A Comparative Analysis of Machine Learning Models for Brain Age Prediction Using the OASIS Dataset

Julia Anantchenko
Faculty of Science
Department of Computer Science
Western University
London ON, Canada
janantch@uwo.ca

Josh Bakelaar
Faculty of Science
Department of Computer Science
Western University
London ON, Canada
jbakela@uwo.ca

Johann Cardenas
Faculty of Science
Department of Computer Science
Western University
London ON, Canada
jcarden4@uwo.ca

Jenessa Lu
Faculty of Science
Department of Computer Science
Western University
London ON, Canada
jlu473@uwo.ca

***Abstract* - This project aims to develop and compare machine learning-based models in predicting brain age from magnetic resonance imaging (MRI) data. As the discrepancy between physical age and brain age can indicate neurological conditions or cognitive decline, accurate prediction models are of clinical relevance.**

The goal of this project is to develop and validate machine learning algorithms capable of accurately identifying the biological age of a brain. Our goal can be split up into the following three steps:

1. **Preprocess the OASIS (Open Access Series of Imaging Studies) consisting of labelled and unlabelled MRI images**
2. **Implement and benchmark various machine learning techniques**
3. **Evaluate and compare the performance of these models to identify the most effective approach to accurately predict brain age.**

Our study will conduct a comprehensive comparison of four different ML models: Support Vector Regression (SVR), Residual Neural Network (ResNet), and Least Absolute Shrinkage and Selection Operator (Lasso). These models were selected for their diverse approaches to pattern recognition and learning capabilities, covering both linear and nonlinear techniques.

The project will evaluate each model's performance based on accuracy, precision, and computational efficiency, aiming to identify the most accurate and effective algorithm for predicting brain age. By providing an in-depth comparative analysis, this paper seeks to contribute valuable insights into neuroimaging and the role of machine learning in medical diagnostics.

The evaluation results show that the SVR model with the linear kernel is the best-performing model in terms of accuracy and computational efficiency, which achieves an MAE of 7.11.

Keywords—machine learning, neuroimaging, MRI, SVR, ResNet, Lasso

I. INTRODUCTION

Neuroimaging-based brain age is a key biomarker for assessing the progression of brain diseases and aging [1]. By applying machine learning techniques to medical resonance imaging (MRI) scans, the biological age of the brain can be predicted. Aging has a significant structural impact on the brain that correlates to a decrease in mental and physical health and an increased risk of neurodegenerative diseases such as Parkinson's and Alzheimer's. The discrepancy between an individual's predicted brain age and actual age is known as the brain-predicted age difference (brainPAD) [2]. This measure indicates whether a brain is aging faster or slower than predicted and is an indicator of brain health. Specifically, a positive brainPAD signifies accelerated brain aging, while a negative value indicates delayed aging [3]. As an empirical metric, brainPAD has been validated as an effective indicator of neurodegeneration and cognitive decline in clinical settings, with a higher brain age relative to chronological age linked to reduced cognitive function, well-being, and overall health, including negative physical and mental health outcomes [4].

Recent publications have demonstrated that MRIs can be used to predict chronological age with reasonably good accuracy. The classical way to predict brain age is to extract brain morphological features from MRIs followed by classification or regression analysis. The extraction of brain features allows the morphological age-related brain changes to be examined in a great variety of disorders and conditions. This includes principal components, cortical thickness and surface curvature, volume of gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) [5]. However, using such feature selection results in a loss of information since features are likely not designed explicitly for extracting information relevant to brain age.

Numerous machine learning studies have been conducted to predict brain age with MRIs and they demonstrate a great variability in methods, including the choice of machine learning model, parameters, sample size, and type of input features. This includes Support Vector Regression (SVR), Relevance Vector Regression (RVR), and Gaussian Process Regression (GPR) [6].

In this study, we will focus exclusively and comparatively evaluate Support Vector Regression (SVR), Residual Neural Network (ResNet), and Last Absolute Shrinkage and Selection Operator (Lasso) on their accuracy in predicting brain age. There, we used publicly available brain MRI scans from the OASIS (Open Access Series of Imaging Studies) dataset. This study aims at comparatively evaluating the efficiency and accuracy of non-traditional and traditional machine learning algorithms as well as whether linear or nonlinear, nonparametric models are more appropriate for future implementations of machine learning-based brain predictors.

II. Background & Related Works

In the area of neuroimaging and machine learning, predicting brain age has emerged as an interesting area of research, highlighting the relationship between brain structure, aging, and overall health. Utilizing machine learning algorithms, researchers have aimed to uncover biomarkers indicative of biological brain age, offering insights into neurological disorders and cognitive decline. Each of these three studies employs distinct methods shedding light on the aging process and its implications for brain health.

A. Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors [10]

A study called Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors [10] looks at least absolute shrinkage and selection operator (Lasso)-predicted brain age and factors that cause deviation. It claims that the effect of aging on the human brain includes cognitive decline and a higher risk of neurological diseases. It studies the so-called "biological age" of the brain as it may be a better predictor of overall health. To predict age, LASSO was used with 10-fold cross-validation to select the best lambda. The design also used bootstrapping to reduce bias due to multicollinearity. The study then measured various biomedical, cognitive, and lifestyle factors of participants to conclude that brain age is most accurately measured by T1-weighted and diffusion-MRI phenotypes. It also saw that medical history, smoking, alcohol and poor cognitive performance were associated with an older-looking brain.

B. Brain Age Prediction: A Comparison between Machine Learning Models Using Brain Morphometric Data [6]

This study comprehensively evaluates 27 machine learning models for their efficacy in predicting brain age from structural morphological features derived from MRI scans. This study not only highlights the variability in performance among different models but also emphasizes the crucial impact of model selection on the accuracy of brain age prediction. Furthermore, the study emphasizes the effectiveness of regularized linear regression algorithms compared to complex nonlinear and ensemble models in brain age prediction. Findings suggest future studies can achieve more accurate models without sacrificing computational power. This work contributes to the field of neuroimaging by proving that there exists more efficient and accessible approaches to diagnosing brain health and detecting abnormal aging processes.

C. Brain age prediction using deep learning uncovers associated sequence variants [9]

The study presents a method for predicting brain age using deep learning techniques, specifically using 3D Convolutional Neural Networks (CNNs), using T1-weighted MRI scans [9]. Through the incorporation of multiple types of images and using transfer learning, the study showed a significant increase in prediction accuracy. This research explored the genetic underpinnings of the predicted age difference (PAD), which revealed associations with specific sequence variants through a genome-wide association study (GWAS). Findings suggest that the use of PAD (which indicates the difference between chronological and biological brain age) could serve as a valuable biomarker for studying the aging of the brain and similar neurological conditions [9]. This work contributes to the growing field of neuroimaging by providing insights into the brain aging process and potential genetic influences.

D. Research gap and analysis

While existing studies for predicting brain age have made a significant impact, there remains a notable research gap regarding evaluations of model performances. The research done in this paper brings a new approach to assessing various machine learning models against a benchmark in a low-performance environment. This presents an opportunity to understand the efficiency and accuracy of these models in real-world scenarios where computational resources are limited, which is crucial for practical application and deployment.

Furthermore, the gap in research lies in the generalizability of findings from low-performance environments to higher-scale computational setups. Bridging the gap will provide valuable insights into the adaptability and robustness of brain age prediction models in diverse computing environments, ultimately enhancing their applicability and utility in research and clinical practices.

III. DATA

The problem to be solved by this project is to accurately predict the biological age of a brain, given by an MRI scan. The dataset is from OASIS (Open Access Series of Imaging Studies) and is shared on the OASIS website [16]. The dataset comprises 416 subjects aged 18 to 96. 100 of the included subjects over the age of 60 have been clinically diagnosed with very mild to moderate Alzheimer's disease (AD). As the dataset contains MRI scans of dementia-affected subjects, this affects the model's performance as scans of demented brains may not accurately reflect the age of the subject.

TABLE I. DATA DICTIONARY

Variable	Variable Explanation
Subject ID	a unique identifier for each participant's MRI scan
age	the age of the participant
sex	the sex of the participant

IV. EXPLORATORY DATA ANALYSIS

The OASIS (Open Access Series of Imaging Studies) dataset provides a valuable resource for studying brain morphology using Voxel-Based Morphometry (VBM). It consists of MRI data from a diverse population, allowing for investigations into various factors influencing gray matter density. To manage memory usage effectively, we utilize a subset of subjects for analysis, with preprocessing already conducted using standard VBM pipelines involving tools like SPM8 and NewSegment.

A. Univariate Analysis

In the context of the OASIS (Open Access Series of Imaging Studies), the VBM dataset entails examining the distribution of ages among subjects and analyzing the distribution of sexes to ensure a balanced representation across male and female subjects.

B. Age Distribution

The age distribution is skewed towards the younger and older age groups with notable peaks in frequency at the youngest age bracket and the 70-80 age range. The distribution indicates a bimodal trend, where two age groups are predominantly more represented than the others. Specifically, the highest frequency is in the youngest age bracket, suggesting a large number of subjects are at the lower end of the age spectrum. Following this, there is a notable decline until we reach the next substantial peak at the 70-80 age range, which could indicate the inclusion of a specific age-related demographic or a cohort effect. Other age groups are represented but with lower frequencies, indicating less representation within the dataset as shown in Figure 1.

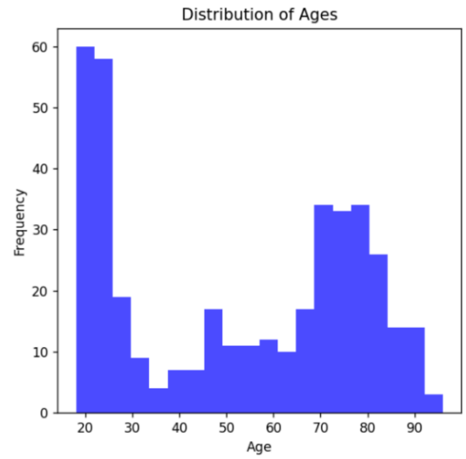

Fig. 1. Distribution of Ages within the OASIS

C. Sex Distribution

The sex distribution (Figure 2) is a straightforward comparison between two categories, Male and Female. The plot shows a significant disparity in the representation of sexes, with the frequency of female subjects being much higher than that of male subjects. This suggests that the dataset might be biased towards female subjects or that the research specifically targeted a female-dominated population.

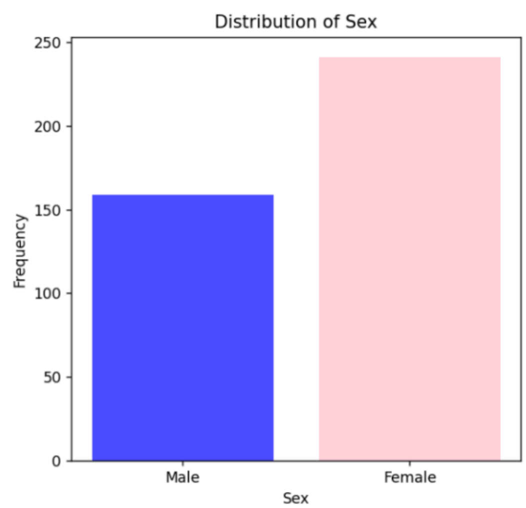

Fig. 2. Distribution of sex within the OASIS dataset

V. METHODS

A. Research Objectives

O1) Predict brain age using different models:
This project aims to build a brain age prediction model using Support Vector Regression (SVR), Residual Neural Network (ResNet), Least Absolute Shrinkage and Selection Operator (Lasso), and

Vision Transformers (ViT). These models were selected for their diverse approaches to pattern recognition and learning capabilities, covering both linear and nonlinear techniques.

O2) *Compare model results using evaluation metrics and the baseline model [8]:*
Utilize Mean Squared Error, R-squared and other evaluation metrics to compare the results of the models.

O3) *Compare model efficiencies:*
Models will be run with limited processing power and memory to test efficiency. Qualitatively analyze different model runtimes and accuracy with limited time to run.

Initially the utilization of the OpenBHB (Open Big Healthy Brains) dataset was used for training and testing. However, due to its extensive size and computational demands, challenges such as efficiently processing and training models within a reasonable timeframe arose. As a result, the OASIS (Open Access Series of Imaging Studies) dataset was chosen, which offered a more manageable size without compromising the quality of the data.

Furthermore, to ensure a fair and equitable evaluation of all models under consideration, a dedicated timeframe of 4 hours to train and evaluate each model was utilized. This approach allowed for equality while comparing the performance of different models, enabling the ability to draw meaningful conclusions regarding their effectiveness and suitability for the task at hand.

B. Research Methodology

1) Core Technical Details

This study was implemented in Python 3.11.0. A variety of libraries were used for the models, data manipulation, and visualization. For each model, the dataset was divided 75/25 train test split.

- Python v. 3.11.0: The programming language used for the implementation.
- Sklearn v. 1.3.2: Used to import models.
- Nilearn v. 0.10.3: Used to import the dataset.
- Numpy v. 1.26.1: Used for data manipulation, including reshaping of data.
- Matplotlib v. 3.8.2: Used for visualization of results.

C. Models

1) Residual Neural Network Regression (ResNet)

ResNet is a deep learning model that effectively handles networks by utilizing residual connection, which helps alleviate the vanishing gradient problem.

The ResNet architecture consists of a series of residual blocks, each containing multiple convolutional layers followed by batch normalization and activation functions [9]. The skip connections throughout each block enable the flow of information throughout the network while mitigating the degradation problem which is associated with training deep neural networks [9]. The typical architecture of ResNet is shown in Figure 3.

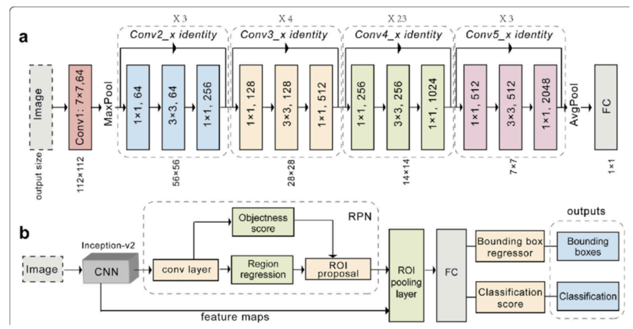

Fig. 3. Typical Architecture of a 101-layer ResNet [13]

The ResNet model was trained using a supervised learning approach, where the input VBM images were paired with their corresponding age labels. The model was trained to minimize the mean absolute error (MAE) loss function, optimizing the disparity between predicted age values and ground truth labels. The model was trained with a dynamic batch size using data generators, enabling efficient processing of the OASIS dataset.

2) Least Absolute Shrinkage and Selection Operator (Lasso)

Lasso (Least Absolute Shrinkage and Selection Operator) regression is a technique used to enhance the accuracy of models. It is typically used with high-dimensional data [11]. Also known as L1 regularization, it adds a penalty term to linear regression which shrinks coefficients. This penalization is denoted by the alpha argument within sklearn's Lasso module.

Lasso aims to minimize MSE. Figure 3 depicts the balance that Lasso performs between total error, variable, and bias. The higher the alpha term, the greater the bias, and the lower the variance. In other words, the higher the alpha value, the larger the penalization. Typically, cross-validation is used to select an alpha term [11]. Through cross-validation, an alpha value of 0.1 was selected out of 0.01, 0.1, 1 and 10.

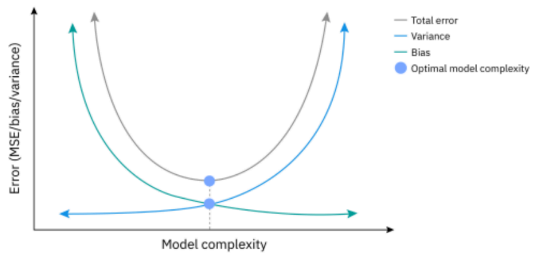

Fig. 3. The goal of Lasso is depicted in a simple graph [11]

3) Support Vector Regression (SVR)

Support Vector Regression (SVR) is one of the more commonly used techniques in neuroimaging and is a variant of SVM (Support Vector Machines). For a training set, where $x_i \in \mathbb{R}$, $y_i \in \mathbb{R}$, SVR aims at finding a regression function that can fit all training samples,

$$fx = w^T \phi x + b \tag{1}$$

where w is a coefficient vector in feature space, Φ(x) is a kernel function to map input x to a vector and b is an intercept. w and b can be solved by solving the optimization problems in literature [14].

While SVM is used for classification, SVR applies the same principles to predict a continuous variable instead of categorizing data into classes. Rather than finding a line of best fit, SVR finds a hyperplane that best fits data points in a continuous space. The function that represents this is chosen from a set of functions defined by a kernel, which can be linear or nonlinear. The parameter epsilon (ε), also known as the margin of tolerance, represents the deviation from the hyperplane that data points are allowed to have. These data points outside the margin are support vectors as they determine the position of the hyperplane [15]. Regularization parameter C is also used to reduce overfitting by balancing the hyperplane complexity (given by the hyperplane's steepness) and the obtained training errors. The structure of support vector regression models is shown in Figure 4 [7].

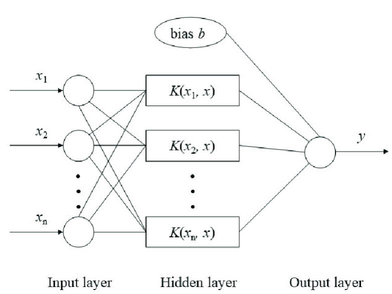

Fig. 4. Support Vector Regression (SVR) Structure

The basic principle of SVR is to map the feature vectors of sample data from low dimension to high dimension and perform regression analysis on them in high dimension by the usage of the kernel function

In this present study, linear kernels and radial basis function kernels are selected to compare the performance between linear and nonlinear methods in building brain age prediction machine learning models.

a. SVR with linear kernel: SVR with a linear kernel was implemented using "SVR" in the scikit-learn package. Grid search was implemented using "GridSearchCV" in the scikit-learn package. The function itself uses k-fold cross-validation to assess the performance of each hyperparameter combination. Grid search was implemented over a search space of 2^{-3},2^{3},2^{7} for C and a search space of 0.01, 1.0, 10 for epsilon. This resulted in C=0.001 and epsilon=0.01. These parameters were then used to train with the model.

b. SVR with radial basis function kernel: SVR with a radial basis function kernel was implemented using "SVR" in the scikit-learn package. All parameters of the radial basis function kernel were used with their default values, for example, an epsilon value of 0.1 and gamma of 1/ (number of features ✕ X.var()).

D. Implementation Strategy

1) Technological Issues and Process Issues

One of the key challenges encountered was related to the size of the datasets, demanding careful consideration in storage management and computational resource allocation. The sheer volume of data generated by Voxel-Based Morphometry (VBM) scans necessitated efficient strategies for dataset storage, retrieval, and processing. Balancing the trade-offs between dataset size and computational feasibility became imperative, requiring innovative approaches to manage and analyze these large datasets effectively.

Moreover, the process issues centred around model training. Training machine learning models on MRI data involved iterative processes that demanded significant computational resources. Furthermore, splitting datasets into manageable batches for learning emerged as a crucial technique to alleviate computational burdens while ensuring continuous model updates and adaptation to new data.

2) Validation

Cross-validation was used to select and test different parameters for the models. It involves sampling data in different ways to reflect the performance of a test dataset.

VI. RESULTS

The difference between the participants' predicted brain age and their chronological age was used to measure the models' predictions at an individual level. This measure is also known as brainPAD. It is calculated as brainPAD = predicted age—chronological age, where a positive brainPAD indicates that the participant's brain age was predicted to be older than their chronological age, while a negative value implies the opposite.

The performance of each model was reported as the mean of the absolute (MAE) values of brainPAD. Additionally, we examined the root mean squared error (RMSE), which is more sensitive to outliers than MAE, the correlation coefficient Pearson's r for chronological age and predicted age, and the prediction R^2. Mean Squared Error (MSE) was utilized as a statistic that underscores the average squared difference between the predicted brain ages and the actual chronological ages. Root Mean Squared Error (RMSE) was also employed as a metric. Its sensitivity to outliers, due to the squaring of prediction errors, makes it a crucial metric for understanding the magnitude of error in predictions. Finally, we used Mean Squared Logarithmic Error (MSLE) as a measure. It mitigates the influence of large outliers in the prediction errors by first applying a logarithmic transformation to the errors before squaring and considers the relative rather than absolute differences between the predicted and actual ages, emphasizing the importance of proportional accuracy across the spectrum of brain ages.

Based on the findings of using ResNet for age prediction, the model has the potential to achieve high accuracy in predicting age. However, this potential comes at the expense of computational efficiency. The training process for ResNet models, especially when dealing with data like 3D MRI images can be significantly time-consuming. This long training process poses a challenge as it may not be feasible for real-time applications or scenarios where computational resources are limited.

Furthermore, the long training time of ResNet models does not allow for a fair comparison with the other models that require less computational resources. When evaluating different models for a given task, it is crucial to ensure that each model is given a comparable amount of training time to fairly assess performances. In situations where training time is a crucial factor, the prolonged training duration of a ResNet model may render it less practical compared to more computationally efficient alternatives. The evaluation of ResNet is shown in Figure 5.

```
Mean Squared Error (MSE): 2684.3461617224557
Root Mean Squared Error (RMSE): 51.81067613651125
Mean Squared Logarithmic Error (MSLE): 5.407200956042537
R-squared (R2) score: -3.3248722872699137
Accuracy within 5 years: 0.00%
```

Fig. 5. Evaluation of the ResNet model

Due to the numbers found when running the model for the same amount of time as the others, the evaluation metrics did not perform to the same calibre as the rest, resulting in the discontinuation of the model.

The evaluation metrics of the models described will be benchmarked against the prediction metrics of another study that uses the same dataset and features the Deep Feature Selection (DFS) model [8] (Figure 6).

Prediction metrics for all independent cohorts.

Cohorts	**Correlation with age**		**MAE (y)**	R^2	**RMSE**
	R	**P-value**			
Before bias correction					
UK Biobank	0.712 (0.007)	**<0.001**	4.19 (0.07)	0.51 (0.03)	5.25 (0.08)
ALFA+	0.448	**<0.001**	4.31	0.20	4.18
ADNI	0.587	**<0.001**	7.21	0.34	5.47
EPAD	0.629	**<0.001**	4.63	0.40	5.62
OASIS	0.733	**<0.001**	6.99	0.54	6.42

Fig. 6. Prediction metrics from Biological brain age prediction using machine learning on structural neuroimaging data: Multi-cohort validation against biomarkers of Alzheimer's disease and neurodegeneration stratified by sex [8]

Furthermore, the differences between the predicted brain and actual brain age were graphed to give a visual understanding of how each model performed individually.

A. Lasso

Figure 7 shows the predicted versus actual brain ages using the Lasso model. True age values were sorted for ease of comparison. There is a general trend of predicted values following the true age line. However, there is noticeable variability throughout.

Figure 8 shows the evaluation metric scores of the Lasso model.

- The MAE using the Lasso model increases from 6.99 to 8.89 compared to the baseline model. This means that Lasso is on average 1.9 times worse than the baseline model.
- The MSE is 129.19. This suggests there is a large variance compared to actual ages, with some

predictions likely deviating more significantly from the actual ages than others.

- The RMSE is 11.37. In comparison, the RMSE of the baseline model is 6.42 which means there is a significant increase from the baseline model. The RMSE being higher than the MAE also indicates the presence of some larger errors in the predictions.
- R^2 score is 0.77 which is higher than the baseline model's 0.54. This means that the model explains 77% of the variation in the response variable. This is a relatively high value, suggesting that the model has a good level of predictive power.

Using these evaluation metrics, it is clear that Lasso does not perform as well in comparison to the baseline model for most metrics. It appears to overfit the data, which is typically penalized by Lasso in comparison to regular linear regression.

In Figure 9, the coefficients plotted on slices of a brain represent the locations of the highest contributing weights in determining brain age. The limited number of dots implies there is only a small section of the brain that has the most say in predicting brain age. This spot looks to be approximately in the middle of the brain. Secondary positions include the upper right part of the brain, however with less influence.

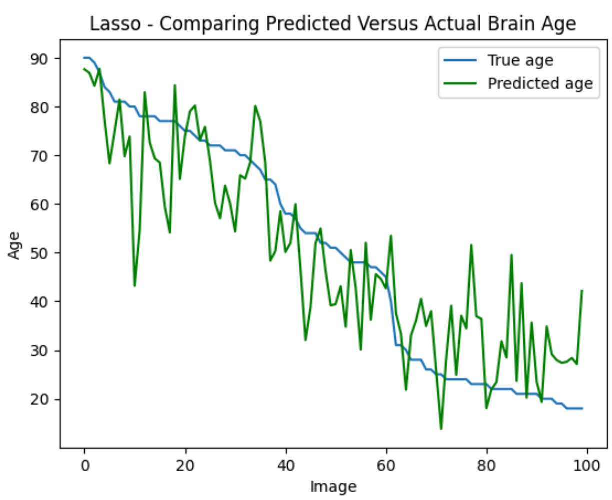

Fig. 7. Predicted and actual brain age values for Lasso

```
Mean absolute error: 8.892644411709286
Mean squared error: 129.19
Root mean squared error: 11.36617789760480
Regression (R^2) score: 0.773080769091533
```

Fig. 8. MAE, MSE, RMSE, R^2 metrics on Lasso

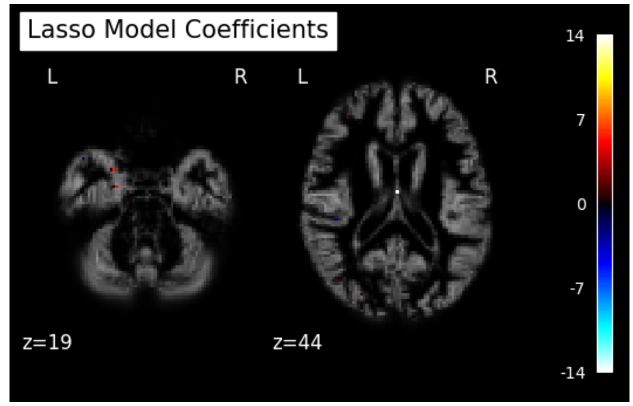

Figure 9: Lasso model coefficients plotted on slices of an MRI. These represent the most impactful areas in brain age prediction.

B. Support Vector Regression (SVR)

1) SVR with linear kernel

Figure 10 shows the MAE, MSE, RMSE, MLSE, and R^2 score according to the prediction result on the test set from the model with the linear kernel SVR, while Figure 11 and Figure 12 are the comparison and residual plot graphs respectively.

- The MAE of the linear SVR model increases from 6.99 to 7.11 compared to the baseline model. 7.11 years suggests that, on average, the model's predictions are about 7.11 years off from the actual brain ages.
- The MSE of the linear SVR model is 76.69 which suggests that there is a variability in the prediction errors, with some predictions potentially being far off from the actual values.
- The RMSE of the linear SVR model increases from 6.42 to 8.86 compared to the baseline model. Similarly to the MAE metric, the RMSE suggests the model is about 8.76 years off from the actual ages.
- The MSLE of the linear SVR model is 0.044 which suggests that on a logarithmic scale, the model's predictions are relatively close to the actual values. This indicates that the model is performing well, especially in datasets where actual ages vary widely
- The R^2 of the linear SVR model increases from 0.54 to 0.86 compared to the baseline model. Around 86.8% of the variance in brain age is explained by the model, which implies the model is effective in capturing the relationship between brain morphometry features and age.

Residual plots should ideally have plots randomly distributed around the horizontal axis. This randomness indicates that the model is correctly capturing the data's patterns without systematic errors. In the case of linear SVR, it appears that the plots are randomly distributed with a small cluster towards the bottom. The model captures the relationships between the brain morphological features and age generally well.

```
Mean Absolute Error: 7.1106426697768
Mean Squared Error: 78.58490891413905
Root Mean Squared Error: 8.864812965547499
Mean Squared Logarithmic Error (MSLE): 0.04480390714963762
Regression (R^2) score: 0.8608516717754079
```

Fig. 10. MAE, MSE, RMSE, MSLE, R^2 metrics on Linear SVR

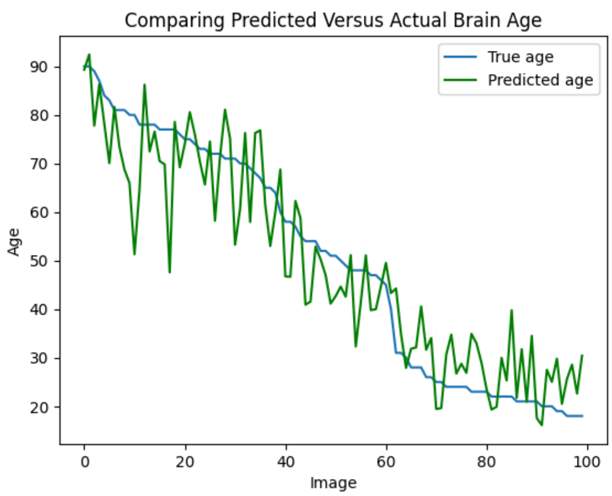

Fig. 11. Predicted and actual brain age values for Linear SVR

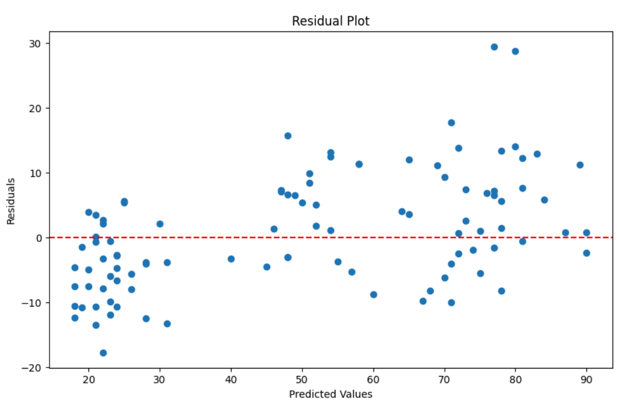

Fig. 12. Residual Plot of Linear SVR's performance

2) SVR with radial basis function kernel (RBF):

Figure 13 shows the MAE, MSE, RMSE, MLSE, R^2 score of the prediction results on the test set on the radial basis function kernel SVR, while Figure 14 and 15 are comparison plots between predicted and actual brain age, and the residual plot respectively.

- The MAE of a RBF SVR model increases from 6.99 to 16.46 compared to the benchmark model. Therefore, the model's age predictions are approximately 16.47 years away from the actual ages. This suggests that the model may not be very accurate in predicting brain age, with a relatively high deviation from the true values.
- The MSE of a RBF SVR model is approximately 384.76 which demonstrates a variance in the prediction errors, suggesting that predictions are significantly off from their actual values.
- The RMSE of a RBF SVR model increases from 6.42 to 19.62 compared to the benchmark model. The model's predictions can be nearly two decades away from the actual age which further emphasizes the model's limited accuracy.
- The MSLE of a RBF SVR model is 0.25. There's still a noticeable discrepancy between predicted and actual ages, although this metric might be less extreme due to the metric's nature of penalizing underestimations.
- The R^2 of the RBF SVR model decreases from 0.54 to 0.31 compared to the benchmark model. Only about 31.9% of the variability in brain age is explained by the model. This indicates a weak predictive power and suggests that the model may not be capturing the complex relationships.

In the predicted versus actual graph, there is an interesting trend regarding predicted age. There seems to be high variability for the first 50 images, then it settles to become less variable for the second half. Since the true age values are sorted from highest to lowest, this implies that the higher the brain age, the more variable the predicted age when using RBF. Furthermore, the trend formed from the predicted age is almost horizontal. This implies that there is a minimal advantage over using a mean value for predictions instead.

In the residual plot, there is a linear arrangement of residuals. This indicates that the relationship between the brain's morphological features and the actual brain age was not captured completely by the model. Therefore, RBF may not be an appropriate model for this dataset or may be due to underfitting where the model is too simple to capture the complex relationships between the data.

The performance of this kernel is quite interesting because previous studies have proven that radial basis function kernels for support vector regression is effective in brain age prediction. Brain age prediction involves understanding complex, nonlinear relationships between the brain morphological features and the actual age of the brain. The RBF kernel is supposedly particularly well-suited for this task due to its flexibility and capability in handling such non-linear relationships [12]. The lack of performance is intriguing and further investigation is recommended for future study.

```
Mean Absolute Error: 16.46847404120417
Mean Squared Error: 384.76164961523347
Root Mean Squared Error: 19.615342199799457
Mean Squared Logarithmic Error (MSLE): 0.25307188554317095
Regression (R^2) score: 0.3187121923448172
```

Fig. 13. MAE, MSE, RMSE, MSLE, R^2 metrics on Radial Basis Function SVR

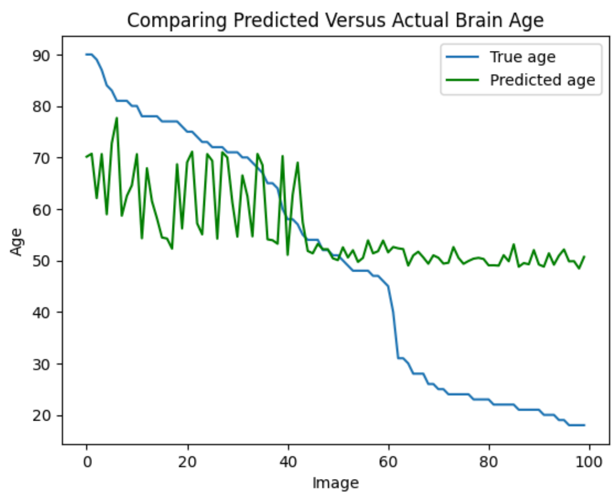

Fig. 14. Predicted and actual brain age values for RBF SVR

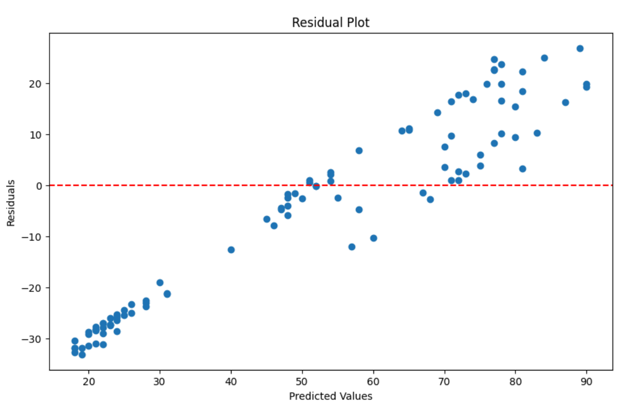

Fig. 15. Residual Plot of RBF SVR's performance

Table 2 below shows the comparisons of the different models for various evaluation methods. The baseline model outperforms all other models in MAE and RMSE. SVR-Linear has the best result for R^2, beating the baseline model by 0.32. For MAE and RMSE, the second best model is SVR - Linear and the third best is Lasso. Overall, the most accurate models seem to be the benchmark, SVR-Linear, Lasso, and SVR-RBF in order.

TABLE II. A Comparison of Evaluation Metrics

	MAE	RMSE	R^2
Benchmark	6.99	6.42	0.54
Lasso	8.89	11.37	0.77
SVR - Linear	7.11	8.86	0.86
SVR - RBF	16.47	19.62	0.31

Table 2: A comparison of evaluation metrics between the benchmark, Lasso, SVR - Linear, and SVR - RBF. The colours represent a gradient from the best value (green) to the worst value (red).

Possible explanations for the ordering may include the fine-tuning of variables and dataset size. With 300 train images, the models trained within this study do not have as much data in comparison to the benchmark. Future work may include working with a larger dataset. With a larger dataset, variables could also undergo more thorough fine-tuning which may improve the results.

C. Novelty and Significance

This study looks at various models compared to a benchmark to assess accuracy and efficiency. The models used had minimal optimization done and have been adapted to a low-performance processor. To the best of our knowledge, there are no other studies that look at model comparisons with a limitation on processing power.

The significance of this work is two-fold. One, knowing which models are most effective in a low-end environment can help professionals with decisions with limited computational resources. Two, these results have some generalizability to higher-scale environments. A higher relative accuracy in this study predicts to some extent the accuracy with more computational resources.

VII. CONCLUSION

The study achieved the research objectives of comparing different machine learning models and comparing model results using evaluation metrics and the baseline model. Overall, with our constraints on computational efficiency, we were only able to train and evaluate Lasso and SVR machine learning models on the OASIS dataset. Otherwise, the O1 objective was completed. In regards to O2 and O3, the SVR model with a linear kernel had the best performance in terms of accuracy and computational efficiency. The Lasso model performed similarly to the linear SVR model, however, at a greater computational cost and loss of accuracy. From these results, we conclude that somewhat accurate brain age prediction is possible with shallow machine-learning techniques and a lack of computational power.

Future work based on the work done in this project includes data management to utilize the OpenBHB dataset and train the models with a greater amount of data to achieve higher accuracy as well as investigating minimal shallow machine learning techniques that can be used to achieve maximal accurate results.

REFERENCES

[1] Cole, J.H.; Franke, K. Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. *Trends Neurosci*. 2017, 40, 681–690. https://www.cell.com/trends/neurosciences/abstract/S0166-2236(17)30187

[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. https://www.sciencedirect.com/science/article/abs/pii/S0925492721000226?via%3Dihub

[3] 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. https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(21)00393-5/fulltext

[4] 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.

[5] Cole, J. H., Leech, R. & Sharp, D. J., Alzheimer's Disease Neuroimaging Initiative. Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Ann. Neurol*. 77, 571–581 (2015).

[6] Han J, Kim SY, Lee J, Lee WH. Brain Age Prediction: A Comparison between Machine Learning Models Using Brain Morphometric Data. *Sensors.* 2022; 22(20):8077. https://doi.org/10.3390/s22208077

[7] Zhang, Yu & Wang, Qifan & Chen, Xiaofei & Yan, Yuchao & Yang, Ruomiao & Liu, Zhentao & Jiahong, Fu. (2022). The Prediction of Spark-Ignition Engine Performance and Emissions Based on the SVR Algorithm. *Processes.* 10. 312. 10.3390/pr10020312.

[8] Cumplido-Mayoral I, García-Prat M, Operto G, Falcon C, Shekari M, Cacciaglia R, Milà-Alomà M, Lorenzini L, Ingala S, Meije Wink A, Mutsaerts HJMM, Minguillón C, Fauria K, Molinuevo JL, Haller S, Chetelat G, Waldman A, Schwarz AJ, Barkhof F, Suridjan I, Kollmorgen G, Bayfield A, Zetterberg H, Blennow K, Suárez-Calvet M, Vilaplana V, Gispert JD; ALFA study; EPAD study; ADNI study; OASIS study. Biological brain age prediction using machine learning on structural neuroimaging data: Multi-cohort validation against biomarkers of Alzheimer's disease and neurodegeneration stratified by sex. *Elife*. 2023 Apr 17;12:e81067. doi: 10.7554/eLife.81067. PMID: 37067031; PMCID: PMC10181824.

[9] Jonsson, B.A., Bjornsdottir, G., Thorgeirsson, T.E. et al. Brain age prediction using deep learning uncovers associated sequence variants. *Nat Commun* 10, 5409 (2019). https://doi.org/10.1038/s41467-019-13163-9

[10] Cole J. H. (2020). Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiology of aging*, 92, 34–42. https://doi.org/10.1016/j.neurobiolaging.2020.03.014

[11] *What is lasso regression? | IBM*. (n.d.). Www.ibm.com. https://www.ibm.com/topics/lasso-regression

[12] Da Costa PF, Dafflon J, Pinaya WHL. Brain-Age Prediction Using Shallow Machine Learning: Predictive Analytics Competition 2019. *Front Psychiatry*. 2020 Dec 2;11:604478. doi: 10.3389/fpsyt.2020.604478. PMID: 33343431; PMCID: PMC7738323.

[13] Tong, Y., Lu, W., Deng, Q., Chen, C., & Shen, Y. (2020). Automated identification of retinopathy of prematurity by image-based deep learning. *Eye and Vision,* 7(1). https://doi.org/10.1186/s40662-020-00206-2

[14] *Support Vector Regression*. (2024, March 5). Science Direct. Retrieved April 8, 2024, from https://www.sciencedirect.com/topics/computer-science/support-vector-regression

[15] SMOLA, A. J., & SCHOLKOPF, B. (n.d.). *A tutorial on support vector regression.*

[16] *OASIS Brains - Open Access Series of Imaging Studies.* (n.d.). Www.oasis-Brains.org. https://www.oasis-brains.org/