3η εργασία στην Τεχνητή Νοημοσύνη

Νικόλαος Αργυρίου 4367

Github link:

https://github.com/nikos258/uni-4-AI.git

Google Colab link: https://colab.research.google.com/drive/1tOmRppxkcYe6GUXZDy69-y06v44Wftke?usp=sharing

Άσκηση 1

| αριθμός επιπέδων | νευρώνες ανά επίπεδο | ρυθμός μάθησης | εποχές | συνάρτηση ενεργοποίησης | ακρίβεια (training) | ακρίβεια (validation) |
| --- | --- | --- | --- | --- | --- | --- |
| 4 | 784-256-256-10 | 0.001 | 10 | tanh | 0.9403 | 0.9370 |
| 4 | 784-256-256-10 | 0.01 | 10 | tanh | 0.9022 | 0.8952 |
| 4 | 784-256-256-10 | 0.0001 | 10 | tanh | 0.8551 | 0.8563 |
| 4 | 784-256-256-10 | 0.001 | 15 | tanh | 0.9481 | 0.9414 |
| 4 | 784-512-256-10 | 0.001 | 10 | tanh | 0.9495 | 0.9430 |
| 4 | 784-512-256-10 | 0.001 | 15 | tanh | 0.9605 | 0.9486 |
| 4 | 784-512-256-10 | 0.005 | 15 | tanh | 0.9479 | 0.9438 |
| 5 | 784-512-256-  128-10 | 0.001 | 15 | tanh | 0.9618 | 0.9503 |
| 5 | 784-512-256-  128-10 | 0.001 | 15 | sigmoid | 0.8450 | 0.8479 |
| 5 | 784-512-256-  128-10 | 0.001 | 15 | relu | 0.9989 | 0.9621 |
| 5 | 784-512-256-  128-10 | 0.0005 | 15 | relu | 0.9993 | 0.9520 |
| 5 | 784-512-256-  128-10 | 0.003 | 15 | relu | 0.9993 | 0.9725 |

Αρχικά δοκίμασα να αλλάξω το ρυθμό μάθησης, ώστε το μοντέλο να μαθαίνει πιο γρήγορα ή πιο αργά, αλλά τα αποτελέσματα ήταν χειρότερα από πριν. Όμως αυξάνοντας τον αριθμό των εποχών και στη συνέχεια τον αριθμό των νευρώνων στο πρώτο κρυφό επίπεδο παρουσιάστηκε βελτίωση στην ακρίβεια του μοντέλου, εφόσον αυτό είχε περισσότερο χρόνο και περισσότερο “υλικό” διαθέσιμα για εκπαίδευση. Ακόμα καλύτερα ήταν τα αποτελέσματα με την προσθήκη ενός τρίτου κρυφού επιπέδου. Τελικά παρατηρούμε ότι η συνάρτηση relu είναι πιο αποδοτική σε σχέση με τη σιγμοειδή και την υπερβολική εφαπτομένη. Επίσης ο ρυθμός μάθησης δεν πρέπει να ξεφεύγει από την άνω γειτονιά του 0.001. Οι καλύτερες παράμετροι που επιτύχαμε παρατίθενται στην τελευταία σειρά του πίνακα.

Άσκηση 2

Με μια γρήγορη αναζήτηση βλέπουμε ότι ένας τύπος νευρωνικών δικτύων κατάλληλος για την ταξινόμηση εικόνων σε κατηγορίες είναι τα συνελικτικά νευρωνικά δίκτυα (cnn). Αυτά βρίσκουν συγκεκριμένα χαρακτηριστικά στις εικόνες και τα χρησιμοποιούν στην αναγνώριση άγνωστων εικόνων (πηγή: <https://course.elementsofai.com/el/5/3>). Το εφαρμοζόμενο μοντέλο καταγράφει ακρίβεια 0.9905, μεγαλύτερη τόσο από αυτή του αρχικού όσο και του βελτιωμένου νευρωνικού δικτύου, αν και χρειάζεται περισσότερο χρόνο για την εκπαίδευσή του.

|  | Ακρίβεια (training) | Ακρίβεια (validation) |
| --- | --- | --- |
| Αρχικό | 0.9403 | 0.9370 |
| Βελτιωμένο (Άσκηση 1) | 0.9993 | 0.9725 |
| CNN | 0.9959 | 0.9905 |

Άσκηση 3

1. Τα δεδομένα της MNIST φαίνεται να είναι καλά για την εκπαίδευση ενός δικτύου. Εξάλλου τα ψηφία παρουσιάζουν μια ποικιλομορφία, όπως άλλωστε και στον πραγματικό κόσμο και θέλουμε το μοντέλο μας να έχει προετοιμαστεί κατάλληλα. Επίσης η βάση δεδομένων MNIST είναι αρκετά γνωστή και δημοφιλής για τη δημιουργία απλών νευρωνικών δικτύων και έχει εφοδιαστεί με έναν σχετικά μεγάλο αριθμό από δείγματα εικόνων.
2. Δεν είναι όλα τα εικονοστοιχεία το ίδιο σημαντικά για την αναγνώριση των ψηφίων. Ειδικά τα pixel που περικλείουν το ψηφίο, δηλαδή τα μαύρα pixel που βρίσκονται εκτός του περιγεγραμμένου παραλληλογράμμου που εσωκλείει τα άσπρα, θα μπορούσαν να απορριφθούν και η κατάταξη του ψηφίου σε μια κλάση θα γινόταν το ίδιο εύκολα.
3. Τα βαθιά νευρωνικά δίκτυα ενδείκνυνται για αναγνώριση εικόνων, δημιουργία κειμένου, αναγνώριση φωνής, δημιουργία τεχνητής φωνής, αλγορίθμους συστάσεων περιεχομένου (π.χ. Netflix, Youtube), αυτόνομη οδήγηση, χρήση στην ιατρική, πρόβλεψη τιμών (μετοχές), καθώς και πολλές εφαρμογές.
4. Η βαθιά μάθηση μπορεί πράγματι να χρησιμοποιηθεί και στις τρεις κατηγορίες της μηχανικής μάθησης. Ένα παράδειγμα χρήσης της στη μάθηση με επίβλεψη είναι ο στόχος αυτής της εργασίας, δηλαδή δεδομένου ενός συνόλου χειρόγραφων ψηφίων και της κλάσης όπου ανήκει το καθένα το μοντέλο μας να μπορέσει να εκπαιδευτεί. Όσον αφορά τις άλλες δύο κατηγορίες, τα GAN είναι μία τεχνική της βαθιάς μάθησης που ανήκει σε αμφότερες. Εδώ ένα μοντέλο αναλαμβάνει τον ρόλο του δημιουργού περιεχομένου και ένα άλλο μοντέλο προσπαθεί να καταλάβει αν αυτό το περιεχόμενο είναι αληθινό ή τεχνητό.