

Prediction of chronological age in healthy elderly subjects using the Extreme Gradient Boosting (XGBoost) algorithm with MRI brain cortical parcellation

JAIME GÓMEZ-RAMÍREZ*

Universidad de Cádiz, Spain

MIGUEL A. FERNÁNDEZ-BLÁZQUEZ

Universidad Complutense de Madrid, Spain

JAVIER J. GONZÁLEZ-ROSA

Universidad de Cádiz, Spain

January 7, 2022

Abstract

Normal aging is associated with changes in volumetric indices of brain atrophy. A quantitative understanding of age-related brain changes can shed light on successful aging. To investigate the effect of age on global and regional brain volumes and cortical surface and thickness, 3,514 magnetic resonance imaging (MRI) scans were analyzed using automated brain segmentation and parcellation methods in elderly healthy individuals (69-88 years of age). Machine learning models to predict the chronological age were built, followed by the relative importance of features when making an age prediction. The Extreme Gradient Boosting (XGBoost) algorithm achieved a mean absolute error of 2 years in predicting the age of new subjects. Feature importance analysis showed that the brain to intracranial volume ratio is the most important feature to predict age, followed by the hippocampi volumes. Additionally, the cortical thickness in temporal and parietal lobes showed superior predictive value than frontal and occipital lobes. Insights from this approach

that integrates model prediction and interpretation may help shorten the current explanatory gap between chronological age and biological brain age.

I. INTRODUCTION

Magnetic Resonance Imaging (MRI) has revolutionized clinical neuroscience, assisting in better diagnostics and playing a crucial role in helping to close the gap between basic and clinical research [Granziera et al., 2021]. The contrast and detail that MRI achieves are unparalleled to detect brain abnormalities, tumors, and micro-hemorrhages. Technological advancements, in particular, software development, are making it possible to assess conditions that went previously undetected [Harisinghani et al., 2019]. Furthermore, faster imaging is helping to alleviate the clinical burden in patients and clinicians alike with AI-based analysis, drastically reducing the time required for image reconstruction [Prakkamakul et al., 2016], [?] zhao2019intelligent In addition, the ever-increasing computational capacity and availability of AI techniques are accelerating the use of automated image post-processing for diagnostics and prognosis assessment in clinics and hospitals. Volumetric analysis aiming at quantifying the volume of brain structures and the thickness and gyration of cortical areas can be particularly effective in flagging brain abnormalities in large datasets [Harisinghani et al., 2019].

MRI is also helping to characterize the neuroanatomy of healthy brain aging across ages and conditions. Several studies show the intricacies of morphological changes visible for the whole brain [MacDonald and Pike, 2021], [Royle et al., 2013] as well as for the cerebral cortex [Salat et al., 2004], subcortical grey matter structures and white matter integrity differences [Bennett et al., 2017].

Changes in functional abilities and brain structural alterations such as atrophy are to be expected during aging. However, there is not always a clear separation line that distinguishes the effects of brain aging from neurodegeneration. For example, white matter hyperintensities or microbleeds are present in neurological conditions such as Alzheimer's disease [Shams et al., 2015], and they are also been found in aging asymptomatic individuals [Prins and Scheltens, 2015]. Both normal aging and neurodegenerative diseases are accompanied by brain morphological changes, notably atrophy or the loss of tissue volume resulting from cellular death. At advanced stages of the disease, the cellular loss or synaptic pruning associated with atrophy is more easily recognizable than at early stages, which makes the early detection of neurodegenerative pathologies particularly challenging.

Whole-brain atrophy intended as a diminution of brain volume normalized to the intracranial volume can be measured with T1-weighted structural

*Corresponding author jd.gomezramirez@gmail.com

MR images. Both grey matter and white matter decline with aging, with at the same time an increase in the size of the ventricles and Cerebrospinal Fluid CSF volume [Grant et al., 1987]. Among the subcortical structures in the limbic system and basal ganglia, the hippocampus is arguably the best characterized anatomic volume concerning aging. Hippocampal volume loss is also a well-recognized biomarker for the diagnosis of Alzheimer's disease in cross-sectional [Maruszak and Thuret, 2014], longitudinal [McRae-McKee et al., 2019], and meta-analysis [Whitwell et al., 2012]. The variation of hippocampal volume as a function of age studied in the UK Biobank Imaging dataset ($N = 19,793$) shows an acceleration in hippocampal volume reduction in middle age (starting around 50 years of age) [Nobis et al., 2019]. Age-related atrophy in other subcortical structures varies, with a more rapid decline in the thalamus and putamen compared to caudate and amygdala [Serbruyns et al., 2015]. However, the software libraries used for automated segmentation can introduce variability in the volume estimates [Gomez-Ramirez et al., 2021].

Post mortem histological studies, and more recently, magnetic resonance imaging (MRI), have identified an association between aging and cortical thinning. Age-related differences show, however, regional variability with more pronounced decreases in cortical thickness in frontal regions and often located in cortical areas involved in decision, recognition, memory, and speech [Fjell et al., 2015]. The observed loss of volume tissue can be caused by neurodegenerative processes linked to cellular loss, as in neurofibrillary tangles toxicity [Cullen and Halliday, 1998] but it could also be related to synaptic pruning which occurs in normal aging [Sakai, 2020].

The "last in first out" hypothesis states that the last brain regions to develop, tend to be the first ones to decline. The hypothesis is rooted in the idea of lifespan composed of two stages, first, a developmental phase then followed by an aging phase characterized by progressive cortical thinning with the onset linked to late-maturing regions of the brain such as the heteromodal association cortices [McGinnis et al., 2011]. Cortical thinning spans widespread cortical regions with unequal effects e.g. more prominent in the prefrontal cortex and less so in the parahippocampal cortex [Sowell et al., 2003], [Salat et al., 2004]. Longitudinal estimates of cortical thinning converge with cross-sectional studies in finding significant thinning in heteromodal association cortex with a later expected atrophy in the primary cortex in line with the "last in first out" hypothesis [Shaw et al., 2016].

Nevertheless, cortical thickness is not the only morphological measure sensitive to aging, gyration, or the ratio of the regional surface area relative to the surface area of a simulated enclosing surface decrease primarily in the parietal cortex [Madan and Kensinger, 2018]. Theoretical measures such as fractal dimensionality have also been used as a marker to quantify age-related brain atrophy [Madan and Kensinger, 2016]. Post mortem histological studies, and more recently, magnetic resonance imaging (MRI), have identified

an association between aging and cortical thinning. Age-related differences show, however, regional variability with more pronounced decreases in cortical thickness in frontal regions and often located in cortical areas involved in decision, recognition, memory, and speech [Fjell et al., 2015]. The observed loss of volume tissue can be caused by neurodegenerative processes linked to cellular loss, as in neurofibrillary tangles toxicity [Cullen and Halliday, 1998] but also to synaptic pruning which occurs in normal aging [Sakai, 2020].

Since, as already mentioned, brain volume decline is associated with age, it is possible at least in principle to use the volumetric measurement of brain atrophy to estimate its age. It is, however, useful to distinguish between biological and chronological age. The gap between both ages could indicate the aging pace of a given subject, that is to say, a subject with biological age larger than chronological age would imply a faster than expected aging decline, on the other hand, a person more chronologically than biologically aged could indicate that that person is aging more slowly than expected.

Neuroimaging-based studies for chronological age prediction use different features, acquisition techniques, and MRI sequences. In [Aycheh et al., 2018] a supervised regression model to predict age using the Destrieux Atlas of cortical parcellation obtained a mean absolute error of 4.05 years [Destrieux et al., 2010]. Cortical thickness combined with diffusion MRI has also been used as input for predicting brain age in multiple regression models [Kondo et al., 2015], [Cherubini et al., 2016]. Studies comparing performance between linear and nonlinear modeling approaches (e.g. neural networks) and combining a wide range of features have shown that performance gain from larger training can be limited if data are obtained with different acquisition protocols [MacDonald et al., 2019].

Although the association between the thinning of cortical areas and aging is recognized, how cortical thinning and overall brain atrophy progress in elderly healthy subjects is an open problem. In this study, we address this issue by building a machine learning model to predict chronological age in a large set of subjects using brain segmentation and cortical parcellation data, collected in a time horizon of 6 years. Different from other studies, we tackled not only the model prediction capabilities but also the model interpretability to assess the importance of the different brain areas, including whole brain and subcortical volumes and cortical thickness of sulcal and gyral areas, for age prediction using SHAP (SHapley Additive exPlanations) [Shapley, 1953].

II. METHODS

i. Study Participants

The dataset used in this study comes from a single-center, observational cohort study [Gómez-Ramírez et al., 2020a], [Gomez-Ramirez et al., 2021]. The partici-

pants are home-dwelling elderly volunteers, 69-88 years of age and without relevant psychiatric, neurological, or systemic disorders. Of the initial 1213 subjects, those diagnosed with mild cognitive impairment (MCI) or dementia were excluded, resulting in a cohort of 890 healthy elderly subjects. The participants signed informed consent and undertook a yearly systematic clinical assessment, including medical history, neurological and neuropsychological examinations, and brain MRI. Apolipoprotein E (APOE) genotype was also studied with total DNA isolated from peripheral blood following standard procedures. The subjects were assessed yearly for six years resulting in 3918 assessments, with the number of yearly visits per subject varying from 2 to 6 visits.

ii. Standard Protocol Approvals, Registrations, and Patient Consents

Ethical approval was granted by the Research Ethics Committee of Instituto de Salud Carlos III, and written informed consent was obtained from all the participants. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation, and with the Helsinki Declaration of 1975 and its later amendments.

iii. MRI Data Acquisition and Preprocessing

For all the subjects, brain MRIs were collected in 6 yearly visits, 948 in the first visit, 744 in the second, 711 in the third, 622 in the fourth, 529 in the fifth, and 364 in the sixth visit. The imaging data were acquired on a 3T General Electric scanner (GE Milwaukee) utilizing the following configuration, T1-weighted inversion recovery, flip angle 12°, 3-D pulse sequence: echo time Min. full, time inversion 600 ms, receiver bandwidth 19.23 kHz, field of view = 24.0 cm, slice thickness 1 mm, Freq. × Phase (288 × 288). The preprocessing of MRI 3 Tesla images consisted of generating an isotropic brain image with non-brain tissue removed. We used the FreeSurfer cortical surface reconstruction pipeline as the initial preprocessing step. The postprocessing was performed with FreeSurfer [Fischl, 2012], version freesurfer-darwin-OSX-ElCapitan-dev-20190328-6241d26 running under Mac OS X, product version 10.14.5 as described in [Gómez-Ramírez and González-Rosa, 2021]. Parcellation was performed using the Destrieux cortical Atlas which is based on the division of the cortex into gyri or sections of the cortex visible on the pial view, and sulci or the hidden parts of the cortex according to the curvature value of the surface [Destrieux et al., 2010]. Figure 1 shows two views of the Destrieux atlas parcellation of the left hemisphere.

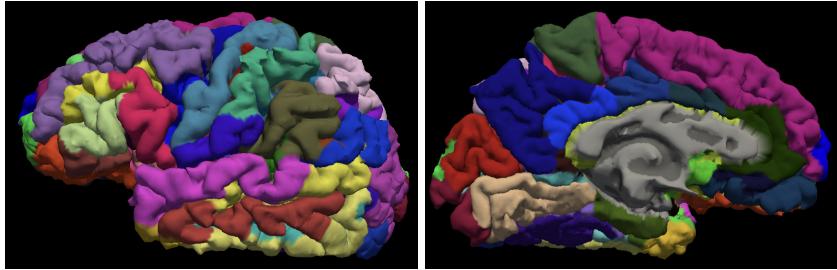


Figure 1: Cortical parcellation of the left hemisphere according to the Destrieux Atlas.

iv. Anomaly detection with the Isolation Forest algorithm

In this study we used brain volume, mean subcortical volumes of seven structures -thalamus, putamen, hippocampus, caudate, pallidum, amygdala, accumbens- and the mean cortical thickness of the 148 regions of interest defined in the Destrieux Atlas. In the first stage, images were manually assessed for quality, and scans considered not suitable for analysis due to artifacts were discarded. In a second step, we performed anomaly detection using the scikit-learn Isolation Forest implementation [Pedregosa et al., 2011], resulting in a total of 3514 scans.

Isolation Forest is an ensemble method [Breiman, 1996] that has been used as an outlier detector in a variety of datasets [Domingues et al., 2018], [Alaverdyan, 2019], [Gómez-Ramírez and González-Rosa, 2021]. The algorithm works by isolating each point in the dataset to assess whether the point is an outlier. The idea behind the algorithm is that points that are outliers are easier to separate than those that they are not. Thus, for each point and each feature, a range comprised between the minimum-maximum values is declared. The algorithm randomly selects a feature and a value and depending on where the value falls in the range, the range is switched upwards to the maximum or downwards to the minimum value. The algorithm proceeds iteratively until the point is isolated, that is, the point is alone inside the range for all features. Finally, the algorithm considers a point as an outlier depending on the number of iterations required to isolate it.

v. Statistical analysis

The demographic (age, sex, educational attainment level) and genetic information (presence of the APOE4 allele) of each subject in the study is summarized in Table 1.

Table 3 shows the results of the segmentation and cortical parcellation analysis. The brain to intracranial volume ratio (Brain2ICV) in % and the estimated volume of seven subcortical structures for both hemispheres in mm^3

$N_{SUBJECTS} = 948$	$\mu \pm \sigma$
Age	74.68 ± 3.85
Sex	F 519, M 280
APOE	$\epsilon(23)\epsilon(23)145, \epsilon(23)\epsilon(4)647, \epsilon(4)\epsilon(4) 7$
Education	None 150, Primary 232, Secondary 203, University 214

Table 1: The table shows the description of the demographic and genetic variables included in the study. M: male, F: female, APOE $\epsilon(23)\epsilon(23)$: lacking allele $\epsilon 4$, $\epsilon(23)\epsilon 4$ one allele $\epsilon 4$ and $\epsilon 4\epsilon 4$ both alleles $\epsilon 4$. Educational attainment level, None: no formal education, Primary: primary education degree, Secondary: secondary high school, University: university studies. The subjects participating in the study had at least one MRI with the resulting volumetric analysis shown in Table 3.

are followed by the average cortical thickness in mm of the regions defined in the surface-based atlas known as Destriau atlas [Destrieux et al., 2010].

The relationship between volumetric and thickness estimates and the variables Sex and APOE4 is investigated using regression analysis. We conduct hypothesis testing on the regression coefficients obtained in the model $Y = \beta_0 + \beta_1 X_{Sex}$ with results shown in the column "Sex p-val" and ANOVA analysis with the three-valued variable APOE4 shown in the last column "APOE PR(>F)". In the first case, the null hypothesis is, $H_0 : \beta_1 = 0$, and the alternative hypothesis, $H_A : \beta_1 \neq 0$. Using the p-value method, H_0 is rejected when the p-value of the test statistic is small (e.g. less than 0.05(*) or 0.01(**)). The hypotheses of interest in ANOVA are $H_0 : \mu_1 = \mu_2 = \mu_3$, and H_1 : Means are not all equal.

The null hypothesis is rejected for Sex in the Brain2ICV variable, all the subcortical structures, and in the large majority of cortical thickness areas. ANOVA using the APOE4 allele, on the other hand, fails to reject the null hypothesis in the large majority of subcortical and cortical regions. These results are not surprising since the volume and thickness were not adjusted for cerebral volume and effect sizes for sex differences are to be expected in either volume estimates and cortical thickness [Leonard et al., 2008], [Lotze et al., 2019].

Carriage of the APOE4 genotype is the main genetic risk factor for developing late-onset Alzheimer's disease. However, the effect of the APOE4 allele on volumetric or cortical thickness estimates in asymptomatic individuals is being investigated. Our results are in agreement with recent studies showing no overall risk effects associated with APOE4 in healthy adults for cortical thickness and subcortical volume [Mole et al., 2020]. Nevertheless, we find that the APOE4 genotype may have a deleterious effect on hippocampal and amygdala volumes which is in agreement with the literature on atrophy of

the hippocampus and amygdala in healthy elderly and impaired memory performance in individuals [Fleisher et al., 2005], [Orihashi et al., 2020].

	Volume(mm^3)	Sex p-val	APOE PR(>F)
Brain2ICV	0.698 ± 0.029(%)	**	> 0.05
LH Th	6000 ± 653	**	> 0.05
RH Th	5848 ± 582	**	> 0.05
LH Pu	3551 ± 387	**	> 0.05
RH Pu	3635 ± 415	**	> 0.05
LH Am	1414 ± 215	**	**
RH Am	1263 ± 193	**	*
LH Pa	1721 ± 252	**	> 0.05
RH Pa	1653 ± 244	**	> 0.05
LH Ca	3249 ± 490	**	> 0.05
RH Ca	3430 ± 252	**	> 0.05
LH Hp	3551 ± 387	**	> 0.05
RH Hp	3634 ± 415	**	**
LH Ac	462 ± 91	**	> 0.05
RH Ac	465 ± 90	**	> 0.05
Avg.Thickness(mm)			
RH STempSuperior	2.3 ± 0.2	**	> 0.05
LH STempSuperior	2.4 ± 0.1	**	> 0.05
RH STempSuperior	2.4 ± 0.1	**	> 0.05
LH STempInferior	2.5 ± 0.1	**	> 0.05
RH STempInferior	2.4 ± 0.1	**	> 0.05
LH SOccTempMedandLingual	2.4 ± 0.2	**	> 0.05
RH SOccTempMedandLingual	2.4 ± 0.2	**	> 0.05
LH SOccTempLat	2.5 ± 0.2	**	> 0.05
RH SOccTempLat	2.5 ± 0.2	**	> 0.05
LH GTempMid	2.9 ± 0.1	**	> 0.05
RH GTempMid	2.9 ± 0.1	> 0.05	> 0.05
LH GTempInf	2.9 ± 1.6	> 0.05	> 0.05
RH GTempInf	2.8 ± 1.7	> 0.05	> 0.05
LH GTempSup	2.5 ± 0.1	**	> 0.05
RH GTempSup	2.5 ± 0.1	**	> 0.05
LH GTempSupPlanPolar	3.1 ± 0.2	*	> 0.05
RH GTempSupPlanPolar	3.0 ± 0.72	**	> 0.05
LH GTempSupLateral	3.0 ± 0.1	**	> 0.05
RH GTempSupLateral	2.4 ± 1.7	> 0.05	> 0.05
LH GTempSupTransv	2.4 ± 1.7	**	> 0.05
RH GTempSupTransv	2.5 ± 1.8	**	> 0.05
RH SIn	2.1 ± 0.1	**	> 0.05
LH SFrontSup	2.1 ± 0.1	**	> 0.05

RH SFrontSup	2.1 ± 0.1	**	> 0.05
LH SFrontMid	2.1 ± 0.1	**	> 0.05
RH SFrontMid	2.1 ± 0.1	0.5781	> 0.05
LH SFrontInf	2.2 ± 0.1	> 0.05	> 0.05
RH SFrontInf	2.1 ± 0.1	> 0.05	**
LH SFrontSup	2.2 ± 0.1	**	> 0.05
RH SFrontSup	2.1 ± 0.1	**	> 0.05
LH GFrontSupp	2.6 ± 1.1	**	> 0.05
RH GFrontSupp	2.6 ± 1.1	**	> 0.05
LH GFrontMid	2.5 ± 1.1	**	> 0.05
RH GFrontMid	2.5 ± 1.1	**	> 0.05
LH GFrontInfTriangul	2.5 ± 1.6	**	
RH GFrontInfTriangul	2.5 ± 1.4	**	
LH GFrontInfOrbital	2.7 ± 0.2	**	> 0.05
RH GFrontInfOrbital	2.6 ± 0.2	**	> 0.05
LH GFrontInfOpercular	2.65 ± 0.13	> 0.05	> 0.05
RH GFrontInfOpercular	2.65 ± 0.13	*	> 0.05
LH GCingPostV	2.4 ± 0.3	*	> 0.05
RH GCingPostV	2.5 ± 0.3	> 0.05	> 0.05
LH SCingMarginalis	2.1 ± 0.1	> 0.05	> 0.05
RH SCingMarginalis	2.1 ± 0.1	**	> 0.05
LH SSubParietal	2.2 ± 0.1	**	> 0.05
RH SSubParietal	2.3 ± 0.1	**	> 0.05
LH SSubOrbital	2.3 ± 0.2	> 0.05	> 0.05
RH SSubOrbital	2.3 ± 0.3	*	**
LH SPreCentralSuperior	2.3 ± 0.1	**	> 0.05
RH SPreCentralSuperior	2.3 ± 0.1	**	> 0.05
LH SPreCentralInferior	2.3 ± 0.1	**	> 0.05
RH SPreCentralInferior	2.3 ± 0.1	> 0.05	> 0.05
LH SPostCentral	2.1 ± 0.1	**	> 0.05
RH SPostCentral	2.1 ± 0.1	**	> 0.05
RH SPeriCallosal	1.8 ± 0.3	**	> 0.05
LH SParietoOcc	2.2 ± 0.1	**	> 0.05
RH SParietoOcc	2.2 ± 0.1	**	> 0.05
LH SOrbMedOlfact	2.1 ± 1.6	**	> 0.05
RH SOrbMedOlfact	2.1 ± 1.5	**	> 0.05
LH SOrbitalLat	2.1 ± 1.1	**	> 0.05
RH SOrbitalLat	2.1 ± 1.1	**	*
LH SOrbitalHShaped	2.6 ± 1.2	**	> 0.05
RH SOrbitalHShaped	2.5 ± 1.2	**	> 0.05
LH SOccMideandLunatus	2.3 ± 1.2	**	> 0.05
RH SOccMideandLunatus	2.3 ± 1.1	**	> 0.05
LH SIntraParietandPariettrans	2.1 ± 0.1	**	

RH SIntraParietandPariettrans	2.1 ± 0.1	**	
LH GParietalSup	2.3 ± 1.2	**	> 0.05
RH GParietalSup	2.3 ± 1.2	**	> 0.05
LH GParietInfSupramar	2.6 ± 1.3	**	> 0.05
RH GParietInfSupramar	2.6 ± 1.2	**	> 0.05
LH GParietInfAngular	2.5 ± 1.2	**	> 0.05
RH GParietInfAngular	2.5 ± 1.3	**	**
LH SCollatTransvPost	2.1 ± 0.1	**	> 0.05
RH SCollTatransvPost	2.1 ± 0.1	**	> 0.05
LH SCollTransvAnt	2.6 ± 0.2	**	> 0.05
RH SCollTransvAnt	2.5 ± 0.2	**	> 0.05
LH PoleOcc	3.3 ± 0.2	> 0.05	> 0.05
RH PoleOcc	3.3 ± 0.2	**	> 0.05
LH GOccSup	2.1 ± 1.1	**	> 0.05
RH GOccSup	2.4 ± 1.3	**	> 0.05
LH GOccMid	2.5 ± 0.1	**	> 0.05
RH GOccMid	2.5 ± 0.1	**	* > 0.05
LH GOccTempMedParahip	3.1 ± 0.6	**	> 0.05
RH GOccTempMedParahip	3.2 ± 0.2	**	> 0.05
LH GOccTempMedLingual	2.1 ± 0.1	> 0.05	> 0.05
RH GOccTempMedLingual	2.1 ± 0.1	> 0.05	> 0.05
LH GOccTempLatFusi	2.8 ± 1.5	> 0.05	> 0.05
RH GOccTempLatFusi	2.8 ± 1.5	**	> 0.05
LH SInsSup	2.4 ± 0.1	> 0.05	> 0.05
RH SInsSup	2.4 ± 0.1	**	**
LH SInsInf	2.6 ± 0.2	**	> 0.05
RH SInsInf	2.5 ± 0.1	> 0.05	> 0.05
LH SCircInsAnt	2.7 ± 0.2	> 0.05	> 0.05
RH SCircInsAnt	2.7 ± 0.2	**	> 0.05
LH GInsularShort	3.4 ± 0.3	**	> 0.05
RH GInsularShort	3.4 ± 0.3	**	> 0.05
LH GCentInsula	3.2 ± 0.3	**	> 0.05
RH GCentInsula	3.3 ± 0.3	**	> 0.05
LH SCentral	2.0 ± 0.1	**	> 0.05
RH SCentral	1.9 ± 0.1	**	> 0.05
LH GPreCentral	2.6 ± 1.7	**	> 0.05
RH GPreCentral	2.6 ± 1.6	**	> 0.05
LH GPostCentral	2.1 ± 1.5	**	> 0.05
RH GPostCentral	2.1 ± 1.5	**	*
LH SCalcarine	1.9 ± 0.1	> 0.05	> 0.05
RH SCalcarine	1.9 ± 0.1	**	> 0.05
LH GRectus	2.5 ± 0.2	> 0.05	> 0.05

RH GRectus	2.5 ± 0.2	> 0.05	> 0.05
LH GOrbital	2.7 ± 1.7	**	> 0.05
RH GOrbital	2.7 ± 1.7	**	> 0.05
LH GCuneus	1.9 ± 0.1	> 0.05	> 0.05
RH GCuneus	1.9 ± 0.1	**	> 0.05
LH LatFisPost	2.3 ± 0.1	**	> 0.05
RH LatFisPost	2.3 ± 0.1	**	*
LH LatFisAntHoriz	2.1 ± 0.1	> 0.05	> 0.05
RH LatFisAntHoriz	2.4 ± 0.1	**	> 0.05

Table 3: The table columns from left to right include the name of the subcortical or cortical area, the gray matter volume ((mm^3)) for subcortical structures, the cortical thickness ((mm)) defined for areas in the surface cortical Atlas used, the p-values of the T-test and the ANOVA test for Sex and APOE4 respectively. The first row of data shows the brain to intracranial volume (Brain2ICV) ratio, followed by the subcortical structures -Thalamus, Putamen, Amygdala, Pallidum, Caudate, Hippocampus, and Accumbens. Next, the parcellation scheme defined in [Destrieux et al., 2010] in which the cortex is divided into gyral and sulcal regions. The brain areas names on the left column are self-descriptive with LH and RH referring to each hemisphere and the first letter S, G for gyral or sulcal thickness. For more information about the cortical atlas and the naming of the areas see [des,].

For the sake of illustration, Figure 2 shows the intracranial volume segmentation and cortical parcellation analysis obtained for four subjects in the study. The summary of the automated segmentation and cortical parcellation of the 3,514 scans included in the study are shown in Table 3.

vi. Age prediction analysis

We aim at predicting the chronological age of elderly healthy adults using structural MRI volumetric analysis and demographic features of interest such as sex, educational attainment level, and the genetic risk factor in dementia (APOE4). We build two types of predicting models, linear and nonlinear, and we assess their performance by testing the model predictions on a held-out dataset of points not used for training.

Schematically, the prediction problem we aim to resolve can be succinctly described as the supervised learning model shown in Equation 1

$$\Gamma(X) \rightarrow Y \quad (1)$$

such that the function Γ maps the Input Space $X = (Sex, Educational\ attainment\ level, APOE4), (Subcortical\ volume\ estimates), (Cortical\ thickness\ estimates)$ into the Output space $Y = (chronological\ age)$.

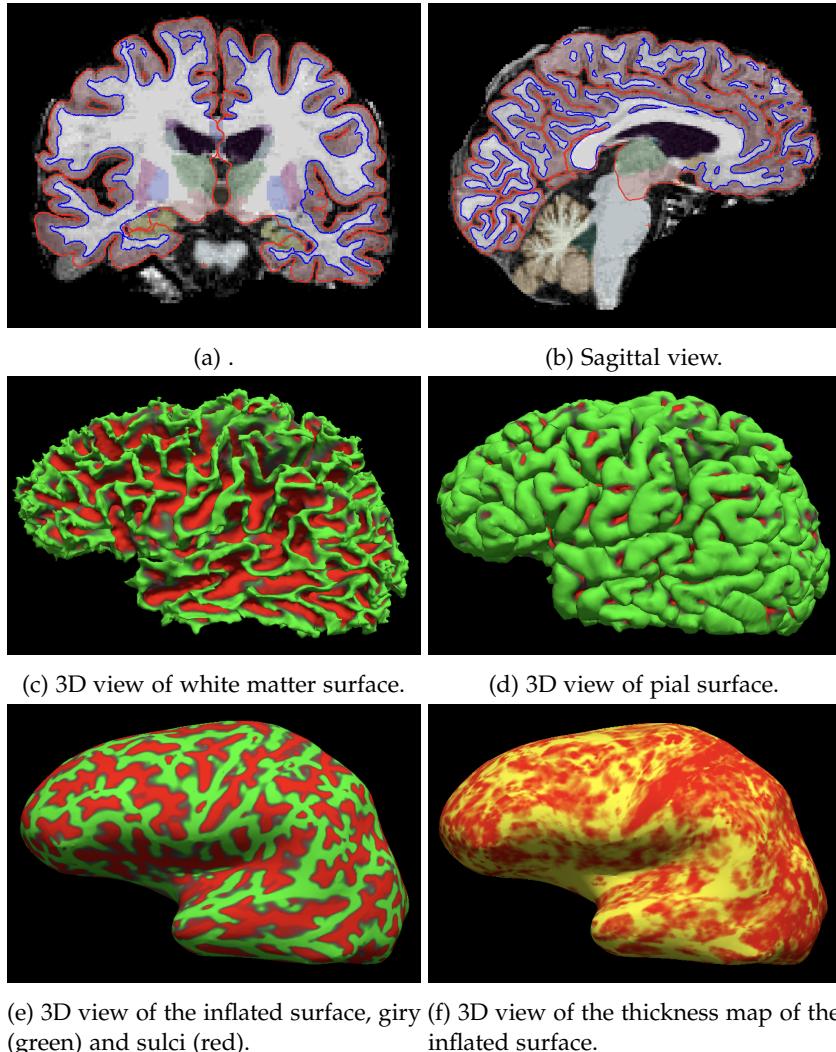


Figure 2: (a) and (b) show the coronal and sagittal segmentation results. The edge color blue indicates the demarcation of the white matter surface, and the red edge the pial surface. Plots (c),(d),(e) and (f) show the three-dimensional view of the surface analysis for the same subject. Surfaces of white matter (a), pial surface (b), inflated map (c) and thickness of the inflated map (d).

The dimensionality of the Input Space X of the model is 153, including the sex of the individuals (female or male), education level, the Apolipoprotein E gene called APOE4 together with 150 brain imaging features, namely the brain to intracranial volume ratio (Brain2ICV), the volume estimates of seven

subcortical structures (caudate, pallidum, putamen, thalamus, hippocampus, amygdala, and accumbens) and the cortical sulci and gyri thickness based on the parcellation defined in the Destrieux cortical atlas (Figure 1).

We build two supervised learning models: Partial Least Squares (PLS) and Extreme Gradient Boosting (XGBoost). The PLS regression model extracts the components of the input X and the output Y that explain the most shared variance among X and Y [Garthwaite, 1994]. Thus, the two components maximally correlated correspond to the first component in feature space X and the target Y .

PLS is based on principal component analysis, so it deals with multicollinearity in the input space by finding a linear transformation W such that the new input space is linearly independent while at the same time maximizing the covariance between Y and X . Thus, X is transformed into X' , by a linear transformation W as in $X' = XW$.

The nonlinear supervised learning model of choice is the decision-tree-based ensemble machine learning algorithm called Extreme Gradient Boosting or XGBoost for short. XGBoost [Chen et al., 2015] is an implementation of gradient boosted trees algorithm that uses a regularized gradient boosting algorithm to accurately predict the target variable, in our case, the chronological age. XGBoost combines weaker models with stronger ones, which are added to improve the overall performance. During training, the gradient descent algorithm minimizes the loss by adding new trees to predict more accurately than weaker decision trees. The models are all here regression trees mapping the input data points (X) to one of the tree's leaf which contains the age (Y) that we want to predict. The objective function that the algorithm tries to minimize combines the difference between the predicted and target outputs and a penalty term for model complexity. Formally, given the dataset tuple (X, Y), the gradient tree boosting algorithm tries to minimize the differentiable loss function L (Equation 2)

$$L(Y, \Gamma(X)) \quad (2)$$

where $\Gamma(X)$ represents an ensemble of n regression trees which are sequentially added to incrementally better predict the residuals of previous trees. Formally,

$$\Gamma_n(X) = \Gamma_{n-1}(X) + \alpha_n g_n(X, \rho_{n-1}) \quad (3)$$

with α_i and ρ_i $i = 1..n$, the regularization and the residual parameter respectively for the i_{th} tree. The function g_i is trained to predict the residuals of the precedent tree in the forest (ρ_{i-1}).

Feature importance analysis provides insight into the workings of the predictive model used, allowing for the interpretability of the results. SHapley Additive exPlanations or SHAP for short was originally developed in cooperative or coalitional game theory and in contrast with non-cooperative Nash equilibrium models [Ritzberger et al., 2002]. Shapley values show how much a

given feature changes the prediction compared to the prediction at the baseline value of that feature. The Shapley value[Shapley, 1953] based approach is being increasingly used by the machine learning community to deal with the interpretable feature subset selection problem [Tripathi et al., 2020]. An important drawback of Shapley values is that they provide additive contributions (attributions) of explanatory variables. If the model is not additive, then the Shapley values may be misleading. For a more in detail description of SHAP values see [Gómez-Ramírez et al., 2020a] and references within.

III. RESULTS AND DISCUSSION

Table ?? shows the results for the age predictors built, linear (PLS) and nonlinear (XGBoost). The performance is evaluated for the holdout set, that is, the dataset is split into train and test sets. We use 75% of data for training and the remaining 25% for testing the model performance on unseen data.

For the partial least squares regression model (PLS), the age of the subjects can be estimated with a maximum residual error (MAE) of 2.57 years. Other metrics such as the maximum residual error (MXE), the mean squared error regression loss (MSE), the mean absolute percentage error (MAPE), the median absolute error (MEDAE), and the explained variation, R^2 , are also shown in Table 4 (first row).

Next, we build, train and tune using cross validation the nonlinear regressor using the Extreme Gradient Boosting algorithm (XGBoost). The optimization or tuning of the hyperparameters $-\eta$, γ , *subsample*, *colsample by tree*, *max depth*, *min child weight* is performed using the Grid search method [Bergstra and Bengio, 2012]. The hyperparameters η and γ correspond to the learning rate and the minimum loss reduction respectively. The learning rate, also called shrinkage rate, η , is used to prevent overfitting. The minimum loss reduction γ acts as a pseudo-regularization hyperparameter in gradient boosting, the higher γ the higher the regularization. The hyperparameters *subsample* and *colsample by tree* control the sampling of the dataset at each boosting round, *subsample* is the fraction of observations (rows) and *colsample by tree* is the fraction of features (columns) used to train each tree. The hyperparameters *max depth*, *min child weight* add constraints on the architecture of the trees, *max depth* is the maximum number of nodes allowed from the root to the farthest leaf (very large max depth can cause overfitting) and *min child weight* is the minimum weight required in order to create a new node in the tree.

The XGBoost model achieves a mean absolute error (MAE) equals to 2.03 which is a 21% improvement relative to the partial least squares PLS model ($MAE = 2.57$) as shown in Table 4 (last row).

For the sake of comparison the performance metrics of the two models built -PLS and XGBoost-are compared with two dummy models, one always tries to predict the mean and the other the median age of the subjects. The

Model	Test Performance measure				
	MAE	MXE	MAPE	MEDAE	R^2
PLS	2.570177	10.2293404	0.03348523	2.15809710	0.353239
XGBoost	2.0301	8.7138485	0.0265	1.74578	0.591519

Table 4: Performance metric of PLS and XGBoost models in the test set (unseen subjects).

Mean Absolute Error (MAE) achieved by the former is 3.258 (PLS MAE = 2.57, XGboost MAE = 2.030) and the Median Absolute Error (MEDAE) by the last is 2.850 (PLS MEDAE = 2.158, XGboost MEDAE = 1.745) (Figure 3)

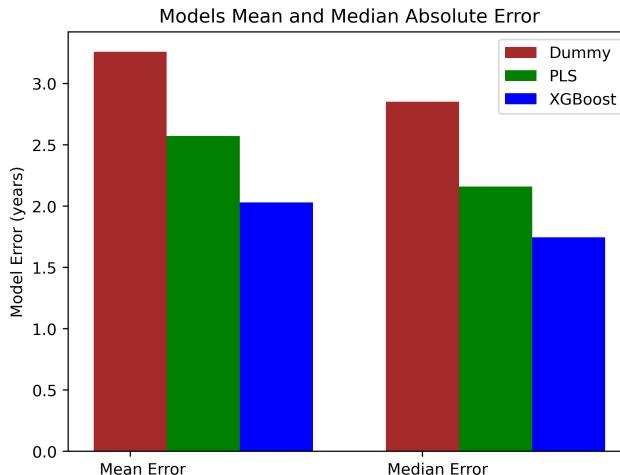


Figure 3: Mean and Median Absolute error for the dummy, PLS and XGBoost models. The XGBoost model shows superior results than the PLS. The R^2 by the dummy models is as expected 0.0, while $R^2=0.353$ for PLS and $R^2=0.5915$ for XGBoost.

Our results show that both linear and nonlinear models achieved a maximum Mean Absolute Error of fewer than three years, which is below previous studies. For example, in [Aycheh et al., 2018] the Mean Absolute Error achieved using a Gaussian Process Regression (GPR) algorithm in a sample of 2,911 cognitively normal subjects (age 45–91 years) was 4.05. Relevance vector machine algorithms have been used as well to predict brain age as in [Franke et al., 2010], [Wang et al., 2014], [Kondo et al., 2015] all with a Mean Absolute Error around 4.5 years. Neuroimaging techniques other than T1-weighted MRI for brain age prediction, for example, diffusion tensor imaging in a cohort of 188 subjects aged 4–85 years, obtained Mean Absolute Error varying depending

on age group and sex between 6 and 10 years [Mwangi et al., 2013]. Neuroanatomical prediction studies of biological maturity using resting-state fMRI [Dosenbach et al., 2010] in a cohort of 238 young subjects 7–30 years, reported 55% of the sample variance, and in multimodal studies combining MRI and diffusion-weighted imaging to predict child age, nonlinear modeling was able to account for more than 92% of the variance in age [Brown et al., 2012].

It has been suggested that the estimates of brain age from neuroanatomical data may suffer from systematic bias manifested as underestimated brain age for older subjects and overestimated for younger ones [Niu et al., 2020]. A plausible explanation for this bias may rely on the age distribution in the study, with nonnormal or skew-normal distributions playing a putative role in introducing a confounding effect [Smith et al., 2019]. Nevertheless, characteristics of the sample data used to train the models can lead to different estimates of Mean Absolute Error in age predictions [Pardoe and Kuzniecky, 2018] with some studies having a wide range of ages [Mwangi et al., 2013],[Schnack et al., 2016], [Cole et al., 2017] while others concentrated on a limited range [Brown et al., 2012]. In a healthy adult population, we may expect both maturation and aging effects in the brain, with an opposing rate of incidence, that is, being maturation prevalent during youth and aging during old age. Of note, the characterization of brain age in this study is intended to assess the effects of aging using voluntary focusing on a specific age structure (elderly population 68 years and older).

i. Feature Importance

In addition to being able to build predictive models of chronological age from age, sex, APOE4 gene, and brain volumetric estimates, it is of particular interest to understand the relative feature importance among predictors. The identification of brain regions and structures may help us characterize the localized effects of normal brain aging and to use the age predictors as potential biomarkers for neurodegenerative diseases [Gomez-Ramirez and Wu, 2014], [Eckerström et al., 2018].

The approach used here for addressing the feature selection problem at hand is Shapley values. Shapley or SHAP values can be used to decompose predictions into the sum of the effect of each feature.

Figure 4 shows the twenty most important features to predict the chronological age of the subjects in the test set according to the Shapley value method. The most important features are the brain to intracranial volume rate (Brain2ICV), followed by the volume of the hippocampi and the sulcal thickness of the occipito-temporal medial and lingual cortical regions. On the left, Figure 4a shows the absolute Shapley values of the most important features, and on the right, Figure 4b depicts the SHAP values for each subject. Thus, Figure 4a is the aggregate plot of Figure 4b.

Every dot in Figure 4b represents a different subject in the test set, colored by red or blue if the feature for that subject tends to push towards the right (more age) or to the left (less age) in the model predictions. For example, large (red) values in the top feature (Brain2ICV) decrease the prediction of age, that is, subjects with a larger brain to ICV ratio will tend to be less aged than subjects with small (blue) values. The same can be said for the rest of the features according to the figure, with however few exceptions such as the thickness of the rectus in the left hemisphere, the thickness of the sulci in the circular anterior insula, and the left caudate volume seem to have the opposite effect (big values seem to push towards more aged).

The total intracranial volume (TCV) changes during the life of a person, including a growth process during young adulthood followed by a steady decline starting at some point in middle adulthood which is accelerated in the presence of neurodegenerative disorders [Wolf et al., 2004], [Caspi et al., 2020]. The TCV is the volume within the cranium, including the cerebrum, meninges, and Cerebrospinal fluid (CSF). Thus, the TCV can be said to set an upper bound for the brain's volume. Accordingly, it is possible to build a proxy of the brain atrophy that an elder person went through her adult life by means of computing the ratio between the brain volume (TBV) and the TCV which represents the upper limit of brain volume [Gómez-Ramírez et al., 2020b].

The brain to total intra-cranial volume ratio or $Brain2ICV = \frac{TBV}{TCV}$ and the hippocampi are according to SHAP values the most important variables to predict chronological age. The key role of the hippocampus in learning and memory makes it a structure of particular interest to study the effects of normal aging in the brain [Kaye et al., 1997], [Bettio et al., 2017]. A recent study with a large normative database confirms that hippocampal volume loss accelerates in middle age [Nobis et al., 2019].

Cortical thickness declines due to normal aging, with a more visible cortical thinning effect on areas responsible for executive processing tasks and episodic memory retrieval, which are also known to be associated with age-related cognitive decline [Habeck et al., 2020].

Interestingly, the brain to intra cranial volume ratio and the hippocampi volume are more important (SHAP value) to predict chronological age than the cortical thickness of individual regions. We argue that the thickness of cortical areas could be more sensitive to the biological age than is to the chronological age. How the current results are subject to change when predicting the biological age of the brain using, for example, DNA methylation directly from the brain [Horvath et al., 2012] or the inflammatory aging clock [Sayed et al., 2021] is a matter of future studies [Schultz et al., 2020].

ii. Feature importance of cortical gyri and sulci

The evaluation of the importance of features for predicting chronological age can also be done by grouping the cortical areas based on their location (hemisphere and lobe) and type (sulci or gyri). We analyze next the relative importance for predicting the chronological age according to location and type of cortical areas segmented in our analysis and describe in the Destriaux Atlas.

Figure 5a depicts the aggregate importance of sulci and gyri cortical areas in both hemispheres. The aggregate importance is computed as the mean average of the SHAP values normalized. According to the SHAP values, the total sulcus thickness in the right hemisphere contains on average more information regarding the chronological age of the subject than gyrus thickness in either hemisphere (0.238, 0.248) and more than sulcus thickness on the left hemisphere (0.169).

Since the Destriaux Atlas used for automated cortical parcellation identifies the location of sulci and gyri in the different brain lobes and the insula, it is possible to plot the relative importance of the cortical areas for predicting chronological age. Figure 5 shows the relative importance according to the SHAP values of sulci and gyri in the different brain lobes and the insula. The sulci areas in the temporal lobe are more important for predicting age than the rest of parcellations. It is important to realize that the estimates depicted in the figure must be interpreted only in relative terms and not in absolute terms, that is to say, the sulci in the temporal lobe are more important for predicting chronological age than sulci and gyri in the frontal lobe. Likewise, the total thickness of sulci and gyri in both frontal and parietal lobes contains, according to this analysis, less information regarding the chronological age than the thickness of sulci and gyri in the temporal lobe.

The human brain presents an astonishing architectonic complexity, in large part as a result of the gyrification developmental process responsible for the distinctive folds found in the cortex. If we follow a topological or surface-based approach, we can characterize the cortex as a continuous sheet of neural tissue that has undergone a folding process, notably during embryonic life.

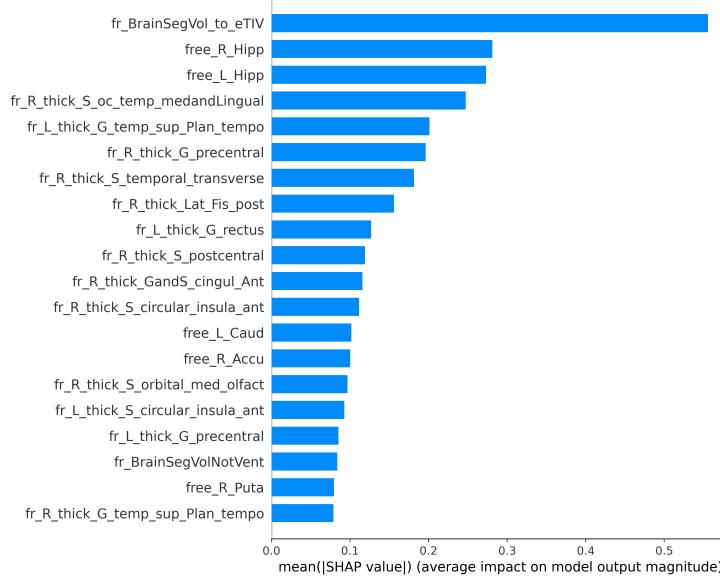
Cortical folds have a dual structure -gyrus or the peak of the fold and sulcus or the gyrus' through. In this study, we use the cortical parcellation defined in the Destriaux atlas which is based on sulco-gyral structure reconstruction [Duvernoy, 1999] where each vertex is classified as sulcus or gyrus and ascribed to one of the 74 labels defined for each hemisphere. As already discussed in [Destrieux et al., 2010], a surface-based cortical parcellation, is preferable to a volume-based parcellation, because a surface coordinate system is consistent with the cortical topology which is not necessarily the case in an orthogonal volume coordinate system. For example, two points nearby in Talairach coordinates may be widely separated on the cortical surface [Fischl et al., 1999]. Furthermore, the Destriaux atlas used here is advantageous to the Desikan-Killiany [Desikan et al., 2006] surface-based atlas which is gyral

based, while the Destriaux uses both gyri and sulci. This is important because as shown in [Van Essen, 2005] the sulci are more abundant in the cortex than the gyri. Thus we can say, without fear of error, that the human brain is more convex (sulci) than concave (gyri).

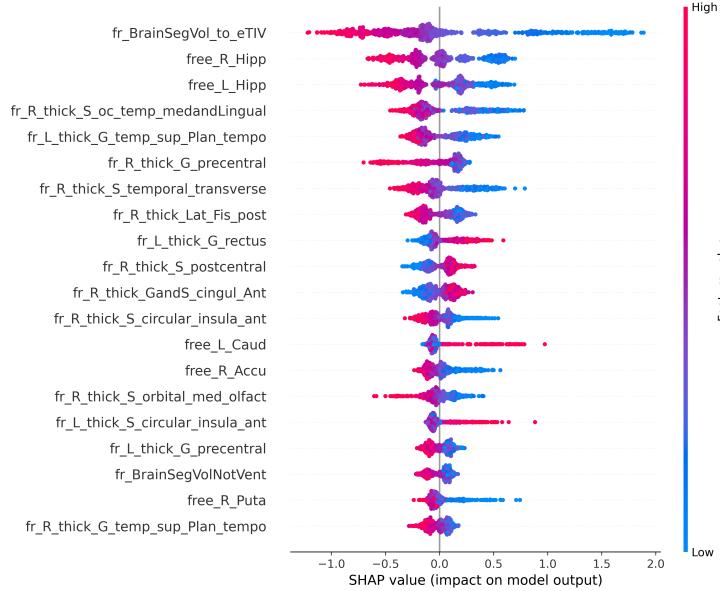
The last in, first out hypothesis implies that late-maturing regions of the brain, such as the heteromodal association cortices in the frontal lobe could be particularly vulnerable to age-related loss of structural integrity [McGinnis et al., 2011]. However, the idea of brain aging following a simple pattern of neural decline with the prefrontal cortex as the region most disrupted is being challenged by studies that suggest a more intricate picture. For example, the prefrontal cortex could play a compensatory role [Cabeza et al., 2016] or as suggested in [Morcom and Henson, 2018], normal aging could reflect non-specific neural responses rather than the predictable decline of target areas or compensation.

Our results indicate that thickness of areas temporal lobe is a better predictor of chronological age than the frontal lobe, and particularly, in cortical sulcal areas. This is in agreement with neuroanatomical evidence that indicates that between one-half to two-thirds of the cortical surface lies in the sulci and the lateral fossa of the brain [Destrieux et al., 2010].

It ought to be remarked that our model predicts chronological age and not biological brain age which can not be directly measured as is the case with the chronological age and can, however, be estimated via proxies such as DNA damage.



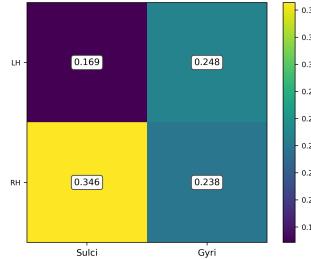
(a) Shapley values averaged for all subjects. The most important feature according to the SHAP values is the brain to intracranial volume rate, followed by the volume of the hippocampi.



(b) Shapley values for each subject. When the point distribution is clustered around 0 indicates that the feature is unimportant, the more spread the distribution is the more SHAP value important for predicting age.

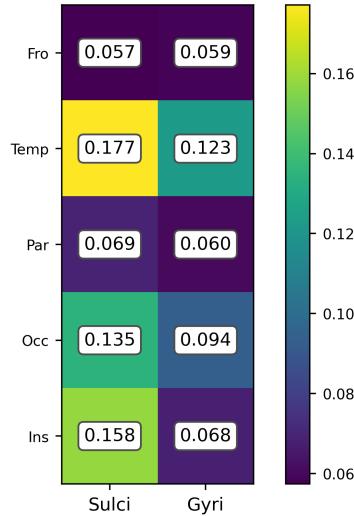
Figure 4: Most important features based on the computation of the SHAP values for each feature and sample of the total of 804 subjects in the test set, both in aggregate (4a) and for each data point (4b). The vertical axis of each figure represents the features ranked by importance (top to bottom) calculated as the sum of the SHAP value magnitudes over all samples (horizontal axis).

SHAP feature importance by Hemisphere & Cortical Sulcus Gyri



(a) SHAP feature importance in relative terms for brain cortical areas depending on the hemisphere and the cortical surface type (sulci, gyri). As shown in the figure, the aggregate importance of sulci in the right hemisphere for predicting chronological age is 0.346 and is computed as the mean of the SHAP values of right sulci areas normalized by the total of sulci and gyri areas in both hemispheres.

SHAP feature importance by Lobe & Cortical Sulcus Gyri



(b) SHAP feature importance in aggregate for brain cortical areas falling in brain lobes -frontal, occipital, parietal, temporal- and the insula. According to the SHAP values calculated, the temporal lobe contains more information for predicting age than the sulci and gyri located regions in the other brain lobes.

Figure 5: SHAP importance grouping cortical areas by hemisphere, lobe and type of fold.

IV. CONCLUSIONS

We trained machine learning models to predict chronological age. We find that the non-linear regression modeling eXtreme Gradient Boosting (XGBoost) obtain better results than the partial least squares PLS model, in particular, XGBoost achieves a mean absolute error of 2 years. Secondly, we find that the best predictor of chronological age is the brain to intracranial volume ratio followed by the hippocampi volume. Thirdly, the thickness of sulci is more important to predict age than the thickness of gyri, and this is particularly so for sulci in the temporal lobe. Our results show that simple volumetric features like the brain to intracranial volume ratio and hippocampal volume are no less important to predict chronological age than the cortical thickness of any specific area in the Destriaux Atlas. Ultimately, these results enable future research in the gap between the brain's biological and chronological age. The operationalization of this gap using the methodology here proposed may derive into a frailty index for healthy individuals or a potential biomarker for neurodegenerative disorders.

V. ACKNOWLEDGEMENTS

The authors acknowledge funding from Ministerio de Ciencia, Innovación y Universidades (CONNECT-AD) RTI2018-098762-B-C31 and Structural Funds ERDF (INTERREG V-A Spain-Portugal (POCTEP) Grant: 0348CIE6E).

VI. ETHICS DECLARATIONS

The authors declare no competing interests.

VII. INFORMATION SHARING STATEMENT

Code and data used in this research are publicly available on the Github repository under an Apache 2.0 license at https://github.com/grjd/chronological_brain_age.

REFERENCES

[des,]

[Alaverdyan, 2019] Alaverdyan, Z. (2019). *Unsupervised representation learning for anomaly detection on neuroimaging. Application to epilepsy lesion detection on brain MRI*. PhD thesis, Université de Lyon.

- [Aycheh et al., 2018] Aycheh, H. M., Seong, J.-K., Shin, J.-H., Na, D. L., Kang, B., Seo, S. W., and Sohn, K.-A. (2018). Biological brain age prediction using cortical thickness data: A large scale cohort study. *Frontiers in aging neuroscience*, 10:252.
- [Bennett et al., 2017] Bennett, I. J., Greenia, D. E., Maillard, P., Sajjadi, S. A., DeCarli, C., Corrada, M. M., and Kawas, C. H. (2017). Age-related white matter integrity differences in oldest-old without dementia. *Neurobiology of aging*, 56:108–114.
- [Bergstra and Bengio, 2012] Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).
- [Bettio et al., 2017] Bettio, L. E., Rajendran, L., and Gil-Mohapel, J. (2017). The effects of aging in the hippocampus and cognitive decline. *Neuroscience & Biobehavioral Reviews*, 79:66–86.
- [Breiman, 1996] Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2):123–140.
- [Brown et al., 2012] Brown, T. T., Kuperman, J. M., Chung, Y., Erhart, M., McCabe, C., Hagler Jr, D. J., Venkatraman, V. K., Akshoomoff, N., Amaral, D. G., Bloss, C. S., et al. (2012). Neuroanatomical assessment of biological maturity. *Current biology*, 22(18):1693–1698.
- [Cabeza et al., 2016] Cabeza, R., Nyberg, L., and Park, D. C. (2016). *Cognitive neuroscience of aging: Linking cognitive and cerebral aging*. Oxford University Press.
- [Caspi et al., 2020] Caspi, Y., Brouwer, R. M., Schnack, H. G., van de Nieuwenhuijzen, M. E., Cahn, W., Kahn, R. S., Niessen, W. J., van der Lugt, A., and Pol, H. H. (2020). Changes in the intracranial volume from early adulthood to the sixth decade of life: A longitudinal study. *NeuroImage*, 220:116842.
- [Chen et al., 2015] Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., et al. (2015). Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4.
- [Cherubini et al., 2016] Cherubini, A., Caligiuri, M. E., Péran, P., Sabatini, U., Cosentino, C., and Amato, F. (2016). Importance of multimodal mri in characterizing brain tissue and its potential application for individual age prediction. *IEEE journal of biomedical and health informatics*, 20(5):1232–1239.
- [Cole et al., 2017] Cole, J. H., Poudel, R. P., Tsagkrasoulis, D., Caan, M. W., Steves, C., Spector, T. D., and Montana, G. (2017). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, 163:115–124.

- [Cullen and Halliday, 1998] Cullen, K. and Halliday, G. (1998). Neurofibrillary degeneration and cell loss in the nucleus basalis in comparison to cortical alzheimer pathology. *Neurobiology of aging*, 19(4):297–306.
- [Desikan et al., 2006] Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980.
- [Destrieux et al., 2010] Destrieux, C., Fischl, B., Dale, A., and Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, 53(1):1–15.
- [Domingues et al., 2018] Domingues, R., Filippone, M., Michiardi, P., and Zouaoui, J. (2018). A comparative evaluation of outlier detection algorithms: Experiments and analyses. *Pattern Recognition*, 74:406–421.
- [Dosenbach et al., 2010] Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., Nelson, S. M., Wig, G. S., Vogel, A. C., Lessov-Schlaggar, C. N., et al. (2010). Prediction of individual brain maturity using fmri. *Science*, 329(5997):1358–1361.
- [Duvernoy, 1999] Duvernoy, H. M. (1999). *The human brain: surface, three-dimensional sectional anatomy with MRI, and blood supply*. Springer Science & Business Media.
- [Eckerström et al., 2018] Eckerström, C., Klasson, N., Olsson, E., Selnes, P., Rolstad, S., and Wallin, A. (2018). Similar pattern of atrophy in early-and late-onset alzheimer’s disease. *Alzheimer’s & Dementia: Diagnosis, Assessment & Disease Monitoring*, 10:253–259.
- [Fischl, 2012] Fischl, B. (2012). Freesurfer. *Neuroimage*, 62(2):774–781.
- [Fischl et al., 1999] Fischl, B., Sereno, M. I., and Dale, A. M. (1999). Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2):195–207.
- [Fjell et al., 2015] Fjell, A. M., Grydeland, H., Krogsrud, S. K., Amlien, I., Rohani, D. A., Ferschmann, L., Storsve, A. B., Tamnes, C. K., Sala-Llonch, R., Due-Tønnessen, P., et al. (2015). Development and aging of cortical thickness correspond to genetic organization patterns. *Proceedings of the National Academy of Sciences*, 112(50):15462–15467.
- [Fleisher et al., 2005] Fleisher, A., Grundman, M., Jack, Clifford R., Petersen, R. C., Taylor, C., Kim, H. T., Schiller, D. H. B., Bagwell, V., Sencakova, D., Weiner, M. F., DeCarli, C., DeKosky, S. T., van Dyck, C. H., Thal, L. J., and

- Study, A. D. C. (2005). Sex, Apolipoprotein E e4 Status, and Hippocampal Volume in Mild Cognitive Impairment. *Archives of Neurology*, 62(6):953–957.
- [Franke et al., 2010] Franke, K., Ziegler, G., Klöppel, S., Gaser, C., Initiative, A. D. N., et al. (2010). Estimating the age of healthy subjects from t1-weighted mri scans using kernel methods: exploring the influence of various parameters. *Neuroimage*, 50(3):883–892.
- [Garthwaite, 1994] Garthwaite, P. H. (1994). An interpretation of partial least squares. *Journal of the American Statistical Association*, 89(425):122–127.
- [Gómez-Ramírez et al., 2020a] Gómez-Ramírez, J., Ávila-Villanueva, M., and Fernández-Blázquez, M. Á. (2020a). Selecting the most important self-assessed features for predicting conversion to mild cognitive impairment with random forest and permutation-based methods. *Scientific Reports*, 10(1):1–15.
- [Gómez-Ramírez et al., 2020b] Gómez-Ramírez, J., Fernandez-Blazquez, M. A., and González-Rosa, J. (2020b). The aging human brain: A causal analysis of the effect of sex and age on brain volume. *bioRxiv*.
- [Gómez-Ramírez and González-Rosa, 2021] Gómez-Ramírez, J. and González-Rosa, J. J. (2021). Intra-and interhemispheric symmetry of subcortical brain structures: a volumetric analysis in the aging human brain. *Brain Structure and Function*, pages 1–12.
- [Gomez-Ramirez et al., 2021] Gomez-Ramirez, J., Quilis-Sancho, J., and Fernandez-Blazquez, M. A. (2021). A comparative analysis of mri automated segmentation of subcortical brain volumes in a large dataset of elderly subjects. *Neuroinformatics*, pages 1–10.
- [Gomez-Ramirez and Wu, 2014] Gomez-Ramirez, J. and Wu, J. (2014). Network-based biomarkers in alzheimer’s disease: review and future directions. *Frontiers in aging neuroscience*, 6:12.
- [Grant et al., 1987] Grant, R., Condon, B., Lawrence, A., Hadley, D., Patterson, J., Bone, I., and Teasdale, G. (1987). Human cranial csf volumes measured by mri: sex and age influences. *Magnetic resonance imaging*, 5(6):465–468.
- [Granziera et al., 2021] Granziera, C., Wuerfel, J., Barkhof, F., Calabrese, M., De Stefano, N., Enzinger, C., Evangelou, N., Filippi, M., Geurts, J. J., Reich, D. S., et al. (2021). Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. *Brain*, 144(5):1296–1311.
- [Habeck et al., 2020] Habeck, C., Gazes, Y., Razlighi, Q., and Stern, Y. (2020). Cortical thickness and its associations with age, total cognition and education across the adult lifespan. *Plos one*, 15(3):e0230298.

- [Harisinghani et al., 2019] Harisinghani, M. G., O’Shea, A., and Weissleder, R. (2019). Advances in clinical mri technology. *Science translational medicine*, 11(523).
- [Horvath et al., 2012] Horvath, S., Zhang, Y., Langfelder, P., Kahn, R. S., Boks, M. P., van Eijk, K., van den Berg, L. H., and Ophoff, R. A. (2012). Aging effects on dna methylation modules in human brain and blood tissue. *Genome biology*, 13(10):1–18.
- [Kaye et al., 1997] Kaye, J. A., Swihart, T., Howieson, D., Dame, A., Moore, M., Karnos, T., Camicioli, R., Ball, M., Oken, B., and Sexton, G. (1997). Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia. *Neurology*, 48(5):1297–1304.
- [Kondo et al., 2015] Kondo, C., Ito, K., Wu, K., Sato, K., Taki, Y., Fukuda, H., and Aoki, T. (2015). An age estimation method using brain local features for t1-weighted images. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 666–669. IEEE.
- [Leonard et al., 2008] Leonard, C. M., Towler, S., Welcome, S., Halderman, L. K., Otto, R., Eckert, M. A., and Chiarello, C. (2008). Size matters: cerebral volume influences sex differences in neuroanatomy. *Cerebral cortex*, 18(12):2920–2931.
- [Lotze et al., 2019] Lotze, M., Domin, M., Gerlach, F. H., Gaser, C., Lueders, E., Schmidt, C. O., and Neumann, N. (2019). Novel findings from 2,838 adult brains on sex differences in gray matter brain volume. *Scientific reports*, 9(1):1–7.
- [MacDonald and Pike, 2021] MacDonald, M. E. and Pike, G. B. (2021). Mri of healthy brain aging: A review. *NMR in Biomedicine*, page e4564.
- [MacDonald et al., 2019] MacDonald, M. E., Williams, R. J., Forkert, N. D., Berman, A. J., McCreary, C. R., Frayne, R., and Pike, G. B. (2019). Inter-database variability in cortical thickness measurements. *Cerebral Cortex*, 29(8):3282–3293.
- [Madan and Kensinger, 2016] Madan, C. R. and Kensinger, E. A. (2016). Cortical complexity as a measure of age-related brain atrophy. *NeuroImage*, 134:617–629.
- [Madan and xKensinger, 2018] Madan, C. R. and xKensinger, E. A. (2018). Predicting age from cortical structure across the lifespan. *European Journal of Neuroscience*, 47(5):399–416.
- [Maruszak and Thuret, 2014] Maruszak, A. and Thuret, S. (2014). Why looking at the whole hippocampus is not enough—a critical role for anteroposterior axis, subfield and activation analyses to enhance predictive value of

- hippocampal changes for alzheimer's disease diagnosis. *Frontiers in cellular neuroscience*, 8:95.
- [McGinnis et al., 2011] McGinnis, S. M., Brickhouse, M., Pascual, B., and Dickerson, B. C. (2011). Age-related changes in the thickness of cortical zones in humans. *Brain topography*, 24(3):279–291.
- [McRae-McKee et al., 2019] McRae-McKee, K., Evans, S., Hadjichrysanthou, C., Wong, M., De Wolf, F., and Anderson, R. (2019). Combining hippocampal volume metrics to better understand alzheimer's disease progression in at-risk individuals. *Scientific reports*, 9(1):1–9.
- [Mole et al., 2020] Mole, J. P., Fasano, F., Evans, J., Sims, R., Kidd, E., Aggleton, J. P., and Metzler-Baddeley, C. (2020). Apoe-ε4-related differences in left thalamic microstructure in cognitively healthy adults. *Scientific reports*, 10(1):1–25.
- [Morcom and Henson, 2018] Morcom, A. M. and Henson, R. N. (2018). Increased prefrontal activity with aging reflects nonspecific neural responses rather than compensation. *Journal of Neuroscience*, 38(33):7303–7313.
- [Mwangi et al., 2013] Mwangi, B., Hasan, K. M., and Soares, J. C. (2013). Prediction of individual subject's age across the human lifespan using diffusion tensor imaging: a machine learning approach. *Neuroimage*, 75:58–67.
- [Niu et al., 2020] Niu, X., Zhang, F., Kounios, J., and Liang, H. (2020). Improved prediction of brain age using multimodal neuroimaging data. *Human brain mapping*, 41(6):1626–1643.
- [Nobis et al., 2019] Nobis, L., Manohar, S. G., Smith, S. M., Alfaro-Almagro, F., Jenkinson, M., Mackay, C. E., and Husain, M. (2019). Hippocampal volume across age: Nomograms derived from over 19,700 people in uk biobank. *NeuroImage: Clinical*, 23:101904.
- [Orihashi et al., 2020] Orihashi, R., Mizoguchi, Y., Imamura, Y., Yamada, S., Ueno, T., and Monji, A. (2020). Oxytocin and elderly mri-based hippocampus and amygdala volume: a 7-year follow-up study. *Brain communications*, 2(2):fcaa081.
- [Pardoe and Kuzniecky, 2018] Pardoe, H. R. and Kuzniecky, R. (2018). Napr: a cloud-based framework for neuroanatomical age prediction. *Neuroinformatics*, 16(1):43–49.
- [Pedregosa et al., 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

- [Prakkamakul et al., 2016] Prakkamakul, S., Witzel, T., Huang, S., Boulter, D., Borja, M. J., Schaefer, P., Rosen, B., Heberlein, K., Ratai, E., Gonzalez, G., et al. (2016). Ultrafast brain mri: clinical deployment and comparison to conventional brain mri at 3t. *Journal of Neuroimaging*, 26(5):503–510.
- [Prins and Scheltens, 2015] Prins, N. D. and Scheltens, P. (2015). White matter hyperintensities, cognitive impairment and dementia: an update. *Nature Reviews Neurology*, 11(3):157–165.
- [Ritzberger et al., 2002] Ritzberger, K. et al. (2002). Foundations of non-cooperative game theory. *OUP Catalogue*.
- [Royle et al., 2013] Royle, N. A., Booth, T., Hernández, M. C. V., Penke, L., Murray, C., Gow, A. J., Maniega, S. M., Starr, J., Bastin, M. E., Deary, I. J., et al. (2013). Estimated maximal and current brain volume predict cognitive ability in old age. *Neurobiology of Aging*, 34(12):2726–2733.
- [Sakai, 2020] Sakai, J. (2020). Core concept: How synaptic pruning shapes neural wiring during development and, possibly, in disease. *Proceedings of the National Academy of Sciences*, 117(28):16096–16099.
- [Salat et al., 2004] Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., Morris, J. C., Dale, A. M., and Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral cortex*, 14(7):721–730.
- [Sayed et al., 2021] Sayed, N., Huang, Y., Nguyen, K., Krejciova-Rajaniemi, Z., Grawe, A. P., Gao, T., Tibshirani, R., Hastie, T., Alpert, A., Cui, L., et al. (2021). An inflammatory aging clock (iage) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging. *Nature Aging*, pages 1–18.
- [Schnack et al., 2016] Schnack, H. G., Van Haren, N. E., Nieuwenhuis, M., Hulshoff Pol, H. E., Cahn, W., and Kahn, R. S. (2016). Accelerated brain aging in schizophrenia: a longitudinal pattern recognition study. *American Journal of Psychiatry*, 173(6):607–616.
- [Schultz et al., 2020] Schultz, M. B., Kane, A. E., Mitchell, S. J., MacArthur, M. R., Warner, E., Vogel, D. S., Mitchell, J. R., Howlett, S. E., Bonkowski, M. S., and Sinclair, D. A. (2020). Age and life expectancy clocks based on machine learning analysis of mouse frailty. *Nature communications*, 11(1):1–12.
- [Serbruyns et al., 2015] Serbruyns, L., Leunissen, I., Huysmans, T., Cuypers, K., Meesen, R. L., van Ruitenbeek, P., Sijbers, J., and Swinnen, S. P. (2015). Subcortical volumetric changes across the adult lifespan: subregional thalamic atrophy accounts for age-related sensorimotor performance declines. *Cortex*, 65:128–138.

- [Shams et al., 2015] Shams, S., Martola, J., Cavallin, L., Granberg, T., Shams, M., Aspelin, P., Wahlund, L., and Kristoffersen-Wiberg, M. (2015). Swi or t2*: which mri sequence to use in the detection of cerebral microbleeds? the karolinska imaging dementia study. *American Journal of Neuroradiology*, 36(6):1089–1095.
- [Shapley, 1953] Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28):307–317.
- [Shaw et al., 2016] Shaw, M. E., Sachdev, P. S., Anstey, K. J., and Cherbuin, N. (2016). Age-related cortical thinning in cognitively healthy individuals in their 60s: the path through life study. *Neurobiology of aging*, 39:202–209.
- [Smith et al., 2019] Smith, S. M., Vidaurre, D., Alfaro-Almagro, F., Nichols, T. E., and Miller, K. L. (2019). Estimation of brain age delta from brain imaging. *Neuroimage*, 200:528–539.
- [Sowell et al., 2003] Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., and Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature neuroscience*, 6(3):309.
- [Tripathi et al., 2020] Tripathi, S., Hemachandra, N., and Trivedi, P. (2020). Interpretable feature subset selection: A shapley value based approach. In *2020 IEEE International Conference on Big Data (Big Data)*, pages 5463–5472. IEEE.
- [Van Essen, 2005] Van Essen, D. C. (2005). A population-average, landmark- and surface-based (pals) atlas of human cerebral cortex. *Neuroimage*, 28(3):635–662.
- [Wang et al., 2014] Wang, J., Li, W., Miao, W., Dai, D., Hua, J., and He, H. (2014). Age estimation using cortical surface pattern combining thickness with curvatures. *Medical & biological engineering & computing*, 52(4):331–341.
- [Whitwell et al., 2012] Whitwell, J. L., Wiste, H. J., Weigand, S. D., Rocca, W. A., Knopman, D. S., Roberts, R. O., Boeve, B. F., Petersen, R. C., Jack, C. R., Initiative, A. D. N., et al. (2012). Comparison of imaging biomarkers in the alzheimer disease neuroimaging initiative and the mayo clinic study of aging. *Archives of neurology*, 69(5):614–622.
- [Wolf et al., 2004] Wolf, H., Julin, P., Gertz, H.-J., Winblad, B., and Wahlund, L.-O. (2004). Intracranial volume in mild cognitive impairment, alzheimer’s disease and vascular dementia: evidence for brain reserve? *International journal of geriatric psychiatry*, 19(10):995–1007.