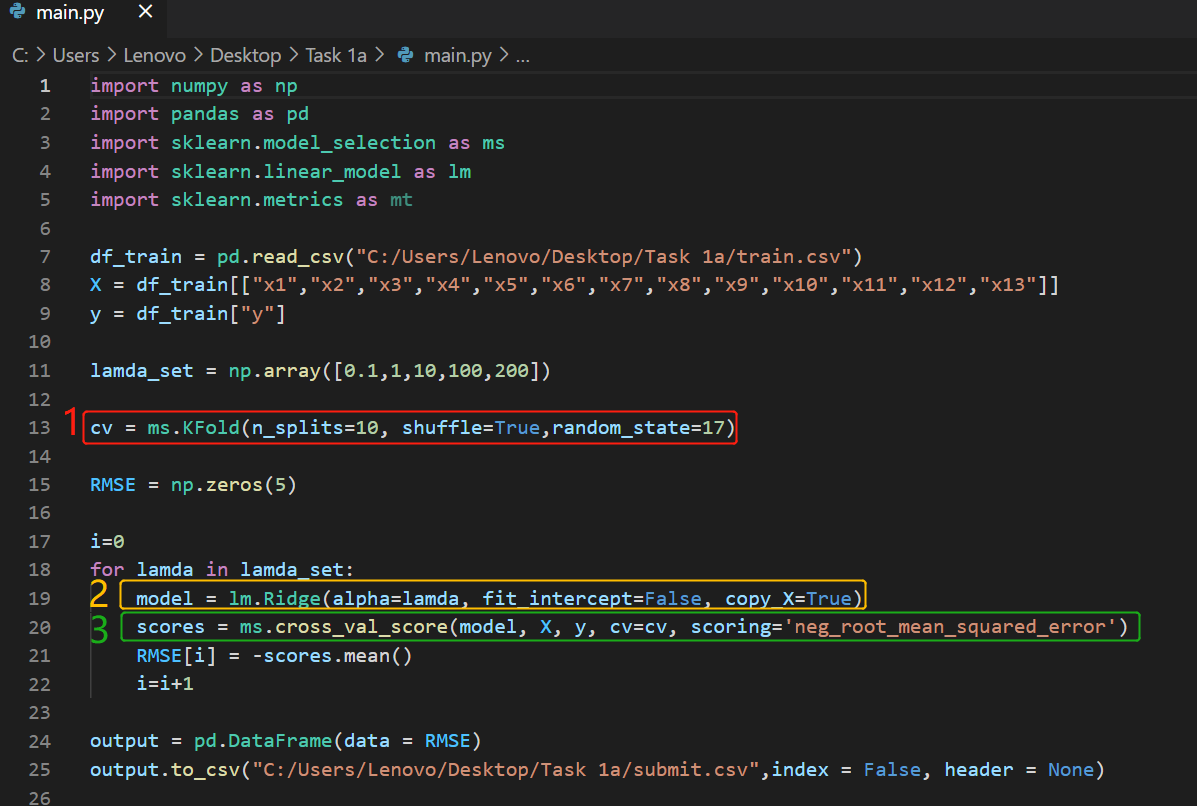
Approach Description for Task 1a

Our code is as followed:



Our main approach is as followed:

1. We use K-Fold cross validation. To this end, we use *KFold* function from *sklearn.model\_selection*, in which we specify 3 arguments: *n\_splits* , *shuffle* and *random\_state*. We set *n\_splits* to 10, which means we have 10 folds. We set *shuffle* to Trueso that this function would randomly pick the datapoints to be training set and test set (The points being chosen to be training set/test set will not be in sequentially in order). We set *random\_state* to a constant, so that every time we run the code, the splits of folds will remain the same (the result can be reproduced).
2. We use ridge regression. The *alpha* is the regularization strength and since our function has no intercept, we don’t need to fit this for our model.
3. Finally, we specify our model, the data to fit, the target variable, our cross-validation splitting strategy and the type of error we want to compute (RMSE) for each fold in the *cross\_val\_score* function.