

Introduction to machine learning

What is machine learning?

Herman Kamper

<http://www.kamperh.com/>

Let's write a spam filter

Simple first attempt in Python:

```
def check_spam(email):  
    spam = False  
    for word in email:  
        if word in ["inheritance", "win", "uncle", ...]:  
            spam = True  
    return spam
```

Let's write a spam filter

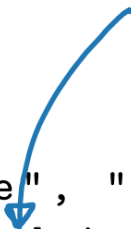
Simple first attempt in Python:

```
def check_spam(email):
    spam = False
    for word in email:
        if word in ["inheritance", "win", "uncle", ...]:
            spam = True
    return spam
```

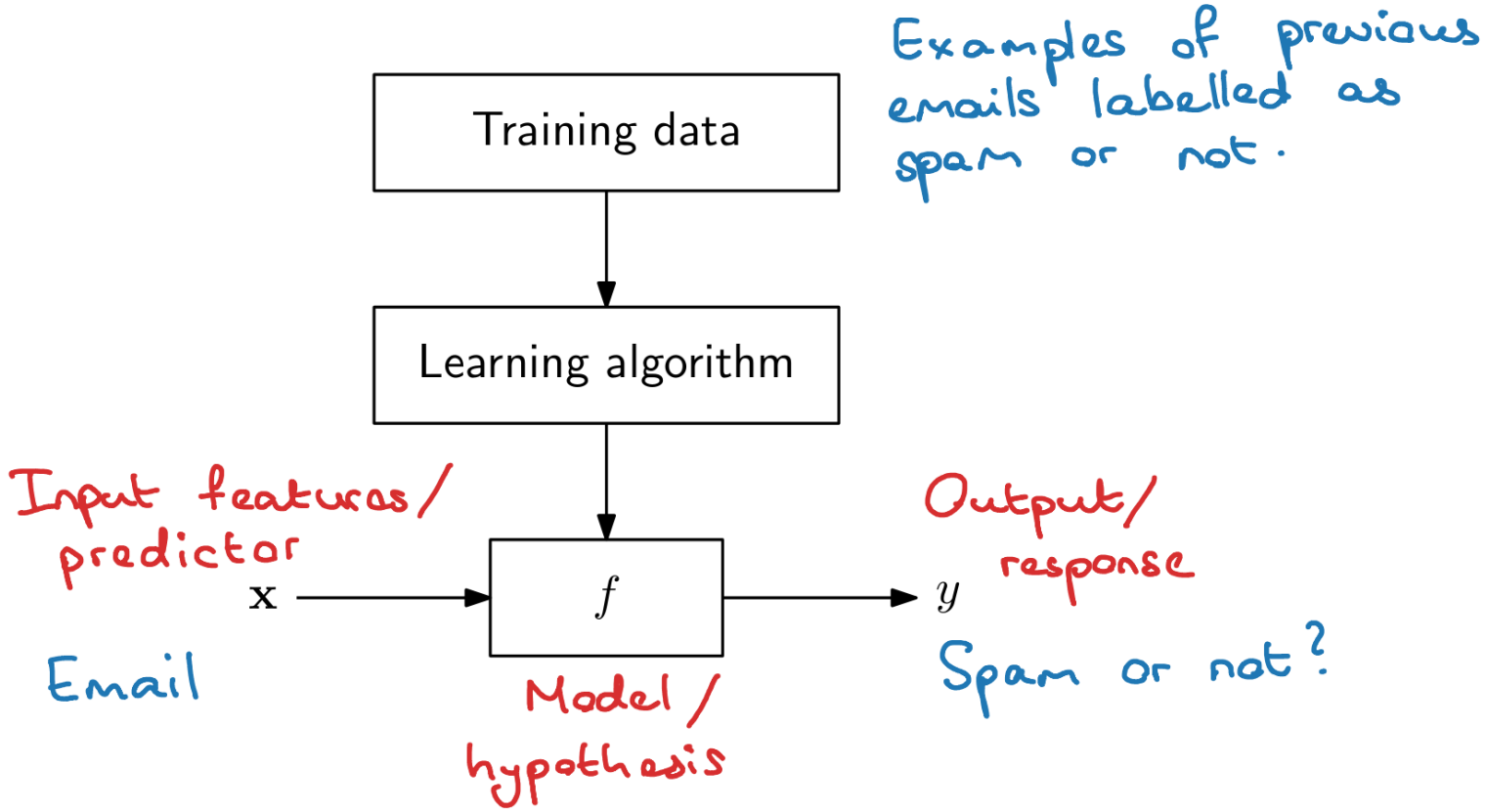
More advanced attempt:

```
def check_spam(email):
    spam_score = 0
    for word in email:
        if word in ["inheritance", "win", "uncle", ...]:
            spam_score += how_bad_is_word[word]
    if spam_score > threshold:
        spam = True
    else:
        spam = False
    return spam
```

How do we set the
word "badness"?



Machine learning



Two types of learning

- **Supervised learning:** Learn from labelled input-output (\mathbf{x}, y) pairs.
- **Unsupervised learning:** Learn from unlabelled data, i.e. only have inputs \mathbf{x} .

Machine learning, statistics, data science/analytics

Machine learning:

- “Field of study that gives computers the ability to learn without being explicitly programmed.” [Arthur Samuel, 1959]
- “The goal of machine learning is to develop methods that can automatically detect patterns in data, and then to use the uncovered patterns to predict future data or other outcomes of interest.” [Kevin P. Murphy]
- Often focusses on prediction.

Statistics:

- “Analysis, interpretation and presentation of data.” [Wikipedia]
- More emphasis on explaining than prediction: “What does the data say?”

Data science/analytics:

- Combination of machine learning and statistics.
- Also engineering infrastructure for acquiring, storing and (pre-)processing data.

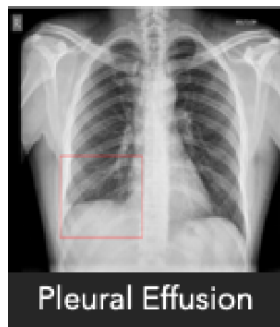
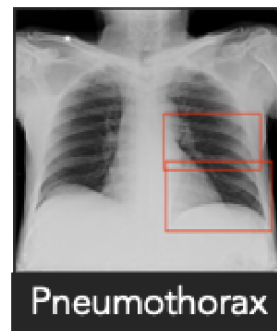
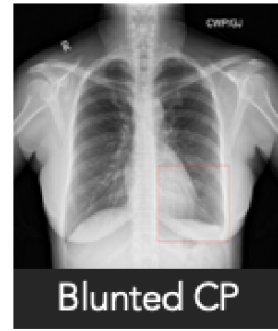
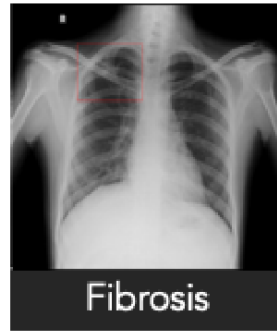
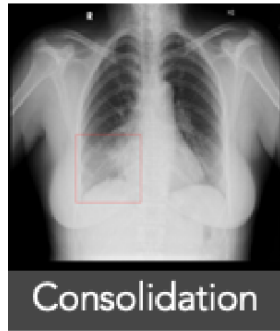
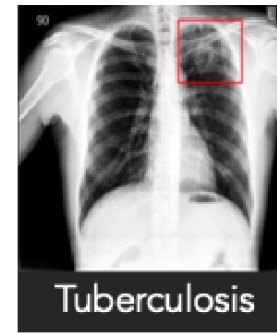
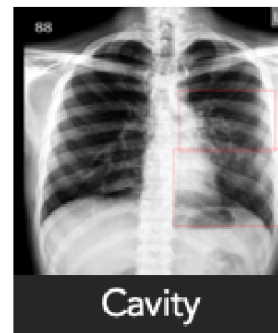
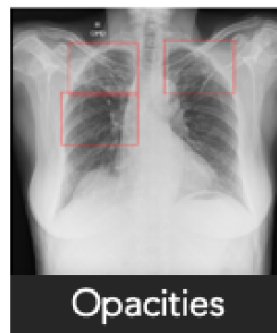
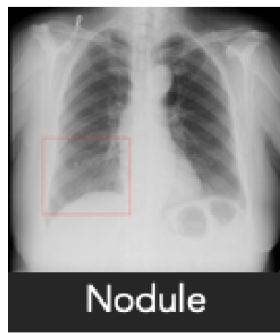
Introduction to machine learning

Supervised and unsupervised learning

Herman Kamper

<http://www.kamperh.com/>

Supervised learning



Covid-19 Risk: High

60%

71%

qXR Interpretation

Abnormal	YES
Lungs	
Opacity	YES
Location: LP, LL, RP, RL	
Atelectasis	NO
Calcification	NO
Cavity	NO
Consolidation	YES
Location: LP, LL, RP, RL	
Fibrosis	NO
Nodule	YES
Location: LU, LP, LL, RP, RL	
Pleura	
Blurred Costophrenic Angle	NO
Pleural Effusion	NO
Mediastinum	
Hilar Enlargement	NO
Heart	
Cardiomegaly	NO

Powered by
qure.ai

Opacity is observed in bilateral mid and lower zones

Projection area of the left lung affected - 71%

Projection area of the right lung affected - 60%

Inhomogeneous Opacity, probable Consolidation is observed in bilateral mid and lower zones

Nodular Opacity observed in left lung and right mid, lower zones

Pleura appears normal

Heart appears normal

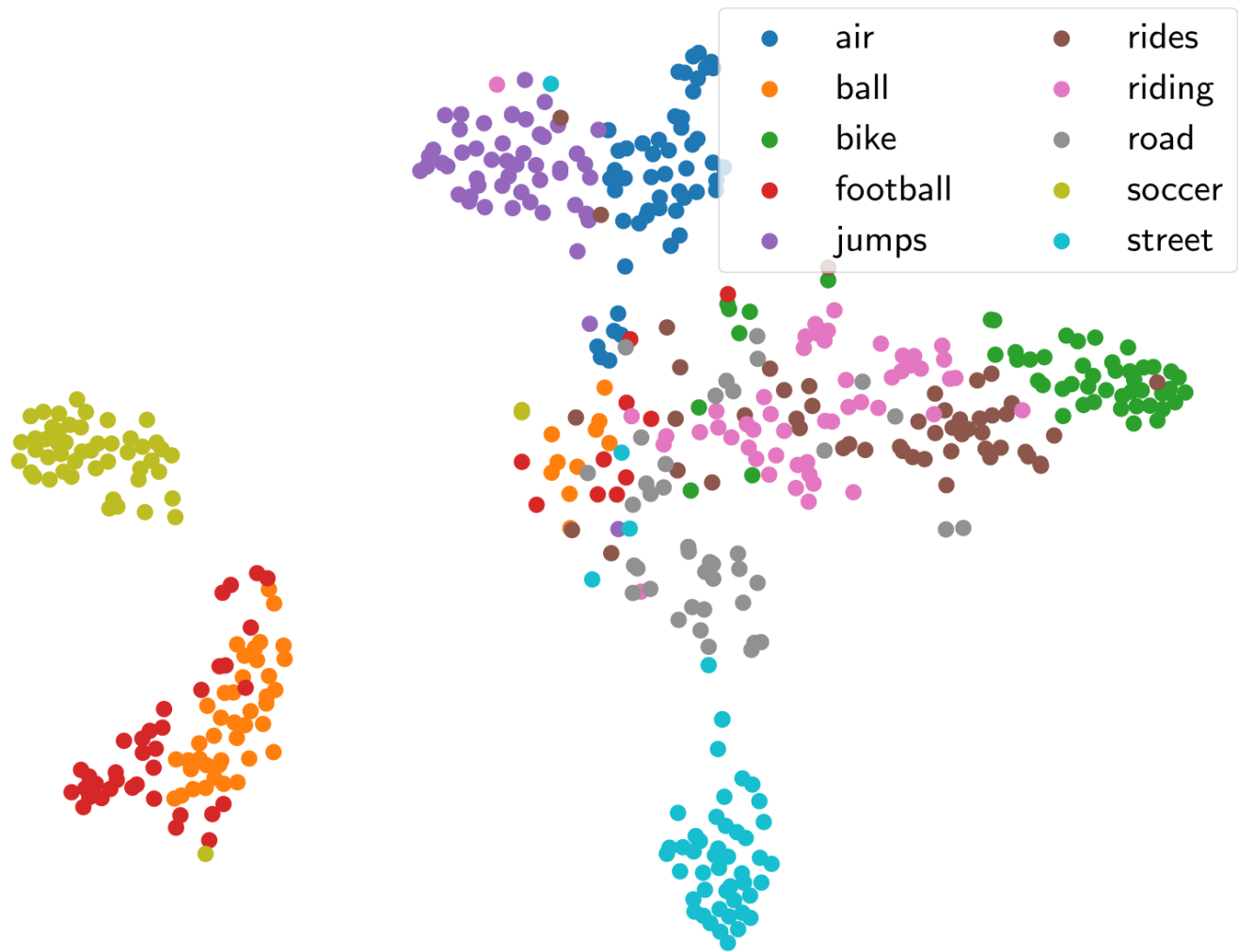
Covid19-Risk: high

Symptoms

N/A

Unsupervised learning





STAYING BIG OR GETTING SMALLER

Expected structural changes in the energy system made possible by the increased use of digital tools

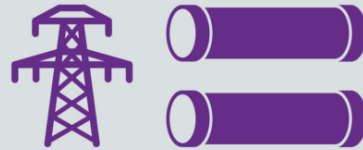
yesterday



few large power plants



centralized, mostly national



based on large power lines and pipelines



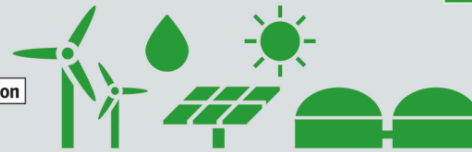
top to bottom



passive, only paying

production

tomorrow



many small power producers



decentralized, ignoring boundaries



including small-scale transmission and regional supply compensation

transmission

distribution



both directions



active, participating in the system

consumer