Problem 1

Recall that in the MNIST data we could have recorded the labels as:

	X	LABEL
Image 1	28x28 image	6
Image 2	28x28 image	3
Image 3	28x28 image	6
Image 4	28x28 image	2

The label column means, for example, that the first image is of the digit 6, the second image of 3 and so on. Suppose we have a new dataset

Name	Sport (Label)	Height	Weight
Brittainey Raven	Basketball	72	162
Simone Biles	Gymnastics	57	104
Asuka Teramoto	Gymnastics	54	66
Irina Miketenko	Track	63	106
Nikki Blue	Basketball	68	163
Elena Delle Donne	Basketball	78	204

Here the label (what we are trying to predict) is the sport. How would we best represent the data in order to train a neural network. For example, one possibility might be:

```
['Basketball', 'Gymnastics', Gymnastics', 'Track', 'Basketball', 'Basketball']
```

What do you think is the best representation?

I want to see the labels of the entire dataset encoded in your representation.

What is this representation called?

Problem 2

Recall that a neuron computes a linear function (z = Wx + b) followed by an activation function.

Suppose the neuron has two inputs:

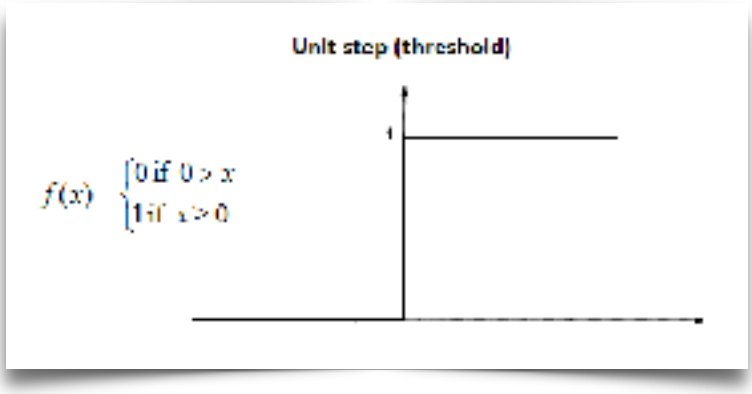
x1 = 12
x2 = 4

The weights are:

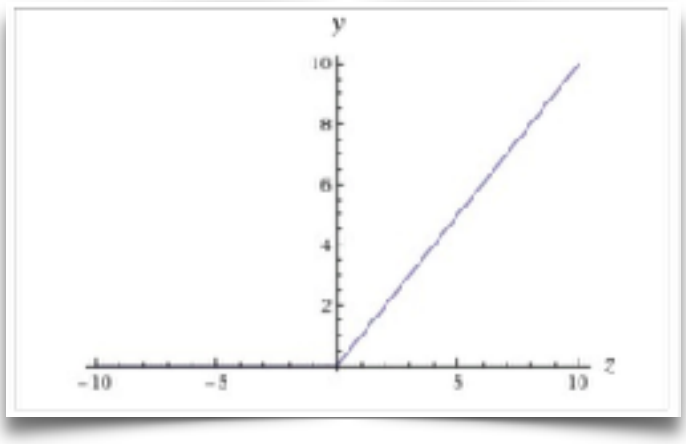
w1 = 0.5
w2 = -1

The bias is zero.

What is the output of the neuron if it uses the 0-1 activation function:



What is the output of the neuron if it uses the ReLU activation function:



What is the output of the neuron if the bias is set to -5? (both for the 0-1 and ReLU activation functions.

Problem 3

The delta rule is

$$\Delta w_k = \sum_i \epsilon x_k^{(i)} (t^{(i)} - y^{(i)})$$

- t^i is the true answer for the i^{th} training example
- y^i is the prediction our neural network made for the i^{th} training example
- x^i is the input value
- epsilon is the learning rate

Suppose we use the 0-1 activation function.

The neuron has two inputs x1 and x2. Our small dataset is

Example	X1	x2	T (the true label)
1	1	0	1
2	1	1	1
3	0	0	0
4	0	1	1
5	0	0	0

Suppose ε is 0.1 and all weights and biases are initially zero. Show the revised values of x1 and x2 after training on each training example. For example,

After example 1: w1 = … w2 = … b = …
After example 2: w1 = … w2 = … b = …
Etc

Problem 4.

Just some general intro questions.

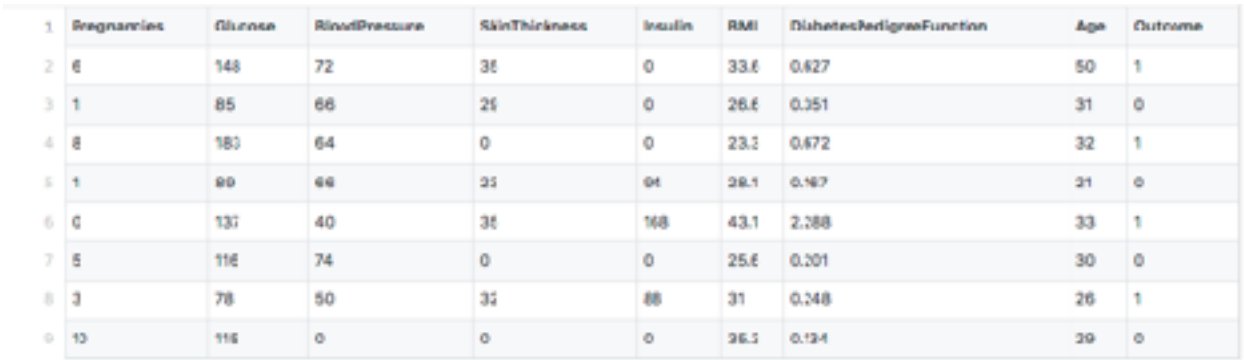
1. What are the reasons that Deep Learning has suddenly taken off in recent years?

2. What does the analogy “AI is the new electricity” refer to?

3. What are the three types of AI we discussed? (their acronyms are ANI, AGI, and ASI). Give specific examples.

Problem 5

Consider the Pima Indian Dataset a small sample of which is shown here.



	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
2	6	148	72	35	0	33.6	0.627	50	1
3	1	85	66	29	0	26.6	0.351	31	0
4	8	183	64	0	0	23.3	0.672	32	1
5	1	89	66	23	94	28.1	0.167	21	0
6	0	137	40	35	168	43.1	2.288	33	1
7	5	116	74	0	0	25.6	0.201	30	0
8	3	78	50	32	88	31	0.248	26	1
9	10	115	0	0	0	35.3	0.134	29	0

What would be the Tensorflow code to define the following variables:

```
X
Y_
W
b
```

Suppose we have the following dataset:

Name	Sport (Label)	Height	Weight
Brittainey Raven	Basketball	72	162
Simone Biles	Gymnastics	57	104
Asuka Teramoto	Gymnastics	54	66
Irina Miketenko	Track	63	106
Nikki Blue	Basketball	68	163
Elena Delle Donne	Basketball	78	204

And we are trying to predict sport.

What would the Tensorflow code for X, Y_, W, and b look like then?

Problem 6

A neuron computes a linear function (z = Wx + b) followed by an activation function,

What line of our code computes the linear function?

What line computes the activation function?

What is the purpose of this line of the code:

```
cross_entropy = -tf.reduce_mean(Y_ * tf.log(Y)) * 1000.0
```

Just curious. In our original code we set the weights and biases initially to zero. What happens if we initially set them to random values using `tf.random_normal`. Does it affect the final accuracy?

Problem 7

We talked about the following activation functions:

0-1
Sigmoid
TanH
ReLU
Softmax

Can you define these? Which ones are commonly used? Any advantages, disadvantages? Can you use them for any neuron? Feel free to Google

Problem 8

Suppose we asks some customers how well they like dream pop music, neotraditional country music, and R&B music and they responded

P

Customer	Dream pop	Neotraditional country	R&B
Mary	1	5	3
Sam	3	1	5
Julie	5	2	3
Lucy	1	5	5

And we had experts rate artists on these same genres:

Q

Artist	Dream pop	Neotraditional country	R&B
Duo Lipa	5	1	4
Camila Cabello	1	`2	5
Bruno Mars	3	1	5
Lorde	4	2	2

Please compute. PQ^T

If you are doing this by hand you only need to do the row matching your team number plus the row (team number + 1) mod whatever.