

Algorithm

Anandu R

5/31/2020

Importing necessary packages/Libraries

```
invisible(library(dplyr))
invisible(library(lubridate))
invisible(library(caTools))
invisible(library(data.table))
```

Generating the dataset

```
set.seed(1)
speed = round(rnorm(1000,50,15),2)
dist_prev = abs(round(rnorm(1000,2,1),2))
dist_next = abs(round(rnorm(1000,2,1),2))
crowd = rpois(1000,25)
booked = rpois(1000,40)
schd_time = sample(seq(strptime('01/01/2018',format = "%d/%m/%Y"),
                        strptime('01/01/2019',format = "%d/%m/%Y"),
                        by="hour"), 1000, replace = T)
arr_time = schd_time+(rnorm(1000,300,350)*-1)
on_time = ifelse(difftime(arr_time,schd_time)<=0,1,0)
data = data.frame(crowd,booked,dist_prev,dist_next,speed,schd_time,arr_time,on_time)
head(select(data,crowd,booked,on_time))
```

```
##   crowd booked on_time
## 1    28     43        0
## 2    26     39        1
## 3    31     35        1
## 4    20     43        1
## 5    27     35        0
## 6    23     36        0
```

Generating an algorithm to label the datasets

Each record is considered as a bus and the label is the indication given to the bus driver whether to maintain speed, decrease speed, or to increase represented by 0,1,2 respectively