**חלק א**

**First try:**

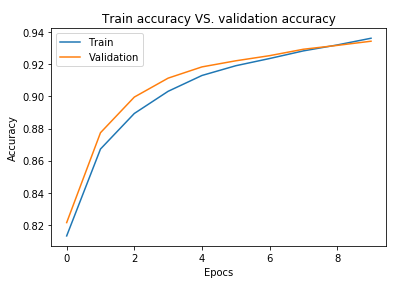
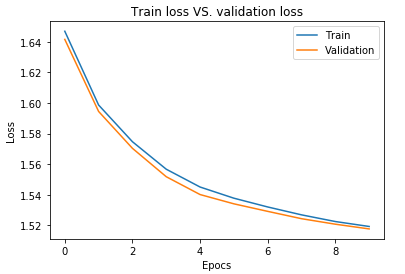
* **Layers architecture:**

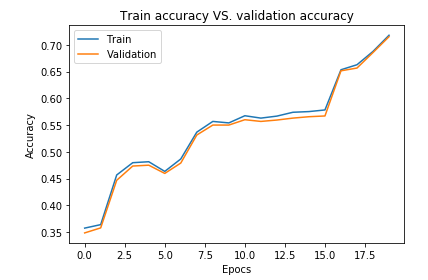
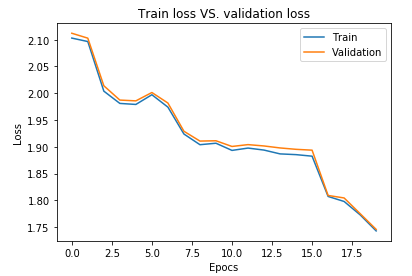
1. Input layer 784,

2. Hidden layer 256

3. Output layer 10

* **Learning rate**: 0.001
* **Optimization algorithm:** Adam optimizer
* **Loss function:** softmax\_cross\_entropy\_with\_logits
* **Batch size**: 60
* **Amount of epochs**: 10
* **Activation functions**: sigmoid
* **Droupout**: 0.5
* **Accuracy**: 0.93



**Second try:**

* **Layers architecture:**

1. Input layer 784,

2. Hidden layer 256

3. Output layer 10

* **Learning rate**: 0.01
* **Optimization algorithm:** Gradient Descent Optimizer
* **Loss function:** softmax\_cross\_entropy\_with\_logits
* **Batch size**: 100
* **Amount of epochs**: 20
* **Activation functions**: relu, softmax
* **Droupout**: 1.0
* **Accuracy**: 0.71

**Third try:**

* **Layers architecture:**

1. Input layer 784,

2. Hidden layer 256

3. Hidden layer 256

4. Hidden layer 128

3. Output layer 10

* **Learning rate**: 0.001
* **Optimization algorithm:** Adam Optimizer
* **Loss function:** softmax\_cross\_entropy\_with\_logits
* **Batch size**: 200
* **Amount of epochs**: 20
* **Activation functions**: sigmoid, softmax
* **Droupout**: 1.0
* **Accuracy**: 0.95