

Ο στόχος της άσκησης είναι να βελτιώσουμε όσο μπορούμε τις επιδόσεις μας στο validation set όπως αυτό ορίζεται στο notebook “**11.1 CIFAR-100, Transfer Learning, Βελτιστοποίηση.ipynb**”.

Το μοντέλο from scratch όπως παρουσιάζεται στο notebook πετυχαίνει accuracy γύρω στο 13% και το VGG16 με λίγη εκπαίδευση λίγο κάτω από 50%.

Πολλά από τα μοντέλα του [leaderboard στο CIFAR-100](#) υπάρχουν έτοιμα είτε απευθείας μέσα στο TensorFlow είτε από κάποιο εξωτερικό αποθετήριο. Τα καλύτερα μοντέλα πετυχαίνουν πάνω από 90% accuracy.

Ξεκινώντας από τα μοντέλα του notebook μπορείτε να ακολουθήσετε δύο πορείες:

α) μοντέλα “from scratch” δηλαδή δικές σας αρχιτεκτονικές που τις εκπαιδεύετε από την αρχή.

β) μεταφορά μάθησης. Στην περίπτωση αυτή, εισάγετε ένα έτοιμο μοντέλο είτε χωρίς τα βάρη του, είτε μαζί με το σύνολο ή μέρος των βαρών του προς εκπαίδευση.

Σε κάθε περίπτωση θα βρείτε όλες τις προτάσεις για βελτίωση στην ενότητα “Βελτίωση της απόδοσης με πειράματα”.

Παρατηρήσεις ως προς τη βελτιστοποίηση:

- Ένα μοντέλο βελτιστοποιείται και ως προς τον εαυτό του, όχι ως μόνο προς τα άλλα μοντέλα. Αν για παράδειγμα φέρετε με μεταφορά μάθησης ένα state-of-the-art μοντέλο, θα εξετάσετε πόσο μπορείτε να βελτιστοποιήσετε το ίδιο και δεν θα το συγκρίνετε με ένα πολύ απλό δίκτυο “from scratch”. Αυτό δεν σημαίνει ότι δεν μπορείτε να παρουσιάσετε συγκρίσεις μεταξύ μοντέλων σε πίνακες και γραφήματα (βλ. και επόμενη παρατήρηση).
- Η βελτιστοποίηση μπορεί να αφορά στην απόδοση ως προς τη μετρική που χρησιμοποιούμε (ορθότητα) στο σύνολο ελέγχου αλλά και σε άλλες ιδιότητες που έχουν να κάνουν με την εκπαίδευση: απαιτούμενη μνήμη, αριθμός παραμέτρων, χρόνος εκπαίδευσης, συμπεριφορά ως προς την υπερεκπαίδευση κ.ο.κ. Μπορείτε να παρουσιάσετε συγκριτικά αποτελέσματα “πριν και μετά” της βελτίωσης αυτών ιδιοτήτων ποσοτήτων ακόμα και για το ίδιο μοντέλο, ασχέτως της απόλυτης απόδοσής του ως προς την ορθότητα σε σχέση με άλλα μοντέλα.
- Προσπαθήστε να μελετήσετε διάφορες από τις δυνατότητες βελτιστοποίησης που αναφέρονται στο notebook όπως για παράδειγμα
  - την επαύξησης δεδομένων (data augmentation)
  - τις μεθοδολογίες περιορισμού της υπερεκπαίδευσης
  - την αλλαγή του πλήθους των επιπέδων που θα εκπαιδευτούν (fine-tuning) κατά τη μεταφορά μάθησης
  - τον ρυθμό μάθησης (learning rate)
  - τον αλγόριθμο βελτιστοποίησης (optimizer)
  - το μέγεθος δέσμης (batch size)
  - το μέγεθος των εικόνων (resize input)