Methods

Our data includes cortical, subcortical, and DTI measurements totaling 66 features for 76 instances in the control group. All data was scaled feature-wise using standard deviations from the mean. Software used includes Python 3.6.8 64bit, scikit-learn, numPy, and matplotlib. Eight models were tested with 5 sets of input features. 5-fold ross validation with random sampling from a shuffled dataset is used for model assessment. This cross validation was done 100 times on each model, with seeded random splits such that each model saw the same splits during training and fitting. For each cross validation, an average score was taken from the 5 configurations. Then, those averages were averaged after 100 iterations.

Results

Bayesian ridge model performed best using the full feature vector, averaging =0.609 over 100 iterations of 5-fold cross validation, showing the best ability to generalize. The final Bayesian ridge model has =0.883. Other notable models using the full feature vector include SVM with RBF kernel (=0.536), ElasticNet (=0.593), and ridge (=0.516).

Subsets of the feature vector were also experimented with. These subsets include the cortical feature vector, subcortical feature vector, DTI feature vector, the top 5 variables linearly correlated with brain age, and the top 20 variables linearly correlated with age. No other subset of data performed as well as the full feature vector with a Bayesian ridge regressor model.

Discussion

Ridge regression vastly outperformed ordinary least squares regression during cross validation, showing that there may be collinearity in the data. This can be expected, as the features represent brain regions which are close spatially. A reading for one area may be influenced by those around it. Ridge regression also shifts the coefficient weights toward 0, which can help stabilize them.

The Bayesian ridge inference model working better than ridge alone implies that there may be priors associated with the input data, which may be induced from the equipment that performed the brain scans to take this data, or the brain’s activity itself. Non-informed priors are used in this implementation of Bayesian ridge regression, assuming a Gaussian distribution around the feature vectors. The Bayesian inference model may also be performing better due to the smaller sample size, allowing for the priors to have significant impact on the resulting predicted age.

|  |  |  |  |
| --- | --- | --- | --- |
| **Feature** | **Linear** | **deg=2** | **deg=3** |
| superiorfrontal | 0.191806974 | 0.190957297 | 0.153828254 |
| GCC | 0.168573645 | 0.185481251 | 0.143588608 |
| parsopercularis | 0.162543893 | 0.148481573 | 0.107115252 |
| medialorbitofrontal | 0.156666067 | 0.144221257 | -0.102068301 |
| Thalamus | 0.151453842 | 0.154285246 | 0.126686093 |
| superiortemporal | 0.138638063 | 0.12829608 | 0.112861574 |
| rostralanteriorcingulate | 0.122889494 | 0.104477196 | 0.062824869 |
| CC | 0.109873238 | 0.11081502 | 0.085506722 |
| BCC | 0.105178546 | 0.099159928 | 0.059506878 |
| Left-Accumbens-area | 0.101227941 | 0.100850562 | -0.441981323 |
| FX | 0.084414875 | 0.075922113 | 0.042183436 |
| caudalmiddlefrontal | 0.078756231 | 0.033309392 | -0.097988119 |
| insula | 0.071186958 | 0.070307634 | 0.058560935 |
| supramarginal | 0.067991748 | 0.051994474 | -0.055696737 |
| frontalpole | 0.062361593 | 0.038629688 | -0.01095099 |
| rostralmiddlefrontal | 0.060488917 | 0.04040489 | -0.046457459 |
| parstriangularis | 0.059285932 | 0.062792963 | 0.027704727 |
| bankssts | 0.052170539 | 0.058758916 | -0.028086754 |
| CR | 0.022410172 | 0.01812194 | 0.010100561 |
| lateralorbitofrontal | 0.015881764 | -0.037889873 | -0.092436024 |

Figure: Table of the top 20 variables linearly related with age

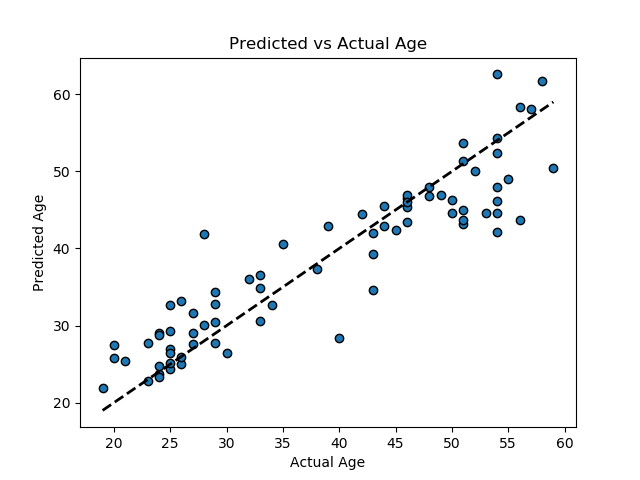


Figure: Predicted vs Actual Age for a Bayesian Ridge Regression model trained on all 76 samples using all 66 feature vectors (=0.883)

|  |  |
| --- | --- |
| Model |  |
| Bayesian Ridge | 0.609 |
| SVM - RBF | 0.536 |
| Linear Regression | -3.647 |
| Decision Tree | -0.37 |
| Ridge | 0.516 |
| Lasso | 0.383 |
| ElasticNet | 0.593 |
| SGD | 0.066 |

Figure: Model scores averaged over 100 iterations of 5-fold cross validation on the full feature vector

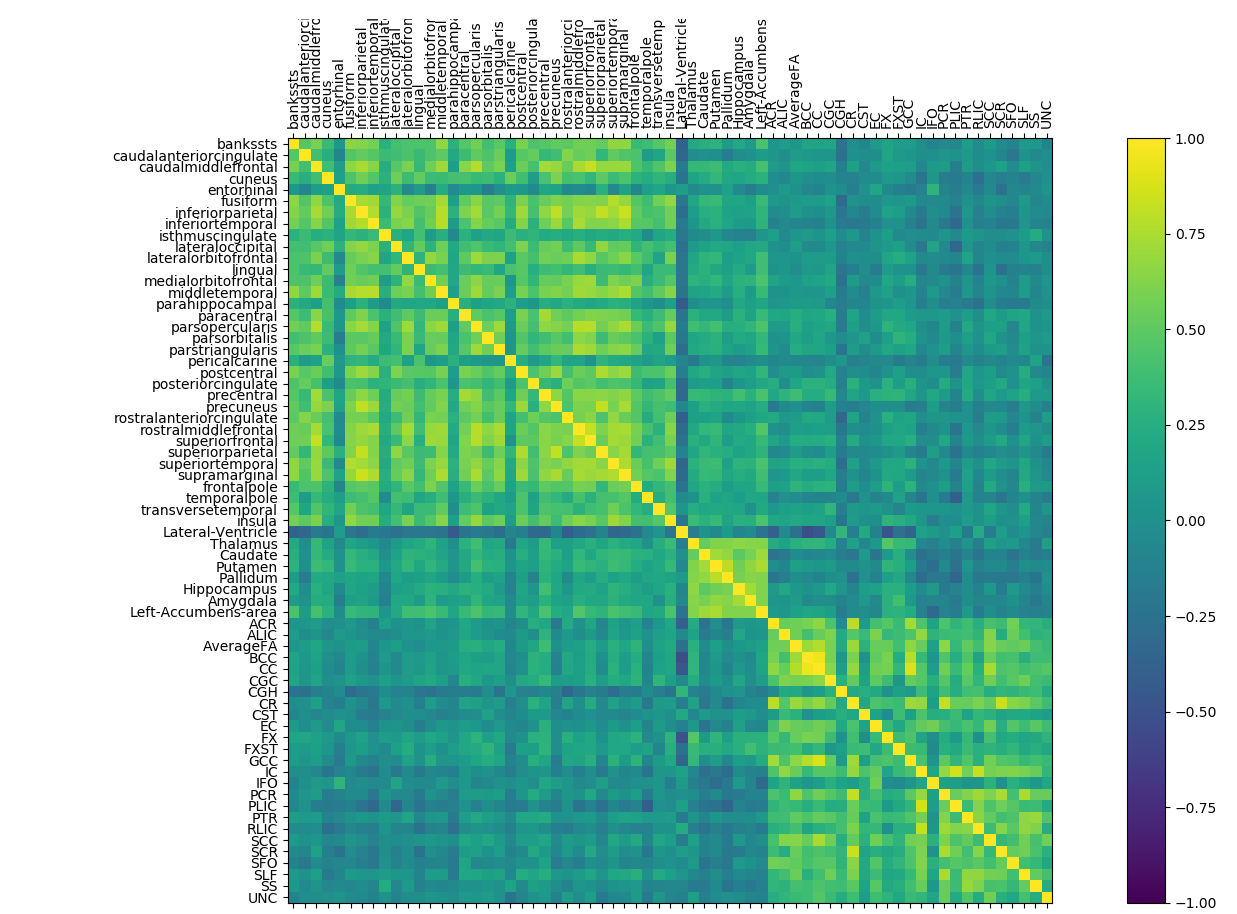


Figure: Covariance matrix of the 66 input features from the DTI, subcortical, and cortical datasets