

P8106_Mtp

Ekta Chaudhary

26/03/2020

```
library(tidyverse)
library(readxl)
library(caret)
library(ModelMetrics)
library(glmnet)
library(gam)
library(mgcv)
library(splines)
library(pdp)
library(earth)
library(dplyr)
library(naniar)
library(bnstruct)
```

Reading the Datasets

```
Diagnosis =
read_csv('./data/Diagnosis.csv')
Prognosis =
  read_csv('./data/Prognosis.csv')
Breast_Cancer =
  read_csv('./data/Breast_Cancer.csv') #,col_names = ['',])
```

```
Breast_Cancer = Breast_Cancer %>% replace_with_na_all(condition = ~.x == '?')
```

```
Breast_Cancer <- knn.impute(as.matrix(Breast_Cancer), k = 10, cat.var = 2:ncol(Breast_Cancer) - 2,
  to.impute = 1:nrow(Breast_Cancer), using = 1:nrow(Breast_Cancer))
```