

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+

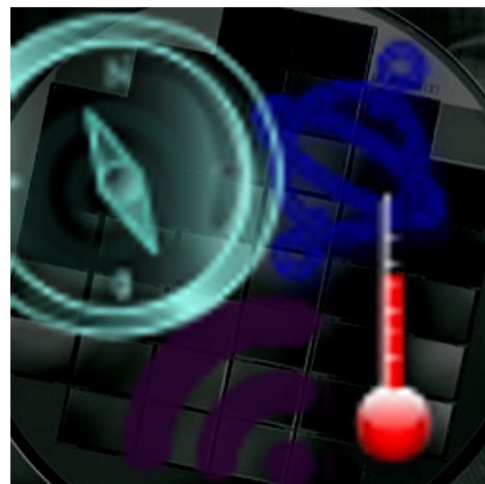
Motion Logger with CSV Export

[Choi Tsz Hei](#)

★★★★★ 5.0 • 10 Ratings

Free

Android




AndroSensor

Fiv Asim Tools

★★★★★ 7,369 

E Everyone

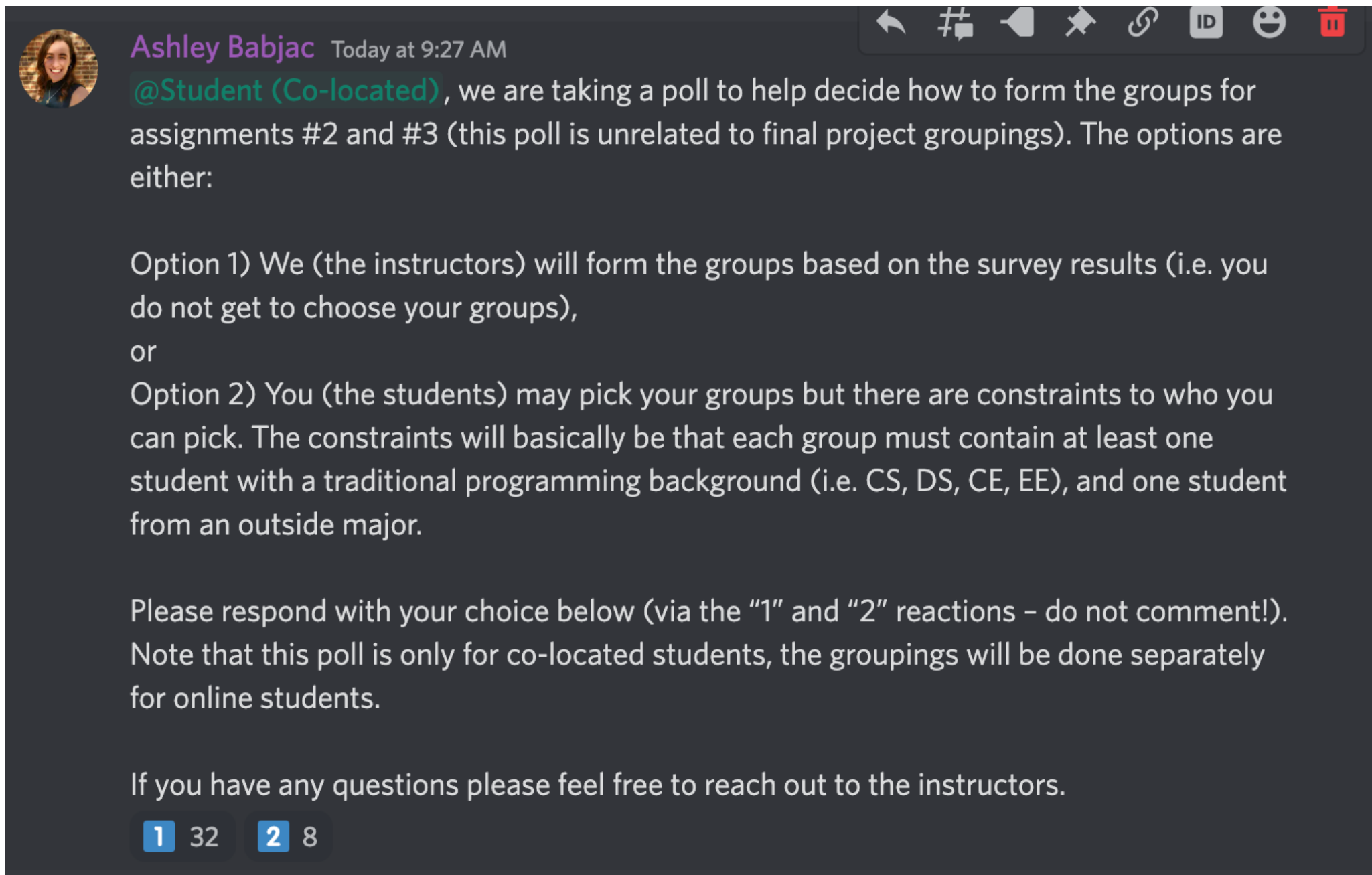
Contains Ads

 You don't have any devices.


You can share this with your family. [Learn more about Family Library.](#)

Installed

Poll on Discord



A screenshot of a Discord message from Ashley Babjac. The message is in a dark-themed window. At the top right of the message area, there is a row of icons: a back arrow, a plus sign with a hash, a speech bubble, a pin, a link, a user icon with 'ID', a smiley face, and a trash can. The message itself starts with a circular profile picture of a woman, followed by the name 'Ashley Babjac' in purple and the timestamp 'Today at 9:27 AM'. The text of the message is in white on a dark background. It mentions '@Student (Co-located)' in green. At the bottom of the message, there are two reaction buttons: a blue square with the number '1' and the count '32', and another blue square with the number '2' and the count '8'.

 **Ashley Babjac** Today at 9:27 AM

@Student (Co-located), we are taking a poll to help decide how to form the groups for assignments #2 and #3 (this poll is unrelated to final project groupings). The options are either:

Option 1) We (the instructors) will form the groups based on the survey results (i.e. you do not get to choose your groups),
or

Option 2) You (the students) may pick your groups but there are constraints to who you can pick. The constraints will basically be that each group must contain at least one student with a traditional programming background (i.e. CS, DS, CE, EE), and one student from an outside major.

Please respond with your choice below (via the "1" and "2" reactions – do not comment!). Note that this poll is only for co-located students, the groupings will be done separately for online students.

If you have any questions please feel free to reach out to the instructors.

1 32 **2** 8

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



WEKA

The workbench for machine learning

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

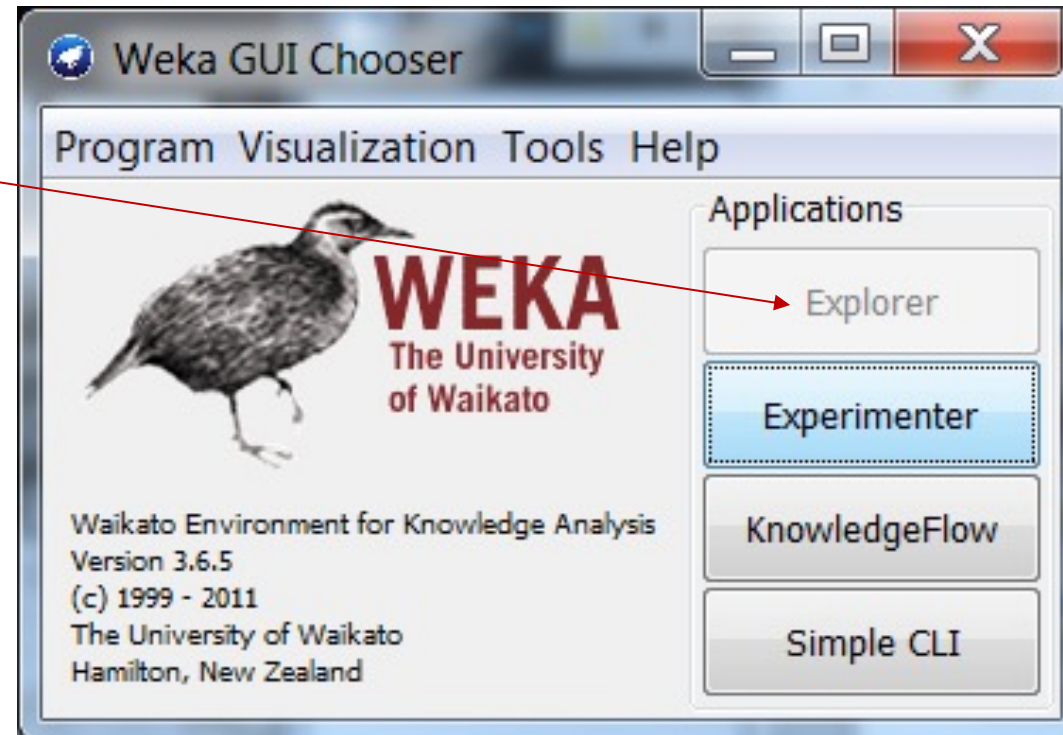
4.6,?,1.5,0.2,Iris-setosa

...

**Data is done as comma
separated values**

Using Weka

- Start Weka
- Open up the Explorer interface



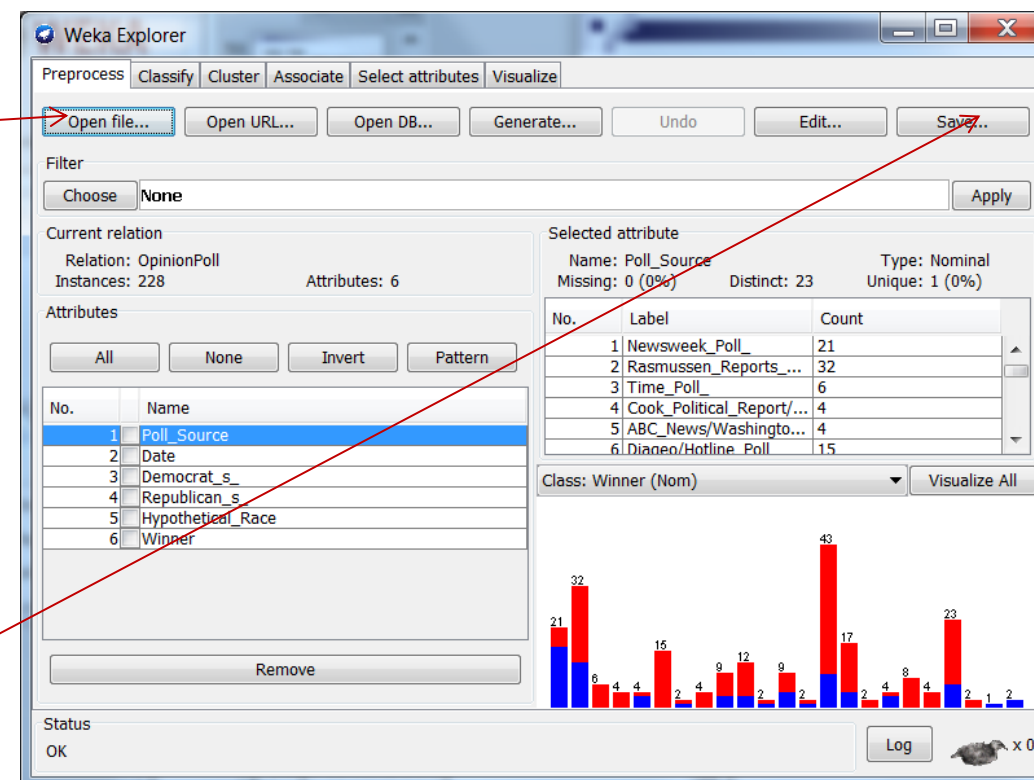
Using Weka

- Click on Open File

Open OpinionPoll.arff
(from Piazza)

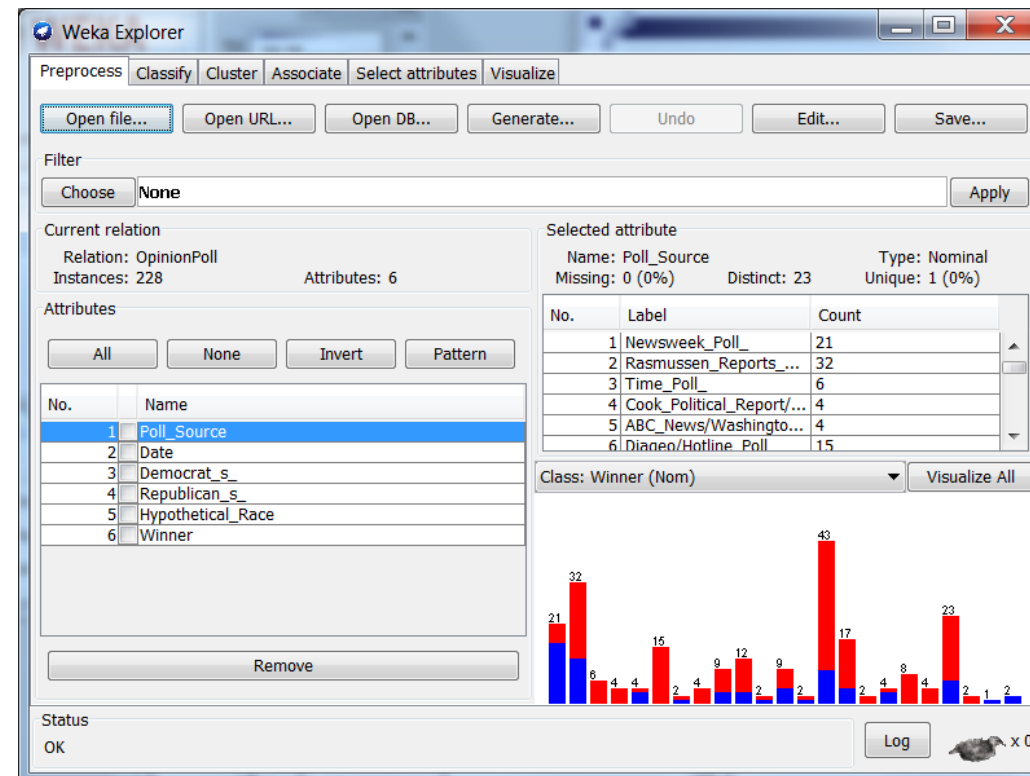
(Can also read
comma separated
value (.csv) files)

You can save
it as a .arff file



Using Weka

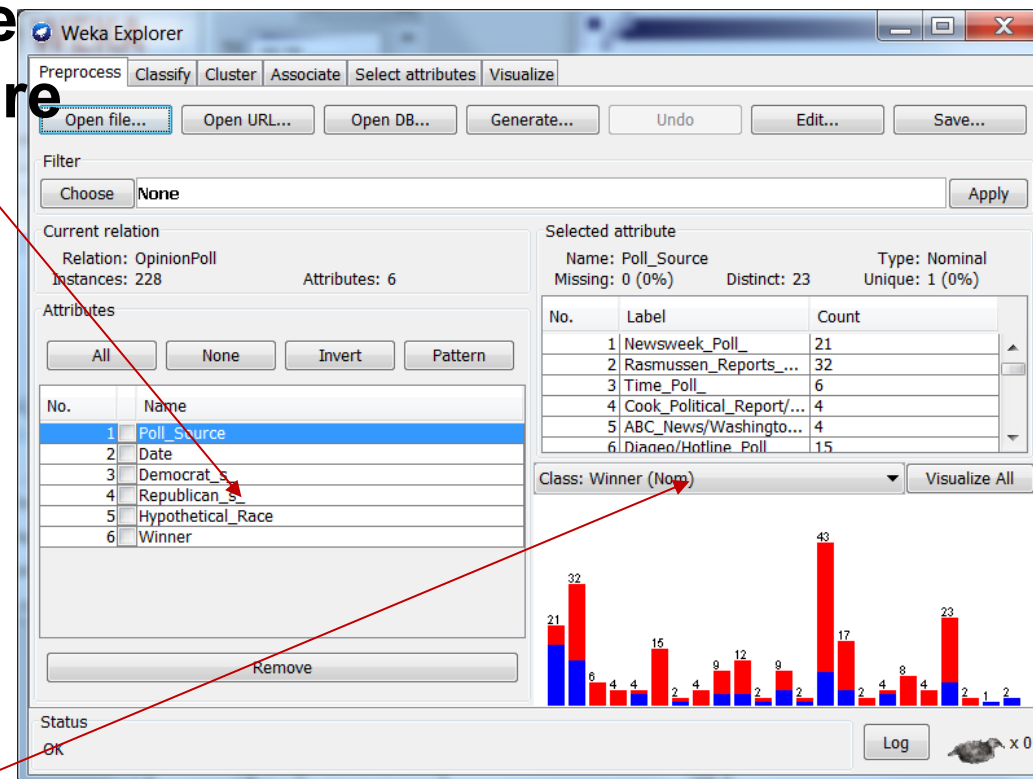
- Summary stats for selected attributes are displayed



Using Weka

- Observe interaction between attributes by selecting in interface

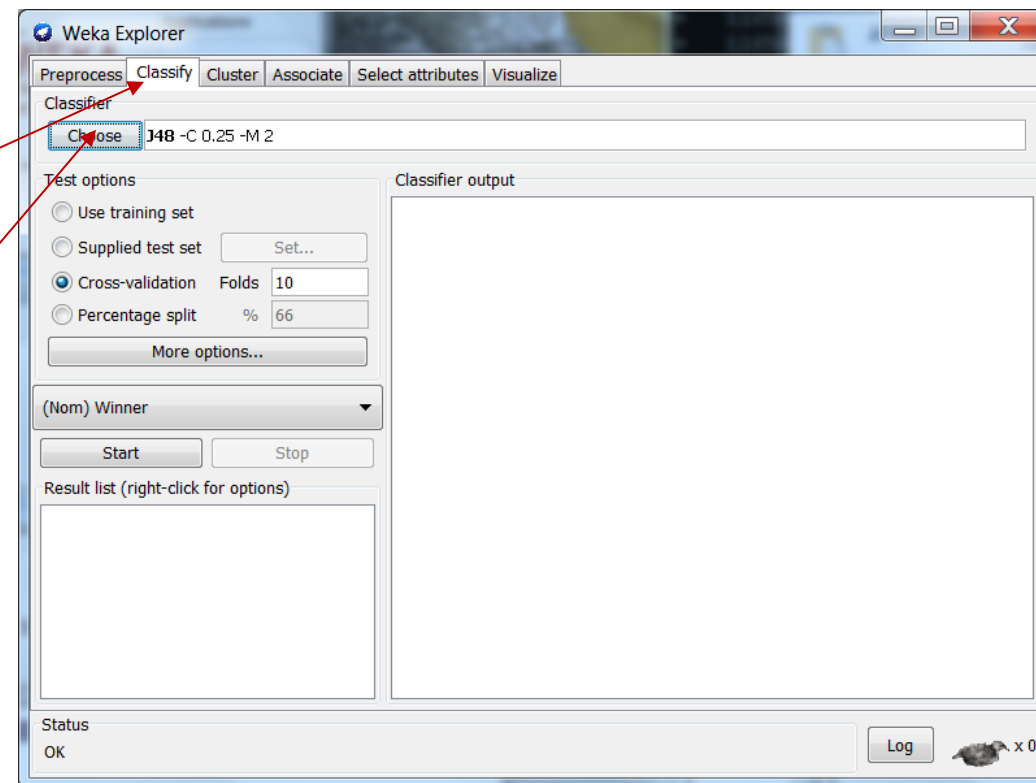
Select one attribute here



Select another attribute here

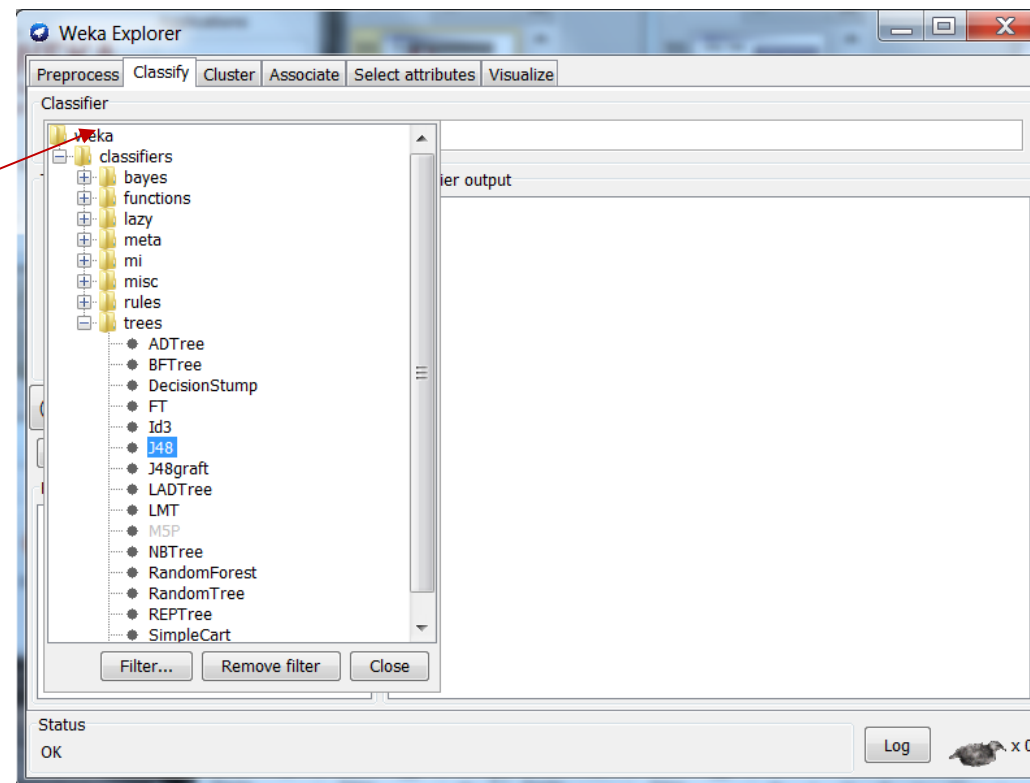
Using Weka

- Go to Classify Panel
- Select a classifier



Using Weka

- Go to Classify Panel
- Select a classifier

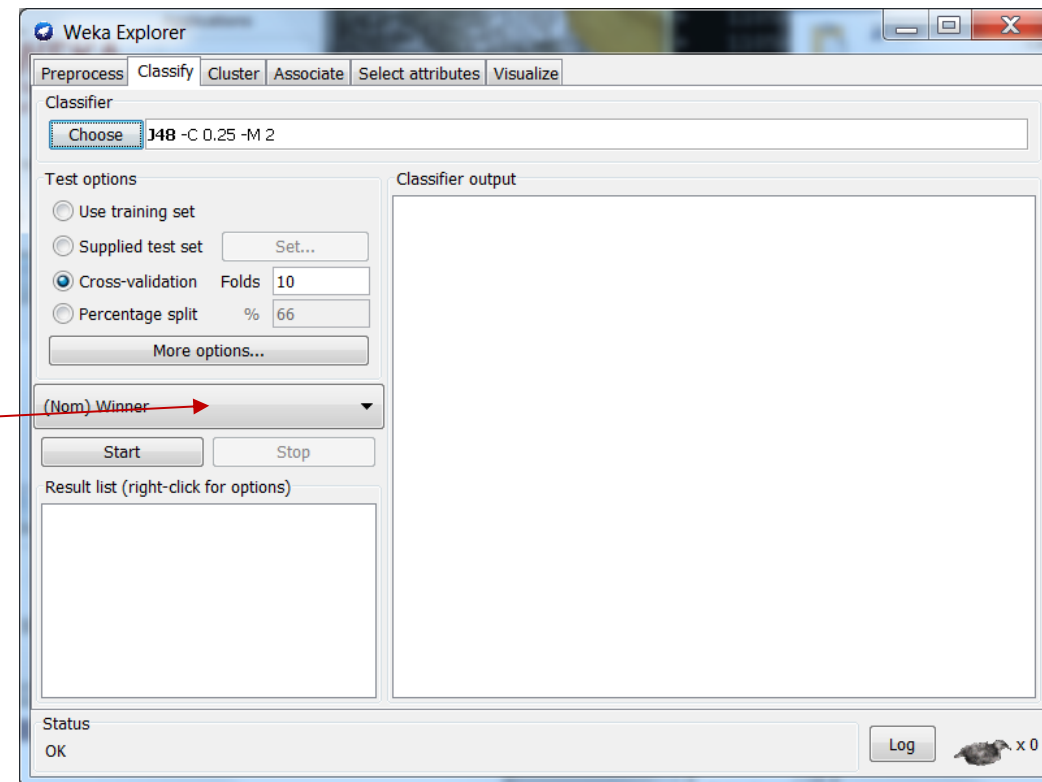


Using Weka

**Go to Classify
Panel**

Select a classifier

**Select the
predicted class**



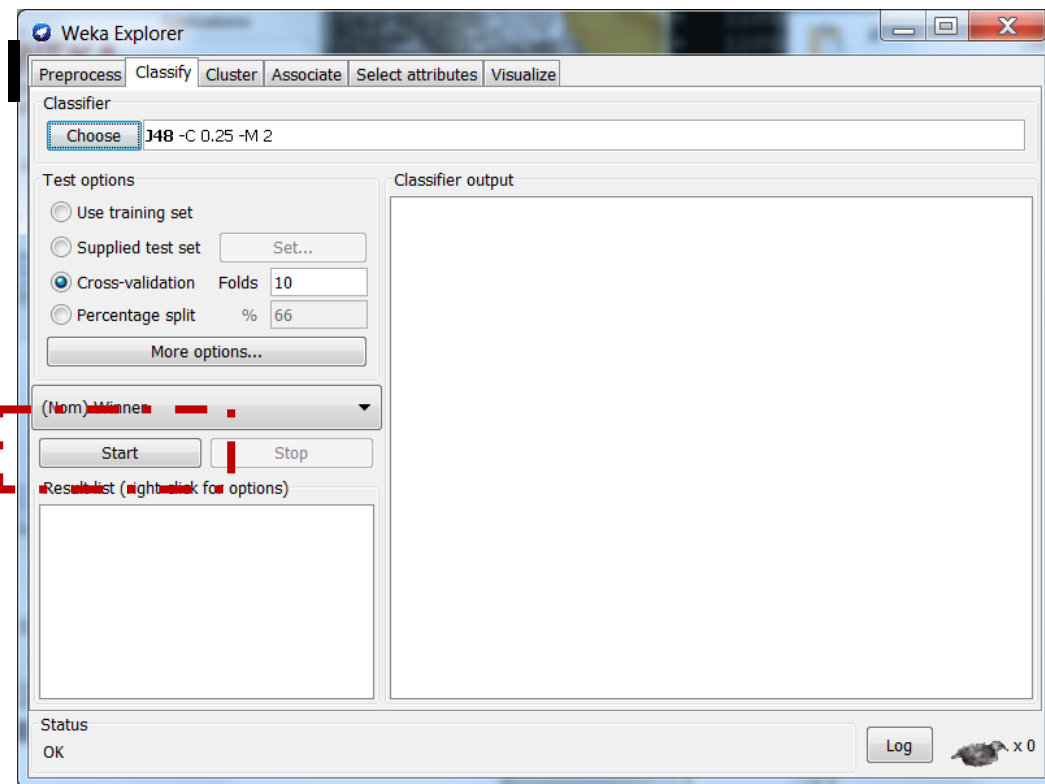
Using Weka

Go to Classify Panel

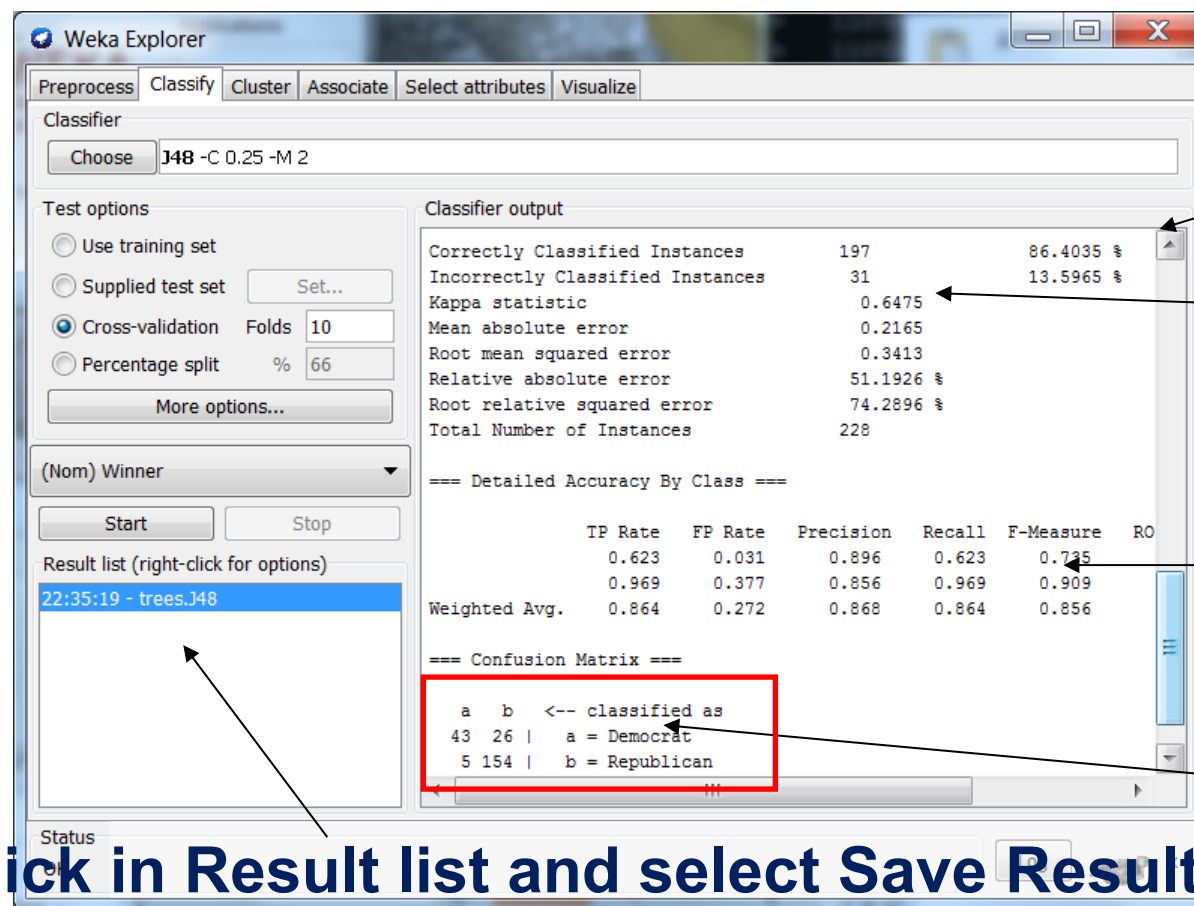
Select a classifier

**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.

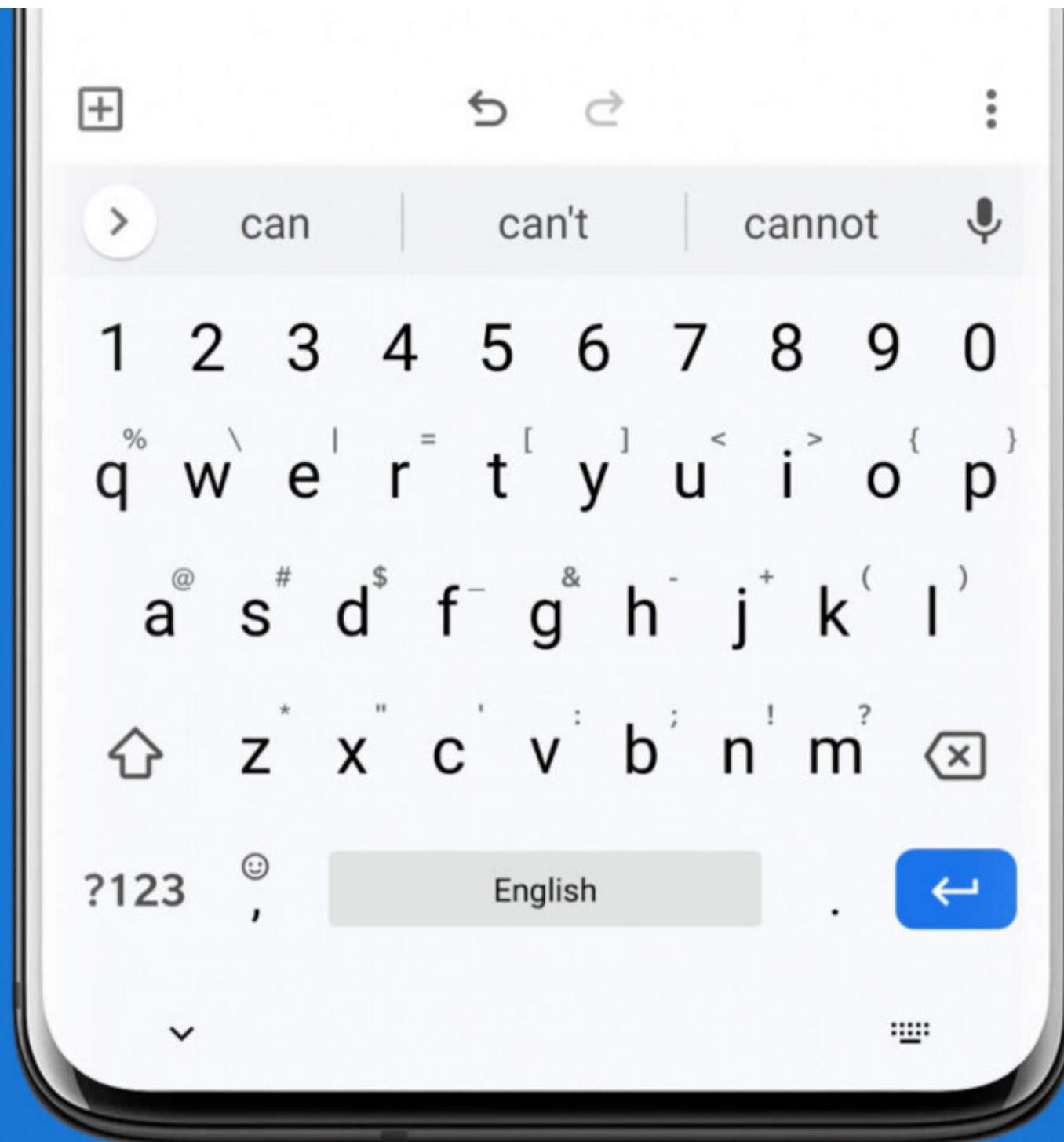
Hint

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



Example #2

