

## Code explanation

The code above performs data analysis on a dataset of pediatric patients who received post-operative mechanical ventilation after surgery. The goal of the analysis is to predict the optimal depth of the tracheal tube for these patients.

The code begins by importing necessary libraries for data analysis, including pandas, numpy, and scikit-learn. It then loads the dataset from a CSV file.

Next, the code performs some initial data exploration, specifically computing the mean height and age stratified by sex. The results are saved in a DataFrame and stored in a pickle file.

After the data exploration, the code proceeds to the analysis phase. It prepares the data for model fitting by selecting the relevant features and the target variable. It then splits the data into training and testing sets with a ratio of 70% training and 30% testing.

The code then fits several machine learning models to the training data, including RandomForest, ElasticNet, SVR, and MLPRegressor. For each model, it predicts the target variable for the testing data and calculates the mean squared error (MSE) as a measure of prediction accuracy.

The results of the MSE calculations for each model are saved in a DataFrame and stored in a pickle file. This table provides a comparison of the different machine learning models in predicting the optimal tracheal tube depth.

Additionally, the code saves some additional results in a separate pickle file. These results include the total number of observations in the dataset and the accuracy scores of each model on the testing data.

Overall, the code performs data exploration, fits machine learning models to predict the optimal tracheal tube depth, and saves the results for further analysis and comparison.