CS 771: Introduction To Machine Learning
Assignment 3

Akshat Agarwal
200081
akshatag20@iitk.ac.in

Muskan Kumari
200610
muskank20@iitk.ac.in

Navya Ratnan
200627
nratnan20@iitk.ac.in

Sejal Sahu
200911
sejals20@iitk.ac.in

Yash Goel
201142
yashgoel20@iitk.ac.in

Answer 1

The linear models used to predict the O3 and NO2 using the method suggested by the manufacturer are:

a) Linear Regression
Linear regression is used to model the relationship between a dependent variable (also known as the outcome or response variable) and one or more independent variables (also known as predictor or explanatory variables) by fitting a linear equation to the observed data. The goal of linear regression is to find the line of best fit that minimizes the distance between the predicted values and the actual values of the dependent variable. This line can be used to make predictions about future values of the dependent variable based on the values of the independent variable(s). The variable you wish to forecast is referred to as the dependent variable. The variable you are using to forecast the value of the other variable is known as the independent variable. It is a popular technique for predictive analysis and modelling.

b) Ridge Regression
Ridge regression is a type of regularization technique used in linear regression models to address the problem of multicollinearity. Multicollinearity occurs when the independent variables in a regression model are highly correlated with each other, leading to unstable and unreliable estimates of the regression coefficients. Ridge regression adds a penalty term to the objective function of the linear regression model, which limits the magnitude of the coefficients and shrinks them towards zero. This penalty term is controlled by a hyperparameter called the regularization parameter, which can be tuned using cross-validation. By reducing the variance of the estimates, ridge regression can improve the predictive accuracy of the model and prevent overfitting. Ridge regression is commonly used in situations where there are many independent variables in the model.

Best result:
$model\_O3 = Ridge(alpha = 0.01, tol = 9e - 2)$
$model\_NO2 = Ridge(alpha = 0.2,\ tol = 1e - 2)$

c) Lasso Regression
Lasso regression is another type of regularization technique used in linear regression models. Like ridge regression, lasso regression also addresses the problem of multicollinearity by adding a penalty term to the objective function of the linear regression model. However, unlike ridge regression, lasso regression adds a penalty term proportional to the absolute value of the coefficients, rather

than the square of the coefficients. This penalty term can result in some of the coefficients being exactly zero, effectively performing variable selection and eliminating some of the independent variables from the model. The amount of shrinkage is controlled by a hyperparameter called the regularization parameter, which can be tuned using cross-validation. Lasso regression is commonly used in situations where there are many independent variables in the model and some of them may be irrelevant or redundant.

Best result:
$model\_O3 = Lasso(alpha = 0.01, tol = 9e - 2, random\_state = 0)$
$model\_NO2 = Lasso(alpha = 0.2, tol = 4e - 2,\ random\_state = 1)$

d) Elastic Net Regression
Elastic net regression is a hybrid regularization technique that combines the strengths of both ridge regression and lasso regression. It adds a penalty term to the objective function of the linear regression model that is a combination of the L1 penalty (used in lasso regression) and the L2 penalty (used in ridge regression). The elastic net penalty allows for some coefficients to be exactly zero, like in lasso regression, and also shrinks the coefficients towards zero, like in ridge regression. The amount of shrinkage and sparsity is controlled by two hyperparameters: the regularization parameter and the mixing parameter that determines the ratio of L1 and L2 penalties. Elastic net regression is commonly used in situations where there are many independent variables in the model, and some of them may be highly correlated, or when there is a need to perform variable selection while also controlling for overfitting.

Best result:
$model\_O3 = ElasticNet(alpha = 0.01, tol = 9e - 2, random\_state = 0)$
$model\_NO2 = ElasticNet(alpha = 0.2, tol = 1e - 2,\ random\_state = 1)$

Observations

Method	**Time Taken (s)**	O_3 **MAE**	NO_2 **MAE**	**Model Size**
Linear Regression	0.00482	6.47355	5.75448	1230
Ridge Regression	0.00564	6.47355	5.75448	1206
Lasso Regression	0.01168	6.22366	5.72295	1409
Elastic Net Regression	0.00637	6.22368	5.72306	1419

Answer 2

Decision tree for Regression
Decision tree builds regression or classification models in the form of a tree structure. It incrementally divides a dataset into smaller and smaller sections while also developing an associated decision tree. The end result is a tree containing leaf nodes and decision nodes. A decision node has two or more branches, each of which represents a value for the attribute tested. A leaf node (for example, Hours Played) represents a numerical target decision. The root node is the topmost decision node in a tree that corresponds to the best predictor. Both category and numerical data can be handled by decision trees.

$model\_O3 = tree.DecisionTreeRegressor(criterion =' absolute\_error', max\_depth = 20, random\_state = 0)$
$model\_NO2 = tree.DecisionTreeRegressor(criterion =' absolute\_error', max\_depth = 20, random\_state = 1)$

Final Values are:

Time Taken (in seconds): 0.03315672499998072
O_3 MAE: 0.3788073
NO_2 MAE: 0.3782636
Model Size: 3347492