Till now we've been using binary classification in our algorithm. (round (1) or not (0), whiskers (1) or not (0))

We use one-hot encoding where we want to classify to more than two options.

One hot encoding

	~~Ear shape~~	Pointy ears	Floppy ears	Oval ears	Face shape	Whiskers	Cat
	~~Pointy~~	1	0	0	Round	Present	1
	~~Oval~~	0	0	1	Not round	Present	1
	~~Oval~~	0	0	1	Round	Absent	0
	~~Pointy~~	1	0	0	Not round	Present	0
	~~Oval~~	0	0	1	Round	Present	1
	~~Pointy~~	1	0	0	Round	Absent	1
	~~Floppy~~	0	1	0	Not round	Absent	0
	~~Oval~~	0	0	1	Round	Absent	1
	~~Floppy~~	0	1	0	Round	Absent	0
	~~Floppy~~	0	1	0	Round	Absent	0

Basically, pointy = [1, 0, 0], oval = [0, 0, 1] and Floppy = [0, 1, 0].

Except for, the correct classification, the whole array is 0.

This is not only used in decision trees but are used in multi-categorical neural network.