# Εισαγωγή στα Νευρωνικά Δίκτυα

Η ιδέα των τεχνητών νευρωνικών δικτύων (Artificial Neural Networks) άρχισε να αναπτύσσεται την δεκαετία του ‘50 από τον Frank Rosenblatt, με επιρροές από προηγούμενες έρευνες των Warren McCulloch και Walter Pitts, και αναφέρονταν ως perceptrons. Τα νευρωνικά δίκτυα είναι υπολογιστικά συστήματα εμπνευσμένα από τα βιολογικά νευρωνικά δίκτυα από τα οποία αποτελείται ο εγκέφαλος, και αποτελούν μια αρκετά απλουστευμένη εκδοχή αυτού. Αυτά τα συστήματα “μαθαίνουν” να εκτελούν ενέργειες για τις οποίες δεν έχουν άμεσα προγραμματιστεί, βασιζόμενα σε ένα πλήθος παραδειγμάτων που έχει αρχικά δοθεί.

Τα ANNs διαφοροποιούνται ανάλογα με την δομή και την λογική στην οποία υλοποιούνται. Στην παρούσα εργασία αναλύονται και χρησιμοποιούνται δυο διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων, το dense (fully connected) neural network και το convolutional neural network.

# Dense Neural Networks

Τα dense neural network αποτελλούν την βασικότερη εκδοχή των νευρωνικών δικτύων. Η αρχιτεκτονική τους είναι fully connected, δηλαδή κάθε νευρώνας ενός layer συνδέεται με όλους τους νευρώνες των γειτονικων layers.



Η γενική ιδέα της λειτουργίας ενός τέτοιου δικτύο αποτελείται από:

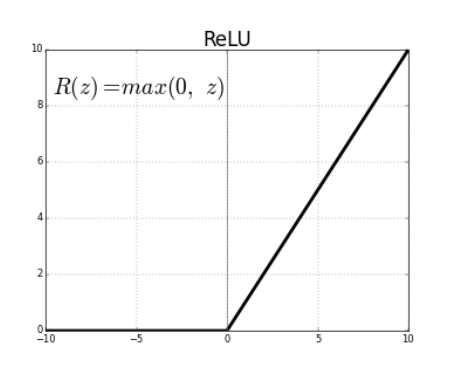
* Feed-forward: Το δίκτυο δέχεται σαν είσοδο το dataset με το οποίο θα εκπαιδευτεί και υπολογίζει μέσω των παραμέτρων του δικτύου. Αυτές αποτελλούνται από τις ακμές (weights) και τους νευρώνες (biases) ανα layer. Μαζί με την συνάρτηση ενεργοποίησης σε κάθε εξόδο του κάθε layer, καταλήγουμε στην έξοδο του δικτύου. Η έξοδος είναι και η εκτίμηση (predict) της αρχικής εισόδου.
* Cost function: Στην έξοδο του δικτύου, κατά την φάση της εκπαίδευσης, εφαρμόζεται μια συνάρτηση κόστους η οποία υπολογίζει το σφάλμα της εκτίμησης του δικτύου για όλα τα δεδομένα με τα οποία εκπαιδεύεται.
* Back propagation: Με την τεχνική της οπισθοδιάδοσης, και βάση του σφάλματος το οποίο βρέθηκε στο προηγούμενο στάδιο, οι παράμετροι ρυθμίζονται και η όλη διαδικασία επαναλαμβάνεται έως το σφάλμα να ελαχιστοποιηθεί. Στην οπισθοδιάδοση (όπως και στην συνάρτηση σφάλματος) μπορούν να υλοποιηθούν αρκετές εναλλακτικές τεχνικές, όπως Gradient Descent, Adam, RMSprop κτλ.

Στην εργασία υλοποιήθηκαν αρκετές παραλλαγές νευρωνικών δικτύων μέχρι να καταλήξουμε στην κατάλληλη αρχιτεκτονική. Το κάθε ένα από αυτά δέχεται σαν είσοδο το κάθε αρχείο ήχου από ένα dataset με πάνω από 23.000 audio clips. Τα δεδομένα αυτά πριν την είσοδο τους στο νευρωνικό δίκτυο πέρασαν από μια διαδικασία preprocessing, η οποία διαφοροποιείται άναλογα με την αρχιτεκτονική του δικτύου.

Οι δομές οι οποίες ερευνήθηκαν ποικίλουν στον αριθμό των νευρώνων ανά layer, στον αριθμό των layer, στην συνάρτηση κόστους αλλά και στον αλγόριθμο του back propagation.

Η τελική μορφή του αποτελλείται από 5 hidden layers (36, 52, 52, 36 και 24 νευρώνες αντίστοιχα). Η συνάρτηση ενεργοποίησης σε κάθε hidden layer είναι η ReLu, ενώ του output layer η softmax, καθώς η τελική έξοδος αποτελείται από 10 εξόδους (μία για κάθε ψηφίο). Για την συνάρτηση κόστους χρησιμοποιήθηκε η cross-entropy, και τέλος για την οπισθοδιάδοση κάναμε χρήση του αλγορίθμου RMSprop.

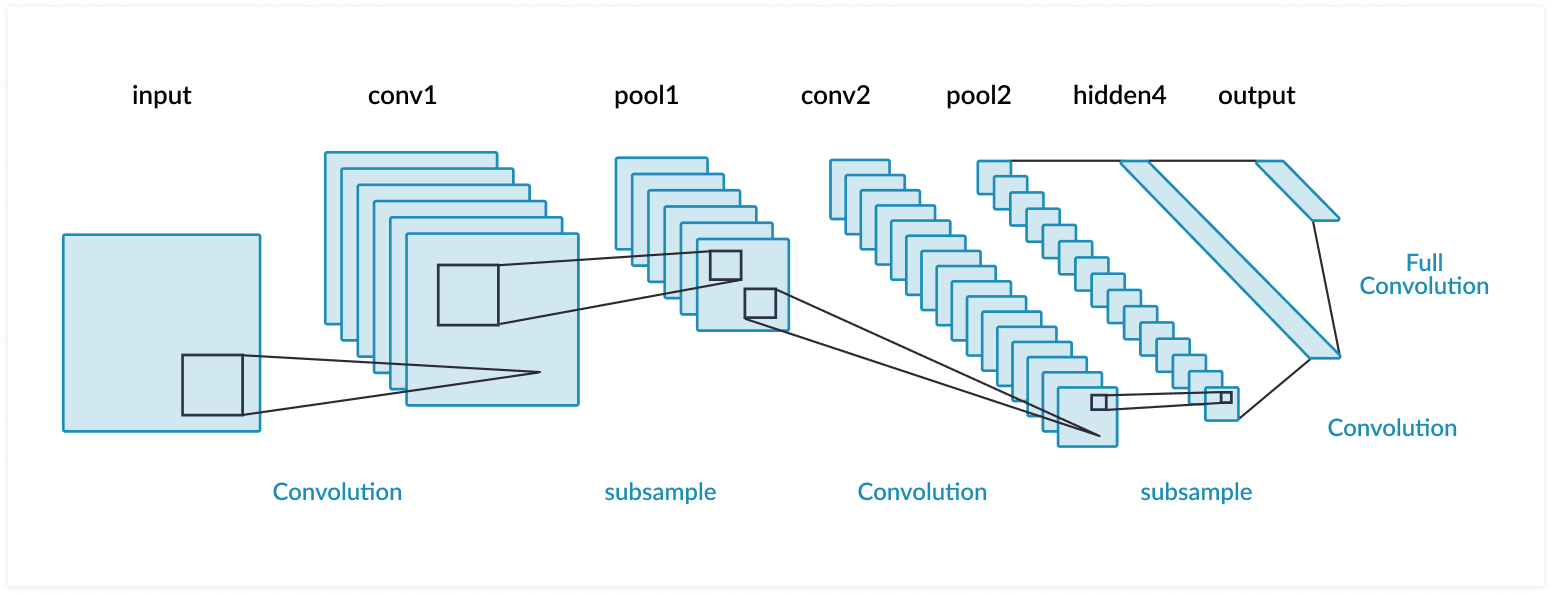
**Cross-entropy cost function: L = - (y \* log(p) + (1 - y) \* log(1 - p))**



Τα αποτελέσματα που είχαμε με είσοδο την διακριτοποίηση της συχνότητας των αρχείων ήχου στα 8000 δείγματα δεν ήταν τόσο καλά. Κάνοντας δοκιμές στην επεξεργασία των δεδομένων μέσω των αλγορίθμων Fourier και MFCC, είχαμε καλύτερα αποτελέσματα. Συγκεκριμένα, κάνοντας preprocessing με τον αλγόριθμο MFCC, πετύχαμε accuracy του train στο 89% και του test στο 82%.

# Convolutional Neural Network

Μετά από έρευνα πάνω στην ανάλυση ήχου καταλήξαμε ότι, χωρίς το κατάλληλο preprocessing, η επιλογή ενός dense νευρωνικού δικτύου για αναγνώριση ψηφίων δεν είναι αρκετά καλή. Έτσι, δοκιμάσαμε το convolutional neural network, μια διαφορετική αρχιτεκτονική δικτύου που χρησιμοποιείται ιδιαίτερα στην ανάλυση και αναγνώριση εικόνας.



Μια βασική διαφορά των convolutional neural network από τα κλασσικά feed forward νευρωνικά είναι η ανοχή τους στα μη προ-επεξεργασμένα δεδομένα. Χωρίς δηλαδή το κατάλληλο preprocessing, τα convolutional καταφέρνουν και ανιχνεύουν τα χαρακτηριστικά των δεδομένων. Αυτό γίνεται τροποποιώντας τα δεδομένα σε μορφή πολύ πιο εύκολη προς επεξεργασία χωρίς να χάνουν σημαντικά features των δεδομένων, κάτι που είναι πολύ σημαντικό για την πρόβλεψη. Μετά την επεξεργασία αυτών των δεδομένων, καταλήγουμε σε ένα dense νευρωνικό που κάνει και την τελική πρόβλεψη των δεδομένων.

Τα αποτελέσματα του convolutional σε σχέση με το dense είναι αρκετά ικανοποιητικά, καθώς φτάνουμε accuracy στο train της τάξης του 97% και 91% στο test.