**ΕΠΛ 442: Μηχανική Μάθηση**

**ΕΡΓΑΣΙΑ 1: Back Propagation  
11/10/2023**

**Νικόλαος Θεοδώρου 1030496  
  
Εισαγωγή**

Σε αυτήν την αναφορά περιγράφονται οι λεπτομέρειες της εκπαίδευσης ενός νευρωνικού δικτύου για την επίλυση του προβλήματος της λογικής πύλης XOR. Περιλαμβάνονται οι χρησιμοποιηθείσες μεταβλητές, οι επιλογές που αφορούν τη δομή του νευρωνικού δικτύου και τα αποτελέσματα που προέκυψαν από την εκπαίδευση. Η εργασία τρέχει με την εντολή python3 Train.py και απαιτεί python 3.6 η νεότερη και τα πακέτα numpy και matplotlib.

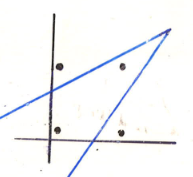
**Μεταβλητές και Επιλογές για το Νευρωνικό Δίκτυο**

Οι επιλογές που έγιναν για τη δομή του νευρωνικού δικτύου είναι οι εξής:

* Πρώτο Επίπεδο: 2 νευρώνες (2 είσοδοι)
* Δεύτερο Επίπεδο: 2 νευρώνες
* Τρίτο Επίπεδο: Δεν υπάρχουν νευρώνες (δεν υπάρχει δεύτερο κρυφό επίπεδο)
* Τέταρτο Επίπεδο (Έξοδος): 1 νευρώνας

Ρυθμός Μάθησης (learningRate): 0.5

* Μεταβλητή Momentum: 0.5
* Αριθμός Επαναλήψεων (maxIterations): 800

Οι παραπάνω παράμετροι καθορίζουν τον τρόπο με τον οποίο το νευρωνικό δίκτυο εκπαιδεύεται και προσεγγίζει την λογική λειτουργία της πύλης XOR. Μπορούμε να προσεγγίσουμε την διάταξη της πύλη XOR στις 2 διαστάσεις χωρίζοντας τις Αληθείς και Ψευδής εξόδους με μόλις μόνο 2 γραμμές, όπως φαίνεται στο πιο κάτω σχήμα.

**Για αυτό το λόγο επέλεξα να χρησιμοποιήσω μόνο ένα κρυφό επίπεδο με 2 νευρώνες.  
Μετά από δοκιμές για να βρώ τη βέλτιστη τιμή του learning rate, παρατήρησα πως η τιμή 0.5, δίνει ακριβής αποτελέσματα στην έξοδο μετά την εκμάθηση, καθώς και το δύκτιο φαίνεται να εκπαιδεύεται γρήγορα.**

**Το momentum το έθεσα 0.5 που είναι αρκετό για να αποφεύγονται τα κολλήματα στα τοπικά ελάχιστα.**

**Ο αριθμός των εποχών εκπαίδευσης είναι αρκετά μεγάλος, ώστε να έχουμε ελάχιστο σφάλμα με το τέλος της εκπαίδευσης και μέγιστο ποσοστό επιτυχίας στα δεδομένα ελέγχου.**

**Αρχείο εκπαίδευσης training.txt:**

**in1 in2 out**

**0 0 0**

**0 1 1**

**1 0 1**

**1 1 0**

**Αρχείο ελέγχου test.txt:**

**in1 in2 out**

**0 0 0**

**0 1 1**

**1 0 1**

**Threshold για κατάταξη**

**Για την αντιστοίχηση της τιμης εξόδου του νευρωνικού δυκτίου με τις τιμές Αληθής, Ψευδής είναι out>0.75 για Ορθο και out<0.25. Πιο συγκεκριμένα για να θεωρείτε σωστή η έξοδος του δυκτιου για target output = 1 , το actual output πρέπει να είναι μεγαλύτερο του 0.75, για target output = 0, το actual output πρέπει να είνα μικρότερο του 0.25.**

**Ανάλυση Αποτελεσμάτων**

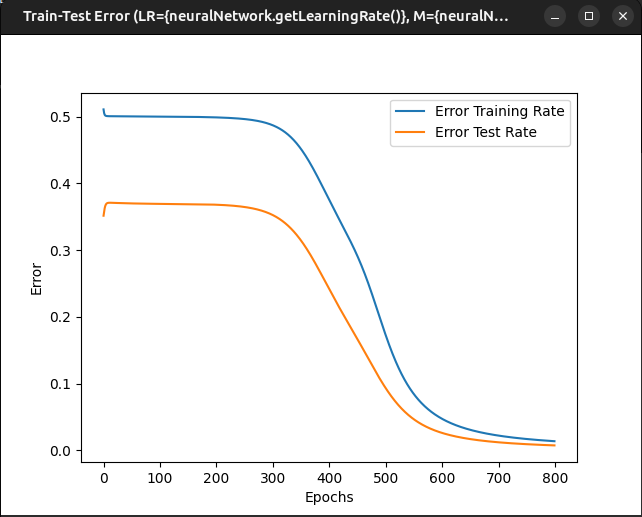
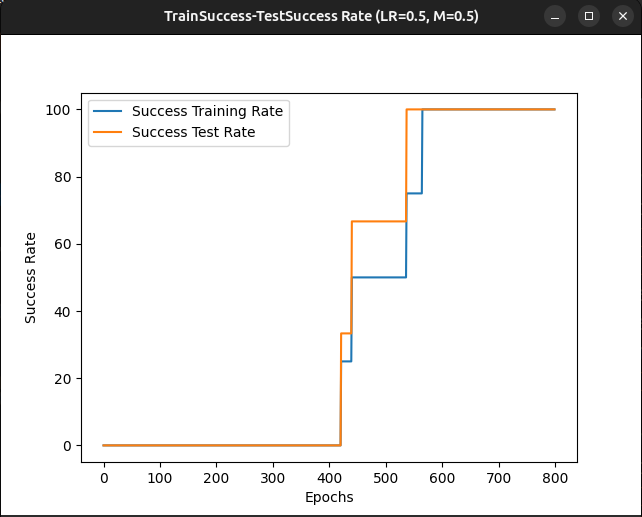
Σύμφωνα με τις πιο κάτω γραφικές, οι προβλεπόμενες εξόδοι σχετίζονται σωστά με τις εισόδους, καταφέρνοντας να επιτύχει το επιθυμητό αποτέλεσμα για όλα τα παραδείγματα του συνόλου δεδομένων.

Τα αποτελέσματα δείχνουν ότι το νευρωνικό δίκτυο κατάφερε να μάθει την πολύπλοκη μη-γραμμική συνάρτηση της πύλης XOR. Οι μεταβλητές του δικτύου ενημερώθηκαν με βάση την αλγοριθμική διαδικασία του backpropagation. Παρατηρούμε πως συνήθως μετά απο 600 εποχές εκμάθησης κατα μέσο όρο, το δύκτιο εκπαιδεύεται επαρκώς χωρίς να κάνει λάθη, τόσο στο σύνολο της εκμάθησης όσο και του ελέγχου. Δεν φαίνεται να κολλούμε σε τοπικά ελάχιστα.

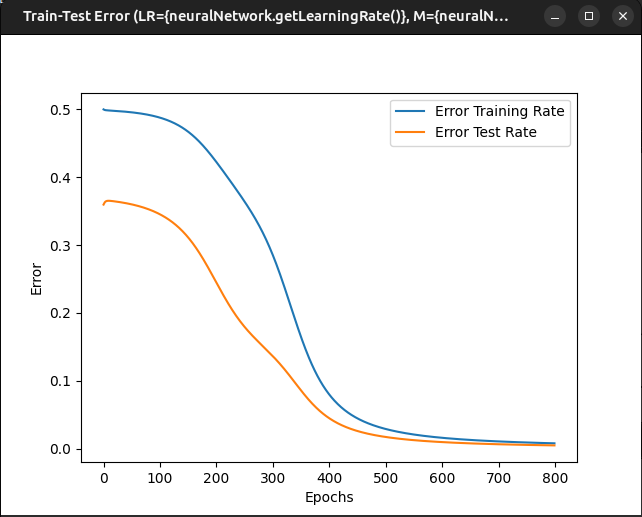
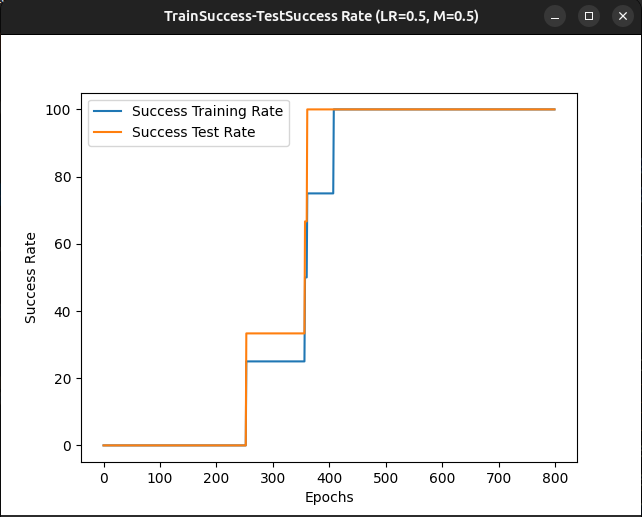
**Γραφική Παράσταση**

Παρακάτω παρουσιάζονται γραφικές παραστάσεις για τα αποτελέσματα της εκπαίδευσης για διάφορες δοκιμές με διαφορετικό seed κατα την αρχικοποίηση των βαρών:

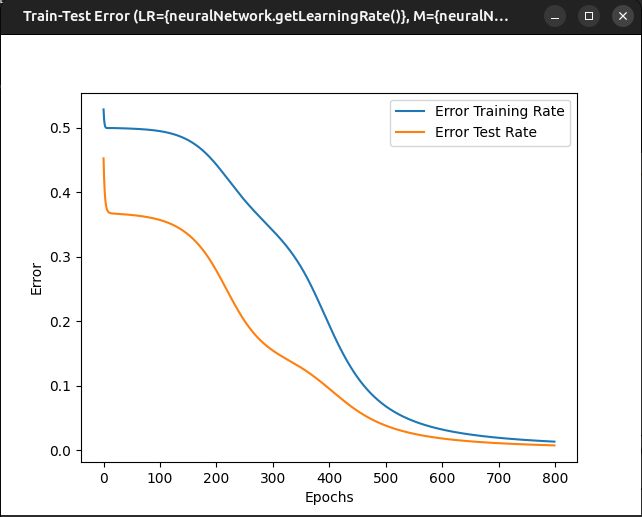
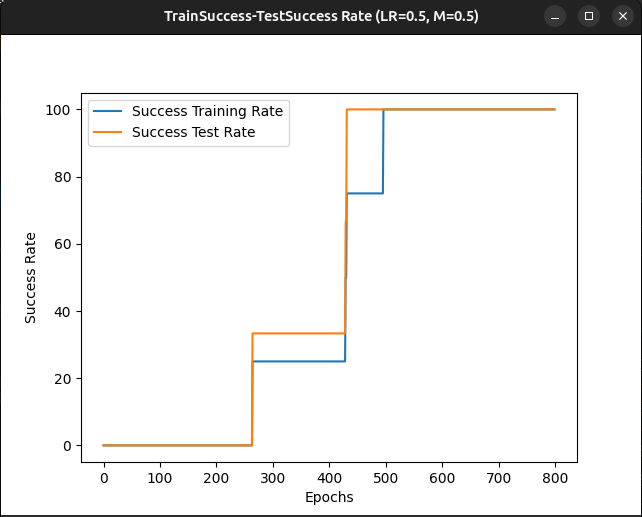
* Γραφική παράσταση του αρχείου errors.txt και success.txt για seed = 6



* Γραφική παράσταση του αρχείου errors.txt και success.txt για seed = 1



* Γραφική παράσταση του αρχείου errors.txt και success.txt για seed = 2



**Συμπεράσματα**

Η επιτυχής εκπαίδευση του νευρωνικού δικτύου να προσεγγίσει τη λογική λειτουργία της πύλης XOR αποδεικνύει την αποτελεσματικότητα του αλγορίθμου backpropagation και της κατάλληλης δομής δικτύου. Αποδεικτικέ πως 1 κρυφό layer με μόνο 2 νευρώνες είναι αρκετό για να λύσει το πρόβλημα. Η γραφική αναπαράσταση των αποτελεσμάτων ενισχύει αυτό το συμπέρασμα, αποτυπώνοντας τη βελτίωση της απόδοσης του δικτύου κατά την πρόοδο της εκπαίδευσης.

**Κλάσεις που χρησιμοποιούνται στον Κώδικα**

Στον παρεχόμενο κώδικα, πολλές προσαρμοσμένες κλάσεις χρησιμοποιούνται για τη δημιουργία και τη διαχείριση ενός νευρωνικού δικτύου, την εκτέλεση feed forward και back propagation και τη διαχείριση λειτουργιών αρχείων. Παρακάτω, θα συζητήσουμε καθεμία από αυτές τις κλάσεις και τους ρόλους τους στην υλοποίηση του νευρωνικού δικτύου.

### 1. ****Neuron****

Η κλάση Neuron αντιπροσωπεύει έναν νευρώνα στο νευρωνικό δίκτυο. Οι νευρώνες είναι οι βασικές μονάδες στο δίκτυο, υπεύθυνες για την επεξεργασία των εισροών, τον υπολογισμό της εξόδου και τη διάδοση του σφάλματος κατά τη διάρκεια της εκπαίδευσης. Τα βασικά χαρακτηριστικά και μέθοδοι αυτής της κλάσης περιλαμβάνουν:

* **Attributes:**
  + id: Identifier for the neuron.
  + connectedFromNeurons: Neurons connected as inputs to this neuron.
  + connectedToNeurons: Neurons connected as outputs to this neuron.
  + output: Output of the neuron after activation.
  + delta: Delta value used in backpropagation.
  + weights: Weights associated with connections to other neurons in the output of this neuron.
  + isBias: Boolean indicating whether the neuron is a bias neuron.
  + isOutput: Boolean indicating whether the neuron is an output neuron.
* **Methods:**
  + calculateOutput(): Calculates the output of the neuron based on its inputs and weights.
  + calculateDelta(): Calculates the delta value of the neuron during backpropagation.
  + updateWeights(): Updates the weights of the neuron based on the calculated delta using a learning rate and momentum.

### 2. ****NeuralNetwork****

Η κλάση NeuralNetwork αντιπροσωπεύει ολόκληρο το νευρωνικό δίκτυο. Ενορχηστρώνει τη δημιουργία στρωμάτων, σύνδεση νευρώνων, αρχικοποίηση βάρους και παρέχει μια διεπαφή για εκπαίδευση. Οι βασικές λειτουργίες περιλαμβάνουν:

* **Attributes:**
  + parameters: Dictionary containing parameters for the neural network.
  + layers: List of layers, each represented as a list of neurons.
* **Methods:**
  + createLayers(): Creates the layers of the neural network based on specified parameters.
  + connectLayers(): Connects neurons in adjacent layers by establishing connections between them.
  + initializeWeights(): Initializes weights for connections between neurons.
  + getLayers(): Returns the layers of the neural network.
  + getLearningRate(): Returns the learning rate from the parameters.
  + getMomentum(): Returns the momentum from the parameters.

### 3. ****FeedForward and BackPropagation****

Αυτές είναι συναρτήσεις για την εκτέλεση των βημάτων feed forward και back propagation στο νευρωνικό δίκτυο. Διευκολύνουν το πέρασμα προς τα εμπρός για τον υπολογισμό της απόδοσης και το πέρασμα προς τα πίσω για την ενημέρωση των βαρών με βάση το σφάλμα.

### 4. ****FileReader and FileWriter****

Αυτές οι κλάσεις χειρίζονται την ανάγνωση και εγγραφή σε αρχεία, που χρησιμοποιούνται ιδιαίτερα για την ανάγνωση δεδομένων εκπαίδευσης και δοκιμής, καθώς και παραμέτρων, και εγγραφή αρχείων καταγραφής σφαλμάτων και ποσοστού επιτυχίας.