How we Categorise Sets of Discrete Features

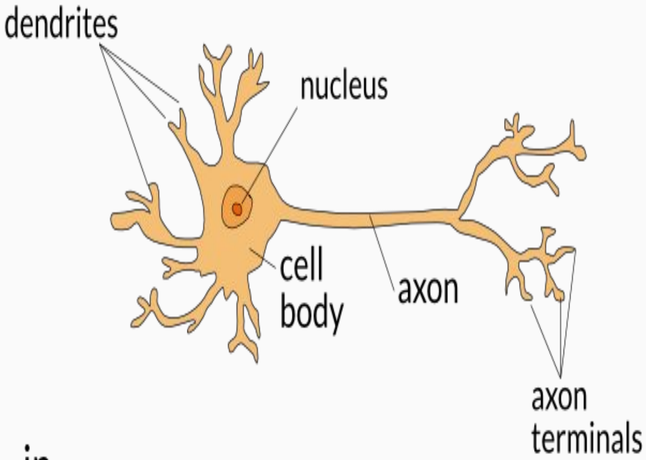
Using an optimally weighted connectionist model as a basis for exploration

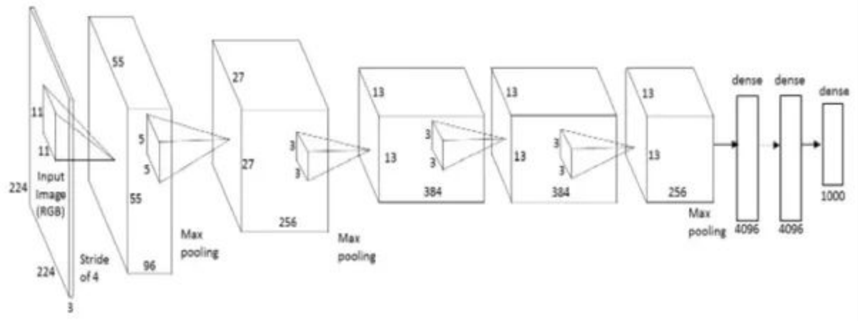
The task

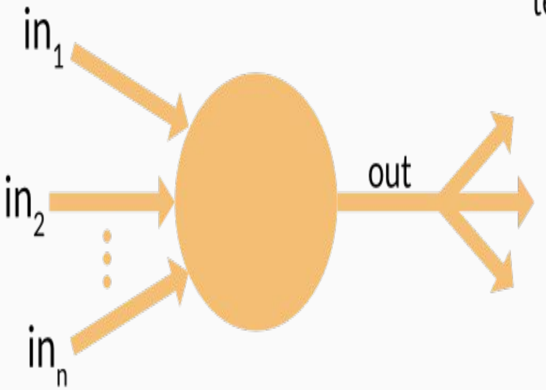
Training item number	Training] item features			Participants learn item is member of disease category or conjunction
	Dim 1	Dim 2	Dim 3	
1	A	X	C	category A
2	A	Y	Y	category A
3	A	A	X	category A
4	Y	A	Y	category A
5	X	A	B	categories A and B
6	A	B	X	categories A and B
7	Z	B	B	category B
8	X	B	B	category B
9	Y	X	B	category B
10	Z	Y	B	category B
11	C	A	Y	category C
12	C	X	B	category C
13	C	Y	C	category C
14	C	A	C	category C
15	C	X	C	category C
16	X	Y	C	category C

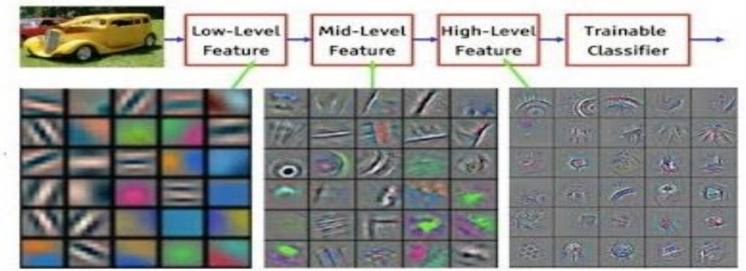
Test item features			Participants asked rate item as member of category or conjunction	Average membership score given for test item in category (range -10 to +10)
Dim 1	Dim 2	Dim 3		
A	B	Y	category A	6.78
C	Y	B	category A	-9.33
Y	A	C	category A	-3.11
X	B	C	category A	-3.06
X	X	B	category A	-1.39
A	B	Y	category B	-2.00
C	Y	B	category B	-3.22
Y	A	C	category B	-5.28
X	B	C	category B	-0.56
X	X	B	category B	7.00
A	B	Y	category C	-9.06
C	Y	B	category C	6.11
Y	A	C	category C	-3.78
X	B	C	category C	0.22
X	X	B	category C	-7.00
A	B	Y	categories A and B	1.17
C	Y	B	categories A and B	-8.56
Y	A	C	categories A and B	-3.72
X	B	C	categories A and B	-2.72
X	X	B	categories A and B	-3.33
A	B	Y	categories A and C	-5.39
C	Y	B	categories A and C	-4.06
Y	A	C	categories A and C	-1.06
X	B	C	categories A and C	-5.06
X	X	B	categories A and C	-7.28
A	B	Y	categories B and C	-7.83
C	Y	B	categories B and C	4.33

Connectionist models



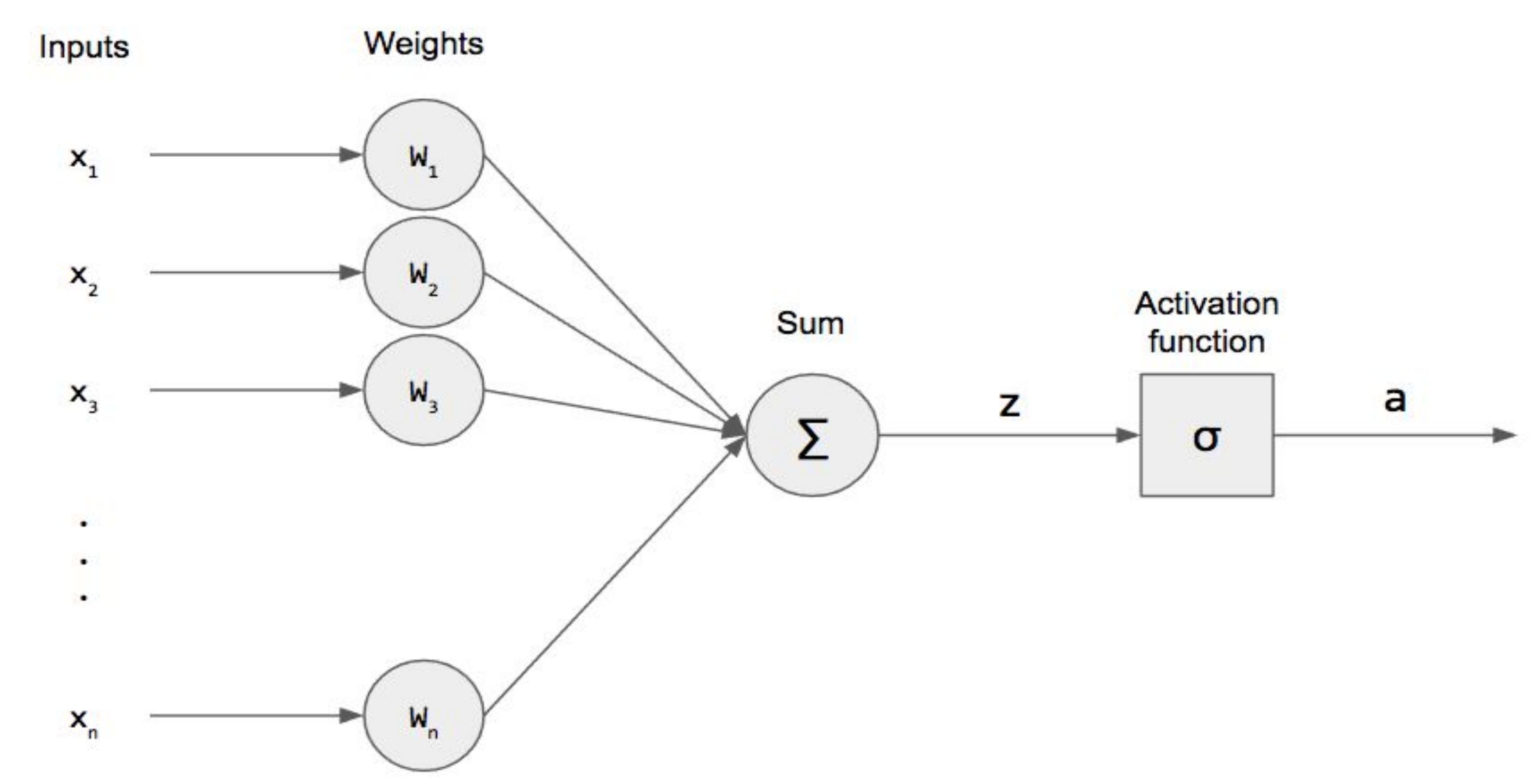




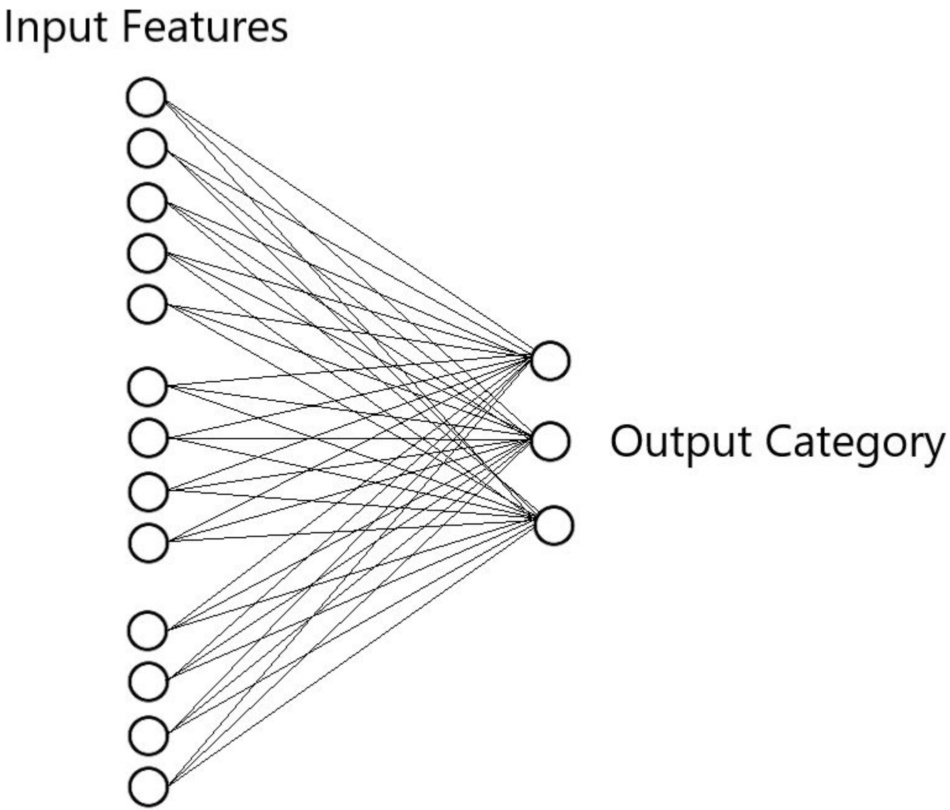


Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

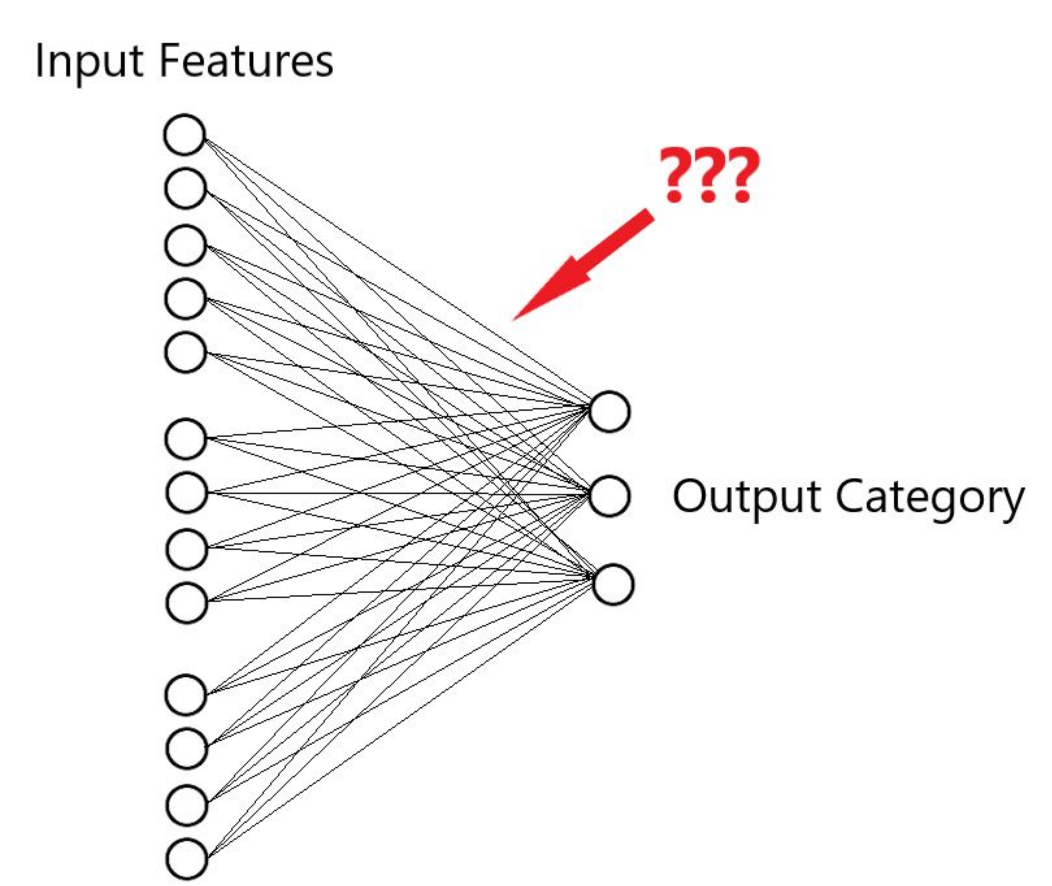
A simple connectionist model for the examined task



A simple connectionist model for the examined task



A simple connectionist model for the examined task



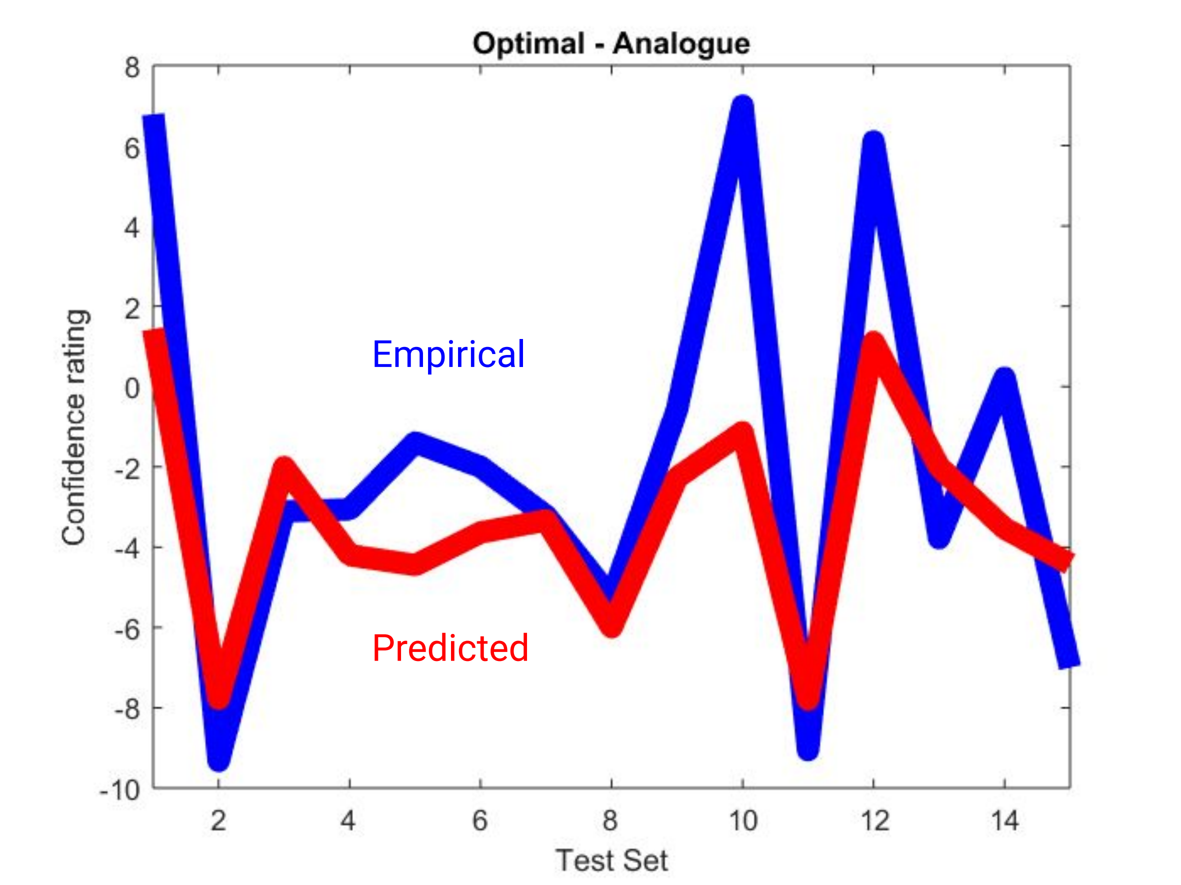
An optimistic choice of connection weights

$$P(\text{Disease}|\text{Symptom})$$

$$=$$

count[sympt. & disease] / count[sympt.]

Results


Optimal - Analogue
8
6
4
2
0
-2
-4
-6
-8
-10
Confidence rating
2
4
6
8
10
12
14
Test Set
Empirical
Predicted

Does ‘digitising’ help?...

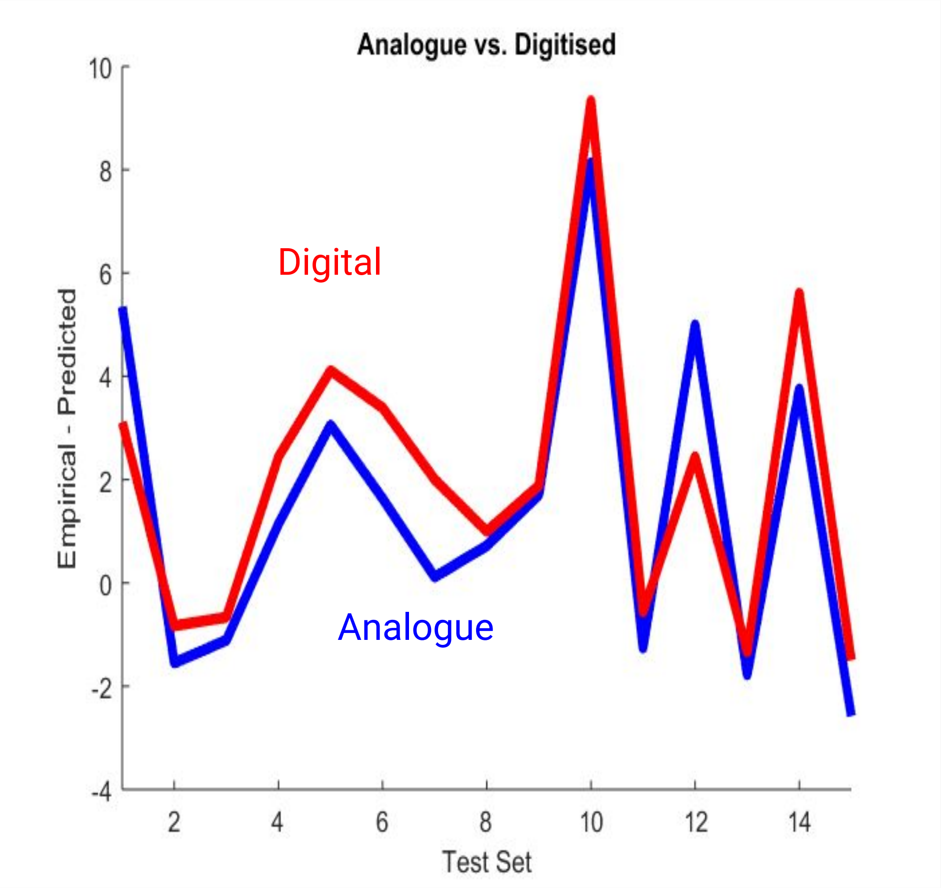
Digitising - Option 1

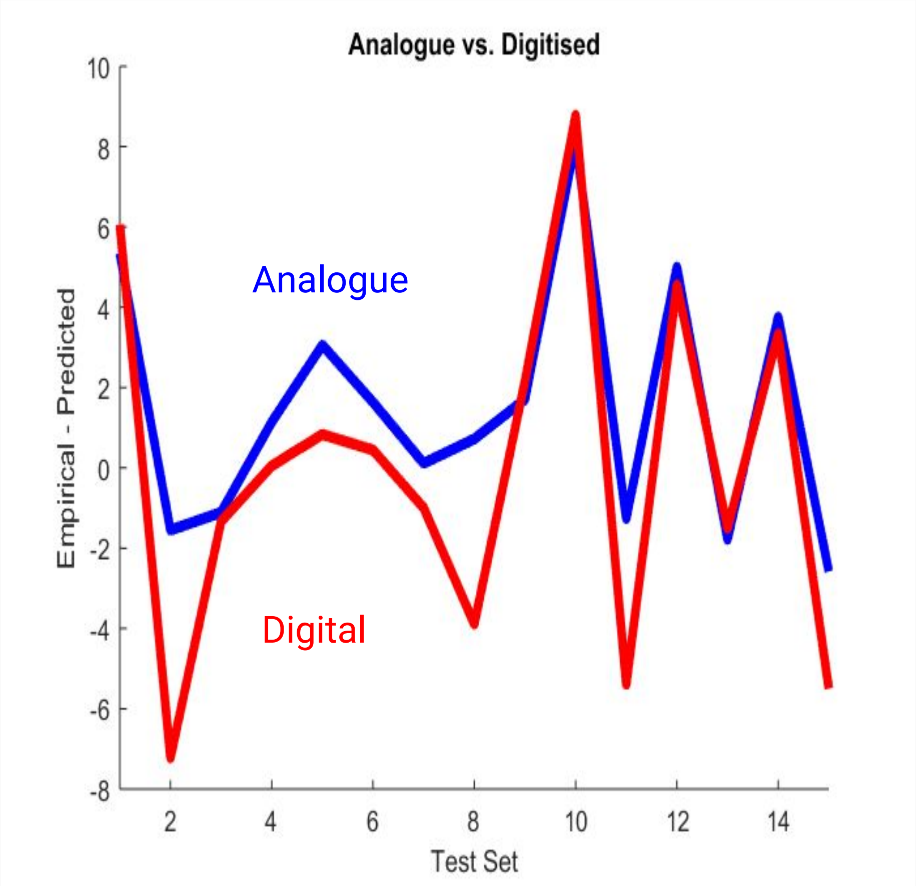
For each individual
Final Output → Closest unique confidence/’belongedness’ rating provided by individual

Digitising - Option 2

Final Output → Roll weighted die to decide whether to accept closest unique weighted confidence/’belongedness’ rating
→ If rejected, try next closest value...wrap if necessary
(weight based on percentage of times that a unique rating appeared in individual’s answers)

...digitising has mixed results **at best**





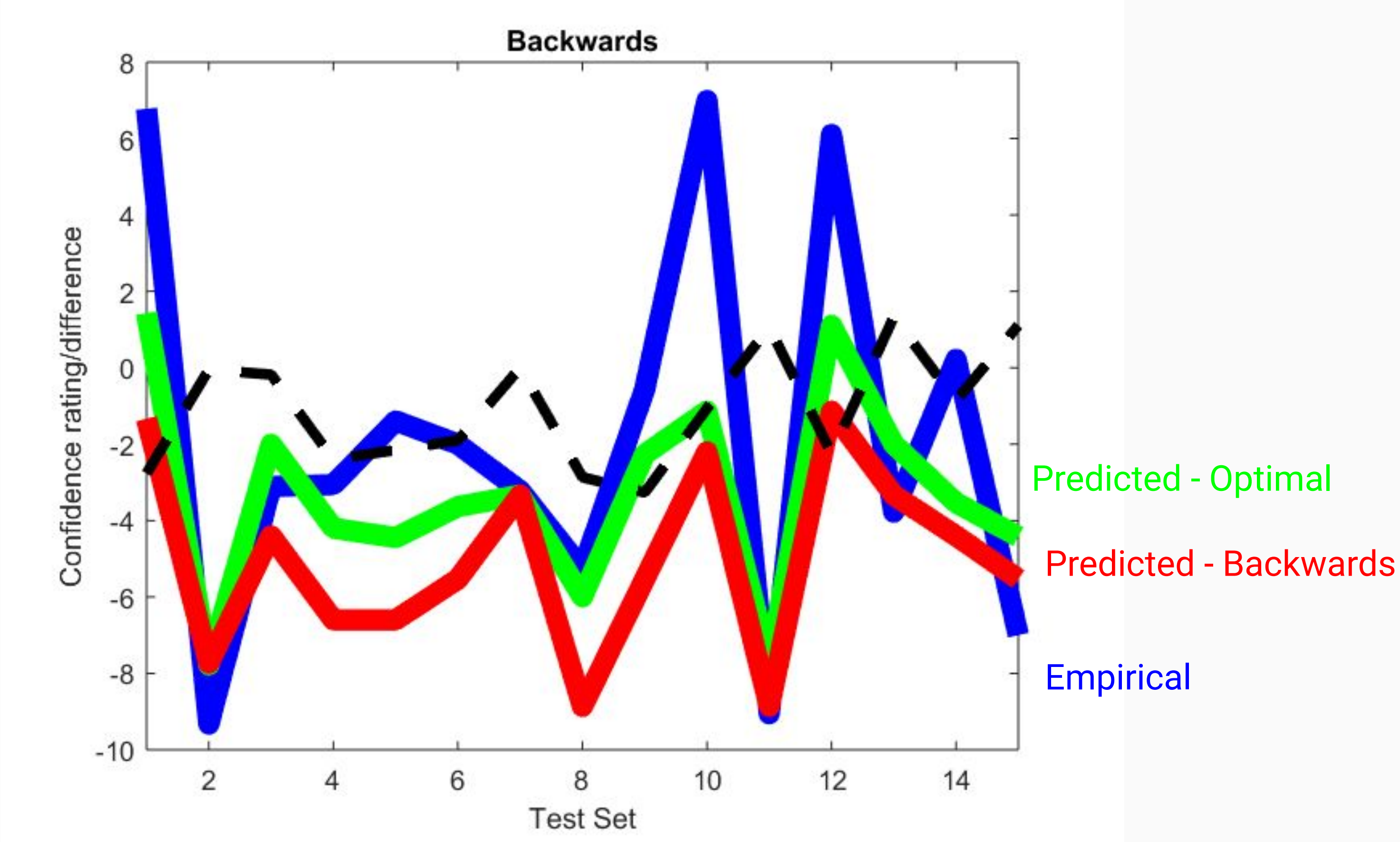
Maybe people are using 'backwards' reasoning.

P(Symptom|Disease)

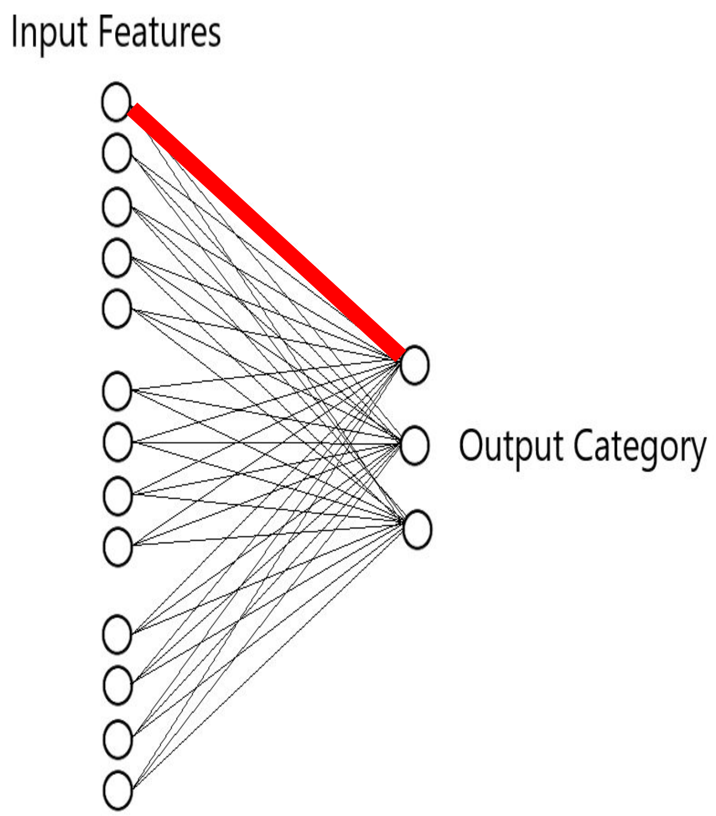
=

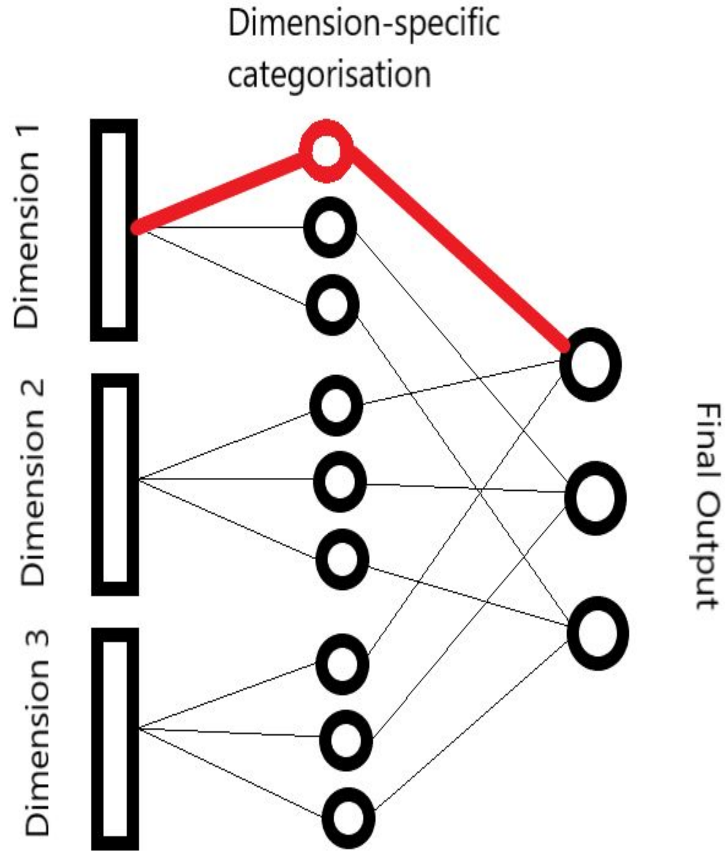
count[sympt. & disease] / count[disease.]

It seems that participants aren't using 'backwards' reasoning.
(side note: adding digitising doesn't help)

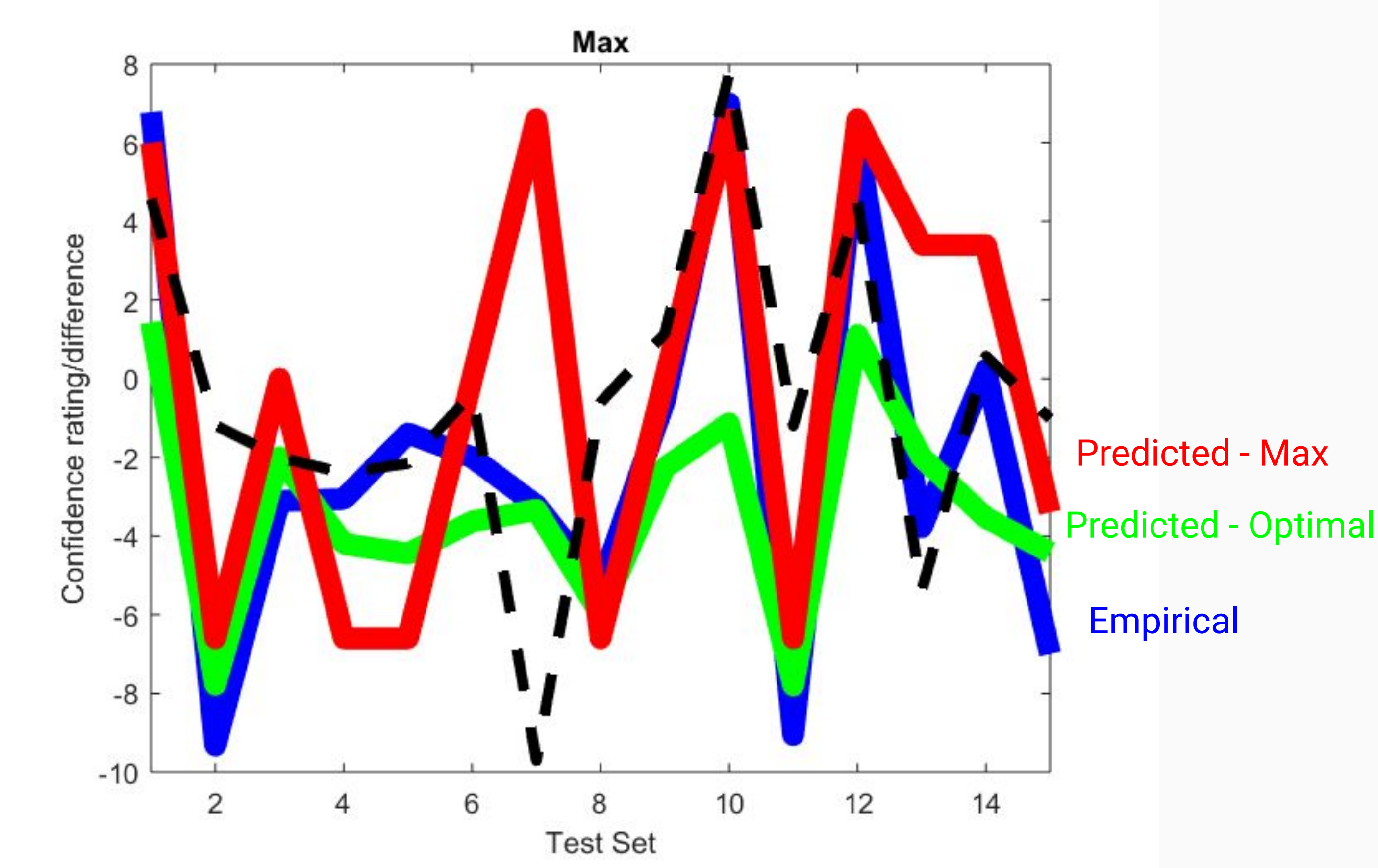


Maybe participants are distracted by - or overweight - particularly strong within-dimension information





The max rule does a good job where the optimum rule fails



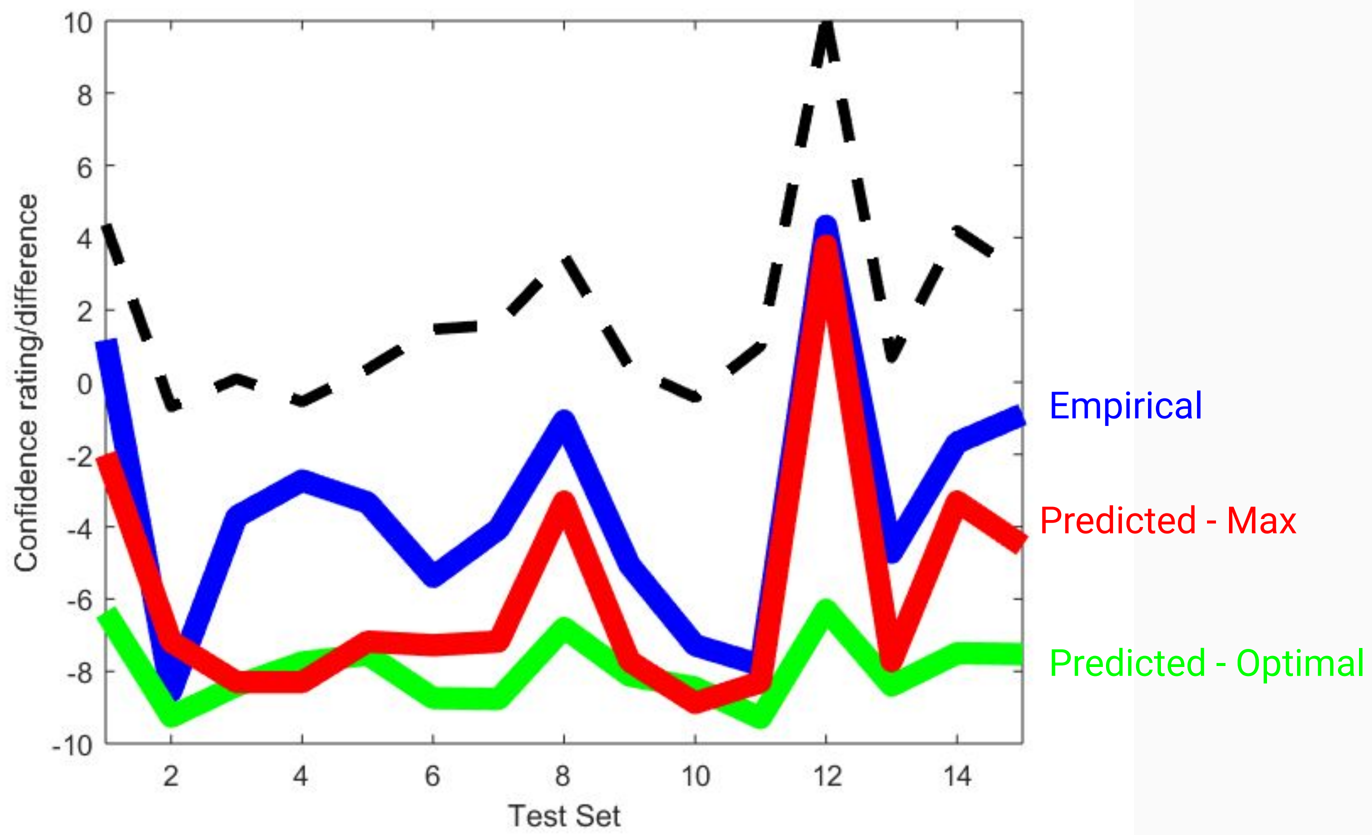
Conclusions so far

- Participants might have used an optimally weighted network, but if a particularly ‘exciting’ feature was present, they ignored other relevant information.

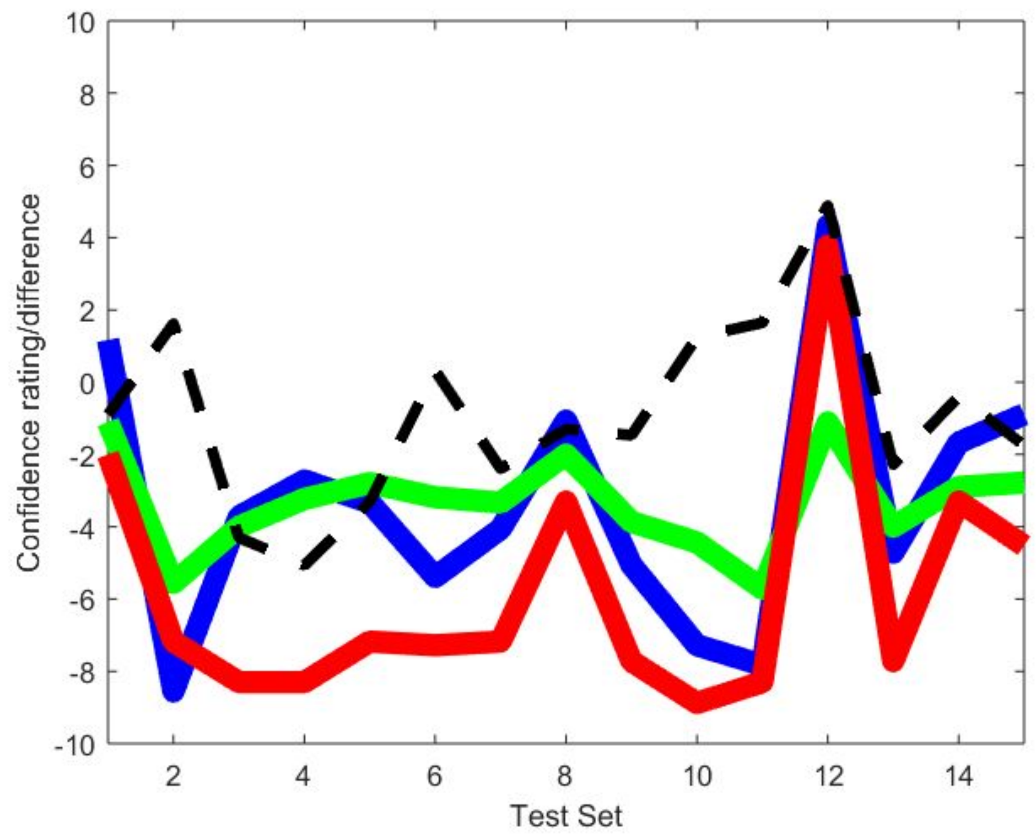
- This might reflect a scenario where participants actually perform closer to optimum when they are less confident in their answer.

- **A simpler explanation might be that participants only learned the easiest classifications, and guessed or chose neutrally when they couldn’t make use of those.**

Conjunctive classification - AND rule



Conjunctive classification - averaging



Conjunctive categorisation: conclusions

- Again, participants might have used an optimally weighted network, and were easily distracted, but a simpler explanation might be that participants only learned the easiest classifications, and guessed (or chose neutrally) when they couldn’t make use of those

- Regardless, participants are too sure of their conjunctive categorisations. This could be due to suboptimal estimations. Or participants fail to use a common scale for both the conjunctive and non-conjunctive categorisation; instead, they rescale for conjunctive categorisation.

Overall conclusions

Participants don't use an optimum strategy for this task:

- They are easily distracted by strong evidence in a single dimension
- They overestimate how confident they should be in conjunctive categorisation

The simplest explanation of the data might be that participants just didn't learn the categories well. With more motivation they might have learned the optimum strategy.

I wonder how well they learned the initial categorisations (was this tested?), and whether some form of reinforcement/feedback would allow them to perform better in conjunction categorisation.

I also wonder whether the same conjunctive categorisation error would be present if the conjunctive task was given first.