



MSC. DATA SCIENCE & ARTIFICIAL INTELLIGENCE

INTRODUCTION TO MACHINE LEARNING

Dr. Michel Riveill & Dr. Diane LINGRAND

---

## Final project: Petfinder

---

*Author:* Joris LIMONIER

[joris.limonier@hotmail.fr](mailto:joris.limonier@hotmail.fr)

*Due:* January 15, 2022

## Contents

1	Problem description	1
2	Exploratory Data Analysis	1
3	Solution	1
4	Evaluation & critical view	1

## List of Figures

## List of Tables

1	Data types per column . . . . .	1
2	Accuracies of first prospect . . . . .	1

# 1 Problem description

The problem we are trying to solve consists of predicting whether an animal will be adopted from a shelter within 30 days, given several pieces of information on this animal. This problem is a clean and reduced version of a [Kaggle competition](#) dating back from 2019.

# 2 Exploratory Data Analysis

CATEGORICAL		NUMERICAL	TEXT	IMAGE
Type	MaturitySize	Age	Description	Images
Gender	FurLength	Fee		
Breed	Vaccinated			
Color1	Dewormed			
Color2	Sterilized			
Color3	Health			

Table 1: Data types per column

# 3 Solution

Classifier	Accuracy
GradientBoostingClassifier	0.629
RandomForestClassifier	0.623
AdaBoostClassifier	0.612
MLPClassifier	0.602
BernoulliNB	0.600
GaussianNB	0.567
DecisionTreeClassifier	0.559
SVC	0.529
KNeighborsClassifier	0.520
GaussianProcessClassifier	0.509
SGDClassifier	0.509

Table 2: Accuracies of first prospect

We did not test XGBoost because is too slow.

# 4 Evaluation & critical view