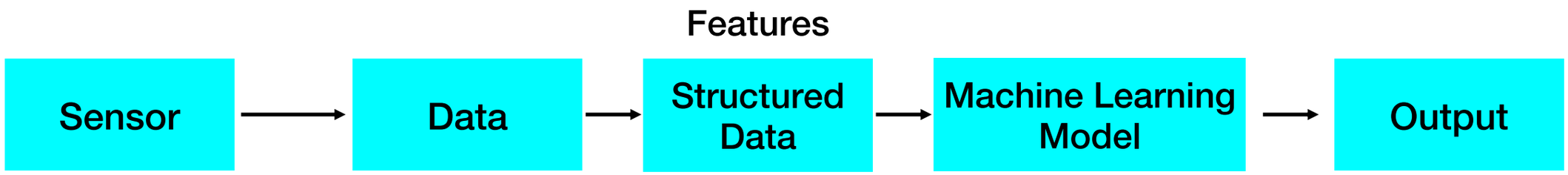

Features
Sensor
Data
Structured Data
Machine Learning Model
Output

Assignment 3

- Collect a dataset of motion sensors on the phone for a set of daily activities

- Make sure to at least include:
  - Walking (at least 4 hours)
  - Climbing up the stairs (at least 20 minutes)
  - Climbing down the stairs (at least 20 minutes)
  - Standing up (at least 50 instances)
  - *one more activity of your choice*

Apps to use for Assignment 3

iOS

Sensor Logger 4+

Motion Logger with CSV Export

Choi Tsz Hei

★★★★★ 5.0 • 10 Ratings

Free

Android

AndroSensor

Fiv Asim Tools

★★★★½ 7,369

Everyone

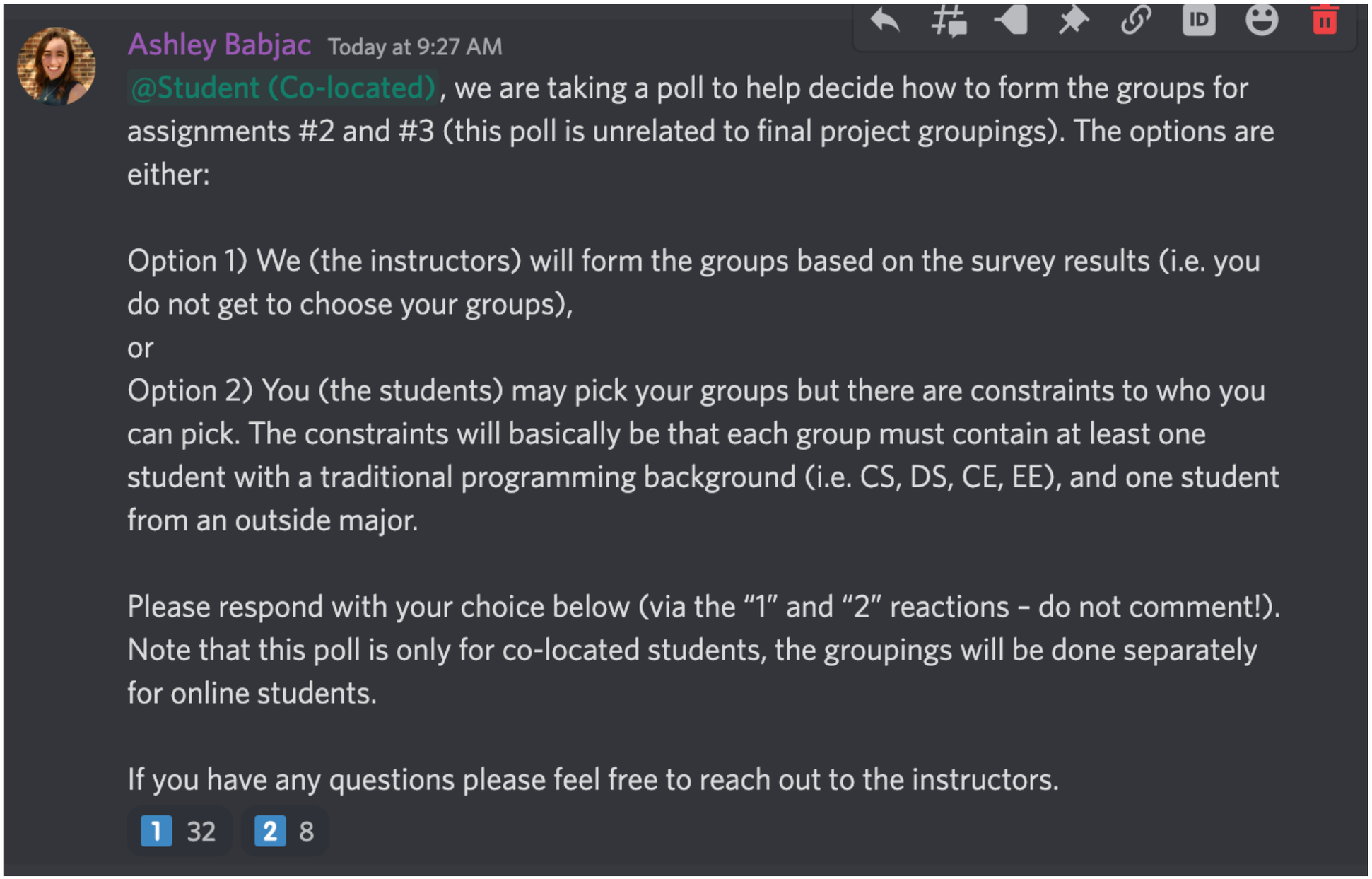
Contains Ads

You don't have any devices.

You can share this with your family. Learn more about Family Library

Installed

Poll on Discord



Using Weka

Provides pretty easy access to lots of different ML algorithms and a framework for trying them out

- https://waikato.github.io/weka-wiki/downloading_weka/
- Open source machine learning toolkit
- Includes Java API

Data for Weka

Assumes data is
a collection of (training) instance vectors

- Features plus a label
- Uses the term “attribute” to refer to an individual data value (either feature or label)

Primary data file format: ARFF (Attribute-Relation File Format)

Sample ARFF File

```
% Iris Plants Database (from R.A. Fischer and Michael Marshall)
%
@RELATION iris
@ATTRIBUTE sepallength NUMERIC
@ATTRIBUTE sepalwidth NUMERIC
@ATTRIBUTE petallength NUMERIC
@ATTRIBUTE petalwidth NUMERIC
@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}
@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,?,1.5,0.2,Iris-setosa
    …
```

Sample ARFF File

```
% Iris Plants Database (from R.A. Fischer and Michael Marshall)
%
@RELATION iris
@ATTRIBUTE sepallength NUMERIC
@ATTRIBUTE sepalwidth NUMERIC
@ATTRIBUTE petallength NUMERIC
@ATTRIBUTE petalwidth NUMERIC
@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}
@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,?,1.5,0.2,Iris-setosa
    …
```

A comment

Sample ARFF File

```
% Iris Plants Database (from R.A. Fischer and Michael Marshall)
%
@RELATION iris
@ATTRIBUTE sepallength NUMERIC
@ATTRIBUTE sepalwidth NUMERIC
@ATTRIBUTE petallength NUMERIC
@ATTRIBUTE petalwidth NUMERIC
@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}
@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,?,1.5,0.2,Iris-setosa
    …
```

Name for this collection of data

Sample ARFF File

```
% Iris Plants Database (from R.A. Fischer and Michael Marshall)
%
@RELATION iris
@ATTRIBUTE sepallength NUMERIC
@ATTRIBUTE sepalwidth NUMERIC
@ATTRIBUTE petallength NUMERIC
@ATTRIBUTE petalwidth NUMERIC
@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}
@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,?,1.5,0.2,Iris-setosa
    …
```

Name and type of each attribute

Sample ARFF File

```
% Iris Plants Database (from R.A. Fischer and Michael Marshall)
%
@RELATION iris
@ATTRIBUTE sepallength NUMERIC
@ATTRIBUTE sepalwidth NUMERIC
@ATTRIBUTE petallength NUMERIC
@ATTRIBUTE petalwidth NUMERIC
@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}
@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,?,1.5,0.2,Iris-setosa
    …
```

Nothing special about this name
But, Weka assumes the *last* attribute is the class unless you tell it otherwise (which you should!)

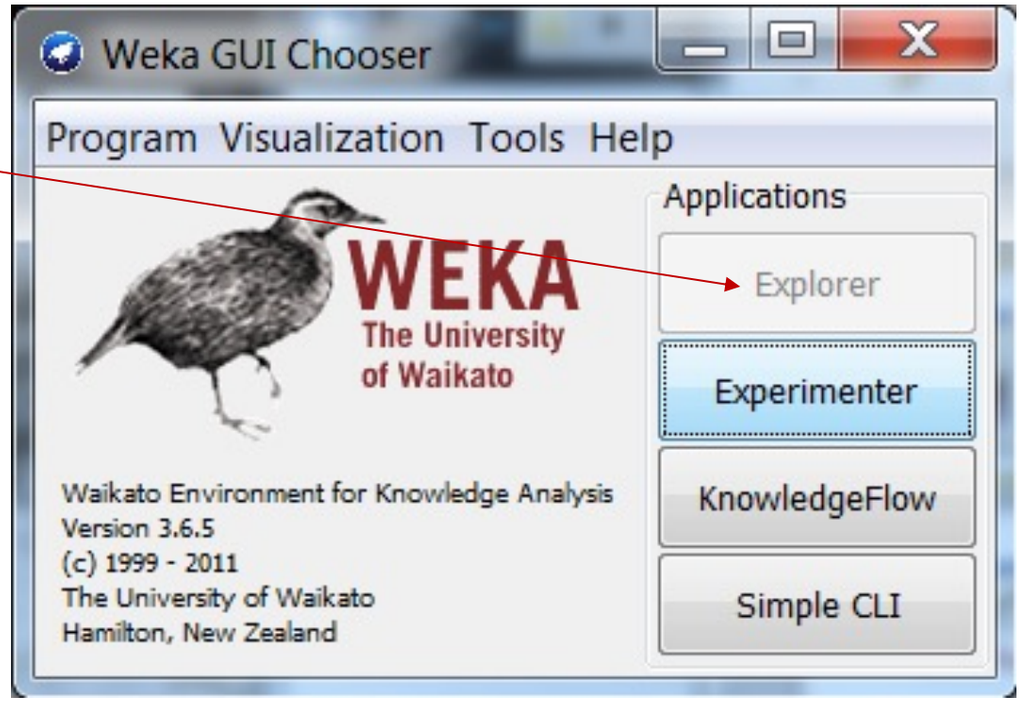
Sample ARFF File

```
% Iris Plants Database (from R.A. Fischer and Michael Marshall)
%
@RELATION iris
@ATTRIBUTE sepallength NUMERIC
@ATTRIBUTE sepalwidth NUMERIC
@ATTRIBUTE petallength NUMERIC
@ATTRIBUTE petalwidth NUMERIC
@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}
@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,?,1.5,0.2,Iris-setosa
      …
```

Data is done as comma separated values

Using Weka

- Start Weka
- Open up the Explorer interface



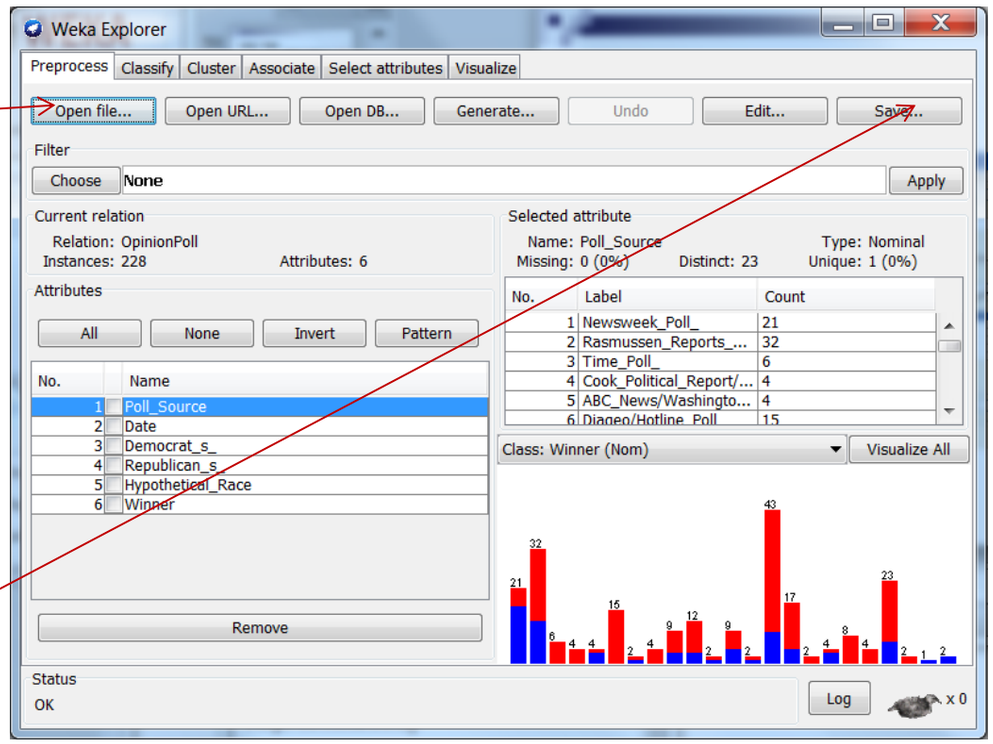
Using Weka

- Click on Open File

Open OpinionPoll.arff
(from Piazza)

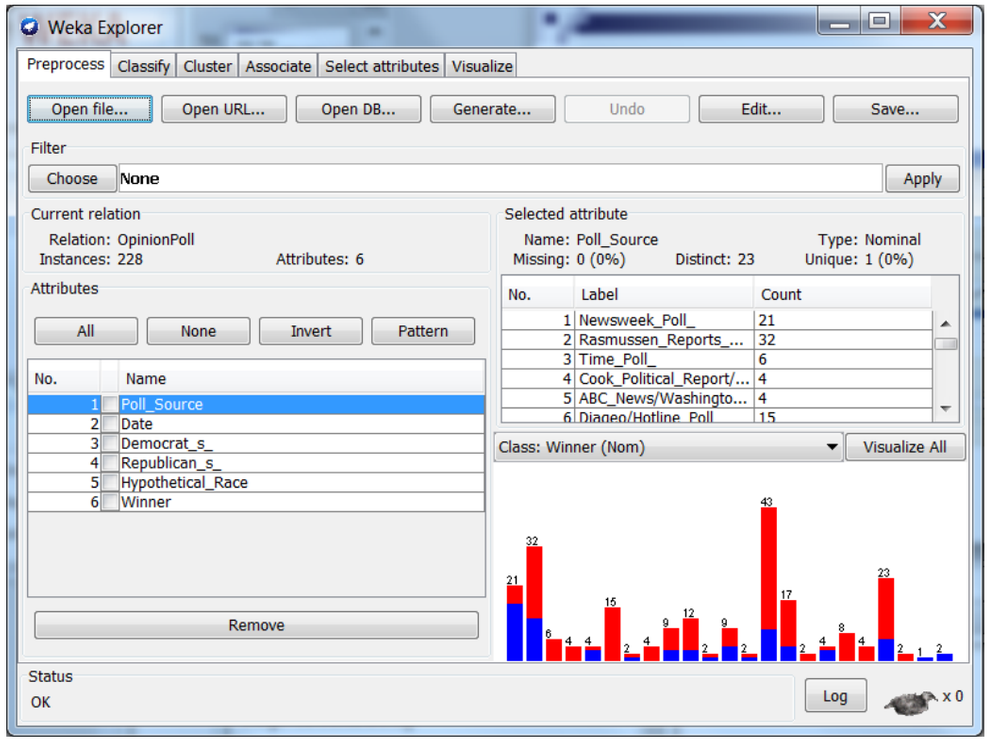
(Can also read
comma separated
value (.csv) files)

You can save
it as a .arff file



Using Weka

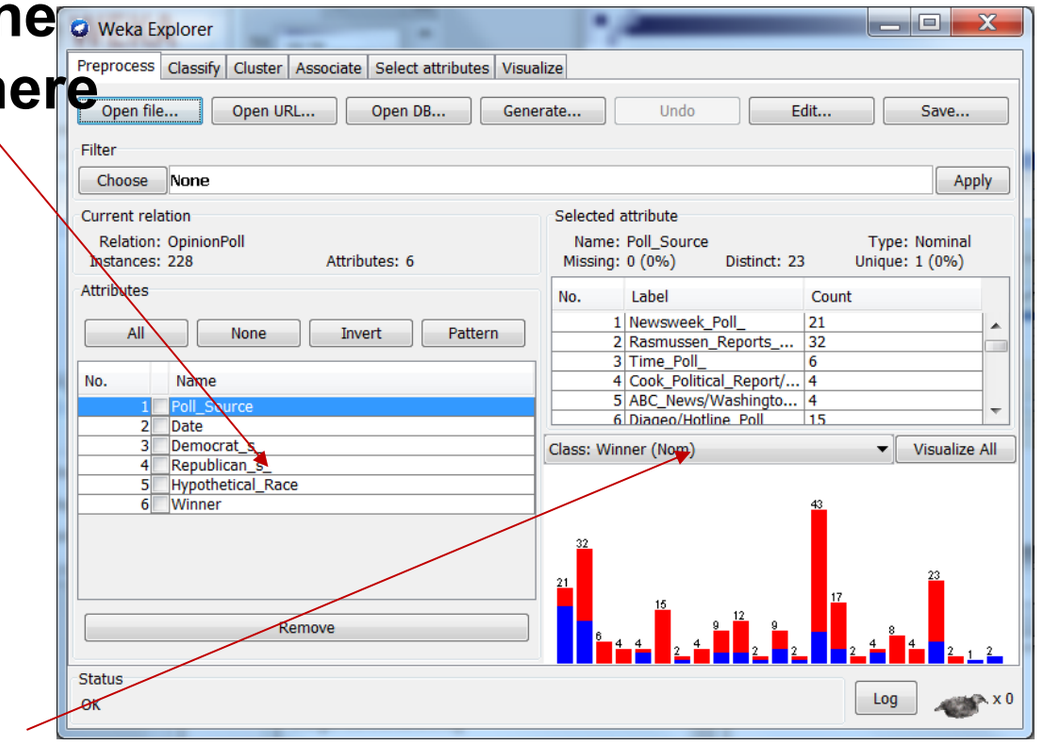
- Summary stats for selected attributes are displayed



Using Weka

- Observe interaction between attributes by selecting in interface

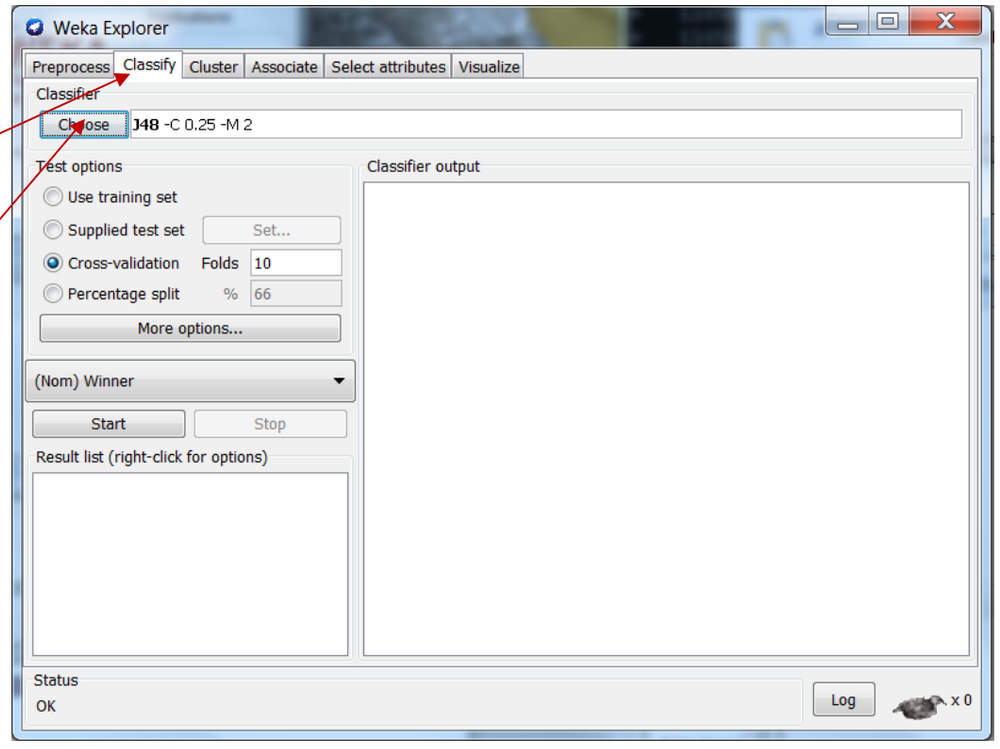
Select one attribute here



Select another attribute here

Using Weka

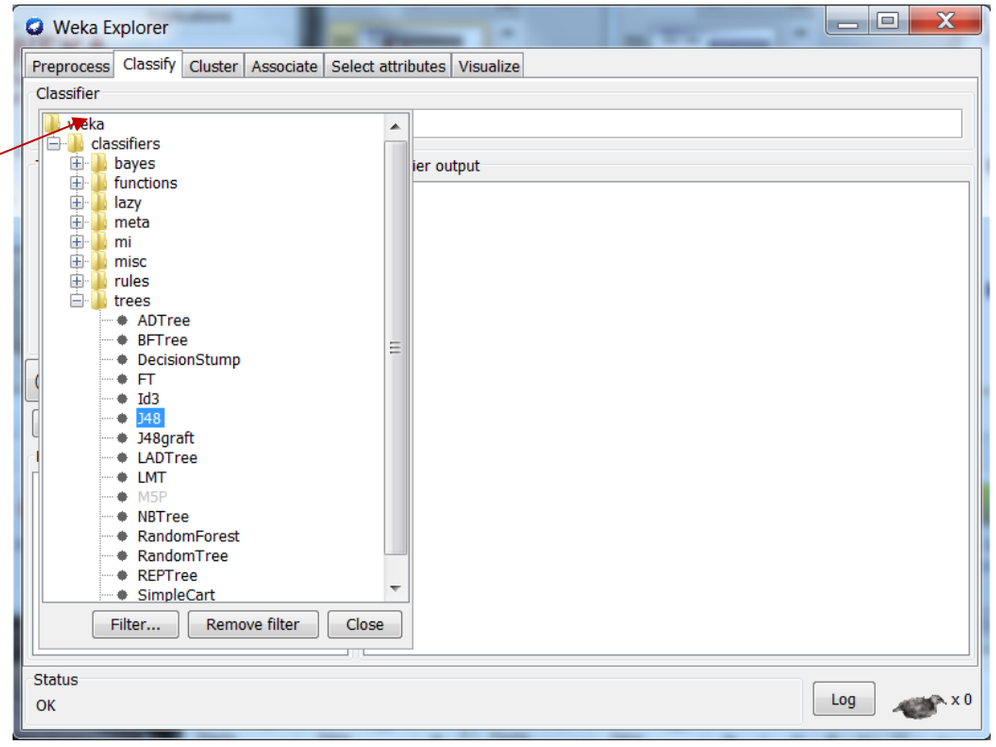
- Go to Classify Panel
- Select a classifier



Using Weka

- Go to Classify Panel

- Select a classifier

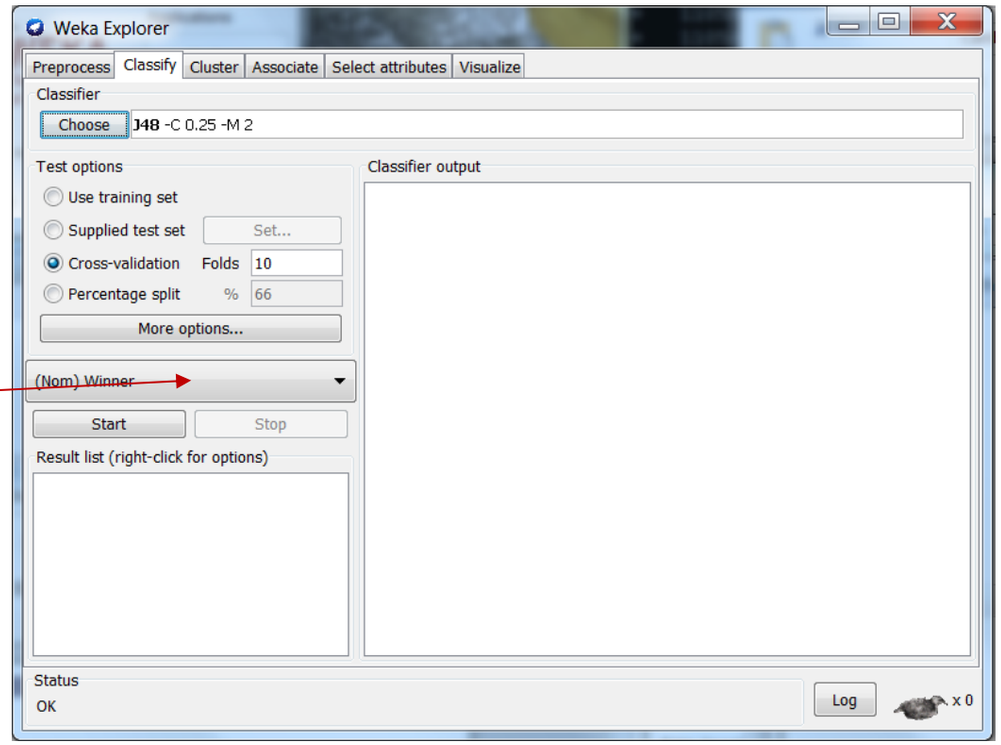


Using Weka

Go to Classify Panel

Select a classifier

Select the predicted class



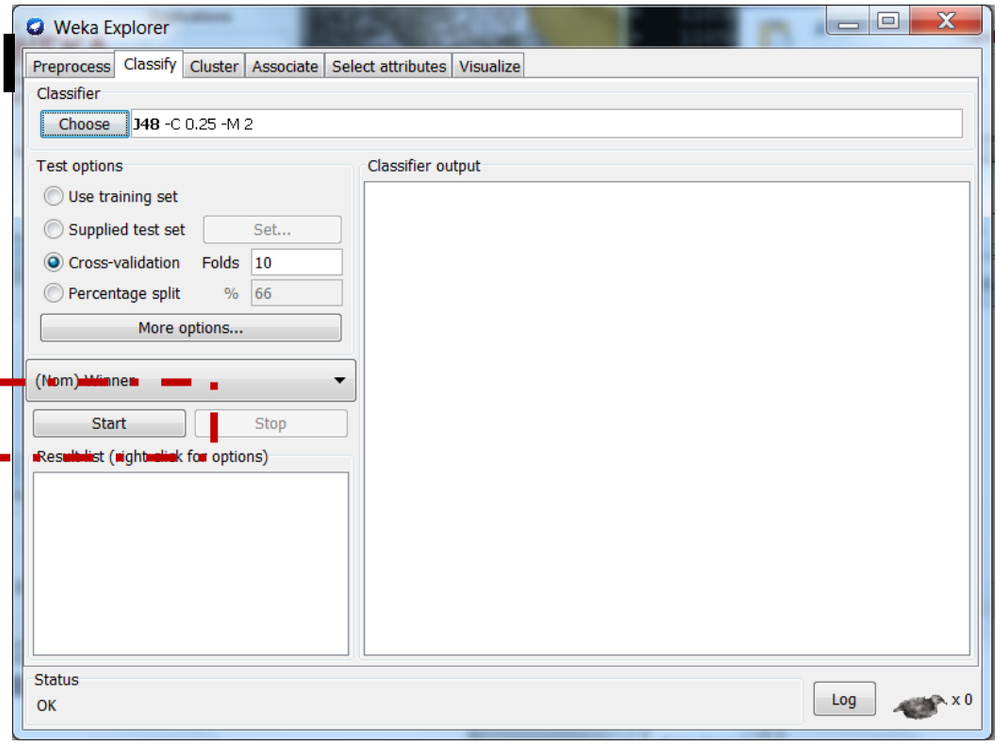
Using Weka

Go to Classify Panel

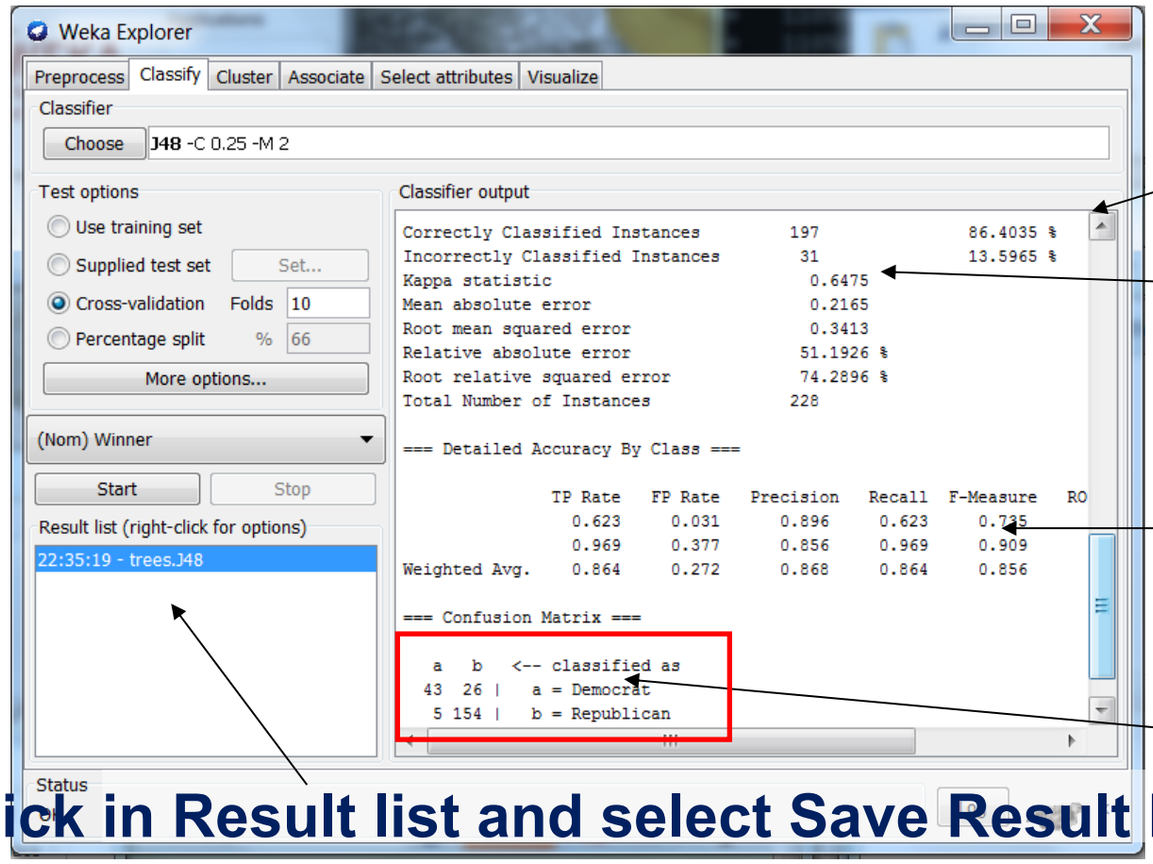
Select a classifier

Select the predicted class

Start the evaluation



Looking at results



Percent correct

Percent correct, controlling for correct by chance

Performance on individual categories

Confusion matrix

Right click in Result list and select Save Result Buffer to save performance stats.

Hint

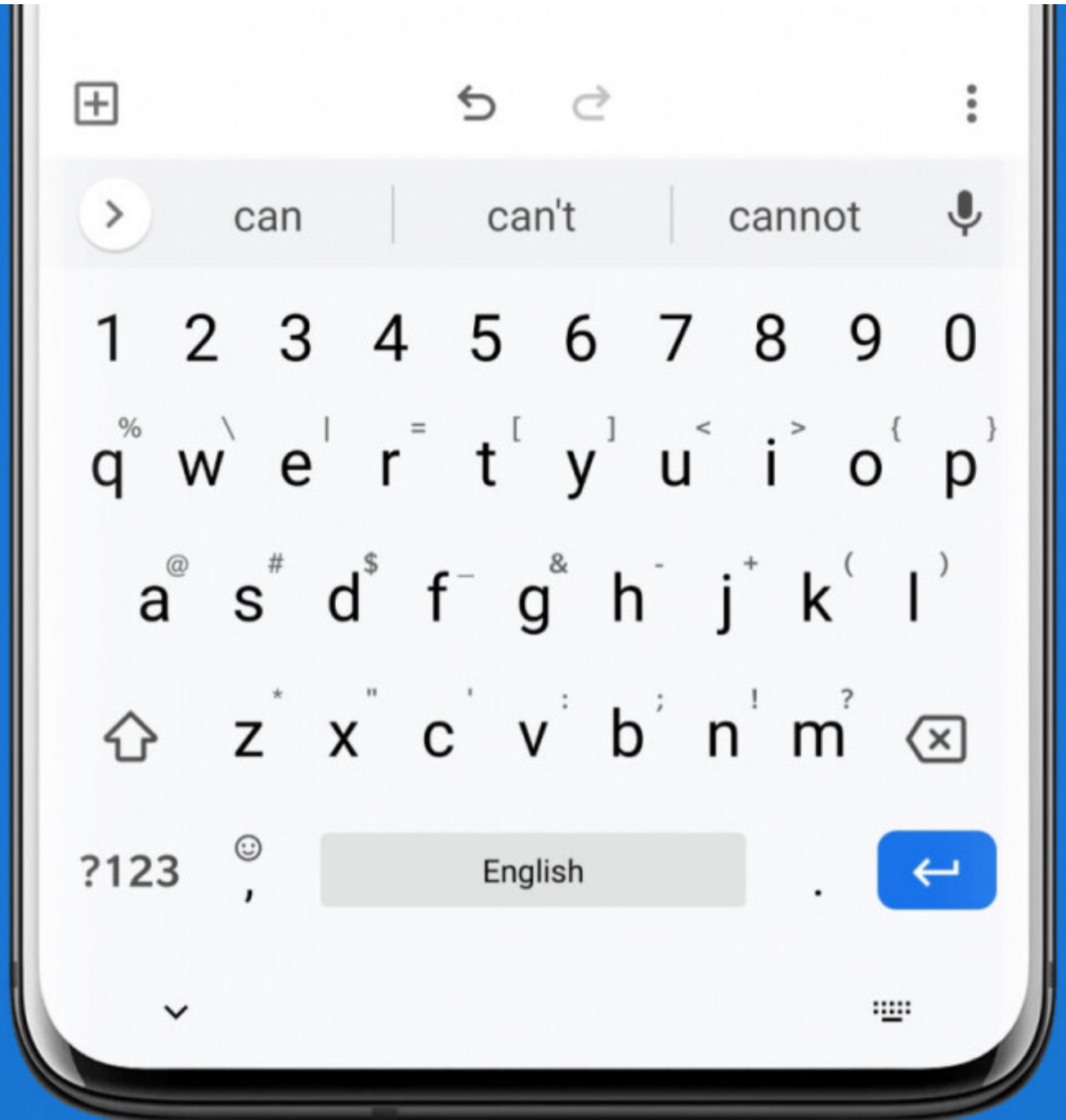
Save results from all your runs

Put additional notes with the textual output!

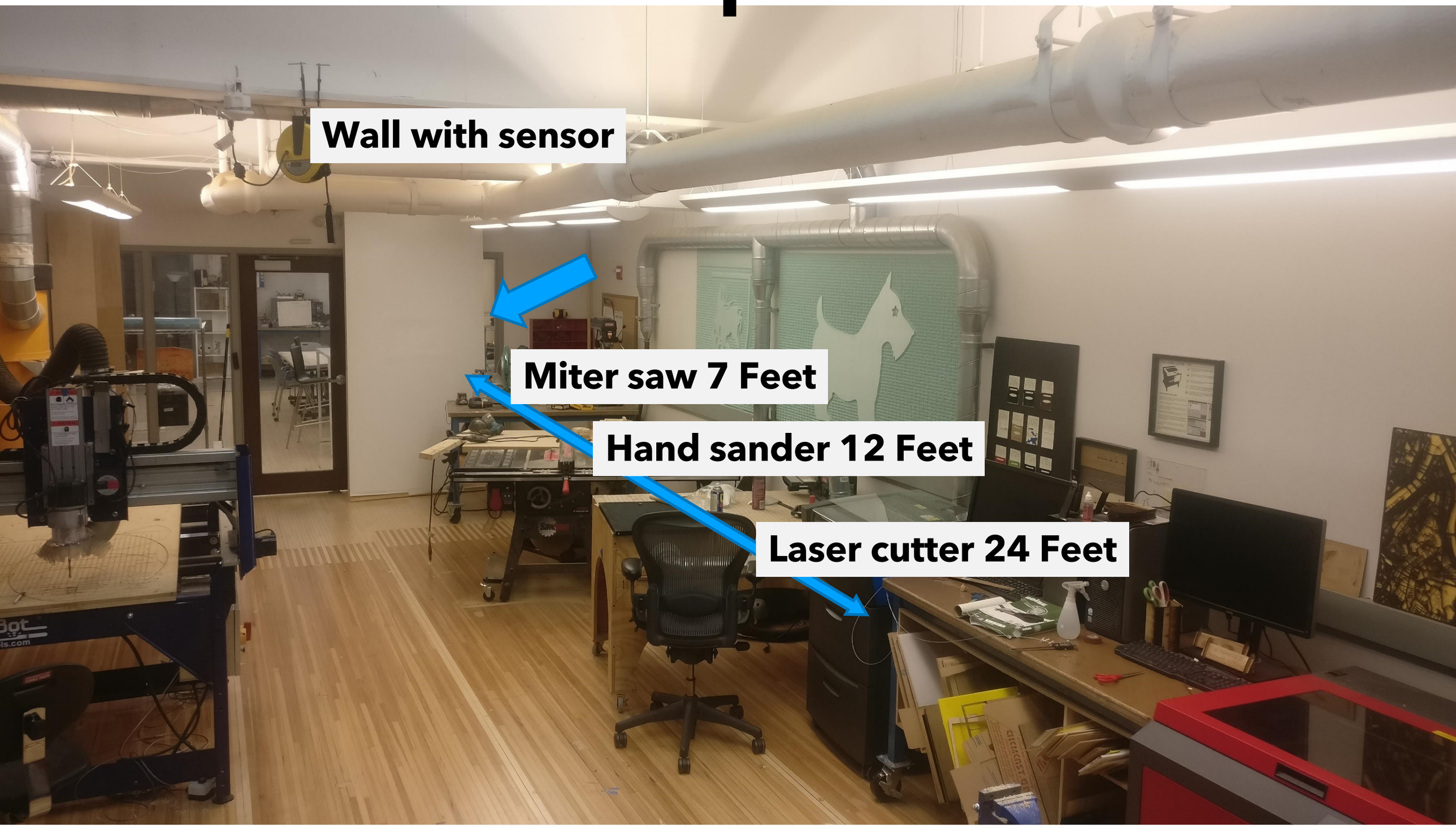
- Not everything you need to reconstruct what you did is in the textual output (can't tell which feature subsets you learned on!)

- You will run a lot of these trying various things and soon forget the details

Example #1


can
can't
cannot
1 2 3 4 5 6 7 8 9 0
q w e r t y u i o p
a s d f g h j k l
z x c v b n m
?123
English

Example #2


Wall with sensor
Miter saw 7 Feet
Hand sander 12 Feet
Laser cutter 24 Feet