**DT + Health and Physical Education**

**Title**

Anti-bullying AI

**Years:**

5–6; 7–8

**Key Terms:**

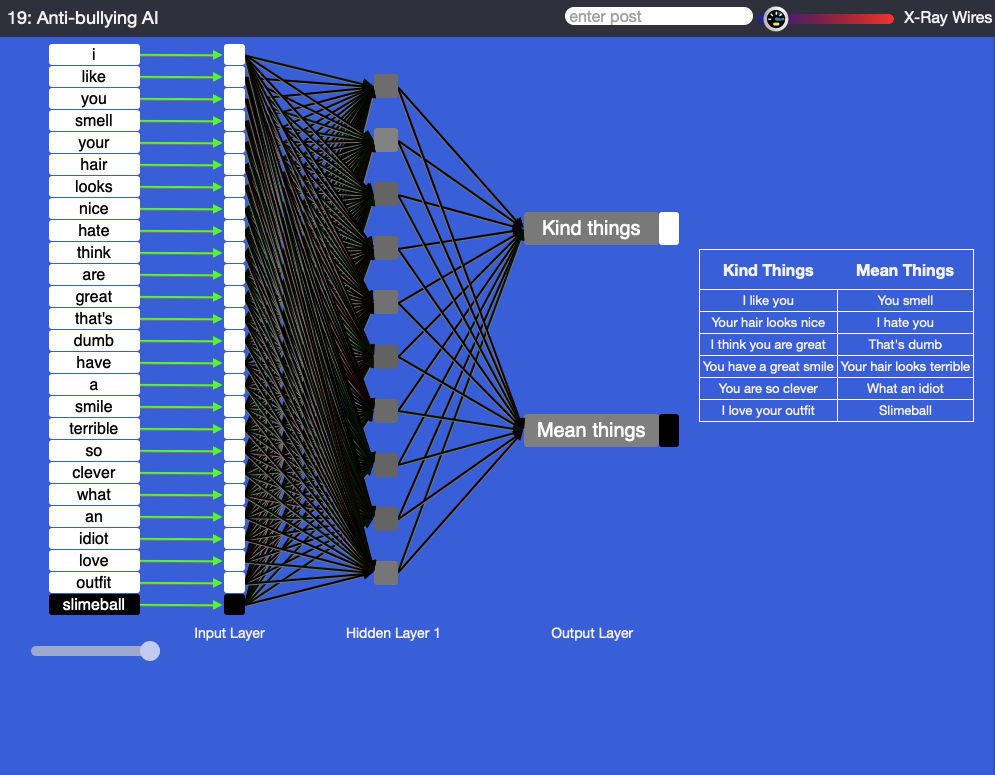
|  |  |
| --- | --- |
| AI | Artificial Intelligence |
| ML | Machine Learning |
| ANN | Artificial Neural Network |

**Summary**

Sometimes we write and post things on social media in a hurry. Such posts can hurt people and even make them feel bullied. Wouldn't it be great if an Artificial Intelligence application could check our posts as we write them, and warn us if they were potentially hurtful?

This lesson was developed by the [Digital Technologies Institute](https://www.digital-technologies.institute/) in collaboration with DT Hub.

# Preliminary notes

[MyComputerbrain](https://mycomputerbrain.net/php/experiments/ai.experiment19b.php) is the tool used in the plugged activity below. You can use this link: <https://mycomputerbrain.net/php/experiments/ai.experiment19b.php> to access this program.

The above image shows the view, consisting of an ANN and a table with expressions that most people would consider either ‘kind’ (left-hand column) or ‘mean’ (right-hand column).

The grey boxes in the input, hidden and output layer are called ‘perceptrons’ and can be likened to neurons in the brain. On the far left-hand side of the ANN, we see all words from the ‘kind’ and ‘mean’ expressions in the table. They are provided to the ANN as learning (or training) data.

# Suggested steps

### Unplugged activity

1. Use a suitable hook to discuss bullying.

* What is bullying? Have you been bullied?
* Why is it important to recognise bullying?
* What effect does bullying have on someone?
* Are cases of bullying always intentional, or can you think of instances where they may not be?
* In daily conversation, what words or phrases might be considered ‘mean’ and potentially a form of bullying? You may wish to create a table on the whiteboard separating out nice and mean words/phrases—you can keep this list up for the plugged activity later.

1. Students increasingly use social media messaging to communicate with their friends.

* What are the social media communication tools that students use?
* In an effort to message back quickly, what do you think might occur? Have you made spelling or grammatical mistakes, or has your message been misunderstood?
* How might sarcasm or irony be interpreted as, or lead to, bullying?

**Socio/emotional differences**

Students who have experienced bullying in the past may require a level of support to engage with this activity. For example, you may decide to pair affected students with a trusted partner.

### Plugged activity

1. Explain that computers can be programmed to be intelligent or at least smart. Ask the class if they think a computer could work out whether something they write is kind or mean? How would it do this?
2. Use the application [My Computer brain](https://mycomputerbrain.net/php/experiments/ai.experiment19b.php), which lets you train an AI application to recognise posts and match them to a particular output. No coding is required. We have provided an example of what it might look like in the image above. Students can enter a total of 12 expressions, including single words and whole sentences.
3. The instructions in the right-hand menu of the application will guide students through the process, which consists of:
   * creating training data
   * training the ANN via the ‘Start Learning’ button
   * testing the ANN by entering short sentences in the input field at the top of the window.

**Additional Scaffolding**

Students may need help with spelling and grammar.

1. As a special feature, as a special feature, the ANN will be constructed in real time, while students enter the training data. This is intended to help students understand how the input data influences the topology of the ANN. It is also an opportunity for the teacher to discuss data input, and more general human sensory perception.

Note: The student input is not stored. When the browser window is reloaded, the experiment is reset. This is to ensure that especially mean expressions are not used for purposes other than the experiment.

1. Share what you have learned about AI and how ‘intelligent’ a computer can be.
2. Students may come up with their own training data and project to test the AI further. Here are some prompts:

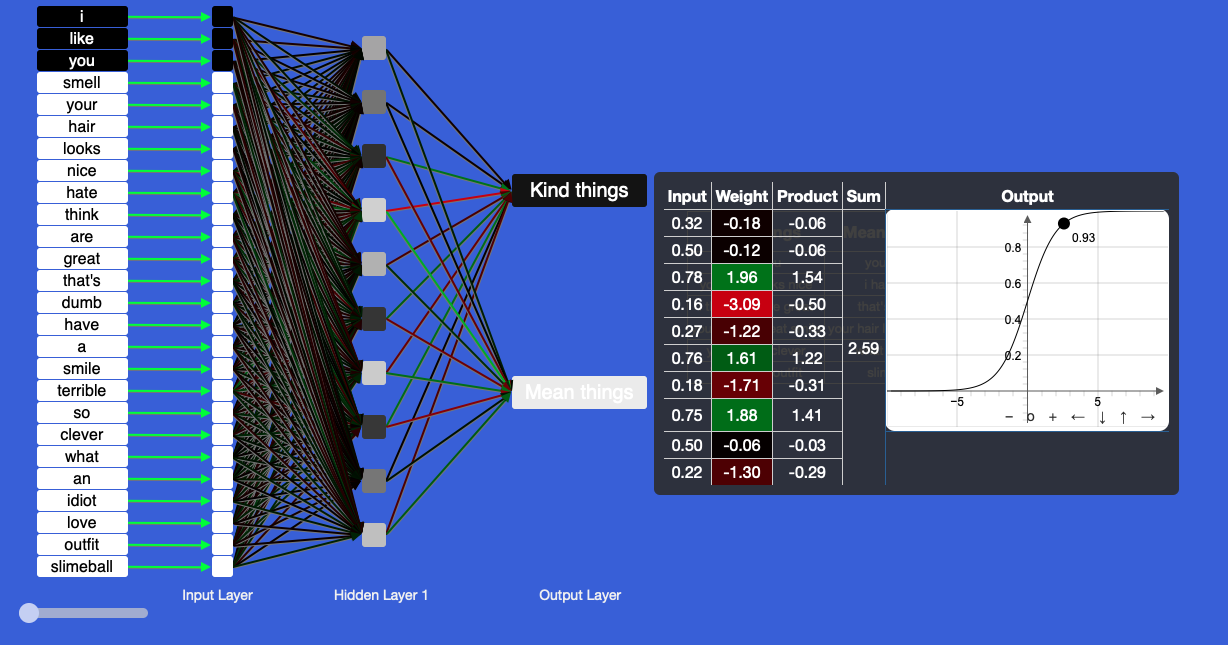
* How will the AI decide when kind and mean words are used in combination?
* Can the AI pick up sarcasm or irony?
* How could the AI be improved (eg by adding a thesaurus that could recognise similar words)?

 **Extension Activity**

For students who finish early, they could choose their own topic (e.g. ‘healthy vs unhealthy ingredients’ or ‘formal vs informal emails’).

# Discussion

1. In the example below, the AI is 0.93 sure that ‘I like you’ is a ‘kind thing’. Discuss the scale used in the ‘Kind things’ and ‘Mean things’ boxes to display the confidence level. A fully black box means the AI was very sure. A light grey box means it was not quite sure. When was the AI very sure of its guess? When was it not so sure?



*Image: confidence level*

1. Look at the table on the right-hand side with ‘Input’, ‘Weight’, ‘Product’, ‘Sum’ and ‘Output’. Each perceptron receives Input data from other perceptrons. Each Input is multiplied by a Weight, leading to a Product. The Products are added up to create a Sum, and then pushed through a function that determines the Output of the perceptron. Note that the mathematics is quite effective, despite its relative ease. Hover your mouse over each perceptron to check its data. Observe that the colour of the lines that connect to the Input of a perceptron correspond to the colours in the Weight column of the table.
2. Observe how the colours of the lines change as the AI learns. Green means positive values, red means negative values, and black means neutral values. These lines mimic the synapses in our brain and regulate the dataflow between the perceptrons.
3. Is green more valuable to the AI’s learning mechanism than red or black? No. They are all equally required to achieve the learning outcome.
4. Note that the words from the training data are provided to the ANN as binary numbers. The ANN’s internal representation[[1]](#footnote-1) of the data is 1’s and 0's.

**Note:** The function depicted above is a Sigmoid function: f(x)= 1 / (1 + e^(-x)). The ANN also uses the following derivative of the Sigmoid function during the learning process: df(x)/dx = f(x) \* (1 - f(x)).

# Why is this relevant?

AI refers to the ability of machines to mimic human capabilities in a way that we would consider intelligent.

Machine learning is an application of AI. With ML, we give the machine lots of examples of data, demonstrating what we would like it to do so that it can figure out how to achieve the goal on its own. The machine learns and adapts its strategy to achieve this goal.

In our example, we are feeding the machine with data in the form of words. The more varied the data we provide, the more likely the AI system is to correctly classify the input as an appropriate output. In ML, the system will give a confidence value of how sure it is of the classification it has provided; in this case:

* a decimal value between 0 and 1
* white, black and shades of grey to the output box.

This lesson focuses on the concept of ‘classification’, which, in the context of AI, is a learning technique used to group data based on attributes or features.

 **Extension Activity**

To learn more about ANNs, explore the other experiments in this course at https://mycomputerbrain.net/php/courses/ai.php. The first experiment is free and provides a good introduction to ANNs.

# Assessment

## Teacher assessment

Choose the task that will best suit your students from these suggested assessment approaches.

|  |  |
| --- | --- |
| Possible tasks | Relevant content descriptor(s) |
| Draw a labelled diagram of the AI in action. Identify the perceptrons and their inputs and outputs:   * Describe the components of the ANN. * Where is the learning stored? * What is the general function of a perceptron? | [ACTDIK014](https://www.australiancurriculum.edu.au/Search/?q=ACTDIK014)/[ACTDIK023](https://www.australiancurriculum.edu.au/Search/?q=ACTDIK023) |
| How does the ANN use numbers and binary data to represent information?  Represent the learning data as a binary matrix. | [ACTDIK015](https://www.australiancurriculum.edu.au/Search/?q=ACTDIK015)/[ACTDIK024](https://www.australiancurriculum.edu.au/Search/?q=ACTDIK024) |
| Describe the steps you used to train your AI to recognise posts.  Imagine how the AI works internally and represent the basic function as pseudo-code or a flowchart. | [ACTDIP017](https://www.australiancurriculum.edu.au/Search/?q=ACTDIP017)/[ACTDIP027](https://www.australiancurriculum.edu.au/Search/?q=ACTDIP027) |
| How might this type of AI be used in our daily lives?  What similar application can you envisage using this kind of classification technology for? | [ACTDIP021](https://www.australiancurriculum.edu.au/Search/?q=ACTDIP021)/[ACTDIP031](https://www.australiancurriculum.edu.au/Search/?q=ACTDIP031) |

# Australian Curriculum alignment

## Technologies – Digital Technologies

**Years 5–6**

Examine the main components of common digital systems and how they may connect together to form networks to transmit data (ACTDIK014)

Examine how whole numbers are used to represent all data in digital systems (ACTDIK015)

Define problems in terms of data and functional requirements drawing on previously solved problems (ACTDIP017)

Explain how student solutions and existing information systems are sustainable and meet current and future local community needs (ACTDIP021)

**Years 7–8**

Investigate how data is transmitted and secured in wired, wireless and mobile networks, and how the specifications affect performance (ACTDIK023)

Investigate how digital systems represent text, image and audio data in binary (ACTDIK024)

Define and decompose real-world problems taking into account functional requirements and economic, environmental, social, technical and usability constraints (ACTDIP027)

Evaluate how student solutions and existing information systems meet needs, are innovative, and take account of future risks and sustainability (ACTDIP031)

## Health and Physical Education

|  |  |  |
| --- | --- | --- |
| **Years 5–6** | **Communicating and interacting for health and wellbeing** | Practise skills to establish and manage relationships (ACPPS055)  Examine the influence of emotional responses on behaviour and relationships (ACPPS056)  Recognise how media and important people in the community influence personal attitudes, beliefs, decisions and behaviours (ACPPS057) |
| **Years 7–8** | Investigate the benefits of relationships and examine their impact on their own and others’ health and wellbeing (ACPPS074)  Analyse factors that influence emotions, and develop strategies to demonstrate empathy and sensitivity (ACPPS075) |

## ICT capability

Typically, by the end of year 6, students:

* **Apply personal security protocols**

Identify the risks to identity, privacy and emotional safety for themselves when using ICT and apply generally accepted social protocols when sharing information in online environments, taking into account different social and cultural contexts

* **Identify the impacts of ICT in society**

Explain the main uses of ICT at school, home and in the local community, and recognise its potential positive and negative impacts on their lives

Typically, by the end of Year 8, students:

* **Apply personal security protocols**

Identify and value the rights to identity, privacy and emotional safety for themselves and others when using ICT and apply generally accepted social protocols when using ICT to collaborate with local and global communities

* **Identify the impacts of ICT in society**

Explain the benefits and risks of the use of ICT for particular people in work and home environments

1. Data representation is a key concept in the Australian Curriculum: Digital Technologies. See also <https://aca.edu.au/curriculum/representation/>. [↑](#footnote-ref-1)