COMP 432 Final Project

AmirHossein Hadinezhad, Axel Solano, Annika Timermanis

Dec 2021

COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Contents

Introduction

Organization of Development

TOOIS

Our data investigation

example
Important Feature of Breas

ancer Wisconsin

Important Fe

Cancer Wisconsin Result of Regression and

)verall

Conclusions: 3. CIFAR: Decision Tree VS CNN

Table of Contents

Title

Table of Contents

Introduction

Organization of Development

Our Goals

Tools

Our data investigation example

Important Feature of Breast Cancer Wisconsin Feature correlation Important Feature of Breast Cancer Wisconsin Result of Regression and Classification

Overall Conclusions: 3. CIFAR: Decision Tree VS CNN

Novelty Component

Outcome

COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Contents

Introduction

Organization of Development

IOOIS

Our data nvestigation example

example Important Feature of Bre

Cancer Wisconsin Feature correlation

Important Fe

Cancer Wisconsin Result of Regression a

Result of Regre Classification

Conclusions: 3. CIFAR: Decision

Introduction

This is Project of Machine Learning course, during which we had to develop a variety of models for different detests

COMP 432 Final Project

AmirHossein Hadinezhad., Axel Solano,, Annika Timermanis

Introduction

Our Goals

- Organization
- Weekly live meetings
- Structure
- Cohesion
- Gain practical experience

COMP 432 Final Project

AmirHossein Hadinezhad., Axel Solano,, Annika Timermanis

Our Goals

Challenges

- We realize we would have less people on our team as one of our teammates dropped the course
- lack of time
- Project and courses workload

COMP 432 Final Project

AmirHossein Hadinezhad., Axel Solano,, Annika Timermanis

Title

Our Goals

Tools

Version Control:

▶ GitHub

IDEs:

Jupyter

Libraries:

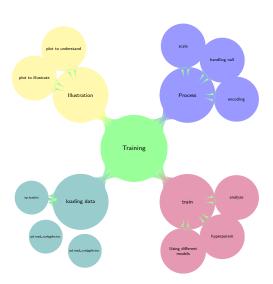
Variety of libraries including matplotlib, numpy, pandas ...

COMP 432 Final Project

AmirHossein Hadinezhad., Axel Solano,, Annika Timermanis

Tools

Training Overview



COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Contents

Introduction

Organization of Development

Tools

Our data investigation example

example
Important Feature of

Cancer Wisconsin Feature correlation

Important Feature Cancer Wisconsin

Result of Regr

Overall

CIFAR: Decision
Tree VS CNN

Our data investigation example

Data preparation and Analysis

23 1057013 8 4 5 1 2 ? 7 3 1 1096800 6 6 6 9 6 ? 7 1184840 1 1 3 1 2 7 2 158 1193683 1 1 2 1 3 2 1 1 1 2

```
1 print(np.count nonzero(dataset-"?"))
 2 print(np.count nonzero(dataset-np.nan))
 3 dataset.isnull().sum()
16
dtype: int64
   dataset.isin(["?"]).sum()
```

```
16
      Ω
dtype: int64
 1 dataset[dataset[6]=='?']
         0 1 2 3 4 5 6 7 8 9 10
```

COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano.. Annika Timermanis

Introduction

Our data investigation example

Important Feature of Breast Cancer Wisconsin

Loading Data

Dataset = pd.read_csv('Regression_Datasets/5_Bike_Sharing/Bike_Sharing.csv', delimiter=',')
Dataset.head()

	instant	dteday	season	yr	mnth	hr	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
0	1	2011- 01-01	1	0	1	0	0	6	0	1	0.24	0.2879	0.81	0.0	3	13	16
1	2	2011- 01-01	1	0	1	1	0	6	0	1	0.22	0.2727	0.80	0.0	8	32	40
2	3	2011- 01-01	1	0	1	2	0	6	0	1	0.22	0.2727	0.80	0.0	5	27	32
3	4	2011- 01-01	1	0	1	3	0	6	0	1	0.24	0.2879	0.75	0.0	3	10	13
4	5	2011- 01-01	1	0	1	4	0	6	0	1	0.24	0.2879	0.75	0.0	0	1	1

Data Prepration And Analysis

data= Dataset.iloc[:,2:].drop(['casual','registered'],axis =1)
data.head()

	season	yr	mnth	hr	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	cnt
0	- 1	0	- 1	0	0	6	0	1	0.24	0.2879	0.81	0.0	16
1	- 1	0	- 1	-1	0	6	0	1	0.22	0.2727	0.80	0.0	40
2	1	0	1	2	0	6	0	1	0.22	0.2727	0.80	0.0	32
3	1	0	1	3	0	6	0	1	0.24	0.2879	0.75	0.0	13
4	1	0	1	4	0	6	0	1	0.24	0.2879	0.75	0.0	1

data.shape

(17379, 13)

data.isnull().sum()

COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Contents

troduction

Organization of Development

Jur Goa

IOOIS

Our data investigation example

Important Feature of Breast Cancer Wisconsin

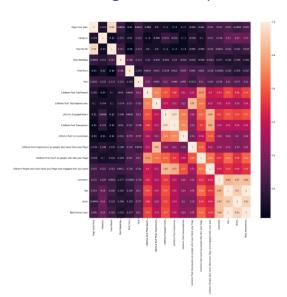
Feature correlation

Cancer Wisconsin
Result of Regression

Overall Conclusions:

CIFAR: Decisio Tree VS CNN

Our data investigation example



COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Titl

Table of Contents

Introduction

Organization of Development

Our Go

1001

Our data investigation example

Important Feature of Brea Cancer Wisconsin

Cancer Wisconsin Feature correlation

Important Feature of Brea Cancer Wisconsin Result of Regression and

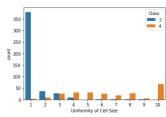
Conclusions: 3.
CIFAR: Decision
Tree VS CNN

Important Feature of Breast Cancer Wisconsin

q	recision	recall	f1-score	support
2 4	0.91	0.94	0.93	66 39
accuracy macro avg weighted avg	0.90	0.89	0.90 0.90 0.90	105 105 105

Uniformity of Cell Size

50]: <AxesSubplot:xlabel='Uniformity of Cell Size', ylabel='count'>



COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

 Fitle

Table of Contents

Introduction

Organization of Development

Tools

Our data investigation example

Important Feature of Br Cancer Wisconsin

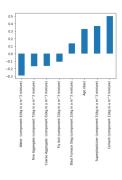
Important Feature of Breast

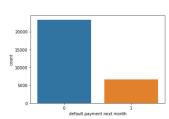
Cancer Wisconsin

Overall

Conclusions: 3. CIFAR: Decision Tree VS CNN

Target distribution and its relation to features





COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Contents

Introduction

Organization of Development

TOOIS

Our data investigation example

Important Feature of B Cancer Wisconsin

Feature correlation

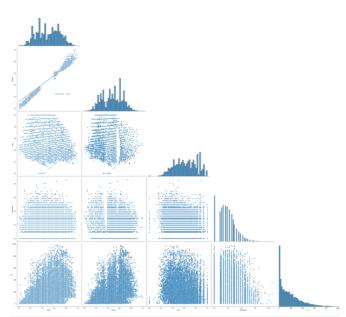
Important Feature of Breast

Cancer Wisconsin Result of Regression a

Overall

Conclusions: 3. CIFAR: Decision Tree VS CNN

Similar Features



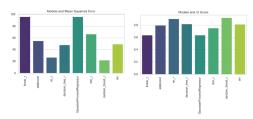
COMP 432 Final Project

AmirHossein Hadinezhad., Axel Solano,, Annika Timermanis

Important Feature of Breast

Cancer Wisconsin

Result of Regression and Classification



model	l acc	w.fl						tr.clas.rec
tree	0.842857	0.841698	0.842857	0.763636	0.953488	0.953293	0.953488	0.913978
knn	0.921429	0.921766	,		1	1	1	1
forest	0.842857	0.841698	0.842857	0.763636	0.953488	0.953293	0.953488	0.913978
adaboost	0.942857	0.94266	0.942857		0.991055	0.991037	0.991055	0.978495
gnb	0.871429	0.872917	0.871429		0.887299	•	0.887299	0.913978
nn	0.921429	0.921766	0.921429	0.927273	0.951699	0.951857	0.951699	0.94086
logistic_r	0.892857	0.894021	0.892857				0.874776	
svc	1 0.9	0.900757	1 0.9	0.927273	0.996422	0.996427	0.996422	1

COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Contents

ntroduction

Organization of Development

Jur Goa

ools

Our data nvestigation xample

mportant Feature of Br Cancer Wisconsin

eature correlation

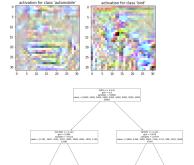
Important Feature Cancer Wisconsin

Result of Regression and Classification

Overall
Conclusions: 3.
CIFAR: Decision
Tree VS CNN

Overall Conclusions: 3. CIFAR: Decision Tree VS CNN

- Most interpretable
 - ► The CNN was more interpretable than the Decision Tree
- ► Reasons::
 - CNN activation maximization figures
 - Decision Tree plot



COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Conten

ntroduction

Organization of Development

ui Goai

ools

Our data nvestigation xample

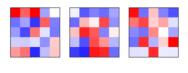
mportant Feature of Brea Cancer Wisconsin Feature correlation mportant Feature of Brea

Cancer Wisconsin
Result of Regression

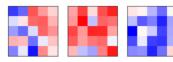
Overall
Conclusions: 3.
CIFAR: Decision
Tree VS CNN

Novelty Component

The novelty component chosen: Tried inspecting the CNN's filters



CNN First Layer Filters



CNN Final Layer filters

COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Contents

Introduction

Organization of Development

IOOIS

Our data nvestigation xample

Important Feature of Breast Cancer Wisconsin

Important Feature Cancer Wisconsin

Overall Conclusions: 3.

Outcome

What We Succeeded VS What We Could Have Improved

- Succeeded
 - Trained and tuned every model
 - Did an analysis, and cross comparison of models
- Improved:
 - ▶ Timing
 - Organization towards the end of the semester
 - Class weights
 - Dummy estimator

COMP 432 Final Project

AmirHossein Hadinezhad,, Axel Solano,, Annika Timermanis

Title

Table of Contents

Introduction

Organization of Development

Tools

Our data investigation

mportant Feature of Bre Cancer Wisconsin

Cancer Wisconsin Feature correlation

Important Featur Cancer Wisconsir

Result of Regression an Classification

Conclusions: 3.