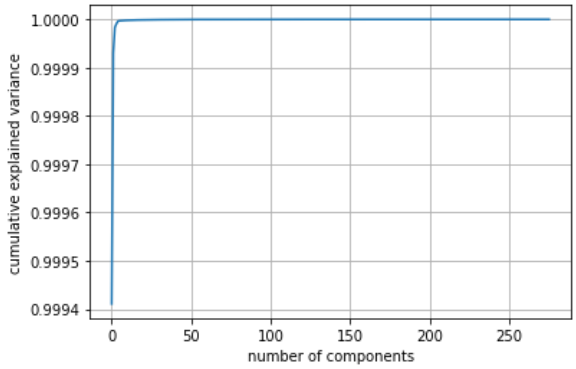
Preparing data

Our data set consist 277 features, we apply PCA on our dataset in order to reduce data. By plotting

"cumulative explained variance" vs "number of components" we find optimum number of component in PCA method.(5 components For our model)



SVR

SVR as a traditional regression model has different kernels for training data, We choose rbf and polynomial kernel from them.

In Svr we have 3 different parameters, epsilon, C as a regularization parameter (for our model C=0.1 has better performance) and gamma that we choose it as a hyperparameter , by plotting validation curve for different range of gamma we can see variation of validation rate and learning rate vs this hyperparameter.

We train our model and calculate loss with mean square error , mean absolute error and r2 square as a metric, By comparing between learning curve and validation curve for this metrics ,we can conclude that with r2 score we have better learning curve and validation curve rather than other metrics.