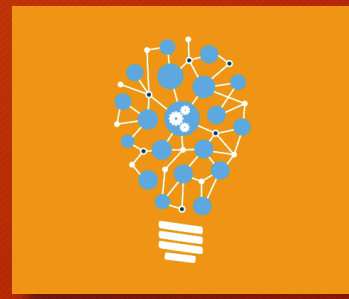
Waikato Environment for Knowledge Analysis

Presentation Prepared By:
Siamul Karim Khan
1105104

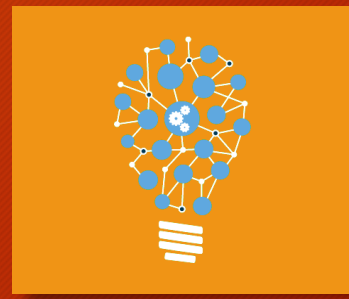
What is WEKA?



- Weka is a collection of machine learning algorithms for data mining tasks. It contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization.
- It is well-suited for developing new machine learning schemes.
- It is open source software issued under the GNU General Public License.

Yes, it is possible to apply Weka to big data!

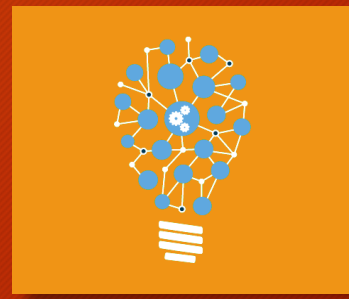
Origin of WEKA



- The name “Weka” was inspired by a flightless bird with an inquisitive nature of the same name which is found only on the islands of New Zealand.
- WEKA was developed by the machine learning group at the University of Waikato; for more details, visit:

http://www.cs.waikato.ac.nz/~ml/weka/contributors.html

How to get WEKA

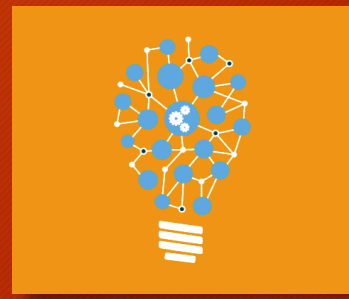


- WEKA is available from:

http://www.cs.waikato.ac.nz/ml/weka

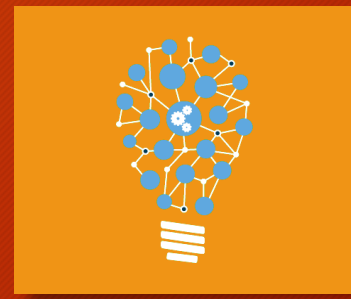
- You can download either a platform-specific installer or an executable Java jar file that you run in the usual way if Java is installed.
- Weka provides a stand-alone software to apply the algorithms directly to data and also provides a API so that it can be used with your own Java code.

Requirements for running WEKA GUI



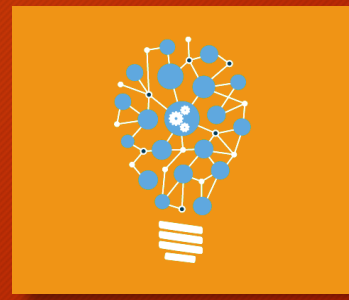
		Java				
		1.4	1.5	1.6	1.7	1.8
WEKA	<3.4.0	X	X	X	X	X
	3.4.x	X	X	X	X	X
	3.5.x	3.5.0-3.5.2	>3.5.2 r2892, 20/02/2006	X	X	X
	3.6.x		X	X	X	X
	3.7.x		3.7.0	>3.7.0 r5678, 25/06/2009	>3.7.13	X

Getting started with WEKA GUI



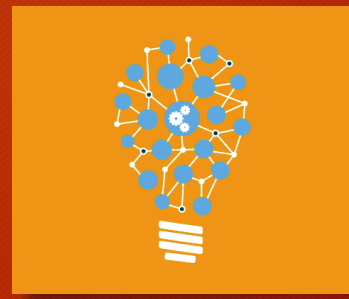
- Start the Weka GUI. The Weka GUI Chooser lets you choose one of the Explorer, Experimenter, KnowledgeExplorer and the Simple CLI (command line interface).

Getting started with WEKA GUI



- WEKA provides a number of small common machine learning datasets that you can use to practice on.
- To select one of the given datasets, click the "***Open file…***" button to open a data set and double click on the "***data***" directory. You will find the datasets inside. You can browse to other directories and load your desired dataset.

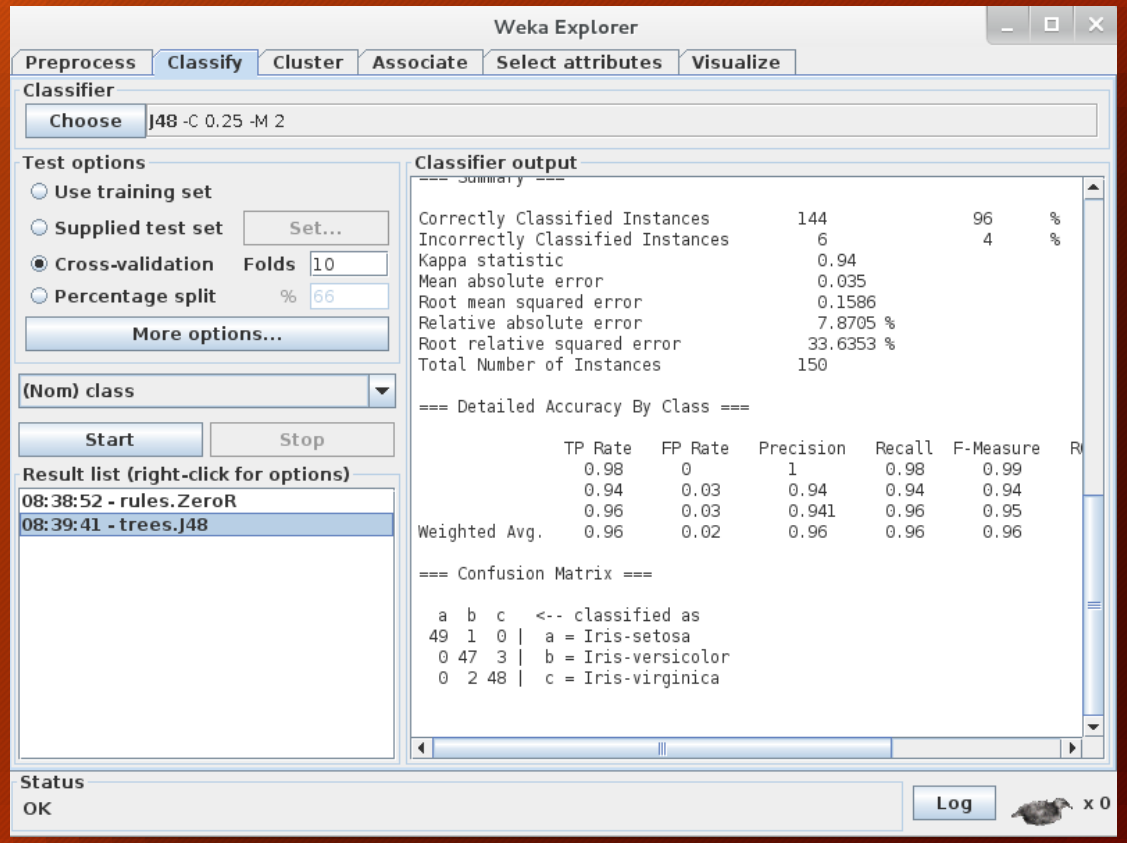
Getting started with WEKA GUI



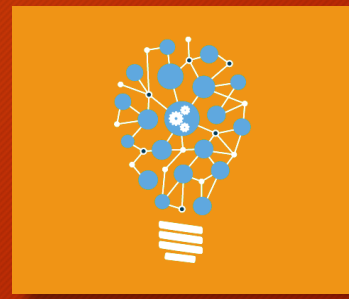
- Once you have loaded a dataset, it's time to choose a machine learning algorithm to model the problem and make predictions.
- Click the "***Classify***" tab. This is the area for running algorithms against a loaded dataset in Weka.
- Click the "Choose" button in the "Classifier" section. Browse and choose the appropriate algorithm.
- Click the "***Start***" button to run the algorithm.

Getting started with WEKA GUI

The image shows WEKA J48 algorithm results on the iris flower dataset.

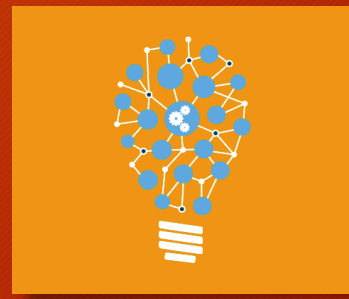


Using WEKA in your Java code



- For including the WEKA library into your java project, you need the JAR file.
- In the download page, download the ZIP archive under "***Other Platforms (Linux, etc.)***"
- Inside the zip there is a file named "***weka.jar***". Include this file as library in your java project to use the weka library with your own java code.

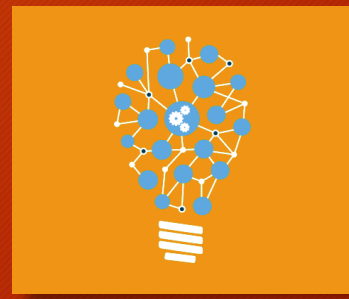
Using WEKA in your Java code



- For info on the different classes available in the library and how to use them, visit: **http://weka.sourceforge.net/doc.stable/**

- **https://weka.wikispaces.com/** is an excellent place to search for what you need; it provides explanations and code snippets for using WEKA.

Learning to use WEKA



- There is also an online appendix on the WEKA software available at:

http://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf

- There is a tutorial series on youtube, hosted by **Rushdi Shams**. (Search for "***Weka Tutorial Rushdi Shams***" in youtube).