



**ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΠΑΤΡΩΝ**  
UNIVERSITY OF PATRAS

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ  
Η/Υ ΚΑΙ  
ΠΛΗΡΟΦΟΡΙΚΗΣ

## **ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ**

Ακαδημαϊκό Έτος 2021-2022

### **Εργαστηριακή Άσκηση Μέρος Β'**

**Γεώργιος Κοντογιάννης**  
**1070908 – Δ' έτος**

**Code Repo Link:**

**[https://github.com/gkontogiannhs/GA\\_From\\_Scratch](https://github.com/gkontogiannhs/GA_From_Scratch)**

## B1. Σχεδιασμός ΓΑ

### *Κωδικοποίηση*

Καθώς ένα άτομο του πληθυσμού αναπαριστά ένα τυχαίο λεξικό του dataset, κάθε άτομο θα πρέπει να αποτελείται από 8.520 στοιχεία. Η κωδικοποίηση θα είναι δυαδική, με το «1» να υποδηλώνει την ύπαρξη της λέξης στο λεξικό ενώ με «0» όχι. Πιο συγκεκριμένα, κάθε άτομο-λεξικό θα είναι μια sparse λίστα από 0 και 1 του οποίου ο δείκτης θέσης αντιπροσωπεύει τη λέξη του λεξικού. Για παράδειγμα, έστω λίστα (λεξικό) ενός ατόμου του οποίου η τιμή στη θέση 69 είναι «1». Ανατρέχοντας το λεξικό `vocabs.txt`, αυτό σημαίνει πως αυτό το άτομο περιέχει τη λέξη «canari».

### *Αρχικός Πληθυσμός*

Έστω `pop_size` το επιθυμητό μέγεθος πληθυσμού. Για `pop_size` επαναλήψεις, θα δημιουργούνται τυχαία άτομα, δηλαδή τυχαία δυαδικά sparse διανύσματα με πιθανότητα εμφάνισης «0» μεγαλύτερη του  $1/2$  ενώ πιθανότητα εμφάνισης «1» μικρότερη του  $1/2$ . Είναι σημαντικό να πολώσουμε ανομοιόμορφα τις δύο αυτές πιθανότητες καθώς θέλουμε τα άτομα να είναι αυστηρά κοντά στο εύρος λέξεων [1000, 2000]. Τέλος, κάθε άτομο του πληθυσμού αποτελεί στοιχείο μιας 2D-λίστας.

### *Διαδικασία επιδιόρθωσης*

i) Η τεχνική της απόρριψης μιας μη νόμιμης λύσης από τον πληθυσμό και αντικατάστασής της από κάποιο άλλο άτομο, είτε τυχαία είτε με ελιτισμό, είναι εύκολη και γρήγορη. Το κυριότερο μειονέκτημα της ωστόσο είναι ότι συγκλίνει γρήγορα σε τοπικό ελάχιστο, καθώς περιορίζουμε συνεχώς το πληθυσμό να έχει πολλά αντίγραφα ατόμων. Αυτό έχει ως αποτέλεσμα οι απόγονοι να είναι αντίγραφα των γονέων το οποίο σημαίνει μικρή πολυμορφία. Για το λόγο αυτό την απορρίπτουμε.

ii) Η τεχνική της επιδιόρθωσης αποτελεί επίσης μια γρήγορη και εύκολη τεχνική αλλά πρακτικά δε προσφέρει κάτι περισσότερο από το συνδυασμό εφαρμογής των γενετικών τελεστών διασταύρωσης και μετάλλαξης. Για το λόγο αυτό την απορρίπτουμε.

## B1. Σχεδιασμός ΓΑ

iii) Η εφαρμογή ποινής σε μη νόμιμη λύση μέσω της συνάρτησης καταλληλότητας είναι η καλύτερη επιλογή για το πρόβλημα μας. Αυτό γιατί διατηρούμε τη πληροφορία του ατόμου αυτού, βεβαιωνόμαστε πώς δε θα κυριαρχήσει (απο το βαθμό καταλληλότητας) και στις επόμενες γενεές είναι πολύ πιθανό να προσδώσει στη δημιουργία ενός άλλου νόμιμου. Με το τρόπο αυτό είναι λιγότερο πιθανό να συγκλίνει γρήγορα σε τοπικό ελάχιστο λόγω έλλειψης πολυμορφίας, όπως στο (i). Οι ποινές θα εφαρμόζονται όταν το πλήθος των λέξεων ενός ατόμου είναι εκτός του εύρους [1000, 2500]. Στη περίπτωση όπου το άτομο έχει λιγότερες απο 1000, η καταλληλότητα θα διαιρείται με τη ποσότητα  $1000/\text{πλήθος λέξεων(άσσω)}$ . Όταν το άτομο έχει περισσότερες απο 2500 λέξεις, θα διαιρούμε τη τιμή καταλληλότητας με το μήκος των «παραπανίστιων» λέξεων. Για παράδειγμα, αν ένα άτομο έχει 2800 λέξεις, τότε η καταλληλότητα διαιρείται με το  $(2800 - 2500) = 300$ . Με το τρόπο αυτό εξασφαλίζουμε πώς οι ποινές είναι ανάλογες της απομάκρυνσης απο τη λύση και ότι θα είναι τα λιγότερο κατάλληλα άτομα του πληθυσμού, ώστε η σύγκλιση να πραγματοποιηθεί στο επιθυμητό εύρος.

### *Υπολογισμός tf-idf*

Για τη συνάρτηση καταλληλότητας αρχικά θα υπολογίσουμε το document term matrix (dtm) των tf\_idf τιμών. Με το τρόπο αυτό κάθε στήλη αντιπροσωπεύει έναν όρο t ενώ κάθε γραμμή ένα κείμενο d. Άρα, το  $\text{dtm}[t_i, d_j]$  μας δίνει τη τιμή του tf\_idf αυτού του συνδυασμού (όρου-κειμένου). Για να πάρουμε τον M.O. tf\_idf κάθε λέξης-όρου, δεν έχουμε παρά να υπολογίσουμε τον μέσο όρο ανά στήλη του μητρώου δίχως να συμπεριλάβουμε τα κείμενα τα οποία δεν περιέχουν τον όρο αυτό.

***Βλέπε Κώδικα, συνάρτηση `calc_tdif_means()` function.***

### *Συνάρτηση καταλληλότητας*

Έχοντας τους M.O. των tfidf στη διάθεση μας, μπορούμε να συνδυάσουμε το κριτήριο (i), για να δούμε πόσο σημαντικές λέξεις περιέχει κάθε άτομο. Άρα, για κάθε άτομο, υπολογίζουμε το άθροισμα των tf\_idf τιμών του κάθε όρου που εμφανίζεται. Ωστόσο, αυτή η ποσότητα καθαυτή δεν είναι καλή μετρική, καθώς όσο περισσότερους όρους του λεξικού περιέχει ένα άτομο, τόσο μεγαλύτερο θα είναι το άθροισμα αυτό. Για το λόγο αυτό, διαιρούμαι τη ποσότητα αυτή με το πλήθος των όρων του ατόμου, έτσι ώστε να μη κυριαρχούν τα πιο πυκνά διανύσματα, συνδυάζοντας έτσι και το δεύτερο ανταγωνιστικό κριτήριο. Τέλος διαιρούμε το μήκος των ατομών (τους άσσους) με το 1000 ώστε η τελική διαίρεση να μη πολώνεται λόγω της μεγάλης διαφοράς τάξης του αριθμητή-παρονομαστή.

## B1. Σχεδιασμός ΓΑ

Ο παραπάνω υπολογισμός συμβαίνει στη περίπτωση όπου το πλήθος των λέξεων των ατόμων βρίσκεται στο εύρος [1000, 2500]. Σε κάθε άλλη περίπτωση εφαρμόζεται η ποινή που προτάθηκε στο παραπάνω ερώτημα. Η μέγιστη τιμή καταλληλότητας που μπορεί να λάβει ένα άτομο είναι η περίπτωση όπου θα περιέχει 1000 ακριβώς λέξεις, εκ των οποίων είναι όλες οι πιο σημαντικές, δηλαδή με τις μεγαλύτερες τιμές  $tf\_idf$ . Στη περίπτωση μας αυτή είναι η τιμή 234.371.

## B1. Σχεδιασμός ΓΑ

### *Γενετικοί Τελεστές*

#### *Επιλογή Πληθυσμού*

(i) Στη ρουλέτα με βάση το κόστος, η πιθανότητα επιλογής ενός ατόμου για να λειτουργήσει ως γονέας είναι ευθέως ανάλογη της τιμής καταλληλότητας του. Επομένως, η πιθανότητα επιλογής ενός γονέα είναι ανάλογη του ποσοστού ποιότητας που συνεισφέρει το άτομο στην συνολική ποιότητα του πληθυσμού. Ωστόσο, η τεχνική αυτή ευνοεί τα άτομα υψηλής καταλληλότητας αλλά το κάνει σε βάρος των λιγότερο ποιοτικών. Αυτή η προτίμηση είναι επιθυμητή αλλά όχι σε τόσο έντονο βαθμό, ώστε να αποτραπούν φαινόμενα πρόωρης σύκλισης του πληθυσμού.

(ii) Στη ρουλέτα με βάση τη κατάταξη, η πιθανότητα επιλογής για ένα άτομο εξαρτάται από τη θέση που κατέχει το άτομο στον ταξινομημένο ως προς την καταλληλότητα πληθυσμό και όχι στην τιμή καταλληλότητας που έχει το ίδιο. Άρα, η τεχνική αυτή μειώνει λίγο τη πιθανότητα επιλογής των ισχυρών ατόμων και ενισχύοντας εκείνη των αδύναμων, χωρίς όμως να χαλάει την ποιοτική κατάταξη. Με το τρόπο αυτό, τυχόν καλά γονίδια σε άτομα κακής ποιότητας, έχουν περισσότερες πιθανότητες να εμπλακούν σε διαδικασία διασταύρωσης και να συνδυαστούν με άλλα καλά γονίδια δίνοντας πιθανώς ακόμη καλύτερα άτομα απογόνους.

(iii) Η επιλογή τουρνουά, δίνει ακόμη μεγαλύτερη πιθανότητα επιλογής ως γονέα σε άτομα χαμηλής καταλληλότητας, καθώς κάθε φορά επιλέγεται τυχαία το καλύτερο άτομο από ένα υποσύνολο του πληθυσμού, μέχρις ότου συγκεντρωθεί ο απαραίτητος αριθμός. Η συγκεκριμένη τεχνική δίνει μεγάλη στοχαστικότητα στην επιλογή της νέας γενιάς.

## B1. Σχεδιασμός ΓΑ

### *Διασταύρωση Πληθυσμού*

(i) Στη διασταύρωση ενός σημείου, ορίζεται με τυχαίο τρόπο ένα σημείο διασταύρωσης σε εσωτερική θέση της αλληλουχίας των bit. Οι δύο δεξιές, απο το σημείο κοπής, υποσυμβολοσειρές γίνονται «ουρές» των των δύο νέων απογόνων. Η τεχνική αυτή είναι υπολογιστικά ανεκτή αλλά με μειονέκτημα πως οι γονείς ανταλλάσσουν πάντα ακραία τμήματα των γονιδίων τους. Παρόμοια, οι αλληλουχίες bit που μπορεί να παραχθούν (ή να καταστραφούν) συνδυάζοντας δύο άτομα, εξαρτώνται σε κάποιο βαθμό από τη θέση τους στο άτομο. Στη δική μας περίπτωση, λόγω του μεγάλου μήκους των ατόμων η τεχνική αυτή δε μας φαίνεται αρκετή.

(ii) Στη διασταύρωση πολλαπλών σημείων ορίζονται με τυχαίο τρόπο N σημεία διασταύρωσης σε εσωτερικές θέσεις της αλληλουχίας των bit, όχι δηλαδή στην αρχή ή στο τέλος. Οι επιμέρους αλληλουχίες bit που ορίζουν τα σημεία διασταύρωσης δίνονται εναλλάξ στους απογόνους. Αυτή η τεχνική διασταύρωσης ίσως ξεπεράσει το λόγο που δε προτιμάται του ενός σημείου.

(iii) Στην ομοιόμορφη διασταύρωση κάθε bit από τα γονίδια του ενός γονέα δίνεται με τυχαίο τρόπο στις αντίστοιχες θέσεις των απογόνων. Στη συνέχεια κάθε απόγονος συμπληρώνει τα bit που του λείπουν απο τα bit αντίστοιχης θέσης του άλλου γονέα. Με άλλα λόγια γίνεται ανταλλαγή bit σε κάποιες θέσεις. Το μειονέκτημα που έχει η τεχνική αυτή είναι ότι δεν αφήνει εύκολα να επικρατήσουν οι υποτιθέμενες «καλές» αλληλουχίες απο bit.

### *Παρατηρήσεις*

Βασιζόμενοι στο βασικό μειονέκτημα της διασταύρωσης ενός σημείου, η διασταύρωση N σημείων ή η ομοιόμορφη διασταύρωση ίσως είναι περισσότερο προτιμότερες τεχνικές. Τέλος να πούμε, πως ενώ ένα απο τα πιο ελκυστικά χαρακτηριστικά της διασταύρωσης ενός σημείου είναι η εξονυχιστική αναζήτηση λύσεων, λόγω του παράγοντα διασποράς β. Βασιζόμενοι στη καμπύλη καταλληλότητας του προβλήματος (απότομη κλίση), αυτό δεν μας είναι τόσο χρήσιμο.

## **B1. Σχεδιασμός ΓΑ**

### ***Μετάλλαξη Πληθυσμού***

(i) Η μετάλλαξη είναι η διαδικασία που μπορεί να βοηθήσει πολύ τον αλγόριθμο στο να μη συγκλίνει γρήγορα, ωστόσο, στο συγκεκριμένο πρόβλημα η μετάλλαξη θέλει προσοχή και η εφαρμογή ελιτισμού είναι απαραίτητη. Τα διανύσματα των ατόμων (γονίδια), έχουν μεγάλο μήκος και ακόμα και με μικρή πιθανότητα μετάλλαξης, όπως το  $1e-2$ , τα νέα άτομα του πληθυσμού επηρεάζονται τόσο ώστε ο αλγόριθμος να συγκλίνει σε μη ικανοποιητική λύση.

### ***Σχεδιαστικές Αποφάσεις***

Έχοντας αναφέρει όλα τα παραπάνω, η καλύτερη προσέγγιση που εκτιμήθηκε είναι να χρησιμοποιηθεί η ρουλέτα με βάση τη κατάταξη για την επιλογή, η διασταύρωση πολλαπλού σημείου ( $N=50$ ) και μετάλλαξη με ελιτισμό.

## **B2. Υλοποίηση ΓΑ**

### ***Βλέπε κώδικα***

### **B3. Αξιολόγηση και Επίδραση Παραμέτρων**

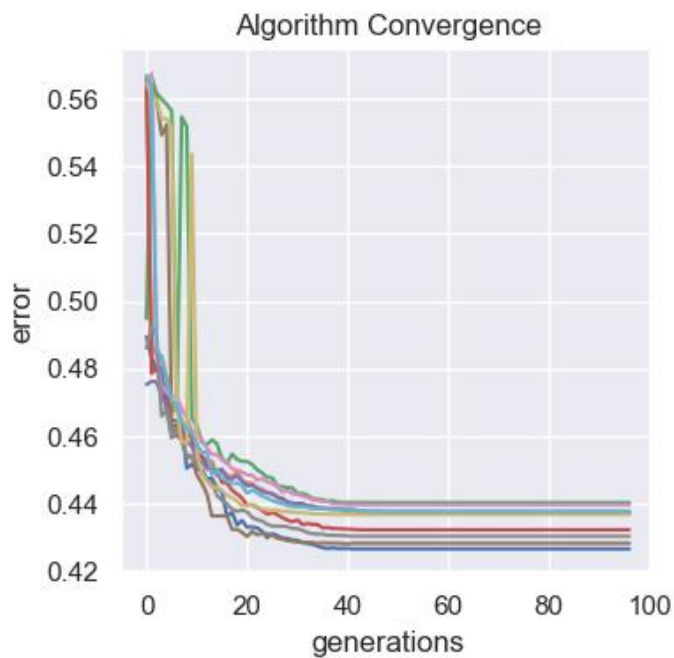
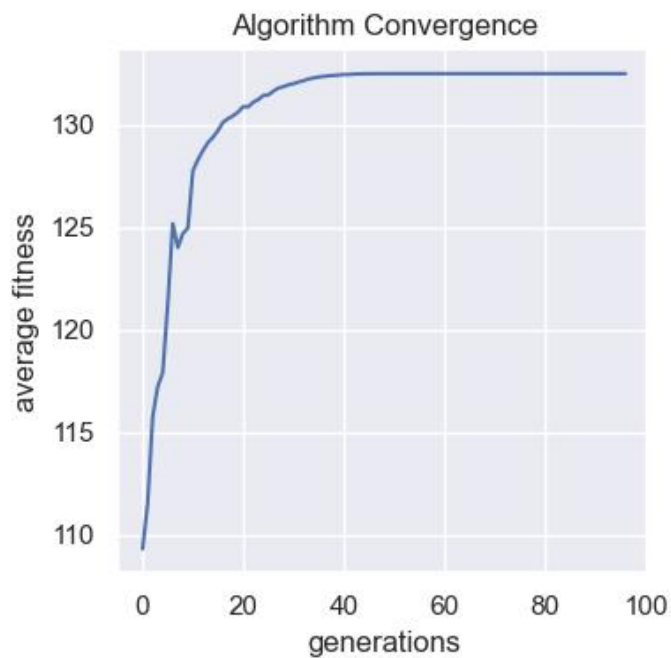
<b>A/A</b>	<b>ΜΕΓΕΘΟΣ ΠΛΗΘΥΣΜΟΥ</b>	<b>ΠΙΘΑΝΟΤΗΤΑ ΔΙΑΣΤΑΥΡΩΣΗΣ</b>	<b>ΠΙΘΑΝΟΤΗΤΑ ΜΕΤΑΛΛΑΞΗΣ</b>	<b>Μ.Τ. ΒΕΛΤΙΣΤΟΥ</b>	<b>Μ.Α. ΓΕΝΕΩΝ</b>
<i>1</i>	<i>20</i>	<i>0.6</i>	<i>0.00</i>	<i>132.482</i>	<i>90.0</i>
<i>2</i>	<i>20</i>	<i>0.6</i>	<i>0.01</i>	<i>128.898</i>	<i>87.1</i>
<i>3</i>	<i>20</i>	<i>0.6</i>	<i>0.10</i>	<i>123.432</i>	<i>84.0</i>
<i>4</i>	<i>20</i>	<i>0.9</i>	<i>0.01</i>	<i>129.109</i>	<i>90.1</i>
<i>5</i>	<i>20</i>	<i>0.1</i>	<i>0.01</i>	<i>128.900</i>	<i>84.3</i>
<i>5.1</i>	<i>20</i>	<i>0.6</i>	<i>0.001</i>	<i>140.294</i>	<i>87.7</i>
<i>5.2</i>	<i>20</i>	<i>0.9</i>	<i>0.001</i>	<i>140.995</i>	<i>88.7</i>
<i>5.3</i>	<i>20</i>	<i>0.9</i>	<i>0.00</i>	<i>134.510</i>	<i>117.7</i>
<i>6</i>	<i>200</i>	<i>0.6</i>	<i>0.00</i>	<i>184.501</i>	<i>289.8</i>
<i>7</i>	<i>200</i>	<i>0.6</i>	<i>0.01</i>	<i>133.027</i>	<i>96.1</i>
<i>8</i>	<i>200</i>	<i>0.6</i>	<i>0.10</i>	<i>124.569</i>	<i>82.5</i>
<i>9</i>	<i>200</i>	<i>0.9</i>	<i>0.01</i>	<i>132.936</i>	<i>87.9</i>
<i>10</i>	<i>200</i>	<i>0.1</i>	<i>0.01</i>	<i>132.146</i>	<i>87.4</i>
<i>10.1</i>	<i>200</i>	<i>0.6</i>	<i>0.001</i>	<i>189.382</i>	<i>365.4</i>
<i>10.2</i>	<i>200</i>	<i>0.9</i>	<i>0.001</i>	<i>194.311</i>	<i>331.8</i>
<i>10.3</i>	<i>200</i>	<i>0.9</i>	<i>0.00</i>	<i>181.483</i>	<i>126.7</i>

❖ *Στις παρακάτω γραφικές, στα αριστερά φαίνεται η απόδοση ανά γενιά, ενώ στα αριστερά της φαίνεται το σφάλμα προσέγγισης ανά γενιά για κάθε ένα απο τα ξεχωριστά τρεξίματα.*

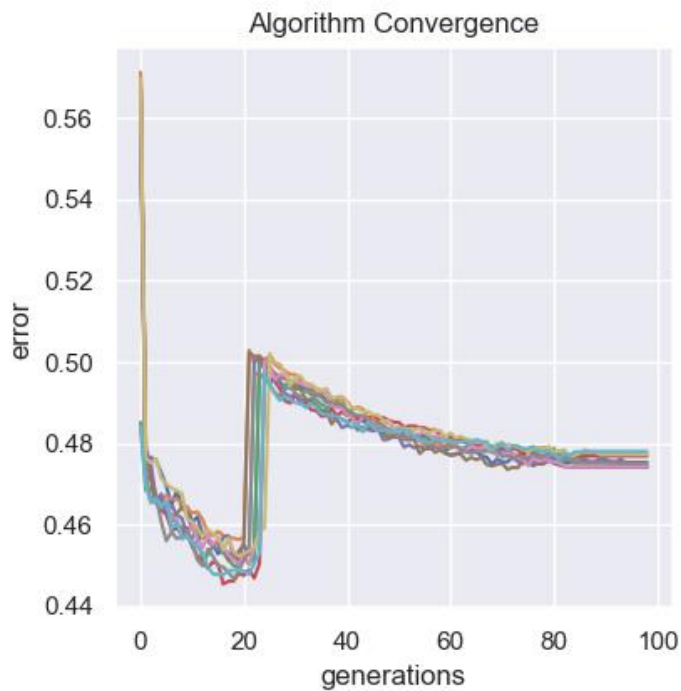
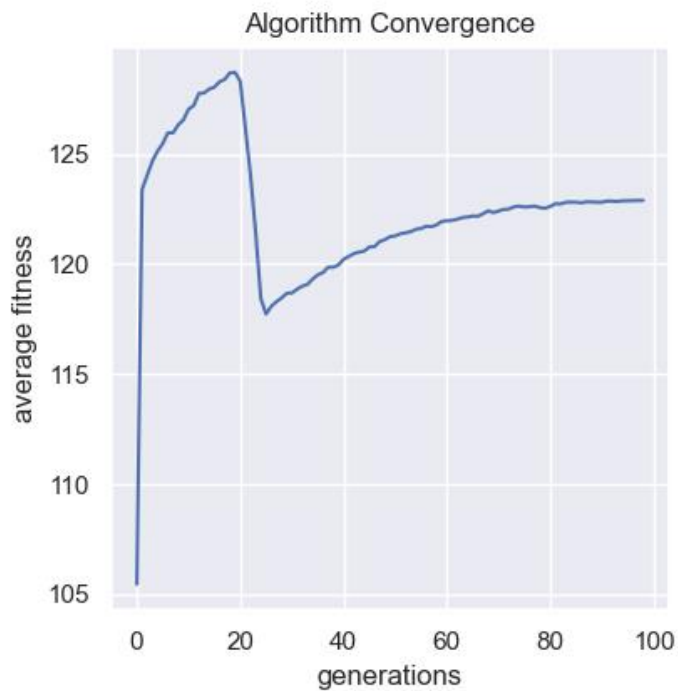


### B3. Αξιολόγηση και Επίδραση Παραμέτρων

*pop\_size=20, pc=0.6, pm=0*

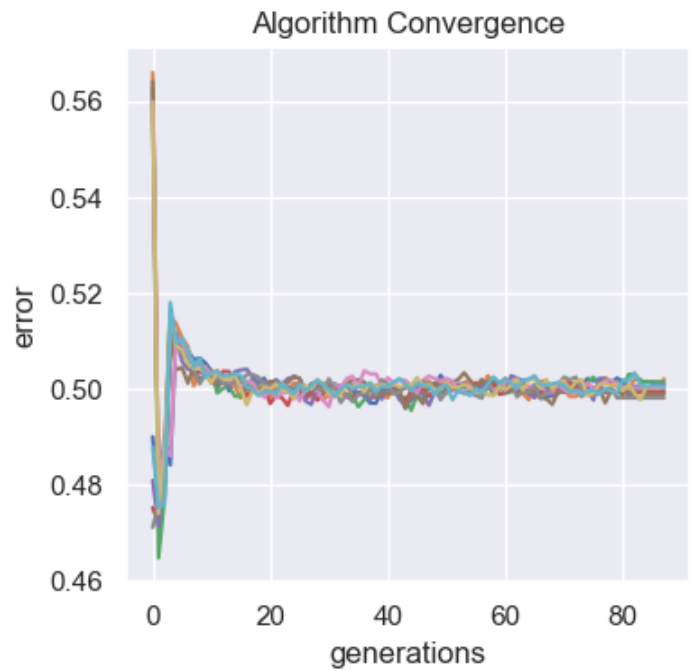
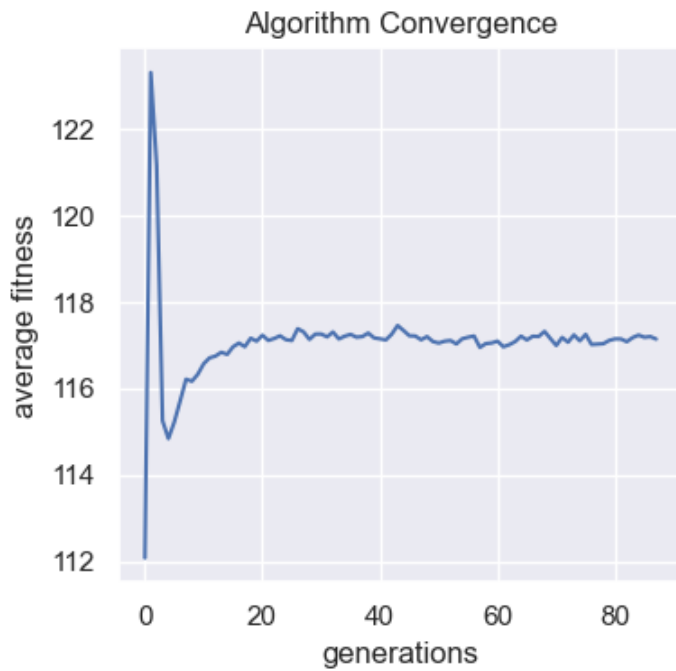


*pop\_size=20, pc=0.6, pm=0.01*

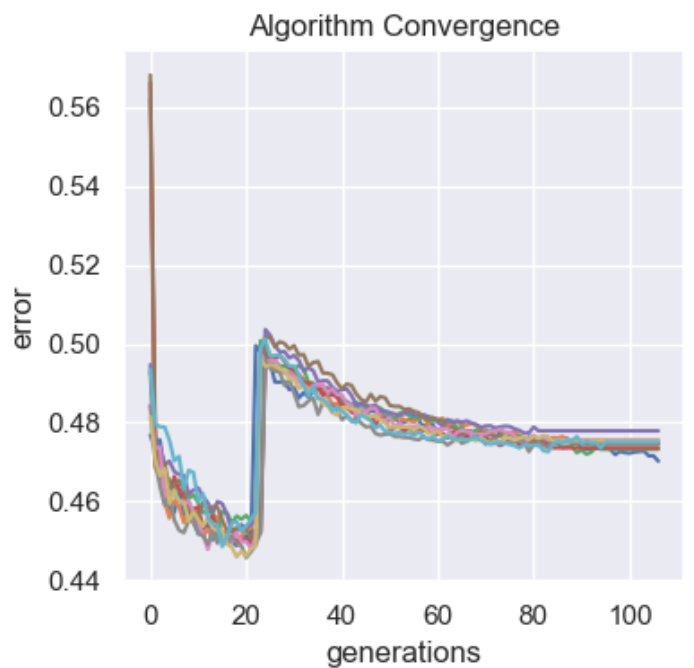
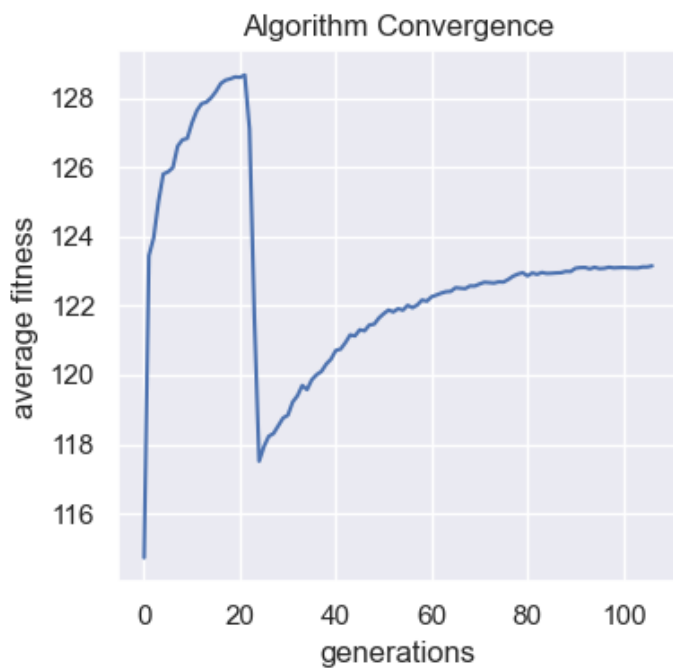


### B3. Αξιολόγηση και Επίδραση Παραμέτρων

*pop\_size=20, pc=0.6, pm=0.1*

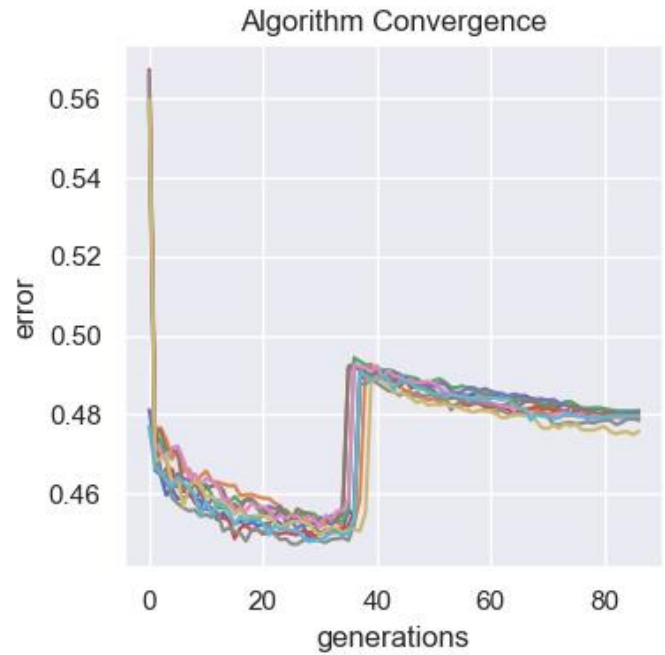
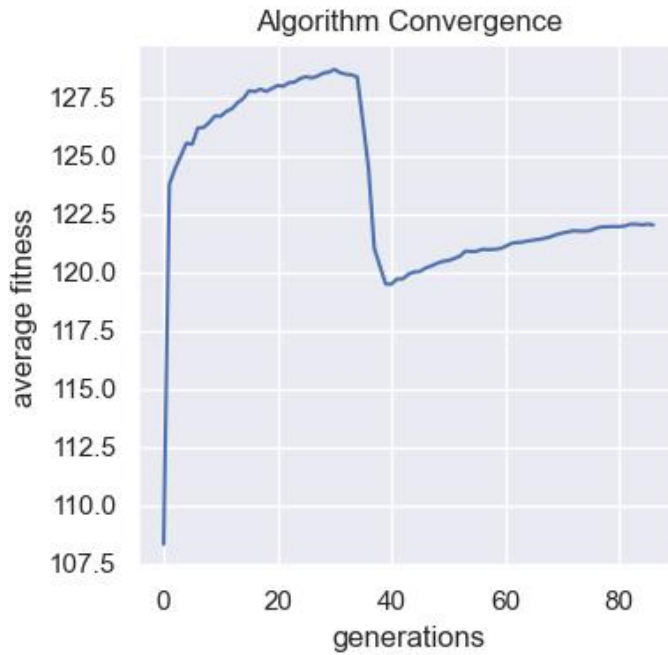


*pop\_size=20, pc=0.9, pm=0.01*

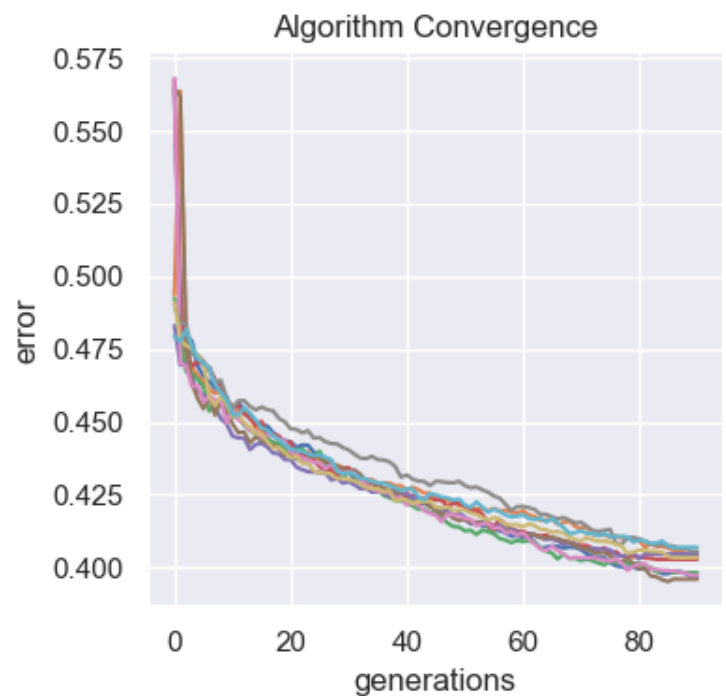
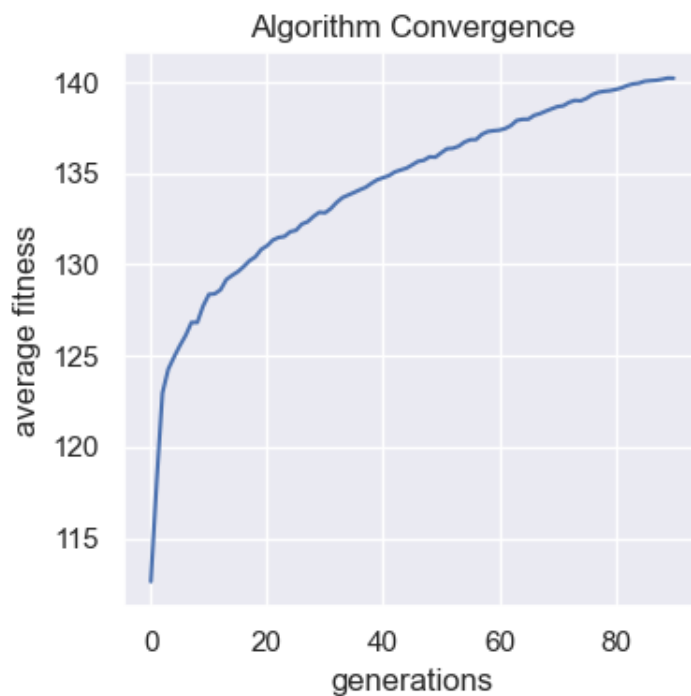


### B3. Αξιολόγηση και Επίδραση Παραμέτρων

*pop\_size=20, pc=0.1, pm=0.01*

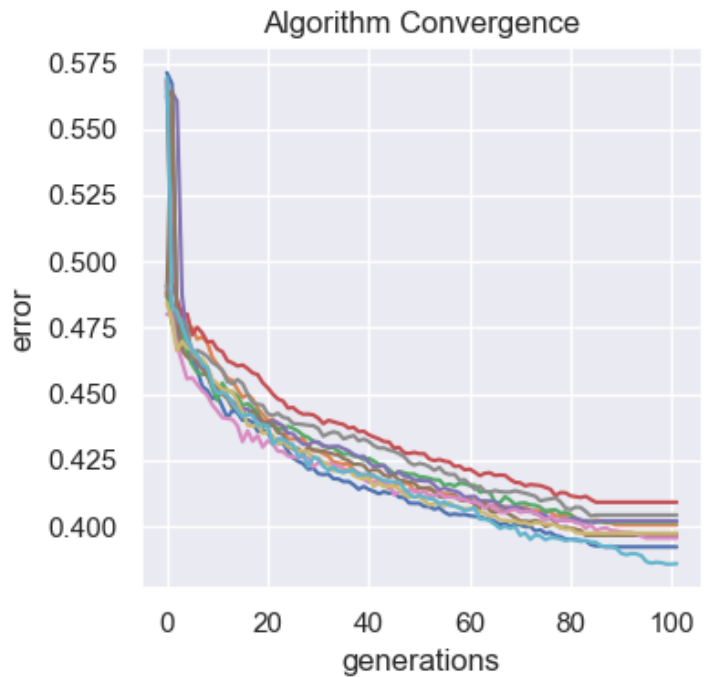
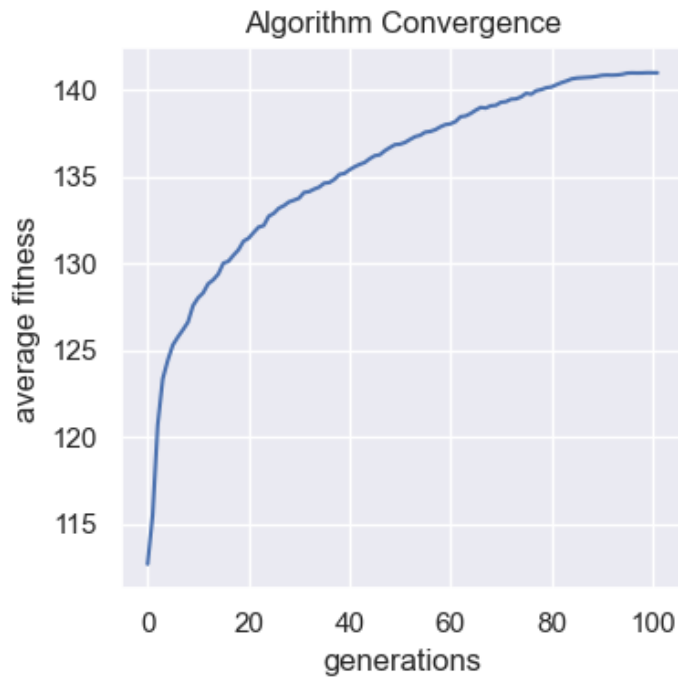


*pop\_size=20, pc=0.6, pm=0.001*

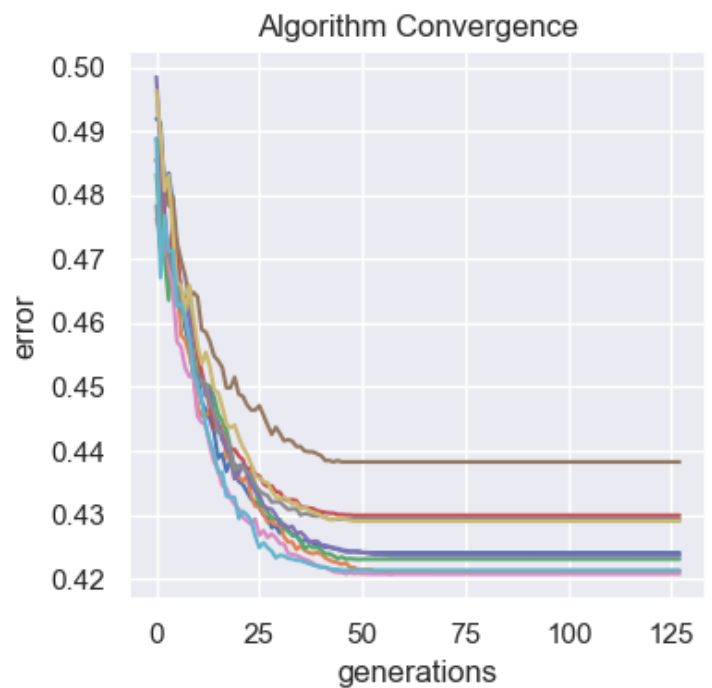
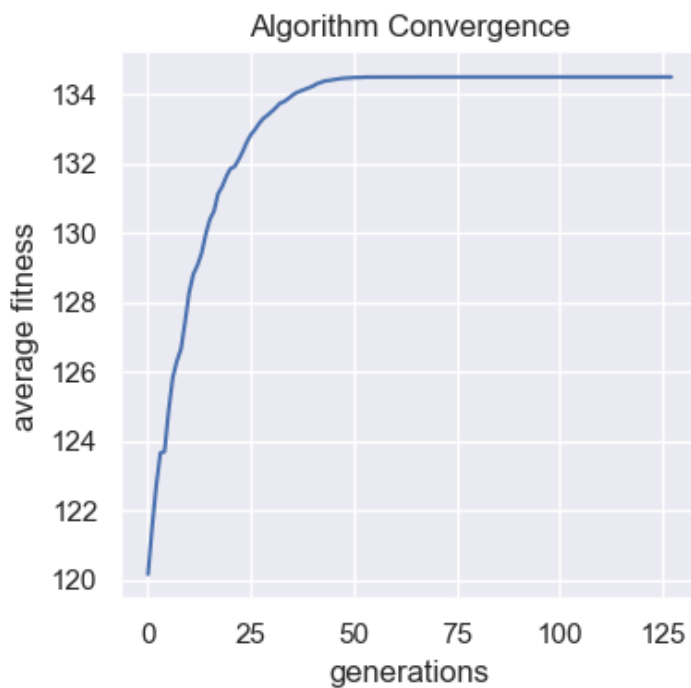


### B3. Αξιολόγηση και Επίδραση Παραμέτρων

*pop\_size=20, pc=0.9, pc=0.001*

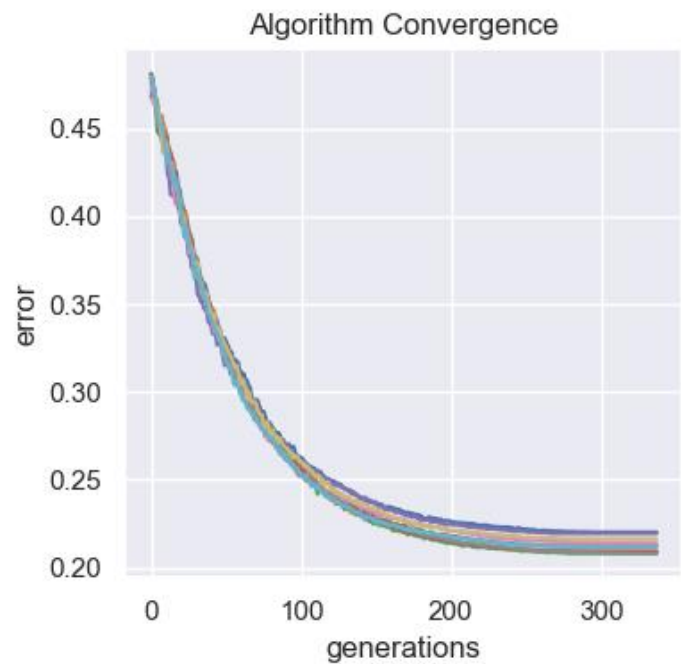
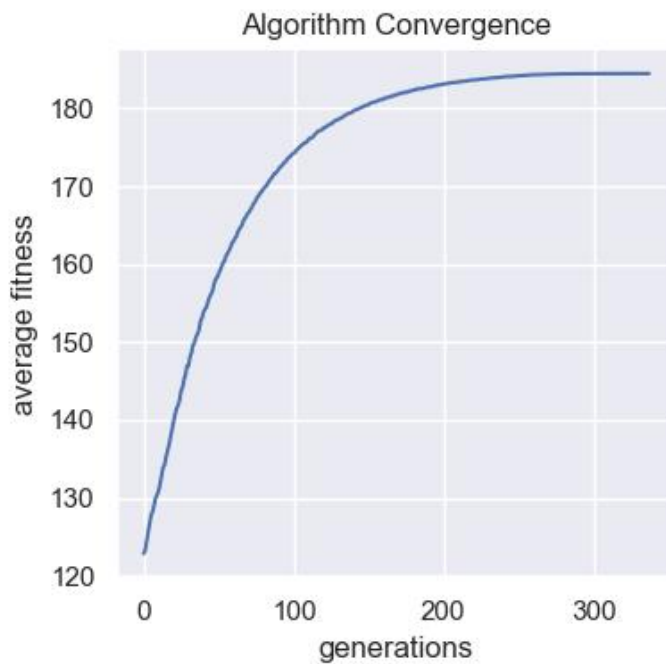


*pop\_size=20, pc=0.9, pm=0.00*

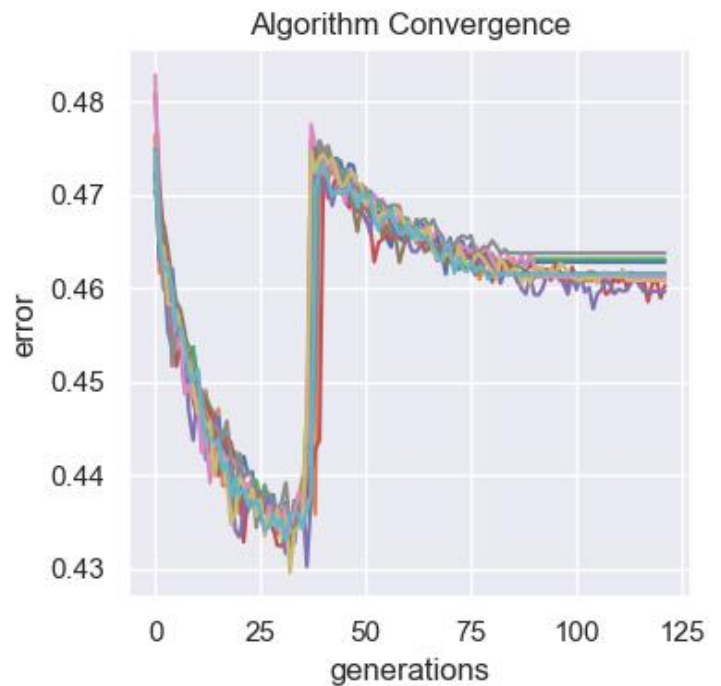
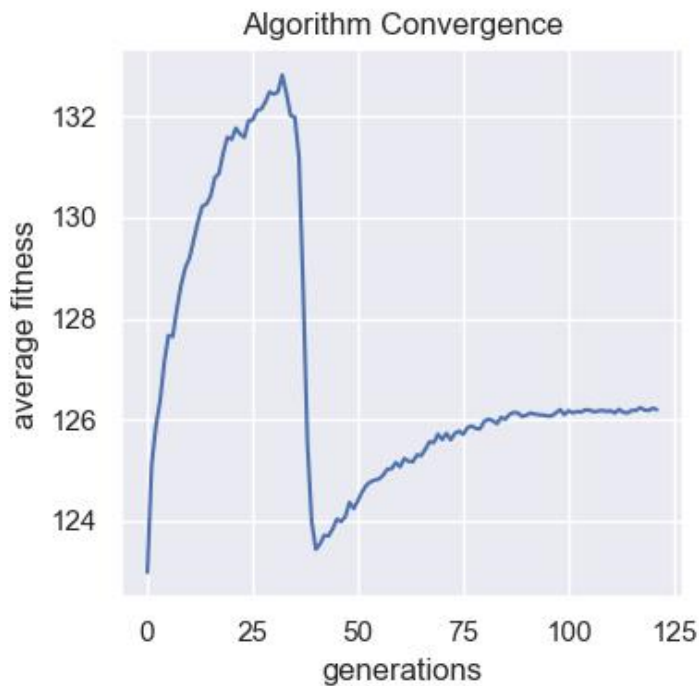


### B3. Αξιολόγηση και Επίδραση Παραμέτρων

*pop\_size=200, pc=0.6, pm=0*

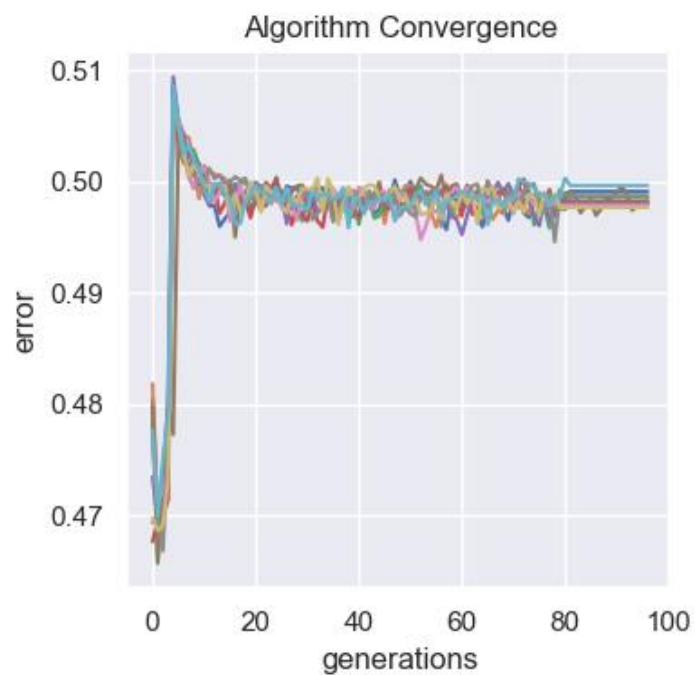
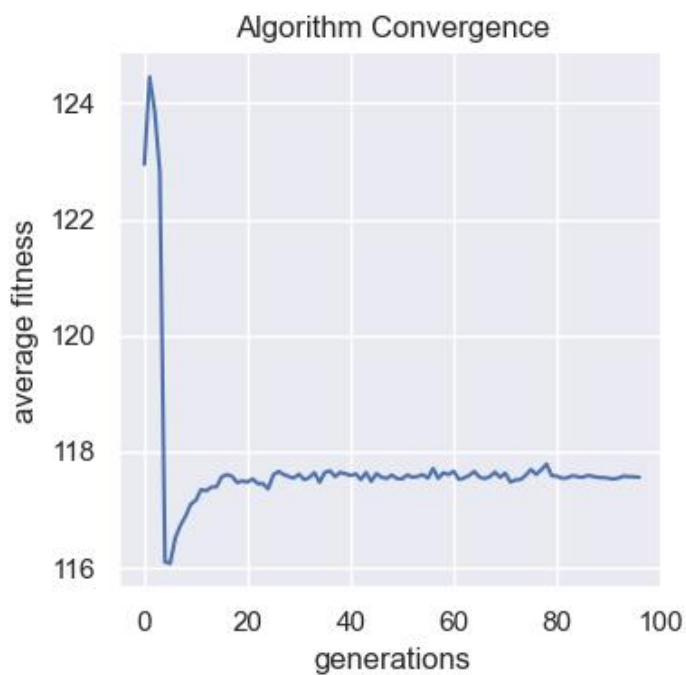


*pop\_size=200, pc=0.6, pm=0.01*

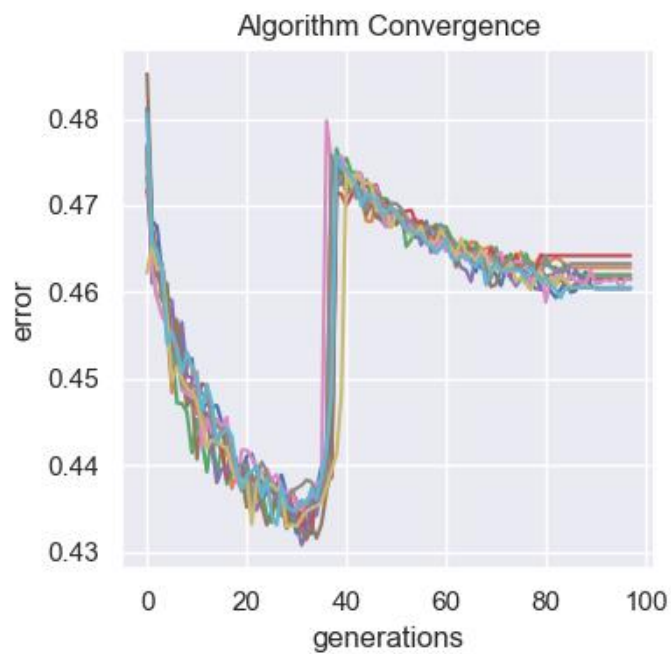
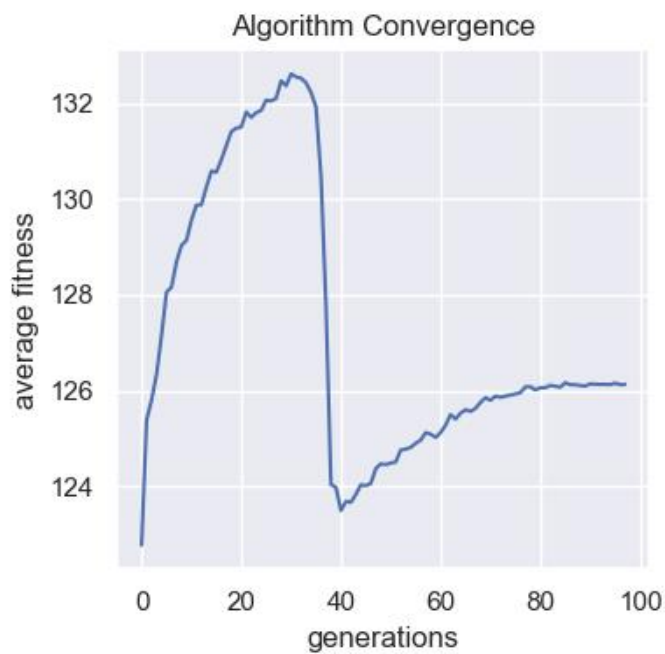


### Β3. Αξιολόγηση και Επίδραση Παραμέτρων

*pop\_size=200, pc=0.6, pm=0.1*

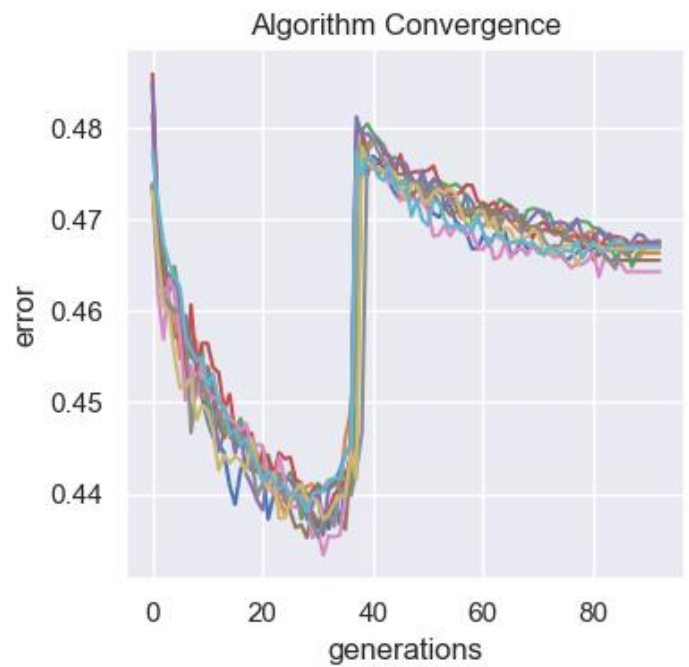
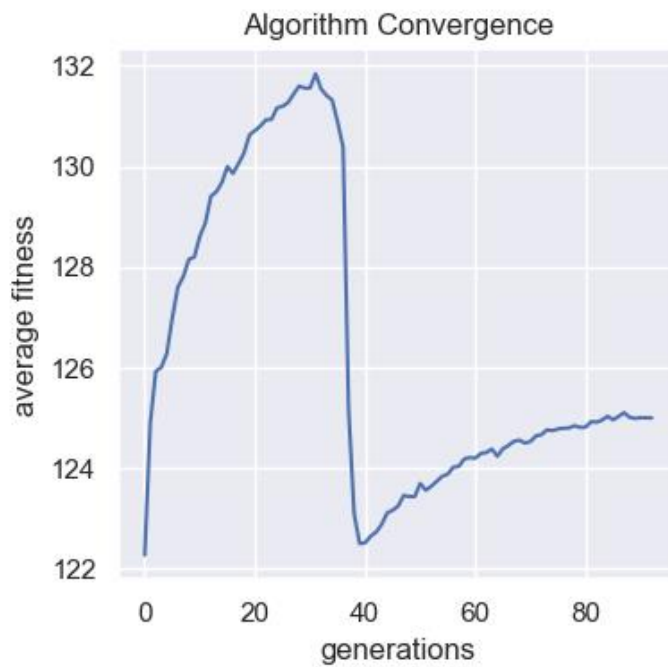


*pop\_size=200, pc=0.9, pm=0.01*

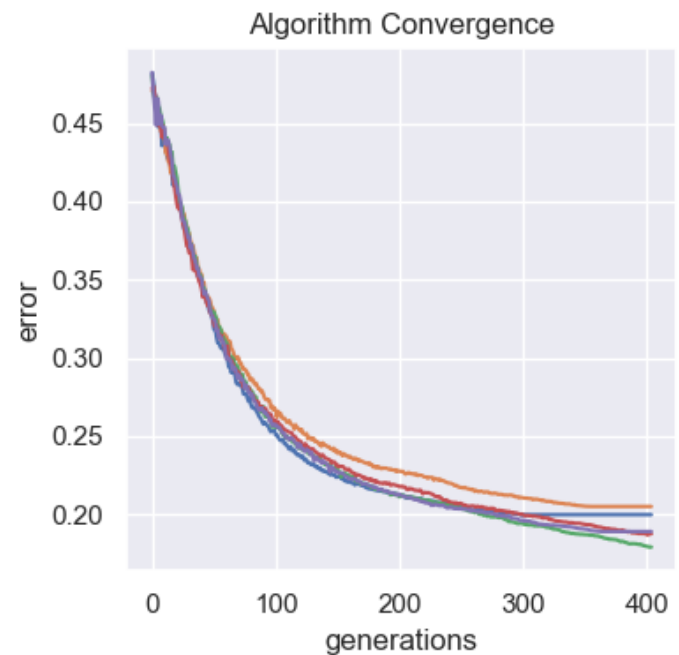
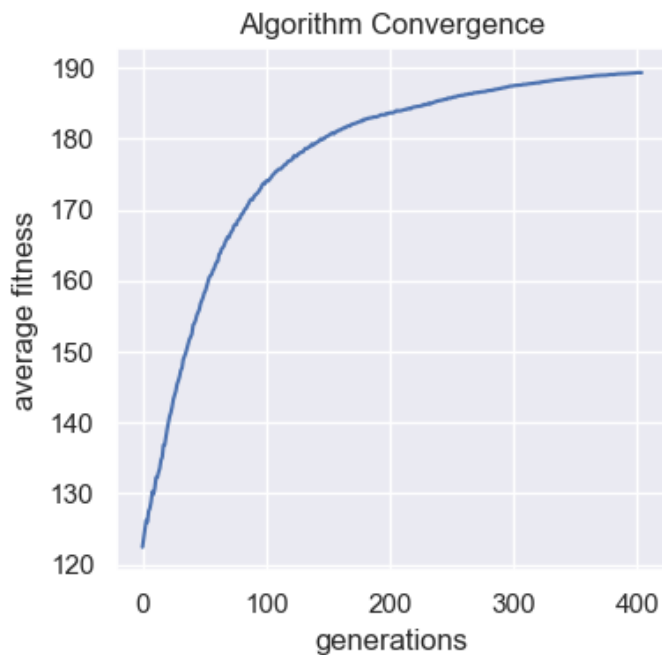


### B3. Αξιολόγηση και Επίδραση Παραμέτρων

*pop\_size=200, pc=0.1, pm=0.01*

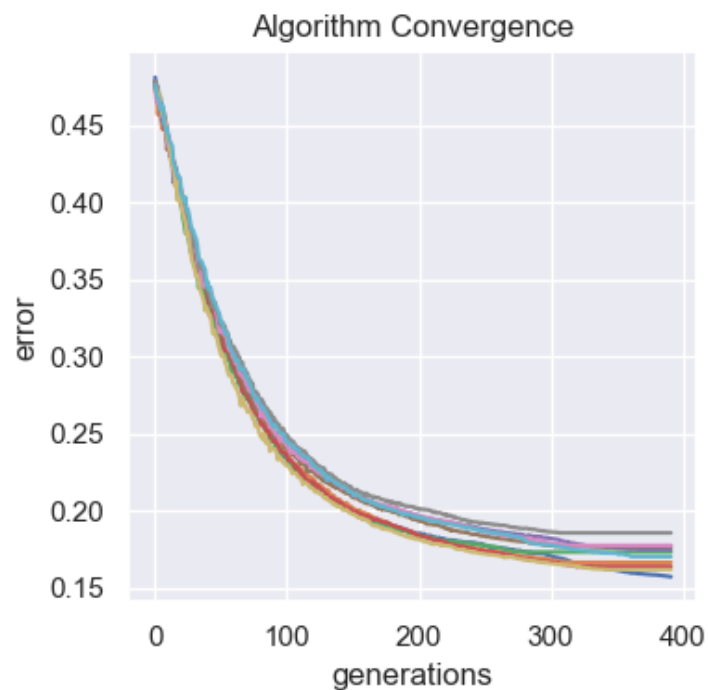
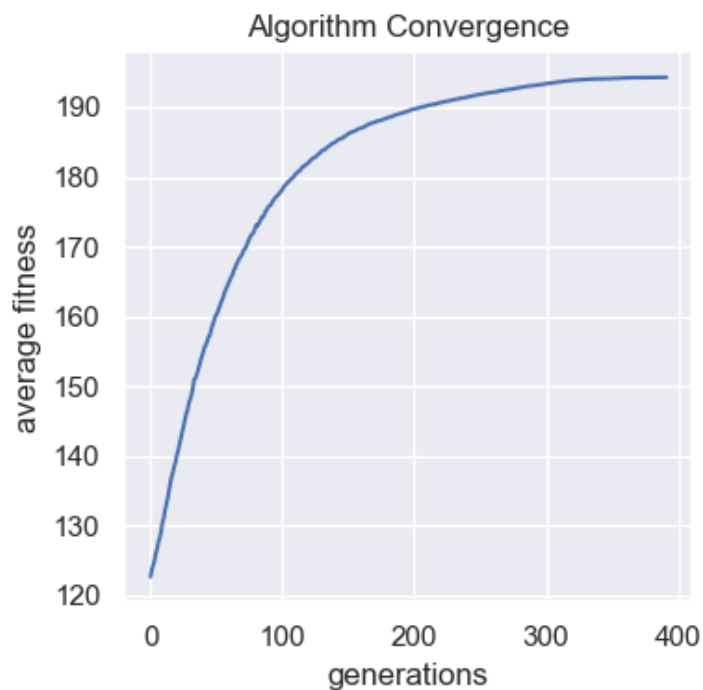


*pop\_size=200, pc=0.6, pm=0.001*

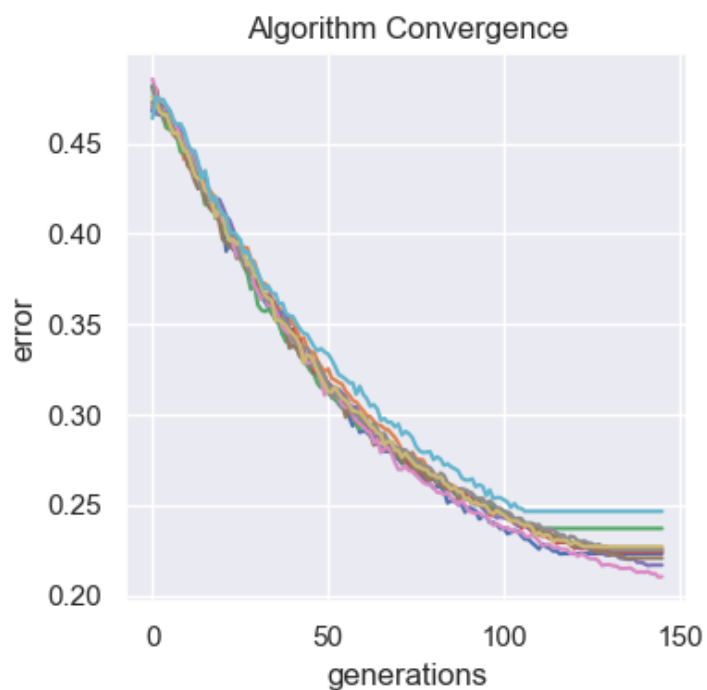
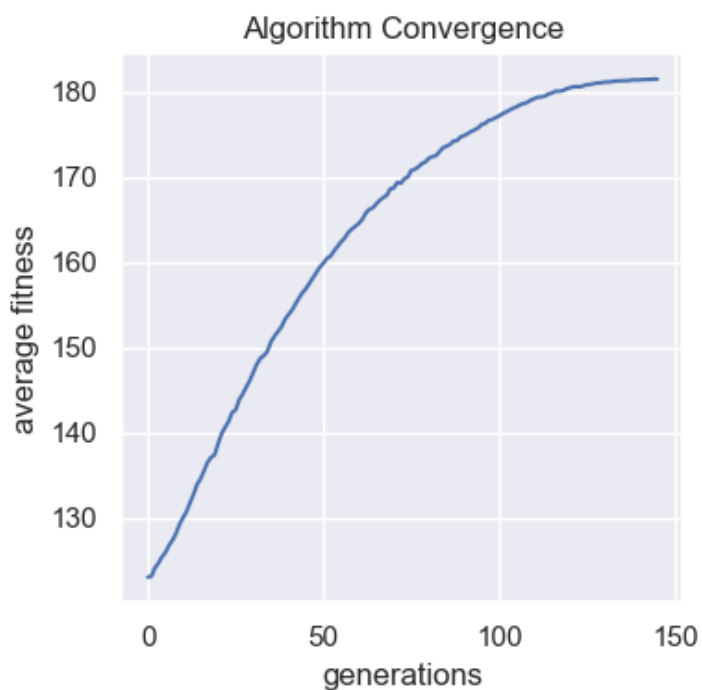


### Β3. Αξιολόγηση και Επίδραση Παραμέτρων

*pop\_size=200, pc=0.9, pm=0.001*



*pop\_size=200, pc=0.9, pm=0.0*





### **B3. Αξιολόγηση και Επίδραση Παραμέτρων**

#### ***Συμπεράσματα και Παρατηρήσεις των πειραμάτων***

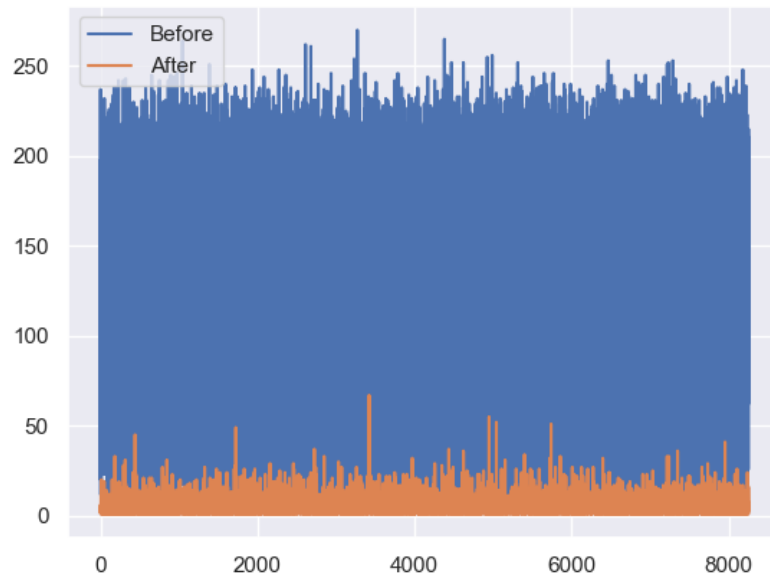
(i) Το αυξημένο μέγεθος του πληθυσμού παίζει σημαντικό και καθοριστικό ρόλο στην εύρεση της λύσης εις βάρος του χρόνου/αριθμού γενεών σύγκλισης. Αυτό είναι λογικό καθώς προσεγγίζει περισσότερο τη παραδοχή της θεωρίας των σχημάτων. Δηλαδή, προκειμένου ο αλγόριθμος να βρεί και να συγκλίνει στη λύση υποθέτει πως η λύση βρίσκεται ήδη μέσα στον αρχικό πληθυσμό. Άρα, αυξάνοντας τον, αυξάνεται και η πιθανότητα να δημιουργηθεί αυτό το άτομο (λύση).

(ii) Όσον αφορά τη παράμετρο της διασταύρωσης, απο τα πειράματα μπορούμε να αποφανθούμε ότι μεγαλύτερη πιθανότητα διασταύρωσης ισοδυναμεί με καλύτερη προσέγγιση της λύσης αλλά και πιο αργή σύγκλιση. Ούτε αυτό θα έπρεπε να μας παραξενεύει καθώς, ειδικά στη περίπτωση μας, για  $N=50$  η διασταύρωση εισάγει μεγάλη πολυμορφία στα άτομα και πιο συγκεκριμένα, λόγω της ρουλέτας με βάση τη κατάταξη, τυχόν καλά γονίδια σε άτομα κακής ποιότητας έχουν μεγάλες πιθανότητες να μετατραπούν ή να συμβάλλουν σε μια καλύτερη λύση.

(iii) Απο το πίνακα των μετρήσεων και τις γραφικές, εύκολα παρατηρείται πώς για πιθανότητα μετάλλαξης  $> 0.001$  η προσέγγιση της λύσης χειροτερεύει και ο αλγόριθμος τερματίζει νωρίς καθώς αποκλίνει ανά γενιά. Όσο μεγαλύτερο είναι το  $pm$ , τόσο χειρότερα αποτελέσματα παίρνουμε. Αυτό συμβαίνει διότι σε κάθε άτομο μεταλλάσσονται κατά μέσο όρο 852 γονίδια για  $pm=0.1$  και 58.2 για  $pm=0.01$ . Αυτό έχει ως αποτέλεσμα τη μεγάλη αύξηση των «1», δηλαδή των λέξεων, το οποίο οδηγεί σε μεγαλύτερη ποινή. Για τις παραπάνω περιπτώσεις των  $pm$ , εύκολα απο τις αντίστοιχες γραφικές τους φαίνεται το σημείο όπου τα άτομα του πληθυσμού ξεπερνούν τις 2500 λέξεις (απότομη αύξηση σφάλματος).

## B4. Επιλογή χαρακτηριστικών TND

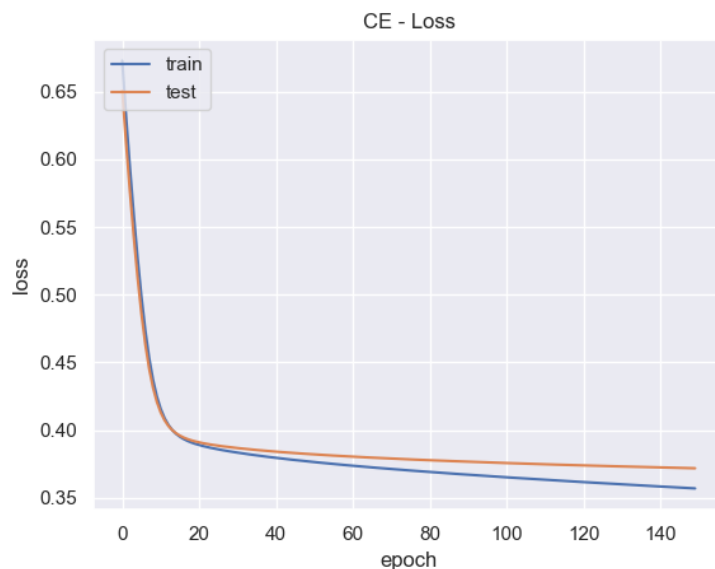
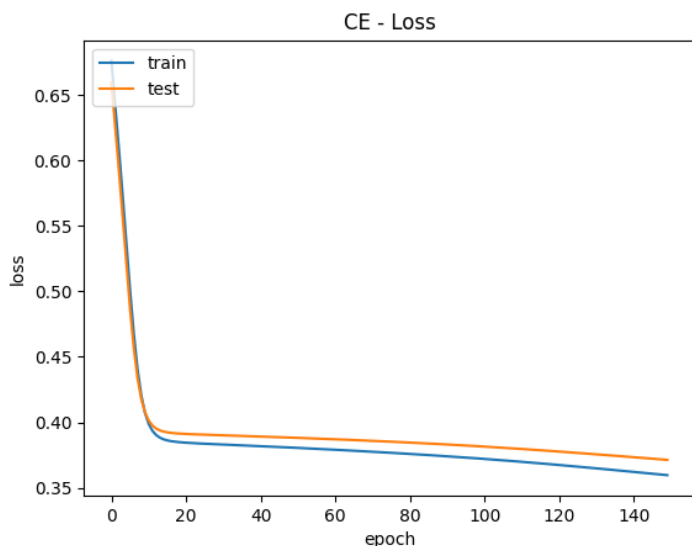
- ❖ Η καλύτερη προσέγγιση του γενετικού ήταν 1002 λέξεις με tfidf~85% του βέλτιστου. Στο παρακάτω γράφημα φαίνονται τα πλήθη λέξεων των κειμένων του dataset πριν και μετά τη μείωση των λέξεων



- ❖ Έχοντας αφαιρέσει τις υπόλοιπες λέξεις από τα κείμενα, τροφοδοτώντας το βέλτιστο νευρωνικό μοντέλο από τη προηγούμενη εργασία με το νέο BoW παίρνουμε τη παρακάτω δεξιά γραφική σύγκριση. Η αριστερά είναι την εργασία Α'.

## B4. Επιλογή χαρακτηριστικών TND

*Σύγκριση σύγκλισης δικτύων ίδιας αρχιτεκτονικής με διαφορετικό # εισόδων*



### *Συμπεράσματα και Παρατηρήσεις*

Όπως φαίνεται και απο τις παραπάνω γραφικές, η απόδοση του δικτύου δεν αλλάζει σχεδόν καθόλου, πράγμα το οποίο μας οδηγεί στο συμπέρασμα πώς η τεχνική του BoW είναι εξαιρετικά αδύναμη για προβλήματα NLP. Τέλος, όσον αφορά τη γενικευτική ικανότητα του δικτύου, απο τις εκτυπώσεις μετρικών των προβλέψεων σε δεδομένα που δεν έχει ξανα δει, φάνηκε να μη μπορεί να προβλέψει τίποτα σωστά όταν το αρχικό δίκτυο με τις 8520 εισόδους πετυχαίνει κάποιο ποσοστό. Αυτό είναι λογικό διότι το νεό δίκτυο δεν έχει αρκετή πολυπλοκότητα ώστε να μπορεί να «ακολουθήσει» τις δύσκολες αυτές επιφάνειες.