Clickbait or Not?

Clickbait tries to draw you in by giving you just enough information in a headline to make you curious.

In this game you will learn some ways in which we can use computers to look at a headline for an article and try to decide: **is this clickbait or not?**

"You Won't Believe What These Hollywood Stars Looked Like As Teenagers!"

"NAPLAN no longer linked to HSC after NSW Government backflip"

"Top 20 Life Hacks Of 2018"

"Europe Saw Fourfold Increase In Measles Cases In 2017"
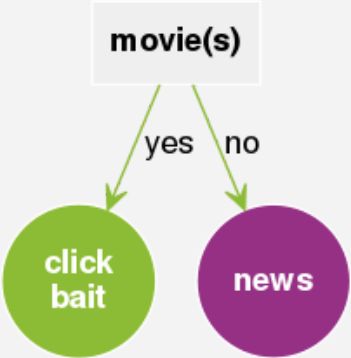
Making clickbait decision trees

1. Find meaningful words or "features".

We filled one document with clickbait headlines, and another with news headlines.

Then we used variants of **TF-IDF** to find the most meaningful words for each document.

```
things,      0.00166862287567
truth,       0.00114717822702
movies,      0.000938600367565
hollywood,   0.000834311437835
video,       0.000834311437835
movie,       0.000730022508106
everyone,    0.000677878043241
didn't,      0.000625733578377
```

2. Build a decision tree using the best features.

We chose the features that help us to separate headlines into clickbait and news.

Then we created graphs that allow us to classify each headline with a maximum of 10 steps.

3. Test the decision tree to see how well it works!

Which tree gets the answer right most often?

We want to choose the best decision tree by trying each one with some example headlines, so we can compare which one guesses correctly most often.

Clickbait	Tree 1	Tree 2	Tree 3
The Longest Movies Ever Made	✔	✔	✔
This is the Cast of 'Sucker Punch' Now	✘	✘	✔
News	**Tree 1**	**Tree 2**	**Tree 3**
ABC apology to Kevin Rudd	✔	✔	✘
Paving The Path For Paid Family Leave	✔	✔	✘
Total Score	3	3	2