Machine Learning For Kids :: Teachers' notes

Worksheet	Phishing
Activity	Create a Python program that can predict if a URL is legitimate
Objective	**Teach a computer to recognize phishing web links** • How computers can be trained to make predictions based on experiences. • How feature selection is choosing what values the computer should learn from • How machine learning systems can be visualised using tree diagrams • How machine learning is used to recognize malicious or suspicious web pages
Difficulty level	Advanced
Time estimate	45 minutes - 1 hour
Summary	Students will train a machine learning model to predict if a URL is for a legitimate webpage or a phishing page. They use this in Python to test new URLs. The project ends with students reviewing AI research papers, which they should be able to understand at a high level if they've followed the project carefully.
Topics	supervised learning, decision trees, feature selection

Setup

Each student will need:

Print-outs	Project worksheet (download from https://machinelearningforkids.co.uk/worksheets)
Access	Username and password for machinelearningforkids.co.uk

Class account will need:

API keys	None

Customizing

If you use **PRIMM** approaches with your class, add a step where students predict how the project template works.
If you want to **increase the amount of coding** involved, delete some of the code from the project template and add steps to the worksheet so students code it themselves.
If you want to **encourage problem solving**, delete some of the detail in the worksheets and provide more general instructions instead.
Worksheets in MS Word format are available so you can **modify them to suit your class**.

Worksheet	https://github.com/IBM/taxinomitis-docs/tree/master/project-worksheets/msword

Help

Potential issues	• The sample code is available on GitHub but you might find it easier to provide the code for your students • The sample code has been tested with Python 3 and will need some modifying if you wish to use Python 2 • The sample code needs third-party libraries **requests, dateutil, whois**. You might find it useful to get these installed first. (e.g. `pip3 install -r requirements.txt`) General troubleshooting and help at https://machinelearningforkids.co.uk/help