

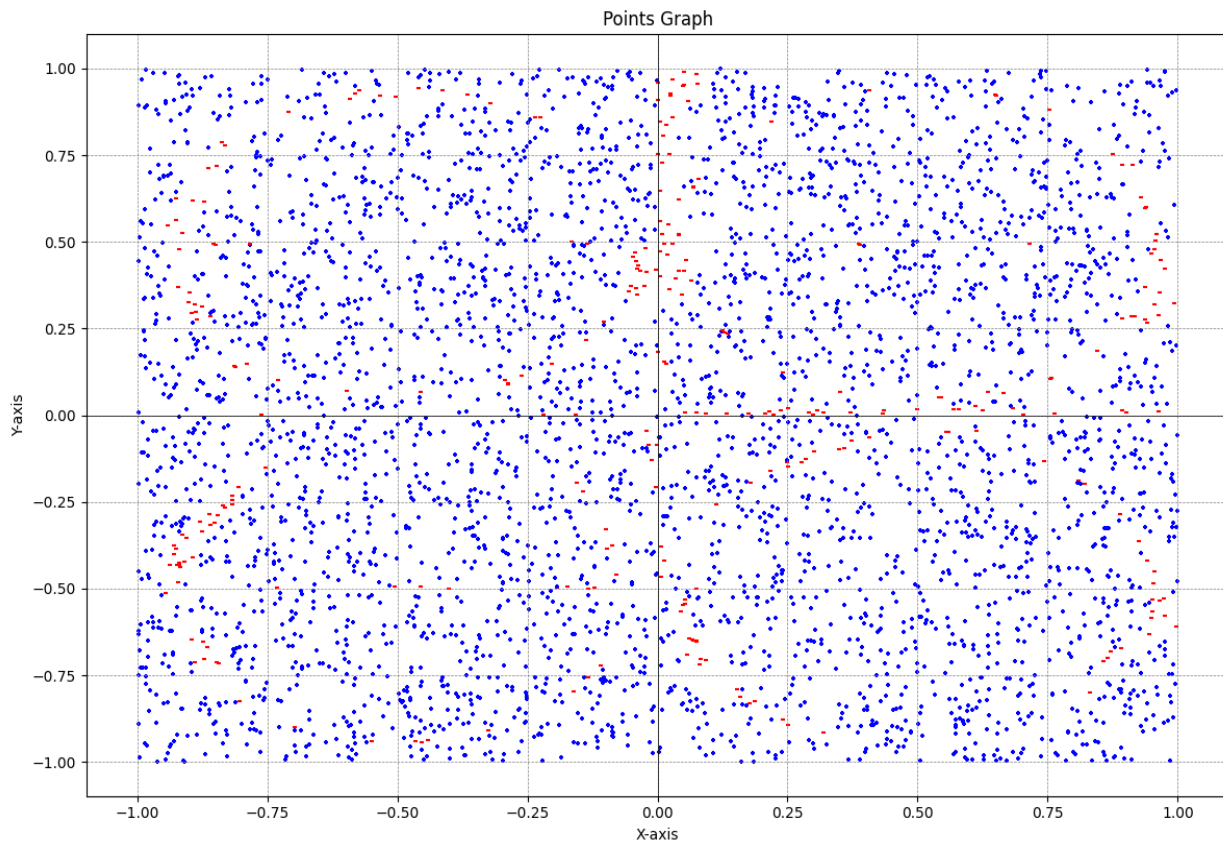
Όν/μο1: Φώτιος Παππάς, ΑΜ: 4773

Όν/μο2: Αργύριος Ζέζος, ΑΜ: 4588

## Αναφορά Πρότζεκτ

### Άσκηση 1

Μετά την δοκιμή 100 τυχαίων συνδυασμών  $H1$  (1 – 15 νευρώνες),  $H2$  (1 – 15 νευρώνες),  $H3$  (0 – 15 νευρώνες) βρέθηκε ότι η καλύτερη γενικευτική ικανότητα (92.4%) πετυχαίνεται με  $H1 = 14$ ,  $H2 = 12$ ,  $H3 = 10$  (3 κρυμμένα επίπεδα), reLU στο τελευταίο κρυμμένο επίπεδο και  $B = 200$ . Παρακάτω φαίνονται τα σημεία από το σύνολο ελέγχου που το συγκεκριμένο δίκτυο κατηγοριοποίησε σωστά (3696/4000) με “+” (μπλε) και τα λάθος (304/4000) με “-” (κόκκινο).



Για τα δίκτυα θεωρήθηκαν τα εξής:

- Χρήση της υπερβολικής εφαπτομένης ( $\tanh[x]$ ) ή reLU για τα κρυμμένα επίπεδα.
- Χρήση της σιγμοειδούς συνάρτησης ( $\sigma[x]$ ) στο επίπεδο εξόδου.
- $B = N/20$  (200) ή  $N/200 = (20)$

Όν/μο1: Φώτιος Παππάς, AM: 4773

Όν/μο2: Αργύριος Ζέζος, AM: 4588

- Ρυθμός μάθησης (Learning Rate) = 0.001
- Κατώφλι τερματισμού (Termination Threshold) = 0.1

Παρατηρώντας τα 20 καλύτερα δίκτυα, φαίνεται ότι το ΠΤ3 δεν υπερτερεί αναγκαστικά:

Αριθμός	Γενικευτικ ή Ικανότητα	H1	H2	H3	Συνάρτηση B Ενεργοποίη σης	
1	0.924	14	12	10	reLU	200
2	0.91525	15	15	0	tahn	200
3	0.88075	13	14	0	reLU	200
4	0.88025	15	12	10	tahn	200
5	0.88	11	15	12	tahn	200
6	0.87575	12	14	6	reLU	200
7	0.86175	10	11	7	reLU	200
8	0.8595	15	11	0	tahn	20
9	0.85225	15	11	0	tahn	20
10	0.851	10	13	5	tahn	200
11	0.85025	10	13	9	reLU	200
12	0.84625	8	10	0	tahn	200
13	0.84025	8	13	0	reLU	200
14	0.83275	15	7	0	tahn	200
15	0.819	15	4	0	tahn	20
16	0.80975	7	13	0	reLU	200
17	0.80575	9	6	0	reLU	200

Όν/μο1: Φώτιος Παππάς, AM: 4773

Όν/μο2: Αργύριος Ζέζος, AM: 4588

18	0.801	10	14	0	tahn	20
19	0.7965	14	4	12	reLU	200
20	0.79	11	15	0	tahn	20

Τα 12/20 δίκτυα στην παραπάνω κατάταξη έχουν 2 κρυμμένα επίπεδα, στην πρώτη δεκάδα 4/10 έχουν 2 κρυμμένα επίπεδα και το δεύτερο καλύτερο δίκτυο που έχει 2 κρυμμένα επίπεδα, έχει ελάχιστη διαφορά από το 1ο, οπότε με βάση αυτά δεν μπορούμε να πούμε ότι κάποιο από τα δύο υπερτερεί. Γενικώς, προκύπτει ότι και τα δύο είδη δικτύων έχουν καλύτερες επιδόσεις με τον αριθμό των νευρώνων σε κάθε κρυμμένο επίπεδο να είναι λίγο μεγαλύτερος από το 10. Όσον αφορά τη συνάρτηση ενεργοποίησης η οποία χρησιμοποιήθηκε στο τελευταίο **κρυμμένο** επίπεδο, 11/20 δίκτυα χρησιμοποίησαν την υπερβολική εφαιπτομένη, οπότε δε φαίνεται η μία να αποδίδει καλύτερα από την άλλη. Αντίθετα, είναι εμφανές ότι όταν η ομαδική ενημέρωση έγινε ανα 200 παραδείγματα, τα αποτελέσματα ήταν καλύτερα.

## Εκτέλεση

Για την εκτέλεση της πρώτης άσκησης:

1) Στο τερματικό εκτελούμε την εντολή: `javac Main_1.java && java Main_1`

2) Στα ακόλουθα ερωτήματα εισάγουμε τις παραμέτρους στο πρόγραμμα:

```
cse94773@dl380ws02:~/projt/ex1$ cse94773@dl380ws02:~/projt/ex1$ javac Main_1.java && java Main_1
Enter the value for H1: 6
Enter the value for H2: 6
Do you want a third hidden layer? (1 for yes, 0 for no): 1
Enter the value for H3: 5
Do you want to use ReLU or Tahn for the last hidden layer? (1 for reLU, 0 for Tahn): 0
Enter the value for B (batch size): 20
```

**Σημείωση:** Το B πρέπει να διαιρεί ακριβώς το 4000 που είναι και το πλήθος των δεδομένων εκπαίδευσης.