Model Codes

# Multiple Linear Regression **without** splitting the data as per Ft##

fit\_reg <- lm(e ~ ., data=f1)

summary(fit\_reg) # show results

coefficients(fit\_reg) # model coefficients

confint(fit\_reg, level=0.95) # CIs for model parameters

fitted(fit\_reg) # predicted values

residuals(fit\_reg) # residuals

anova(fit\_reg) # anova table

vcov(fit\_reg) # covariance matrix for model parameters

influence(fit\_reg) # regression diagnostics

p <- predict(fit\_reg, f1[,2:4])

p[1:10]

# summarize results

#Root mean square error##

rmse1 <- mean((f1$e - p)^2)

print(rmse1)

library(caret)

library(klaR)

confusionMatrix(p,y\_test)

# K-fold cross-validation

library(DAAG)

cv.lm(f1, fit\_reg, m=5)

# Multiple Linear Regression with splitting the data as per Ft##

k2 <- read.csv("k1.csv")

Obs\_pps <- cbind(k2$Ft, k2$Mk, k2$MS, f1)

reg\_spt <- cbind(k2$Ft, f1)

write.csv(reg\_spt, "reg\_data.csv")

rg <- read.csv("reg\_data\_1.csv")

df = data.frame(rg)

install.packages("cv.glm")

library(sqldf)

install.packages("tcltk")

petrol <- sqldf("SELECT \* FROM rg WHERE Ft = 'petrol'")

## Model for Petrol ##

fit\_petrol <- lm(e ~ car\_at + car\_engine + car\_body, data=petrol)

summary(fit\_petrol)

p1 <- predict(fit\_petrol, petrol[,3:5])

p1[1:10]

# summarize results

#Root mean square error##

rmse\_petrol <- mean((petrol$e - p1)^2)

print(rmse\_petrol)

#cv.lm(petrol, fit\_petrol, m=3)

diesel <- sqldf("SELECT \* FROM rg WHERE Ft = 'diesel'")

fit\_diesel <- lm(e ~ car\_at + car\_engine + car\_body, data=diesel)

summary(fit\_diesel)

p2 <- predict(fit\_diesel, diesel[,3:5])

p2[1:10]

# summarize results

#Root mean square error##

rmse\_diesel <- mean((diesel$e - p2)^2)

print(rmse\_diesel)

levels(k2$Ft)

biodiesel <- sqldf("SELECT \* FROM rg WHERE Ft = 'biodiesel'")

fit\_biodiesel <- lm(e ~ car\_at + car\_engine + car\_body, data=biodiesel)

summary(fit\_biodiesel)

p3 <- predict(fit\_biodiesel, biodiesel[,3:5])

p3[1:10]

# summarize results

#Root mean square error##

rmse\_biodiesel <- mean((biodiesel$e - p3)^2)

print(rmse\_biodiesel)

e85 <- sqldf("SELECT \* FROM rg WHERE Ft = 'e85'")

fit\_e85 <- lm(e ~ car\_at + car\_engine + car\_body, data=e85)

summary(fit\_e85)

p4 <- predict(fit\_e85, e85[,3:5])

p4[1:10]

# summarize results

#Root mean square error##

rmse\_e85 <- mean((e85$e - p4)^2)

print(rmse\_e85)

hydrogen <- sqldf("SELECT \* FROM rg WHERE Ft = 'hydrogen'")

fit\_hydrogen <- lm(e ~ car\_at + car\_engine + car\_body, data=hydrogen)

summary(fit\_hydrogen)

p5 <- predict(fit\_hydrogen, hydrogen[,3:5])

p5[1:10]

# summarize results

#Root mean square error##

rmse\_hydrogen <- mean((hydrogen$e - p5)^2)

print(rmse\_hydrogen)

lpg <- sqldf("SELECT \* FROM rg WHERE Ft = 'lpg'")

fit\_lpg <- lm(e ~ car\_at + car\_engine + car\_body, data=lpg)

summary(fit\_lpg)

p5 <- predict(fit\_lpg, lpg[,3:5])

p5[1:10]

# summarize results

#Root mean square error##

rmse\_lpg <- mean((lpg$e - p5)^2)

print(rmse\_lpg)

ng\_biomethane <- sqldf("SELECT \* FROM rg WHERE Ft = 'ng biomethane'")

fit\_ng\_biomethane <- lm(e ~ car\_at + car\_engine + car\_body, data=ng\_biomethane)

summary(fit\_ng\_biomethane)

p6 <- predict(fit\_ng\_biomethane, ng\_biomethane[,3:5])

p6[1:10]

# summarize results

#Root mean square error##

rmse\_ng\_biomethane <- mean((ng\_biomethane$e - p6)^2)

print(rmse\_ng\_biomethane)

petrolgas <- sqldf("SELECT \* FROM rg WHERE Ft = 'petrol gas'")

fit\_petrolgas <- lm(e ~ car\_at + car\_engine + car\_body, data=petrolgas)

summary(fit\_petrolgas)

p7 <- predict(fit\_petrolgas, petrolgas[,3:5])

p7[1:10]

# summarize results

#Root mean square error##

rmse\_petrolgas <- mean((petrolgas$e - p7)^2)

print(rmse\_petrolgas)