Predicting Brain Age Using Machine Learning Algorithms

Madhu Yadav Kasa, Lokesh Avala, , Jayakrishna Dontharaboina, Ashish Saraf
Mrs.Anuradha reddy,(Assistant professor),
Department of CSE,
MALLA REDDY INSTITUTE OF TECHNOLOGY AND SCIENCE, Telangana, Hyderabad.

Abstract:

Predicting an individual's age based on brain morphological characteristics may aid in the diagnosis of aberrant aging since brain structural morphology changes along the aging trajectory. A person's brain health may be measured by how much their brain age deviates from a typical brain aging trajectory, which is based on neuroimaging. With the wide range of machine learning algorithms available, accurate brain age prediction is becoming more complex, but not impossible, with the help of machine learning methodologies. In this study, we sought to evaluate how well several machine learning models estimated brain age from structural MRI images of the brain by comparing their morphological metrics. The Human Connective Project (HCP) included 1113 participants ranging in age from 22 to 37; Cam-CAN included 601 participants ranging in age from 18 to 88; and the Information extraction from Images (IXI) included 567 participants ranging in age from 19 to 86. We assessed 27 machine learning models that had been applied to these three separate datasets. Each sample's performance was evaluated with the use of cross-validation and an unseen test set.

1. Introduction

The use of neuroimaging to determine brain age has several potential uses as a biomarker for monitoring the progression of neurological diseases and the impact of becoming older [1]. The current gold standard for determining a person's biological age (or "brain age") is to use a machine learning methodology to extrapolate their chronological age from MRI data. The disparity between a person's reported chronological age and their anticipated brain age is known as the brain-expected age difference (brainPAD) [2,3], brain age gap [4,5] or brain age delta [6]. Data deviation from anticipated age trajectories is a standard metric for brain health [1]. A positive result on the brainPAD indicates that the patient is suffering from accelerated aging, in which their brain age is higher than their actual age [5]. If the brainPAD is negative, it means that the brain-predicted age is lower, which is also called delayed aging [5]. As a clinical diagnostic tool for neurodegeneration and cognitive impairment, this population-based empirical measure of brainPAD has shown potential [7–10]. Cognitive decline, decreased happiness, and poor general health have all been associated with increasing brain age relative to chronological age [11], in addition to unfavorable characteristics in physical and mental health [8, 12].

The results support the hypothesis that brain-predicted age might be a valuable biomarker for evaluating brain health when considered together. Some machine learning studies have tried to predict brain aging using structural MRI data [2,13-16]. Brain morphological features obtained from structural MRI scans have seen extensive use due to the capacity to study aging-related morphological alterations in a variety of diseases and disorders [17–21]. Many other methods are mentioned in the literature [5,13]. The parameters and selection of machine learning algorithms, sample size and composition, and input characteristics are all part of this. A number of machine learning methods exist for the purpose of predicting brain age.

Support Vector Regression (SVR), Relevance Vector Regression (RVR), or Gaussian Process Regression (GPR) has historically been the only machine learning method utilized for brain age prediction [8,14,15,22,23]. A number of studies [2,14-16] have examined the topic of multiple ML models that were trained on the same dataset. There is a large gap in the accuracy of several machine learning algorithms for predicting brain aging, and no comprehensive study of their relative performance exists.

Our only objective is to compare several ML algorithms for brain age prediction using morphological factors acquired from structural MRI. Information extraction from Images (IXI), the Human Connective Project (HCP), and the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) were used as publicly available samples of healthy individuals for this aim. Three independent datasets were used to guarantee that the results were unaffected by the composition of the samples. We examined the

performance of 27 ML algorithms on a separate hold-out test set using the identical morph metric data for each sample. The approaches that were investigated included both parametric and nonparametric models, as well as linear and nonlinear models, Bayesian models, and models based on kernels and trees. Using machine learning models to predict brain age from brain morphometric data has great benefits in age-related diseases; this study aims to provide a framework for accomplishing precisely that.

2. Materials and Methods

2.1. Datasets

The three datasets that were taken into account were the following: the Human Connectome Project (HCP) S1200 release, which included 1,113,606 females ranging in age from 22 to 37 years [24], the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) with 601,302 females ranging in age from 18 to 88 years [25], and the Information extraction from Images (IXI) with 567,316 females ranging in age from 19 to 86 years (https://brain-development.org, accessed on 1 September 2020). To make sure no one had a history of serious mental, neurological, or physical issues, we screened them using local study procedures. Using conventional T1-weighted MRI sequences, 1.5T or 3T scanners were used to get T1-weighted MRI images. For HCP [26], Cam-CAN [25], and IXI (https://brain-development.org, accessed on 1 September 2020), the acquisition process and pipelines are detailed elsewhere. Deidentified data was culled from publically accessible databases. Participation in an exchange of research data was subject to local ethical clearances and informed consent processes.

2.2. Processing Images and Extracting Features

According to what was already said, structural T1-weighted pictures were processed [2]. Brain morph metric measures were extracted from all three datasets using the identical preprocessing pipeline in Free Surfer 6.0 (http://surfer.nmr.mgh.harvard.edu, accessed on 1 September 2020). In a nutshell, the following procedures were used to rebuild the cortical surface from each subject's T1-weighted image: dissecting the skull into its component parts, identifying and separating the cerebral hemispheres and sub cortical structures, and finally, creating a seamless model of the border between the cerebral hemispheres and the pial surface.
Other sources [27,28] provided further technical information on the pipeline.

2.3. Algorithms for Machine Learning

The open-source, low-code machine learning tool known as PyCaret was used to automate the machine learning process and perform the brain age prediction [31]. We opted for the PyCaret library since it streamlines the training process for machine learning models with a minimum of code and provides a pipeline with all the essential blocks of functions and modules. As shown below, we used twenty-seven machine learning algorithms.

2.3.1. Algorithms for Parametric Data

By minimizing the residual sum of squares between the observed value and the value predicted by the ordinary least squares regression model, a linear regression (LR) strategy is used to construct a linear model . Assuming that the total absolute value of the coefficients is less than a constant, the Least Absolute Shrinkage and Selection Operator (Lasso) Regression linear method minimizes the residual sum of squares. It is common for this procedure to generate zero-value coefficients. When dealing with data that is multi-collinear, one model tuning strategy is Ridge Regression . Here, L2-norm regularization is used. Whenever multi-co linearity is a problem, the variance is substantial and the least squares are fair. Reduced model complexity and multi-co linearity are two benefits of this algorithm's shrinking coefficients. Relevance Vector Regression (RVR) is a Bayesian framework for learning sparse regression models that is part of the 2.3.2. Nonparametric Algorithms Linear Models area. While RVR and SVR have identical functional forms, the Bayesian formulation of RVR avoids using the free parameters that SVR collects. The sparsity of the RVR is produced in a Bayesian setting by the hyper priors on the model parameters, in line with the maximum a posteriori (MAP) principle. The RVR's behavior is controlled by the sort of kernel that is chosen; the learning approach automatically calculates all other parameters.
Theil-Sen Regression is a nonparametric method that calculates the slope of the regression line by averaging all potential line slopes over the data points. With the use of a multidimensional extension of the median, it provides an alternative to least squares for basic linear regression that is robust against multivariate outliers.

Methodology for Estimating Brain Decades (2.4)

We used 27 distinct ML algorithms using the identical procedures in all three samples (HCP, Cam-CAN, and IXI). Using a conditionally random approach, each sample was divided into a training set (80%) and a test set (20%) to verify that the age and sex distributions were statistically identical. You can find more information on the samples and demographics for all three of them in Supplementary

Table S3. Before building the model, we standardized all morphological metrics, so the data had a mean of 0 and a standard deviation of 1. We utilized hyper parameter tuning with 10-fold cross-validation to train the parameters and validate the approach. When training each algorithm, grid search was used to find the most accurate parameters. For each method, we looked at its mean absolute error (MAE) and the Pearson correlation coefficient (r) between predicted brain age and chronological age [6]. We further included weighted MAE for the purpose of comparing studies with different sample ages. Dividing the MAE by the age range of the hold-out test group yielded the weighted MAE value [13]. Finally, we tracked the time it took to train the model using 10-fold cross-validation on the training data to quantify each algorithm's computational efficiency. An AMD Ryzen 9 5900X CPU with 32 GB of RAM was used to train all of our models in Python.

Part 2.5: Suppressing Ageism

We took the participants' chronological age and subtracted their predicted brain age to get BrainPAD for each approach. Regression analysis's general statistical characteristics lead to an inflated BrainPAD in younger people and an underinflated one in older ones [55]. Addressing age bias was done using a strategy that was suggested by de Lange and colleagues [6]. Y represents the predicted age in the model, with a and b denoting the slope and intercept, respectively, with regard to chronological age (W). This equation was then used to carry out a process of rectification. To adjust the predicted brain age using the a and b coefficients, which generate a linear fit, the formula "corrected predicted brain age" was utilized, which is "predicted brain age" + [W ⌕ (aW + b)]. The method for computing a bias-free brainPAD was "corrected predicted brain age" ~ "chronological age."

Part 2.6: Looking at Different Algorithms

All three algorithms—HCP, Cam-CAN, and IXI—were compared using Pearson's correlation analysis and hierarchical clustering using Ward's minimal variance techniques for Euclidean distances, with the aim of gauging how well they predicted brain ages within their respective samples [2]. An analysis of variance (ANOVA) and post hoc analyses using Tukey's honestly significant difference were used to conduct statistical comparisons of the algorithms at the 5% level of significance.

The Significance of Features (2.7)

In order to identify which morphological features contribute to brain age prediction, we chose three very accurate model types: a linear model, a nonlinear model, and an ensemble model. To examine brain-PAD, or regional morphological factors that impact model prediction error, for all three of the leading techniques, we used kernel Shapley additive explanation (SHAP) [56]. To determine which of the three models were most important, we calculated SHAP values for all samples. When all features' SHAP values are added together, the total is equal to the predicted output less the expected model output, which is the deviation from the whole training data. Based on the 10 nearest neighbors in the training sample, we defined the baseline set and then computed the age-specific feature significance values for each test subject [57]. This led to the creation of a "model error explanation" matrix of size subject _ feature. Each column in this matrix represents the relative importance of a geographical factor to that individual's brainPAD as compared to training samples that were matched for age. In a similar vein, the significance of each attribute for a certain person is shown in each row. The sum of all SHAP values across characteristics is known as an individual's brainPAD, which stands for model prediction error.

3. Results

3.1. The Accuracy of Algorithms for Brain Age Estimation

Results from the assessments of the 27 algorithms included in the HCP, Cam-CAN, and IXI samples are shown in Tables 1-3. The training sets were used to measure model performance, and the hold-out test sets were used to measure prediction performance. Predictions produced by regression algorithms differed in terms of accuracy. The supplementary materials include Figures S2-S4, which show the associations between chronological age and brain age estimates from 27 different algorithms for the individual samples. Results from the approaches used to train the model utilizing structural information from the HCP individuals are provided in Table 1, together with the hold-out test data (n = 223) and the model data (n = 890).

Algorithm	Model Performance			Prediction Performance		
	r	MAE	Weighted MAE	r	MAE	Weighted MAE
Lasso	0.4921	2.6444	0.1763	0.4258	2.7565	0.1838
Lasso LAR	0.4921	2.6444	0.1763	0.4258	2.7565	0.1838
SVR	0.4515	2.6981	0.1799	0.4268	2.7756	0.1850
LAR	0.4723	2.6933	0.1796	0.4124	2.7896	0.1860
Elastic Net	0.4714	2.6737	0.1782	0.4199	2.7919	0.1861
Bayesian Ridge	0.4712	2.6745	0.1783	0.4182	2.7927	0.1862
Ridge	0.4698	2.6797	0.1786	0.4255	2.7941	0.1863
ARD	0.4973	2.6373	0.1758	0.3991	2.8251	0.1883
Random Forest	0.4245	2.7785	0.1852	0.4131	2.8304	0.1887
PAR	0.4563	2.7251	0.1815	0.4010	2.8322	0.1888
CatBoost	0.4282	2.7631	0.1842	0.4069	2.8328	0.1889
RVR	0.4498	2.7148	0.1810	0.4021	2.8371	0.1891
LightGBM	0.4273	2.7457	0.1830	0.4016	2.8418	0.1895
GBM	0.4458	2.7149	0.1810	0.4000	2.8437	0.1896
kNN	0.3768	2.8367	0.1891	0.3801	2.8591	0.1906
AdaBoost	0.3982	2.8003	0.1867	0.4188	2.8595	0.1906
Extra Trees	0.4224	2.7738	0.1849	0.4197	2.8674	0.1912
XGBoost	0.4201	2.7726	0.1848	0.3859	2.8771	0.1918
Kernel Ridge	0.4417	2.7495	0.1833	0.3878	2.8775	0.1918
GPR	0.4735	2.7199	0.1813	0.3689	2.9420	0.1961
MLP	0.4744	2.7216	0.1814	0.3675	2.9450	0.1963
OMP	0.4790	2.6927	0.1795	0.3590	2.9457	0.1964
LR	0.4736	2.7244	0.1816	0.3679	2.9474	0.1965
Huber	0.4705	2.7366	0.1824	0.3674	2.9484	0.1966
Theil-Sen	0.4663	2.7544	0.1836	0.3398	2.9724	0.1982
RANSAC	0.4553	2.8094	0.1873	0.3627	3.0015	0.2001
Decision Tree	0.1694	3.0653	0.2044	0.1122	3.1206	0.2080

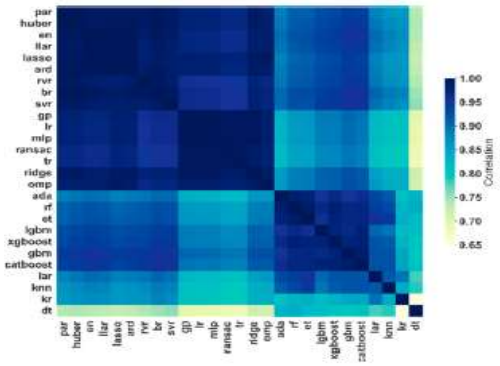
Table 2 shows the results of the algorithms that were used to predict outcomes in the hold-out test data (n = 101) and to train the model (n = 500) using structural characteristics from the Cam-CAN people that were input into the model.

Algorithm	Model Performance			Prediction Performance		
	r	MAE	Weighted MAE	r	MAE	Weighted MAE
Lasso LAR	0.8952	6.6767	0.0954	0.8589	7.0830	0.1012
ARD	0.8992	6.5372	0.0934	0.8585	7.1040	0.1015
Lasso	0.8943	6.6898	0.0956	0.8567	7.1757	0.1025
Elastic Net	0.8960	6.6632	0.0952	0.8548	7.1816	0.1026
Huber	0.8938	6.7060	0.0958	0.8455	7.4663	0.1067
Bayesian Ridge	0.8927	6.7091	0.0967	0.8445	7.4698	0.1067
RVR	0.8824	6.9355	0.0991	0.8378	7.5311	0.1076
PAR	0.8877	6.9834	0.0998	0.8395	7.5762	0.1082
Ridge	0.8906	6.8230	0.0975	0.8432	7.5865	0.1084
OMP	0.8827	7.0357	0.1005	0.8437	7.6179	0.1088
GPR	0.8839	7.0175	0.1003	0.8377	7.7190	0.1103
LR	0.8826	7.0582	0.1008	0.8366	7.7432	0.1106
MLP	0.8831	7.0570	0.1008	0.8364	7.7450	0.1106
SVR	0.8887	6.8523	0.0979	0.8309	7.7551	0.1108
RANSAC	0.8789	7.2202	0.1031	0.8282	7.8652	0.1124
Theil-Sen	0.8791	7.1771	0.1025	0.8366	7.8698	0.1124
GBM	0.8681	7.3435	0.1049	0.8368	7.9222	0.1132
CatBoost	0.8667	7.3767	0.1054	0.8230	8.1285	0.1161
XGBoost	0.8552	7.5686	0.1081	0.8167	8.3920	0.1199
LightGBM	0.8646	7.1822	0.1026	0.8040	8.4686	0.1210
Kernel Ridge	0.876	7.2091	0.1030	0.7022	8.6938	0.1242
Extra Trees	0.8565	7.7880	0.1111	0.8050	8.8377	0.1263
Random Forest	0.8410	8.0043	0.1143	0.7955	8.9883	0.1284
AdaBoost	0.8405	8.0458	0.1149	0.7725	9.4055	0.1344
LAR	0.8378	8.3740	0.1196	0.7577	9.5307	0.1362
kNN	0.8234	8.7403	0.1249	0.7709	9.6734	0.1382
Decision Tree	0.7259	9.7473	0.1392	0.6430	10.5017	0.1500

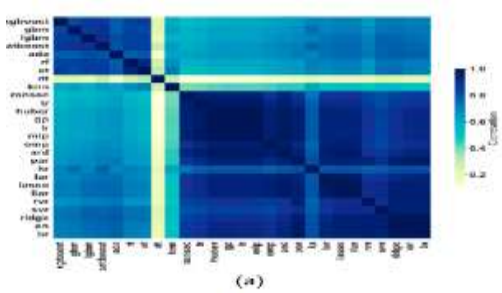
3.2. Evaluation of Brain Age Prediction Algorithms

Pair wise correlations in estimated brain ages varied greatly throughout HCP methods, ranging from 0.1 to 0.97 (Table 1; Figure 1a). Figure 1b displays the outcomes of the hierarchical clustering of the anticipated ages of the brain. The third cluster consists of seven linear models: RANSAC, Theil-Sen, Huber, Linear Regression, OMP, ARD, and PAR; the second cluster contains ensemble models and kNN; while the third cluster has GPR and MLP. Along with eight linear models—LAR, Lasso, Lasso LAR, RVR, SVR, Ridge, Elastic Net, and Bayesian Ridge—a third grouping has kernel ridge regression. The Cam-CAN discovered pair wise correlations in estimated brain ages among approaches ranging from 0.64 to 0.99, as shown in Table 2 and Figure 1c. Hierarchical clustering of the predicted brain ages of the individuals created two groups, whereas the nine linear models (PAR, Huber, Elastic Net, Lasso LAR,

Lasso, ARD, RVR, Bayesian Ridge, and SVR) were grouped together in one (Figure 1d). Five linear models—LR, RANSAC, Theil-Sen, Ridge, and OMP—along with GPR and MLP formed another cluster. Through the use of ensemble models, KNN, and LAR, a third cluster was generated. Decision tree stood up as the strategy that was least similar to the others in a fourth cluster that also included kernel ridge regressions. In the IXI, the algorithms' predicted brain ages had pair wise correlations ranging from 0.66 to 0.99, as shown in Table 3 and Figure 1e. Figure 1f shows that hierarchical clustering analysis revealed that ensemble models, kNN, OMP, kernel ridge, and MLP all contributed to the formation of a common cluster. The second set of linear models included Elastic Net, LAR, RVR, SVR, Ridge, and Bayesian Ridge, among ten less. In a third group of linear models, GPR was included alongside RANSAC, Theil-Sen, Huber, and LR. Decision tree regression was the method that differed the most from the others, according to results from both the HCP and Cam-CAN samples.

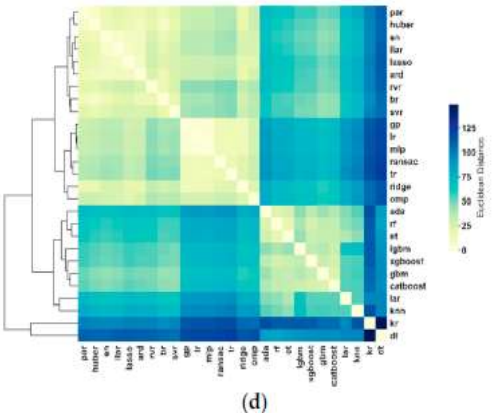


(c)

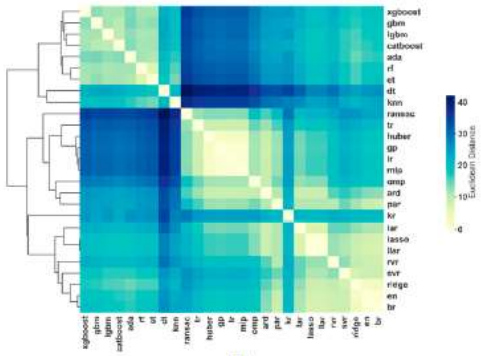


(a)

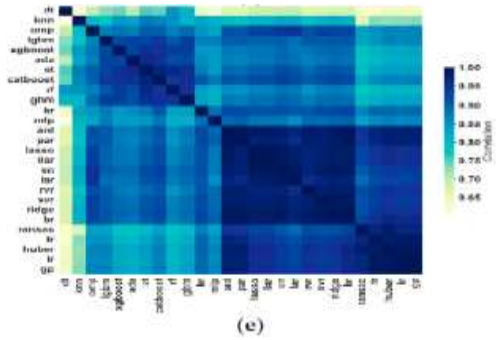
Figure 1. *Cont.*



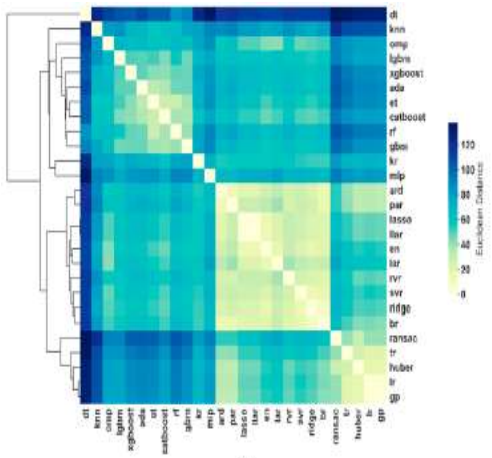
(d)



(b)



(e)



(f)

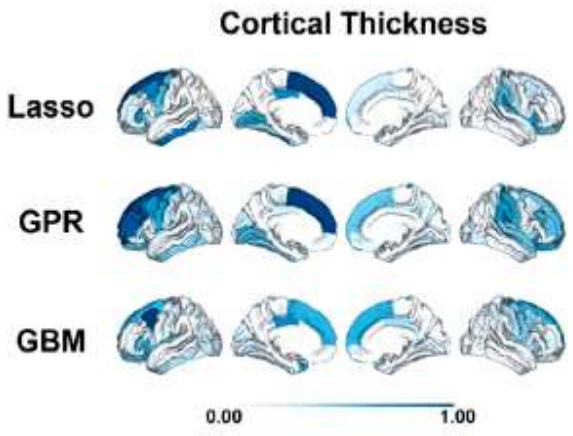
the results of 27 different algorithms. Here are the findings for the HCP sample: (a) a similarity matrix showing the correlations between the 27 methods for predicting brain age, and (b) a distance matrix and dendrogram developed from the hierarchical clustering of these data. Here are the findings for the Cam-CAN sample: (c) a similarity matrix showing the correlations between the 27 methods for predicting brain age, and (d) a distance matrix and dendrogram made from the hierarchical clustering of those results. Here are the findings for the XI sample: (e) a similarity matrix showing the relationships between the 27 methods for predicting brain age, and (f) a distance matrix and dendrogram made from the hierarchical clustering of these data. A lasso is an operator that minimizes absolute shrinkage and selection, whereas a llar is an operator that minimizes the angle between lassos.

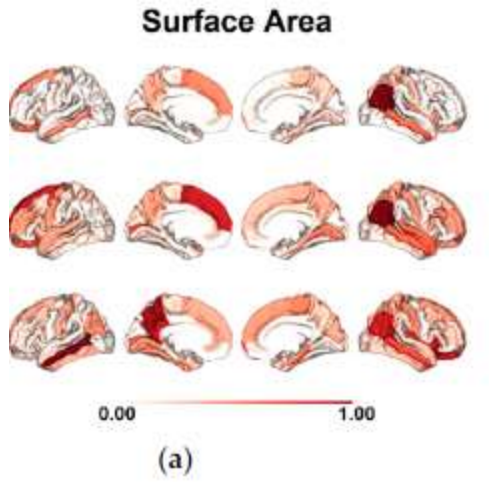
Picture 1. The hold-out test sets for the HCP, Cam-CAN, and IXI samples were predicted to have a similar age distribution across 27 different methods. Here are the findings for the HCP sample: (a) a similarity matrix showing the correlations between the 27 methods for predicting brain age, and (b) a distance matrix and dendrogram made from the hierarchical clustering of these data. Using the data from the Cam-CAN sample, we can see (c) the similarity matrix that shows the correlations between the 27 methods in terms of the anticipated brain ages of individuals, and (d) the distance matrix and dendrogram that come from the hierarchical clustering of those ages. Here are the findings for the XI sample: (e) a similarity matrix showing the correlations between the 27 methods for predicting individual brain ages, and (f) a distance matrix and dendrogram made from the hierarchical clustering of those results. The following acronyms stand for various statistical methods: lasso, llar, svr, lar, ridge, ard, par, cat, rvr, and automatic relevance determination. Light Gradient Boosting Machine (lgbm), Gradient Boosting Machine (gbm), K-Nearest Neighbors (knn), Adaptive Boosting Regression (ada), Extra Trees Regression (et), Extreme Gradient Boosting (xgb), Kernel Ridge Regression (kr), Gaussian Processes Regression (gp), Multi-layer Perceptron Regression (mlp), Orthogonal Matching Pursuit (omp), Linear Regression (lr), Huber Regression (huber), Theil-Sen Regression (tr), Random Sample Consensus (ransac), and Decision Tree Regression (dt).

3.3. Factors Influencing Brain Age Prediction by Region

To determine the relative importance of regional characteristics in brainPAD (brain age prediction error), we computed SHAP values. Using subject-averaged mean absolute SHAP values, Figure 3 displays the relative significance of each model's regional features to brainPAD.

On average, the following HCP features contributed the most to brainPAD across all models: total intracranial volume, cortical thickness of the left superior frontal gyrus and left caudal middle fontal gyrus, surface area of the right inferior parietal lobule, and sub cortical regions of the left pallidum and right putamen. Characteristics such as total intracranial volume, cortical thickness of areas in the left superior frontal gyrus, the left precuneus, and the left supramarginal gyrus, surface area of areas in the left superior fontal gyrus and the right precentral gyrus, and sub cortical regions in the left thalamus and the left amygdale were the most important in the Cam-CAN regarding brainPAD for all models.





section 3. We may evaluate the significance of the SHAP feature by calculating the mean absolute SHAP value for the HCP, Cam-CAN, and IXI samples. The SHAP value is determined by a combination of the GRADIT Boosting Machine, LASSO Regression, and GPR methods. It is the average absolute feature significance for subcortical volume, surface area, and regional cortical thickness across all individuals. When talking about model prediction error or brainPAD, features with darker colors are seen as more important.

4.DISCUSSION

To determine an approximate brain age, this research used 27 distinct machine learning algorithms trained on morphological characteristics of the brain. We tested several machine learning algorithms on three separate datasets to see how well they performed. We showed that there was a large discrepancy in the anticipated brain age when using various ML algorithms on the same brain morphological data. Using 27 different regression techniques and three different datasets, the same result was obtained.

Our earlier research shown that algorithm selection significantly affected morphological feature-based brain age estimation [2]. Our previous work was supplemented in this research by assessing 27 ML algorithms and demonstrating their computational efficiency. We also confirmed our earlier findings in the Cam-CAN and IXI datasets, which included older individuals, in addition to the young adult HCP participants. Even when applied to the identical morphological data, we discovered that algorithm choice produced different brain age estimations. The models obtained a correlation value of 0.42 and an MAE ranging from 2.75 to 3.12 in the HCP. The models obtained a correlation coefficient ranging from 0.64 to 0.85 and an MAE ranging from 7.08 to 10.50 in the Cam-CAN dataset. The models obtained an MAE ranging from 8.04 to 9.86 and a correlation coefficient ranging from 0.63 to 0.79 in the IXI dataset. Regression models with regularization (weighted MAE = 0.10-0.20) performed as well as those with nonlinear and ensemble features (weighted MAE = 0.11-0.20), according to three separate datasets. When compared to other ensemble models, our findings revealed that Lasso LAR, Lasso, and ARD fared the best, but there were only little variations in accuracy. Ensemble models do not always outperform regularized regression models, as shown by our findings. Predicting brain age using a small subset of brain morphological data is possible because regularized algorithms reduce feature weights that aren't relevant to the problem to zero.

Incorporating the regularization (or penalty) term into the models allows one to regulate the complexity of the models; for instance, Elastic Net uses both the L1-norm and the L2-norm, Lasso uses the L2-norm, and Ridge uses the L2-norm. Because of this, the models are more able to withstand the effects of co linearity among the variables used to predict [2]. Ensemble approaches, however, may help with both building models with less variance and higher robustness. Model stability and robustness are achieved when many models are aggregated since the combined results are always less noisy than the individual models. On the other hand, the added complexity that comes with applying ensemble techniques makes it harder to understand the models. Having separate predictors improves the performance of the ensemble models. This meant that regularized models performed similarly to ensemble models.

All samples showed that decision tree algorithms had the worst accuracy. The 27 regression models were tested in three large samples of healthy individuals from the HCP, Cam-CAN, and IXI. The linear regression models (e.g., Lasso, Lasso LAR, RVR,

and SVR) and the ensemble models (e.g., AdaBoost, Cat Boost, GBM, LightGBM, RF, and XGBoost) consistently clustered together, and the results were reproducible in terms of similarity. The decision tree regression approach has the lowest correlation with the other methods for predicting an individual's brain age. The sensitivity of the machine learning models to various sample variables was also assessed. The results showed that the accuracy varied among the age groups represented in the exam. Based on these findings, it seems that the age distribution of a sample is a key factor affecting the generalizability of a model to unknown populations [6].

5. Conclusions

We proved that method selection substantially impacts deduced brain age by running 27 ML algorithms on identical morphological characteristics in the HCP, Cam-CAN, and IXI datasets. In a study including three datasets, we demonstrated that, while preserving accuracy, regularized linear models might potentially be computationally much more efficient than nonlinear and ensemble models when it comes to estimating brain age from morphological traits.

References

1. Cole, J.H.; Franke, K. Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. Trends Neurosci. ***2017****, 40,681–690.*

2. Lee, W.H.; Antoniades, M.; Schnack, H.G.; Kahn, R.S.; Frangou, S. Brain age prediction in schizophrenia: Does the choice of machine learning algorithm matter? Psychiatry Res. Neuroimaging ***2021****, 310, 111270.*

3. Wrigglesworth, J.; Yaacob, N.;Ward, P.;Woods, R.L.; McNeil, J.; Storey, E.; Egan, G.; Murray, A.; Shah, R.C.; Jamadar, S.D.; et al. Brain-Predicted age difference is associated with cognitive processing in later-Life. Neurobiol. Aging ***2022****, 109, 195–203*

4. Anaturk, M.; Kaufmann, T.; Cole, J.H.; Suri, S.; Griffanti, L.; Zsoldos, E.; Filippini, N.; Singh-Manoux, A.; Kivimaki, M.;Westlye, L.T.; et al. Prediction of brain age and cognitive age: Quantifying brain and cognitive maintenance in aging. Hum. Brain Mapp. ***2021****, 42, 1626–1640. [CrossRef]*

5. Baecker, L.; Garcia-Dias, R.; Vieira, S.; Scarpazza, C.; Mechelli, A. Machine learning for brain age prediction: Introduction to methods and clinical applications. EBioMedicine ***2021****, 72, 103600. [CrossRef]*

6. de Lange, A.G.; Anaturk, M.; Rokicki, J.; Han, L.K.M.; Franke, K.; Alnaes, D.; Ebmeier, K.P.; Draganski, B.; Kaufmann, T.; Westlye, L.T.; et al. Mind the gap: Performance metric evaluation in brain-age prediction. Hum. Brain Mapp. ***2022****, 43, 3113–3129.*

7. Gonneaud, J.; Baria, A.T.; Pichet Binette, A.; Gordon, B.A.; Chhatwal, J.P.; Cruchaga, C.; Jucker, M.; Levin, J.; Salloway, S.; Farlow, M.; et al. Accelerated functional brain aging in pre-clinical familial Alzheimer's disease. Nat. Commun. ***2021****, 12, 5346. [CrossRef]*

8. Cole, J.H.; Ritchie, S.J.; Bastin, M.E.; Valdes Hernandez, M.C.; Munoz Maniega, S.; Royle, N.; Corley, J.; Pattie, A.; Harris, S.E.; Zhang, Q.; et al. Brain age predicts mortality. Mol. Psychiatry ***2018****, 23, 1385–1392. [CrossRef]*

9. Smith, S.M.; Elliott, L.T.; Alfaro-Almagro, F.; McCarthy, P.; Nichols, T.E.; Douaud, G.; Miller, K.L. Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. Elife ***2020****, 9, e52677. [CrossRef]*

10. Hogestol, E.A.; Kaufmann, T.; Nygaard, G.O.; Beyer, M.K.; Sowa, P.; Nordvik, J.E.; Kolskar, K.; Richard, G.; Andreassen, O.A.; Harbo, H.F.; et al. Cross-Sectional and Longitudinal MRI Brain Scans Reveal Accelerated Brain Aging in Multiple Sclerosis. Front. Neurol. ***2019****, 10, 450. [CrossRef]*