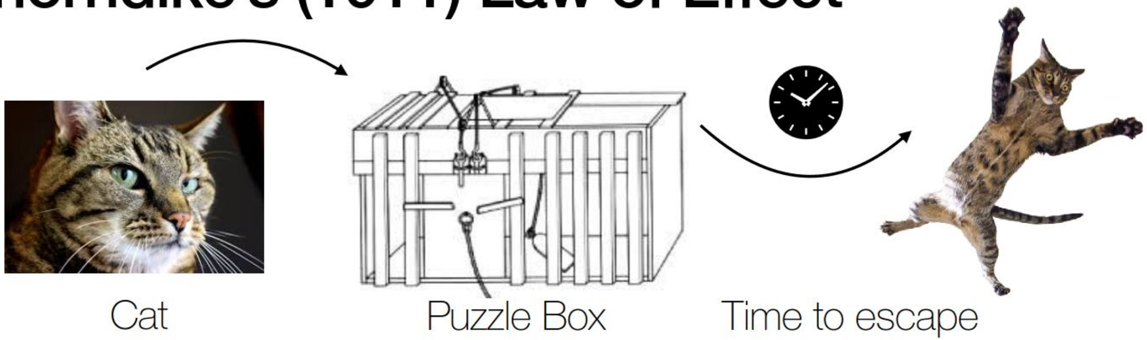
General Principles of Human and Machine Learning

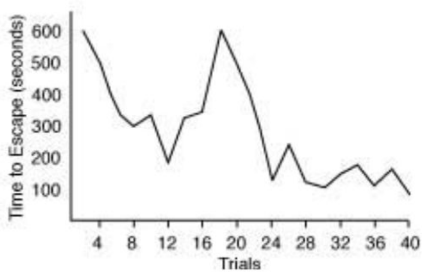
Tutorial 2: Origins of biological and artificial learning

Tutorial Questions

How far can we get with Thorndike's law of effect?

Thorndike's (1911) Law of Effect





Actions associated with satisfaction are strengthened, while those associated with discomfort become weakened.

Tutorial Questions

How far can we get with Thorndike's law of effect?

What are the *benefits*? What are the *limitations*?

Benefits:

- Errors decrease over time
- Openess to trying new solutions
- Basis for all modern reinforcement learning (RL)

Limitations:

- Dangerous when some errors are fatal
- Lacks creativity and generalizastion of past solutions
- No formalism between behavior and outcome….

Tutorial Questions

What are examples of complex behaviors learned through the law of effect?

Tutorial Questions

What are examples of complex behaviors learned through the law of effect?

- Getting a traffic ticket when running a red light →learn to obey the traffic laws
- Missing your flight →arrive at the airport ~2 hours before the flight
- Code runs successfully after hours of debugging →learn to be patient when coding

Tutorial Questions

What are examples of Pavlovian conditioning in our daily lives?

Tutorial Questions

What are examples of Pavlovian conditioning in our daily lives?

- food aversion
- fear of dogs
- craving for popcorn at movies or glühwein at a christmas market

Tutorial Questions

How do advertisers take advantage of us via Pavlovian conditioning?

Tutorial Questions

How do advertisers take advantage of us via Pavlovian conditioning?

- Netflix sound
- coca cola commercial where they open a bottle dramatically

Tutorial Questions

Complete the following table using the Rescorla-Wagner learning rules

Reward Estimation:

Weight Update:

Excitatory association: a single CS_1 paired with a reward of 100 on each trial. Assume w_i starts at 0 and eta = .1

Trial	r_hat	RPE (r - r_hat)	w_1
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Tutorial Questions

Complete the following table using the Rescorla-Wagner learning rules

Reward Estimation:

$$\hat{r}_t = \sum_i \mathrm{CS}_i^t w_i$$

Weight Update:

$$w_i \leftarrow w_i + \eta(r_t - \hat{r}_t)$$

Trial	r_hat	RPE (r - r_hat)	w_1
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Excitatory association: a single CS_1 paired with a reward of 100 on each trial. Assume w_i starts at 0 and eta = .1

Tutorial Questions

Complete the following table using the Rescorla-Wagner learning rules

Reward Estimation:

$$\hat{r}_t = \sum_i \mathrm{CS}_i^t w_i$$

Weight Update:

$$w_i \leftarrow w_i + \eta(r_t - \hat{r}_t)$$

Trial	r_hat	RPE (r - r_hat)	w_1
1	0	100	10
2	10	90	19
3	19	81	27.1
4	27.1	72.9	34.39
5	34.39	65.61	40.951
6	40.951	59.049	46.8559
7	46.8559	53.1441	52.17031
8	52.17031	47.82969	56.953279
9	56.953279	43.046721	61.2579511
10	61.2579511	38.7420489	65.13215599

Excitatory association: a single CS_1 paired with a reward of 100 on each trial. Assume w_i starts at 0 and eta = .1

Tutorial Questions

Complete the following table using the Rescorla-Wagner learning rules

Reward Estimation:

$$\hat{r}_t = \sum_i \text{CS}_i^t w_i$$

Weight Update:

$$w_i \leftarrow w_i + \eta(r_t - \hat{r}_t)$$

Overshadowing: same as before, but with two CS_1 and CS_2. CS_1 and CS_2 are paired with a reward of 100 from trial 1 to 10. What reward expectations are present on trial 11 if only CS_1 or CS_2 are present? As before, assume all weights start at 0 and eta = .1

Trial	r_hat	RPE (r - r_hat)	w_1	w_2
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11 with CS_1				
11 with CS_2				

Tutorial Questions

Complete the following table using the Rescorla-Wagner learning rules

Reward Estimation:

$$\hat{r}_t = \sum_i CS_i^t w_i$$

Weight Update:

$$w_i \leftarrow w_i + \eta(r_t - \hat{r}_t)$$

Overshadowing: same as before, but with two CS_1 and CS_2. CS_1 and CS_2 are paired with a reward of 100 from trial 1 to 10. What reward expectations are present on trial 11 if only CS_1 or CS_2 are present? As before, assume all weights start at 0 and eta = .1

Trial	r_hat	RPE (r - r_hat)	w_1	w_2
1	0	100	10	10
2	20	80	18	18
3	36	64	24.4	24.4
4	48.8	51.2	29.52	29.52
5	59.04	40.96	33.616	33.616
6	67.232	32.768	36.8928	36.8928
7	73.7856	26.2144	39.51424	39.51424
8	79.02848	20.97152	41.611392	41.611392
9	83.222784	16.777216	43.2891136	43.2891136
10	86.5782272	13.4217728	44.63129088	44.63129088
11 with CS_1	44.63129088			
11 with CS_2	44.63129088			

Tutorial Questions

Complete the following table using the Rescorla-Wagner learning rules

Reward Estimation:

$$\hat{r}_t = \sum_i \mathrm{CS}_i^t w_i$$

Weight Update:

$$w_i \leftarrow w_i + \eta(r_t - \hat{r}_t)$$

Blocking: same as before, with CS_1 and CS_2. CS_1 is now paired with a reward of 100 on each trial, but then CS_2 is introduced on trial 5. As before, assume all weights start at 0 and eta = .1

Trial	r_hat	RPE (r - r_hat)	w_1	w_2
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11 with CS_1				
11 with CS_2				

Tutorial Questions

Complete the following table using the Rescorla-Wagner learning rules

Reward Estimation:

$$\hat{r}_t = \sum_i \text{CS}_i^t w_i$$

Weight Update:

$$w_i \leftarrow w_i + \eta(r_t - \hat{r}_t)$$

Blocking: same as before, with CS_1 and CS_2. CS_1 is now paired with a reward of 100 on each trial, but then CS_2 is introduced on trial 5. As before, assume all weights start at 0 and eta = .1

Trial	r_hat	RPE (r - r_hat)	w_1	w_2
1	0	100	10	0
2	10	90	19	0
3	19	81	27.1	0
4	27.1	72.9	34.39	0
5	34.39	65.61	40.951	6.561
6	47.512	52.488	46.1998	11.8098
7	58.0096	41.9904	50.39884	16.00884
8	66.40768	33.59232	53.758072	19.368072
9	73.126144	26.873856	56.4454576	22.0554576
10	78.5009152	21.4990848	58.59536608	24.20536608
11 with CS_1	58.59536608			
11 with CS_2	24.20536608			

Tutorial Questions

What are examples of Operant conditioning?

Tutorial Questions

What are examples of Operant conditioning?

- training a puppy
- rewarding a child for cleaning their room

Tutorial Questions

What are examples of where companies or governments use operant conditioning to shape behavior?

Tutorial Questions

What are examples of where companies or governments use operant conditioning to shape behavior?

- Getting likes/retweets on social media posts
- Getting citations as a scientist
- Tip culture as a service worker
- Social credit score in China

Tutorial Questions

Perceptron activation function is:

In the perceptron below, what will the output be when the input is (0, 0)? What about inputs (0, 1), (1, 1) and (1, 0)? What if we change the bias weight to -0.5?

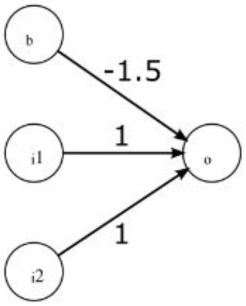


Figure 1: Single Layer Perceptron. b = 1

Tutorial Questions

Perceptron activation function is:

$$\sigma(\mathbf{w}^\top \mathbf{x} + b) = \begin{cases} 1 & \text{if} \quad \mathbf{w}^\top \mathbf{x} + b \geq 0 \\ 0 & \text{else} \end{cases}$$

In the perceptron below, what will the output be when the input is (0, 0)? What about inputs (0, 1), (1, 1) and (1, 0)? What if we change the bias weight to -0.5?

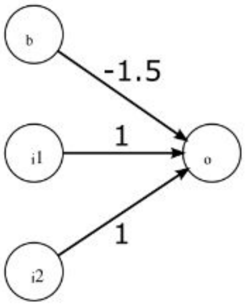


Figure 1: Single Layer Perceptron. b = 1

Tutorial Questions

Perceptron activation function is:

$$\sigma(\mathbf{w}^\top \mathbf{x} + b) = \begin{cases} 1 & \text{if} \quad \mathbf{w}^\top \mathbf{x} + b \geq 0 \\ 0 & \text{else} \end{cases}$$

In the perceptron below, what will the output be when the input is (0, 0)? What about inputs (0, 1), (1, 1) and (1, 0)? What if we change the bias weight to -0.5?

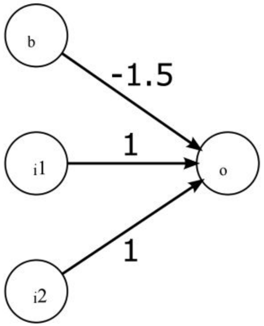


Figure 1: Single Layer Perceptron. b = 1

Answer:

Bias $= -1.5$			Bias $= -0.5$		
Input	Weighted sum	Output	Input	Weighted sum	Output
$(0,0)$	-1.5	0	$(0,0)$	-0.5	0
$(0,1)$	-0.5	0	$(0,1)$	0.5	1
$(1,0)$	-0.5	0	$(1,0)$	0.5	1
$(1,1)$	0.5	1	$(1,1)$	1.5	1

Tutorial Questions

Bias = -1.5 perceptron:

AND gate:

I_1: is_friday

I_2: is_4_15_pm

O: tutorial_time

Bias = -0.5 perceptron:

OR gate:

I_1: is_not_friday

I_2: is_5_30_pm

O: tutorial_over

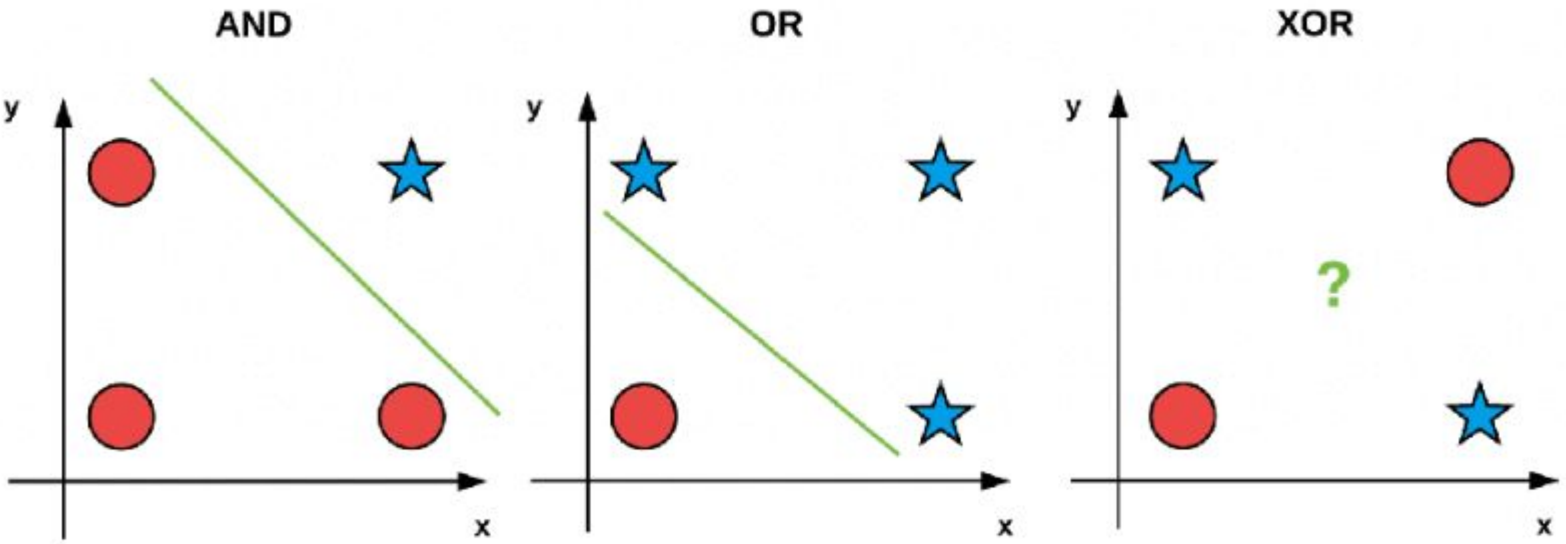
Tutorial Questions

What about XOR gates? Can a perceptron model it? Why or why not?

Tutorial Questions

What about XOR gates? Can a perceptron model it? Why or why not?

NO! XOR inputs aren't linearly separable.



Tutorial Questions

Design a perceptron to model predict whether or not you would like a certain food, movie, etc… based on a set of continuous features with binary outcomes. First, draw a table with a set of features and an outcome label. Come up with about 2+ features and 5-6 examples.

Tutorial Questions

Design a perceptron to model predict whether or not you would like a certain food, movie, etc… based on a set of continuous features with binary outcomes. First, draw a table with a set of features and an outcome label. Come up with about 2+ features and 5-6 examples.

Food	Sweet	Savory	Bitter	Enjoyment (outcome)
Chocolate	0.8	0	0.2	1
Whisky	0.1	0	0.2	1
Spaghetti	0.1	0.9	0.1	1
Tennis ball	0	0	0.4	0
Hot Garbage	0.1	0.2	0.3	0

Tutorial Questions

draw/program a perceptron and perform the update rule

Food	Sweet	Savory	Bitter	Enjoyment (outcome)
Chocolate	0.8	0	0.2	1
Whisky	0.1	0	0.2	1
Spaghetti	0.1	0.9	0.1	1
Tennis ball	0	0	0.4	0
Hot Garbage	0.1	0.2	0.3	0

Algorithm 1: Perceptron Learning Algorithm

Input: Training examples $\{\mathbf{x}_i, y_i\}_{i=1}^m$.

Initialize $\mathbf{w}$ and b randomly.

while *not converged* **do**

    # # # Loop through the examples.
    for $j = 1, m$ **do**
        # # # Compare the true label and the prediction.
        $error = y_j - \sigma(\mathbf{w}^T\mathbf{x}_j + b)$
        ### If the model wrongly predicts the class, we update the weights and bias.
        if *error != 0* **then**
            ### Update the weights.
            $\mathbf{w} = \mathbf{w} + error \times x_j$
            ### Update the bias.
            $b = b + error$
    Test for convergence

Output: Set of weights $\mathbf{w}$ and bias b for the perceptron.

Tutorial Questions

Food	Sweet	Savory	Bitter	Enjoyment (outcome)
Chocolate	0.8	0	0.2	1
Whisky	0.1	0	0.2	1
Spaghetti	0.1	0.9	0.1	1
Tennis ball	0	0	0.4	0
Hot Garbage	0.1	0.2	0.3	0

Training Step	w_1	w_2	w_3	b	Outcomes
0 (Chocolate)	0	0	0	0	0
1 (Whisky)	0.8	0	0.2	1	1
2 (Spaghetti)	0.8	0	0.2	1	1
3 (Tennis ball)	0.8	0	0.2	1	1
4 (Hot Garbage)	0.8	0	-0.2	0	1

Tutorial Questions

Food	Sweet	Savory	Bitter	Enjoyment (outcome)
Chocolate	0.8	0	0.2	1
Whisky	0.1	0	0.2	1
Spaghetti	0.1	0.9	0.1	1
Tennis ball	0	0	0.4	0
Hot Garbage	0.1	0.2	0.3	0

Training Step	w_1	w_2	w_3	b	Outcomes
5 (Chocolate)	0.7	-0.2	-0.5	-1	0
6 (Whisky)	1.5	-0.2	-0.3	0	1
7 (Spaghetti)	1.5	-0.2	-0.3	0	0
8 (Tennis ball)	1.6	0.7	-0.2	1	1
9 (Hot Garbage)	1.6	0.7	-0.6	0	1

Tutorial Questions

Food	Sweet	Savory	Bitter	Enjoyment (outcome)
Chocolate	0.8	0	0.2	1
Whisky	0.1	0	0.2	1
Spaghetti	0.1	0.9	0.1	1
Tennis ball	0	0	0.4	0
Hot Garbage	0.1	0.2	0.3	0

Training Step	w_1	w_2	w_3	b	Outcomes
10 (Chocolate)	1.5	0.5	-0.9	-1	1
11 (Whisky)	1.5	0.5	-0.9	-1	0
12 (Spaghetti)	1.6	0.5	-0.7	0	1
13 (Tennis ball)	1.6	0.5	-0.7	0	0
14 (Hot Garbage)	1.6	0.5	-0.7	0	1

Tutorial Questions

Food	Sweet	Savory	Bitter	Enjoyment (outcome)
Chocolate	0.8	0	0.2	1
Whisky	0.1	0	0.2	1
Spaghetti	0.1	0.9	0.1	1
Tennis ball	0	0	0.4	0
Hot Garbage	0.1	0.2	0.3	0

Training Step	w_1	w_2	w_3	b	Outcomes
15 (Chocolate)	1.5	0.3	-1	-1	1
16 (Whisky)	1.5	0.3	-1	-1	0
17 (Spaghetti)	1.6	0.3	-0.8	0	1
18 (Tennis ball)	1.6	0.3	-0.8	0	0
19 (Hot Garbage)	1.6	0.3	-0.8	0	0

Tutorial Questions

Can you test it on new data to see how well it performs in predicting your preferences?

Food	Sweet	Savory	Bitter	Enjoyment (outcome)
Chocolate	0.8	0	0.2	1
Whisky	0.1	0	0.2	1
Spaghetti	0.1	0.9	0.1	1
Tennis ball	0	0	0.4	0
Hot Garbage	0.1	0.2	0.3	0
Chips	0.2	0.9	0.2	1
Wood	0	0	0.1	0

Testing Step	w_1	w_2	w_3	b	Outcomes
1 (Chips)	1.6	0.3	-0.8	0	1
2 (Wood)	1.6	0.3	-0.8	0	0