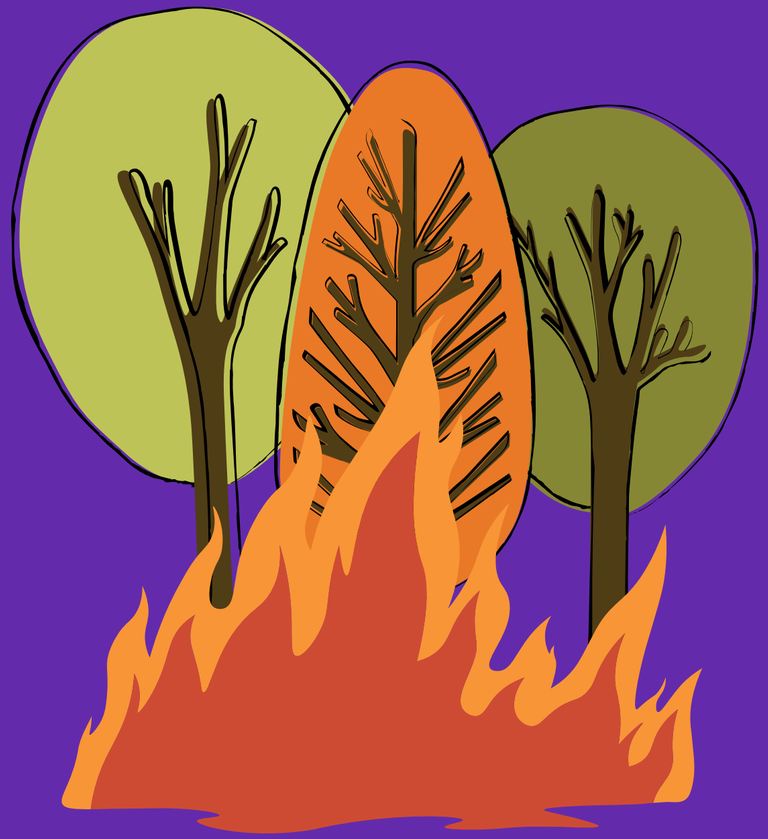
WILDFIRES

Practical Assignment for Data Mining I



By: Sofia Malpique & Yannik Bauer

DATA PREPARATION

extinction_date
extinction_hour

region

firstInterv_date
firstInterv_hour

lat
lon

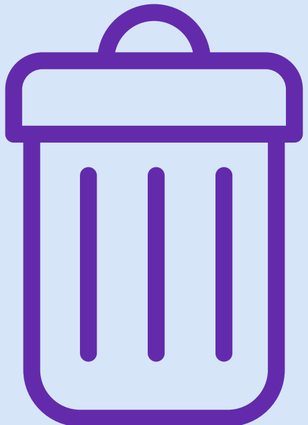
village_veget_area

municipality

alert_source

total_area

parish

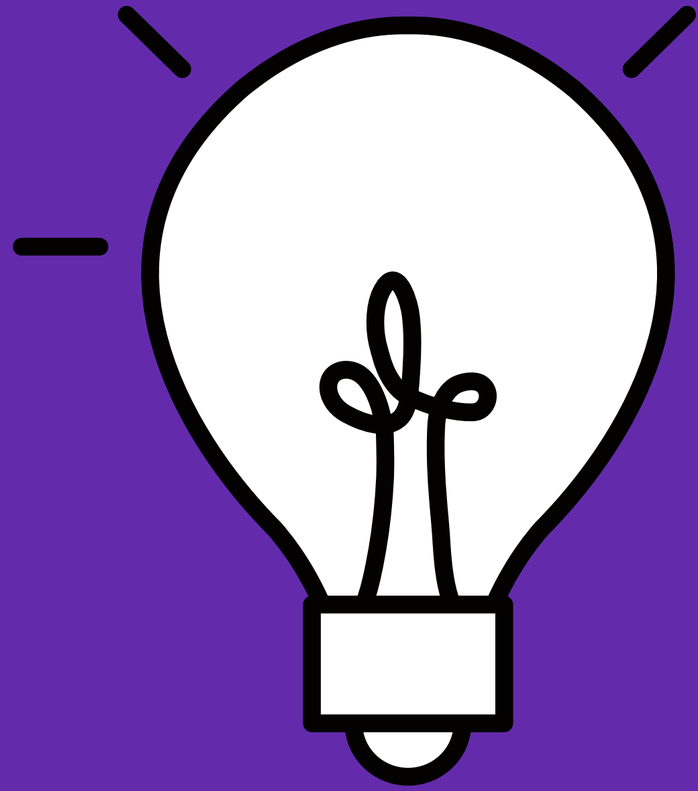


DATA PREPARATION

tmax

Season

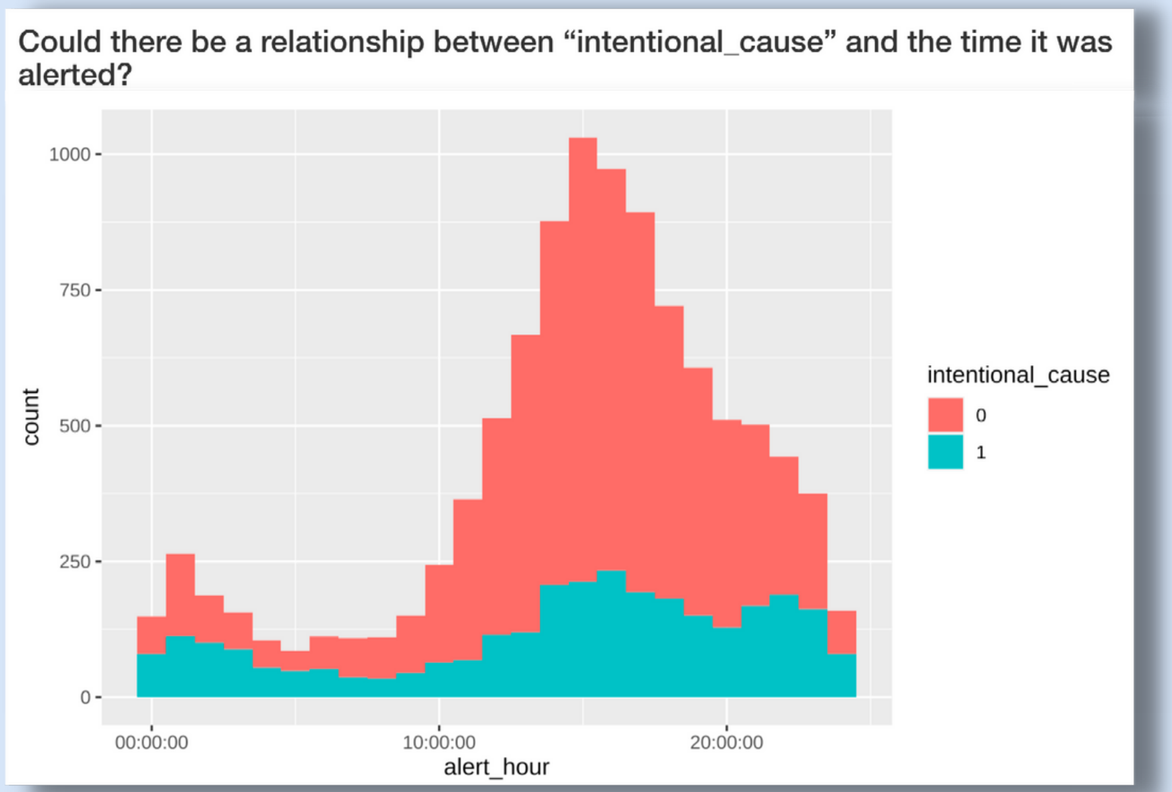
inLitoral

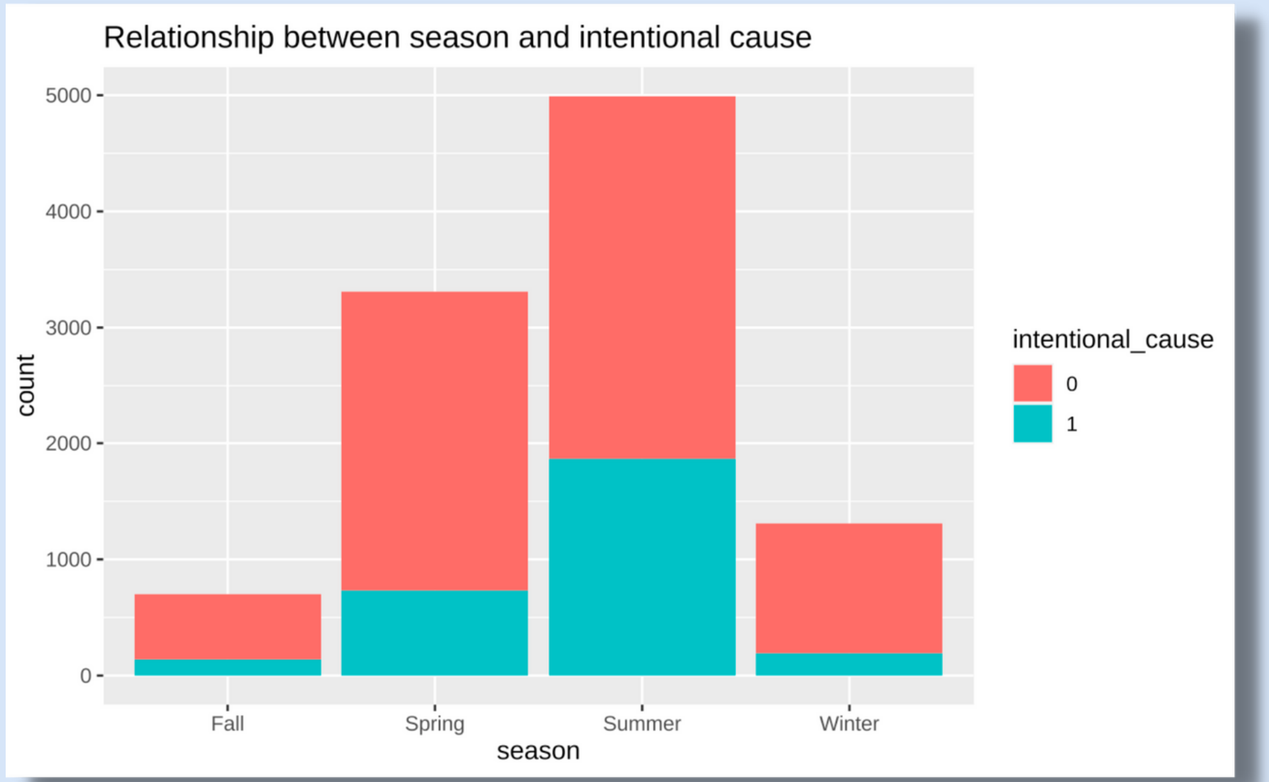


Weekday

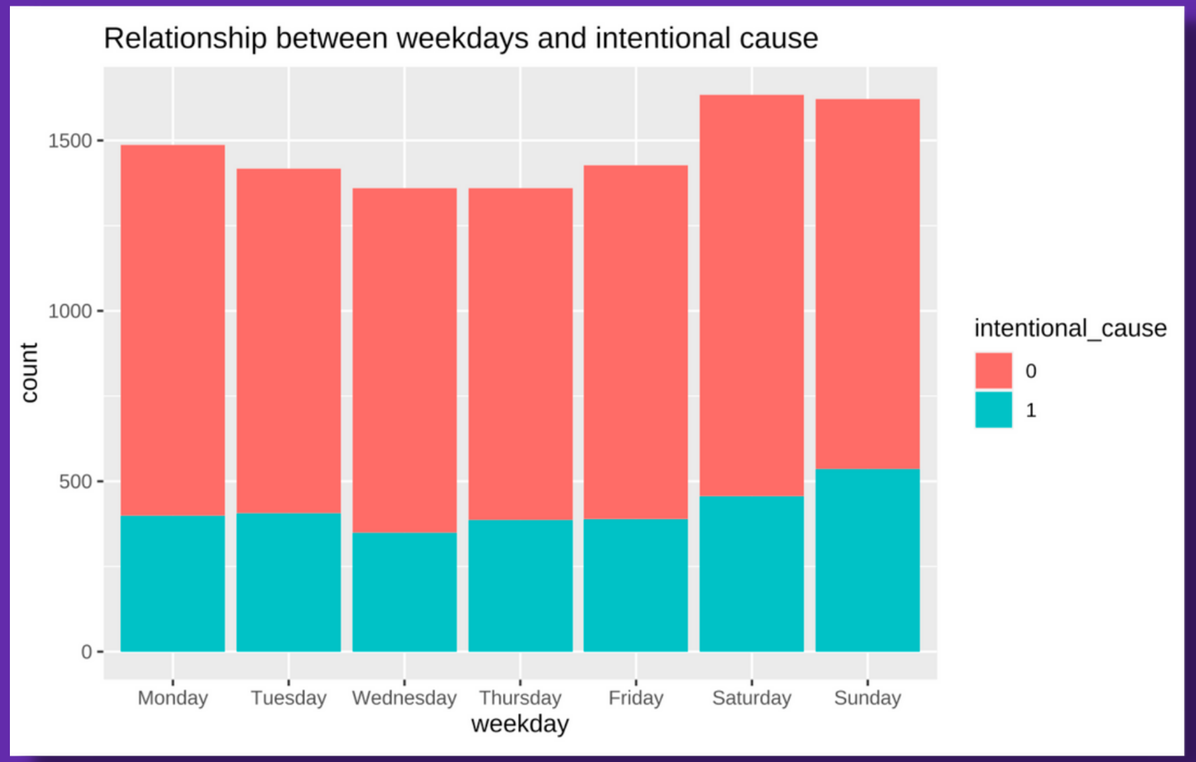
timeOfDay

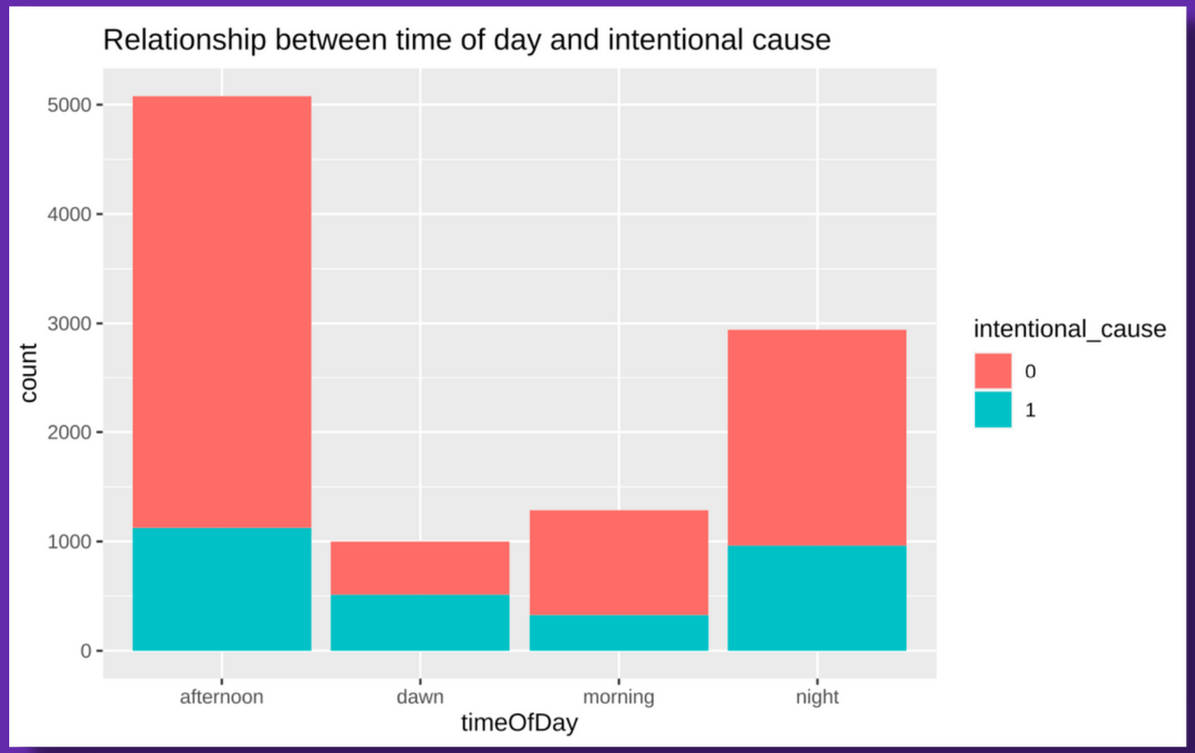
DATA EXPLORATORY ANALYSIS





DATA EXPLORATORY ANALYSIS





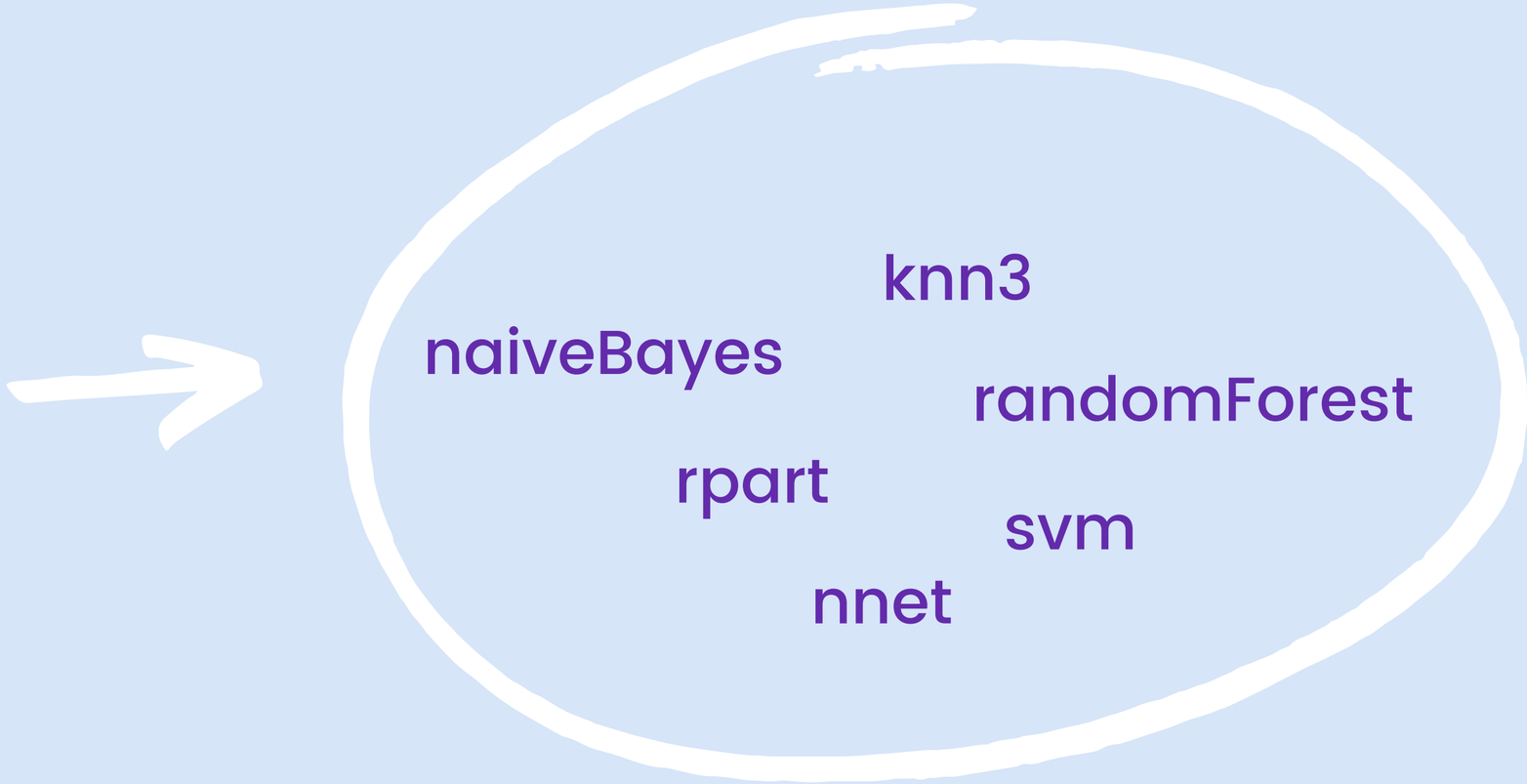
BEGINNING OF PREDICTIVE MODELLING

normlaized data

VS

non-normlaized data

1 x 10 Fold Cross Validation



What's the best ML model?

normalized

randomForest

nnet

svm

non-normalized

randomForest

rpart

TESTING...

Complete

-timeOfDay

-(season, timeOfDay)

-(alert_date, timeOfDay)

-(alert_hour, alert_date, district)

...

? randomForest ?

FINAL RESULT

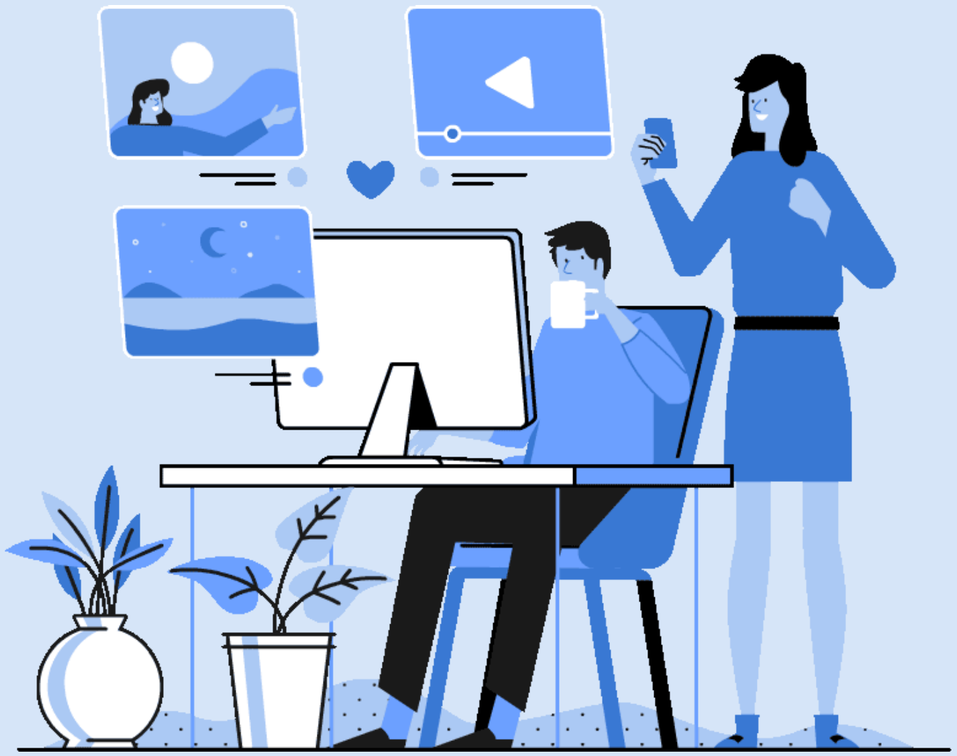
ML model:

randomForest

Dataset attributes:

-(timeOfDay)

Fazit
(Conclusion)



We were not able to use:

- bagging
- boosting
- xgboost

Dataset with a few limitations

But overall 😎