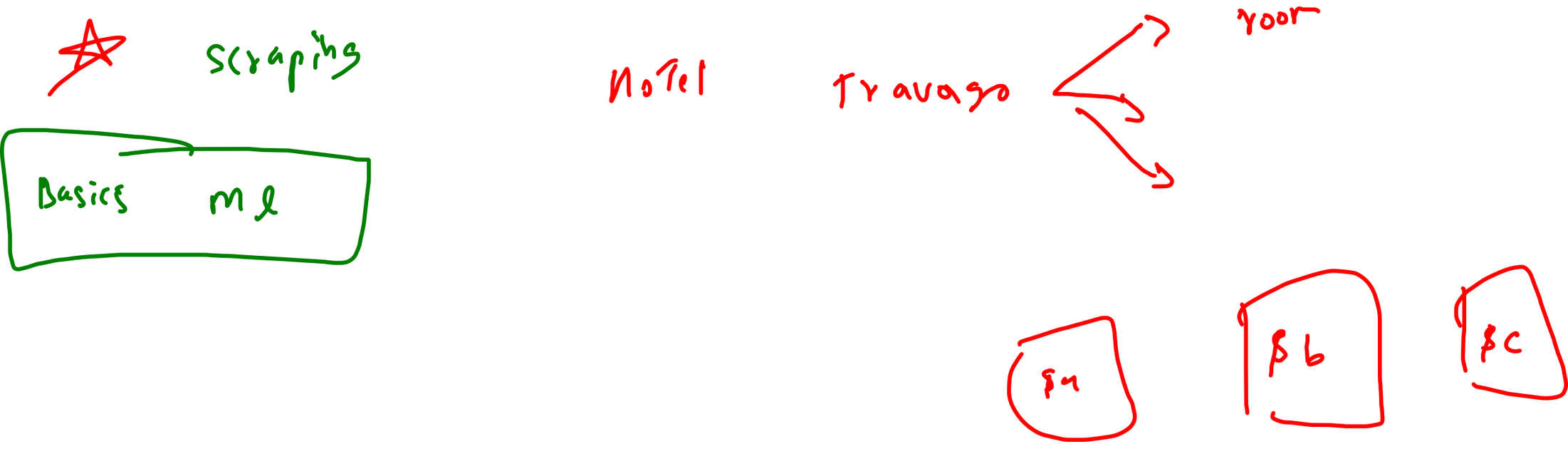

Scraping
Basics
ML
Hotel
Travago
yoor
$m
$b
$c

}

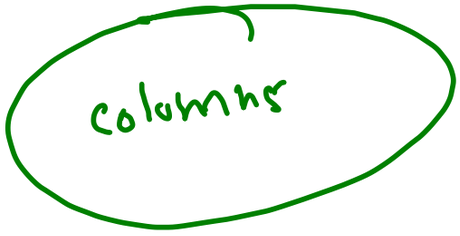
<a>

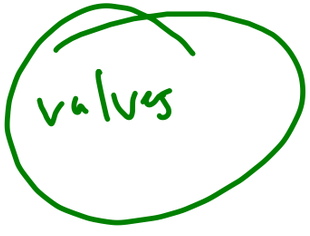
bs.findAll('a')

bs.find(div).findAll('a')

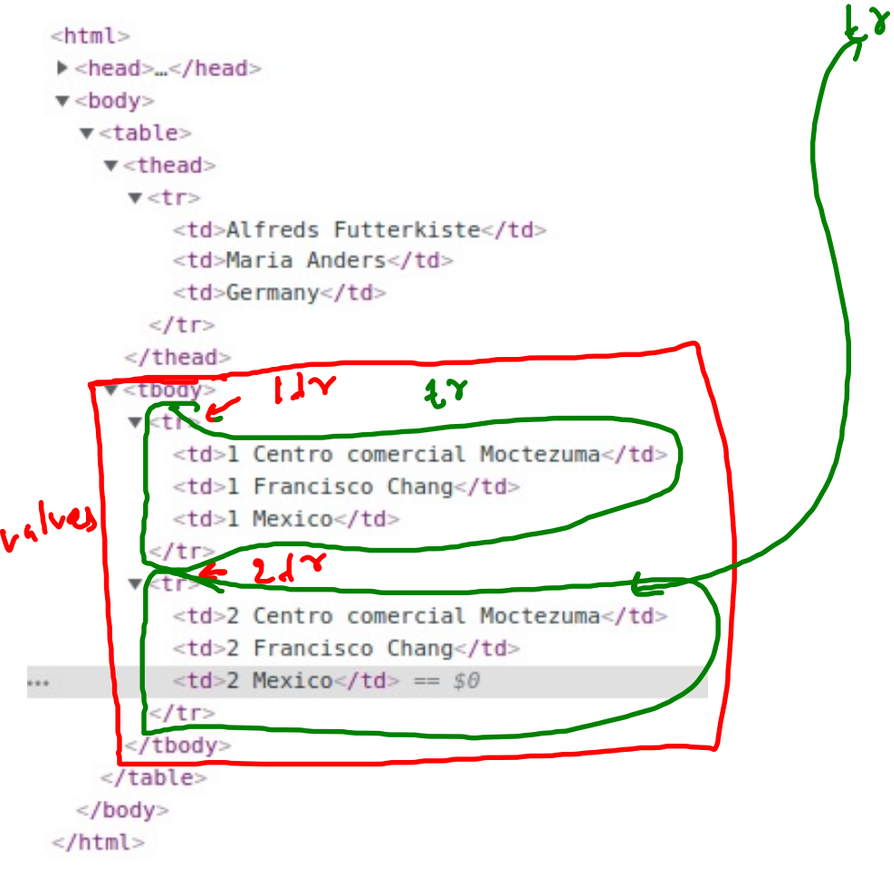
attrs = {'attr': value}
name

Data frame






```
<html>
▶ <head>…</head>
▼ <body>
  ▼ <table>
    ▼ <thead>
      ▼ <tr>
          <td>Alfreds Futterkiste</td>
          <td>Maria Anders</td>
          <td>Germany</td>
        </tr>
      </thead>
    ▼ <tbody>
      ▼ <tr>
          <td>1 Centro comercial Moctezuma</td>
          <td>1 Francisco Chang</td>
          <td>1 Mexico</td>
        </tr>
      ▼ <tr>
          <td>2 Centro comercial Moctezuma</td>
          <td>2 Francisco Chang</td>
          <td>2 Mexico</td> == $0
        </tr>
      </tbody>
    </table>
  </body>
</html>
```

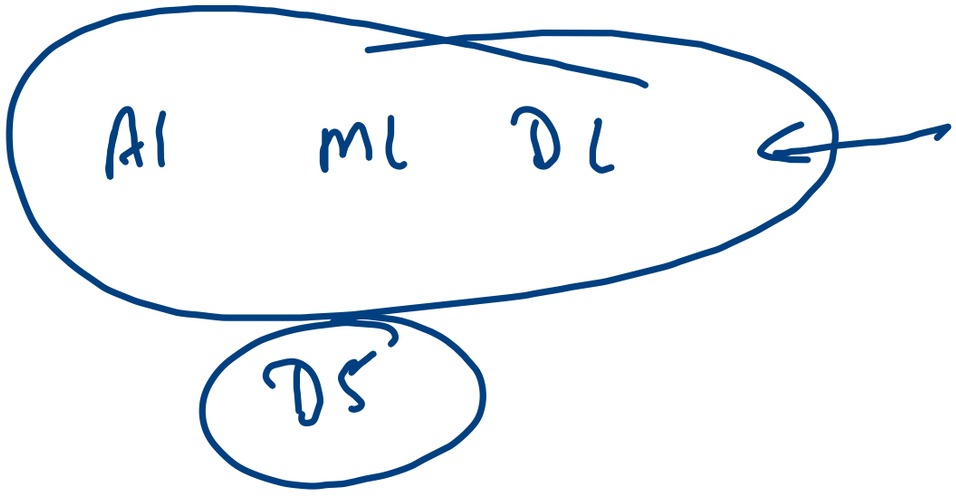



```
values = bs.tbody
values
```

```
pd_values = []

for tr in values.findAll('tr'):
    row = []
    for td in tr.findAll('td'):
      row.append(td.text)

    pd_values.append(row)

pd_values
```

```
[['1 Centro comercial Moctezuma', '1 Francisco Chang', '1 Mexico'],
 ['2 Centro comercial Moctezuma', '2 Francisco Chang', '2 Mexico']]
```

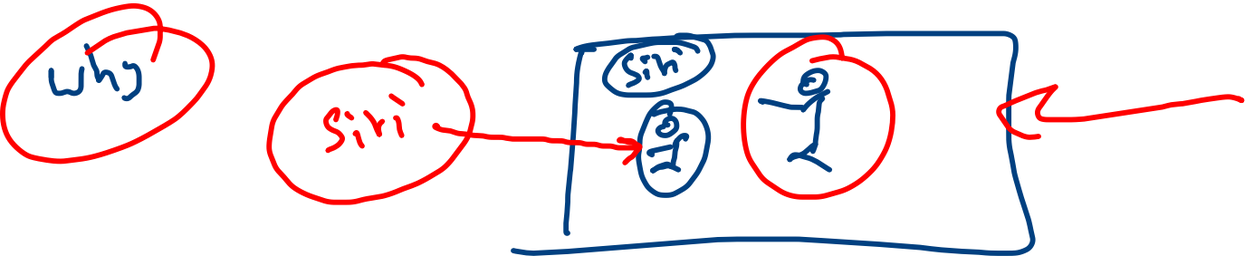
pd_values = [[1centro, 1franc., 1mexico], [2cent, 2franc., 2mexico]]



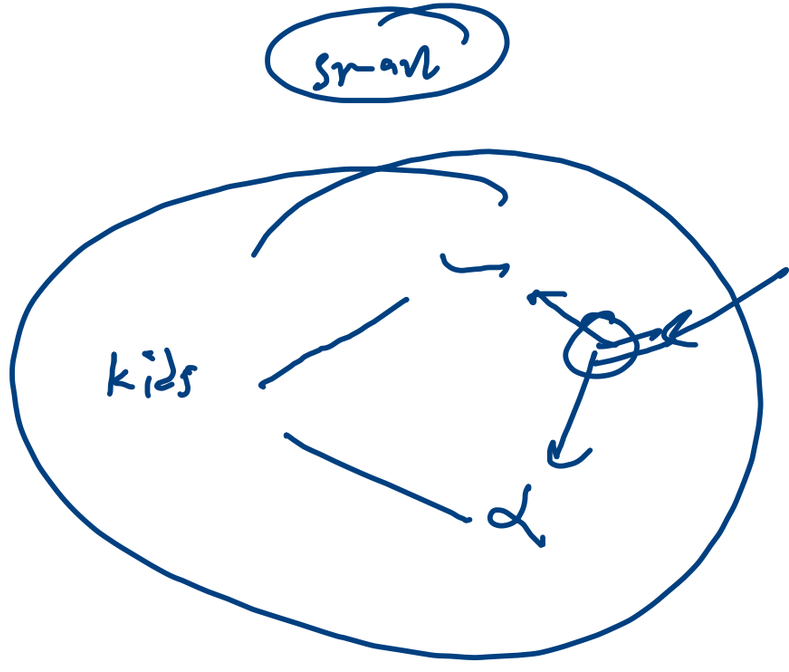
AI → Machine Intelligence
Artificial Intelligence

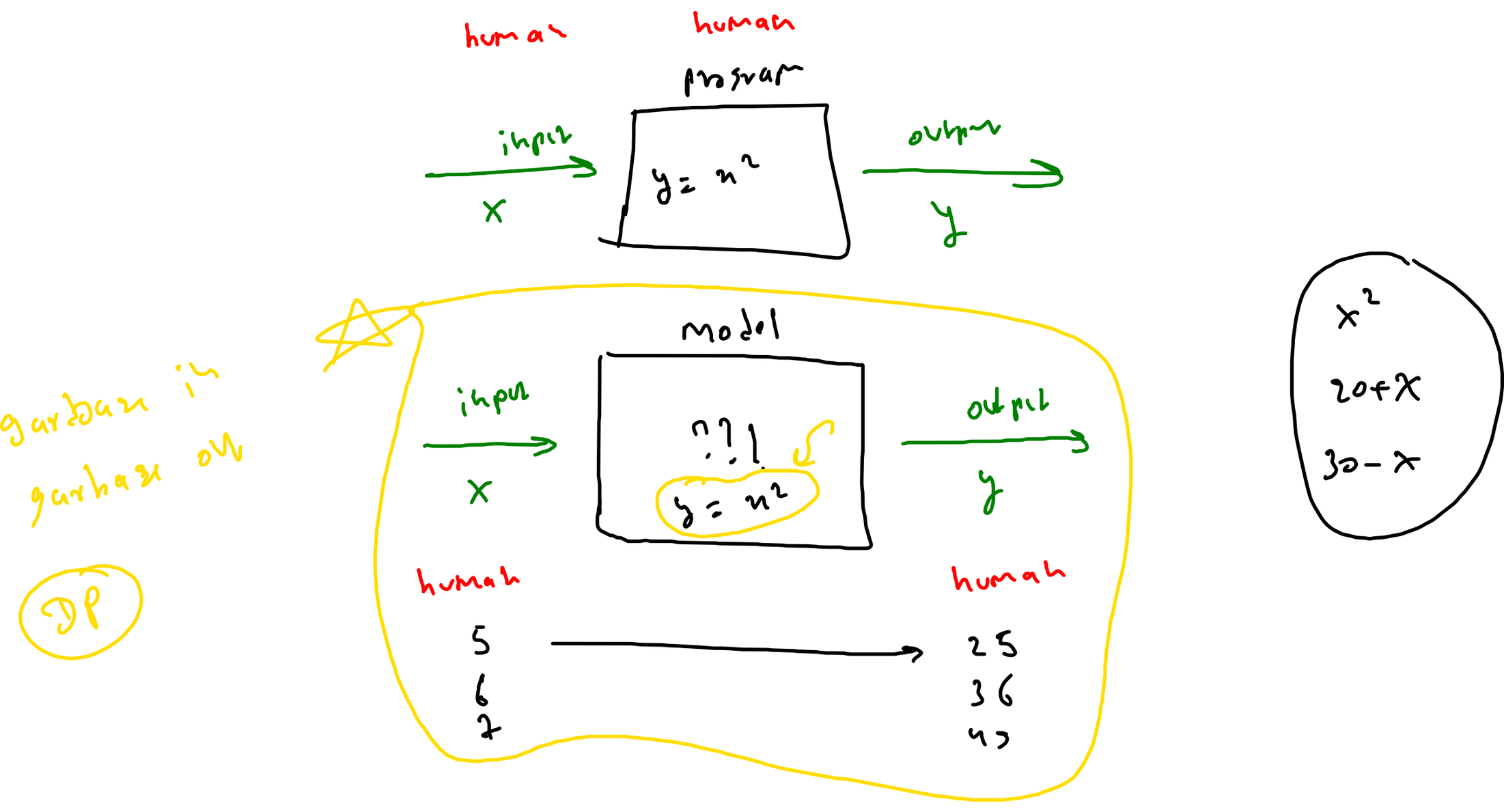
ML → ? stats

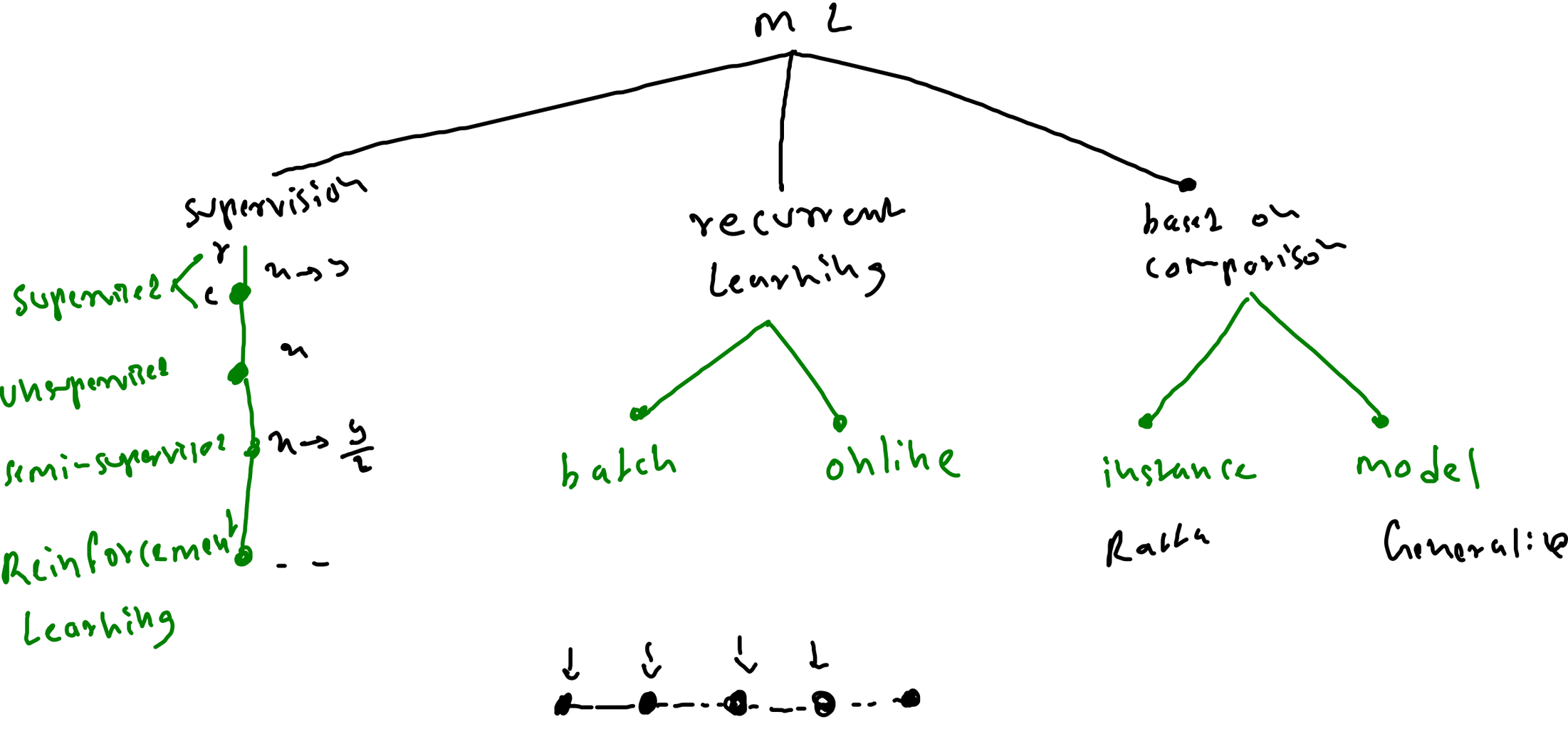
DL → neural networks



elaboration




human
human
program
input
x
$y = x^2$
output
y
model
input
x
?!!
$y = x^2$
output
y
human
human
5
6
7
25
36
49
x^2
$20+x$
$30-x$
garbage in
garbage out
DP


ML
Supervision
recurrent Learning
based on comparison
Supervised
Unsupervised
semi-supervised
Reinforcement Learning
batch
online
instance
model
Rule
Generalize

Supervised

label

X

Y

y

infinite set

finite set

Regression

[4, 5, 7, 8, 5.5]

5.1, 5.2, 5.3

[10k, 20k, 30k, 40k)

yy

finite set

classification

{ cat, dog, horse }

{ male, female }

{ 1, 2, 3, 4, 5, 6 }

X Iris

SL, SW, PL, PW

y

S
V
V

Un-Supervised

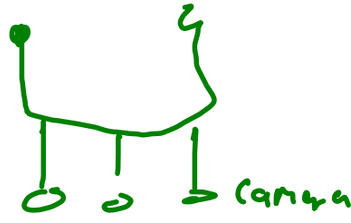
dog, cat, horse

input

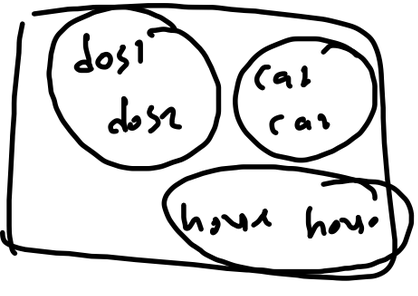
X

mnist

784



clustering
Grouping



Dimensional
deduction

784
⋮
2, 3

Anomaly
detection

outlier
inquiry

Bank
Fraud

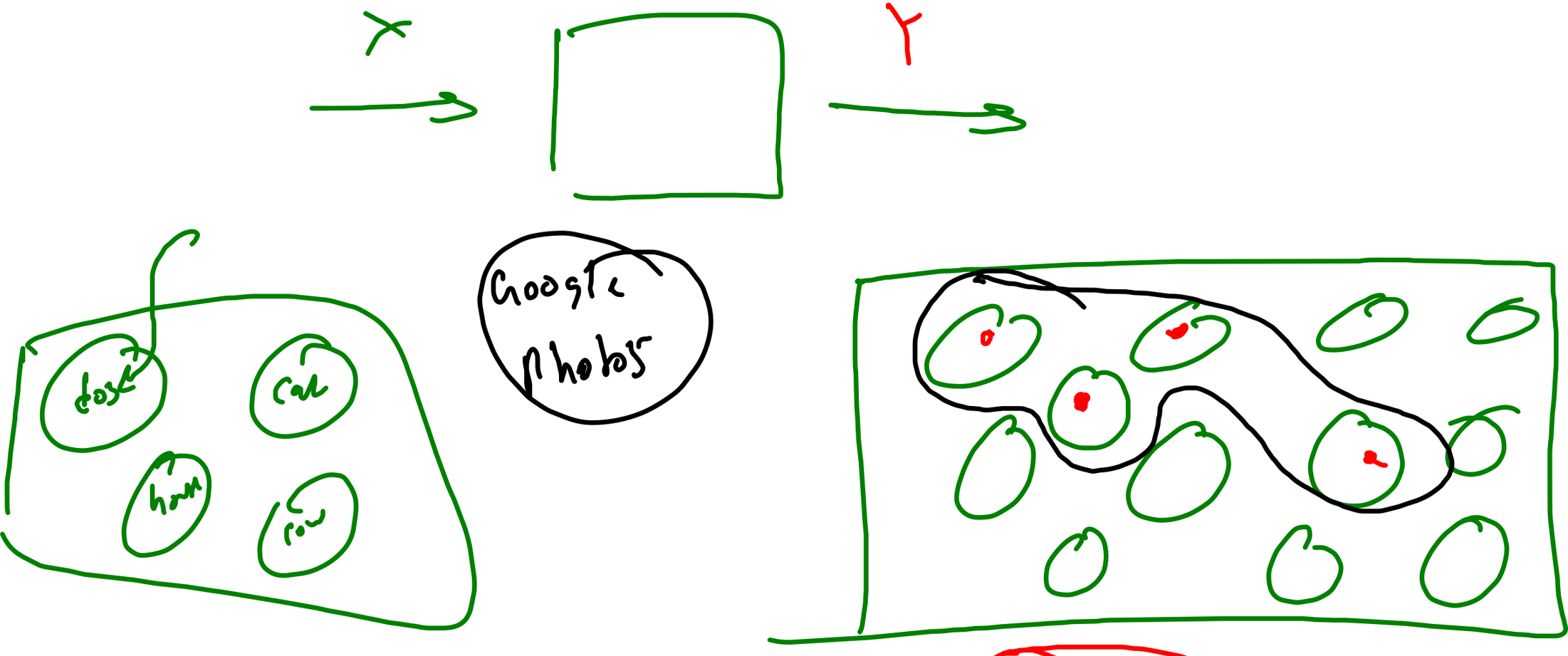
Association
rule learning

Walmart

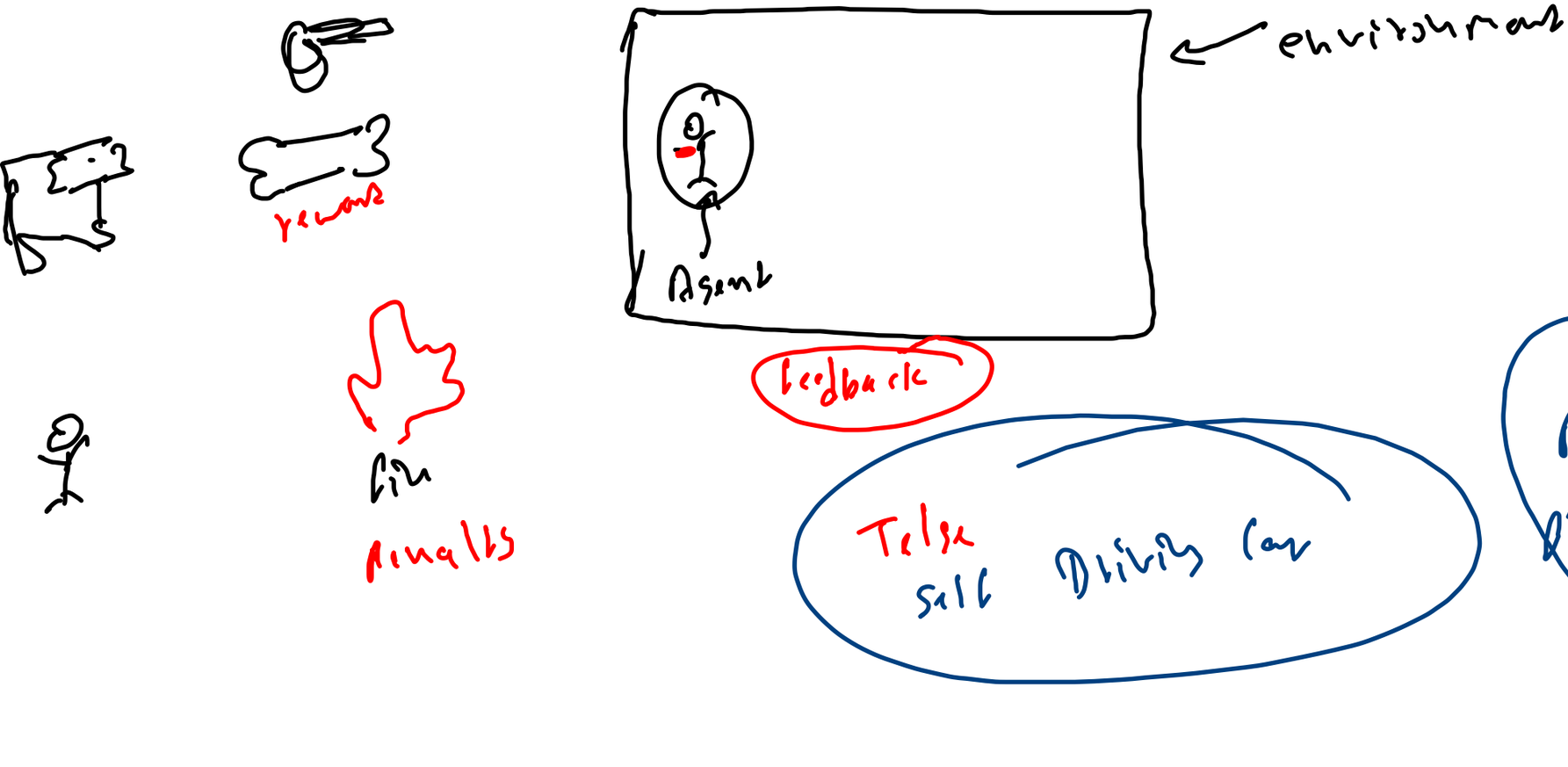
Beer - Diaper

IT, tool, etc.
electronics

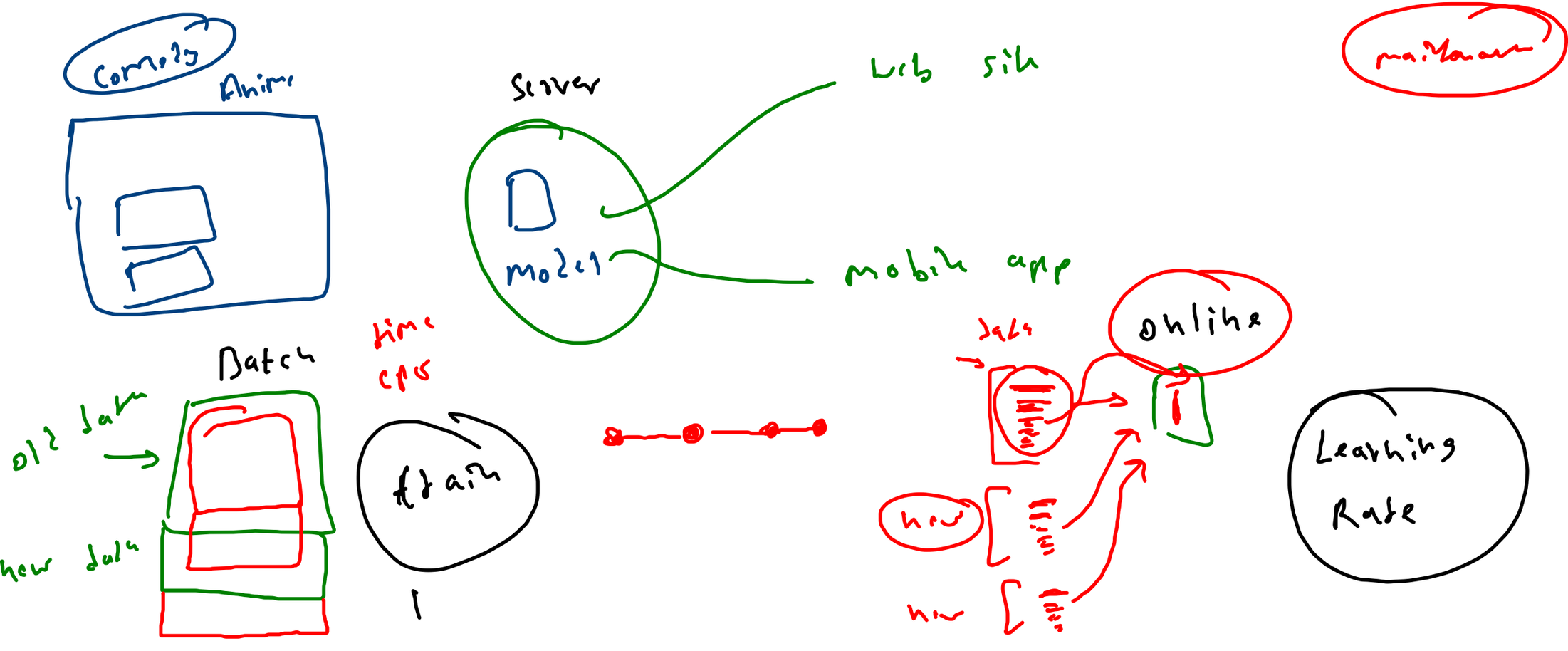
Semi- supervised



Reinforcement Learning



Recurrent Learning



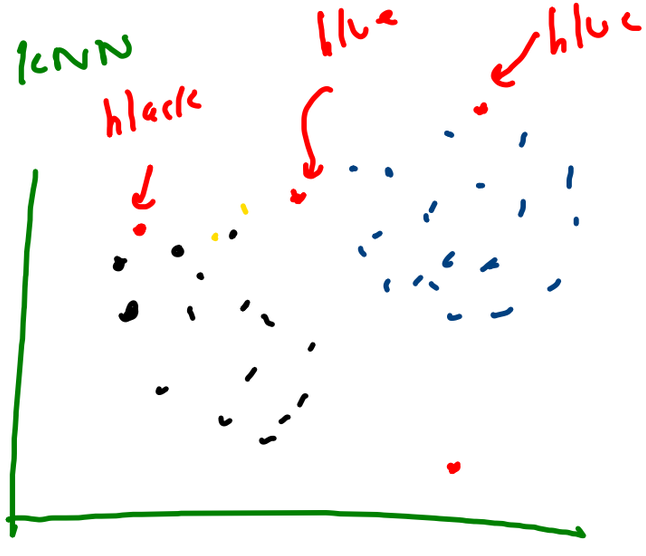
Based on Comparison

Instance

Model

KNN

Rabbit



Logic, Generalize