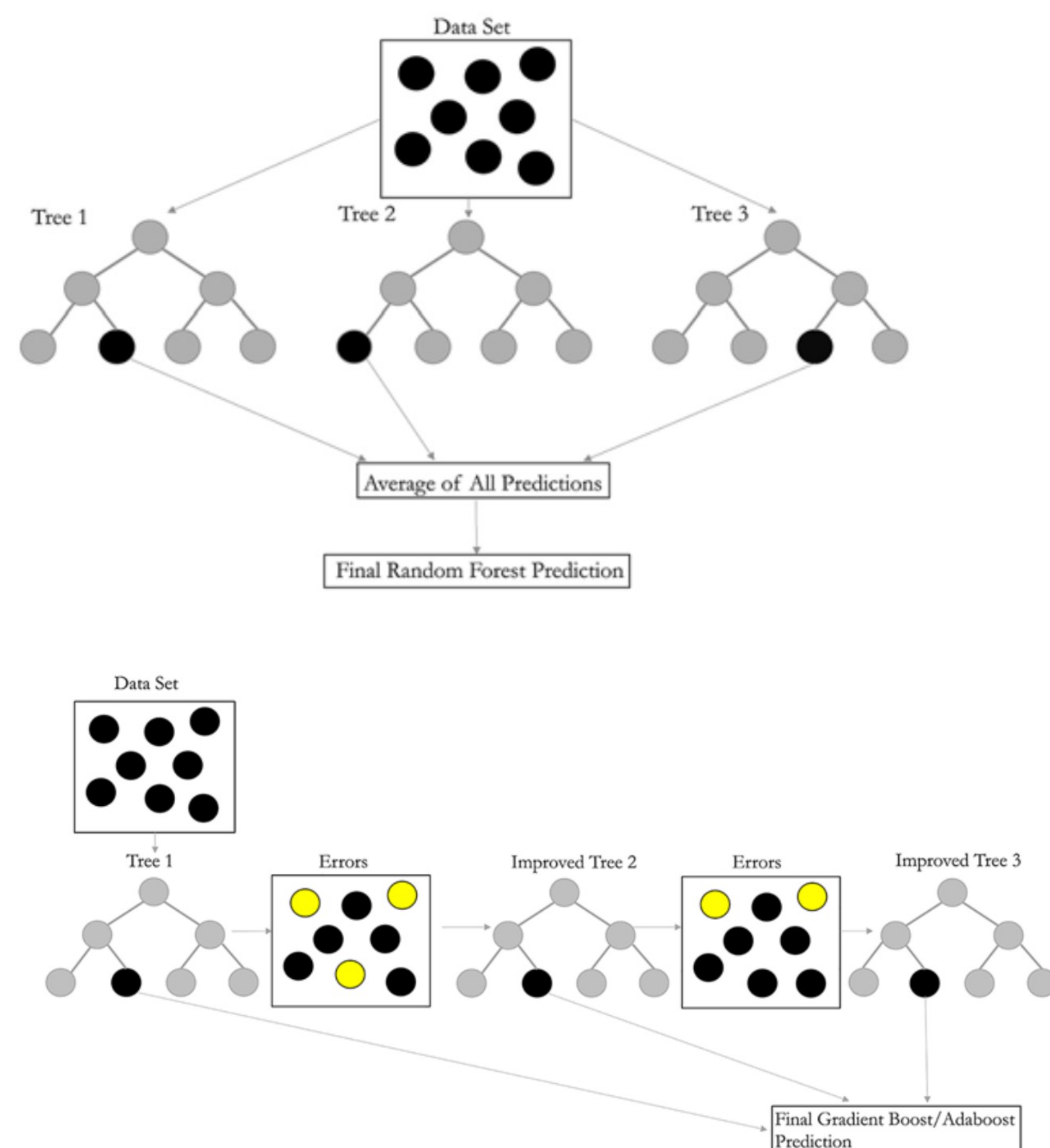
Machine Learning for Predicting Hemorrhage In Pediatric Patients with Brain Arteriovenous Malformation

Satvir Saggi, Ethan Winkler MD PhD, Nalin Gupta MD PhD

Introduction

- Ruptured brain arteriovenous malformations (bAVMs) in a child are associated with substantial morbidity and mortality.
- Prior studies investigating predictors of hemorrhagic presentation of a bAVM during childhood are limited
- Machine learning (ML), which has high predictive accuracy when applied to large data sets, can be a useful adjunct for predicting hemorrhagic presentation.
- The goal of this study was to use ML in conjunction with a traditional regression approach to identify predictors of hemorrhagic presentation in pediatric patients based on a retrospective cohort study design.

Random Forest and Gradient Boost Methodologies



Demographics and Traditional Regression

TABLE 1. Demographics of 186 pediatric patients with bAVMs

Variable	No Hemorrhage	Hemorrhage	p Value
Total pts	70 (38%)	116 (62%)	
Demographic characteristics			
Mean age, yrs (SD)	12.9 (4.7)	11.6 (4.0)	0.06*
Male sex	37 (53%)	65 (56%)	0.67†
Clinical presentation			
Seizure‡	33 (47%)	12 (10%)	<0.0001†
Neurological deficit‡	29 (41%)	41 (35%)	>0.99†
Concurrent arterial aneurysms			
Any	6 (9%)	39 (34%)	<0.0001§
Intranidal	2 (3%)	26 (22%)	<0.001§
Morphological characteristics			
Mean bAVM size, cm (SD)	3.6 (1.5)	2.6 (1.6)	<0.0001*
Deep venous drainage	39 (56%)	76 (66%)	0.18†
Lt laterality	31 (44%)	67 (58%)	0.07†
Eloquence	38 (54%)	66 (57%)	0.73†
Diffuse	28 (40%)	44 (38%)	0.78†
Location			
			0.36§
Cortical	40 (57%)	56 (48%)	
Cortical/subcortical	16 (23%)	24 (21%)	
Subcortical	6 (9%)	20 (17%)	
Posterior fossa	8 (11%)	16 (14%)	
Deep location	6 (9%)	23 (20%)	0.06§
Supratentorial location	53 (76%)	64 (55%)	<0.01†

TABLE 2. Multivariable regression analyses testing risk of ICH at initial presentation

Variable	OR	95% CI	p Value
Age, yrs	0.95	0.87–1.03	0.21
Male sex	1.22	0.60–2.51	0.58
Any concurrent arterial aneurysms	5.27	1.48–23.00	0.02
Intranidal aneurysm	1.61	0.23–14.37	0.64
bAVM size, cm	0.56	0.42–0.74	<0.0001
Deep venous drainage	1.55	0.68–3.59	0.30
Eloquence	1.63	0.72–3.74	0.24
Diffuse	1.05	0.47–2.35	0.91
Lt laterality	2.21	1.07–4.72	0.04
Location			
Cortical vs cortical/subcortical	1.10	0.40–3.03	0.85
Cortical vs posterior fossa	0.27	0.04–1.68	0.18
Cortical vs subcortical	0.38	0.06–2.02	0.26
Deep location	1.12	0.19–6.46	0.90
Supratentorial bAVM location	0.26	0.04–1.20	0.10

Comparison of Algorithms

TABLE 3. Comparison of ML algorithms based on most important factors

ML Algorithm	Most Important Factors*
Random forest	bAVM size, presenting age, aneurysm
GBDT	Presenting age, bAVM size, location
AdaBoost	Presenting age, bAVM size, location

* The three factors with the highest importance are listed, ranked from highest to lowest.

Conclusions

- By using an ML approach, we were able to find predictors of hemorrhagic presentation that were not identified using a conventional regression approach.
- Results from the ML analyses demonstrated bAVM size, presenting age, the presence of a concurrent arterial aneurysm, and bAVM location to be important features for predicting pediatric hemorrhagic presentation.

Machine Learning Results

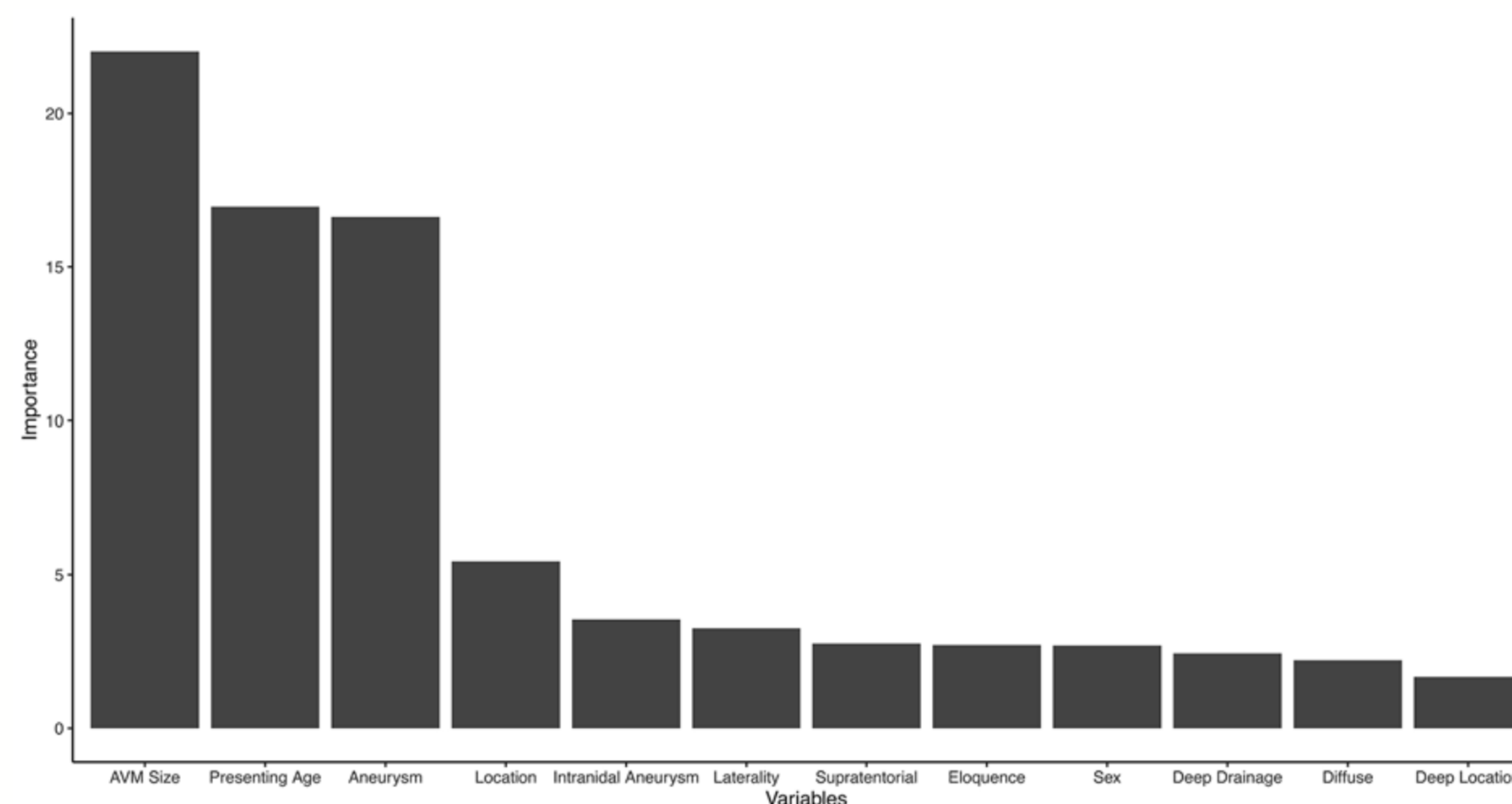


FIG. 3. Absolute values for the importance of variables in regard to predicting hemorrhagic presentation based on the random forest model. Importance of variables is ranked from left to right.

Future Directions

- By using an ML approach with larger, multiinstitutional data sets in the future, determinants of pediatric bAVM rupture can continue to be elucidated to improve care of children with bAVMs.
- We hope to implement these machine learning algorithms to create a calculator that can predict hemorrhage using pre-operative variables.

UCSF

University of California
San Francisco