



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

**ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ**

**Αναφορά Εργαστηριακής Άσκησης
Μέρος Β΄
Υπολογιστική Νοημοσύνη**

Κωνσταντίνος Τσάκωνας
Α.Μ.: 1059666

Ακαδημαϊκό έτος 2020-21
Εαρινό Εξάμηνο

Περιεχόμενα

Repository Κώδικα

Για την ανάπτυξη της άσκησης χρησιμοποιήθηκε η γλώσσα Python με τις βιβλιοθήκες Tensorflow, numpy, pandas, dear και matplotlib. Παρακάτω υπάρχει το repository του κώδικα στο github

[github link](#)

Σχεδιασμός ΓΑ

1. Τα άτομα του αρχικού πληθυσμού θα αναπαραστηθούν ως δυαδικές συμβολοσειρές. Ο λόγος που θα ακολουθηθεί αυτή η κωδικοποίηση προέρχεται από το σκεπτικό ότι, αυτό που θέλουμε να κάνουμε είναι να μειώσουμε το είσοδος από τα 784 pixels, δηλαδή να μηδενισούμε πολλές από αυτές. Δημιουργώντας λοιπόν άτομα του πληθυσμού ως πίνακες 784 pixels που περιέχουν δυαδικά στοιχεία, μπορούμε με ένα πολλαπλασιασμό element-wise να κρατήσουμε τις εισόδους που θέλουμε.
2. Ο αρχικός πληθυσμός θα είναι N τυχαίοι πίνακες 28x28 και οι τιμές που θα περιέχουν θα είναι 0 και 1.
3. Η συνάρτηση καταλληλότητας που επιλέχθηκε είναι οι εξής:
 - κάθε φορά που θα γίνεται έλεγχος στο νευρωνικό σύμφωνα με τις εισόδους που προκύπτουν από κάθε άτομο του πληθυσμού θα κάνουμε ταξινόμηση των πρώτων 10.000 εικόνων και θα συγκρίνουμε την ταξινόμηση αυτή με βάση τα labels για να δούμε πόσο ακριβής είναι. Το αποτέλεσμα που θα προκύπτει από αυτό θα είναι μια τιμή μεταξύ του διαστήματος $[0, 1]$ και στόχος του γενετικού αλγορίθμου θα είναι να πλησιάσει όσο το δυνατό πιο κοντά στο 1.
 - Επίσης θα επιβάλεται μία ποινή σε άτομα του πληθυσμού που έχουν μεγάλο αριθμό εισόδων, συγκεκριμένα άτομα που έχουν περισσότερες από 392 εισόδους θα αφαιρείται μία τιμή η οποία θα είναι ανάλογη με το ποσοστό που δεν ταξινομήθηκε σωστά επί το πλήθος των παραπάνω εισόδων που έχει σε σχέση με αυτές που έχουμε θέσει ως επιθυμητό άνω όριο.

Ο λόγος που επιλέχθηκε η μεγιστοποίηση του ποσοστού ταξινόμησης έχει να κάνει με το γεγονός ότι κατά την δημιουργία και την εκπαίδευση του μοντέλου στην προηγούμενη εργαστηριακή άσκηση σε πάρα πολύ

μικρές τιμές του loss η ταξινόμηση δεν ήταν πάντα σωστή και δεν συμβάδιζε με το accuracy του μοντέλου. Ακόμα ο λόγος που εφαρμόζουμε ποινή σε άτομα που έχουν περισσότερες από 392 εισόδους είναι γιατί θα έχουμε μία μείωση 50% αλλά για λιγότερες εισόδους από αυτές θα χάνουμε χαρακτηριστικά που χρειαζόμαστε αφού το σχήμα των ψηφιών στις εικόνες συνεχώς μεταβάλλεται, με αποτέλεσμα να μην έχουμε καλή ταξινόμηση.