

ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ

Διδάσκοντες: Σ. Λυκοθανάσης, Δ. Κουτσομητρόπουλος
Ακαδημαϊκό Έτος 2024-2025

Εργαστηριακή Άσκηση Μέρος Β΄

Στοιχεία Φοιτητή:

Ον/μο: ΚΟΥΤΡΟΥΜΠΕΛΑΣ ΒΑΣΙΛΕΙΟΣ

ΑΜ: 1093397

Εξάμηνο: 8ο

email: up1093397@ac.upatras.gr

Code Repository

<https://github.com/vasiliskoutroumpelas/computational-intelligence>

Περιεχόμενα

B1. Σχεδιασμός ΓΑ.....	2
α) Κωδικοποίηση.....	2
β) Αρχικός πληθυσμός.....	2
γ) Συνάρτηση καταλληλότητας.....	2
δ) Γενετικοί Τελεστές.....	2
B2. Υλοποίηση ΓΑ.....	5
B3. Αξιολόγηση και Επίδραση Παραμέτρων.....	6
α) Πίνακας αποτελεσμάτων.....	6
β) Καμπύλες εξέλιξης.....	7
γ) Σχολιασμός.....	12
B4. Αξιολόγηση ΤΝΔ.....	13
α) Σύγκριση χωρίς επανεκπαίδευση.....	13
β) Σύγκριση με επανεκπαίδευση.....	14

B1. Σχεδιασμός ΓΑ

α) Κωδικοποίηση

Κάθε άτομο στον πληθυσμό του ΓΑ αναπαρίσταται ως ένα δυαδικό διάνυσμα 34 bits. Κάθε bit αντιστοιχεί σε ένα χαρακτηριστικό του ΤΝΔ. Με '1' το χαρακτηριστικό χρησιμοποιείται στο ΤΝΔ, ενώ με '0' το χαρακτηριστικό αγνοείται.

β) Αρχικός πληθυσμός

Μπορεί να παραχθεί με τυχαία δημιουργία ατόμων. Για κάθε άτομο, δημιουργείται ένας δυαδικός πίνακας 34 θέσεων με πιθανότητα $p=0.5$ ανά θέση να είναι 1 (παραμένει το χαρακτηριστικό).

γ) Συνάρτηση καταλληλότητας

$$fitness = A - \lambda \left(\frac{k}{K} \right)$$

Όπου A είναι το accuracy του ΤΝΔ με τις επιλεγμένες εισόδους, k ο αριθμός των ενεργών χαρακτηριστικών, K ο συνολικός αριθμός των χαρακτηριστικών και λ ο συντελεστής βαρύτητας της ποινής.

δ) Γενετικοί Τελεστές

i. Επιλογή

Για την ρουλέτα με βάση το κόστος, κάθε άτομο παίρνει ένα κομμάτι από τον τροχό ανάλογο με το fitness του. Το πρόβλημα είναι πως εξαρτάται απόλυτα από τα μεγέθη του fitness, αν υπάρχει ένα άτομο με πολύ καλό fitness, απορροφά όλη την πιθανότητα και οδηγεί σε απώλεια ποικιλομορφίας. Επίσης, το fitness στο συγκεκριμένο πρόβλημα είναι συνδυασμός ακρίβειας - ποινής, που μπορεί να δημιουργήσει αρνητικά fitness ή πολύ κοντινές τιμές.

Για την ρουλέτα με βάση την κατάταξη, η πιθανότητα επιλογής εξαρτάται από τη θέση στην κατάταξη και όχι απόλυτες τιμές fitness. Αν και υπάρχει σταθερή επιλογή, ακόμα κι αν τα fitness είναι πολύ κοντά, το πρόβλημα είναι πως οι

διαφορές στο fitness δεν αντικατοπτρίζονται ακριβώς στην επιλογή και χάνεται πληροφορία. Μπορεί δηλαδή να υπάρχουν δύο άτομα με πολύ μεγάλη διαφορά στο fitness, αλλά να χειρίζονται ως σχεδόν ίδια.

Για την επιλογή μέσω τουρνουά, διαλέγουμε τυχαία 2 ή 3 άτομα από τον πληθυσμό και επιλέγεται το καλύτερο ως γονέας, η διαδικασία επαναλαμβάνεται για τον 2ο γονέα. Η συγκεκριμένη επιλογή είναι αποδοτική υπολογιστικά, δουλεύει καλά με οποιοδήποτε είδος fitness και με την επιλογή μεγέθους τουρνουά μπορούμε να ελέγξουμε το πόσο ισχυρή είναι η επιλογή.

Καταλήγουμε πως η επιλογή βάση τουρνουά είναι η καταλληλότερη για το συγκεκριμένο πρόβλημα.

ii. Διασταύρωση

Η διασταύρωση μονού σημείου είναι απλός και φυσικός τρόπος αναπαραγωγής που συντηρεί την δομή από τους γονείς. Ωστόσο, επηρεάζει μεγάλο τμήμα του χρωμοσώματος με βάση μόνο ένα σημείο, είναι λιγότερο τυχαίος και μπορεί να κολλήσει σε περιορισμένες δομές.

Στην διασταύρωση πολλαπλού σημείου επιλέγονται 2 ή περισσότερα σημεία και τα τμήματα εναλλάσσονται μεταξύ γονέων. Υπάρχει μεγαλύτερο ανακάτεμα πληροφορίας από ότι στην μονού σημείου, ωστόσο ο διαχωρισμός παραμένει σε σταθερά σημεία ενώ, η δυαδική αναπαράσταση μπορεί να μην παράγει σταθερούς αριθμούς “1” που μπορεί να είναι πρόβλημα στα πλαίσια της άσκησης.

Με την ομοιόμορφη διασταύρωση κάθε bit του απογόνου επιλέγεται με πιθανότητα p από έναν από τους δύο γονείς και δεν εξαρτάται από τη θέση. Με αυτόν τον τρόπο έχουμε πλήρη ανακάτεμα πληροφορίας και δεν εξαρτάται από σταθερά σημεία. Αν και δεν διατηρεί μεγάλη ενότητα με τους γονείς, καταλήγουμε πως η ομοιόμορφη διασταύρωση είναι κατάλληλη για το πρόβλημα μας, καθώς είναι καλή για επιλογή χαρακτηριστικών, αφού φτιάχνει νέα σύνολα εισόδων χωρίς να καταστρέφει πλήρως την δομή.

iii. Ελιτισμός

Ο ελιτισμός είναι στρατηγική διατήρησης των καλύτερων λύσεων στο επόμενο γενεαλογικό βήμα χωρίς καμία αλλαγή. Εγγυάται ότι η καλύτερη λύση δεν θα χαθεί και επιταχύνει τη σύγκλιση του ΓΑ, αφού διατηρούμε “καλές” λύσεις.

B2. Υλοποίηση ΓΑ

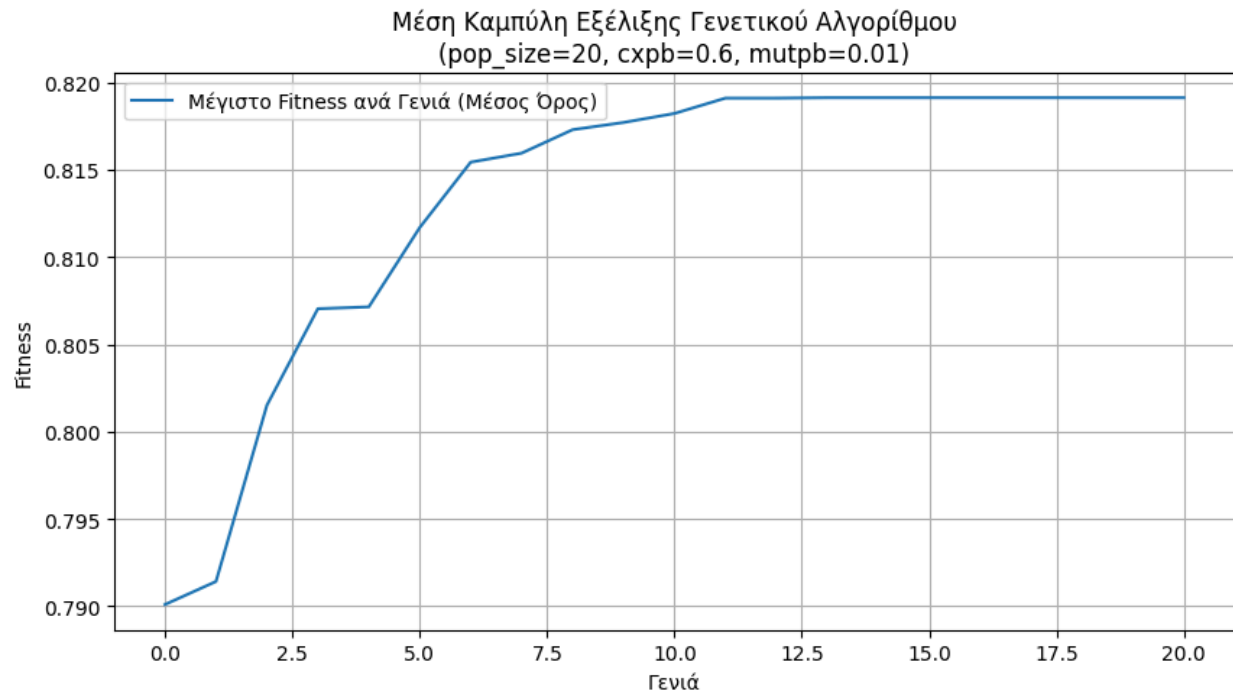
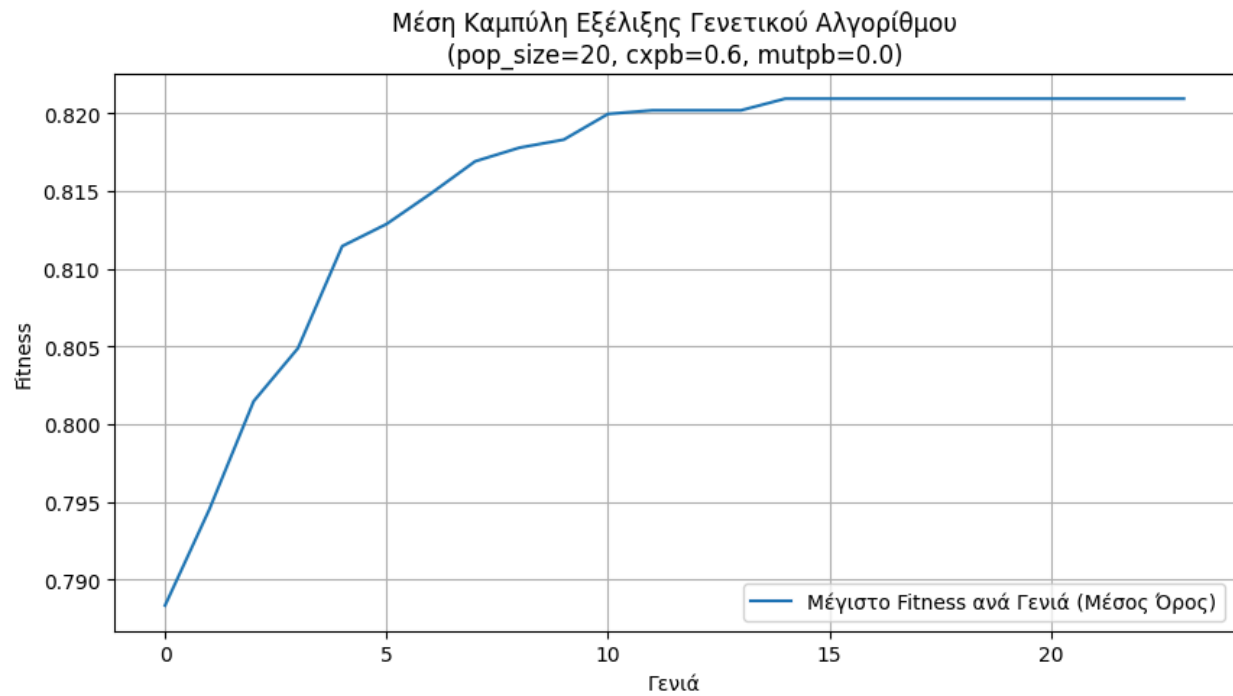
Το πλήρες πρόγραμμα που υλοποιήθηκε βρίσκεται στο code repository. Συνοπτικά, το πρόγραμμα τρέχει το νευρωνικό δίκτυο από το Μέρος Α και αποθηκεύει από το καλύτερο fold το μοντέλο και τα test data. Ορίζουμε την συνάρτηση καταλληλότητας όπως παραπάνω με $\lambda = 0.01$ και χαρακτηριστικά όσα της εισόδου X . Ρυθμίζουμε το εργαλείο DEAP βάση των απαντήσεων του ερωτήματος B1. Έπειτα, ορίζουμε 2 συναρτήσεις `run_ga` και `multiple_run_ga`. Η πρώτη τρέχει τον ΓΑ βάση των επιθυμητών παραμέτρων και εφαρμόζει early stop σε περίπτωση που δεν υπάρχει βελτίωση για 20 γενιές. Η δεύτερη τρέχει την πρώτη για n φορές και κρατάει στατιστικά για κάθε τρέξιμο του αλγορίθμου σχεδιάζοντας την μέση καμπύλη εξέλιξης του ΓΑ. Τέλος, για κάθε τρέξιμο της `multiple_run_ga` κρατάμε τον καλύτερο απόγονο ώστε να συγκριθεί με τα υπόλοιπα τρεξίματα της.

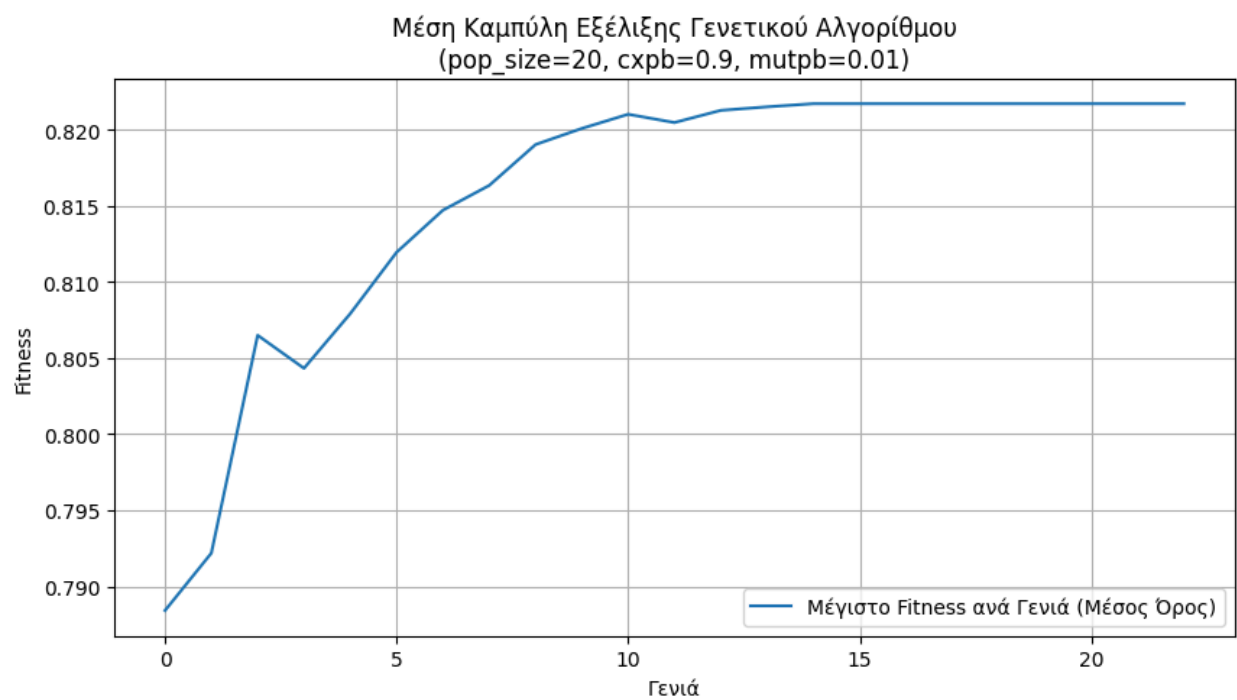
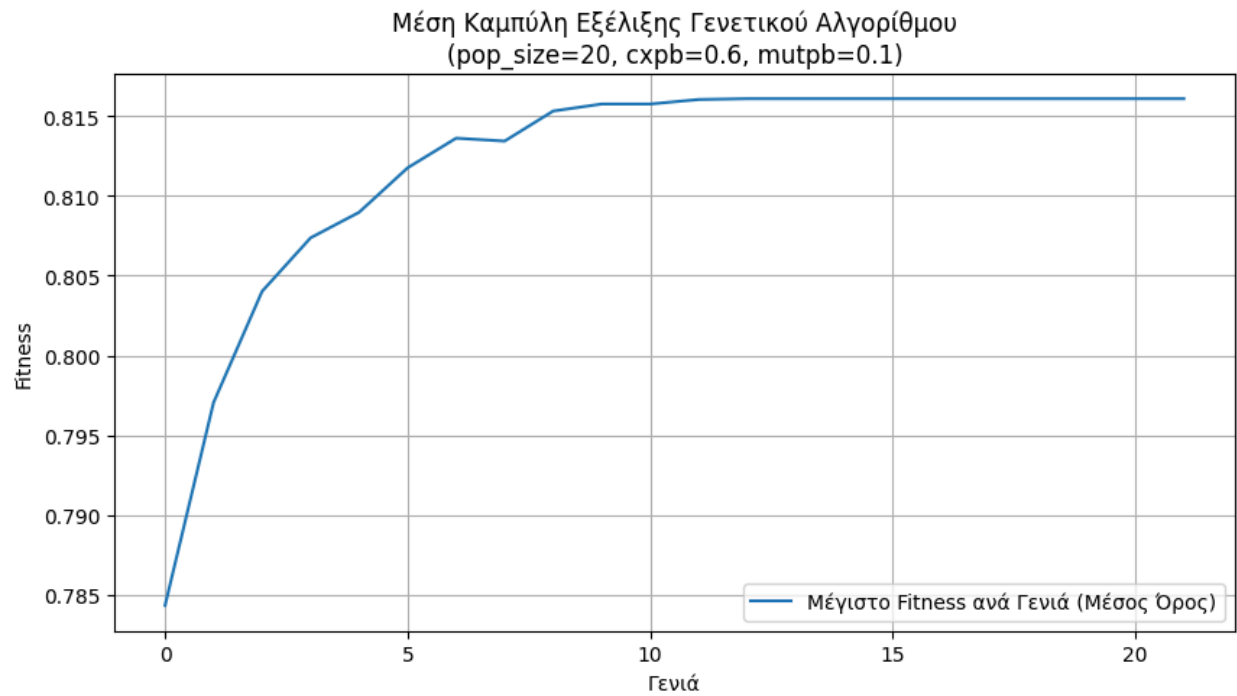
B3. Αξιολόγηση και Επίδραση Παραμέτρων

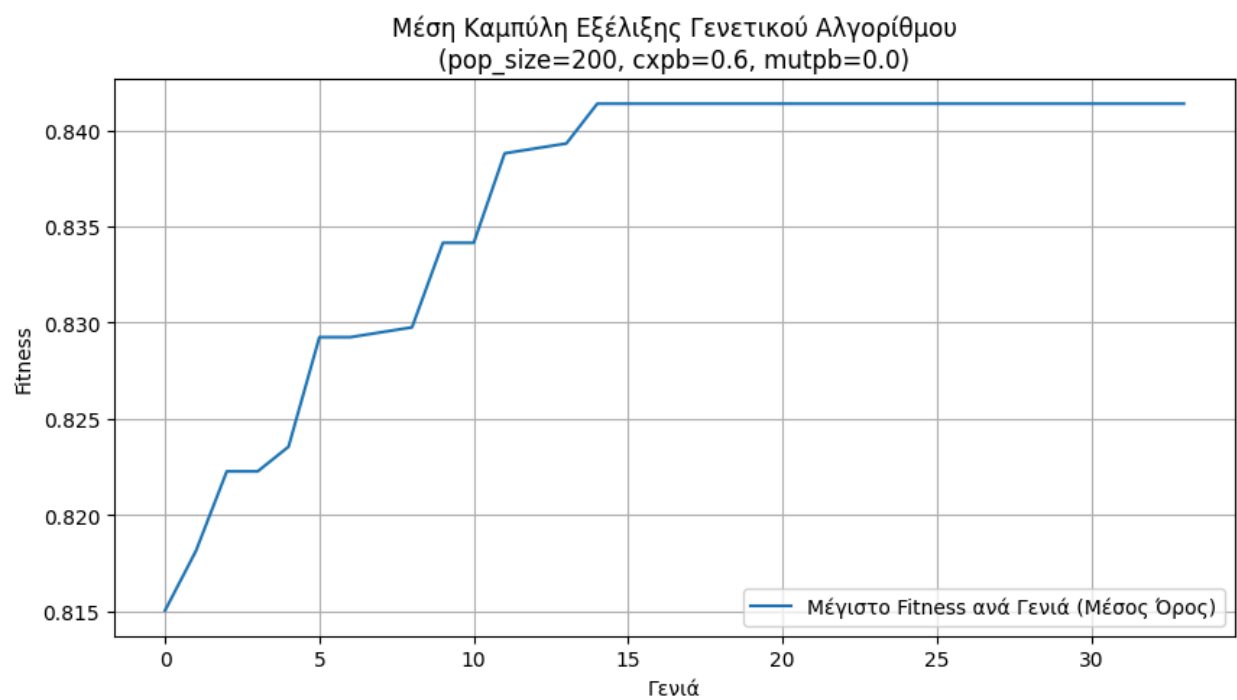
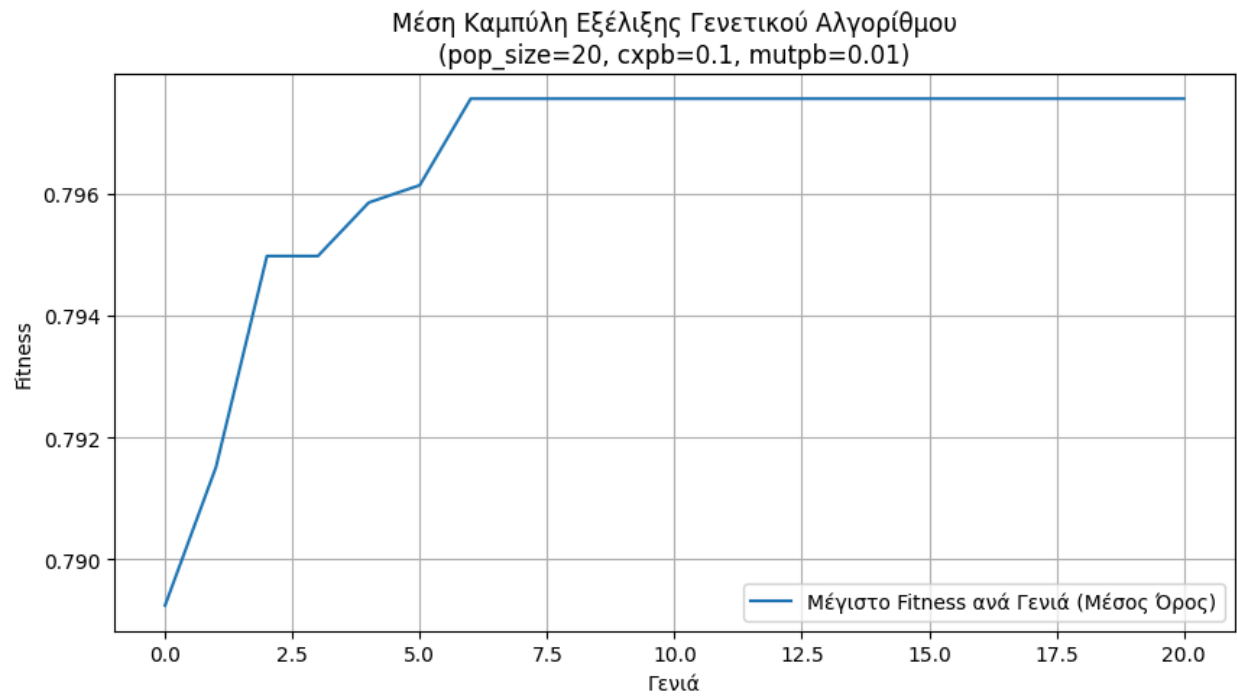
α) Πίνακας αποτελεσμάτων

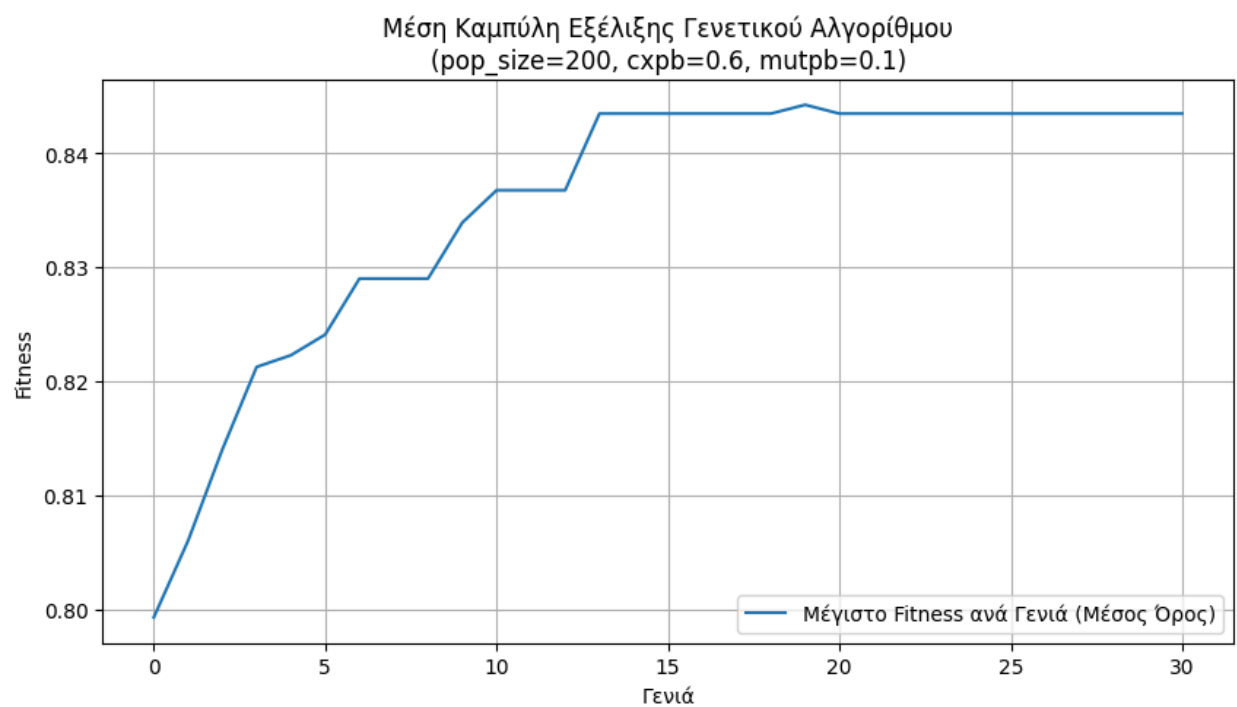
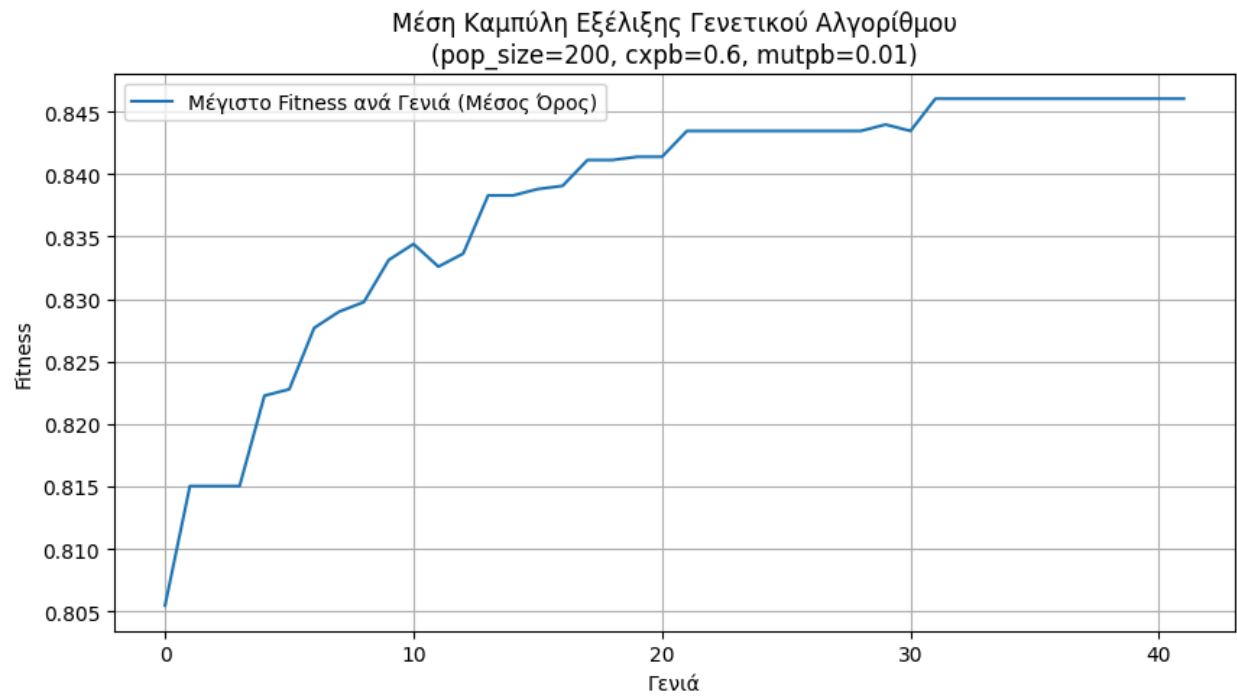
A/A	ΜΕΓΕΘΟΣ ΠΛΗΘΥΣΜΟΥ	ΠΙΘΑΝΟΤΗΤΑ ΔΙΑΣΤΑΥΡΩΣΗΣ	ΠΙΘΑΝΟΤΗΤΑ ΜΕΤΑΛΛΑΞΗΣ	ΜΕΣΗ ΤΙΜΗ ΒΕΛΤΙΣΤΟΥ	ΜΕΣΟΣ ΑΡΙΘΜΟΣ ΓΕΝΕΩΝ
1	20	0.6	0.00	0.8209	27.9
2	20	0.6	0.01	0.8191	24.9
3	20	0.6	0.10	0.8163	24.7
4	20	0.9	0.01	0.8221	25.7
5	20	0.1	0.01	0.7976	20.9
6	200	0.6	0.00	0.8414	33
7	200	0.6	0.01	0.8460	41
8	200	0.6	0.10	0.8442	30
9	200	0.9	0.01	0.8458	32
10	200	0.1	0.01	0.8339	36

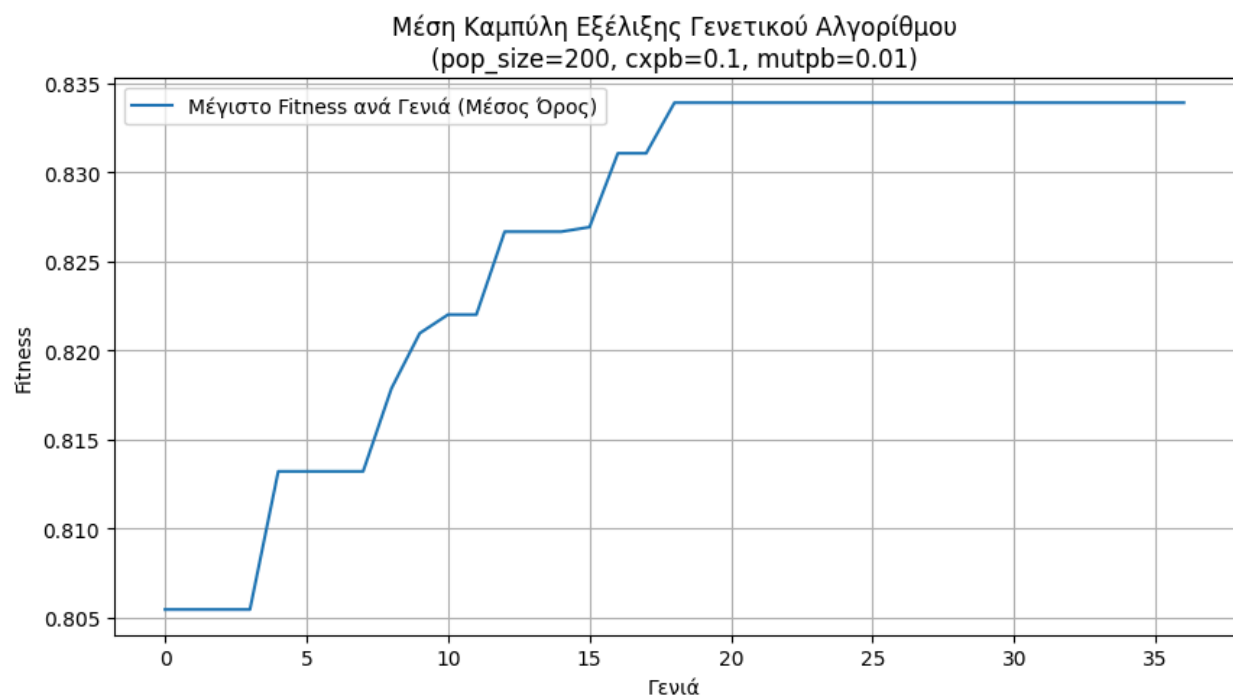
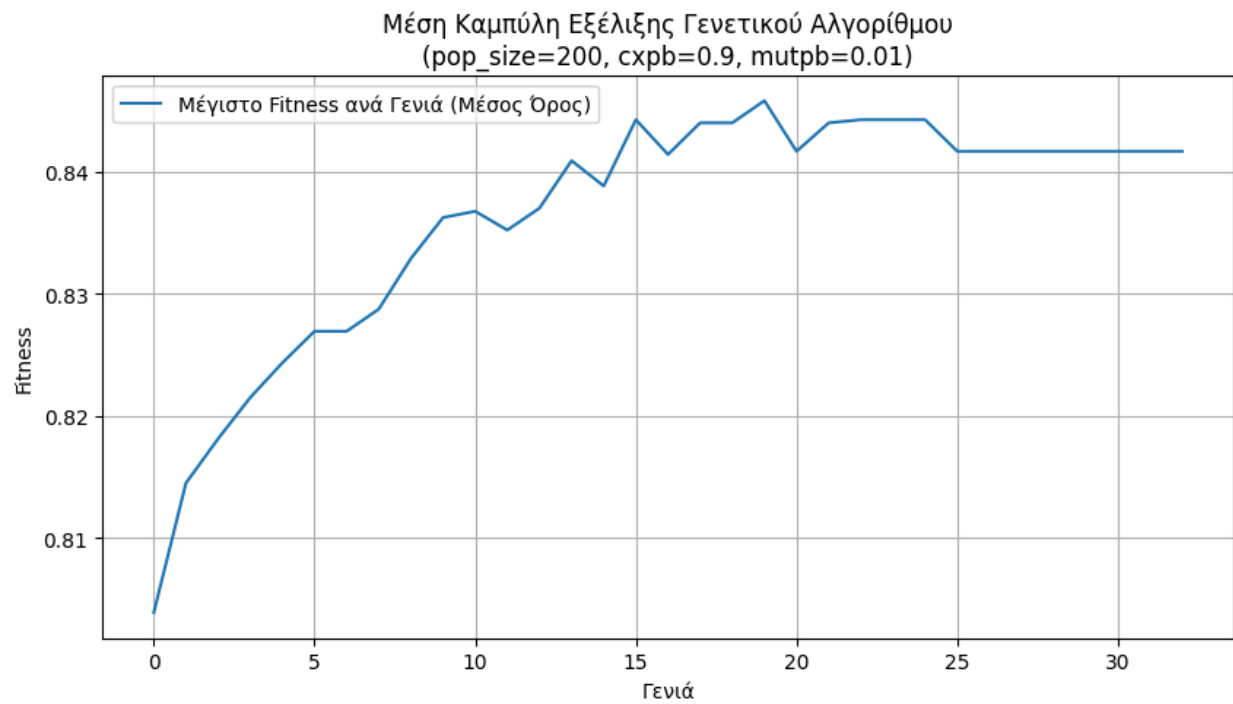
β) Καμπύλες εξέλιξης











γ) Σχολιασμός

Σημειώνεται πως για τα τρεξίματα 6-10 όπου το μέγεθος του πληθυσμού είναι 200 εκτελέστηκαν 1 φορά αντί 10 για οικονομία χρόνου, αφού οι απαραίτητοι υπολογισμοί είναι πολλοί περισσότεροι.

Με την αλλαγή του population size από 20 σε 200 παρατηρούμε μέσω της καμπύλης σύγκλισης και του πίνακα πειραμάτων πως ο αλγόριθμος αργεί μερικές γενιές παραπάνω να φτάσει σε σύγκλιση, αλλά βρίσκει καλύτερες λύσεις. Η συμπεριφορά αυτή εξηγείται από το γεγονός ότι ένας μεγαλύτερος πληθυσμός εισάγει μεγαλύτερη ποικιλομορφία στις λύσεις, επιτρέποντας στον αλγόριθμο να εξερευνήσει πληρέστερα τον χώρο λύσεων και να αποφύγει την πρόωρη σύγκλιση σε τοπικά άκρα. Ωστόσο, η αυξημένη ποικιλία επιβραδύνει την ταχύτητα σύγκλισης, καθώς η διαδικασία βελτίωσης ανά γενιά είναι πιο σταδιακή.

Με την αλλαγή της πιθανότητας διασταύρωσης από 0.6 σε 0.9 ο ΓΑ φαίνεται να βελτιώνει ελαφρώς τη μέση τιμή του βέλτιστου fitness. Αυτό οφείλεται στο ότι η διασταύρωση συμβάλλει σημαντικά στη δημιουργία νέων συνδυασμών χαρακτηριστικών, επιταχύνοντας την εύρεση καλύτερων λύσεων.

Με την αλλαγή της πιθανότητας διασταύρωσης σε 0.1 μειώνεται αισθητά η απόδοση (0.7976) και τρέχει για λιγότερες γενιές, διότι ο αλγόριθμος βασίζεται κυρίως στη μετάλλαξη και λιγότερο στον ανασυνδυασμό των καλύτερων ατόμων, κάτι που περιορίζει την εξελικτική πρόοδο.

Παρατηρώντας την επίδραση της πιθανότητας μετάλλαξης στα αποτελέσματα του Γενετικού Αλγορίθμου, διαπιστώνουμε ότι η εισαγωγή χαμηλής τιμής (0.01) συμβάλλει στη μείωση του αριθμού γενεών που απαιτούνται για σύγκλιση (από 27.9 σε 24.9), χωρίς σημαντική μείωση της απόδοσης (από 0.8209 σε 0.8191). Αυτό υποδηλώνει ότι μια ελαφριά μετάλλαξη βοηθά την εξερεύνηση του χώρου λύσεων και αποτρέπει την πρόωρη σύγκλιση. Αντίθετα, υψηλότερη πιθανότητα μετάλλαξης (0.10) οδηγεί σε περαιτέρω μείωση της απόδοσης (0.8163), χωρίς σημαντικά οφέλη στη σύγκλιση.

B4. Αξιολόγηση ΤΝΔ

α) Σύγκριση χωρίς επανεκπαίδευση

i. Γενικευτική ικανότητα

Η σύγκριση της ακριβείας δείχνει ότι το μοντέλο με τα επιλεγμένα χαρακτηριστικά (0.8512) παρουσιάζει καλύτερη απόδοση από το αρχικό μοντέλο (0.8163), γεγονός που υποδηλώνει αυξημένη γενικευτική ικανότητα. Αυτό οφείλεται πιθανώς στο ότι η επιλογή χαρακτηριστικών απομάκρυνε άσχετες ή θορυβώδεις μεταβλητές, επιτρέποντας στο δίκτυο να επικεντρωθεί στις πιο σημαντικές για τη διάκριση των κλάσεων.

ii. Επίδραση μείωσης χαρακτηριστικών

Συγκρίνοντας τις 2 ακρίβειες με μείωση χαρακτηριστικών (20 χαρακτηριστικά από τα 39 συνολικά) και με το αρχικό μοντέλο παρατηρούμε αύξηση στην πρώτη περίπτωση (0.8512) σε σχέση με την άλλη (0.8163)

iii. Υπερπροσαρμογή στα δεδομένα ελέγχου

Παρά τη βελτιωμένη απόδοση με τα επιλεγμένα χαρακτηριστικά, υπάρχει πιθανότητα υπερπροσαρμογής στα δεδομένα ελέγχου, καθώς το ίδιο validation set χρησιμοποιήθηκε επανειλημμένα κατά την εξέλιξη του γενετικού αλγορίθμου. Το γεγονός αυτό μπορεί να έχει οδηγήσει τον αλγόριθμο στην εύρεση υποσυνόλων χαρακτηριστικών που ταιριάζουν πολύ καλά στα δεδομένα ελέγχου, χωρίς να εγγυώνται ανάλογη απόδοση σε άγνωστα δεδομένα.

β) Σύγκριση με επανεκπαίδευση

Μέρος Α:

Fold 0: Loss=0.5142, MSE=0.1456, Accuracy=0.8163

Fold 1: Loss=0.5358, MSE=0.1546, Accuracy=0.7884

Fold 2: Loss=0.5332, MSE=0.1555, Accuracy=0.7814

Fold 3: Loss=0.4927, MSE=0.1396, Accuracy=0.8209

Fold 4: Loss=0.5322, MSE=0.1543, Accuracy=0.7832

Average Loss: 0.5216

Average MSE: 0.1499

Average Accuracy: 0.7980

Μέρος Β:

Fold 0: Loss=0.5082, MSE=0.1546, Accuracy=0.7814

Fold 1: Loss=0.4818, MSE=0.1437, Accuracy=0.8163

Fold 2: Loss=0.4816, MSE=0.1434, Accuracy=0.8047

Fold 3: Loss=0.4674, MSE=0.1358, Accuracy=0.8163

Fold 4: Loss=0.4967, MSE=0.1490, Accuracy=0.8112

Average Loss: 0.4871

Average MSE: 0.1453

Average Accuracy: 0.8060

i. Γενικευτική ικανότητα

Το μοντέλο του Μέρους Β, το οποίο εκπαιδεύτηκε με επιλεγμένα χαρακτηριστικά, παρουσίασε ελαφρώς καλύτερη μέση ακρίβεια (0.8060 έναντι 0.7980 του αρχικού μοντέλου) και μικρότερη μέση τιμή απωλειών (Loss: 0.4871 έναντι 0.5216). Αυτό υποδηλώνει μια μικρή βελτιωμένη γενικευτική ικανότητα.

ii. Επίδραση μείωσης χαρακτηριστικών

Η μείωση των χαρακτηριστικών από 39 σε 20 οδήγησε σε βελτίωση της απόδοσης του δικτύου τόσο από πλευράς ακρίβειας όσο και στατιστικών σφαλμάτων (Loss, MSE). Η επιλογή χαρακτηριστικών μέσω Γενετικού Αλγορίθμου φαίνεται να

ενίσχυσε τη μάθηση του δικτύου, απομακρύνοντας περιττές ή θορυβώδεις μεταβλητές που δε συνέβαλε ουσιαστικά στη διάκριση των κατηγοριών.

iii. Υπερπροσαρμογή στα δεδομένα ελέγχου

Το δίκτυο που εκπαιδεύτηκε με επιλεγμένα χαρακτηριστικά σε σχέση με το δίκτυο με όλα τα χαρακτηριστικά απέδωσε καλύτερα στα δεδομένα επικύρωσης χωρίς ωστόσο να έχουν σημαντικές διαφορές ώστε να αποφανθούμε πως κάποιο από τα 2 υπερπροσαρμόζεται.