Case Studies 2022L
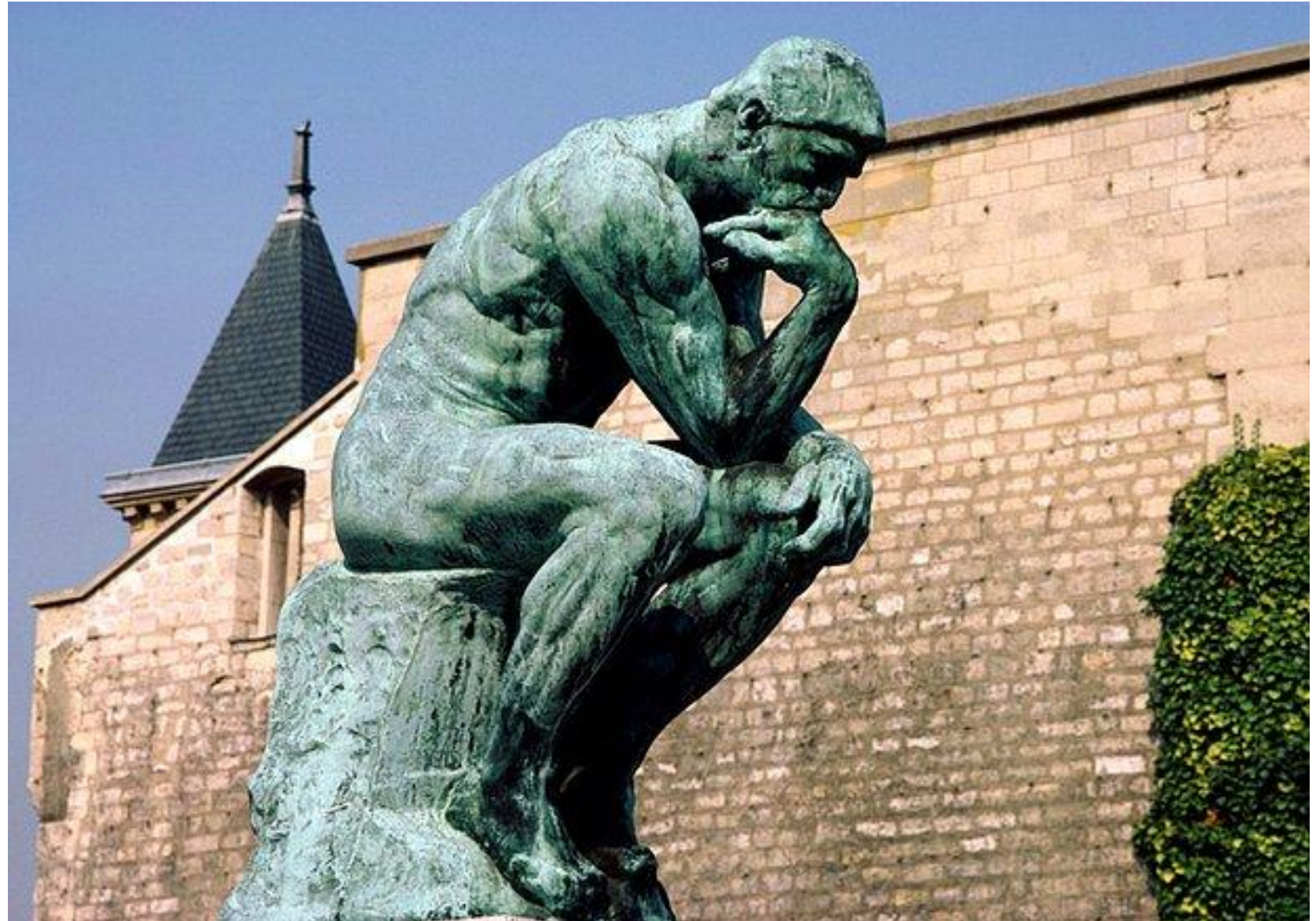
ML Philosophy and XAI
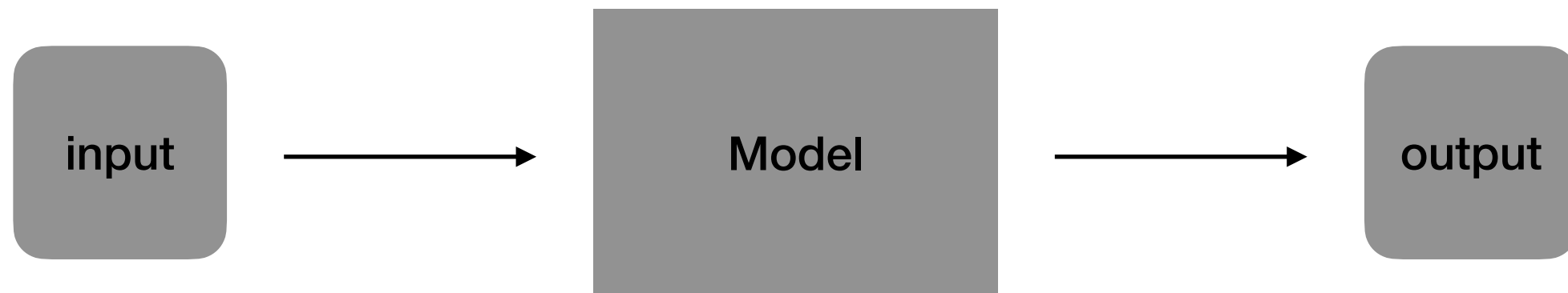
Mar 3, 2022

Do you like philosophy?

Le Penseur, Auguste Rodin (1840-1917)
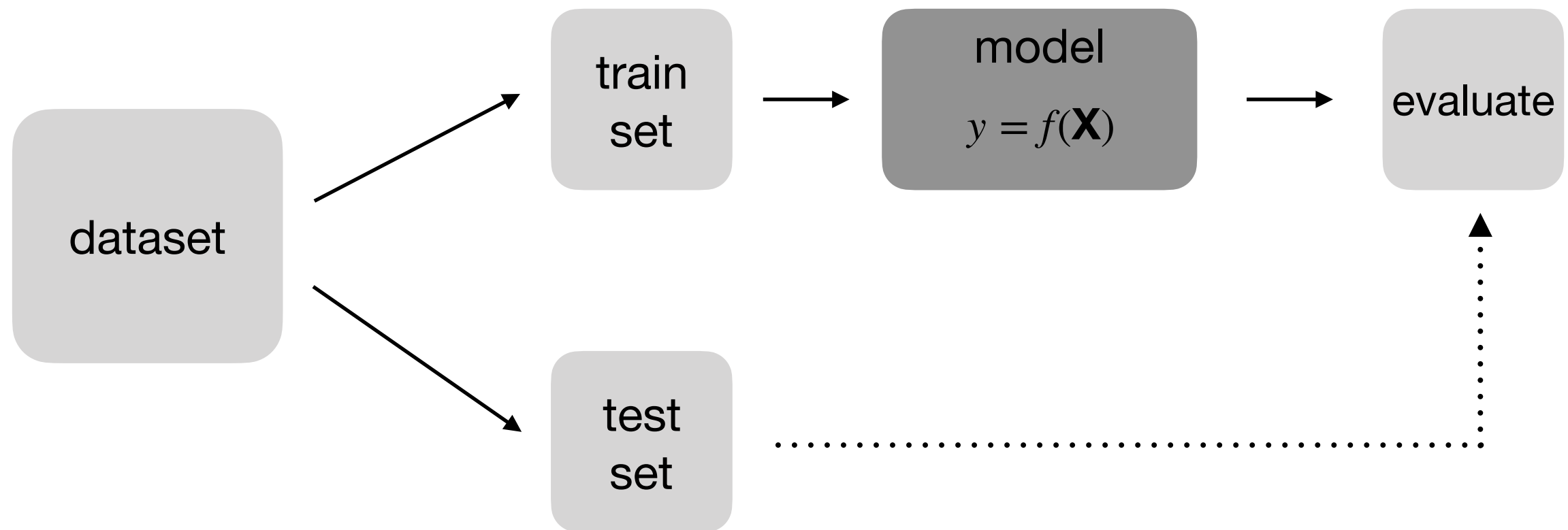
Learning Process

input → Model → output

Machine Learning

- A set of tools for making predictions from the data.

- Intersection of statistics and computer science.

- Ability to learn from the data without programmed.

- Learns patterns from data and applies it to new data.

- Quality of the process depends on the data.

ML Workflow

Train set: data that model learn from

Test/validation set: new data to measure/validate the performance of the model

Main types of ML problems

1. Supervised Learning

   • Regression

   • Classification

2. Unsupervised Learning

   • Clustering

3. Deep Learning

4. Reinforcement Learning

…

dataset

```
> data(apartments)
>
> head(apartments)
  m2.price construction.year surface floor no.rooms    district
1     5897              1953      25     3        1 Srodmiescie
2     1818              1992     143     9        5     Bielany
3     3643              1937      56     1        2       Praga
4     3517              1995      93     7        3      Ochota
5     3013              1992     144     6        5     Mokotow
6     5795              1926      61     6        2 Srodmiescie
```

```
> glimpse(apartments)
Rows: 1,000
Columns: 6
$ m2.price          <dbl> 5897, 1818, 3643, 3517, 3013, 5795, 29…
$ construction.year <dbl> 1953, 1992, 1937, 1995, 1992, 1926, 19…
$ surface           <dbl> 25, 143, 56, 93, 144, 61, 127, 105, 14…
$ floor             <int> 3, 9, 1, 7, 6, 6, 8, 8, 6, 9, 5, 7, 1,…
$ no.rooms          <dbl> 1, 5, 2, 3, 5, 2, 5, 4, 6, 4, 2, 3, 3,…
$ district          <fct> Srodmiescie, Bielany, Praga, Ochota, M…
```
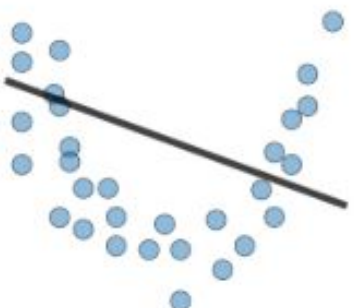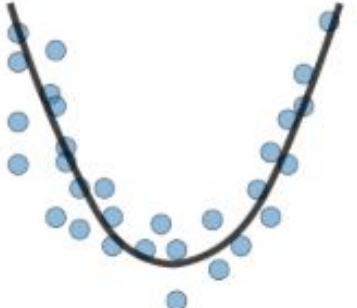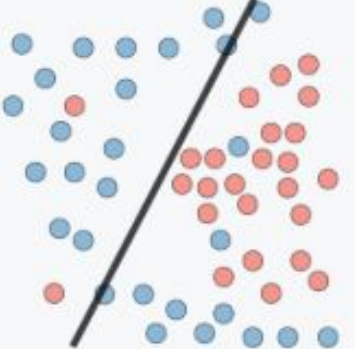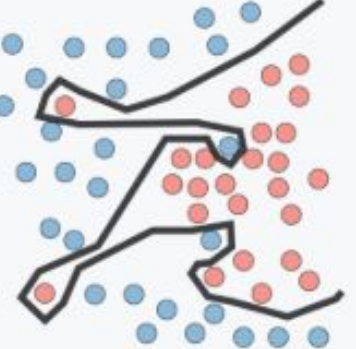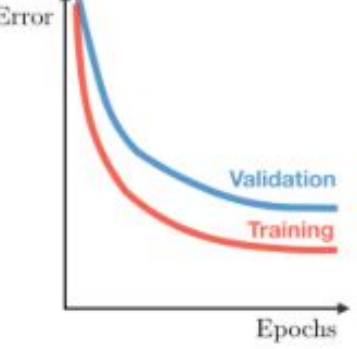
Performance metrics

Regression
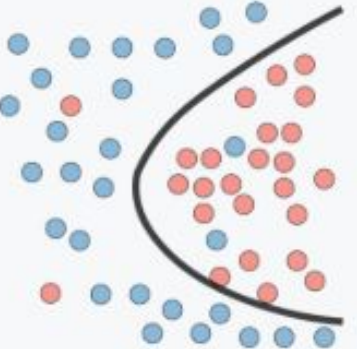
- Root mean square error (RMSE)

- R-squared

- Mean absolute error (MAE)

Classification

- Recall: TP / (TP + FN)

- Precision: TP / (TP + FP)

- Accuracy: (TP + TP) / (TP + FP + FN + TN)

True Class

	Positive	Negative
Positive	TP	FP
Negative	FN	TN

Predicted Class

Problems in ML

	Underfitting	Just right	Overfitting
Symptoms	• High training error • Training error close to test error • High bias	• Training error slightly lower than test error	• Very low training error • Training error much lower than test error • High variance
Regression illustration			
Classification illustration			
Deep learning illustration			
Possible remedies	• Complexify model • Add more features • Train longer		• Perform regularization • Get more data

💡 compare the model performance on the train and test set

Problem?

When your kitten
thinks he's a
rabbit too

Explainable artificial intelligence
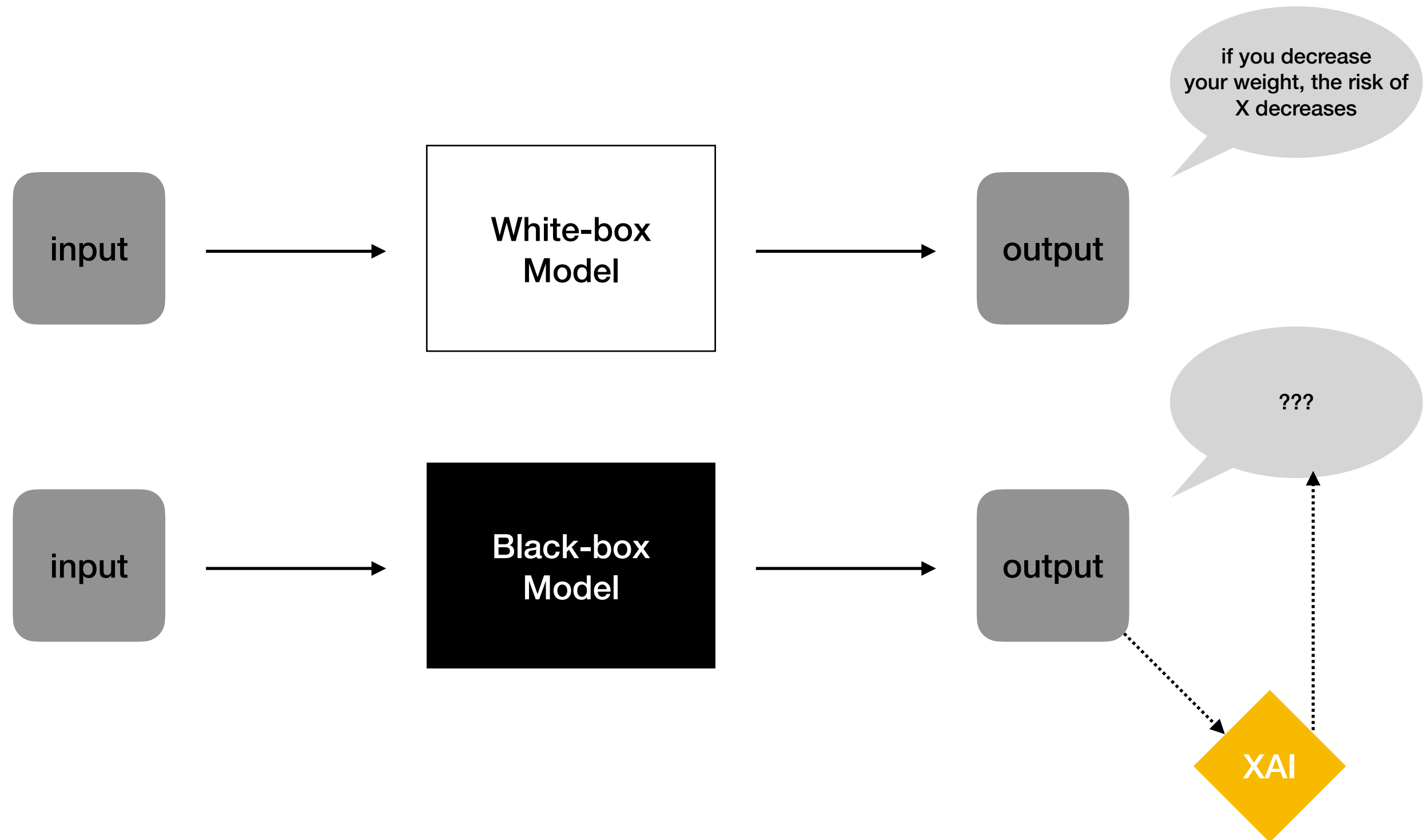
Please feel free to send e-mail about your questions!

📨 mustafa.cavus@pw.edu.pl