

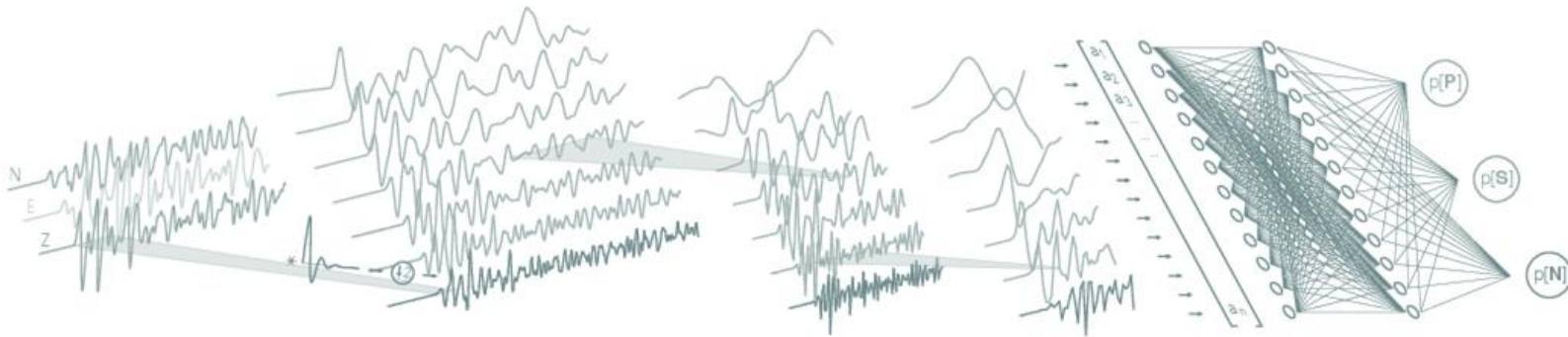


ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

**Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών**

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΓΕΩΛΟΓΙΑΣ ΚΑΙ ΓΕΩΠΕΡΙΒΑΛΛΟΝΤΟΣ
ΤΟΜΕΑΣ ΓΕΩΦΥΣΙΚΗΣ- ΓΕΩΘΕΡΜΙΑΣ

Προχωρημένη Σεισμολογία (ΓΣ-Υ04)

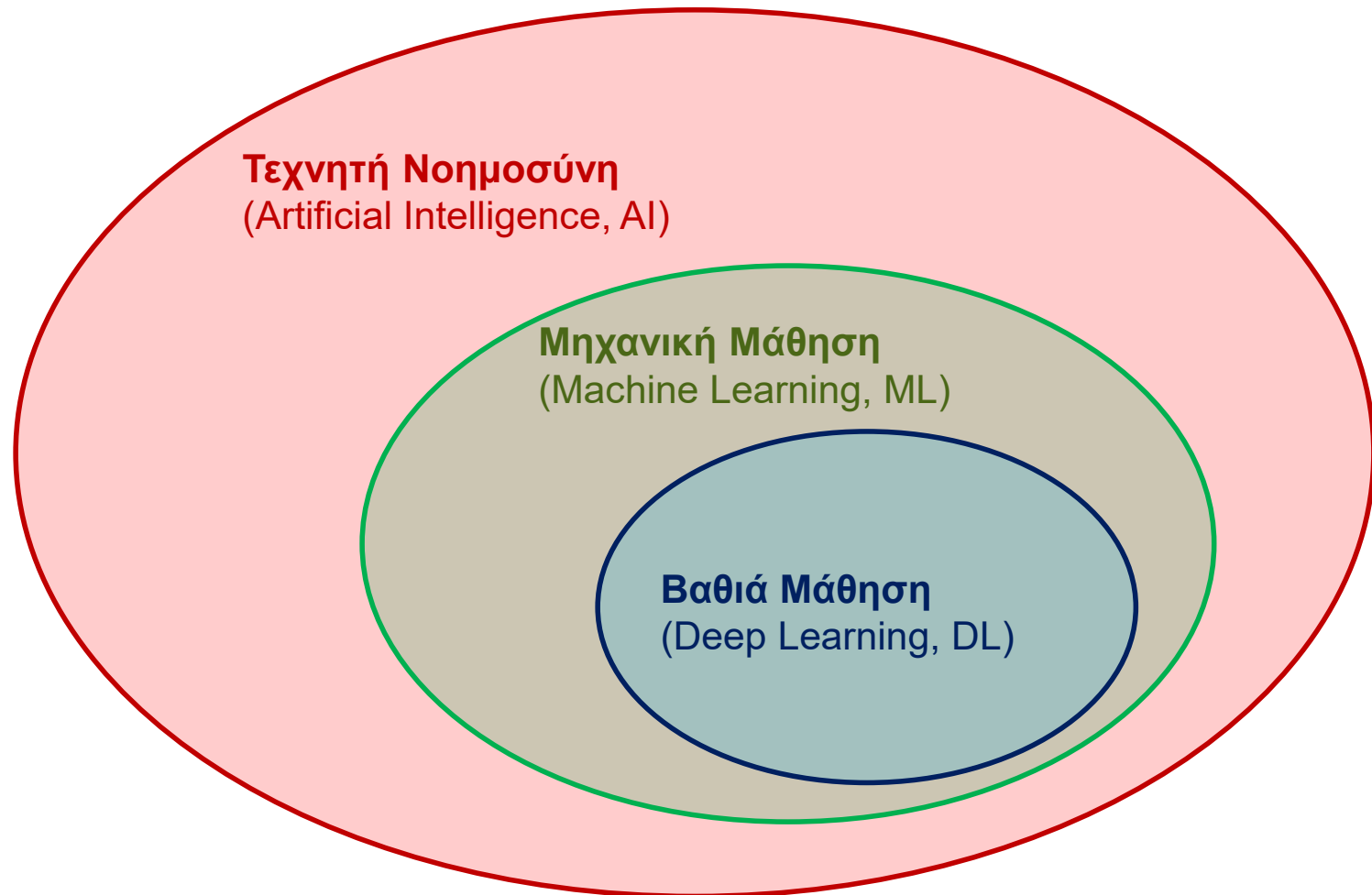


Εφαρμογές Μηχανικής Μάθησης στη Σεισμολογία

Περιεχόμενα του μαθήματος

- Τεχνητή Νοημοσύνη – Μηχανική Μάθηση – Βαθιά Μάθηση
- Τεχνητά νευρωνικά δίκτυα (ΤΝΔ)
- Το μοντέλο Perceptron
- Βασικές έννοιες ΤΝΔ
- Εκπαίδευση μοντέλου ΤΝΔ
- Σεισμολογικές εφαρμογές
 - Αναγνώριση χρόνου άφιξης σεισμικών κυμάτων
 - Συσχετισμός φάσεων με σεισμικά γεγονότα
 - Αναγνώριση πολικότητας πρώτης απόκλισης επιμήκων κυμάτων
 - Αναγνώριση τύπου σεισμικών σημάτων
 - Διαχωρισμός σήματος από θόρυβο

Τεχνητή νοημοσύνη – Μηχανική Μάθηση – Βαθιά Μάθηση



Τι είναι η τεχνητή νοημοσύνη

(Artificial Intelligence, ή AI)

- Η **Τεχνητή Νοημοσύνη** είναι ένας κλάδος της επιστήμης των υπολογιστών που στοχεύει στη δημιουργία μηχανών ικανών για ευφυή συμπεριφορά.
- Στην ουσία, η Τεχνητή Νοημοσύνη έχει ως στόχο να κάνει τους υπολογιστές να εκτελούν εργασίες που κανονικά θα απαιτούσαν ανθρώπινη νοημοσύνη.
- Αυτό περιλαμβάνει δυνατότητες όπως:
 - η εκμάθηση από την εμπειρία
 - η αναγνώριση αντικειμένων
 - η κατανόηση και η ανταπόκριση στη γλώσσα
 - η λήψη αποφάσεων
 - η επίλυση προβλημάτων



Τι είναι η Μηχανική Μάθηση

(Machine Learning, ή ML)

- Η **Μηχανική Μάθηση** είναι ένας κλάδος της τεχνητής νοημοσύνης που επικεντρώνεται στη δημιουργία συστημάτων τα οποία μπορούν να «μαθαίνουν» από δεδομένα και να λαμβάνουν αποφάσεις με βάση αυτά.
- Αντί να προγραμματίζονται ρητά για να εκτελέσουν μια εργασία, τα συστήματα ML εκπαιδεύονται χρησιμοποιώντας μεγάλους όγκους δεδομένων, τους οποίους αξιοποιούν για να κάνουν προβλέψεις ή να λαμβάνουν αποφάσεις.
- Υπάρχουν μέθοδοι **τεχνητής νοημοσύνης** που **δεν** εντάσσονται στον κλάδο της μηχανικής μάθησης, π.χ.:
 - Αναγνώριση άφιξης σεισμικών σημάτων μέσω χαρακτηριστικών συναρτήσεων (π.χ., μέθοδος STA/LTA).
 - Γενικευμένη αντιστροφή (π.χ., λύση εστιακών παραμέτρων, σεισμική τομογραφία).
 - Αριθμητική προσομοίωση διάδοσης σεισμικών κυμάτων (π.χ., μέθοδος πεπερασμένων διαφορών, μέθοδος πεπερασμένων στοιχείων, κτλ.).

Διαφορά με συμβατικά μοντέλα

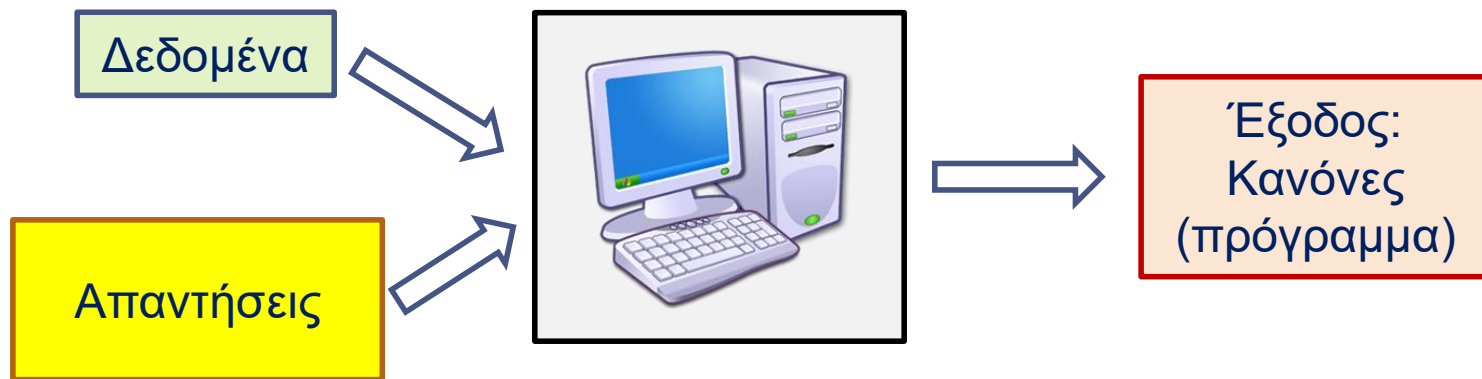
- Στα **συμβατικά μοντέλα**, φτιάχνουμε εμείς κάποιους κανόνες (πρόγραμμα), δίνουμε τα δεδομένα, και το μοντέλο παράγει κάποια αποτελέσματα.
- Δηλαδή πρέπει εμείς να βρούμε κάποιον τρόπο με τον οποίο να υπολογίζουμε τα επιθυμητά αποτελέσματα.



- Αυτή η κλασική διαδικασία μπορεί να εφαρμοστεί σε σχετικά απλά προβλήματα, στα οποία είναι γνωστοί οι απαραίτητοι υπολογισμοί που πρέπει να γίνουν για να παραχθούν τα αποτελέσματα.

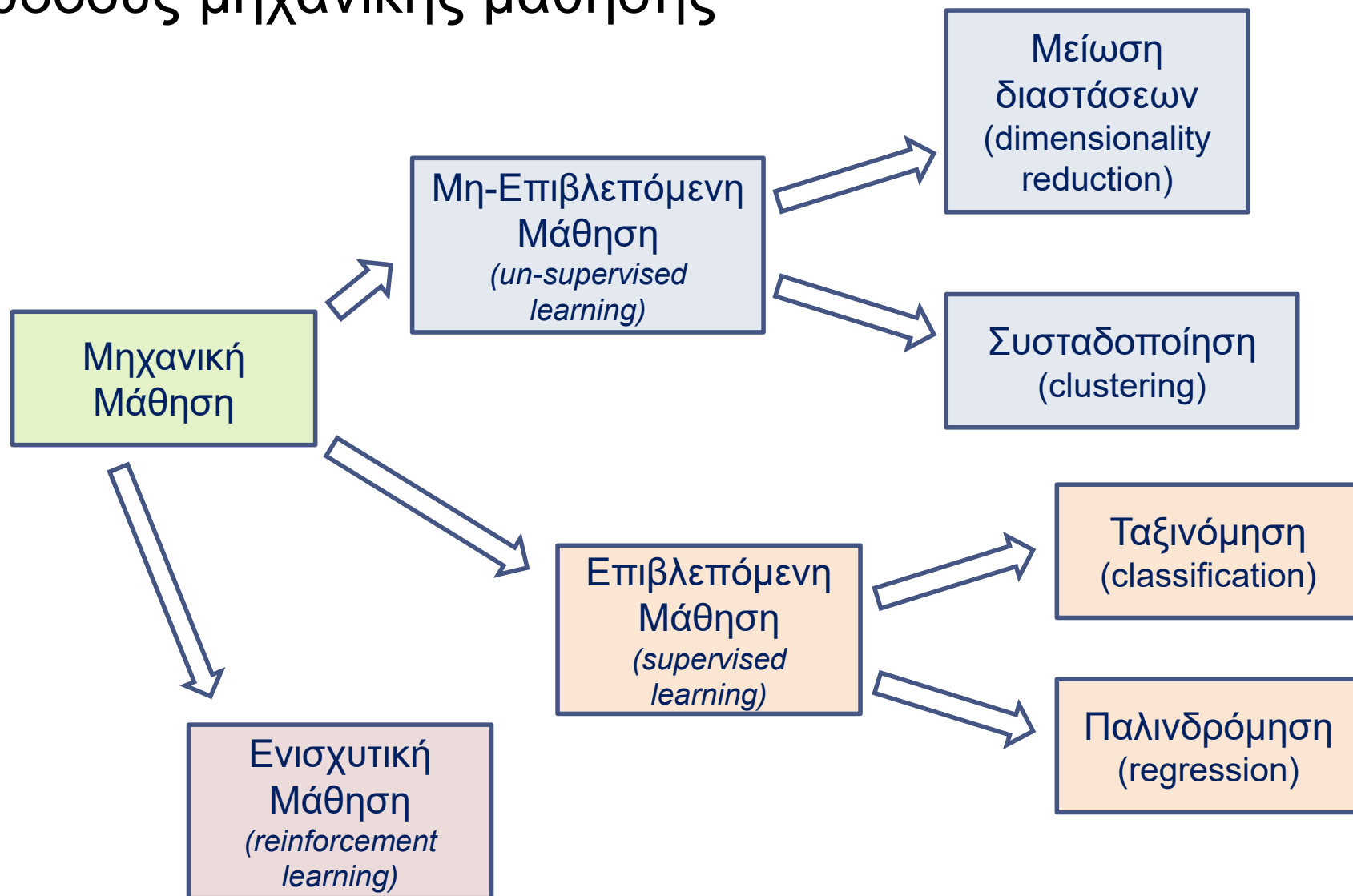
Διαφορά με συμβατικά μοντέλα

- Στα **μοντέλα μηχανικής μάθησης**, εμείς δίνουμε τα δεδομένα (και συνήθως την επιθυμητή έξοδο / απαντήσεις), και ο υπολογιστής φτιάχνει το «πρόγραμμα».
- Δηλαδή, στα μοντέλα μηχανικής μάθησης ο υπολογιστής βρίσκει τον (άγνωστο) αλγόριθμο ο οποίος από τα δεδομένα παράγει τα επιθυμητά αποτελέσματα.



- Τα μοντέλα μηχανικής μάθησης βρίσκουν λύσεις σε προβλήματα για τα οποία δεν υπάρχει σαφής τρόπος να επιλυθούν με συμβατικές μεθόδους, ή επιτυγχάνουν μεγαλύτερη ακρίβεια και καλύτερες επιδόσεις σε σύγκριση με συμβατικές μεθόδους.

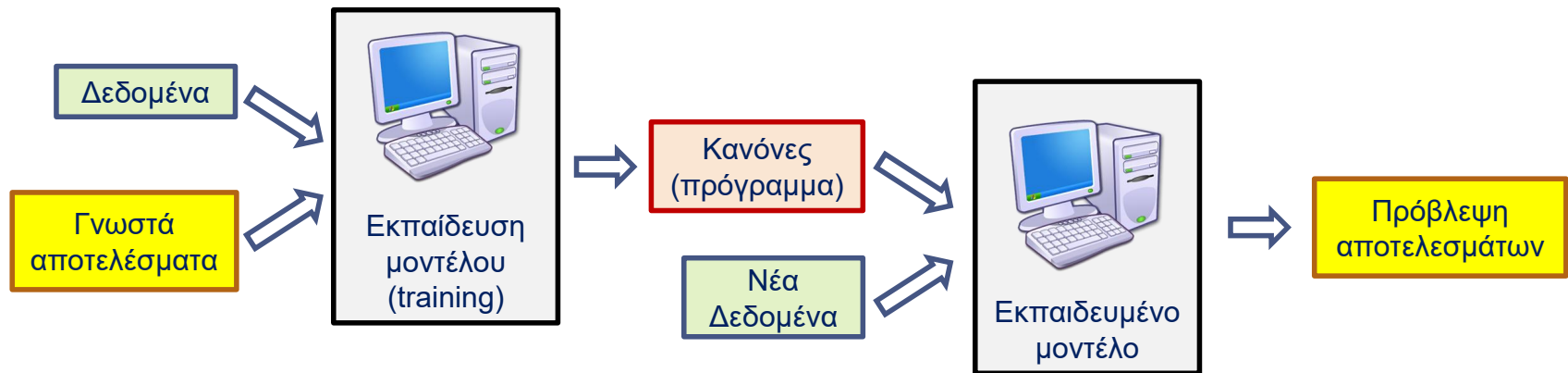
Γενικοί τύποι προβλημάτων που επιλύονται με μεθόδους μηχανικής μάθησης



Επιβλεπόμενη μάθηση

(supervised learning)

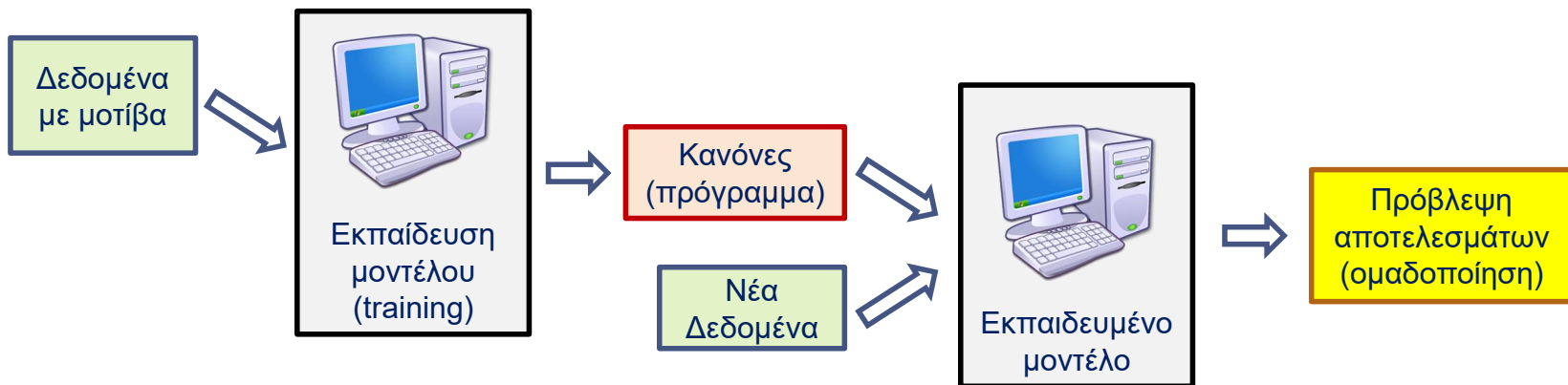
- Στην υποκατηγορία **μοντέλων μηχανικής μάθησης** με επιβλεπόμενη μάθηση, δίνουμε αρχικά στο μοντέλο κάποια δεδομένα για τα οποία γνωρίζουμε ήδη τα επιθυμητά αποτελέσματα, ώστε με αυτά να «εκπαιδεύσουμε» το μοντέλο.
- Στη συνέχεια, μπορούμε να δώσουμε στο μοντέλο νέα δεδομένα με άγνωστα αποτελέσματα, και το μοντέλο να μας δώσει μια πρόβλεψη για τα αποτελέσματα που θα είχαμε αν εξετάζαμε εμείς τα δεδομένα.



Μη-επιβλεπόμενη μάθηση

(un-supervised learning)

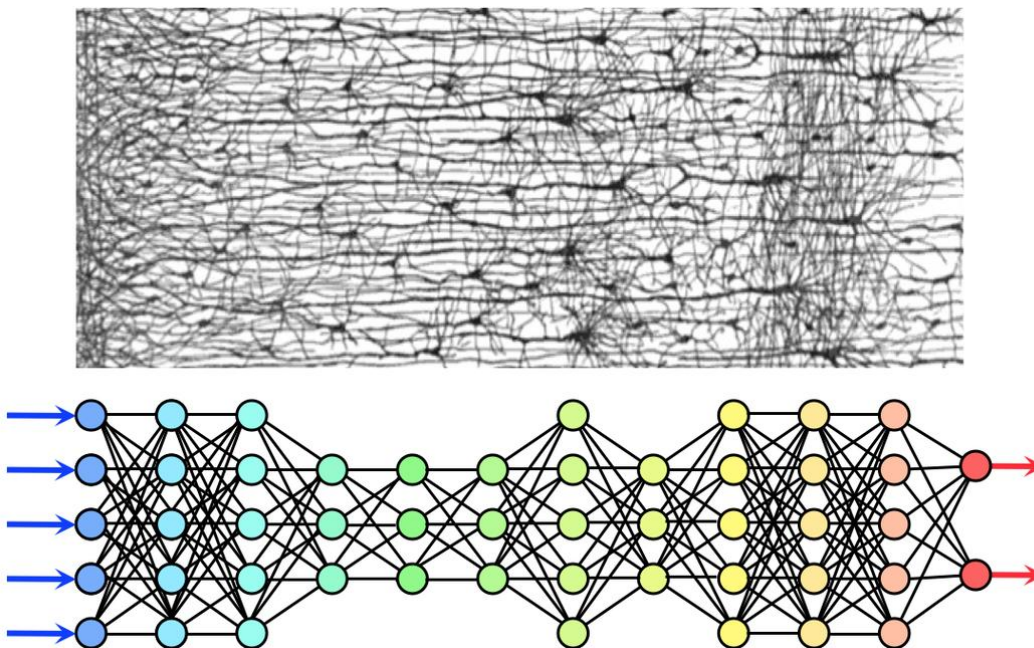
- Στην υποκατηγορία **μοντέλων μηχανικής μάθησης** με μη-επιβλεπόμενη μάθηση, δίνουμε αρχικά στο μοντέλο κάποια δεδομένα τα οποία θεωρούμε ότι αποτελούνται από ομάδες με κάποια χαρακτηριστικά που τις διαφοροποιούν.
- Το μοντέλο εκπαιδεύεται ώστε να αναγνωρίζει διάφορα κοινά μοτίβα που χαρακτηρίζουν τα δεδομένα και τα ομαδοποιεί με βάση αυτά.
- Στη συνέχεια μπορούμε να δώσουμε στο μοντέλο νέα δεδομένα, και το μοντέλο να τα ομαδοποιήσει με βάση τα κοινά τους χαρακτηριστικά με τα προηγούμενα δεδομένα με τα οποία το μοντέλο είχε ήδη εκπαιδευτεί.
- Μπορούμε, εκ των υστέρων, να θέσουμε «ετικέτες» (labels) στις διάφορες ομάδες που έχει αναγνωρίσει αυτόματα το μοντέλο.



Τι είναι η Βαθιά Μάθηση

(Deep Learning, ή DL)

- Η **Βαθιά Μάθηση** είναι ένα υποσύνολο της μηχανικής μάθησης, στο οποίο αλγόριθμοι, εμπνευσμένοι από τη δομή και τη λειτουργία του εγκεφάλου (που ονομάζονται **τεχνητά νευρωνικά δίκτυα**, ή ΤΝΔ), «μαθαίνουν» από μεγάλους όγκους δεδομένων.
- Αυτοί οι αλγόριθμοι μιμούνται τον τρόπο με τον οποίο μαθαίνουν οι άνθρωποι, βελτιώνοντας σταδιακά την ακρίβεια και την κατανόησή τους.
- Χαρακτηριστικό των μοντέλων βαθιάς μάθησης είναι ότι αποτελούνται από **πολλαπλά** (κρυμμένα) **επίπεδα τεχνητών νευρώνων** που είναι διασυνδεδεμένα μεταξύ τους.



Τι είναι η Βαθιά Μάθηση

(Deep Learning, ή DL)

- Παραδείγματα μοντέλων **μηχανικής μάθησης** που **δεν** ανήκουν στην κατηγορία της βαθιάς μάθησης:
 - Κ-πλησιέστεροι γείτονες (K-nearest Neighbors, KNN, π.χ. για εφαρμογές συσταδοποίησης σεισμικών υποκέντρων ή όμοιων σεισμών)
 - Τυχαία Δάση (Random forests, π.χ. για εφαρμογές σεισμικής επικινδυνότητας)
 - Ανάλυση κύριων συνιστωσών (Principal Component Analysis, PCA, π.χ. για ανάλυση πόλωσης της κίνησης υλικού σημείου)
 - Μοντέλα Γκαουσιανού Μείγματος (Gaussian Mixture Models, GMM, π.χ. για συσχέτισμό φάσεων με σεισμικά γεγονότα)
 - Μπεϋζιανά δίκτυα (Bayesian networks, π.χ. για μοντέλα σεισμικού κινδύνου)

Τι είναι η Βαθιά Μάθηση

(Deep Learning, ή DL)

- Παραδείγματα σεισμολογικών εφαρμογών **Βαθιάς Μάθησης**:
 - Αναγνώριση σεισμικών σημάτων
 - Προσδιορισμός χρόνων άφιξης σεισμικών φάσεων
 - Προσδιορισμός εστιακών παραμέτρων
 - Κατηγοριοποίηση σεισμικών σημάτων (π.χ., τεκτονικοί σεισμοί, ηφαιστειακοί σεισμοί, επιφανειακά φαινόμενα, ανθρωπογενείς θόρυβοι)
 - Υπολογισμός σεισμικού μεγέθους
 - Διαχωρισμός σεισμικού σήματος από τον θόρυβο υποβάθρου
 - Έγκαιρη προειδοποίηση σεισμών (γρήγορη αναγνώριση σεισμών και αυτόματη λήψη αποφάσεων)

Πλεονεκτήματα χρήσης μοντέλων βαθιάς μάθησης

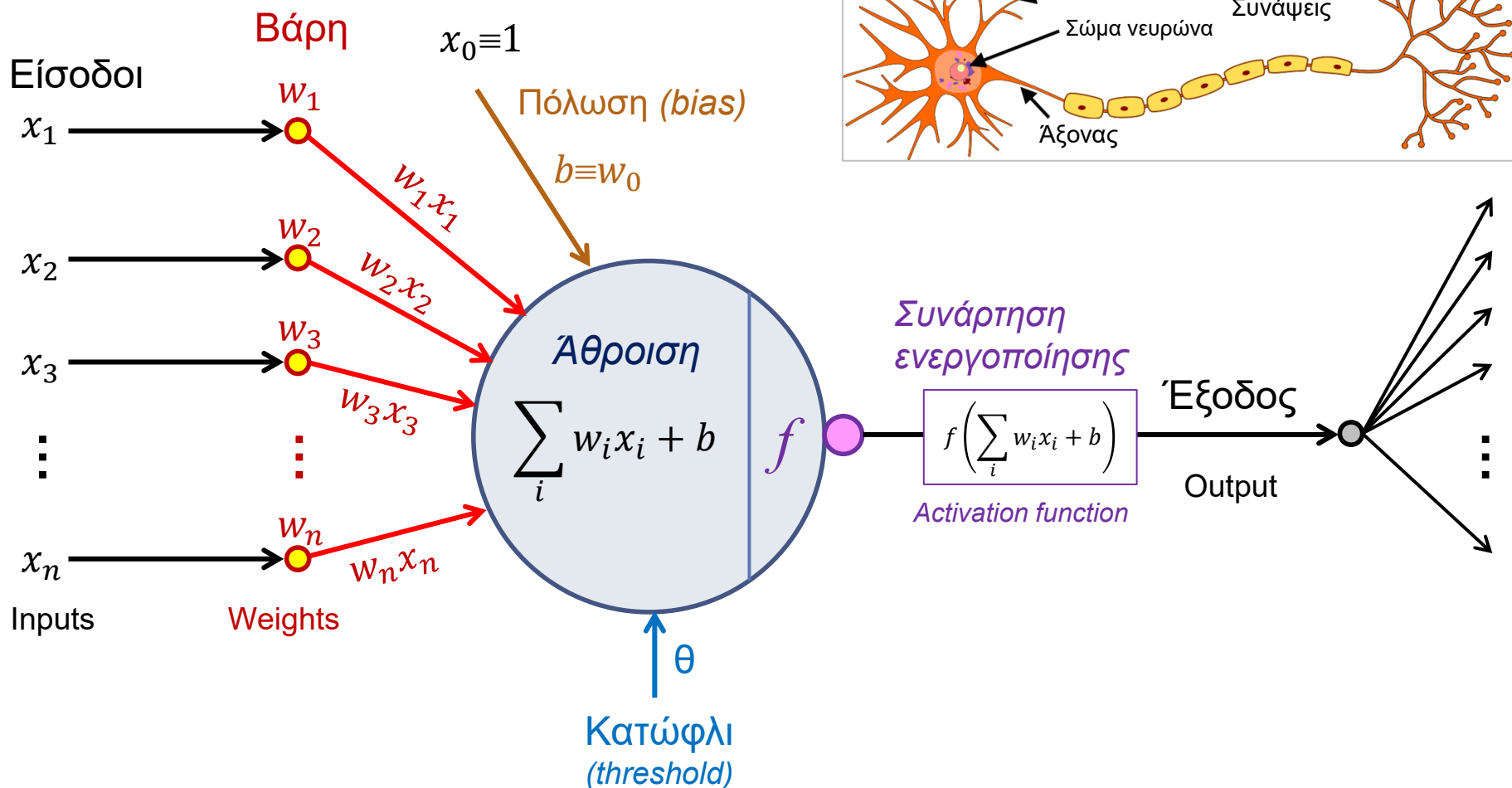
(Deep Learning, ή DL)

- **Υψηλή ακρίβεια αποτελεσμάτων:** τα αποτελέσματα της ανάλυσης με μεθόδους βαθιάς μάθησης είναι συγκρίσιμα σε ακρίβεια με τα αποτελέσματα χειρακτικής (manual) ανάλυσης από ειδικευμένο προσωπικό.
- **Διαχείριση μεγάλου όγκου δεδομένων σε ελάχιστο χρόνο:** παράλληλη επεξεργασία δεδομένων με χρήση GPU.
- **Αναγνώριση μοτίβων:** τα μοντέλα βαθιάς μάθησης είναι πολύ αποδοτικότερα στην αναγνώριση μοτίβων σε σύγκριση με συμβατικά μοντέλα, ακόμα και υπό συνθήκες χαμηλού λόγου σήματος-προς-θόρυβο.
- **Προσαρμοστικότητα:** τα μοντέλα βαθιάς μάθησης που είναι εκπαιδευμένα με δεδομένα μίας περιοχής, μπορούν να εύκολα εφαρμοστούν και σε άλλες περιοχές, σε δεδομένα με τα οποία δεν έχουν εκπαιδευτεί.

Τεχνητός νευρώνας

(Artificial neuron)

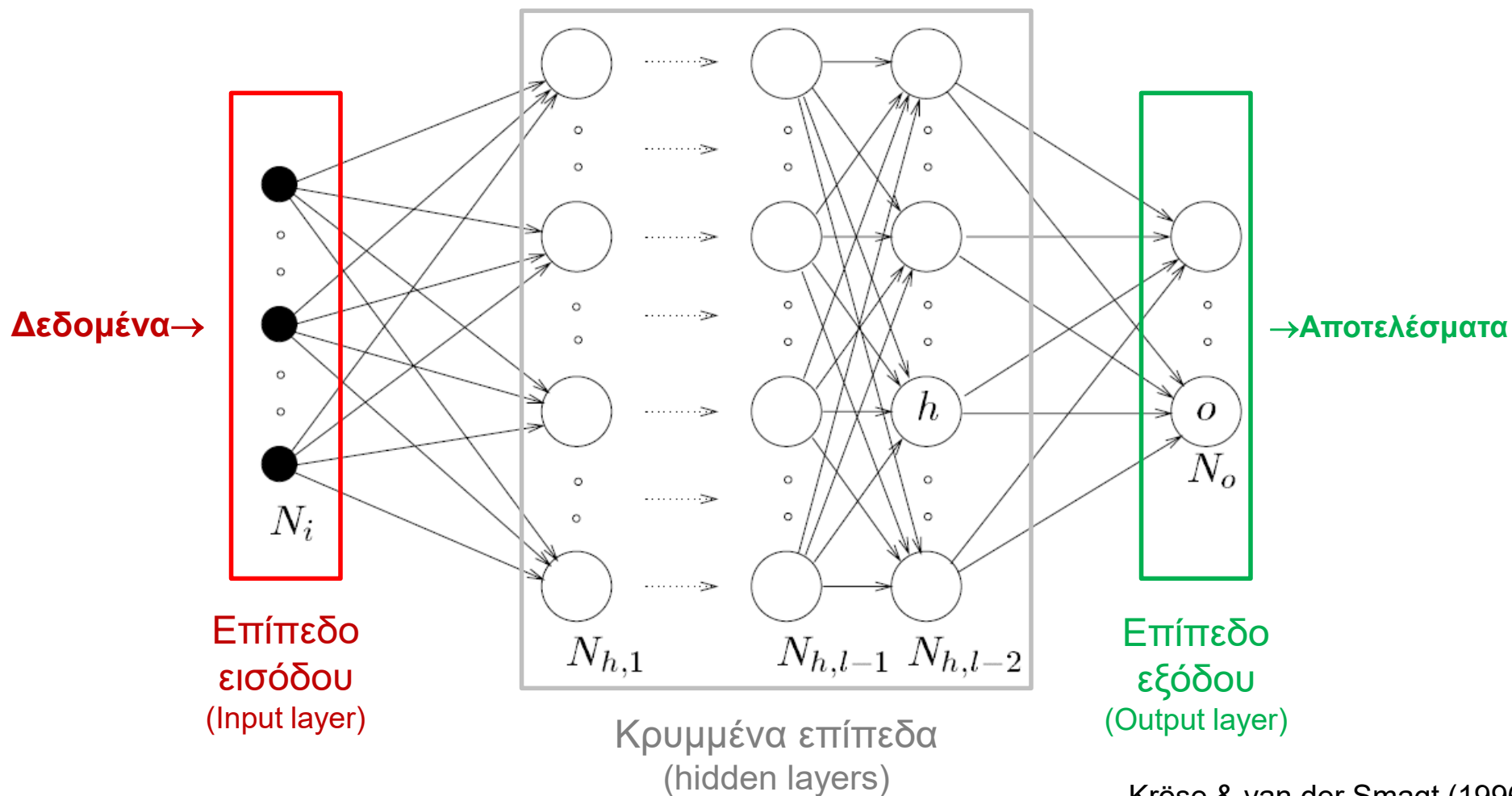
- Τα βάρη (w_i) δείχνουν το πόσο σημαντικό ρόλο παίζει η κάθε είσοδος (x_i) στην ενεργοποίηση ενός τεχνητού νευρώνα.



Τεχνητά νευρωνικά δίκτυα (ΤΝΔ)

(Artificial Neural Networks – ANNs)

- Χαρακτηριστικό των μοντέλων βαθιάς μάθησης είναι ότι αποτελούνται από πολλά κρυμμένα επίπεδα (hidden layers).



Συναρτήσεις ενεργοποίησης

(Activation functions)

$$f\left(\sum_i w_i x_i + b\right)$$

- Οι συναρτήσεις ενεργοποίησης εισάγουν **μη γραμμικότητα** στο τεχνητό νευρωνικό δίκτυο.
- Η έξοδος της συνάρτησης ενεργοποίησης μπορεί είτε να τροφοδοτήσει έναν άλλο νευρώνα που ανήκει σε άλλο επίπεδο, είτε να δώσει ένα τελικό αποτέλεσμα.
- Κάθε είδος συνάρτησης ενεργοποίησης έχει πλεονεκτήματα και μειονεκτήματα και βρίσκει εφαρμογή σε διαφορετικά προβλήματα.

Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = \max(0, z)$	Multi-layer Neural Networks	
Rectifier, softplus	$\phi(z) = \ln(1 + e^z)$	Multi-layer Neural Networks	

Υπερπαράμετροι ΤΝΔ

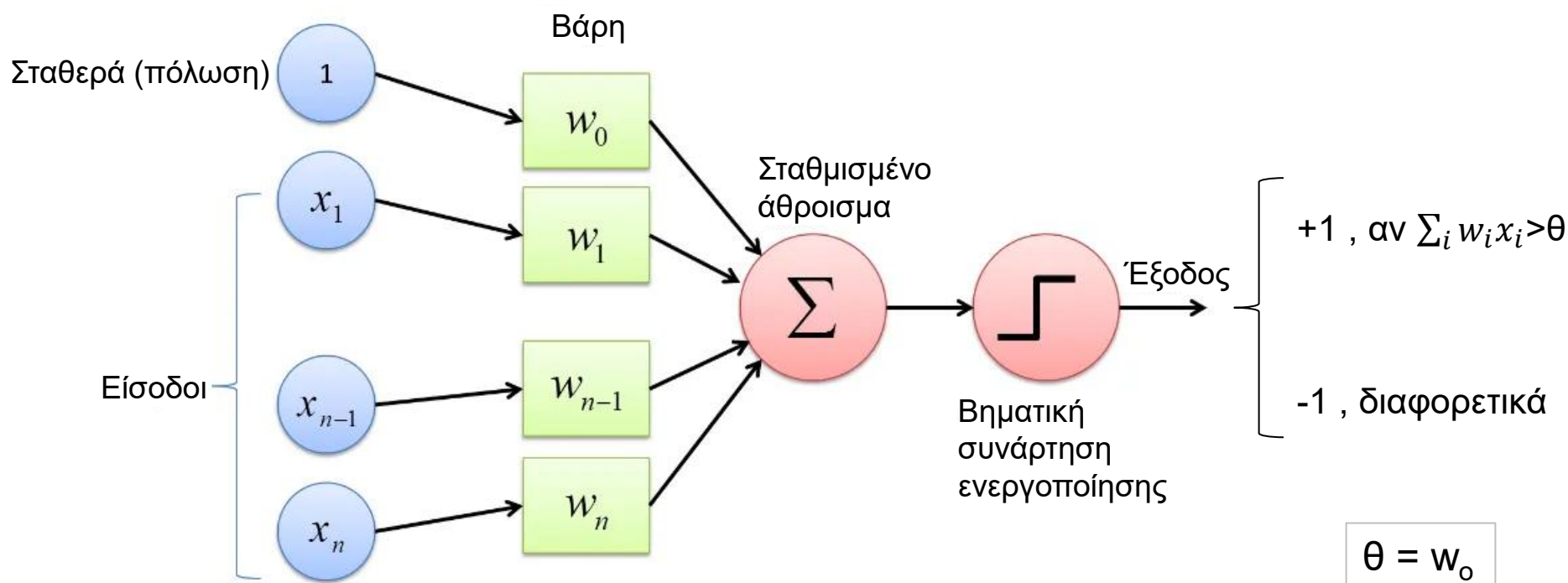
(Hyperparameters)

Οι υπερπαράμετροι ενός ΤΝΔ είναι κάποιες παράμετροι που τίθενται από τον χρήστη/προγραμματιστή κατά την κατασκευή και την εκπαίδευση του ΤΝΔ.

- Αρχιτεκτονική ΤΝΔ (architecture):
 - Αριθμός κρυμμένων επιπέδων
 - Αριθμός νευρώνων ανά επίπεδο
 - Συναρτήσεις ενεργοποίησης
 - Τύπος αρχιτεκτονικής ΤΝΔ (π.χ. ΤΝΔ πρόσθιας τροφοδότησης, επαναληπτικά ΤΝΔ, συνελκτικά ΤΝΔ, κτλ.)
- Εκπαίδευση ΤΝΔ (training):
 - Ρυθμός εκμάθησης (learning rate)
 - Μέγεθος πακέτου (batch size)
 - Αριθμός εποχών εκπαίδευσης (epochs)
 - Αλγόριθμος βελτιστοποίησης (optimizer)
 - Συνάρτηση απώλειας (loss function)
- Ομογενοποίηση (regularization):
 - Ρυθμός εγκατάλειψης (dropout rate)
 - Μείωση βάρους (weight decay)
 - Κριτήρια πρόωρης διακοπής εκπαίδευσης

Το μοντέλο Perceptron

- Ένα από τα πρώτα απλά ΤΝΔ που προτάθηκαν είναι το μοντέλο **Perceptron**.
- Το μοντέλο Perceptron αποτελείται από:
 - Ν εισόδους (x_i) (inputs) με οποιεσδήποτε αριθμητικές τιμές
 - Έναν (1) τεχνητό νευρώνα / μονάδα επεξεργασίας (unit)
 - Κανένα (0) κρυμμένο επίπεδο (hidden layer)
 - Βηματική συνάρτηση (step function) ενεργοποίησης, τύπου κατωφλίου (threshold)
- Ο στόχος του Perceptron είναι από τα στοιχεία των Ν εισόδων να ταξινομεί ανάμεσα σε δύο κατηγορίες (0 / 1, ή συνήθως -1 / +1)



Εκπαίδευση ΤΝΔ: Perceptron

(training)

- Προκειμένου ένα ΤΝΔ να μπορέσει να δώσει τα επιθυμητά αποτελέσματα θα πρέπει να «εκπαιδευτεί» με κατάλληλα προσαρμοσμένα δεδομένα.
- Η **εκπαίδευση** ενός ΤΝΔ είναι η διαδικασία με την οποία ρυθμίζονται τα **βάρη** w_i σε όλες τις εισόδους των μονάδων επεξεργασίας (τεχνητών νευρώνων) του ΤΝΔ, σε όλα τα επίπεδα (κρυμμένα και μη) που το αποτελούν.
- Σε ένα Perceptron:
 1. Αρχικά, τα βάρη τυχαιοποιούνται (random weights).
 2. Εφαρμόζεται το Perceptron για κάθε διάνυσμα εισόδου από το σετ εκπαίδευσης.
 3. Όταν το Perceptron βγάζει ένα λανθασμένο αποτέλεσμα, γίνεται αναπροσαρμογή των βαρών w_i σύμφωνα με τον **κανόνα εκπαίδευσης του Perceptron**.
 4. Η διαδικασία επαναλαμβάνεται από το βήμα (2), μέχρι όλα τα παραδείγματα από το σετ εκπαίδευσης να δίνουν σωστά αποτελέσματα.

Κανόνας εκπαίδευσης Perceptron: $\Delta w_i = \eta * ((t-o) * x_i)$

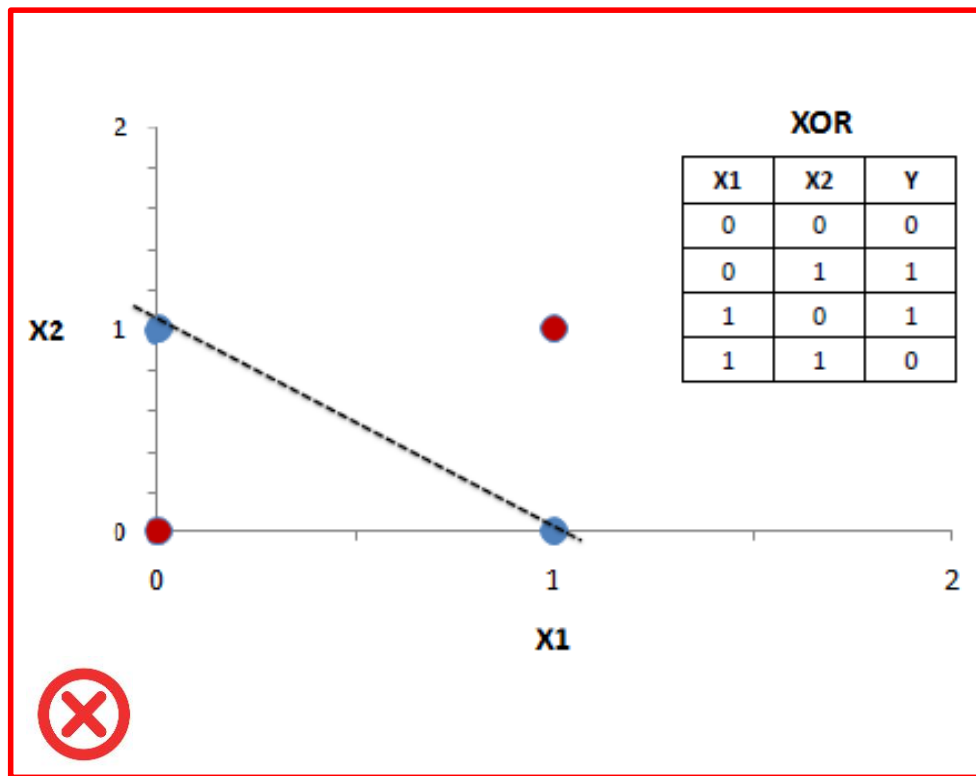
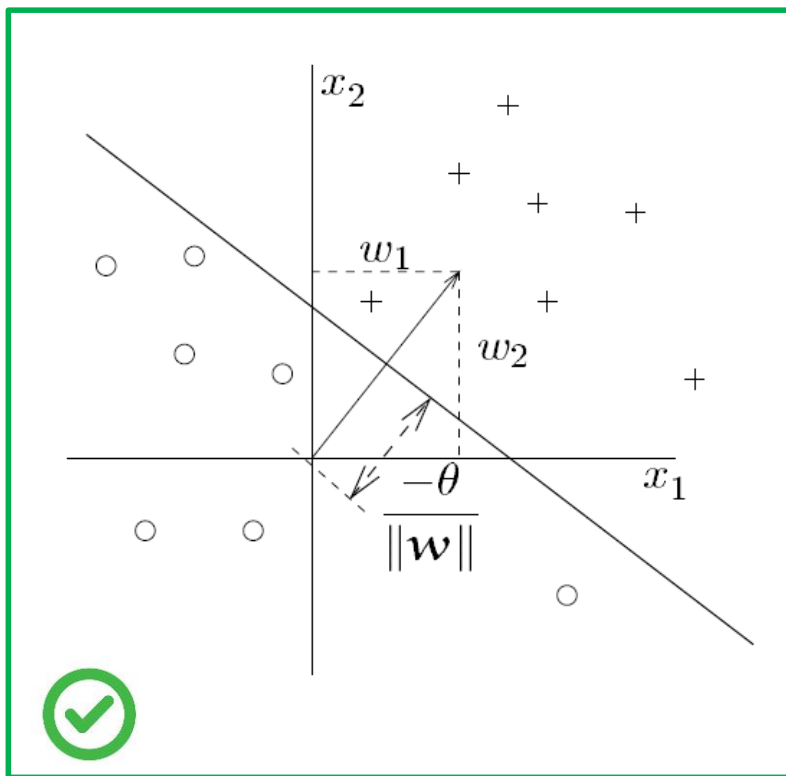
- Δw_i : διόρθωση του βάρους w_i της εισόδου x_i ,
- η : ρυθμός εκμάθησης (learning rate)
- t : σωστό/επιθυμητό αποτέλεσμα/στόχος (π.χ. -1 ή +1)
- o : πραγματική τιμή εξόδου Perceptron (π.χ. -1 ή +1)

Σωστό αποτέλεσμα: $\Delta w_i = 0$

Λάθος αποτέλεσμα: π.χ. $\Delta w_i = \pm \eta \cdot 2 \cdot x_i$

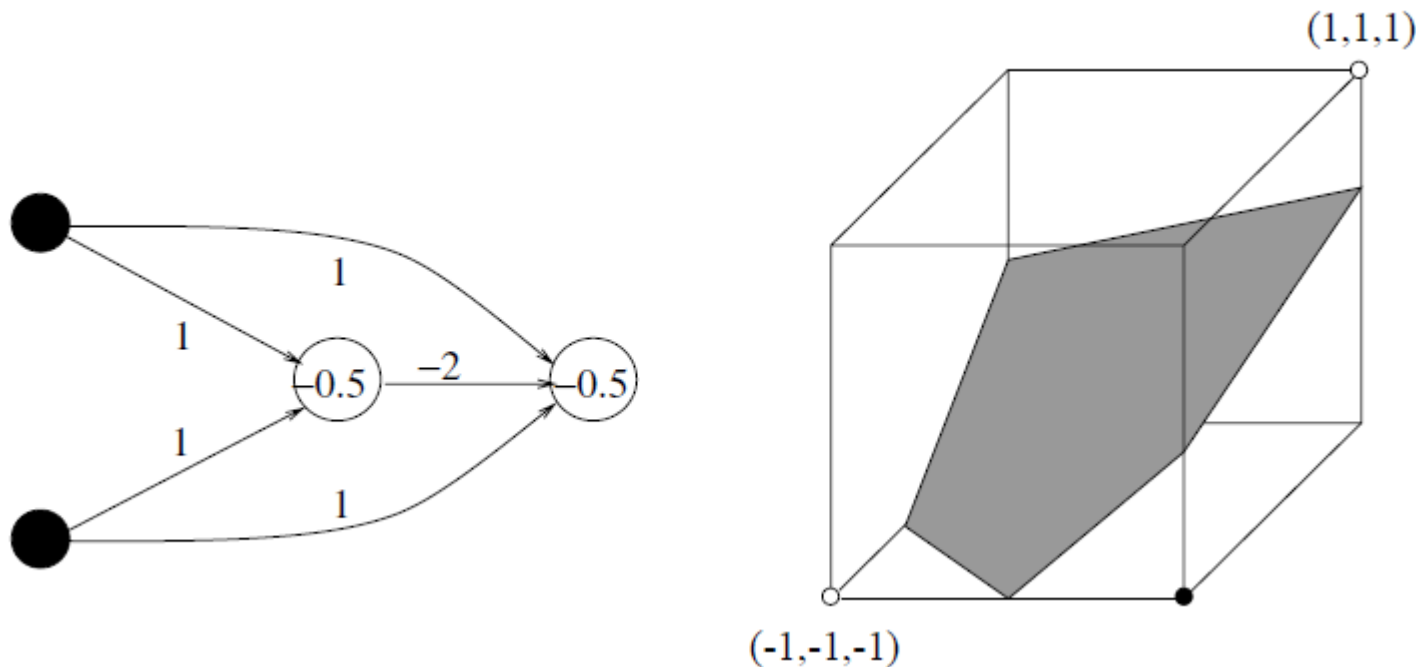
Το μοντέλο Perceptron

- Το μοντέλο Perceptron μπορεί να επιλύσει προβλήματα **δύο (2) κλάσεων** (binary classification), τα οποία είναι **γραμμικά διαχωρίσιμα**, μέσω μιας ευθείας σε 2 διαστάσεις ή ενός κατάλληλου (υπερ)επιπέδου σε N διαστάσεις.
- Μπορεί να μιμηθεί τους λογικούς τελεστές σύζευξης, AND, και διάζευξης, OR (γραμμικά διαχωρίσιμοι).
- ΔΕΝ μπορεί να μιμηθεί τον λογικό τελεστή αποκλειστικής διάζευξης, XOR (γραμμικά μη-διαχωρίσιμος).



Το πολυεπίπεδο μοντέλο Perceptron (MLP)

- Το πρόβλημα του λογικού τελεστή XOR (eXclusive OR) μπορεί να λυθεί αν στο μοντέλο Perceptron προστεθεί ένα **κρυμμένο επίπεδο** (hidden layer).
- Το κρυμμένο επίπεδο προσθέτει μια τρίτη διάσταση, οπότε τα δεδομένα κατανέμονται στον χώρο των 3 διαστάσεων, και έτσι το πρόβλημα μπορεί να επιλυθεί με ένα κατάλληλο 2Δ επίπεδο να διαχωρίζει τις δύο κλάσεις.

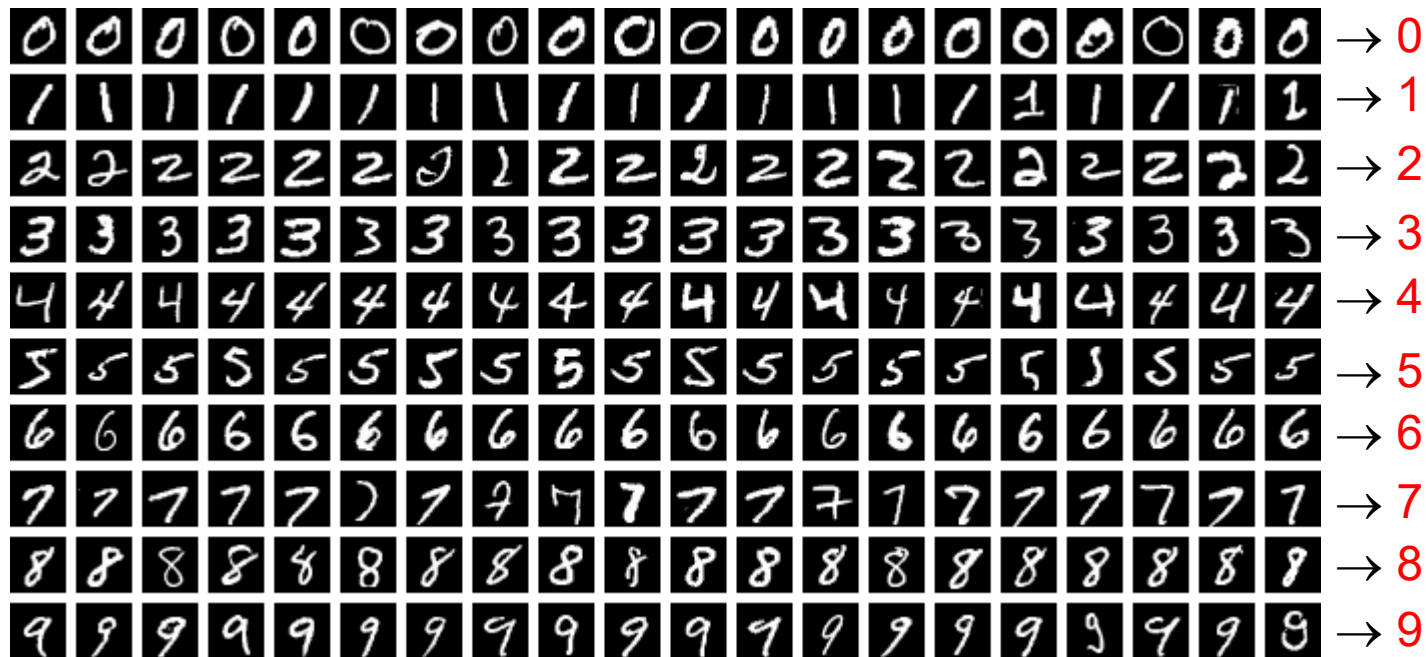


Επισημασμένα δεδομένα για εκπαίδευση ΤΝΔ

(labeled datasets)

- Προκειμένου να εκπαιδευτεί ένα ΤΝΔ με μέθοδο επιβλεπόμενης μάθησης, είναι απαραίτητο να διατίθεται ένα ευρύ σύνολο από επισημασμένα δεδομένα.
- Επισημασμένα δεδομένα (labeled data) είναι δεδομένα εισόδου για τα οποία γνωρίζουμε εκ των προτέρων τη σωστή απάντηση που θέλουμε να δώσει για αυτά το ΤΝΔ.

Απάντηση:



Επισημασμένο σύνολο δεδομένων **MNIST** για ΤΝΔ οπτικής αναγνώρισης χαρακτήρων (OCR).

Συνάρτηση απώλειας

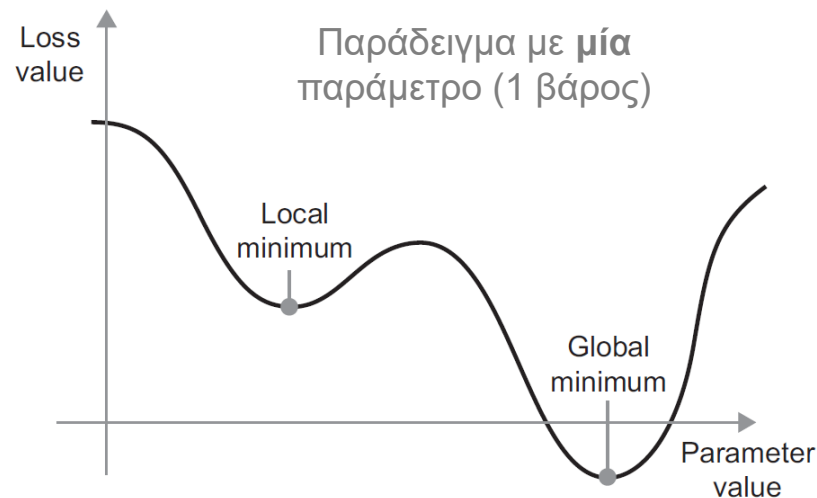
(loss function)

- Η συνάρτηση απώλειας (loss function) είναι μια μαθηματική συνάρτηση που ποσοτικοποιεί τη διαφορά μεταξύ των προβλεπόμενων τιμών από ένα ΤΝΔ και των πραγματικών τιμών που θα έπρεπε να προβλέψει.
- Είναι ένα κρίσιμο εργαλείο στην εκπαίδευση μοντέλων ΤΝΔ, καθώς παρέχει έναν τρόπο μέτρησης για το πόσο καλά το μοντέλο επιτελεί το έργο του.
- Ο σκοπός είναι να ρυθμιστούν κατάλληλα τα βάρη των μονάδων του ΤΝΔ ώστε να ελαχιστοποιηθεί η συνάρτηση απώλειας.
- Πρέπει δηλαδή με κάποιον τρόπο (με έναν αλγόριθμο βελτιστοποίησης) να βρεθεί το **ολικό ελάχιστο** της συνάρτησης απώλειας.

Παραδείγματα συναρτήσεων απώλειας:

- Μέσο τετραγωνικό σφάλμα (Mean Squared Error, MSE)
- Μέσο απόλυτο σφάλμα (Mean Absolute Error, MAE)

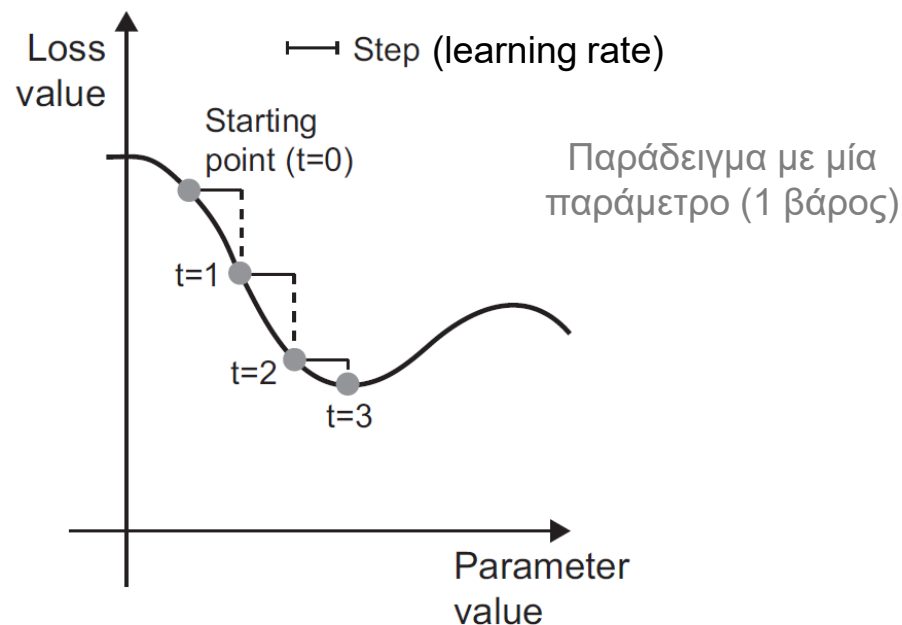
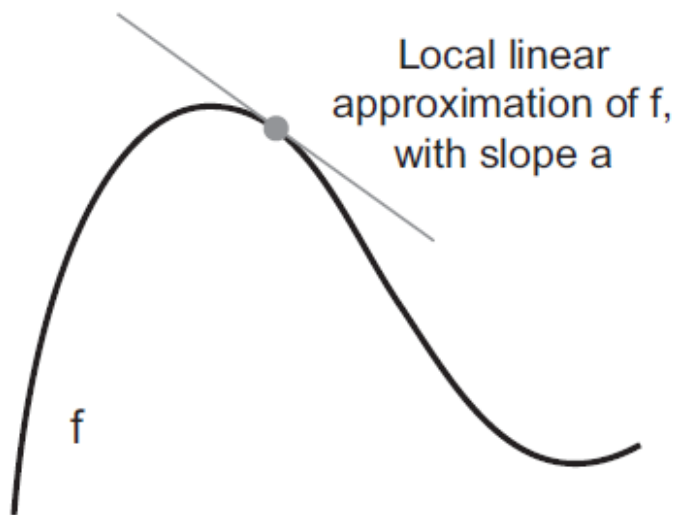
$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$



Κατάβαση κλίσης

(gradient descent)

- Η μέθοδος της **κατάβασης κλίσης** είναι ένας αλγόριθμος βελτιστοποίησης με σκοπό την εύρεση του ολικού ελαχίστου της συνάρτησης απώλειας.
- Η (μερική) παραγωγή της συνάρτησης απώλειας ως προς καθεμία από τις παραμέτρους του μοντέλου (βάρη) δημιουργεί ένα **διάνυσμα διόρθωσης** το οποίο «δείχνει την κατεύθυνση» προς την οποία πρέπει να διορθωθεί η καθεμία από τις παραμέτρους του μοντέλου ώστε να μειωθεί η απώλεια.
- Το πόσο ισχυρή θα είναι αυτή η διόρθωση καθορίζεται από τον **ρυθμό εκμάθησης** (learning rate).

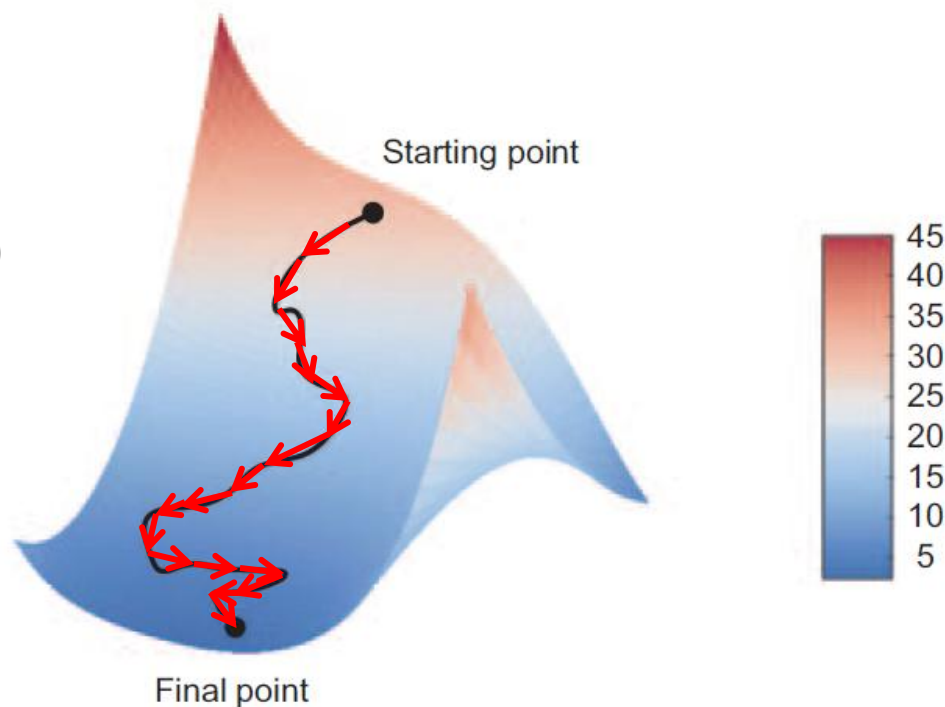


Κατάβαση κλίσης

(gradient descent)

- Καθώς ο αριθμός των βαρών ενός ΤΝΔ είναι συνήθως πολύ μεγάλος (π.χ., μεγάλα γλωσσικά μοντέλα σήμερα έχουν τρισεκατομμύρια βάρη), η συνάρτηση απώλειας βρίσκεται σε έναν χώρο **πολλών διαστάσεων**.
- Η ελαχιστοποίηση της συνάρτησης απώλειας γίνεται με μια επαναληπτική διαδικασία η οποία σε κάθε βήμα (εποχή ή epoch) θα πρέπει να οδηγεί σε σταδιακή μείωση της απώλειας (ή συνάρτησης κόστους – cost function).

Παράδειγμα με **δύο**
παραμέτρους (2 βάρη)



Εκπαίδευση ΤΝΔ: Μέθοδος οπισθο-διάδοσης του σφάλματος (back-propagation of the error)

- Η πιο διαδεδομένη μέθοδος εκπαίδευσης ΤΝΔ είναι η μέθοδος της **οπισθο-διάδοσης του σφάλματος** (back-propagation).
- Η μέθοδος αυτή αποτελεί έναν γρήγορο τρόπο υπολογισμού μερικών παραγώγων για τη ρύθμιση των βαρών στους νευρώνες των **κρυμμένων επιπέδων**, με βάση τις επιδόσεις του δικτύου στην πρόβλεψη των σωστών απαντήσεων σε επισημασμένα σετ δεδομένων.
- Ο αλγόριθμος της οπισθο-διάδοσης λειτουργεί ως εξής:
 - Δίνουμε δεδομένα στο αρχικό ΤΝΔ (αρχικά με τυχαία βάρη) και παίρνουμε ένα αποτέλεσμα **απόκρισης** του ΤΝΔ (*πρόσθια διάδοση*).
 - Υπολογίζουμε τη **συνάρτηση απώλειας**, L (σήμα σφάλματος).
 - Υπολογίζουμε και διαδίδουμε το **σήμα σφάλματος** (error signal) από την έξοδο προς τα πίσω (οπισθο-διάδοση).
 - Κατά την οπισθο-διάδοση του σήματος σφάλματος, υπολογίζεται μια **διόρθωση** Δw_{ij} σε **κάθε βάρος** w_{ij} που συνδέει έναν νευρώνα i με έναν νευρώνα j του προηγούμενου επιπέδου, προς την κατεύθυνση η οποία τείνει να **μειώσει τη συνάρτηση απώλειας**.

Εκπαίδευση ΤΝΔ: Μέθοδος οπισθο-διάδοσης του σφάλματος (back-propagation of the error)

- Ξεκινάμε από το επίπεδο εξόδου υπολογίζοντας τις μερικές παραγώγους της L ως προς τις συναρτήσεις ενεργοποίησης του επιπέδου εξόδου.
- Στη συνέχεια, με τον **κανόνα της αλυσίδας** (chain rule) της παραγωγίσιμης σύνθετων συναρτήσεων, υπολογίζονται οι μερικές παράγωγοι της ενεργοποίησης ως προς καθεμία από τις παραμέτρους εισόδου του αντίστοιχου επιπέδου.

$$L = \frac{1}{2}(y - \hat{y})^2$$

Συνάρτηση απώλειας

(π.χ. μέσο τετραγωνικό σφάλμα)

y : επιθυμητή τιμή-στόχος

$\hat{y}$: πραγματική τιμή απόκρισης ΤΝΔ

$$\frac{\partial L}{\partial \hat{y}} = \hat{y} - y$$

Μερική παράγωγος της συνάρτησης απώλειας ως προς την έξοδο του ΤΝΔ

$$z = \sum_{i=1}^n w_i x_i + b$$

$$f'(z) = \sigma(z)(1 - \sigma(z))$$

Παράγωγος της συνάρτησης ενεργοποίησης $f(z)$, π.χ. για τη σιγμοειδή, $\sigma(z)$.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

$$\delta^{(2)} = (\hat{y} - y)f'(z^{(2)})$$

Σήμα σφάλματος στο επίπεδο εξόδου.
π.χ. το $\delta^{(2)}$ για ΤΝΔ με δύο επίπεδα.

Εκπαίδευση ΤΝΔ: Μέθοδος οπισθο-διάδοσης του σφάλματος (back-propagation of the error)

- Υπολογίζουμε το σφάλμα $\delta^{(1)}$ στο **προηγούμενο κρυμμένο επίπεδο** (το επίπεδο ⁽¹⁾ για ΤΝΔ με δύο επίπεδα).

$$\delta^{(1)} = (W^{(2)})^T \delta^{(2)} \cdot f'(z^{(1)})$$

Διάδοση σφάλματος στο προηγούμενο επίπεδο, όπου το $\delta^{(2)}$ είναι το σφάλμα που υπολογίστηκε στο επίπεδο εξόδου.

Όπου: $(W^{(2)})^T$

Διάνυσμα βαρών επιπέδου εξόδου ⁽²⁾

$f'(z^{(1)})$

Παράγωγος της συνάρτησης ενεργοποίησης στο κρυμμένο επίπεδο ⁽¹⁾

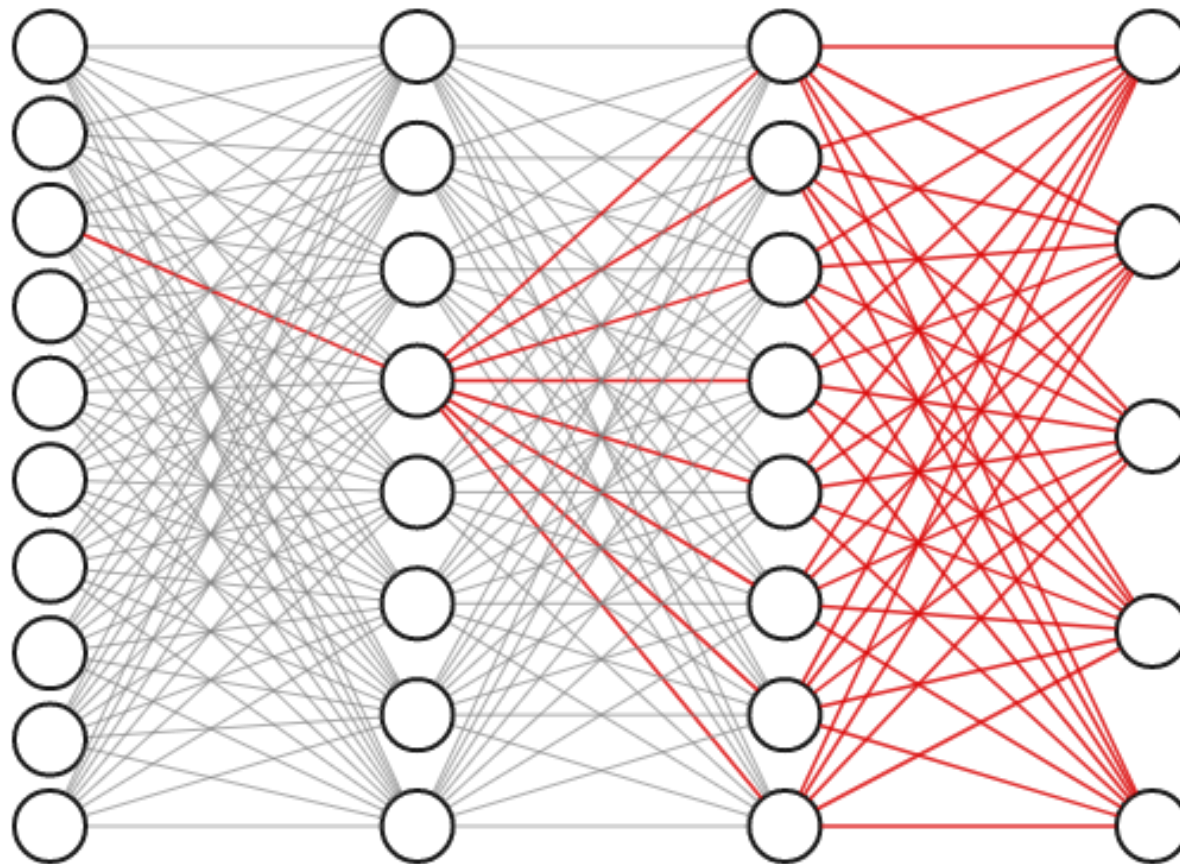
- Διορθώνουμε τα βάρη στο επίπεδο εξόδου ⁽²⁾: $W^{(2)} = W^{(2)} - \eta \cdot \delta^{(2)} \cdot a^{(1)}$
- Διορθώνουμε τα βάρη στο κρυφό επίπεδο ⁽¹⁾: $W^{(1)} = W^{(1)} - \eta \cdot \delta^{(1)} \cdot x$

Όπου η ο ρυθμός εκμάθησης, $a^{(1)} = f(z^{(1)})$ η ενεργοποίηση του κρυμμένου επιπέδου ⁽¹⁾.

- Επαναλαμβάνουμε τη διαδικασία όσες φορές χρειάζεται μέχρι να ολοκληρωθεί η εκπαίδευση του δικτύου.

Εκπαίδευση ΤΝΔ: Μέθοδος οπισθο-διάδοσης του σφάλματος (back-propagation of the error)

- Η αλλαγή βάρους σε μία σύνδεση ενός νευρώνα από το πρώτο επίπεδο με έναν νευρώνα στο δεύτερο επίπεδο επηρεάζει στη συνέχεια όλους τους νευρώνες του 3^{ου} επιπέδου, κ.ο.κ., και τελικά όλες τις εξόδους.
- Γι'αυτόν τον λόγο, οι διορθώσεις βαρών γίνονται από την έξοδο προς την είσοδο.



Συναρτήσεις ενεργοποίησης (activation functions)

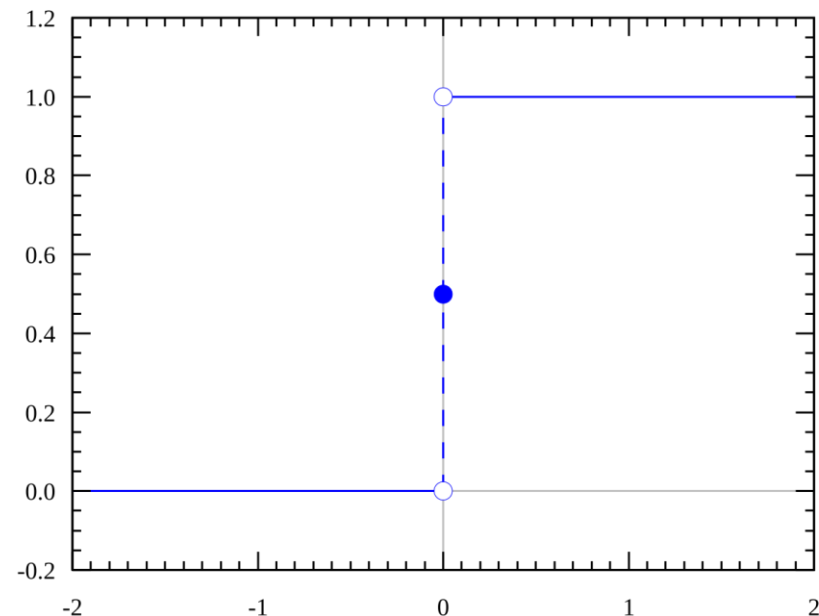
$$f\left(\sum_i w_i x_i + b\right)$$

- Προκειμένου να μπορεί να εφαρμοστεί η διαδικασία οπισθο-διάδοσης για την εκπαίδευση ενός ΤΝΔ, πρέπει οι συναρτήσεις ενεργοποίησης σε κάθε κόμβο να είναι **παντού παραγωγίσιμες**.
- Θα πρέπει επίσης η παράγωγος να **μην είναι μηδενική**, αλλιώς δεν θα υπάρχει διόρθωση βάρους.

Βηματική συνάρτηση: (ή συνάρτηση Heaviside)

$$H(x) := \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- Χρησιμοποιείται στο μοντέλο Perceptron: θέτει την έξοδο =0 για αρνητικές τιμές της σταθμισμένης άθροισης και =1 για μη-αρνητικές.
- Ακατάλληλη για εκπαίδευση με τη μέθοδο οπισθο-διάδοσης.
 - **$H'(x) = 0$** για $x < 0$ ή $x > 0$
 - $H'(x=0)$: **δεν ορίζεται**



Συναρτήσεις ενεργοποίησης (activation functions)

$$f\left(\sum_i w_i x_i + b\right)$$

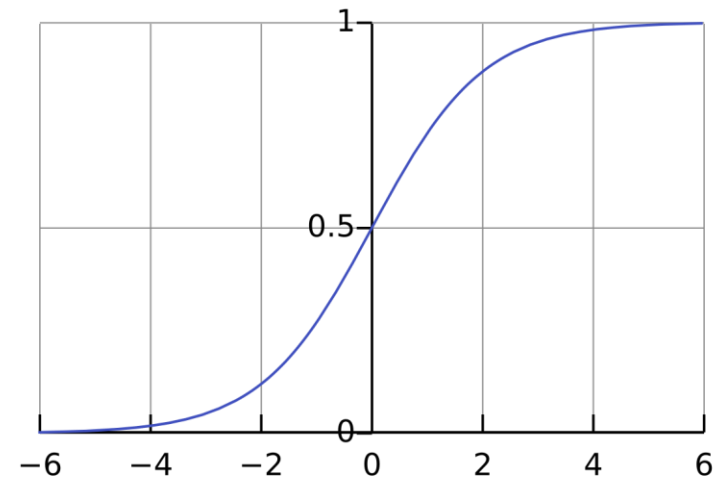
- Για να αντιμετωπιστούν τα προηγούμενα προβλήματα, εισήχθησαν νέες συναρτήσεις ενεργοποίησης οι οποίες είναι παραγωγίσιμες και έχουν μη μηδενική παράγωγο.

Σιγμοειδής συνάρτηση:

(sigmoid function ή logistic function)

- Ομαλή και παραγωγίσιμη παντού.
- Περιορίζει την έξοδο μεταξύ 0 και 1 (κατάλληλη για αναπαράσταση πιθανοτήτων)
- Χρησιμοποιείται κατά κόρον σε ΤΝΔ στο επίπεδο εξόδου για προβλήματα δυαδικής ταξινόμησης, αντί της βηματικής.
- Πρόβλημα: για μεγάλες τιμές x , η παράγωγος $f'(x)$ παίρνει πολύ μικρές τιμές (vanishing gradient)

$$f(x) = \frac{1}{1 + e^{-x}}$$



Συναρτήσεις ενεργοποίησης (activation functions)

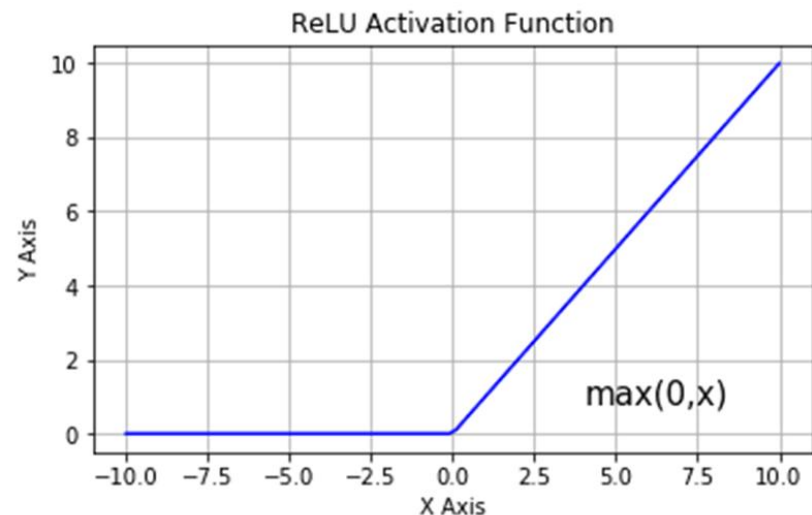
$$f\left(\sum_i w_i x_i + b\right)$$

- Κάποιες συναρτήσεις ενεργοποίησης είναι περισσότερο κατάλληλες για χρήση σε κρυμμένα επίπεδα, όπως η συνάρτηση ReLU.

Συνάρτηση ReLU (Rectified Linear Unit): (Ανορθωμένη γραμμική μονάδα)

$$\text{ReLU}(x) = x^+ = \max(0, x) = \frac{x + |x|}{2} = \begin{cases} x & \text{if } x > 0, \\ 0 & x \leq 0 \end{cases}$$

- Παρά τη δι-γραμμική της μορφή, εισάγει **μη-γραμμικότητα** στα ΤΝΔ.
- Μηδενίζει κόμβους με $x < 0$ απλοποιώντας τους υπολογισμούς (αραίωση / sparsity).
- Χρησιμοποιείται κατά κόρον σε κρυμμένα επίπεδα ΤΝΔ.
- Έχει μη μηδενική παράγωγο για $x > 0$
- Προβλήματα: η παράγωγος **δεν ορίζεται** για $x=0$ και **μηδενίζεται** για $x < 0$
- Οι **μηδενικές τιμές για $x < 0$** μπορεί να προκαλέσουν πρόβλημα κατά την εκπαίδευση του ΤΝΔ.



Συναρτήσεις ενεργοποίησης (activation functions)

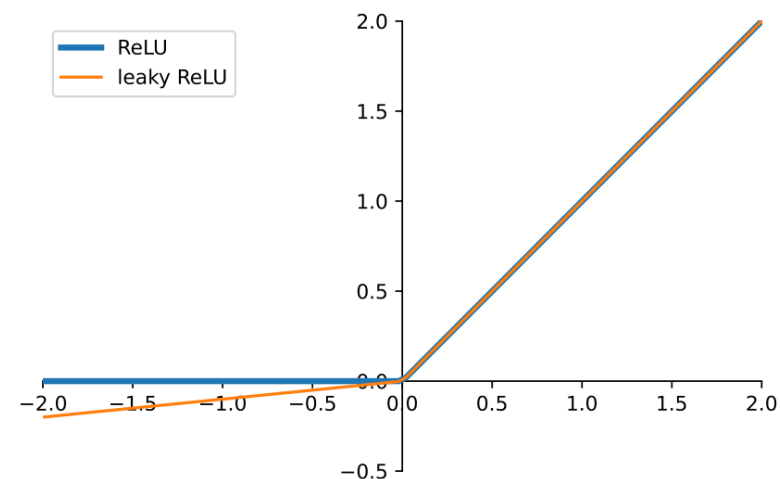
$$f\left(\sum_i w_i x_i + b\right)$$

Συνάρτηση *Leaky ReLU*:

(Διαρρέουσα ανορθωμένη γραμμική μονάδα)

- Δι-γραμμική, όπως η ReLU, αλλά με μη-μηδενικές τιμές για $x < 0$.
- Βελτιώνει τα προβλήματα της μηδενικής παραγώγου και μηδενικών τιμών για $x < 0$.
- Πρόβλημα: Παραμένει μη-παραγωγίσιμη για $x=0$, αλλά αυτό δεν θεωρείται τόσο σημαντικό πρόβλημα.

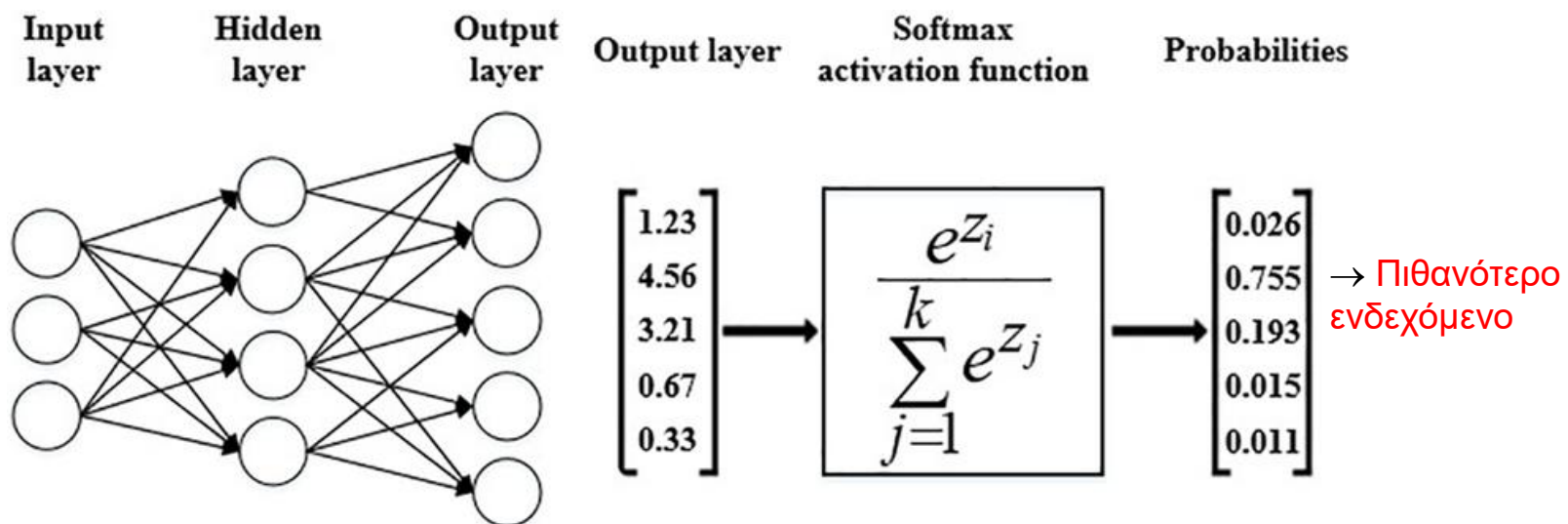
$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ \alpha x, & \text{if } x \leq 0 \end{cases}$$



Συναρτήσεις ενεργοποίησης (activation functions)

$$f\left(\sum_i w_i x_i + b\right)$$

- Κάποιες συναρτήσεις ενεργοποίησης είναι περισσότερο κατάλληλες για χρήση σε επίπεδα εξόδου, όπως η συνάρτηση **Softmax**.
- Η **Softmax** μετατρέπει αποτελέσματα τύπου «σκορ» (score) σε τιμές πιθανότητας.
- Φροντίζει ώστε τα αποτελέσματα των εξόδων να είναι μη-μηδενικά και το άθροισμά τους να ισούται με τη μονάδα.
- Χρησιμοποιείται στο επίπεδο εξόδου για προβλήματα ταξινόμησης σε πολλαπλές κλάσεις (multi-class classification).



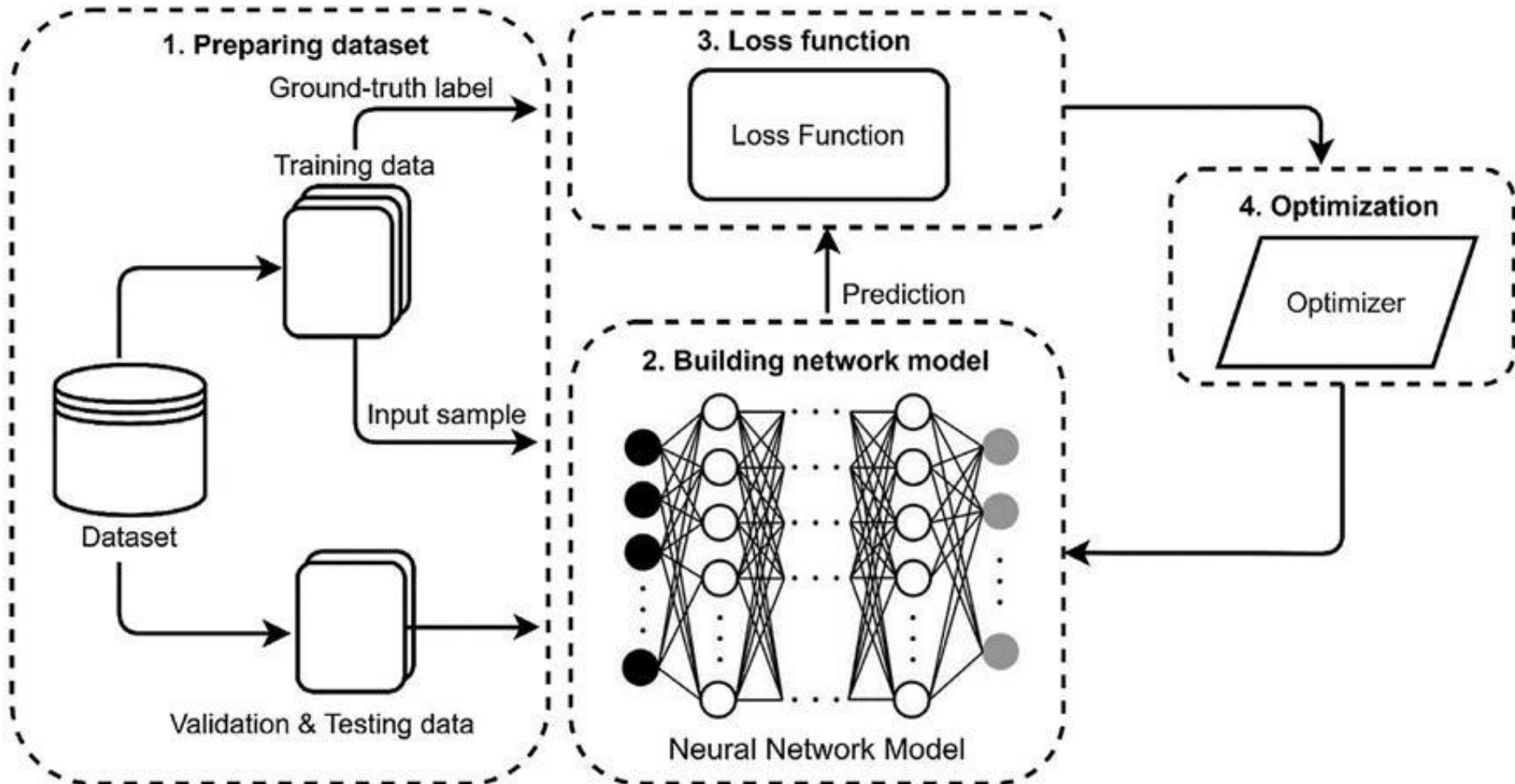
Επαλήθευση ή επικύρωση

(validation)

- Σε ένα ΤΝΔ, τα βάρη των νευρώνων ρυθμίζονται με βάση τις επιδόσεις του ΤΝΔ πάνω σε ένα **σετ δεδομένων εκπαίδευσης** (training dataset).
- Το ΤΝΔ μπορεί έτσι να γίνει πολύ καλό στο να δίνει σωστές απαντήσεις για το συγκεκριμένο σετ εκπαίδευσης, αλλά σημασία έχει να μπορεί να δίνει απαντήσεις και σε άλλα σετ δεδομένων με τα οποία το ΤΝΔ δεν έχει εκπαιδευτεί (γενίκευση μοντέλου).
- Γι' αυτό, σε κάθε εποχή (epoch) εκπαίδευσης, πέρα από τις δοκιμές με το σετ εκπαίδευσης, με βάση το οποίο προσαρμόζονται τα βάρη, ελέγχονται και οι επιδόσεις σε ένα **σετ επικύρωσης**, ή επαλήθευσης (validation dataset).
- Η επικύρωση γίνεται για να εξεταστεί κατά πόσο πρέπει ή όχι να συνεχίσει η εκπαίδευση του μοντέλου, εφόσον βελτιώνονται οι επιτυχείς απαντήσεις **και** στο σετ επικύρωσης (μειώνεται η συνάρτηση κόστους σε αυτά τα δεδομένα).
- Αν τα ποσοστά επιτυχίας στις απαντήσεις του σετ επικύρωσης σταματήσουν να βελτιώνονται ή αρχίσουν να χειροτερεύουν, ενώ στο σετ εκπαίδευσης συνεχίζουν να βελτιώνονται, αυτό σημαίνει πως το ΤΝΔ έχει αρχίσει να παθαίνει **υπερπροσαρμογή** (overfitting) στα δεδομένα εκπαίδευσης.

Εκπαίδευση / Επικύρωση

(training / validation)



Υπερπροσαρμογή

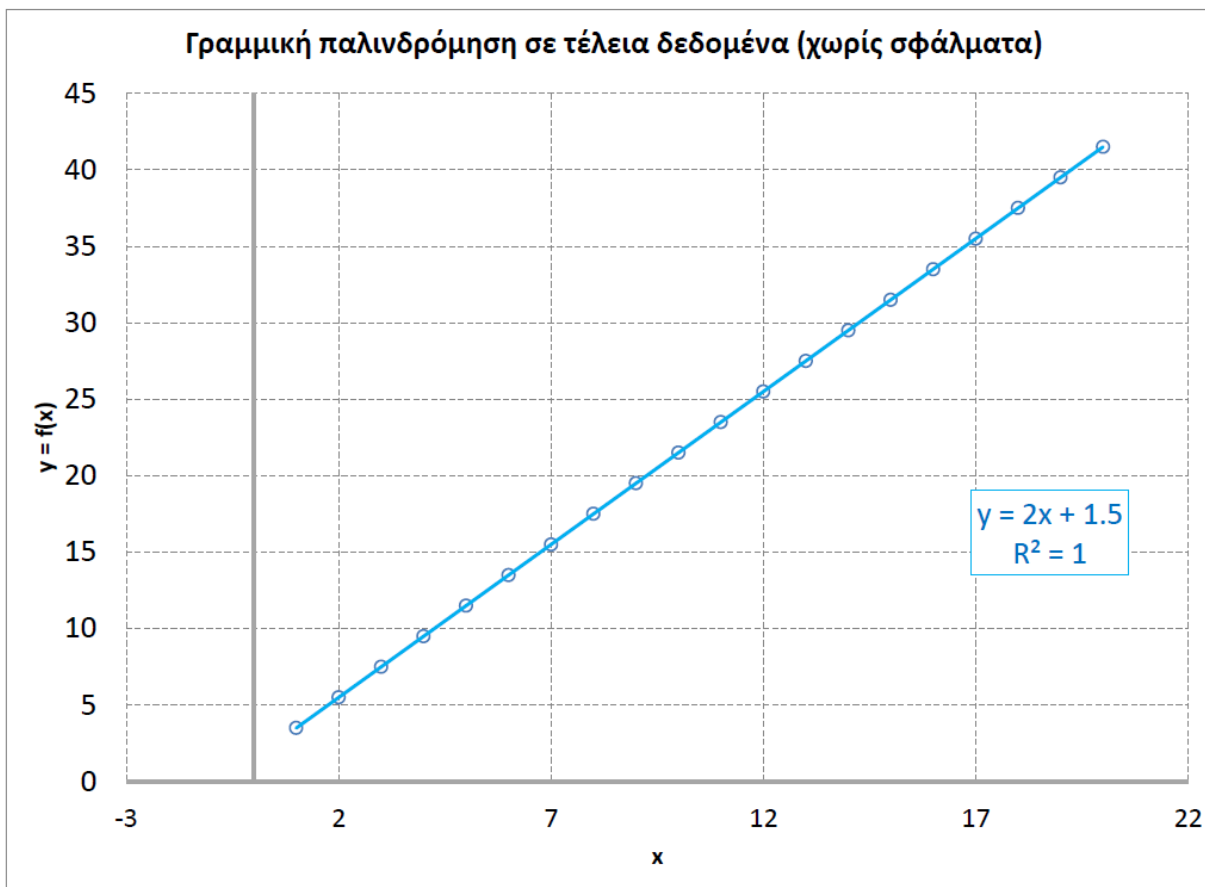
(overfitting)

- Εφόσον υπάρχει σύγκλιση σε ένα ολικό ελάχιστο της συνάρτησης κόστους, το ΤΝΔ μπορεί να εκπαιδευτεί ώστε να δίνει σωστές απαντήσεις σε κάθε δείγμα που του παρουσιάζεται στην είσοδο με σκοπό την εκπαίδευση του ΤΝΔ.
- Ωστόσο, η επιθυμητή λειτουργία ενός ΤΝΔ είναι να μπορεί να κάνει σωστές προβλέψεις σε δείγματα με τα οποία **δεν** έχει εκπαιδευτεί.
- Δηλαδή, πρέπει να μπορεί να γίνει **γενίκευση** (generalization) του μοντέλου σε ευρύτερα σύνολα δεδομένων.
- Όταν ένα μοντέλο έχει καλές επιδόσεις μόνο στα δεδομένα με τα οποία έχει εκπαιδευτεί αλλά αποτυγχάνει στη γενίκευση, τότε λέμε ότι το μοντέλο έχει υποστεί **υπερπροσαρμογή** (overfitting) στα δεδομένα εκπαίδευσης.

Υπερπροσαρμογή σε συμβατικά μοντέλα

Γραμμική παλινδρόμηση

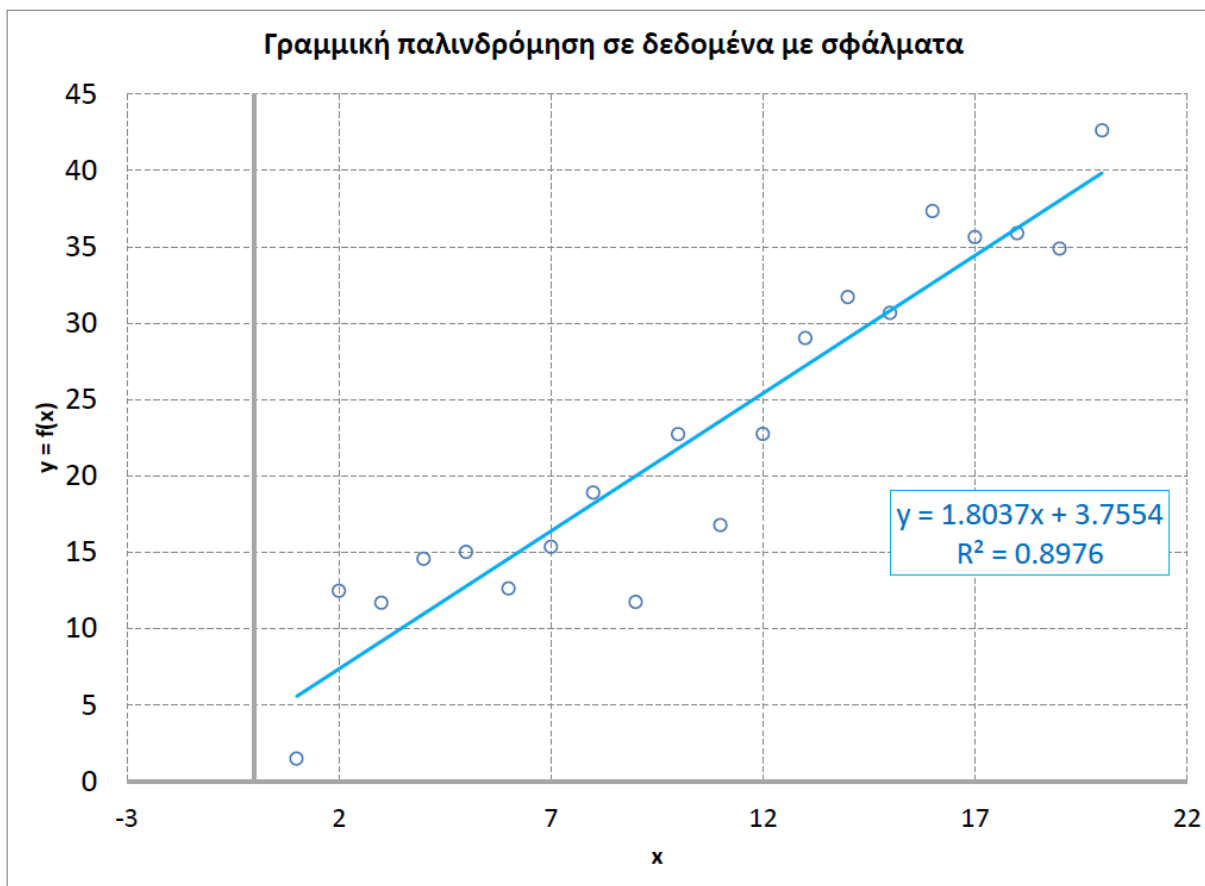
- Έστω ότι έχουμε ένα σύνολο τιμών x , y που κανονικά έχουν μεταξύ τους γραμμική συσχέτιση ($y=ax+b$).



Υπερπροσαρμογή σε συμβατικά μοντέλα

Γραμμική παλινδρόμηση

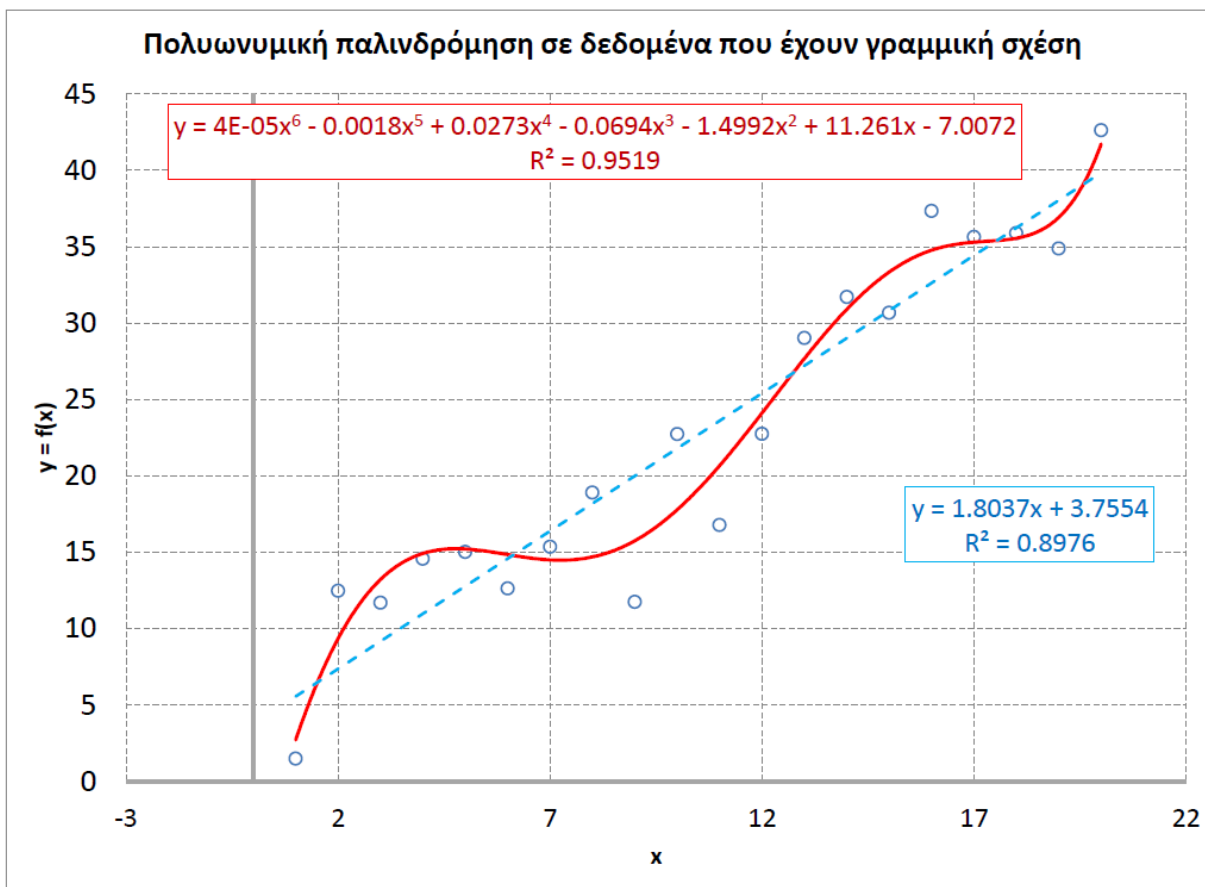
- Σε πραγματικά δεδομένα (με σφάλματα), η ευθεία γραμμή ελαχίστων τετραγώνων δεν περνάει ακριβώς από τα σημεία x, y , αλλά είναι το βέλτιστο μοντέλο για την πρόβλεψη τιμών y για τυχαία x .



Υπερπροσαρμογή σε συμβατικά μοντέλα

Πολυωνυμική παλινδρόμηση

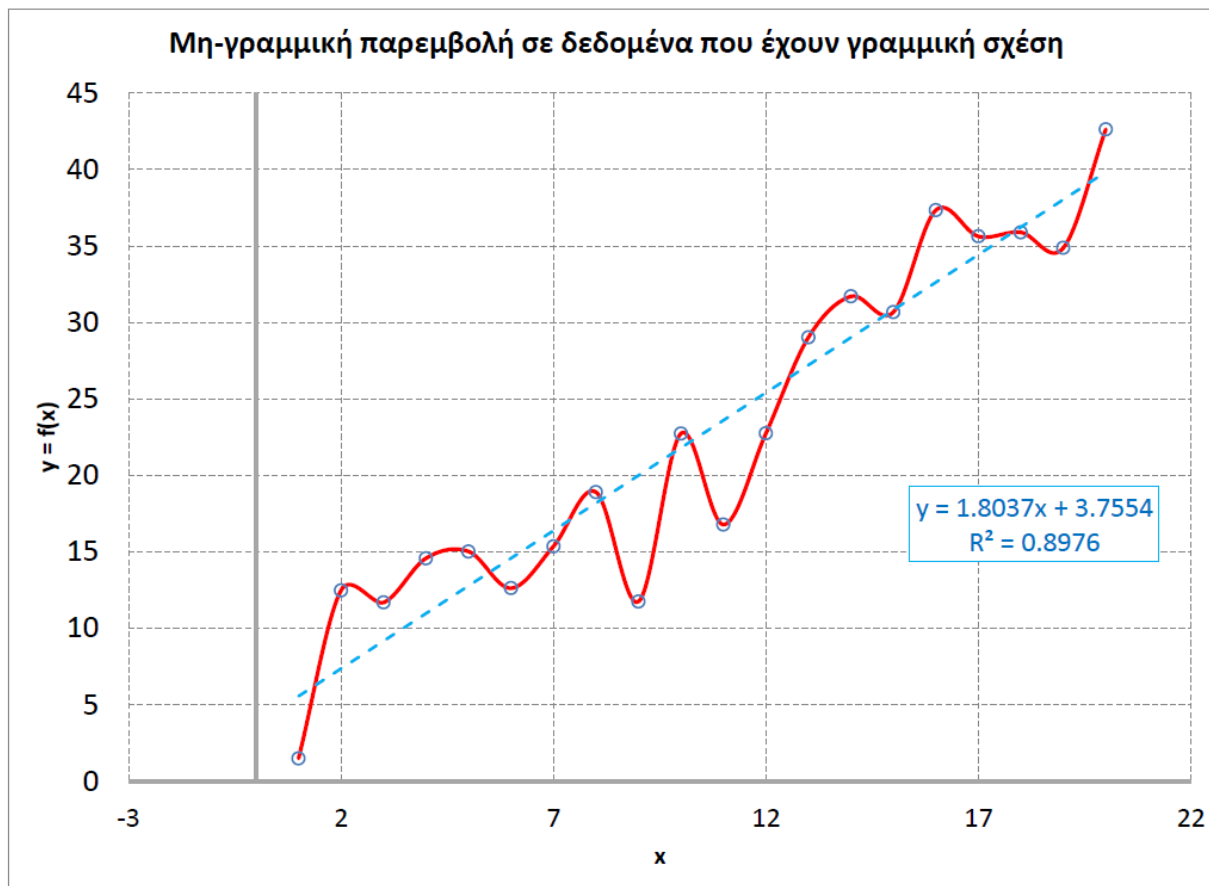
- Θα μπορούσε να βρεθεί ένα πολυωνυμικό μοντέλο που να περνάει καλύτερα από τα ίδια δεδομένα x, y .
- Ωστόσο με ένα τέτοιο μοντέλο, οι προβλέψεις των y για τυχαία x θα είχαν μεγαλύτερη απόκλιση από τις πραγματικές τιμές.



Υπερπροσαρμογή σε συμβατικά μοντέλα

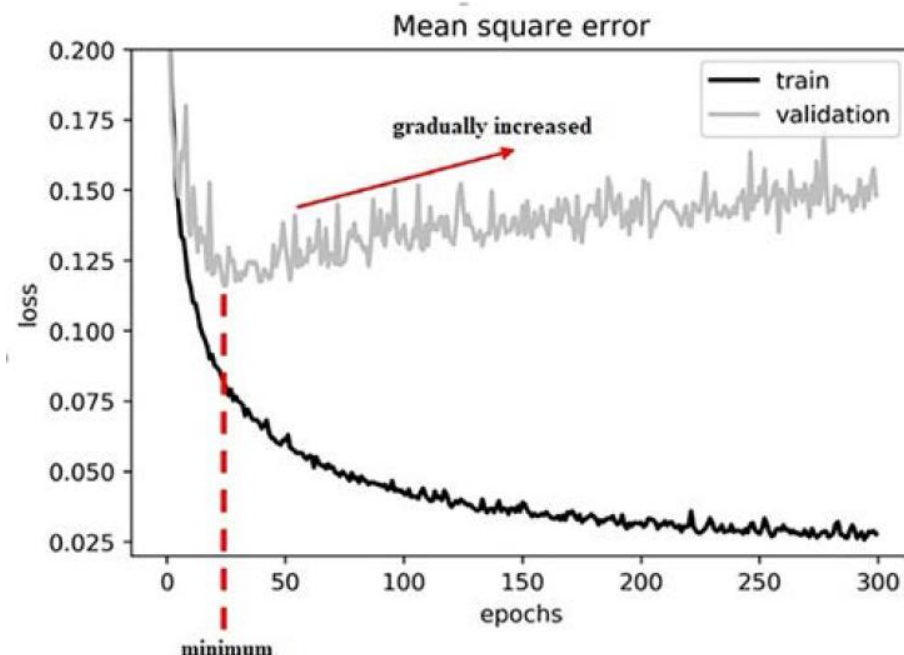
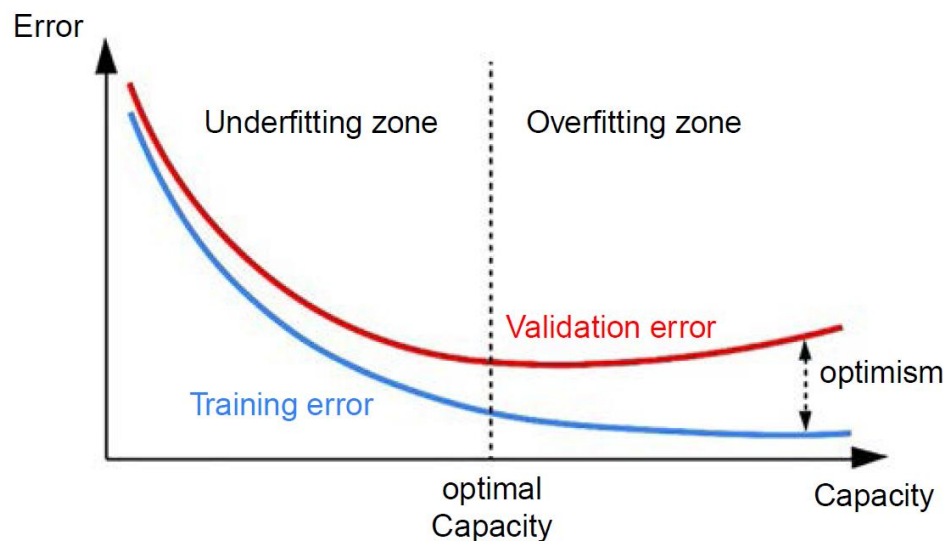
Μη-γραμμική παρεμβολή

- Ένα ακόμα πιο ακραίο παράδειγμα θα ήταν μια μη-γραμμική παρεμβολή, με τη γραμμή να διέρχεται **ακριβώς** από τα πειραματικά σημεία (τα οποία έχουν σφάλματα).
- Αυτή είναι μια εικόνα υπερπροσαρμογής στα δεδομένα εκπαίδευσης, και εμποδίζει τη γενίκευση.



Αποφυγή υπερπροσαρμογής

- Προκειμένου να αποφευχθεί η υπερπροσαρμογή, πρέπει η εκπαίδευση να σταματάει όταν πλέον παύει να μειώνεται ή αρχίζει να αυξάνεται το κόστος στα δεδομένα επικύρωσης.



Διασταυρωμένη Επικύρωση

(cross-validation)

- Η εκπαίδευση των ΤΝΔ βαθιάς μάθησης με επίβλεψη απαιτεί ένα μεγάλο σύνολο επισημασμένων δεδομένων που πολλές φορές είναι δύσκολο να βρεθεί.
- Επιπλέον, τα διαθέσιμα επισημασμένα δεδομένα πρέπει να διαχωριστούν σε:
 - Σετ εκπαίδευσης (training set)
 - Σετ επικύρωσης (validation set)
 - Σετ δοκιμής (testing set)
- Ένας τρόπος να αυξήσουμε τα δεδομένα εκπαίδευσης χωρίς τον κίνδυνο υπερπροσαρμογής, είναι να εφαρμόσουμε **διασταυρωμένη επικύρωση** (cross-validation).
- Με τη διασταυρωμένη επικύρωση, το μοντέλο επανεκπαιδεύεται κάθε φορά σε ένα διαφορετικό τμήμα των διαθέσιμων δεδομένων, ενώ ένα άλλο τμήμα παίζει τον ρόλο του σετ επικύρωσης και ένα τρίτο (γενικά αχρησιμοποίητο) τον ρόλο του σετ δοκιμής.
- Η διασταυρωμένη επικύρωση βοηθάει και στον περιορισμό της υπερπροσαρμογής αλλά και στη γενίκευση του μοντέλου βαθιάς μάθησης.

Διασταυρωμένη Επικύρωση

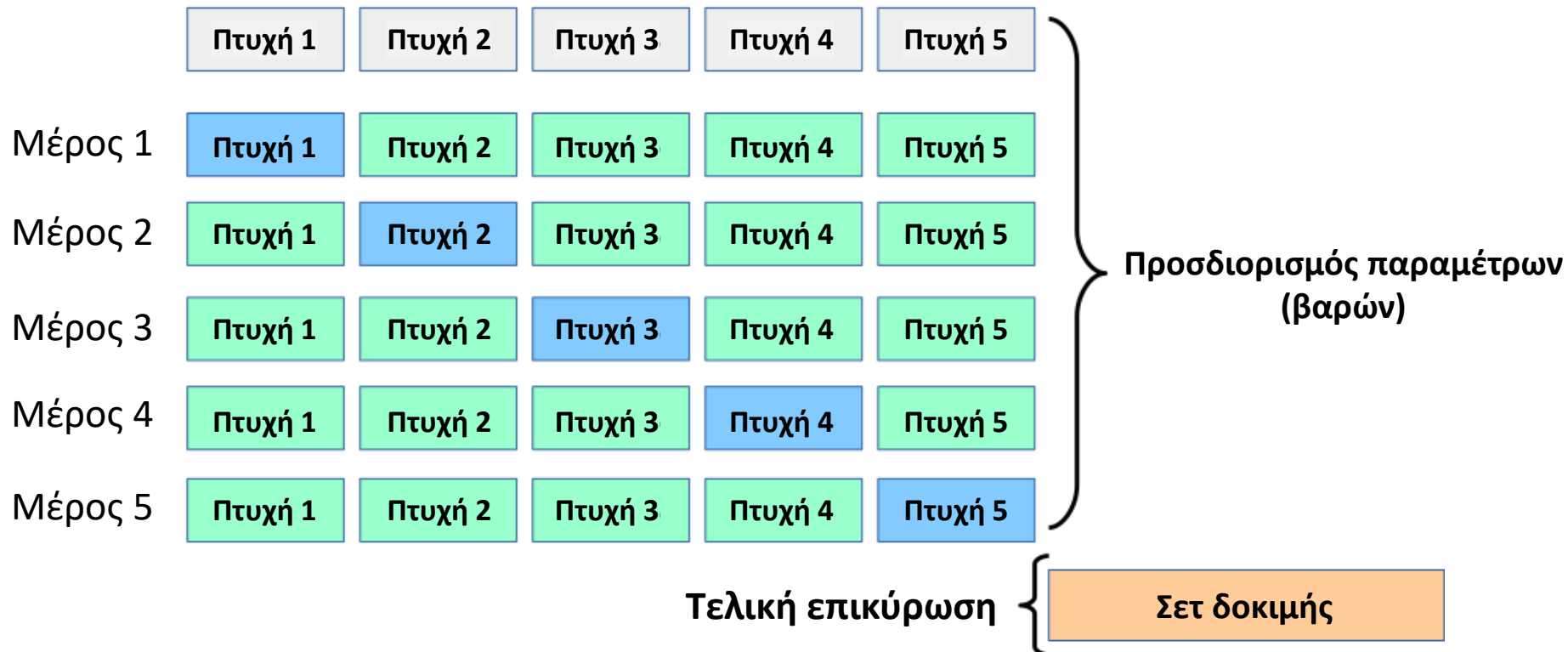
(cross-validation)

Σύνολο διαθέσιμων επισημασμένων δεδομένων (labeled data)

Δεδομένα εκπαίδευσης και διασταυρωμένης επικύρωσης

Σετ δοκιμής

- Διαχωρισμός, π.χ., σε 5 πτυχές
- Κάθε φορά, **τέσσερις** πτυχές χρησιμοποιούνται **για εκπαίδευση** και η **μία για επικύρωση**

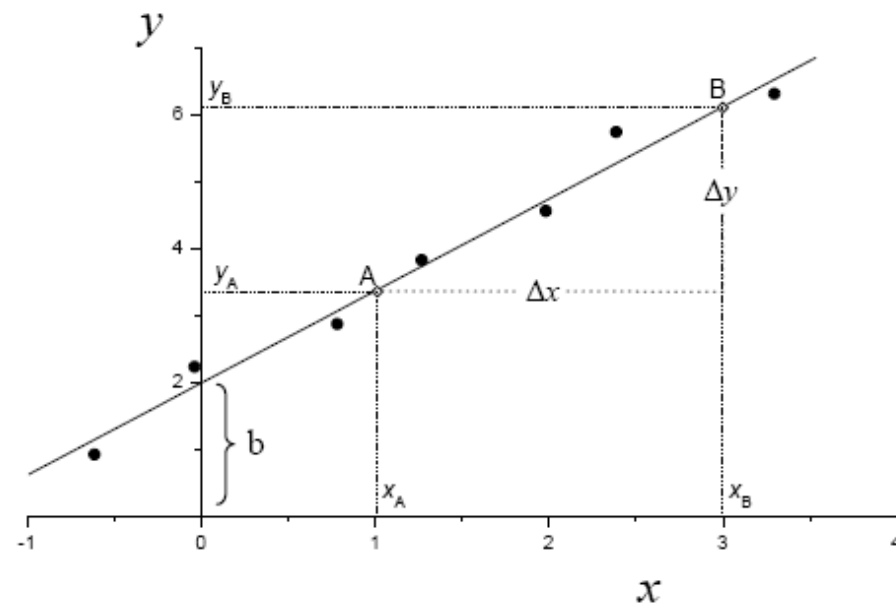


Σεισμολογικές εφαρμογές

Απλό παράδειγμα παλινδρόμησης (regression)

Γραμμική παλινδρόμηση:

- Διαθέτουμε δεδομένα x και y τα οποία θεωρούμε ότι έχουν γραμμική συσχέτιση μεταξύ τους.
- Με τη μέθοδο ελαχίστων τετραγώνων βρίσκουμε τις σταθερές k , b που ικανοποιούν τη σχέση $y(x)=kx+b$ για μια ευθεία η οποία ελαχιστοποιεί τα τετράγωνα των υπολοίπων ως προς τα πειραματικά σημεία.
- Στη συνέχεια, μπορούμε για ένα **τυχαίο “ x ”** να βρούμε την **προβλεπόμενη** τιμή **“ y ”** βάσει της σχέσης $y=kx+b$



$$k = \frac{\sum_{i=1}^n x_i y_i - n\bar{x} \cdot \bar{y}}{\sum_{i=1}^n (x_i^2) - n\bar{x}^2}$$

$$b = \frac{\bar{y} \cdot \sum_{i=1}^n x_i^2 - \bar{x} \cdot \sum_{i=1}^n x_i y_i}{\sum_{i=1}^n (x_i^2) - n\bar{x}^2}$$

Παλινδρόμηση με μηχανική μάθηση

Παράδειγμα: Σχέσεις απόσβεσης

(Ground Motion Prediction Equations, GMPE)

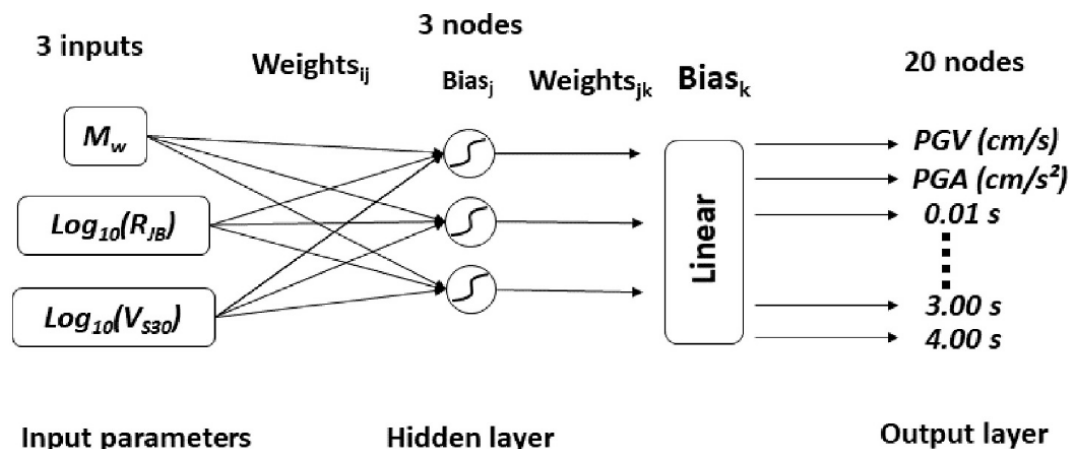
- Ένα σύνθετο πρόβλημα παλινδρόμησης είναι αυτό της εκτίμησης κάποιων μέτρων έντασης (PGA, PGV, PGD, κ.α.) συναρτήσει των παρακάτω μεταβλητών:
 - Μέγεθος σεισμού
 - Επικεντρική απόσταση
 - Εστιακό βάθος
 - Εδαφικές συνθήκες (π.χ., V_{s30})
 - Μηχανισμός γένεσης
- Μπορεί να εκπαιδευτεί ένα μοντέλο μηχανικής μάθησης με δεδομένα σεισμών που έχουν δώσει γνωστές (μετρούμενες) τιμές PGA, PGV ή PGD, κ.α., για μια περιοχή μελέτης με γνωστές (δεδομένες) εδαφικές συνθήκες.
- Στη συνέχεια, μπορούμε να δώσουμε στο μοντέλο τυχαίες μεταβλητές (μέγεθος, επικεντρική απόσταση, κτλ), με τις οποίες δεν έχει εκπαιδευτεί, και το μοντέλο να προβλέψει τις τιμές PGA, PGV, PGD, κ.α.
- Οι σχέσεις απόσβεσης (GMPE) είναι σημαντικές για μελέτες σεισμικής επικινδυνότητας.

Παλινδρόμηση με μηχανική μάθηση

Παράδειγμα: Σχέσεις απόσβεσης

(Ground Motion Prediction Equations, GMPE)

- Μια πρόσφατη εφαρμογή μοντέλου μηχανικής μάθησης για σχέσεις απόσβεσης στον Ελλαδικό χώρο:
- 3 πηγαίοι κόμβοι (είσοδοι):
 - Μέγεθος M_w
 - Απόσταση Joyner-Boore (km)
 - Ταχύτητα V_{s30} (m/s)
- 1 κρυμμένο επίπεδο 3 κόμβων.
- 20 κόμβοι εξόδου:
 - PGV (cm/s)
 - PGA (cm/s²)
 - 0.01 s
 - ...
 - 3.00 s
 - 4.00 s

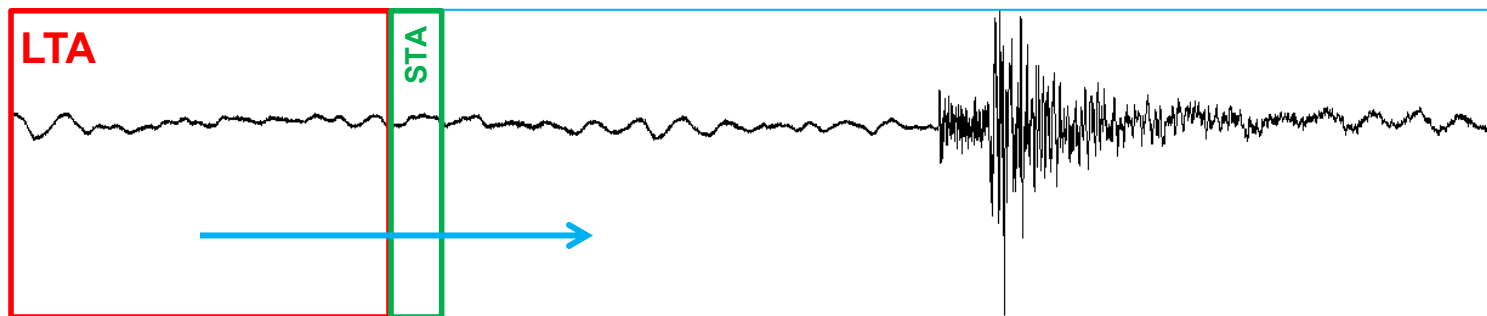


Αναγνώριση άφιξης σεισμικών κυμάτων (phase picking)

Συμβατική μέθοδος STA/LTA:

STA (Short Time Average): μέση τιμή πλάτους σε χρονικό παράθυρο μικρού μήκους

LTA (Long Time Average): μέση τιμή πλάτους σε χρονικό παράθυρο μεγάλου μήκους



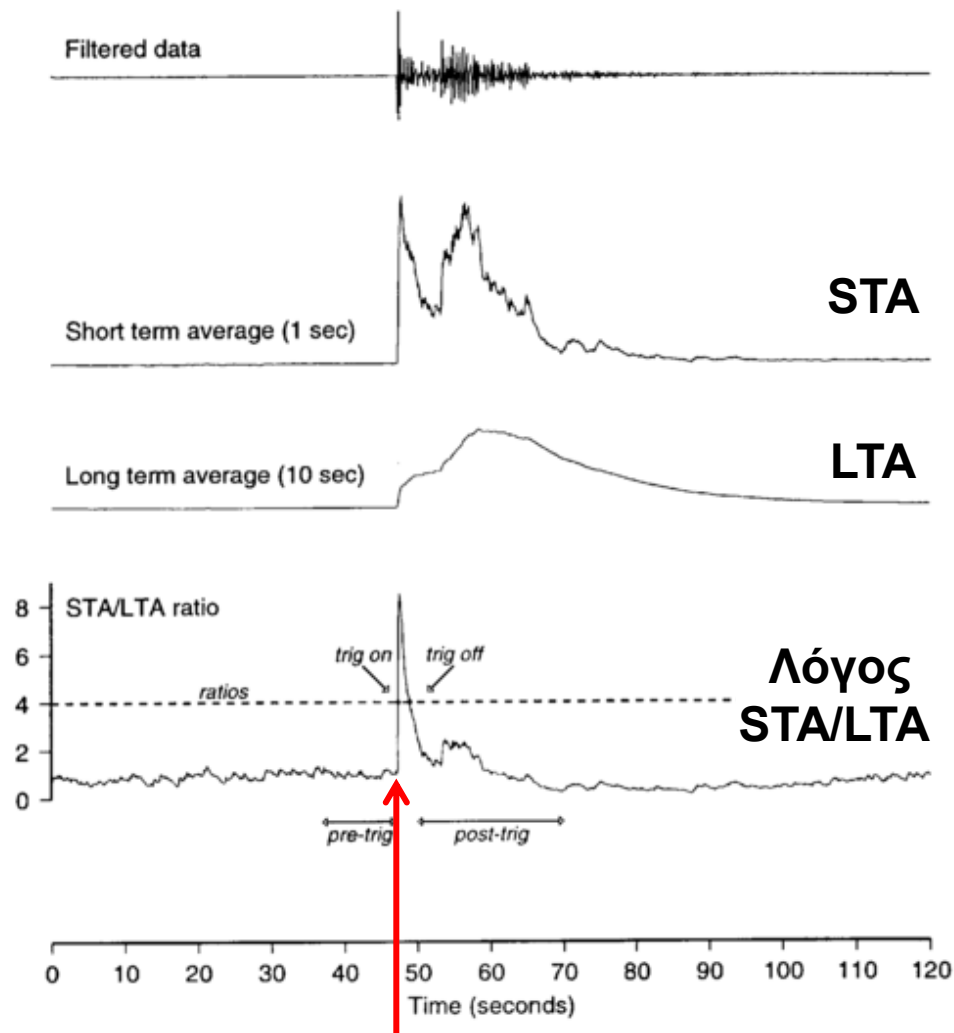
Λόγος STA/LTA: μας δίνει μια εκτίμηση του λόγου σήματος-προς-θόρυβο (Signal-To-Noise Ratio ή **SNR**)

- Η μέτρηση του λόγου STA/LTA γίνεται σε **κυλιόμενα χρονικά παράθυρα** πάνω στην κυματομορφή.
- Το παράθυρο **LTA**, επειδή είναι μεγάλου μήκους ανταποκρίνεται **αργά** όταν σε αυτό μπει ένα σεισμικό σήμα (αφού περιλαμβάνει κυρίως θόρυβο υποβάθρου, χαμηλού πλάτους).
- Αντιθέτως, το παράθυρο **STA**, που τοποθετείται στο δεξί άκρο του παραθύρου LTA, καθώς είναι μικρότερου μήκους και προπορεύεται του LTA, επηρεάζεται **άμεσα** και έντονα όταν σε αυτό εισέλθει το σεισμικό σήμα.

Αναγνώριση άφιξης σεισμικών κυμάτων (phase picking)

Συμβατική μέθοδος STA/LTA:

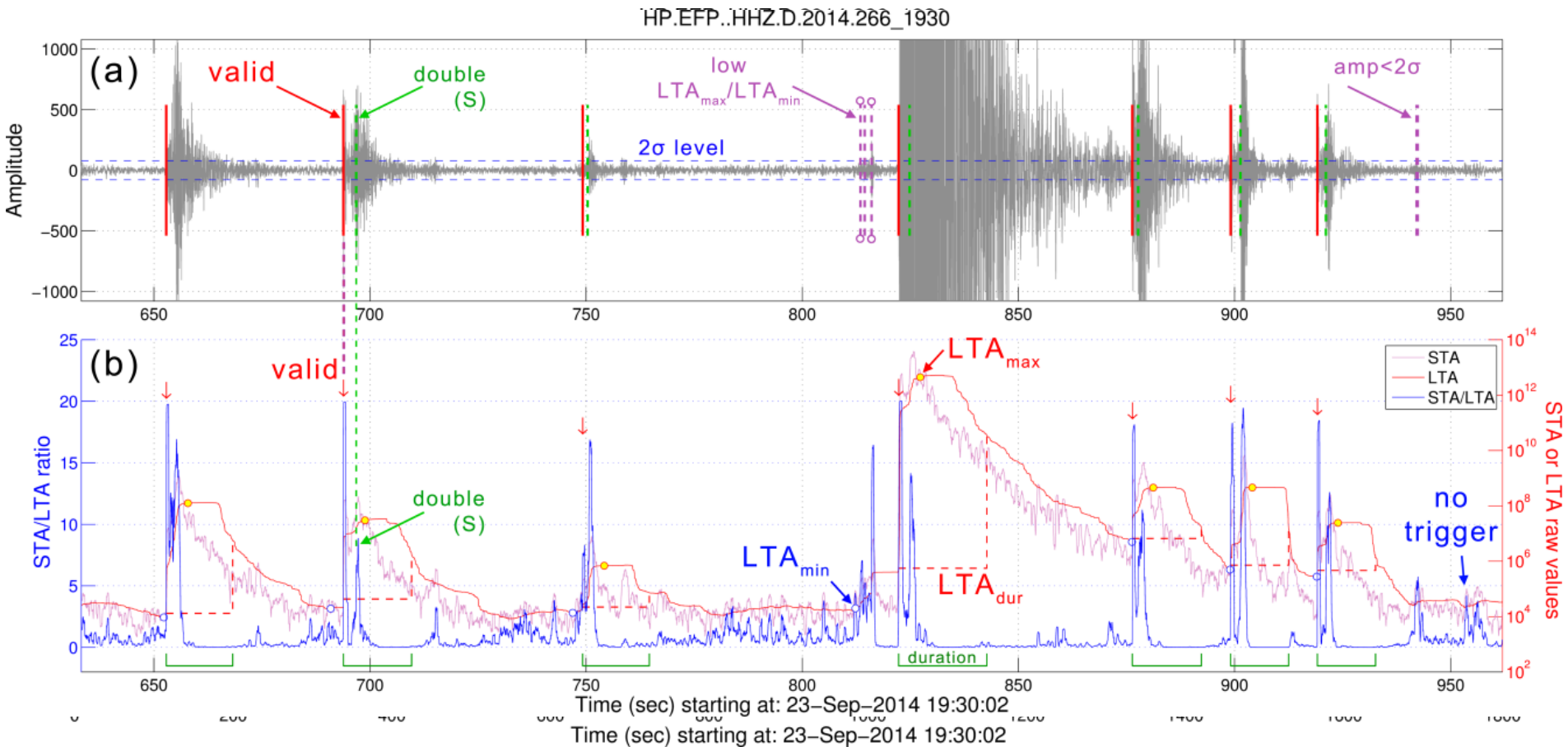
- Όσο τα παράθυρα STA και LTA περιέχουν και τα δύο θόρυβο, ο λόγος **STA/LTA** θα είναι κοντά στο 1.
- Όταν στο παράθυρο STA μπει ένα σεισμικό σήμα, ο λόγος STA/LTA θα αυξηθεί απότομα, με τη μέγιστη τιμή να αντιπροσωπεύει το **SNR** του σήματος (πόσο ισχυρότερο είναι το σήμα από το επίπεδο του θορύβου).
- Η χρονική στιγμή κατά την οποία ο λόγος STA/LTA ξεπερνάει μια ορισμένη τιμή κατωφλίου (π.χ., 4) θεωρείται ως ο χρόνος άφιξης του σεισμικού κύματος (**P**).
- Στα κύματα S ο λόγος STA/LTA έχει συνήθως μικρότερη τιμή, λόγω παρουσίας του κυματοσυρμού των P.



Εντοπισμός σήματος όταν ο λόγος STA/LTA ξεπεράσει μια ορισμένη τιμή κατωφλίου (trigger on)

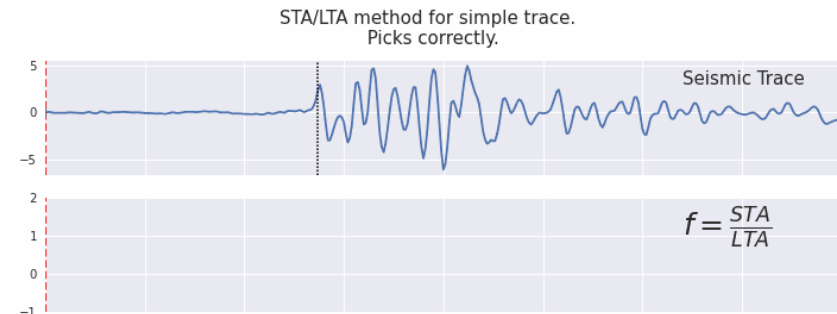
Αναγνώριση άφιξης σεισμικών κυμάτων (phase picking)

Συμβατική μέθοδος STA/LTA: Αναγνώριση σημάτων



Αναγνώριση άφιξης σεισμικών κυμάτων (phase picking)

Συμβατική μέθοδος STA/LTA:



Προτερήματα μεθόδου STA/LTA:

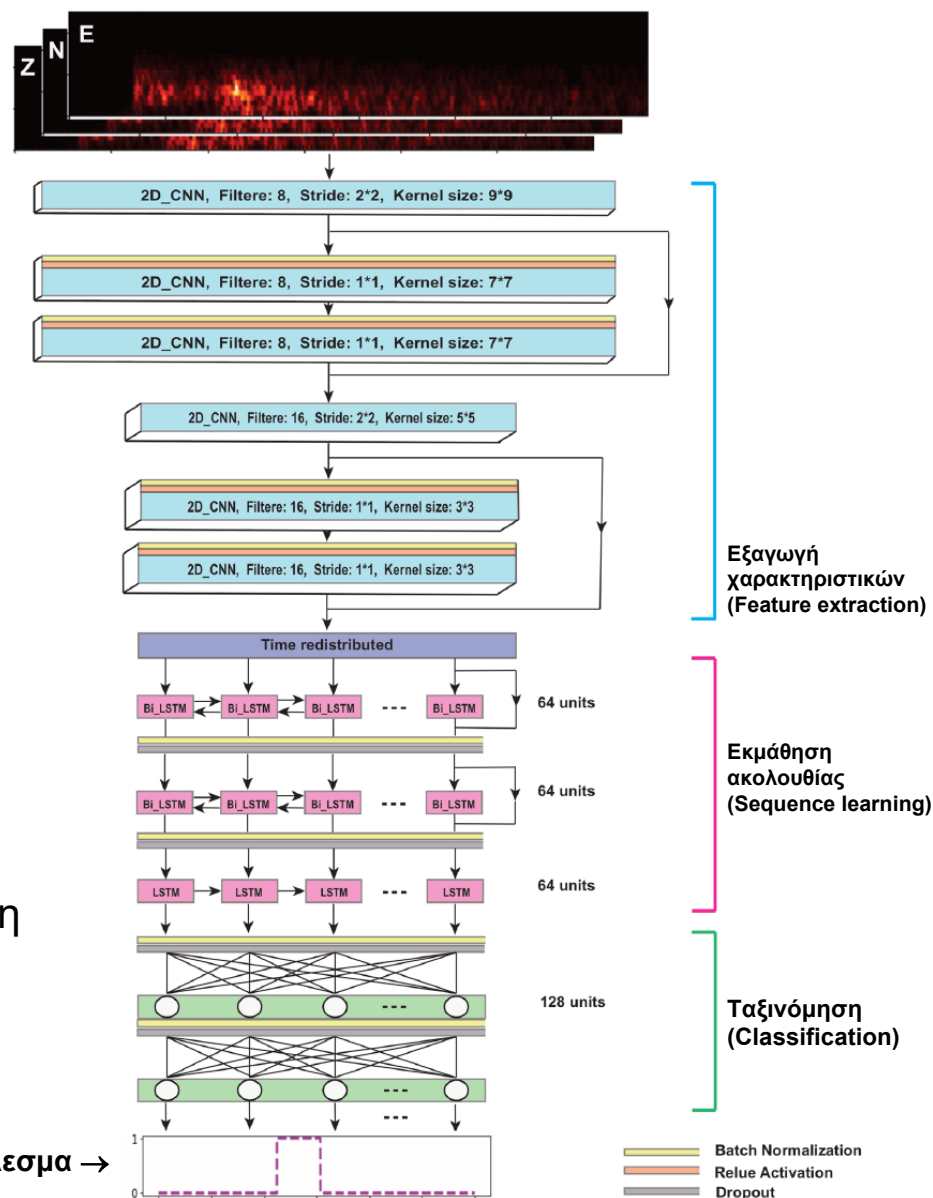
- Εύκολη στον προγραμματισμό.
- Χρήσιμη για γρήγορη αναγνώριση σημάτων και για σύγκριση με άλλες μεθόδους.

Προβλήματα μεθόδου STA/LTA:

- Δεν ενδείκνυται για προσδιορισμό της άφιξης των κυμάτων S.
- Για σήματα σεισμών σε περιφερειακές αποστάσεις (μεγάλη διαφορά $T_s - T_p$) μπορεί να εντοπίσει τα κύματα S σαν να ήταν κύματα P κάποιου «δεύτερου» σεισμού.
- Το ίδιο μπορεί να συμβεί αν τα κύματα P είναι πολύ ασθενή και διακρίνονται μόνο τα κύματα S.
- Η δυνατότητά της να ανιχνεύσει σεισμούς περιορίζεται σε σήματα με υψηλό λόγο σήματος-προς-θόρυβο (SNR).
- Η ακρίβειά της περιορίζεται από τις διαστάσεις του παραθύρου STA και τον τύπο του πρώτου παλμού (κρουστικός ή αναδυσόμενος).
- Χρειάζεται εφαρμογή φίλτρου (π.χ. Butterworth) προκειμένου να βελτιωθεί ο λόγος SNR.
- Δυσκολεύεται να ανιχνεύσει σήματα μικρών σεισμών που βρίσκονται στην «ουρά» ενός μεγαλύτερου, ή σε αλληλοεπικαλυπτόμενα σήματα σεισμών (π.χ. σε σεισμικές ακολουθίες).
- Η σταθερή τιμή κατωφλίου δεν προσαρμόζεται στις συνθήκες θορύβου, οι οποίες μπορεί να μεταβάλλονται.
- Δεν διακρίνει σεισμούς από άλλα είδη παροδικών σημάτων (π.χ., αυτοκίνητα).

Αναγνώριση σεισμικών γεγονότων: CRED (event detection)

- Το μοντέλο βαθιάς μάθησης CRED (Cnn-Rnn Earthquake Detector) είναι φτιαγμένο για να αναγνωρίζει σεισμικά γεγονότα.
- Είναι εκπαιδευμένο με ένα σύνολο 500,000 επισημασμένων σειсмоγραμμάτων:
 - 250,000 επιβεβαιωμένοι τεκτονικοί **σεισμοί** από τη Βόρεια Καλιφόρνια
 - 250,000 επιβεβαιωμένοι **θόρυβοι**
- Οι συγγραφείς το δοκίμασαν σε συνεχείς σεισμικές καταγραφές από το Κεντρικό Αρκάνσας, διάρκειας ενός μήνα:
 - Το μοντέλο εντόπισε πάνω από 800 επαγόμενους μικροσεισμούς με μεγέθη M_L έως και -1.3.
- Το μοντέλο CRED συνδυάζει συνελκτικά (CNN) και επαναληπτικά (RNN) ΤΝΔ.



Αποτέλεσμα →

Αναγνώριση σεισμικών γεγονότων: CRED

(event detection)

- Παραδείγματα επισημασμένων δεδομένων για την εκπαίδευση του μοντέλου CRED.

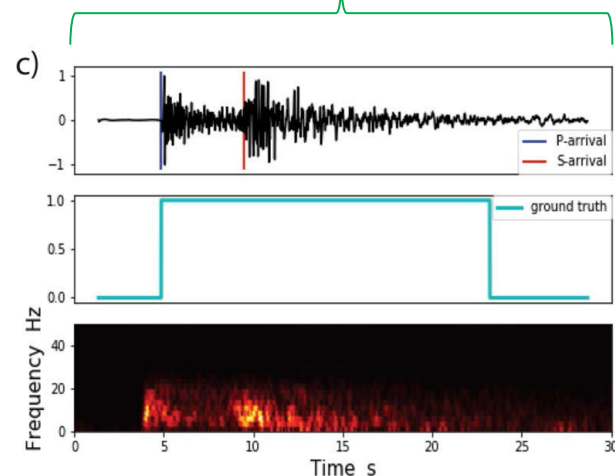
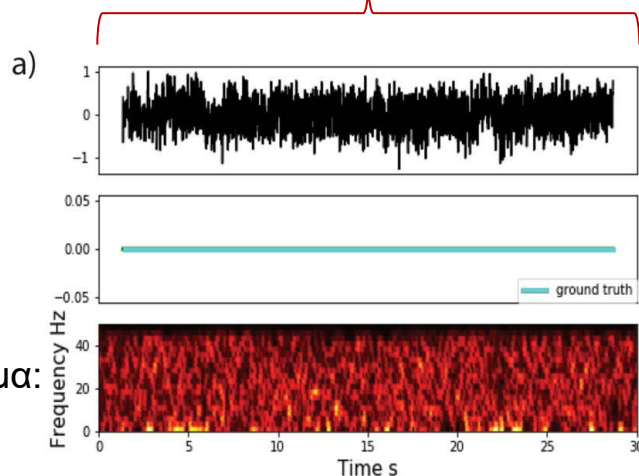
Θόρυβοι

Σεισμοί

Χρονοσειρά:

“ΕΤΙΚΕΤΑ”:
(ground truth)

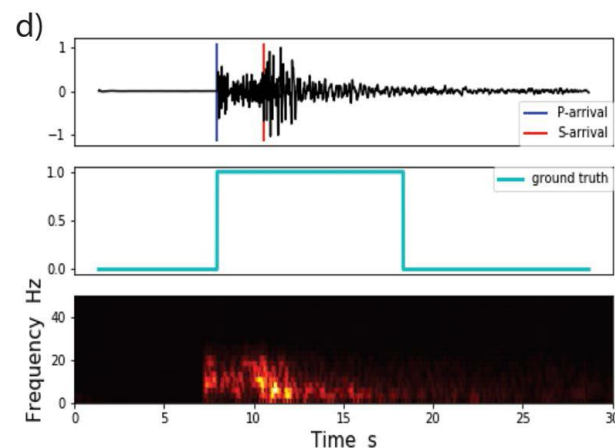
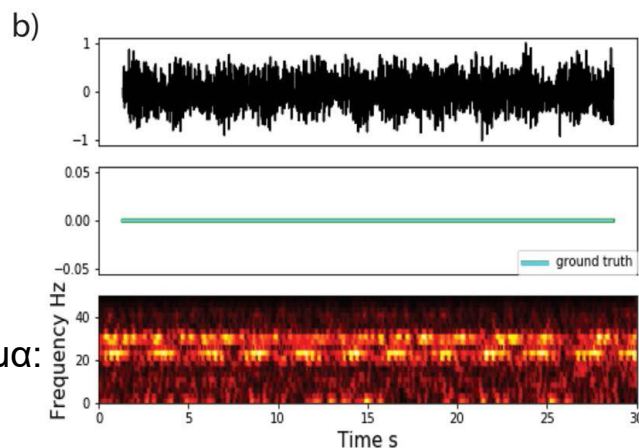
Φασματογράφημα:



Χρονοσειρά:

“ΕΤΙΚΕΤΑ”:
(ground truth)

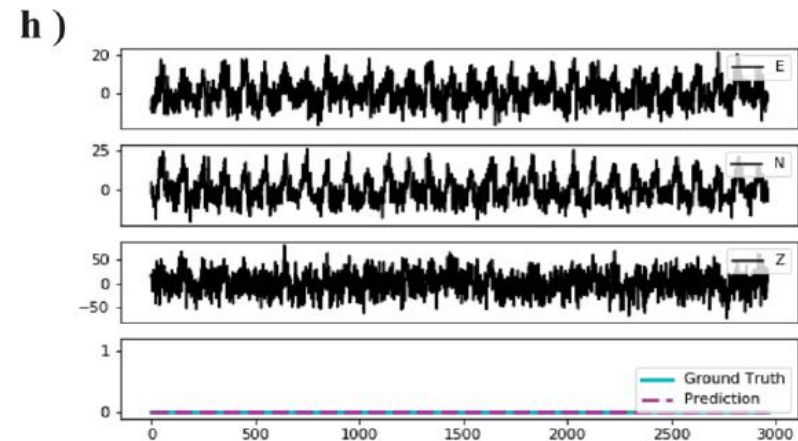
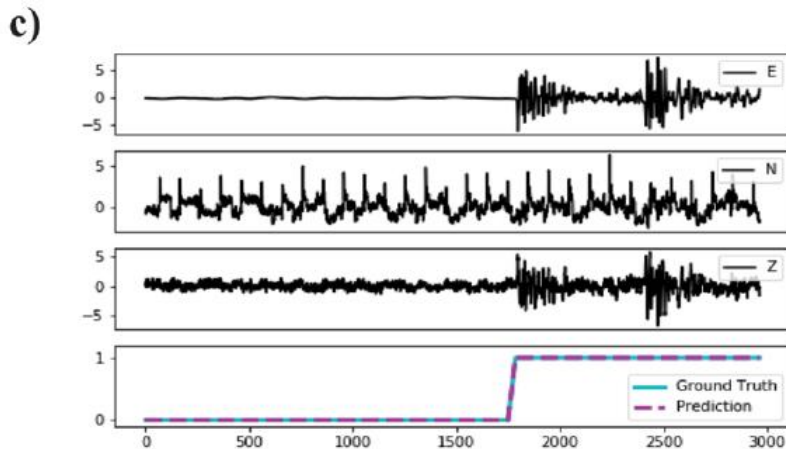
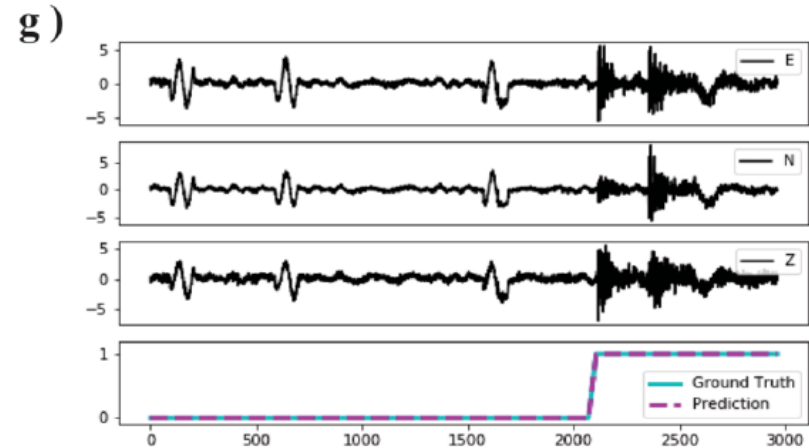
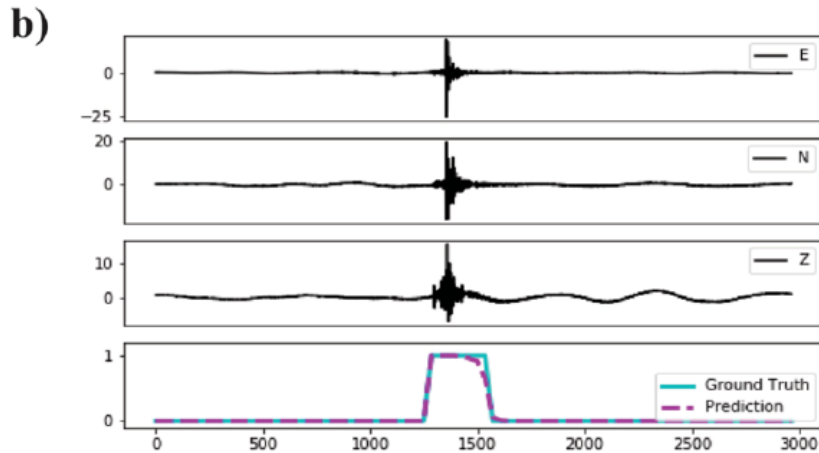
Φασματογράφημα:



Αναγνώριση σεισμικών γεγονότων: CRED

(event detection)

- Παραδείγματα εφαρμογής μοντέλου CRED.



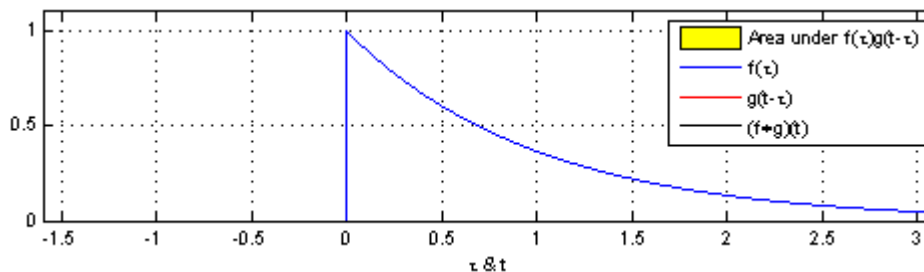
Συνελικτικά Νευρωνικά Δίκτυα

(Convolutional Neural Networks, CNN)

- Όταν τα δεδομένα έχουν χωρική συσχέτιση μεταξύ τους, όπως π.χ. εικόνες (με 1 ή περισσότερα «χρώματα») ή άλλες 2Δ απεικονίσεις (π.χ. φασματογραφήματα), για την αναγνώριση μοτίβων χρησιμοποιούνται συνήθως κάποια **επίπεδα συνέλιξης**.
- Τα ΤΝΔ που χρησιμοποιούν επίπεδα συνέλιξης ονομάζονται **Συνελικτικά Νευρωνικά Δίκτυα** (Convolutional Neural Networks, ή CNN).
- Στα δίκτυα αυτά εφαρμόζεται η πράξη της **συνέλιξης** (convolution) ανάμεσα στα επίπεδα εισόδου και σε κάποια μικρά μπλοκ (πίνακες) που ονομάζονται «πυρήνες» (kernels), με σκοπό την ανίχνευση διαφόρων μοτίβων μέσα στην απεικόνιση που δίνεται στο προηγούμενο επίπεδο (π.χ., το επίπεδο εισόδου)

Παράδειγμα: Συνέλιξη * μεταξύ χρονοσειρών, στο πεδίο του χρόνου:

$$(f * g)(t) := \int_{-\infty}^{\infty} f(\tau)g(t - \tau) d\tau$$

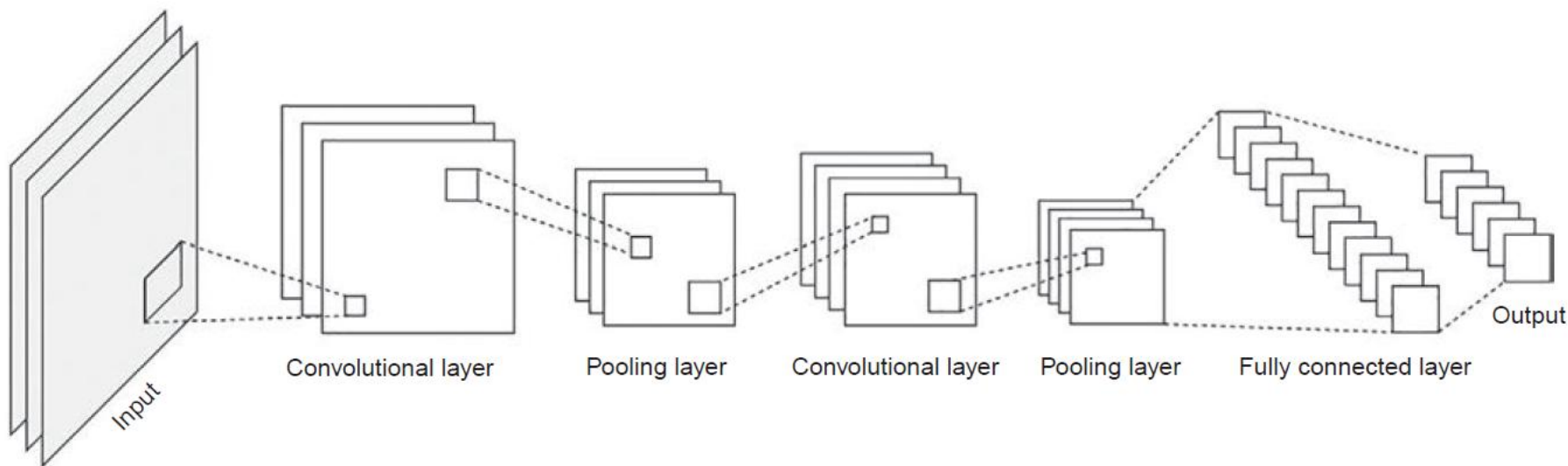


Η συνέλιξη ισοδυναμεί με πολλαπλασιασμό στο πεδίο συχνοτήτων (π.χ., φίλτρα).

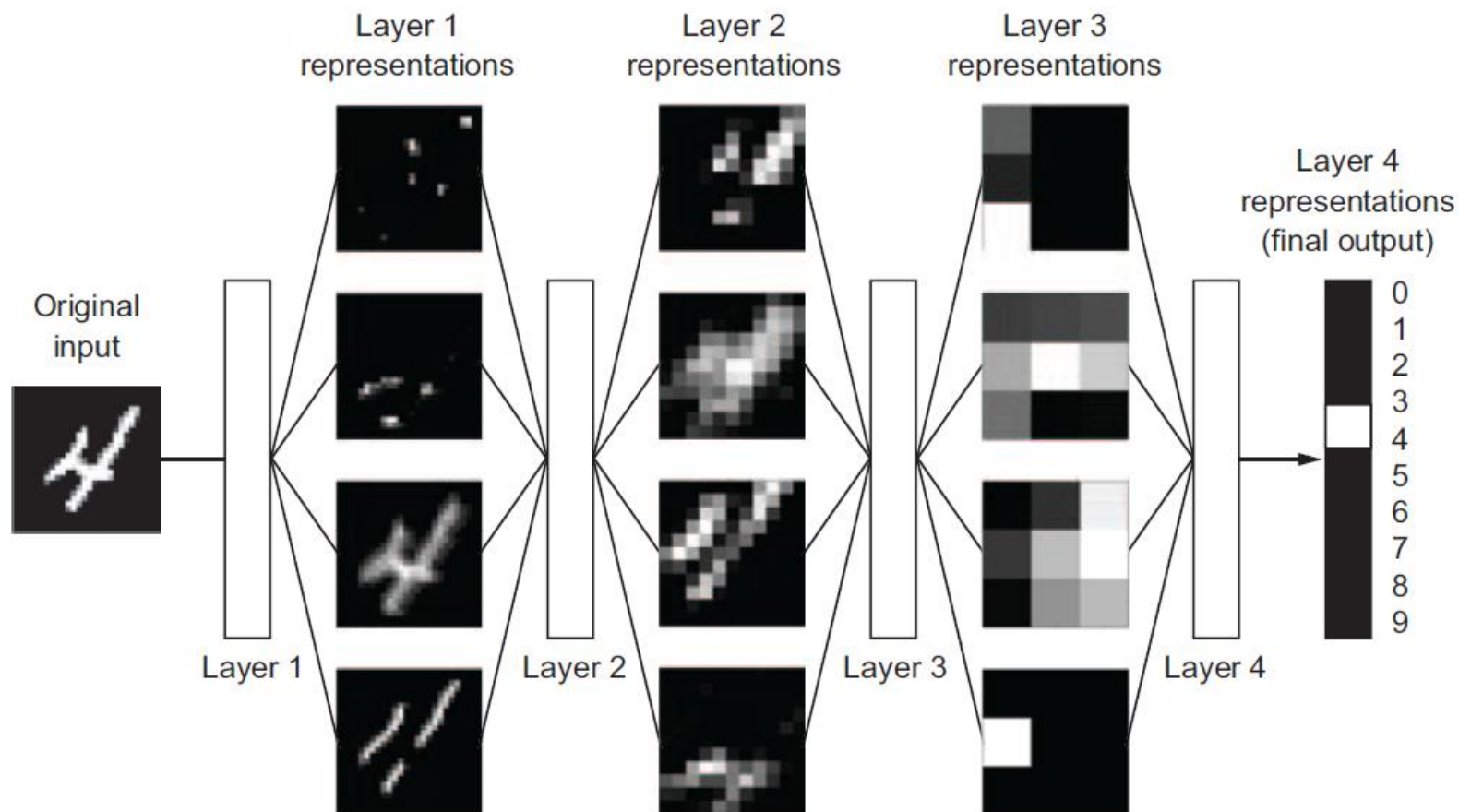
Συνελικτικά Νευρωνικά Δίκτυα

(Convolutional Neural Networks, CNN)

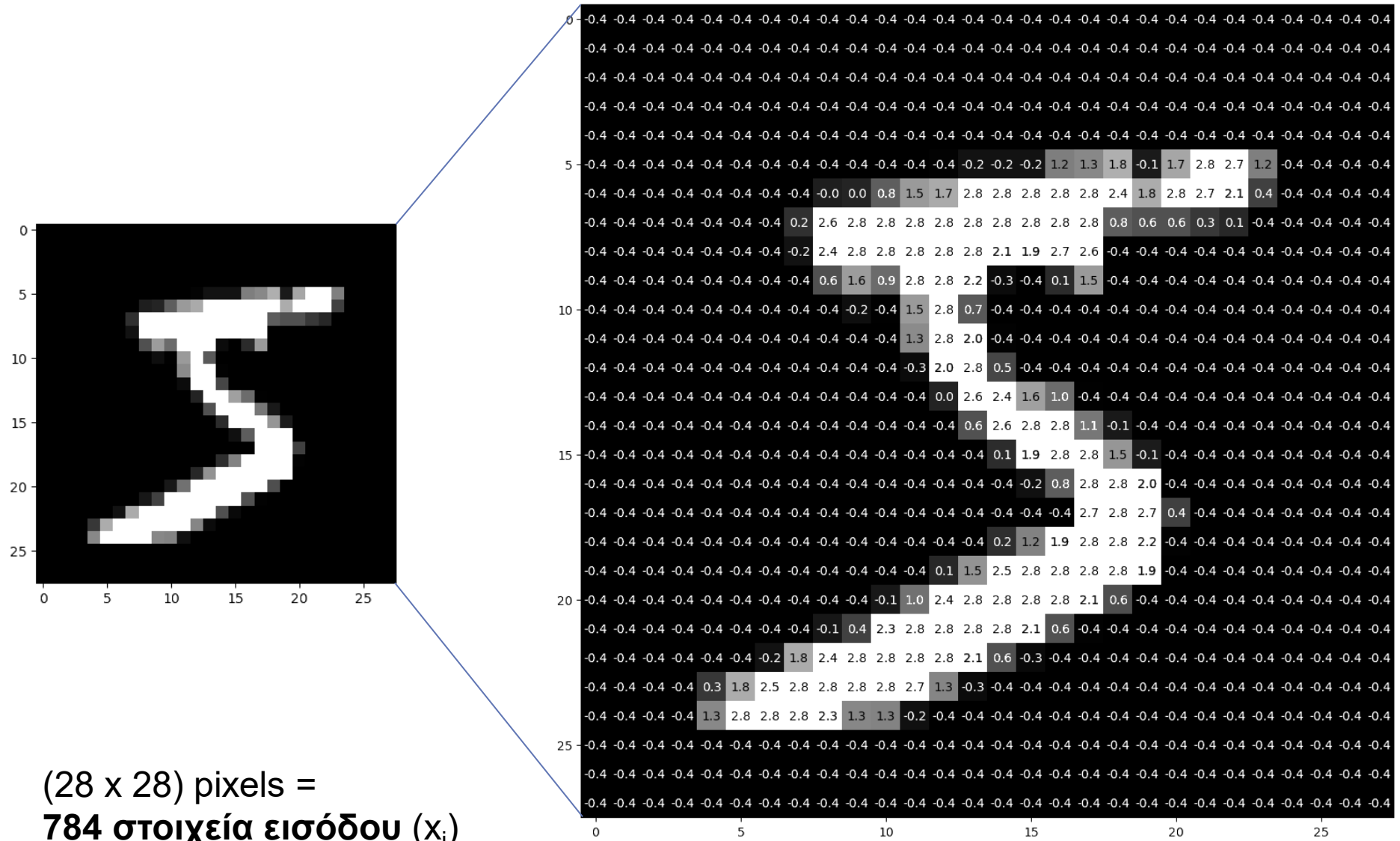
- Τα επίπεδα συνέλιξης δημιουργούν αναπαραστάσεις (representations) ανάμεσα στα αρχικά επίπεδα (π.χ. εικόνες με 3 επίπεδα RGB, ή φασματογραφήματα των τριών συνιστωσών) και τους πυρήνες (kernels), οι οποίοι περιέχουν απλά μοτίβα.
- Παραδείγματα μοτίβων: αναγνώριση αιχμών (edge detection), ή παρουσία κάποιου «αντικειμένου» μέσα στην εικόνα.
- Κάθε νευρώνας του συνελικτικού επιπέδου συνδέεται μόνο με νευρώνες μιας **περιορισμένης περιοχής** του επιπέδου εισόδου.
- Στη συνέχεια ακολουθεί ένα επίπεδο συλλογής (pooling layer) το οποίο «συμπιέζει» την αρχική εικόνα σε μια μικρότερη, η οποία περιέχει τα χαρακτηριστικά που ανιχνεύθηκαν.
- Η διαδικασία επαναλαμβάνεται με εναλλαγές επιπέδων συνέλιξης και συλλογής.
- Στο τέλος, για την τελική **ταξινόμηση**, ακολουθεί ένα **πλήρως διασυνδεδεμένο** επίπεδο και στη συνέχεια το **επίπεδο εξόδου** με τις κλάσεις.



Αναπαράσταση στα επίπεδα ενός ΤΝΔ

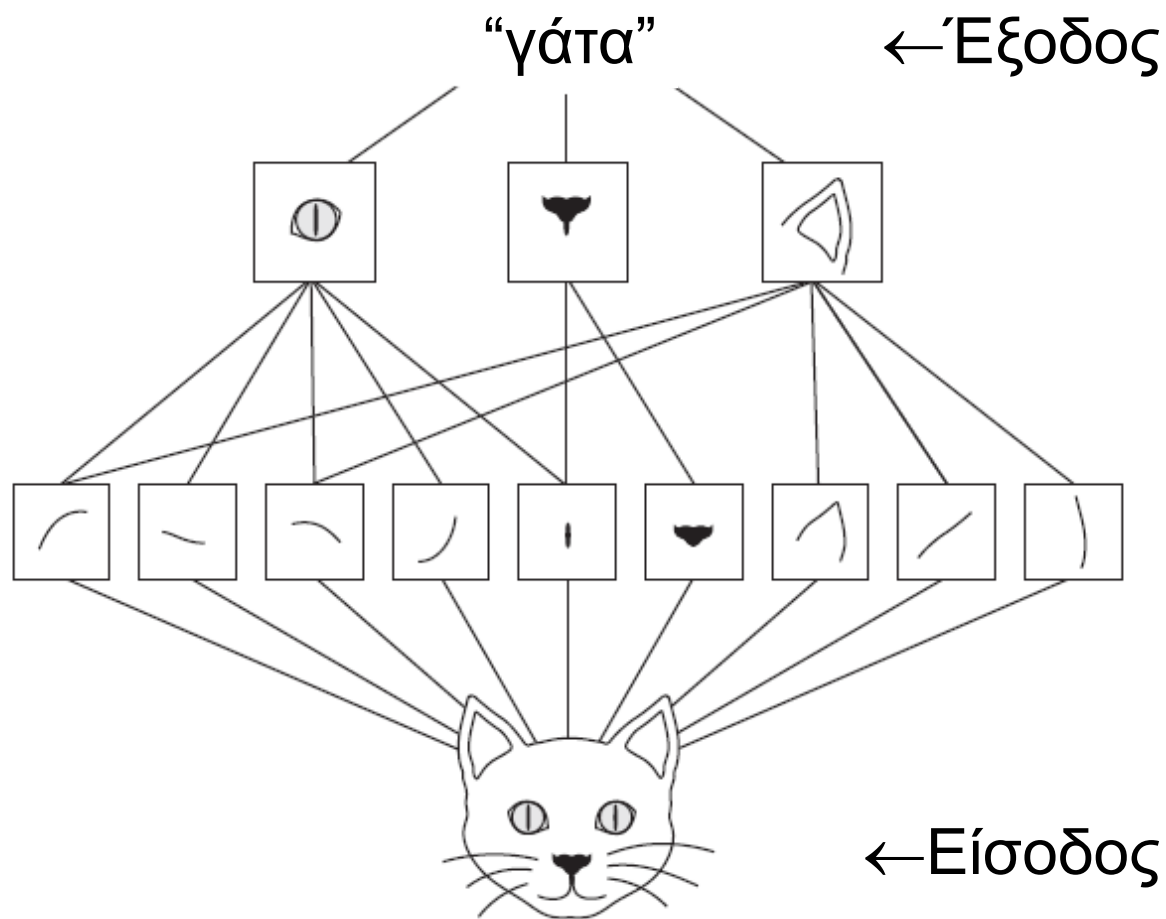
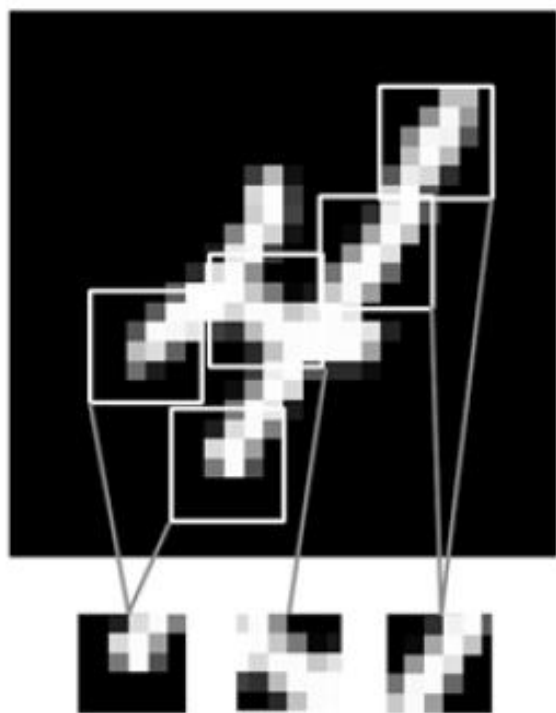


Δεδομένα εισόδου



(28 x 28) pixels =
784 στοιχεία εισόδου (x_i)

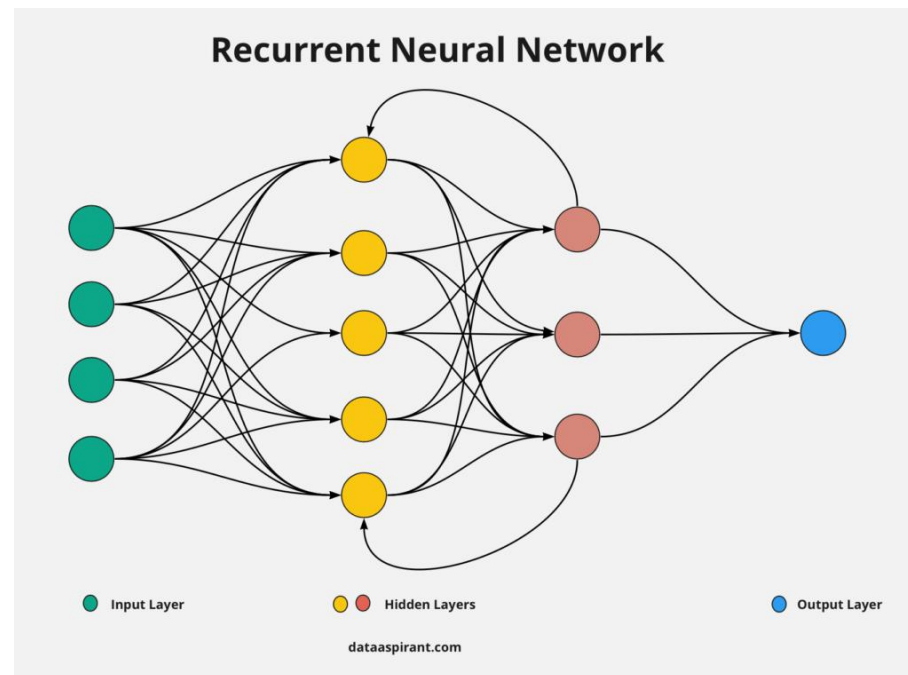
Αναπαράσταση στα επίπεδα ενός συνελικτικού ΤΝΔ



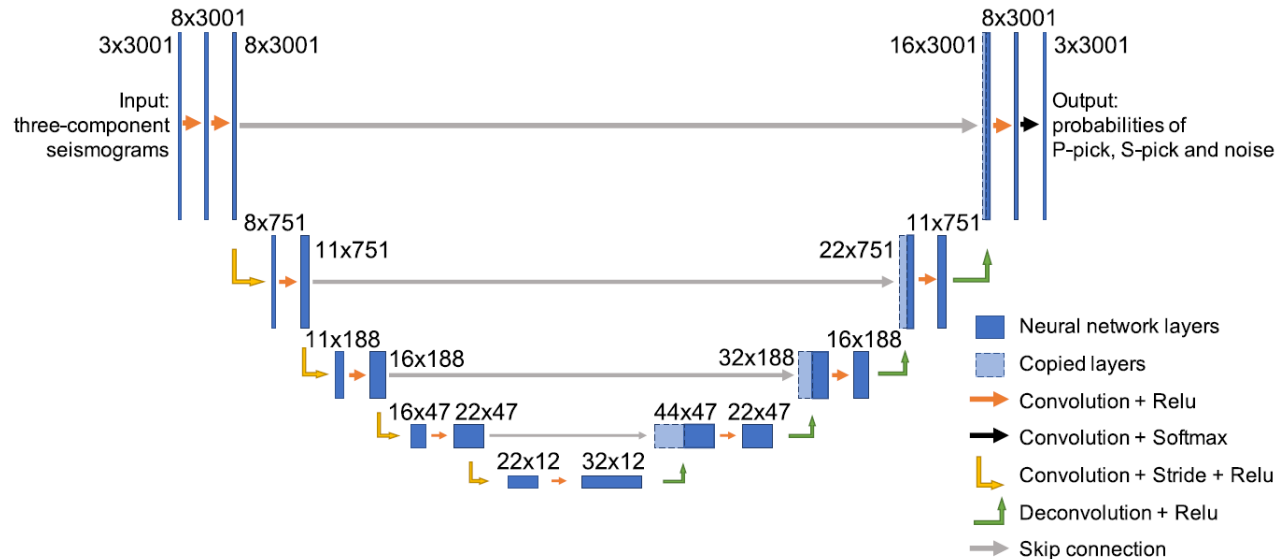
Επαναληπτικά Νευρωνικά Δίκτυα

(Recurrent Neural Networks, RNN)

- Σε αντίθεση με τα ΤΝΔ εμπρόσθιας τροφοδότησης (feed-forward), στα επαναληπτικά νευρωνικά δίκτυα (Recurrent Neural Networks, RNN) η έξοδος ορισμένων νευρώνων καταλήγει στην είσοδο νευρώνων προηγούμενου επιπέδου ή ακόμα και στον ίδιο τον νευρώνα που έδωσε την έξοδο.
- Τα RNN χρησιμοποιούνται για δεδομένα τα οποία αποτελούν ακολουθίες με μια χρονική αλληλουχία, όπως οι χρονοσειρές, οι προτάσεις (π.χ., για μετατροπή φωνής σε κείμενο), κτλ.
- Τα RNN εμφανίζουν φαινόμενο «μνήμης», και χρησιμοποιούν αποτελέσματα του προηγούμενου βήματος κατά την επεξεργασία του επόμενου βήματος μιας ακολουθίας.



Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet (P & S phase picking)

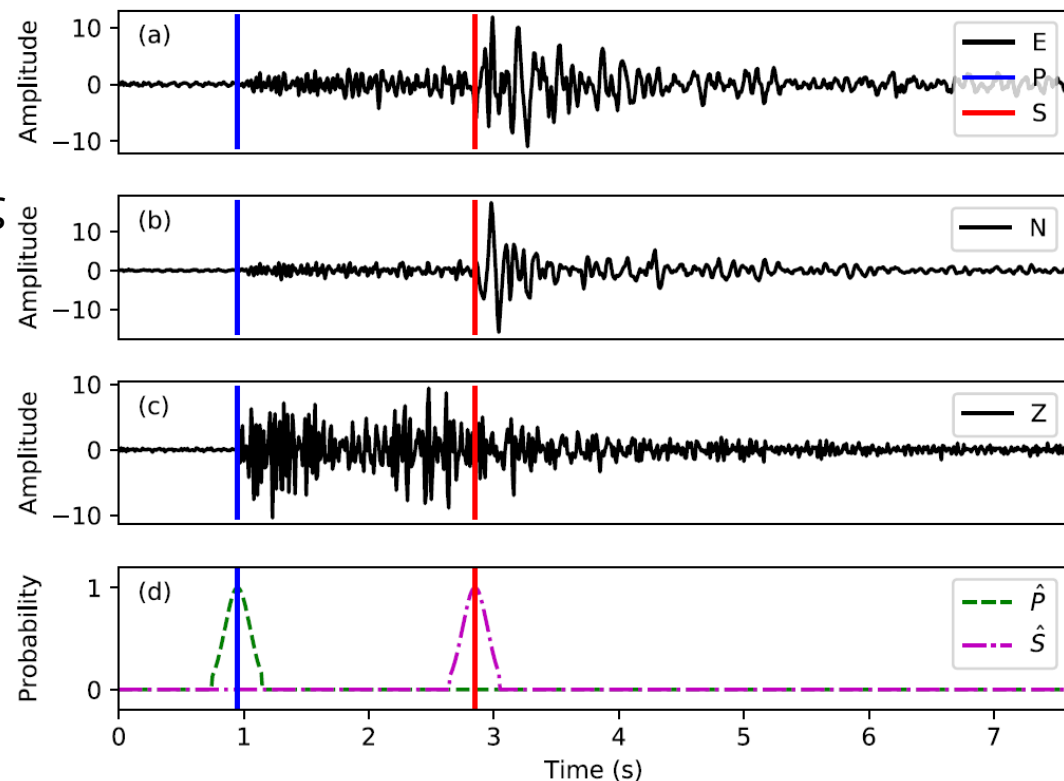


- Το μοντέλο βαθιάς μάθησης **PhaseNet** είναι φτιαγμένο για να αναγνωρίζει την άφιξη κυμάτων P ή κυμάτων S σε μια συνεχή σεισμική καταγραφή.
- Είναι εκπαιδευμένο με ένα σύνολο άνω των 700,000 επισημασμένων σειсмоγραμμάτων από πάνω από 230,000 σεισμούς που εκδηλώθηκαν σε μια περίοδο 30 ετών στη Βόρεια Καλιφόρνια.
- Το μοντέλο διαβάζει τμήματα κυματομορφών μήκους 3000 δειγμάτων, και χρησιμοποιεί συνελικτικά επίπεδα για να εξαγάγει χαρακτηριστικά.
- Στη συνέχεια, επιστρέφει στην έξοδό του χρονοσειρές ίσου μήκους (3000 δειγμάτων) με μόνο τα αναγνωρισμένα χαρακτηριστικά:
 - Την πιθανότητα σε μια δεδομένη χρονική στιγμή να εντοπίζεται κύμα P.
 - Την πιθανότητα σε μια δεδομένη χρονική στιγμή να εντοπίζεται κύμα S.

Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet

(P & S phase picking)

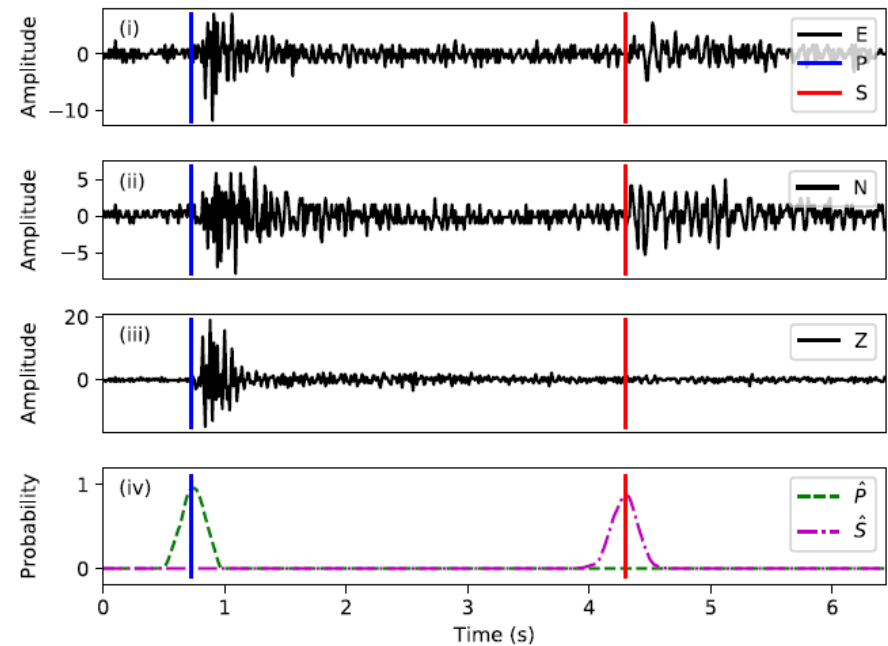
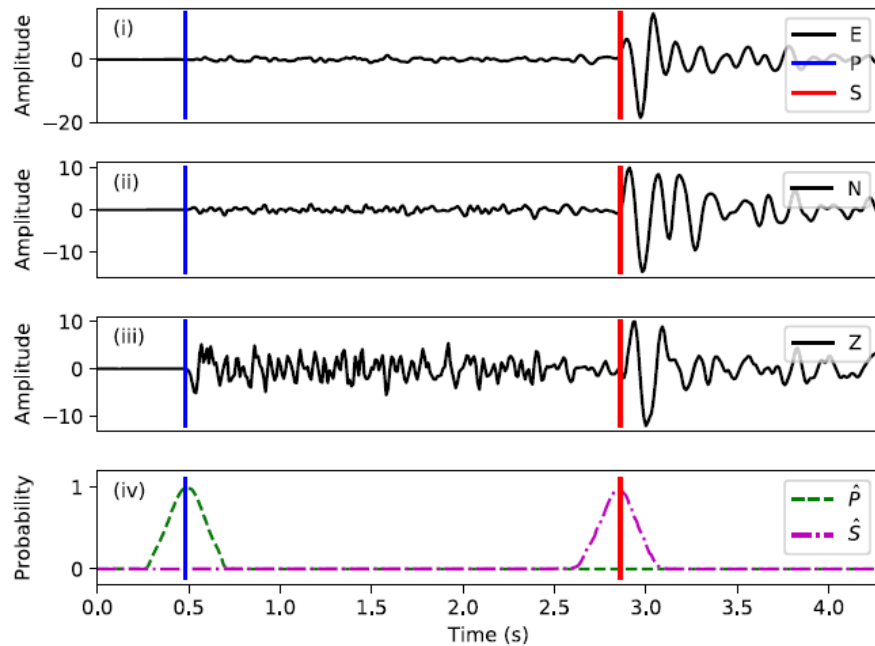
- Παράδειγμα επισημασμένων δεδομένων για την εκπαίδευση του μοντέλου PhaseNet.
- Τα δεδομένα είναι οι χρονοσειρές 3 συνιστωσών σε κάθε έναν σταθμό με καταγραφή ενός σεισμού.
- Ως ετικέτες στα επισημασμένα δεδομένα λαμβάνονται οι χρόνοι άφιξης των κυμάτων P και των κυμάτων S στον σταθμό, οι οποίοι έχουν μετρηθεί από αναλυτές.
- Ως επιθυμητά χαρακτηριστικά για την έξοδο του δικτύου θεωρούνται οι κατανομές πιθανότητας, οι οποίες είναι Γκαουσιανές με μέση τιμή μ τον εκάστοτε χρόνο άφιξης και τυπική απόκλιση $\sigma=0.1$ sec.
- Ο χρόνος άφιξης των P ή S βρίσκεται στο ολικό μέγιστο της κάθε Γκαουσιανής κατανομής.



Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet

(P & S phase picking)

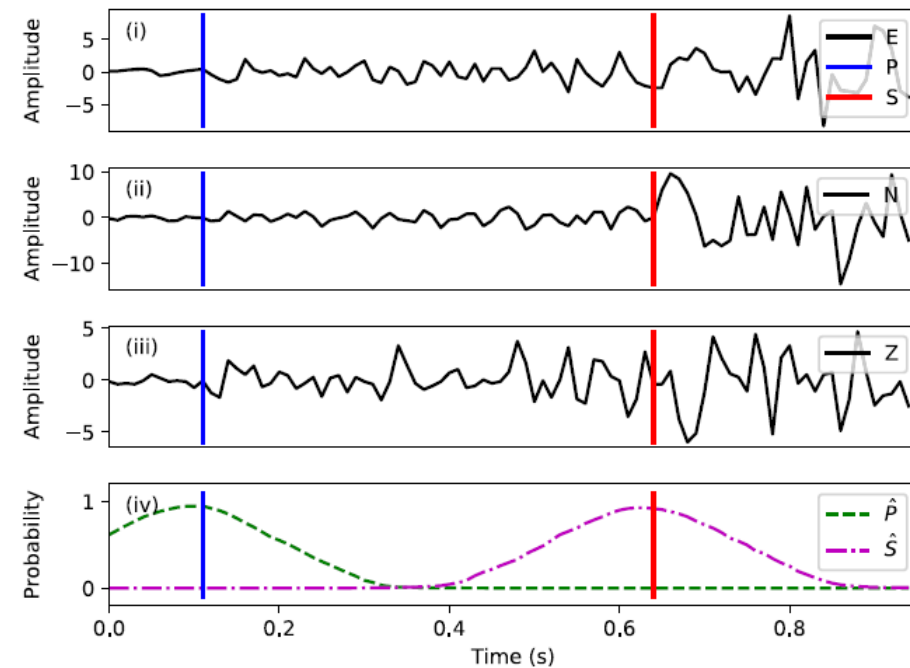
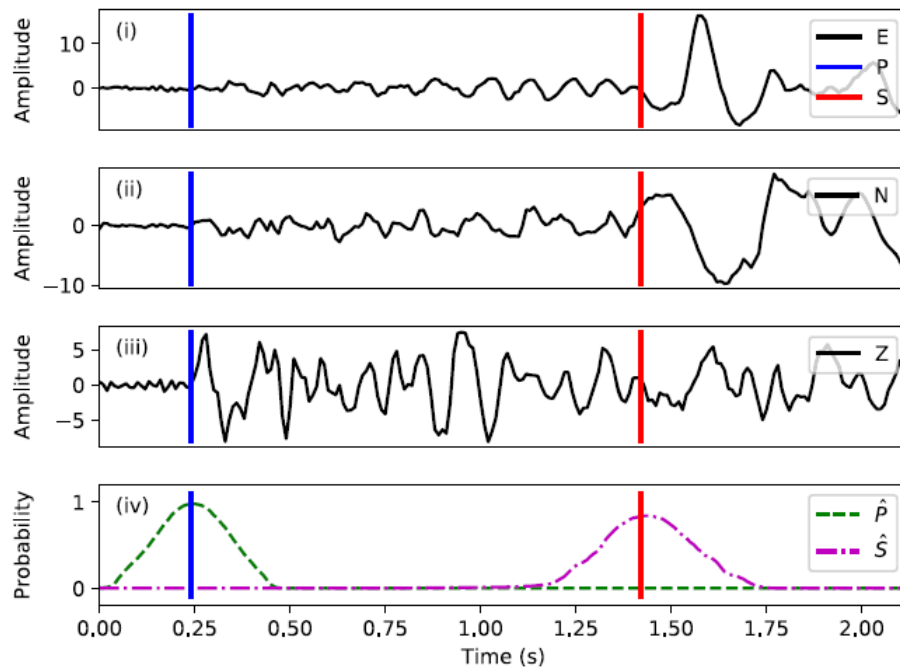
- Αποτελέσματα εφαρμογής του μοντέλου PhaseNet.



Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet

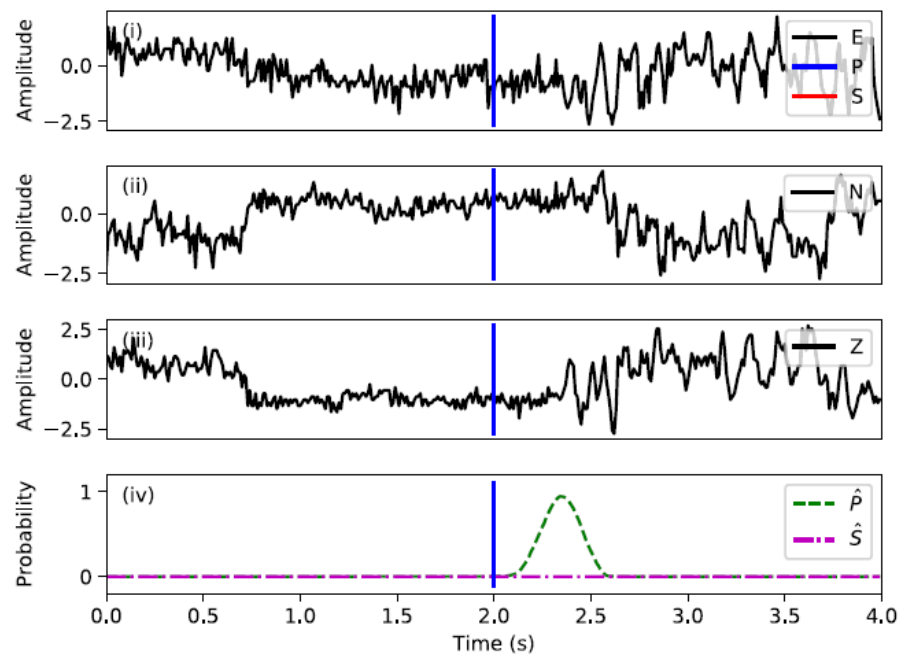
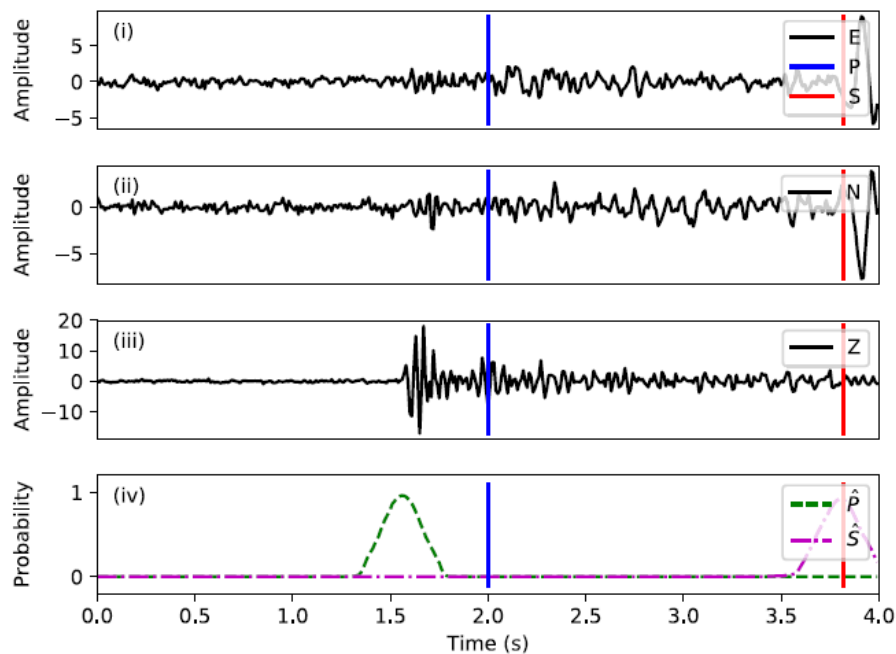
(P & S phase picking)

- Αποτελέσματα εφαρμογής του μοντέλου PhaseNet.



Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet (P & S phase picking)

- Παραδείγματα στα οποία κάποια επισημασμένα δεδομένα **δεν** είχαν καλή ακρίβεια στους χρόνους άφιξης.

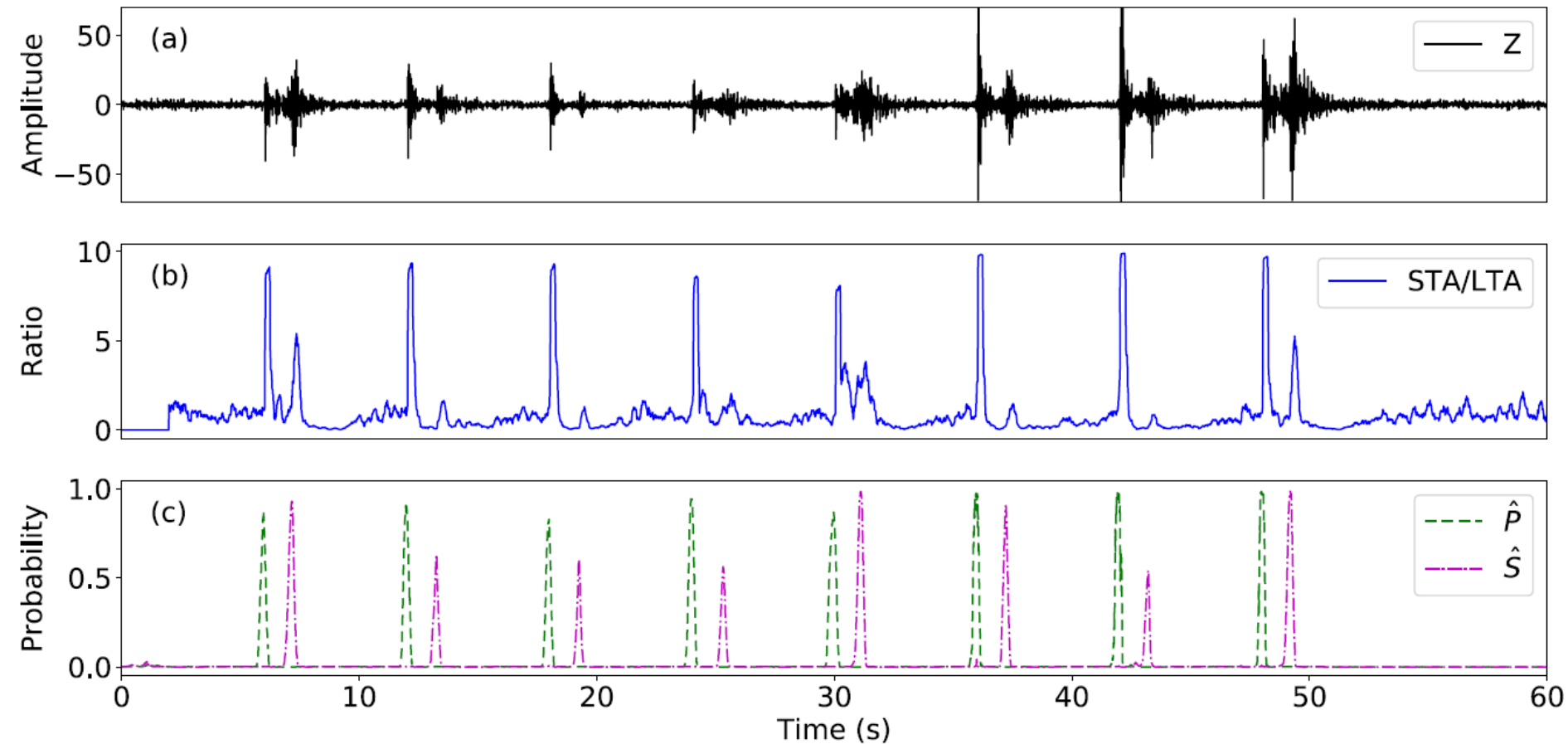


- Το μοντέλο PhaseNet βρίσκει σωστά τους χρόνους άφιξης.

Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet

(P & S phase picking)

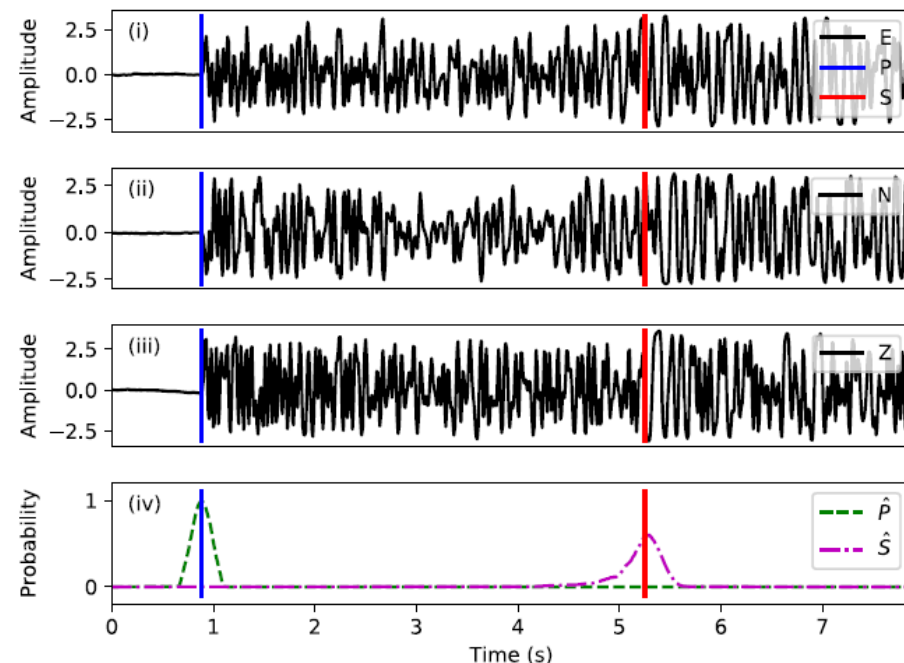
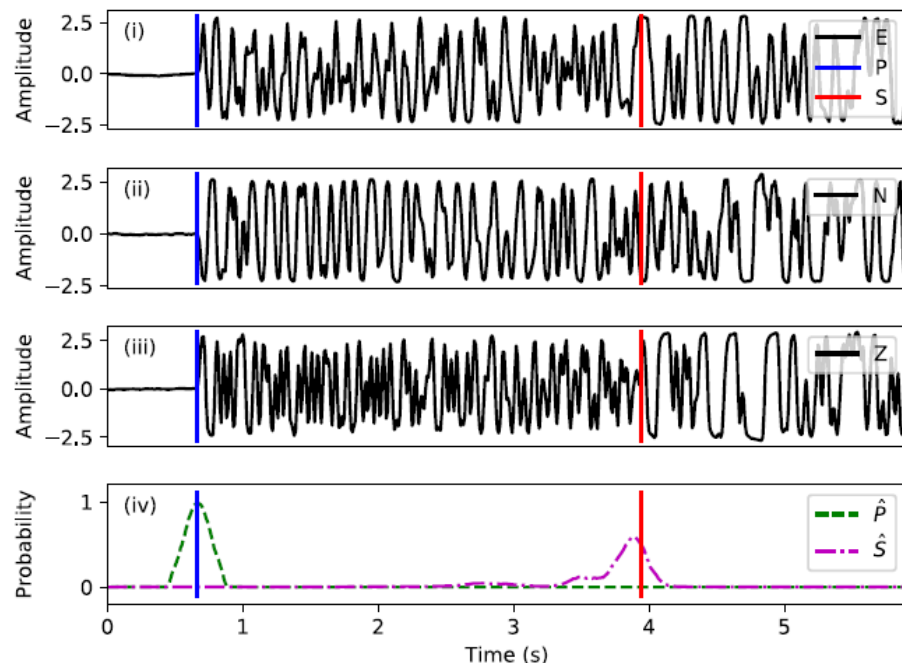
- Σύγκριση αποτελεσμάτων μοντέλου PhaseNet με μέθοδο STA/LTA.



Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet

(P & S phase picking)

- Αποτελέσματα μοντέλου PhaseNet σε δεδομένα με ψαλιδισμένα (clipped) πλάτη.
- Αυτό μπορεί να συμβεί κατά την καταγραφή ισχυρών σεισμών από τοπικούς σταθμούς.

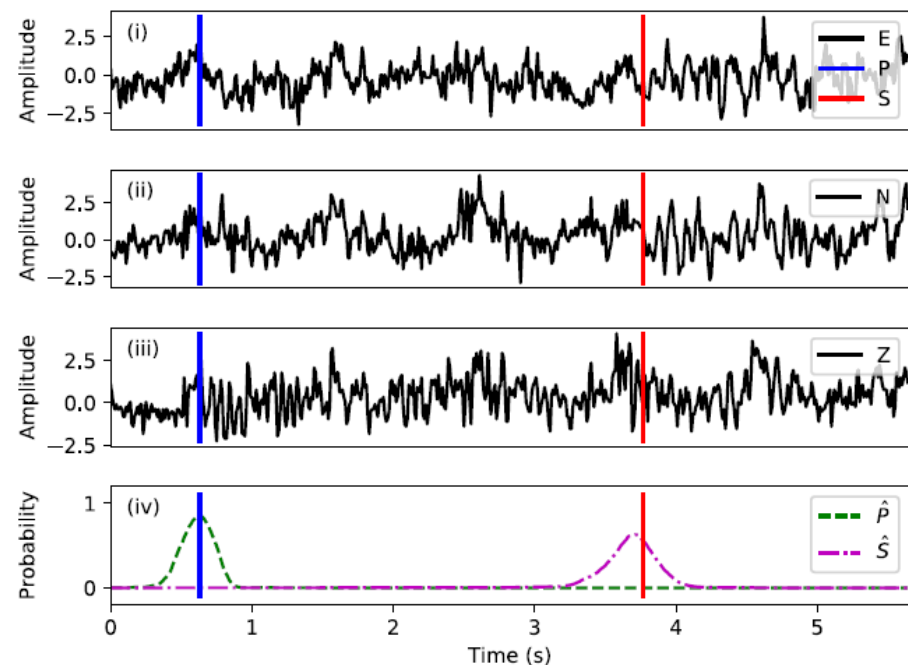
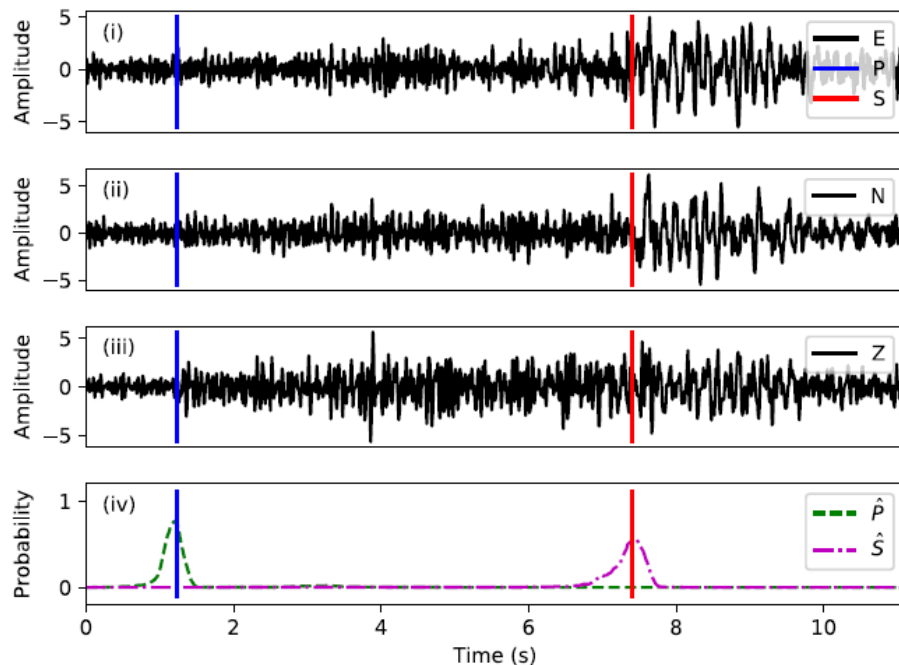


- Οι αναλυτές διακρίνουν την άφιξη των S από την αλλαγή στην περίοδο των κυμάτων.
- Το μοντέλο PhaseNet μπορεί επίσης να τα διακρίνει.

Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet

(P & S phase picking)

- Αποτελέσματα μοντέλου PhaseNet σε δεδομένα με χαμηλό λόγο σήματος-προς-θόρυβο (SNR).

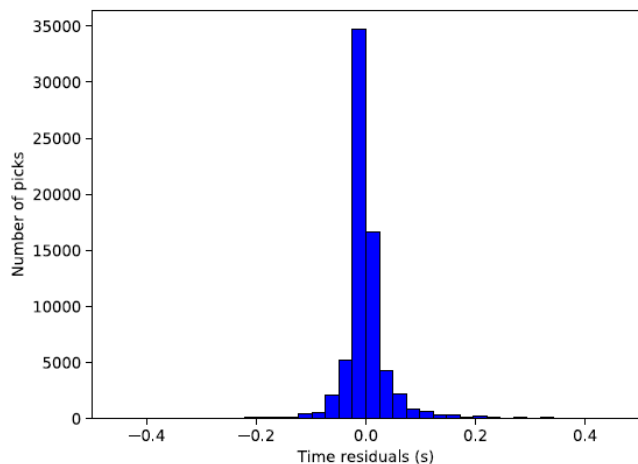


- Το μοντέλο PhaseNet μπορεί να διακρίνει τα P και S κύματα, ακόμα και χωρίς την εφαρμογή κάποιου φίλτρου (π.χ. Butterworth) για την ενίσχυση του SNR.

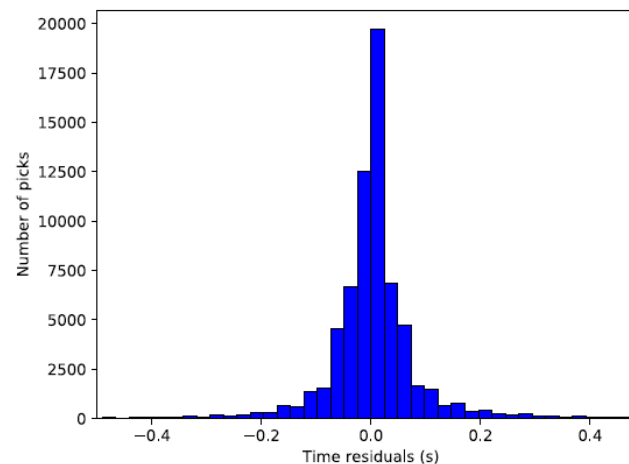
Αναγνώριση άφιξης σεισμικών κυμάτων P και S: PhaseNet

(P & S phase picking)

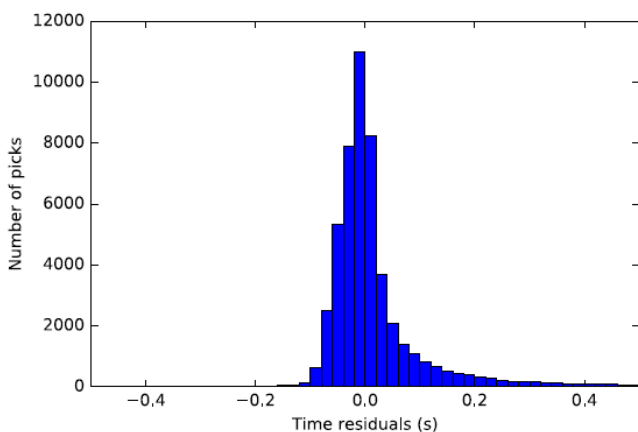
- Ακρίβεια μετρήσεων χρόνων άφιξης με μοντέλο PhaseNet ως προς τους αντίστοιχους χρόνους άφιξης από το σετ δοκιμών, σε σύγκριση με συμβατικά μοντέλα αυτοπαλινδρόμησης (auto-regression, ή AR).



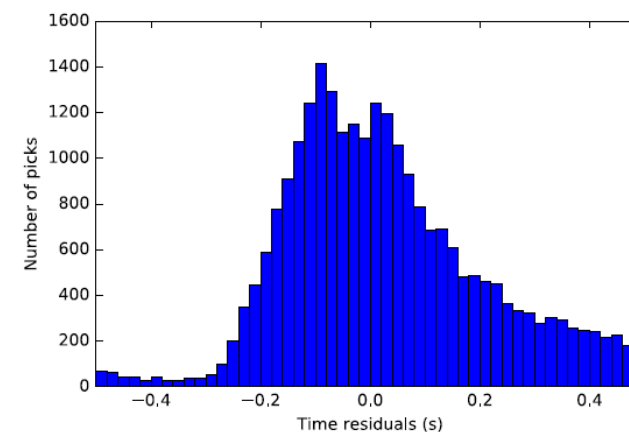
(a) P picks of PhaseNet



(b) S picks of PhaseNet



(c) P picks of AR picker

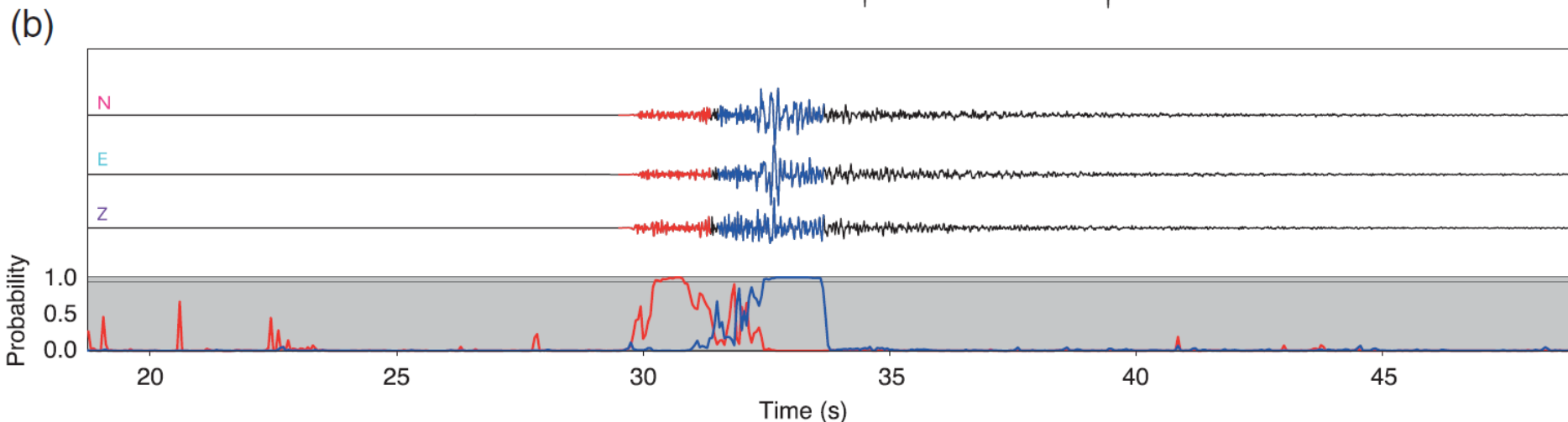
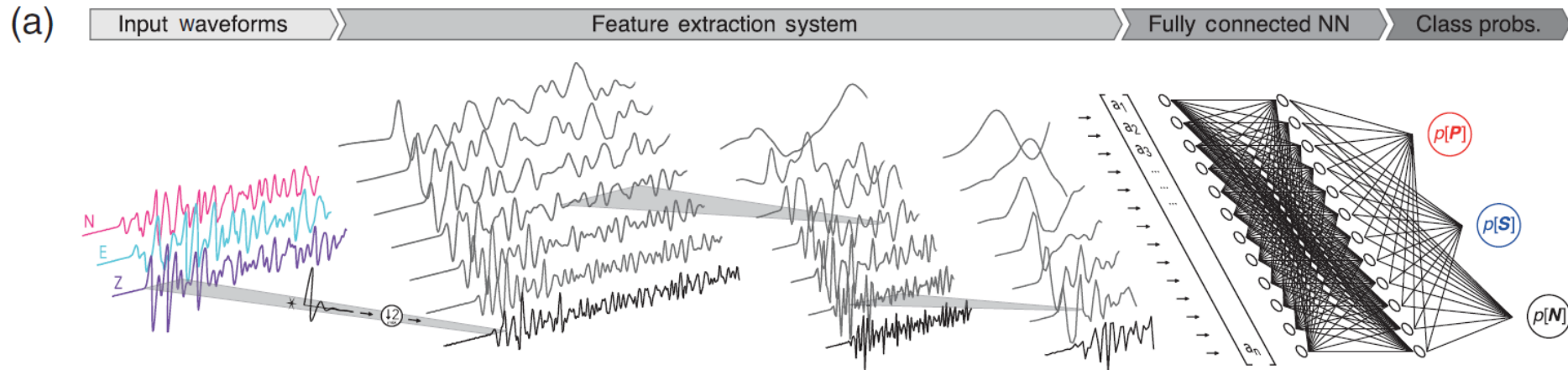


(d) S picks of AR picker

Αναγνώριση άφιξης σεισμικών κυμάτων P και S: GPD

(P & S phase picking)

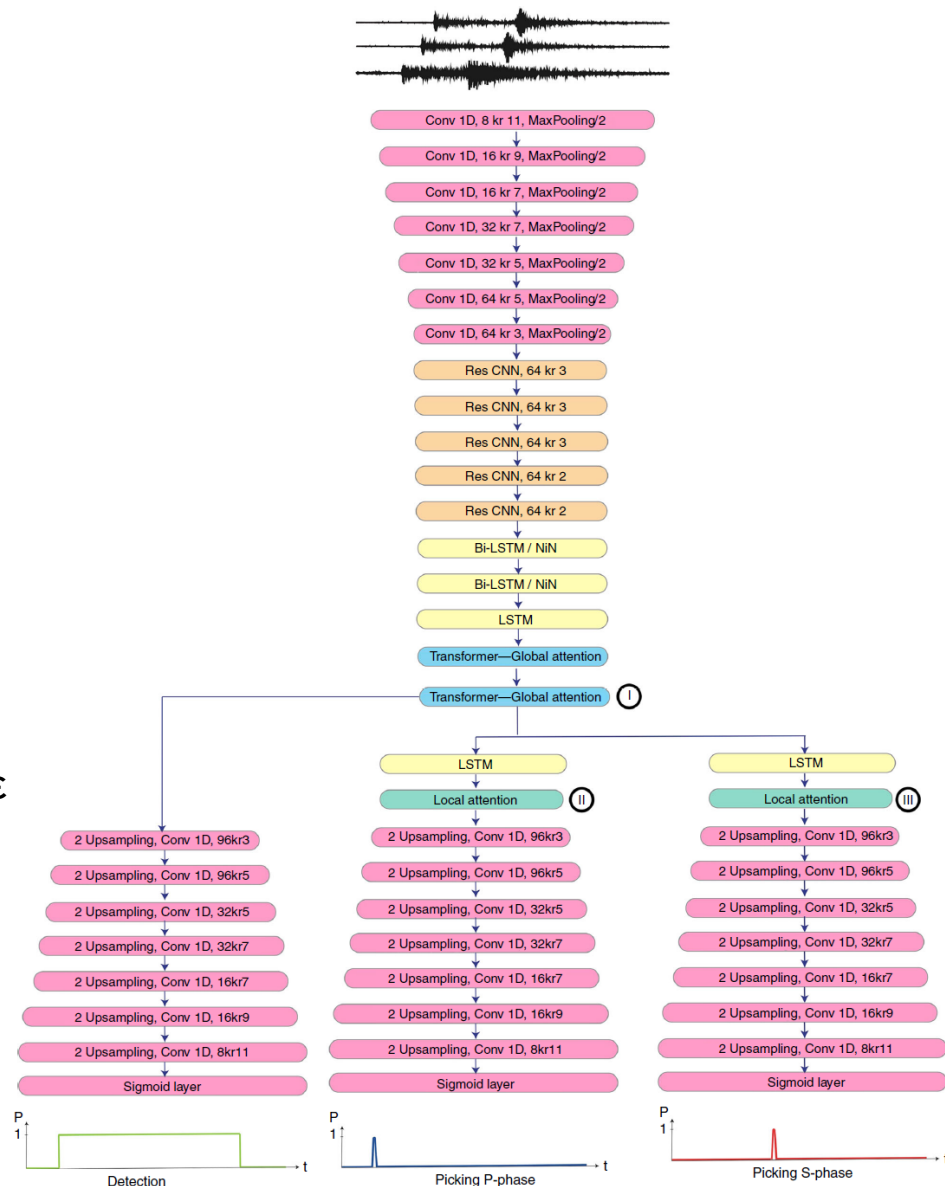
- Ένα παλαιότερο μοντέλο, παρόμοιας λογικής με το PhaseNet, είναι το μοντέλο GPD (Generalized seismic Phase Detection).



- Το μοντέλο αυτό διαθέτει στην έξοδο μια **τρίτη** κλάση για τον χαρακτηρισμό θορύβου (**N**: Noise).

Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

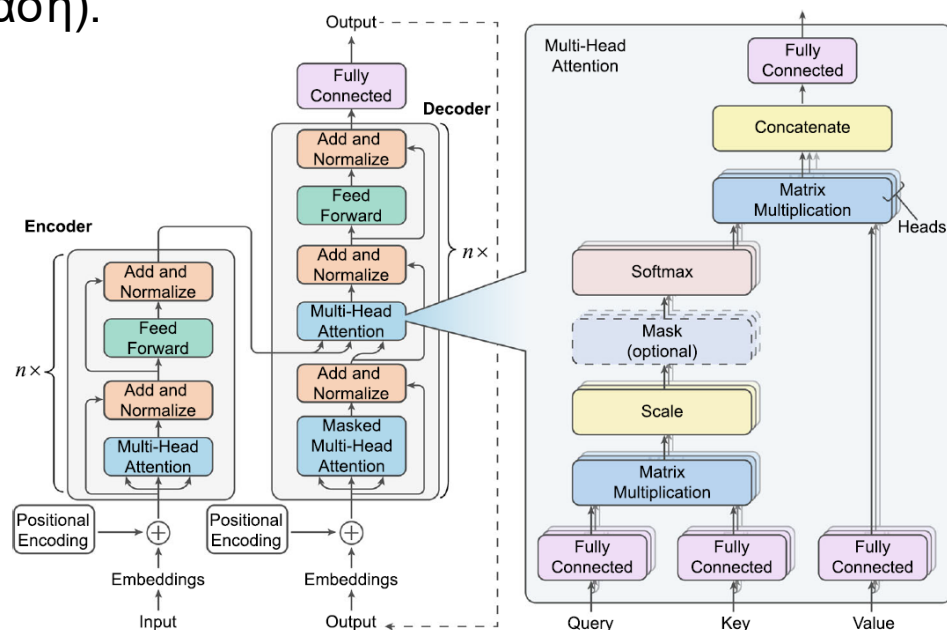
- Το μοντέλο βαθιάς μάθησης **EQTransformer** είναι φτιαγμένο για:
 - Να αναγνωρίζει σήματα που ανήκουν σε σεισμικά γεγονότα (event detection).
 - Να μετράει χρόνους άφιξης των P και S κυμάτων στα γεγονότα που αναγνώρισε.
- Για την εκπαίδευση του μοντέλου, χρησιμοποιήθηκε το επισημασμένο σετ δεδομένων “STEAD”, με ~1,000,000 καταγραφές ~450,000 σεισμών από διάφορες περιοχές της Γης σε τοπικές αποστάσεις (έως 300 km).
- Το μοντέλο χρησιμοποιεί επίπεδα συνέλιξης (όπως και το PhaseNet), αλλά τα συνδυάζει με έναν **μηχανισμό προσοχής** (attention mechanism), ανάλογο με τα μπλοκ **μετασχηματισμού** (transformer) που χρησιμοποιούνται σε μεγάλα γλωσσικά μοντέλα.



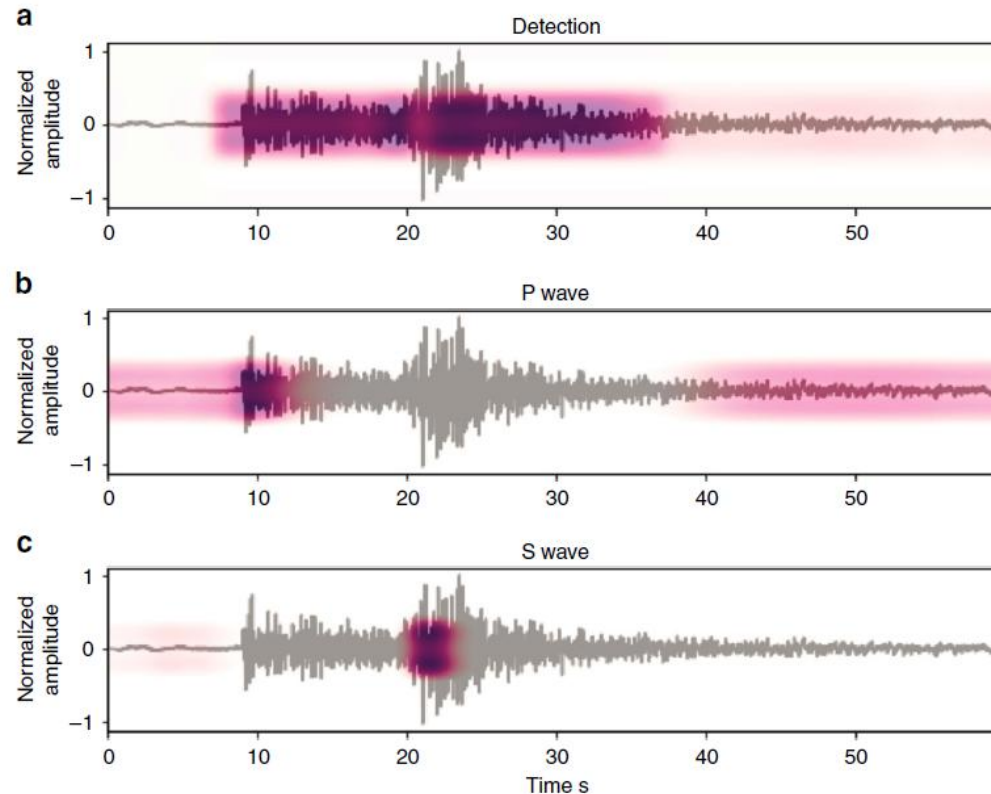
Μετασχηματιστές – Μηχανισμός Προσοχής

(Transformers – Attention Mechanism)

- Οι **μετασχηματιστές** (transformers) είναι ειδικά μπλοκ που περιλαμβάνονται κυρίως σε μεγάλα γλωσσικά μοντέλα (Large Language Models, LLM) μηχανικής μάθησης (π.χ. ChatGPT, DeepSeek, κτλ.).
- Οι μετασχηματιστές χρησιμοποιούνται σε δεδομένα που παρουσιάζονται υπό τη μορφή **ακολουθίας** (sequential data). Εξετάζουν ολόκληρη την ακολουθία ως έχει και βρίσκουν τον τρόπο με τον οποίο το κάθε τμήμα της συνδέεται με τα υπόλοιπα (όπως οι λέξεις σε μια πρόταση).
- Ο στόχος τους είναι να εστιάζουν την «προσοχή» του ΤΝΔ σε ορισμένα τμήματα της εισόδου που παρουσιάζουν **μεγαλύτερο «ενδιαφέρον»** από **άποψη πληροφορίας**.



Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

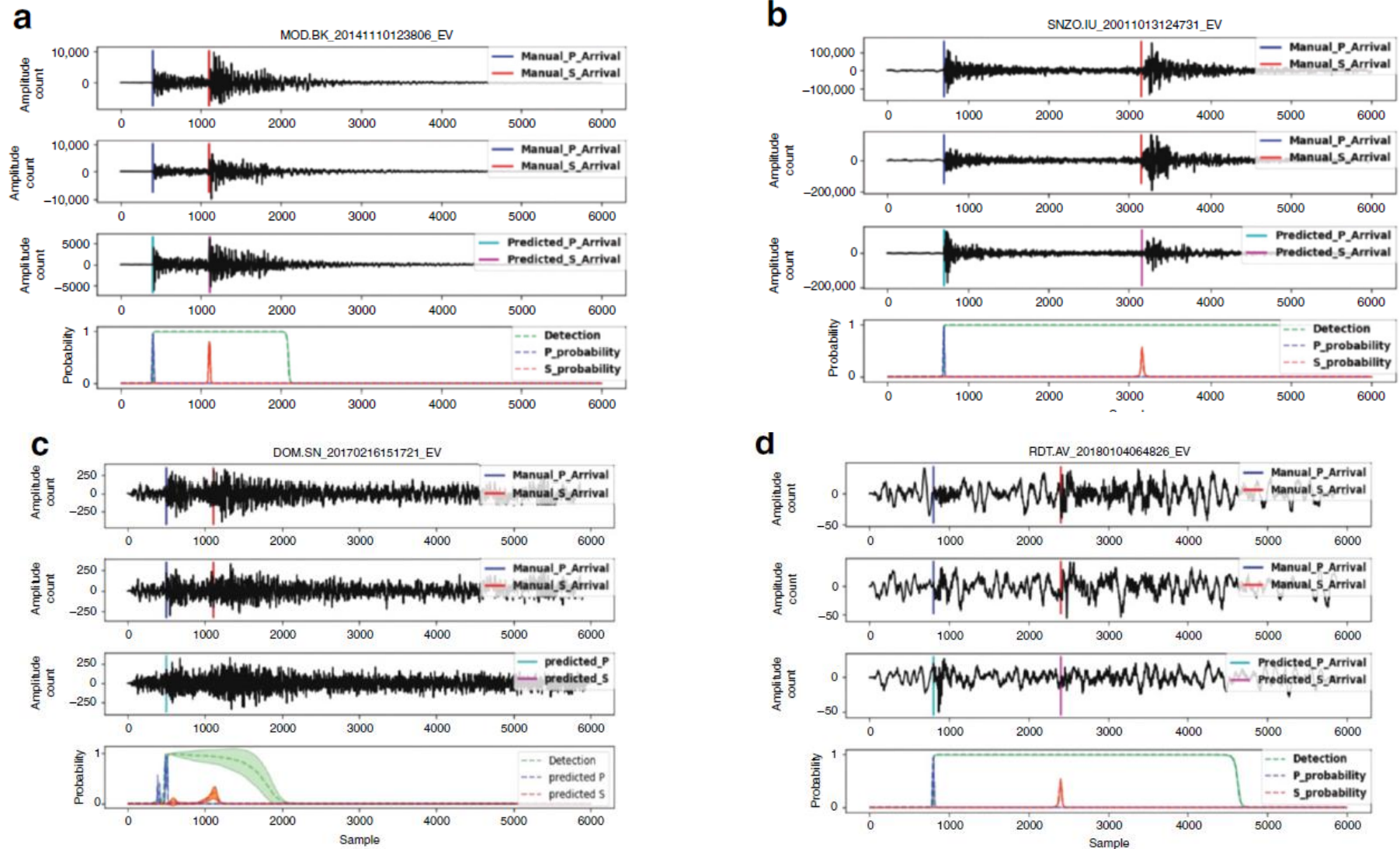


Mousavi et al. (2020)

- Ο μηχανισμός προσοχής στο EQTransformer δρα σε δύο κλίμακες:
 - Στη **συνολική σεισμική καταγραφή**, εστιάζοντας την προσοχή σε σήματα που αναγνωρίζονται ως **σεισμικά γεγονότα**.
 - Σε **τοπικό επίπεδο** γύρω από το κάθε σήμα σεισμικού γεγονότος, εστιάζοντας την προσοχή στο πρώτο μέρος που αφορά στο κύμα P και στο δεύτερο μέρος που αφορά στο κύμα S.
- Κατά αυτό τον τρόπο, το μοντέλο EQTransformer λειτουργεί όπως ένας (άνθρωπος) αναλυτής όταν αναλύει μια σεισμική καταγραφή.

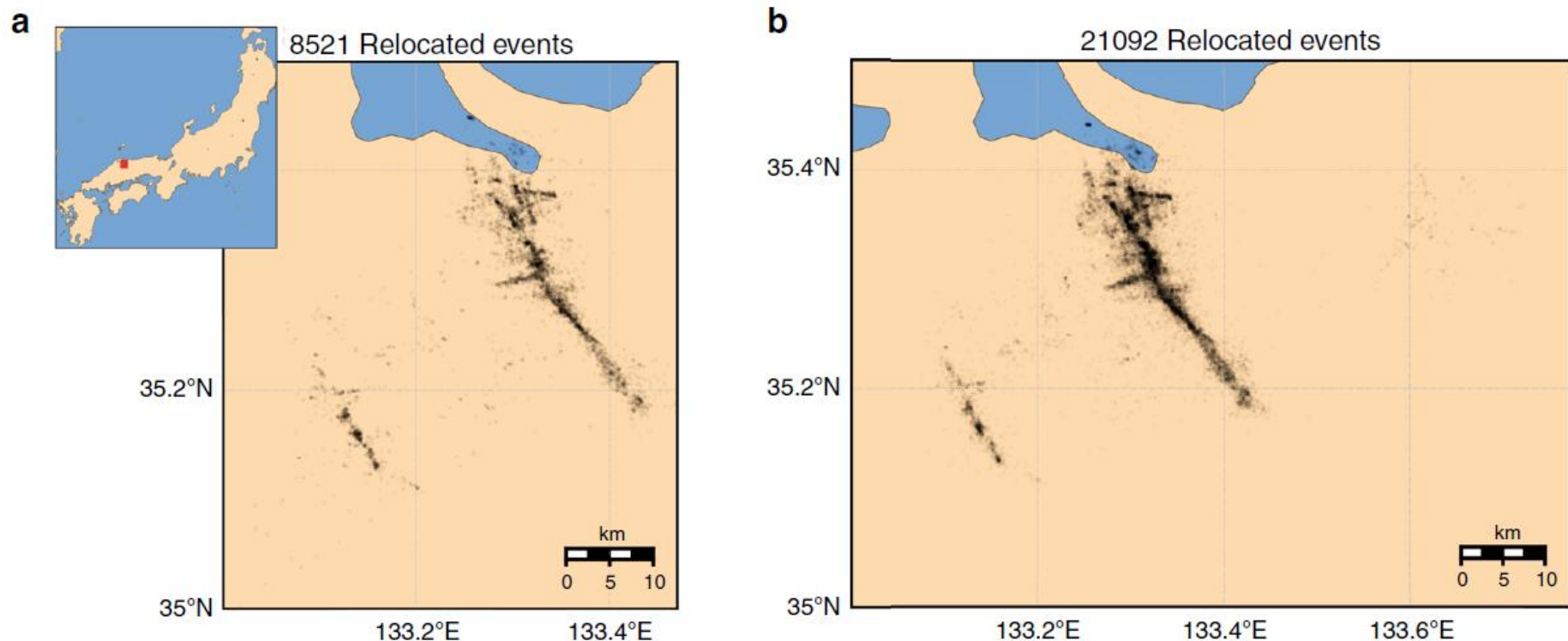
Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

- Παραδείγματα εφαρμογής του μοντέλου EQTransformer.



Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

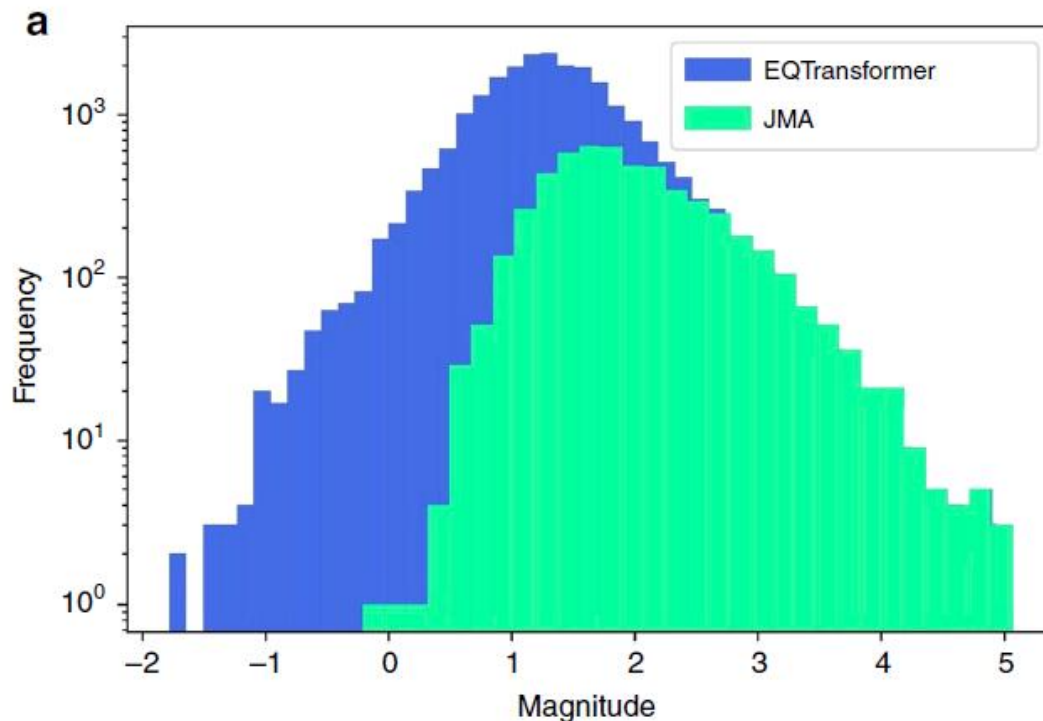
- Εφαρμογή του μοντέλου EQTransformer στη σεισμική ακολουθία της περιοχής Tottori στη Ν. Ιαπωνία (6 Οκτωβρίου – 17 Νοεμβρίου 2000).



- Το μοντέλο EQTransformer δεν διέθετε καθόλου επισημασμένα δεδομένα από σεισμούς στην Ιαπωνία (παράδειγμα γενίκευσης μοντέλου).
- Το EQTransformer εντόπισε υπερδιπλάσιο αριθμό σεισμών απ'ότι το ινστιτούτο JMA, με λιγότερο από το 1/3 των διαθέσιμων σταθμών.

Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

- Εφαρμογή του μοντέλου EQTransformer στη σεισμική ακολουθία της περιοχής Tottori στη Ν. Ιαπωνία (6 Οκτωβρίου – 17 Νοεμβρίου 2000).

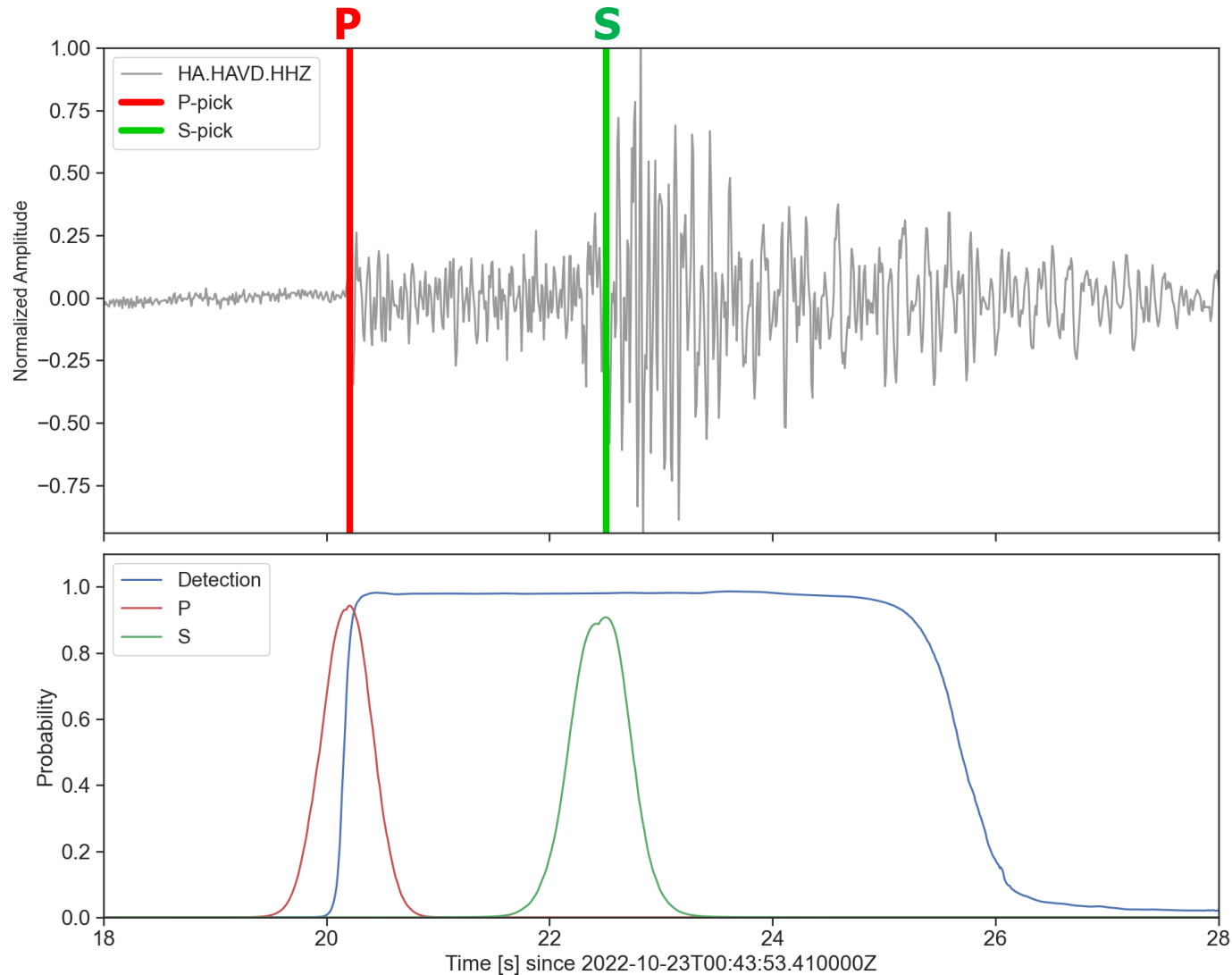


Mousavi et al. (2020)

- Το μοντέλο EQTransformer εντόπισε σεισμούς με πολύ μικρότερα μεγέθη από αυτούς που είχε εντοπίσει το JMA.
- Το μέγεθος πληρότητας του καταλόγου JMA ήταν $M_c=1.82$, ενώ του καταλόγου που δημιουργήθηκε με το EQTransformer είναι $M_c=1.50$.
- Σε άλλες εφαρμογές, η πληρότητα μπορεί να βελτιωθεί ακόμα και κατά μία τάξη μεγέθους με χρήση μοντέλων βαθιάς μάθησης.

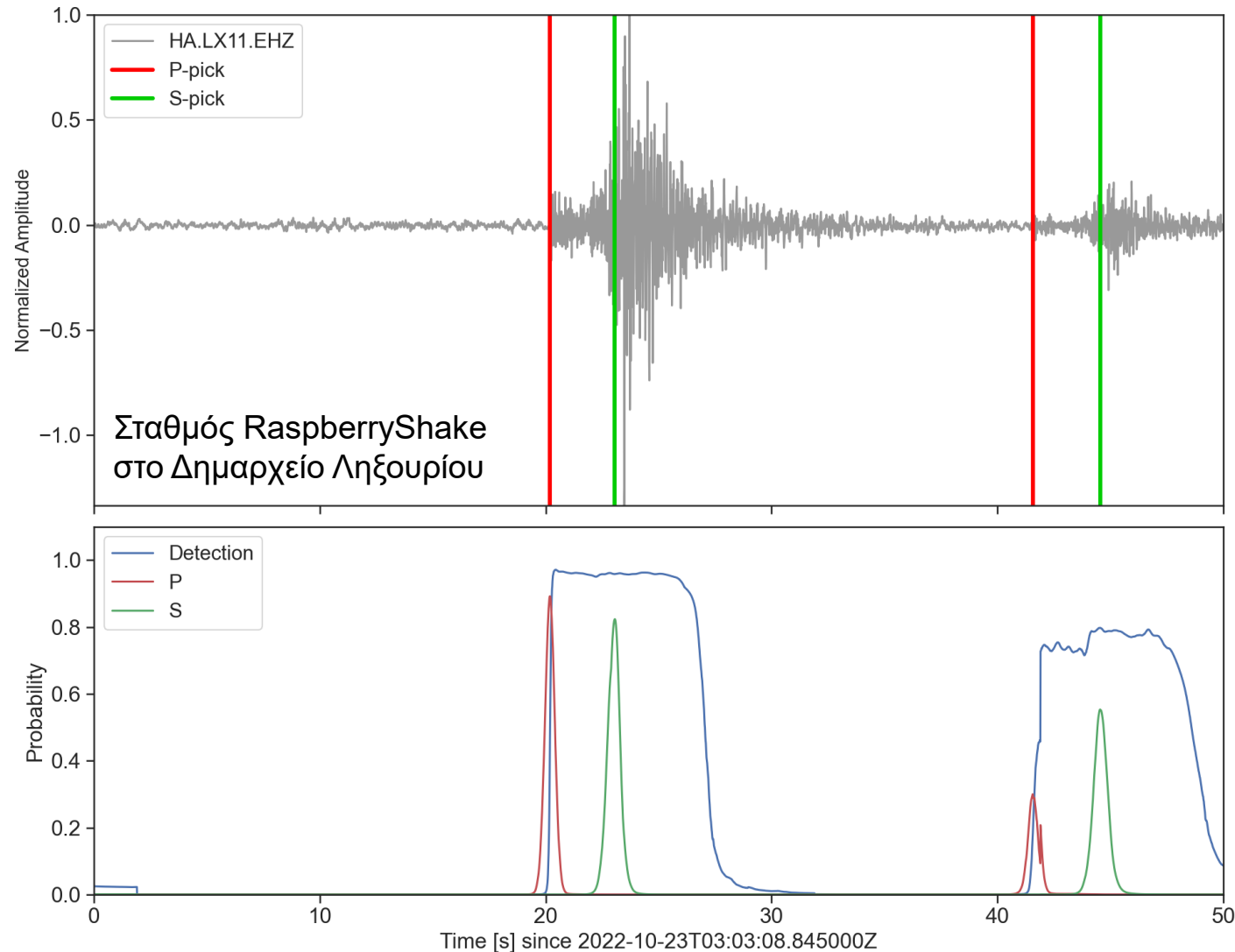
Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

- Εφαρμογή του μοντέλου EQTransformer σε δεδομένα σεισμολογικών σταθμών στην Κεφαλονιά.



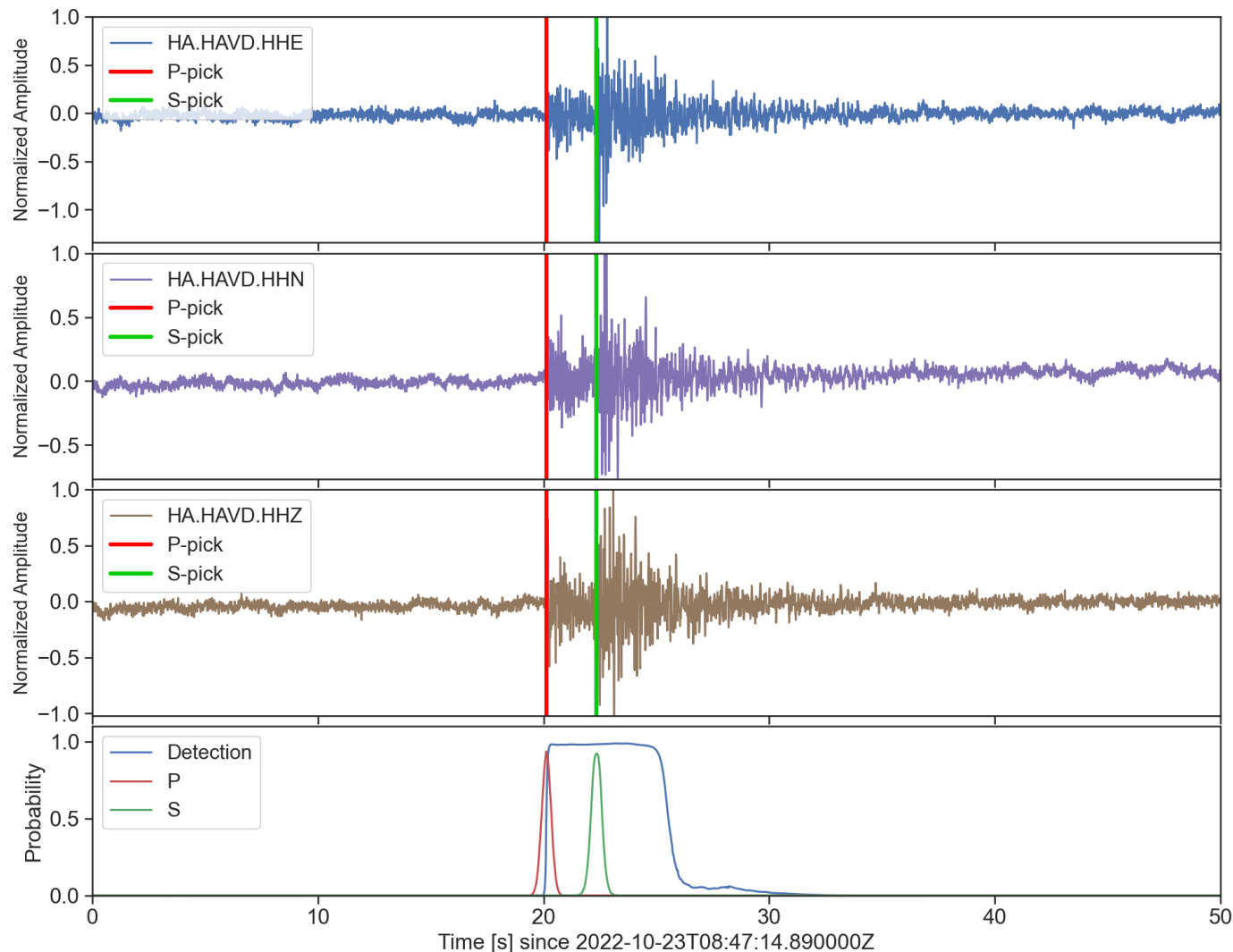
Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

- Εφαρμογή του μοντέλου EQTransformer σε δεδομένα σεισμολογικών σταθμών στην Κεφαλονιά.



Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

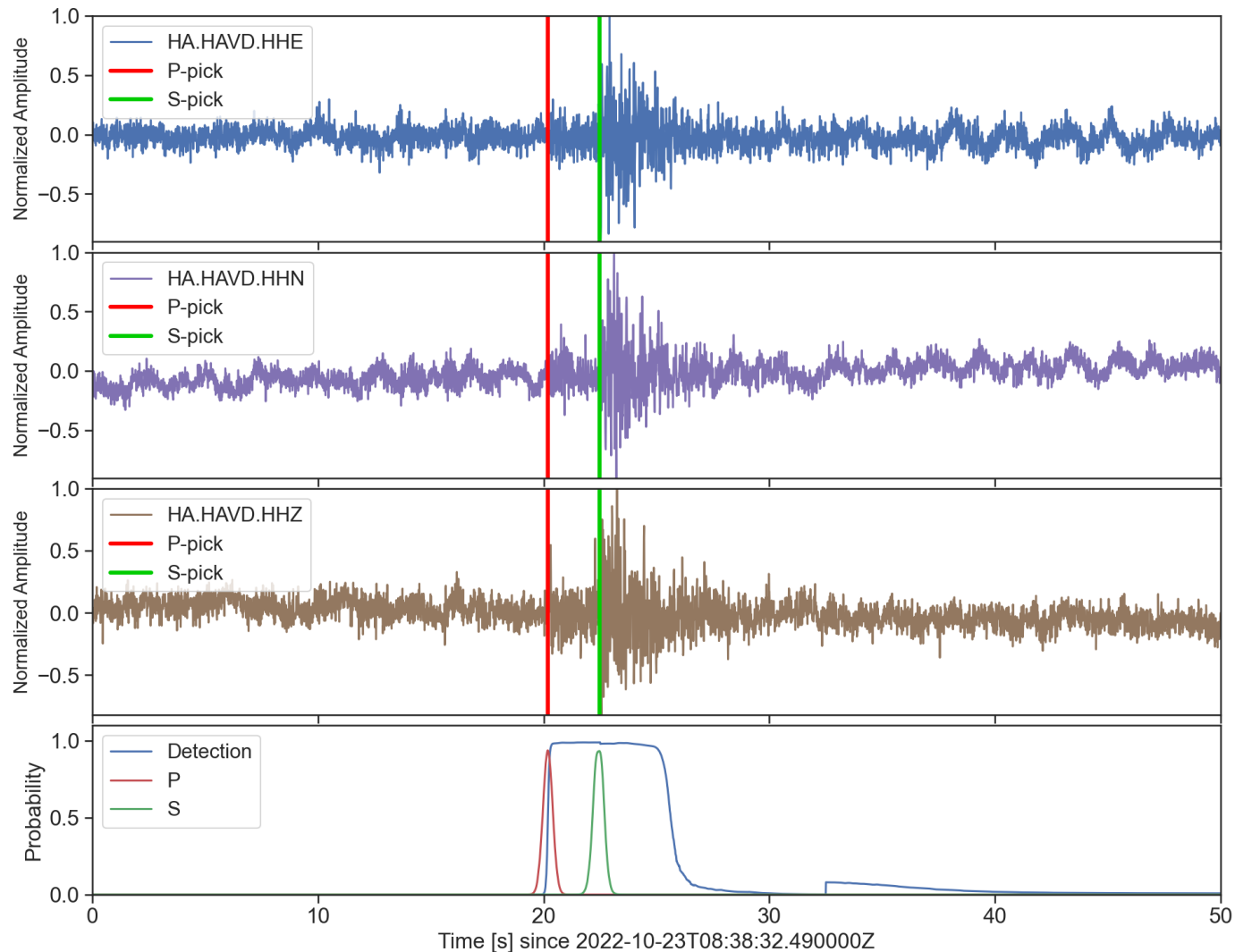
- Εφαρμογή του μοντέλου EQTransformer σε δεδομένα σεισμολογικών σταθμών στην Κεφαλονιά, με σταδιακά χαμηλότερο λόγο SNR.



Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: **EQTransformer**

(event detection + P&S phase picking)

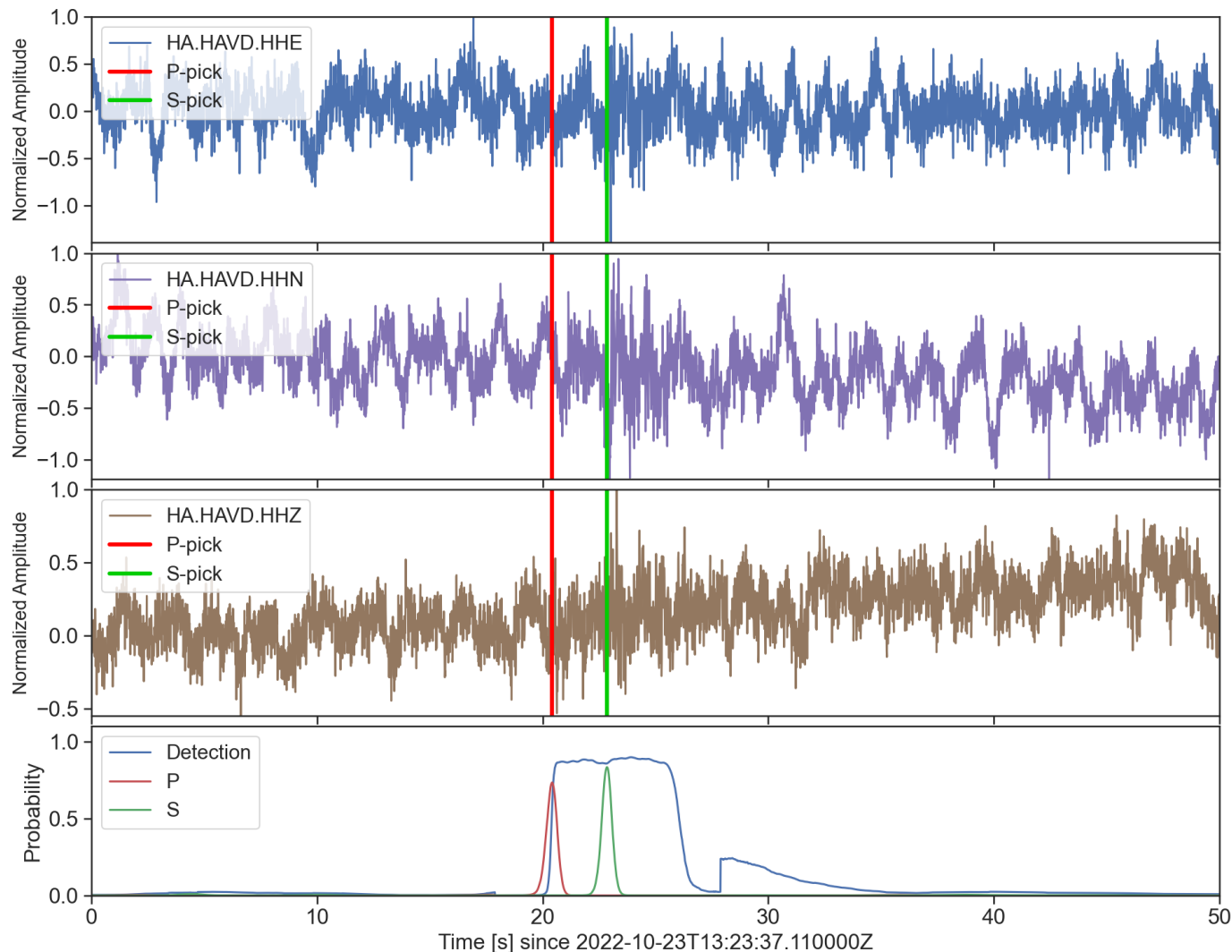
- Εφαρμογή του μοντέλου EQTransformer σε δεδομένα σεισμολογικών σταθμών στην Κεφαλονιά, με σταδιακά χαμηλότερο λόγο SNR.



Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: **EQTransformer**

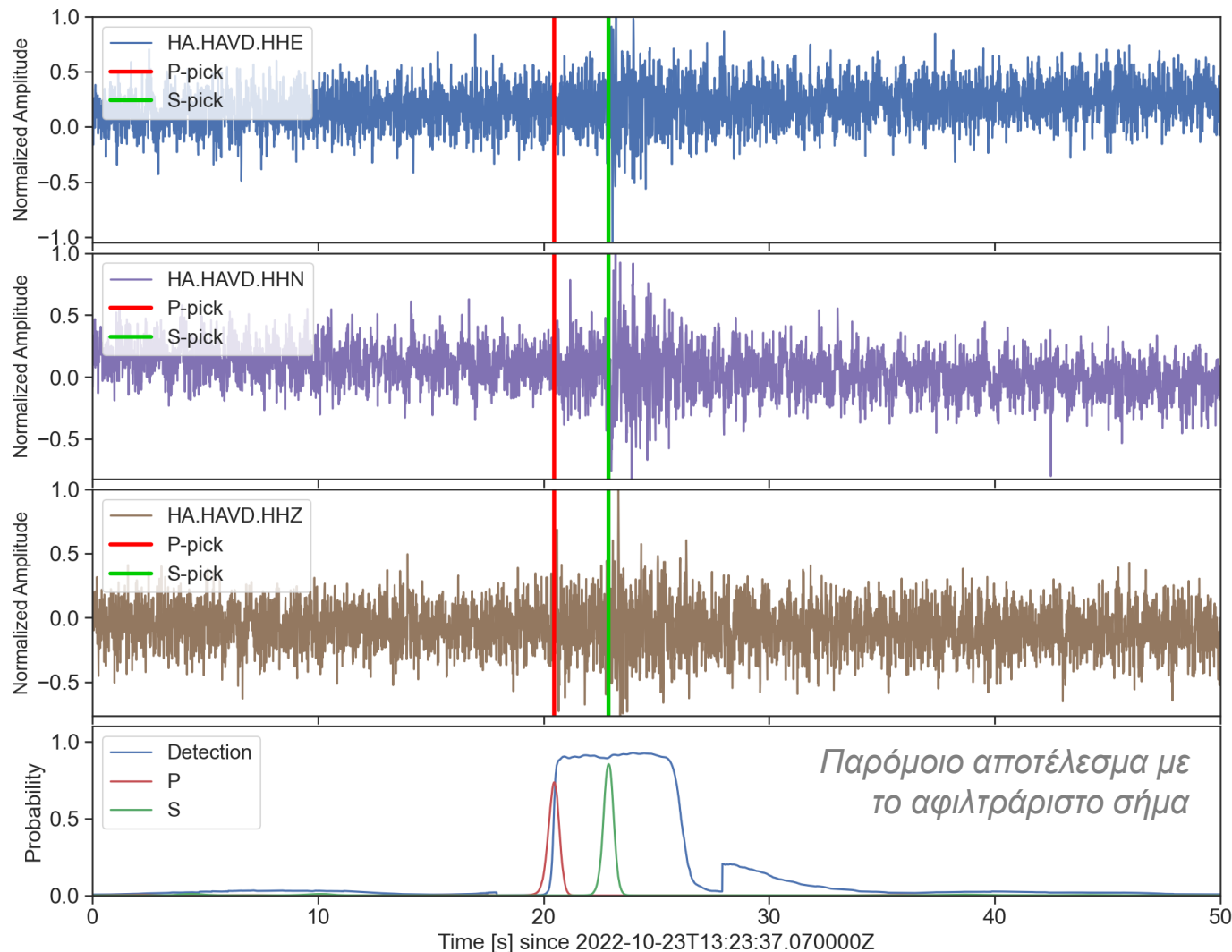
(event detection + P&S phase picking)

- Εφαρμογή του μοντέλου EQTransformer σε δεδομένα σεισμολογικών σταθμών στην Κεφαλονιά, με σταδιακά χαμηλότερο λόγο SNR.



Αναγνώριση σεισμικών γεγονότων και άφιξης σεισμικών κυμάτων P και S: EQTransformer (event detection + P&S phase picking)

- Εφαρμογή του μοντέλου EQTransformer σε δεδομένα σεισμολογικών σταθμών στην Κεφαλονιά, με σταδιακά χαμηλότερο λόγο SNR. [εφαρμογή υπερατατού φίλτρου Butterworth >1Hz]

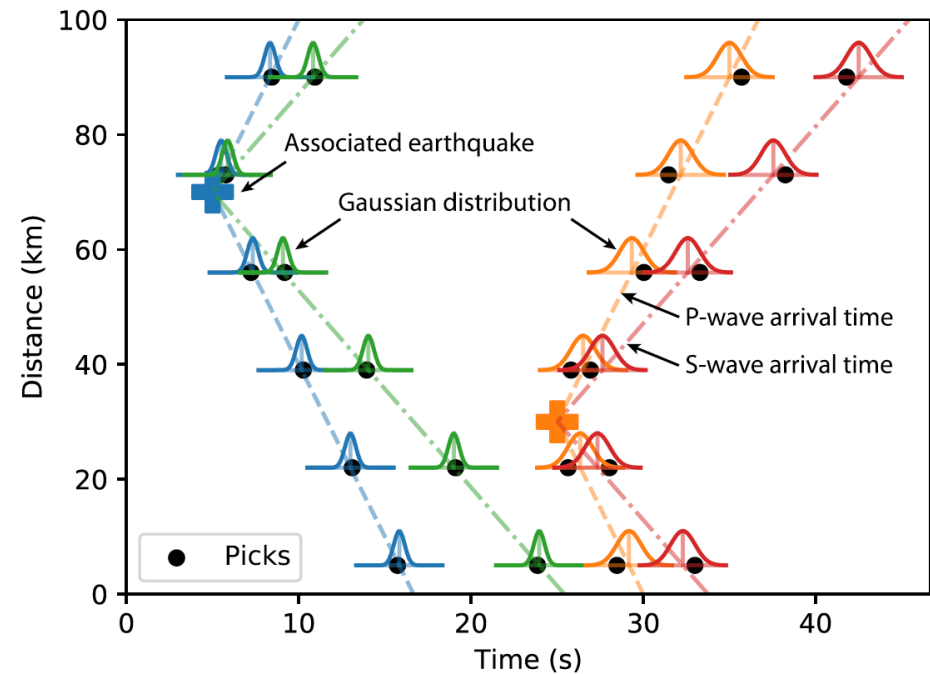


Συσχετισμός φάσεων με σεισμικά γεγονότα (phase/event association)

- Τα μοντέλα PhaseNet και EQTransformer ασχολούνται με τη μέτρηση των χρόνων άφιξης των κυμάτων P και S.
- Χρησιμοποιώντας τα μοντέλα αυτά μπορούμε να φτιάξουμε πίνακες με χρόνους άφιξης των P και S σε κάθε σταθμό.
- Ωστόσο, για να μπορούμε να προσδιορίσουμε τις εστιακές παραμέτρους ενός σεισμού, χρειάζεται να γνωρίζουμε ποιοι χρόνοι άφιξης P και S από κάθε σταθμό αντιστοιχούν στο **ίδιο** σεισμικό γεγονός.
- Η διαδικασία της εύρεσης και συσχετισμού των χρόνων άφιξης σεισμικών φάσεων με μια ορισμένη **πηγή προέλευσης** (origin) ονομάζεται **συσχετισμός φάσεων** (phase- ή event-association).
- Ο πιο απλός τρόπος είναι να επιλέξουμε τις φάσεις P και S που εντοπίζονται σε πολλούς σταθμούς σε στενό χρονικό παράθυρο.
- Ωστόσο, αυτό δεν εξασφαλίζει ότι οι φάσεις αυτές ανήκουν στο ίδιο σεισμικό γεγονός, ιδιαίτερα όταν υπάρχει μια σεισμική ακολουθία σε εξέλιξη στην (ευρύτερη) περιοχή μελέτης.

Συσχετισμός φάσεων με σεισμικά γεγονότα: μέθοδος **GaMMA** (phase/event association)

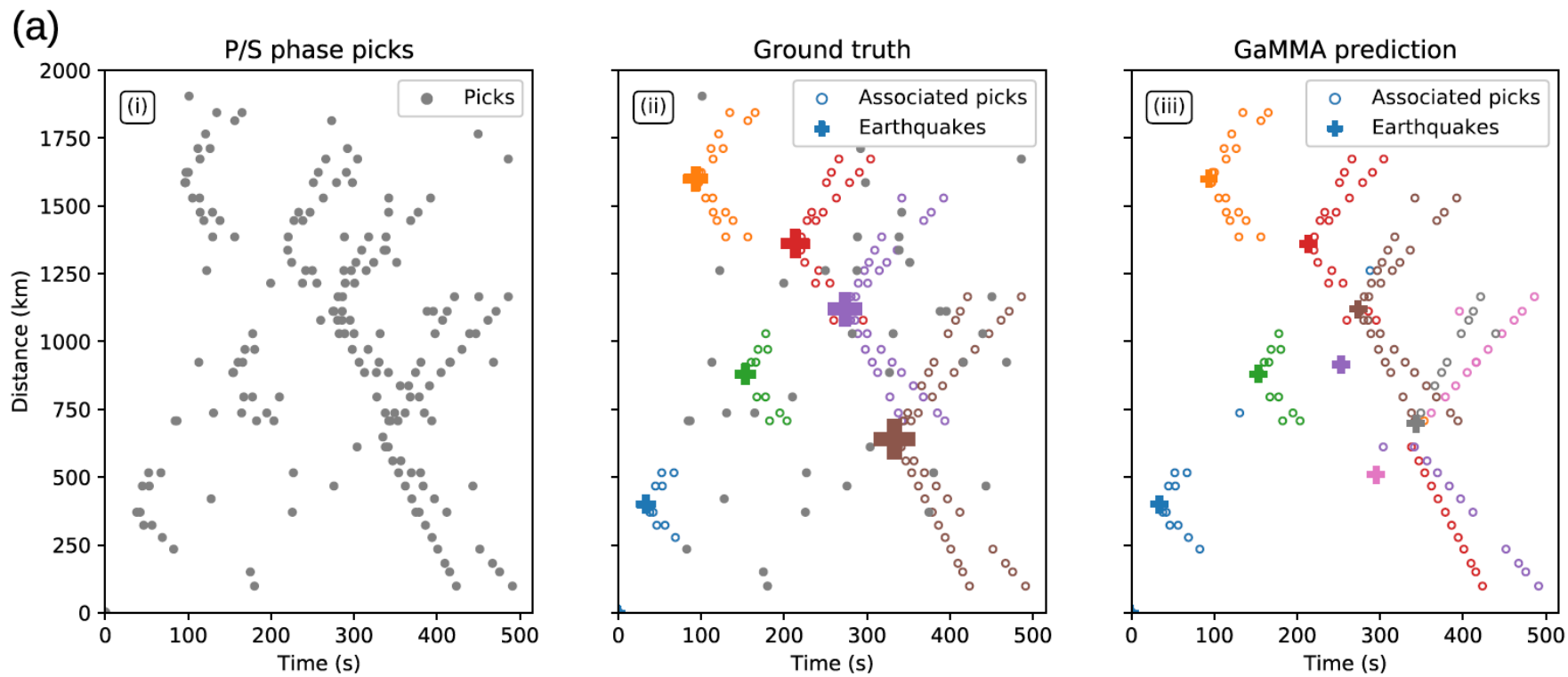
- Η μέθοδος GaMMA (Gaussian Mixture Model Association) είναι μια μέθοδος μη επιβλεπόμενης μηχανικής μάθησης, αλλά όχι βαθιάς μάθησης.
- Συνδυάζει ένα μοντέλο Γκαουσιανού μείγματος με τον προσδιορισμό των εστιακών παραμέτρων (χρόνο γένεσης, επίκεντρο, εστιακό βάθος και μέγεθος).



- Η μέθοδος λειτουργεί θεωρώντας πως ο κάθε σεισμός αντιστοιχεί σε μια συστάδα χρόνων άφιξης των P και S κυμάτων, με χρόνους άφιξης που εξαπλώνονται υπερβολικά συναρτήσει της επικεντρικής απόστασης και του χρόνου διαδρομής, ενώ και τα πλάτη των κυμάτων φθίνουν με αυξανόμενη επικεντρική απόσταση.

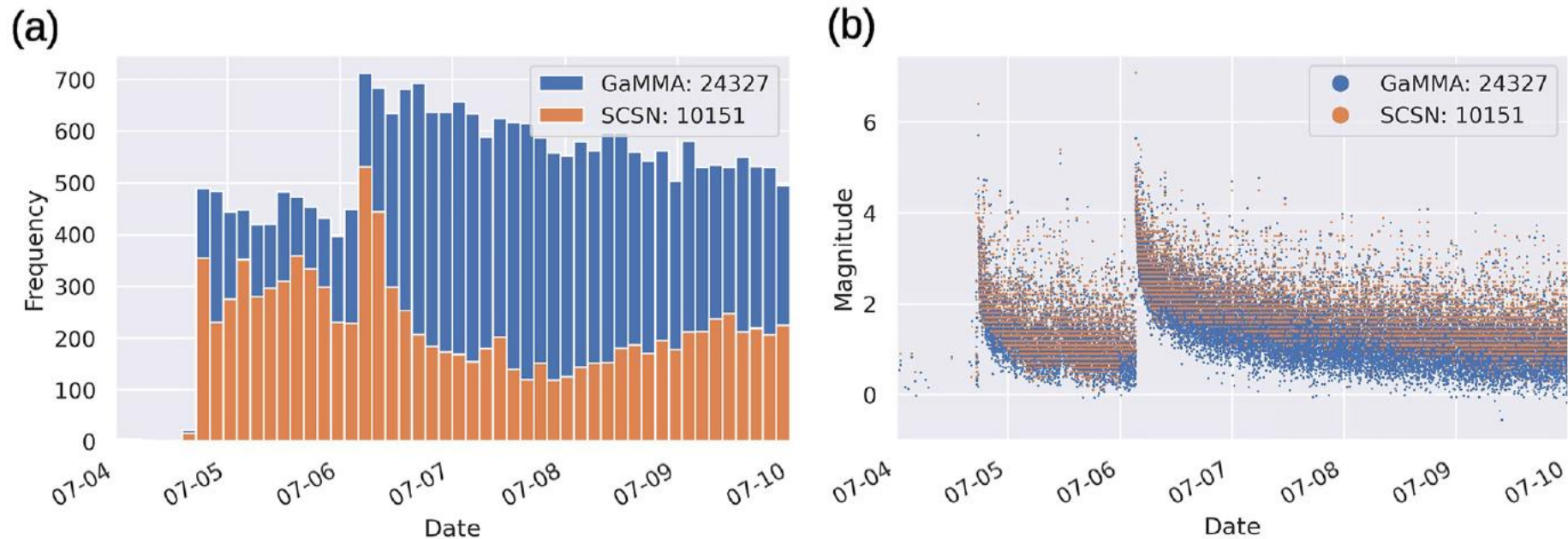
Συσχετισμός φάσεων με σεισμικά γεγονότα: μέθοδος **GaMMA** (phase/event association)

- Απεικόνιση του τρόπου με τον οποίο λειτουργεί η μέθοδος GaMMA για τον συσχετισμό φάσεων σε περίπτωση πολλαπλών σεισμικών γεγονότων.
- Στο παράδειγμα χρησιμοποιούνται μόνο χρόνοι άφιξης P και S (χωρίς πλάτη).



Συσχετισμός φάσεων με σεισμικά γεγονότα: μέθοδος **GaMMA** (phase/event association)

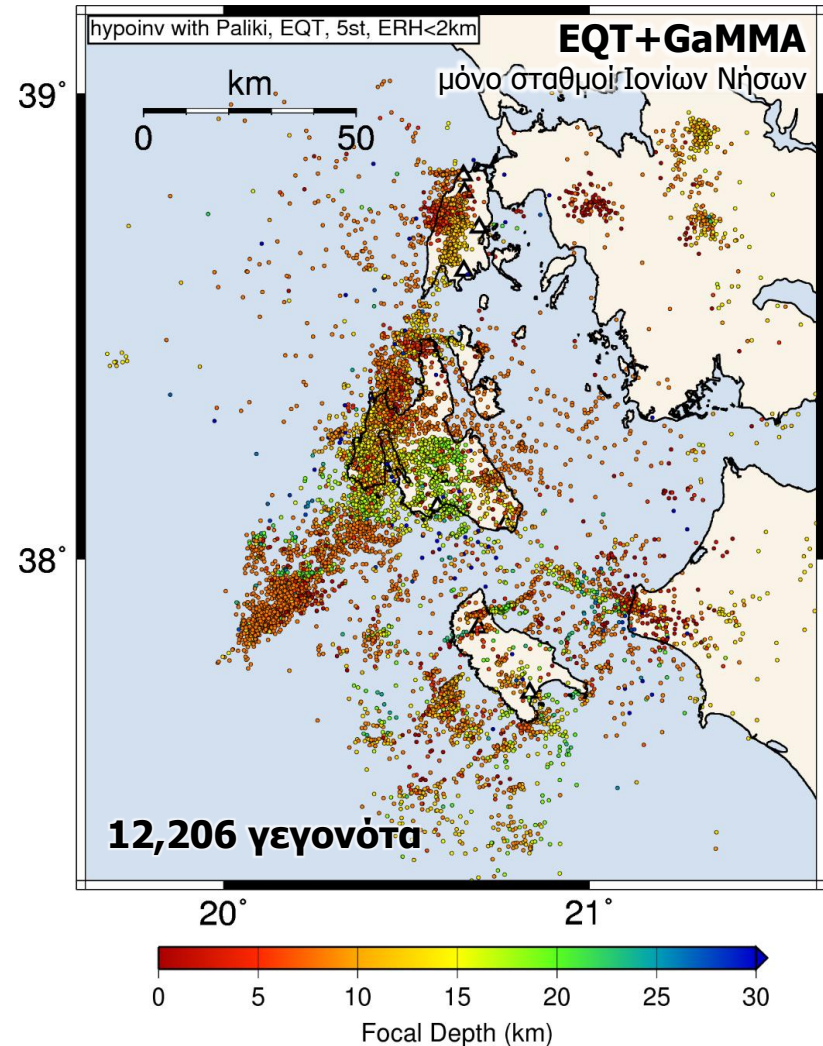
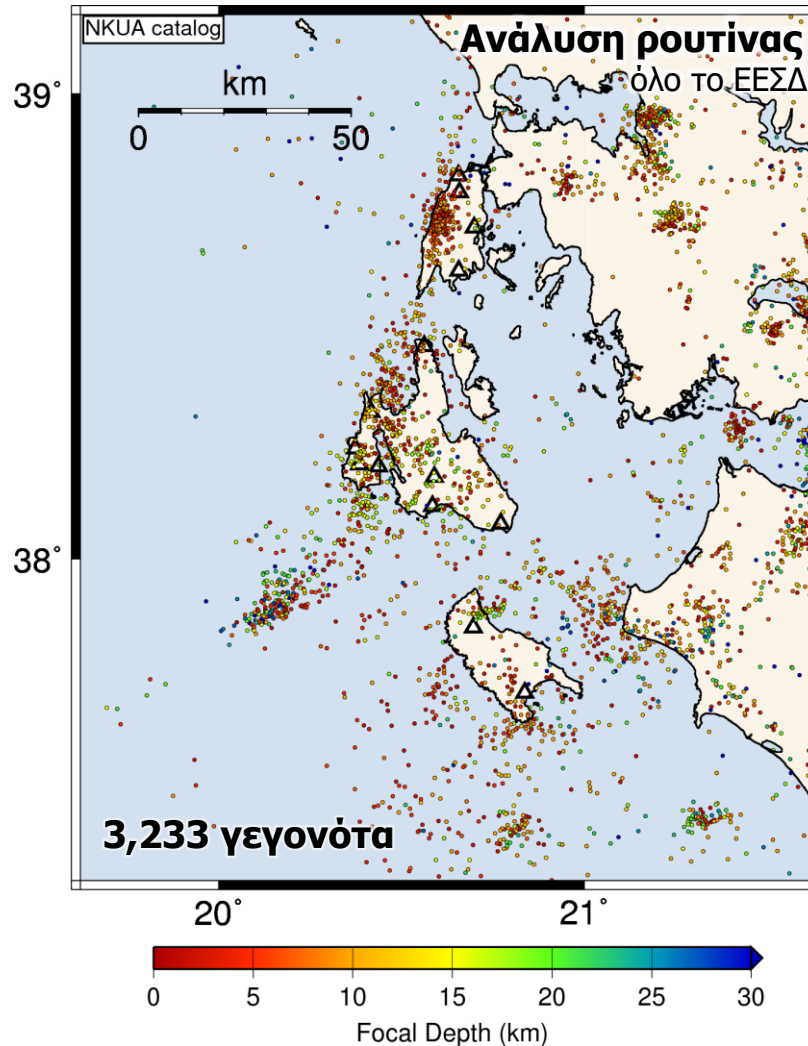
- Παράδειγμα συνδυασμού μοντέλου PhaseNet (για χρόνους άφιξης P και S) με μέθοδο GaMMA (για συσχετισμό φάσεων), σε ένα σετ δεδομένων 6 ημερών κατά τη σεισμική ακολουθία του 2019 στο Ridgecrest, Καλιφόρνια.



- Με τη συνδυαστική χρήση PhaseNet και GaMMA εντοπίστηκε υπερδιπλάσιος αριθμός σεισμών σε σύγκριση με τον κατάλογο του SCSN.

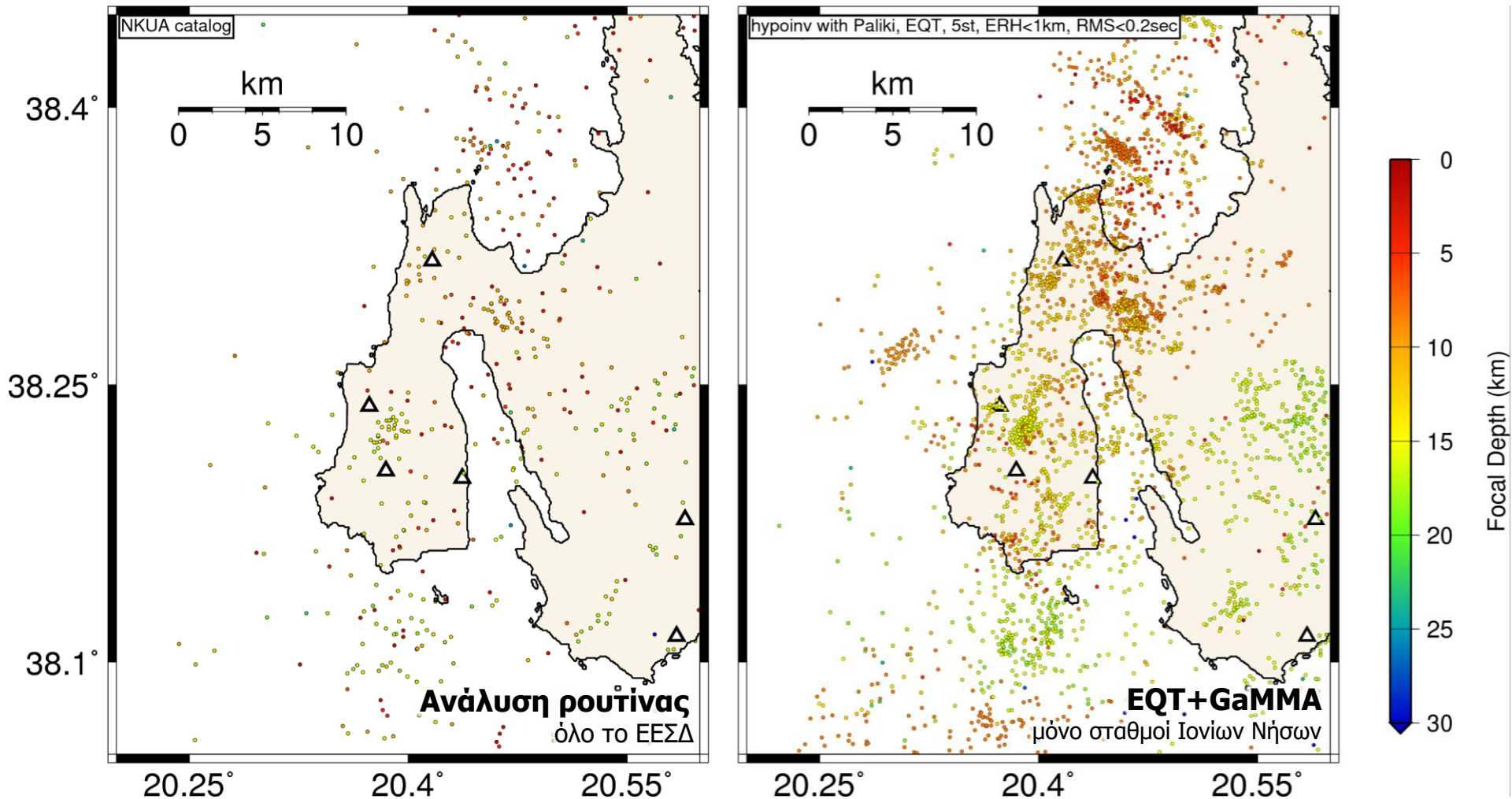
Αναγνώριση άφιξης σεισμικών κυμάτων P και S σε συνδυασμό με συσχετισμό φάσεων με σεισμικά γεγονότα: **EQTransformer + GaMMA**

- Εφαρμογή του μοντέλου EQTransformer και GaMMA σε δεδομένα σεισμολογικών σταθμών στα Ιόνια Νησιά κατά την περίοδο 14/4/2022 έως 30/6/2023.

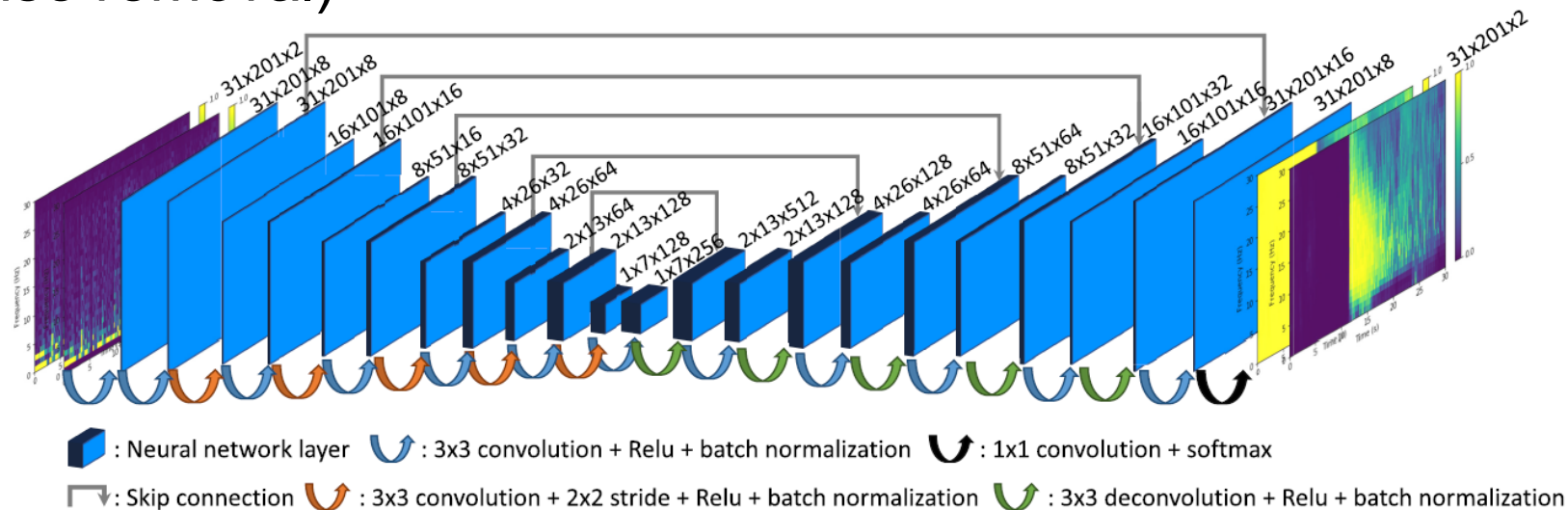


Αναγνώριση άφιξης σεισμικών κυμάτων P και S σε συνδυασμό με συσχετισμό φάσεων με σεισμικά γεγονότα: **EQTransformer + GaMMA**

- Εφαρμογή του μοντέλου EQTransformer και GaMMA σε δεδομένα σεισμολογικών σταθμών στα Ιόνια Νησιά κατά την περίοδο 14/4/2022 έως 30/6/2023.



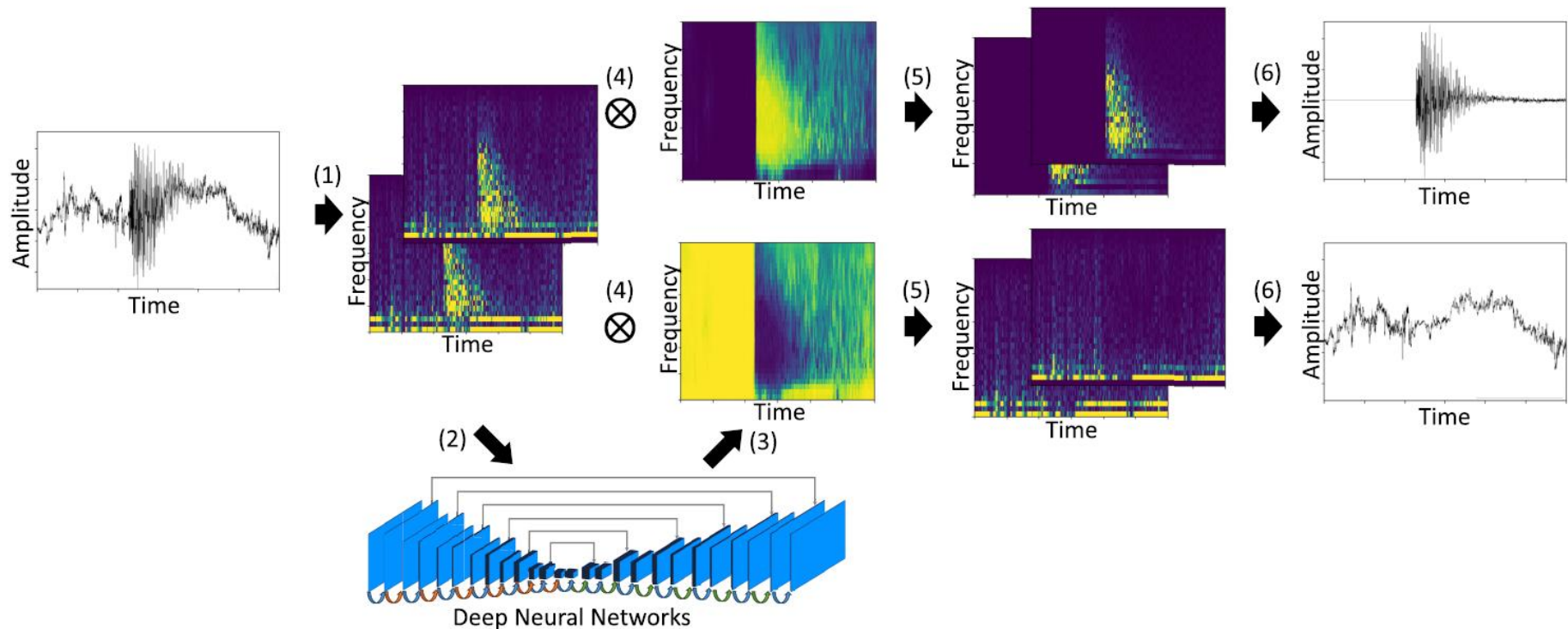
Διαχωρισμός σήματος από θόρυβο: DeepDenoiser (noise removal)



- Το μοντέλο βαθιάς μάθησης DeepDenoiser είναι φτιαγμένο για να διαχωρίζει το σεισμικό σήμα από τον θόρυβο υποβάθρου.
- Το μοντέλο αποτελείται από πολλαπλά επίπεδα συνέλιξης, κατά τα οποία γίνεται η αναγνώριση μοτίβων πάνω σε απεικονίσεις φασματογραφημάτων και σταδιακά μειώνεται η ανάλυση.
- Στη συνέχεια, ακολουθούν πολλαπλά επίπεδα αποσυνέλιξης, κατά τα οποία στην έξοδο επανασυντίθενται δύο απεικονίσεις: η μία για καθαρό θόρυβο και η δεύτερη για καθαρό σεισμικό σήμα (αν υπάρχει).

Διαχωρισμός σήματος από θόρυβο: DeepDenoiser (noise removal)

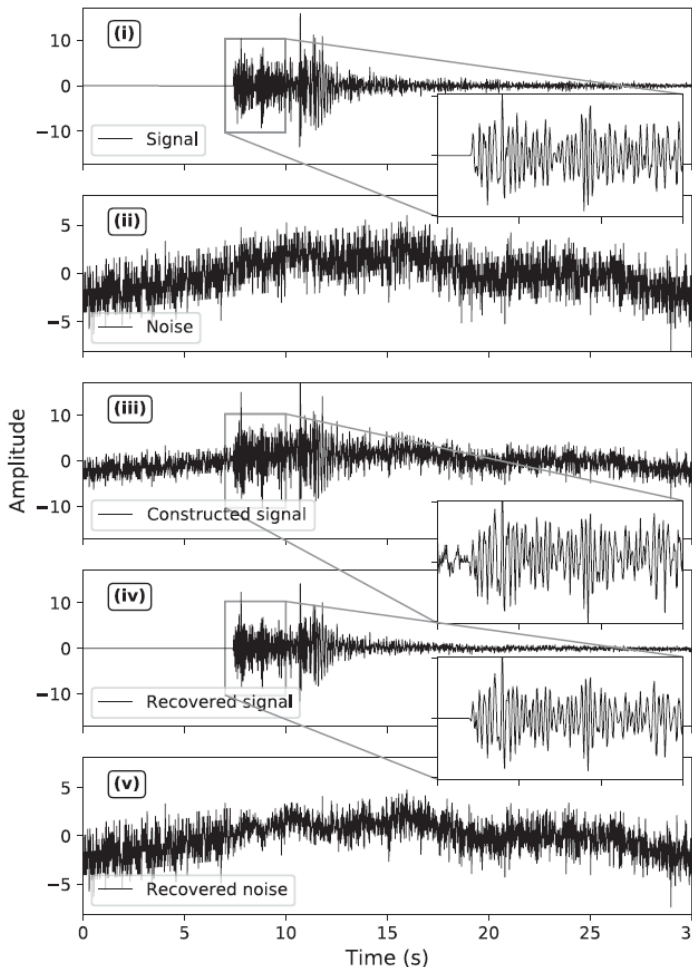
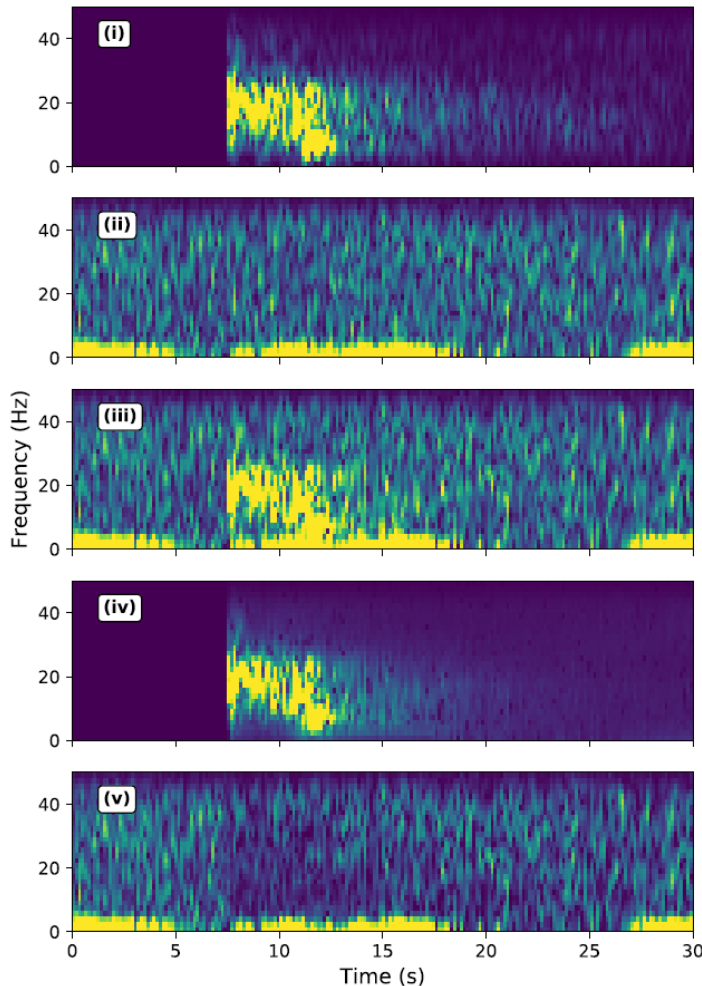
- Σχηματική λειτουργία του μοντέλου DeepDenoiser.



- Η αρχική καταγραφή στην είσοδο περιέχει σήμα και θόρυβο.
- Το μοντέλο δημιουργεί μια «μάσκα» στο φασματογράφημα, η οποία διακρίνει το σεισμικό σήμα από τον θόρυβο.
- Στην έξοδο δημιουργούνται δύο χρονοσειρές, μία για το σεισμικό σήμα και μια για τον θόρυβο.

Διαχωρισμός σήματος από θόρυβο: **DeepDenoiser** (noise removal)

- Παράδειγμα εφαρμογής του μοντέλου DeepDenoiser.



Αρχικό «καθαρό»
σεισμικό σήμα
(υψηλού SNR)

Αρχικός «καθαρός»
θόρυβος
(δεν περιέχει σεισμό)

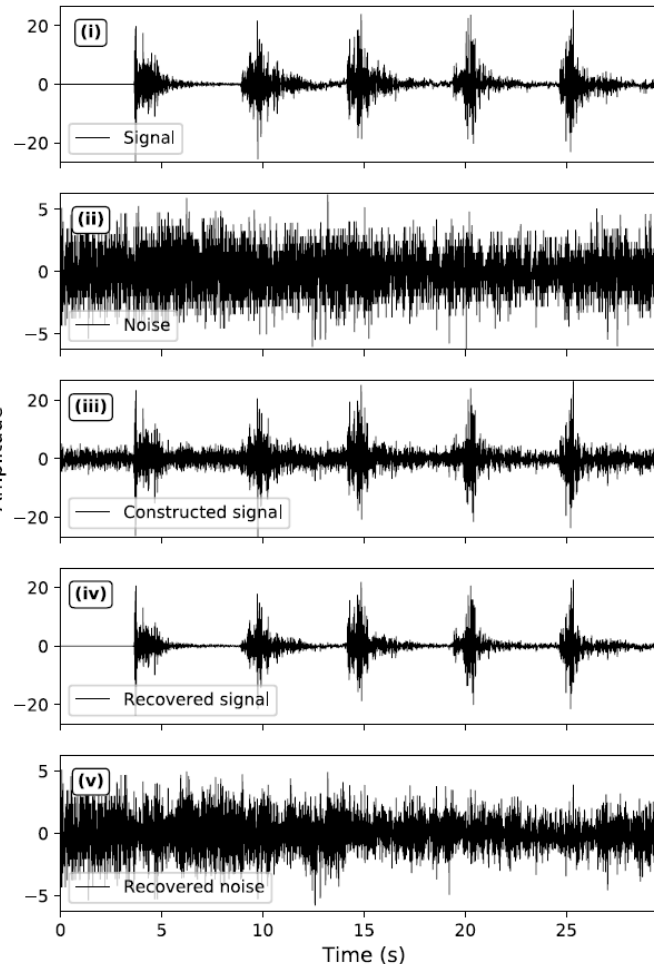
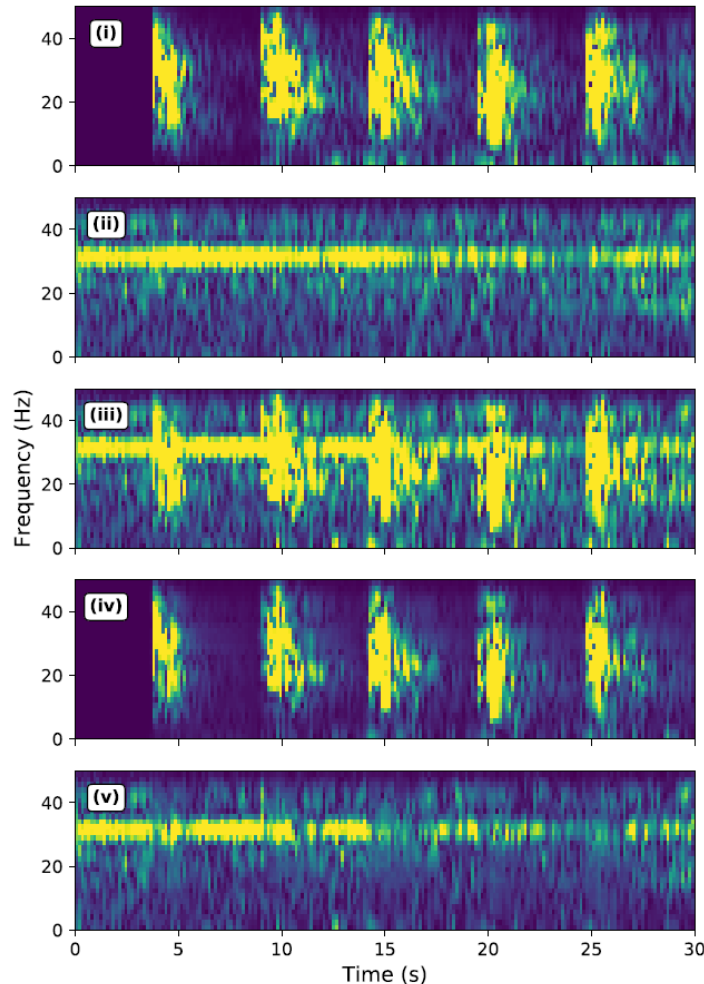
Υπέρθωση (i) και (ii)
(«συνθετικό» σήμα με
θόρυβο)

Έξοδος DeepDenoiser:
διαχωρισμένο σεισμικό
σήμα

Έξοδος DeepDenoiser:
διαχωρισμένος
θόρυβος

Διαχωρισμός σήματος από θόρυβο: **DeepDenoiser** (noise removal)

- Παράδειγμα εφαρμογής του μοντέλου DeepDenoiser.



Αρχικό «καθαρό»
σεισμικό σήμα
(υψηλού SNR)

Αρχικός «καθαρός»
θόρυβος
(δεν περιέχει σεισμό)

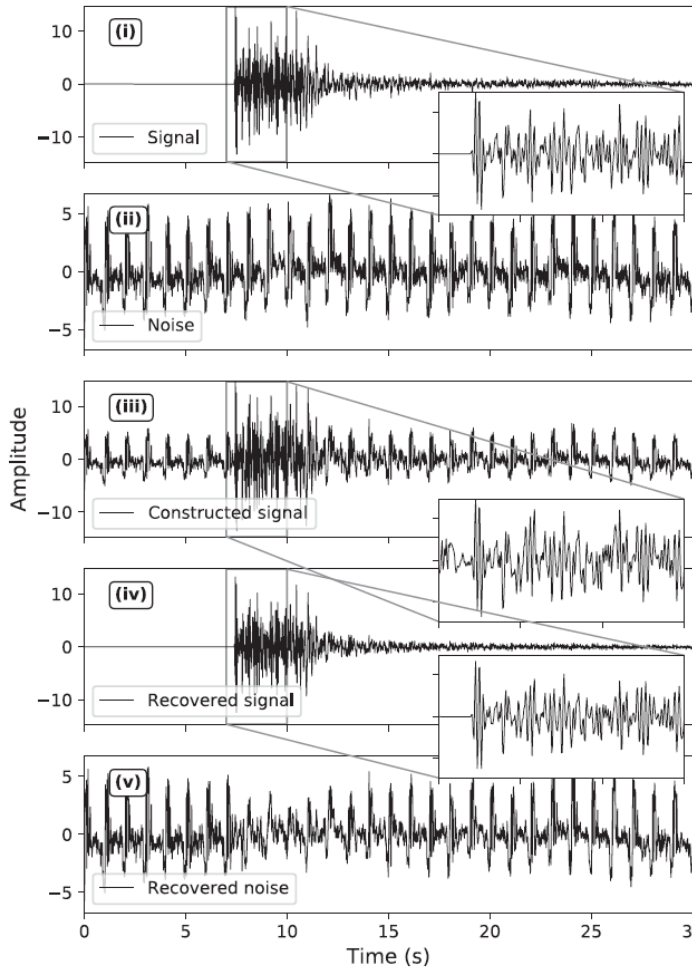
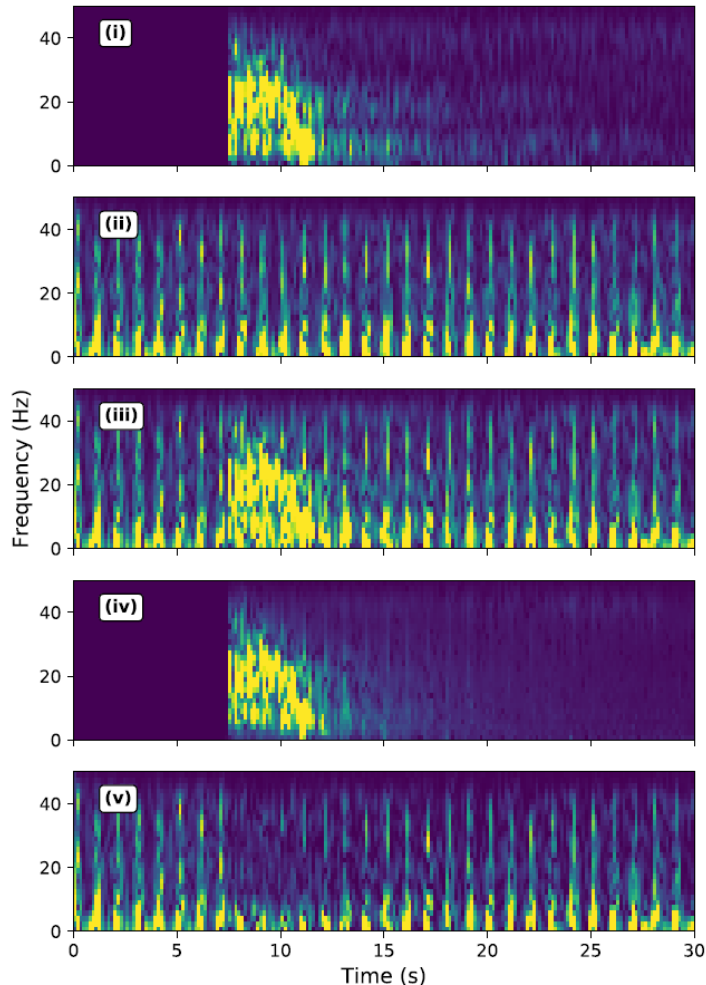
Υπέρθεση (i) και (ii)
(«συνθετικό» σήμα με
θόρυβο)

Έξοδος DeepDenoiser:
διαχωρισμένο σεισμικό
σήμα

Έξοδος DeepDenoiser:
διαχωρισμένος
θόρυβος

Διαχωρισμός σήματος από θόρυβο: **DeepDenoiser** (noise removal)

- Παράδειγμα εφαρμογής του μοντέλου DeepDenoiser.



Αρχικό «καθαρό»
σεισμικό σήμα
(υψηλού SNR)

Αρχικός «καθαρός»
θόρυβος
(δεν περιέχει σεισμό)

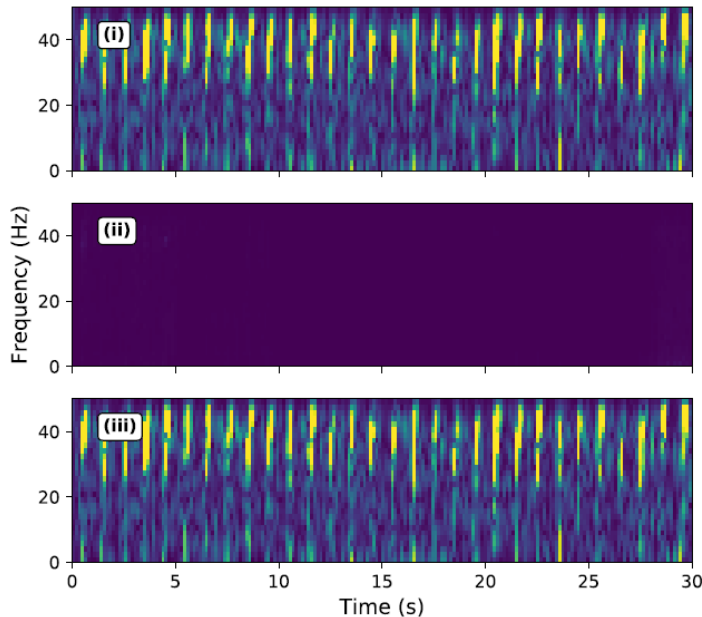
Υπέρθωση (i) και (ii)
(«συνθετικό» σήμα με
θόρυβο)

Έξοδος DeepDenoiser:
διαχωρισμένο σεισμικό
σήμα

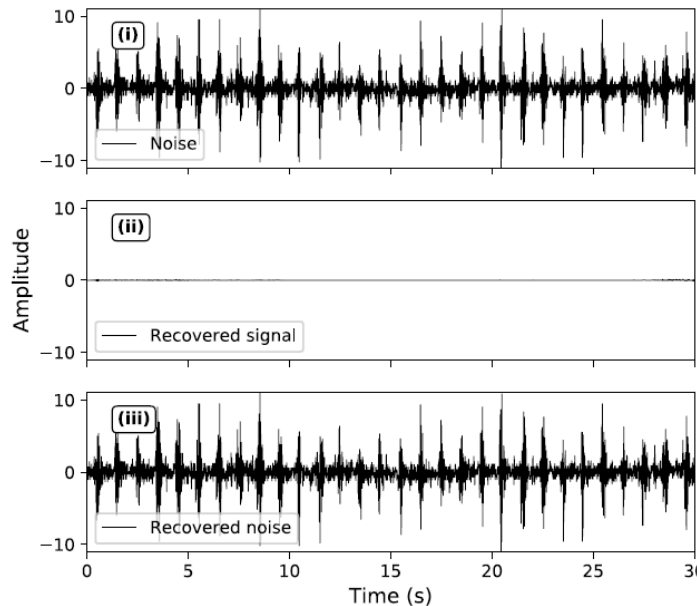
Έξοδος DeepDenoiser:
διαχωρισμένος
θόρυβος

Διαχωρισμός σήματος από θόρυβο: **DeepDenoiser** (noise removal)

- Παράδειγμα εφαρμογής του μοντέλου DeepDenoiser σε καταγραφή με σκέτο θόρυβο.



(c)



(d)

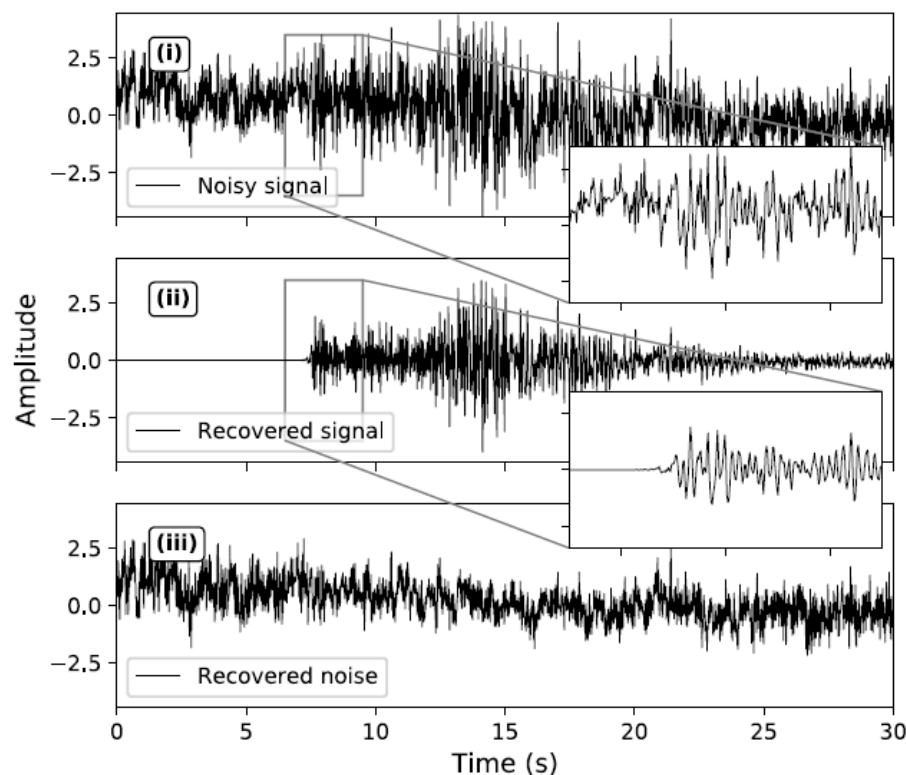
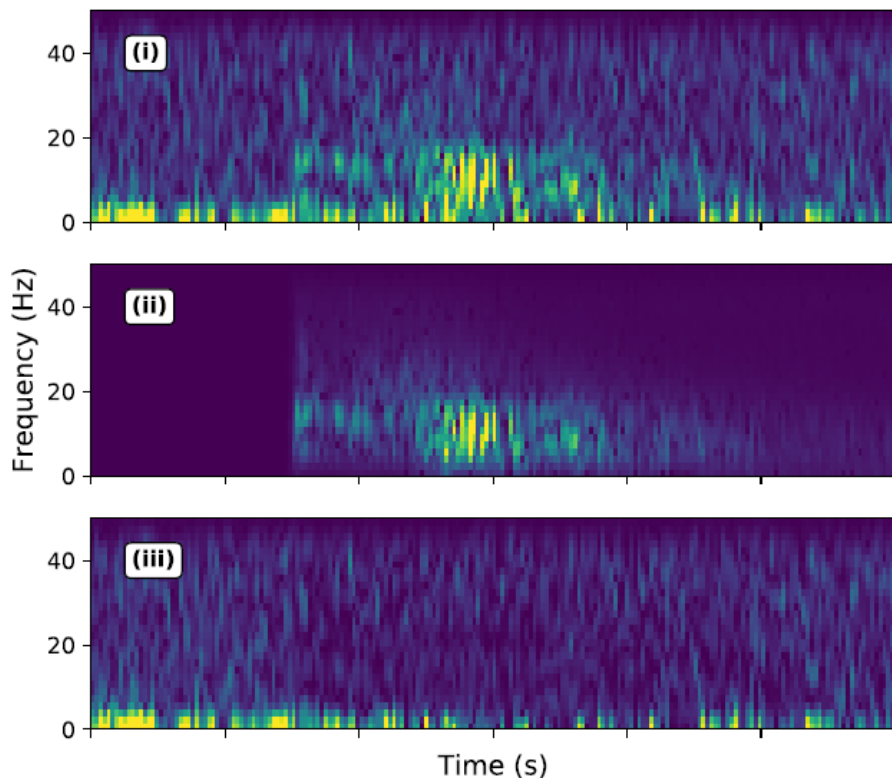
Αρχικός «καθαρός»
θόρυβος
(δεν περιέχει σεισμό)

Έξοδος DeepDenoiser:
διαχωρισμένο σεισμικό
σήμα (κενό)

Έξοδος DeepDenoiser:
διαχωρισμένος
θόρυβος

Διαχωρισμός σήματος από θόρυβο: **DeepDenoiser** (noise removal)

- Παράδειγμα εφαρμογής του μοντέλου DeepDenoiser σε καταγραφή σεισμού χαμηλού SNR, με πραγματικό θόρυβο.

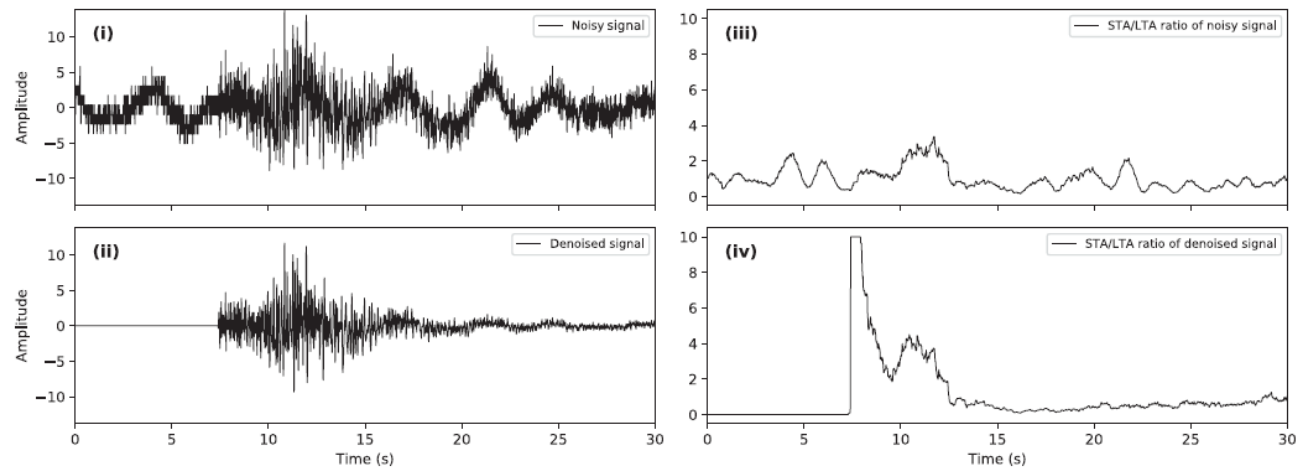


- Το διαχωρισμένο σήμα είναι ανακατασκευή από το μοντέλο με βάση τα χαρακτηριστικά που (θεωρεί πως) αναγνώρισε, και όχι αποτέλεσμα εφαρμογής κάποιου κλασικού φίλτρου (π.χ., Butterworth).

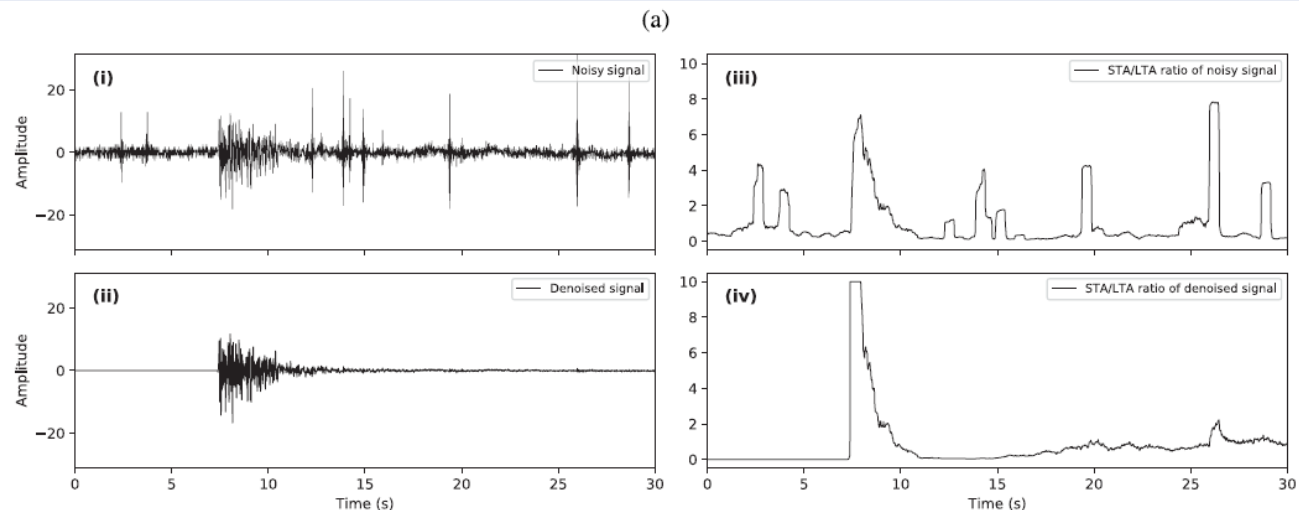
Διαχωρισμός σήματος από θόρυβο: DeepDenoiser (noise removal)

- Παραδείγματα εφαρμογής του μοντέλου DeepDenoiser και βελτίωσης του SNR όπως προκύπτει από την εφαρμογή της μεθόδου STA/LTA.

- Αρχική, θορυβώδης καταγραφή
- Διαχωρισμένο σήμα με DeepDenoiser



- Αρχική, θορυβώδης καταγραφή
- Διαχωρισμένο σήμα με DeepDenoiser



Λογισμικό SeisBench

- Το λογισμικό ανοιχτού κώδικα **SeisBench** (σε γλώσσα Python) είναι μια πλατφόρμα η οποία συγκεντρώνει πολλά διαφορετικά μοντέλα βαθιάς μάθησης, π.χ. για τον προσδιορισμό του χρόνου άφιξης των κυμάτων P και S.
- Στα διαθέσιμα μοντέλα του SeisBench περιλαμβάνονται τα PhaseNet, EQTransformer, CRED, GPD και DeepDenoiser.

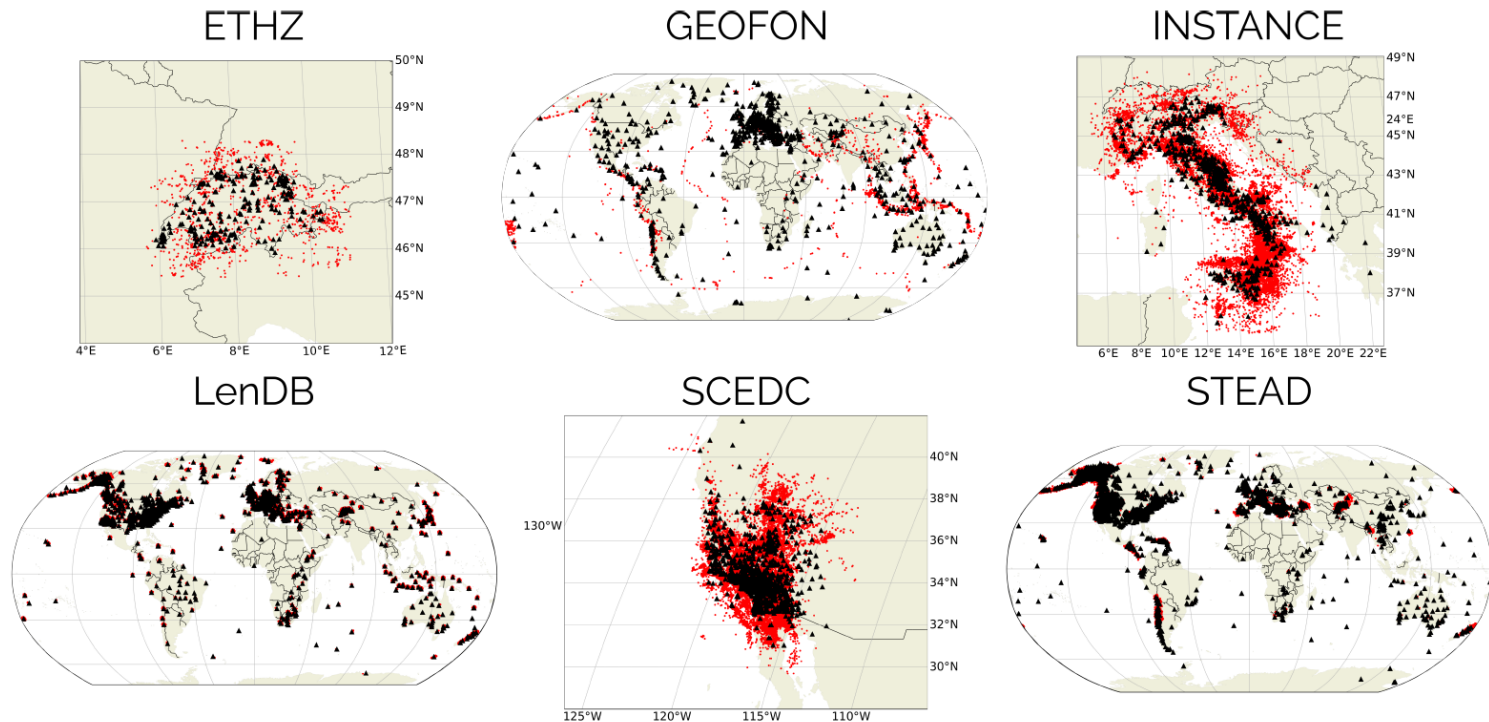


Integrated models	Task
BasicPhaseAE	Phase Picking
CRED	Earthquake Detection
DPP	Phase Picking
DepthPhaseNet	Depth estimation from depth phases
DepthPhaseTEAM	Depth estimation from depth phases
DeepDenoiser	Denoising
EQTransformer	Earthquake Detection/Phase Picking
GPD	Phase Picking
LFEDetect	Phase Picking (Low-frequency earthquakes)
PhaseNet	Phase Picking
PhaseNetLight	Phase Picking
PickBlue	Earthquake Detection/Phase Picking
OBSTransformer	Earthquake Detection/Phase Picking
VariableLengthPhaseNet	Phase Picking

<https://seisbench.readthedocs.io/en/stable/index.html>

Λογισμικό SeisBench

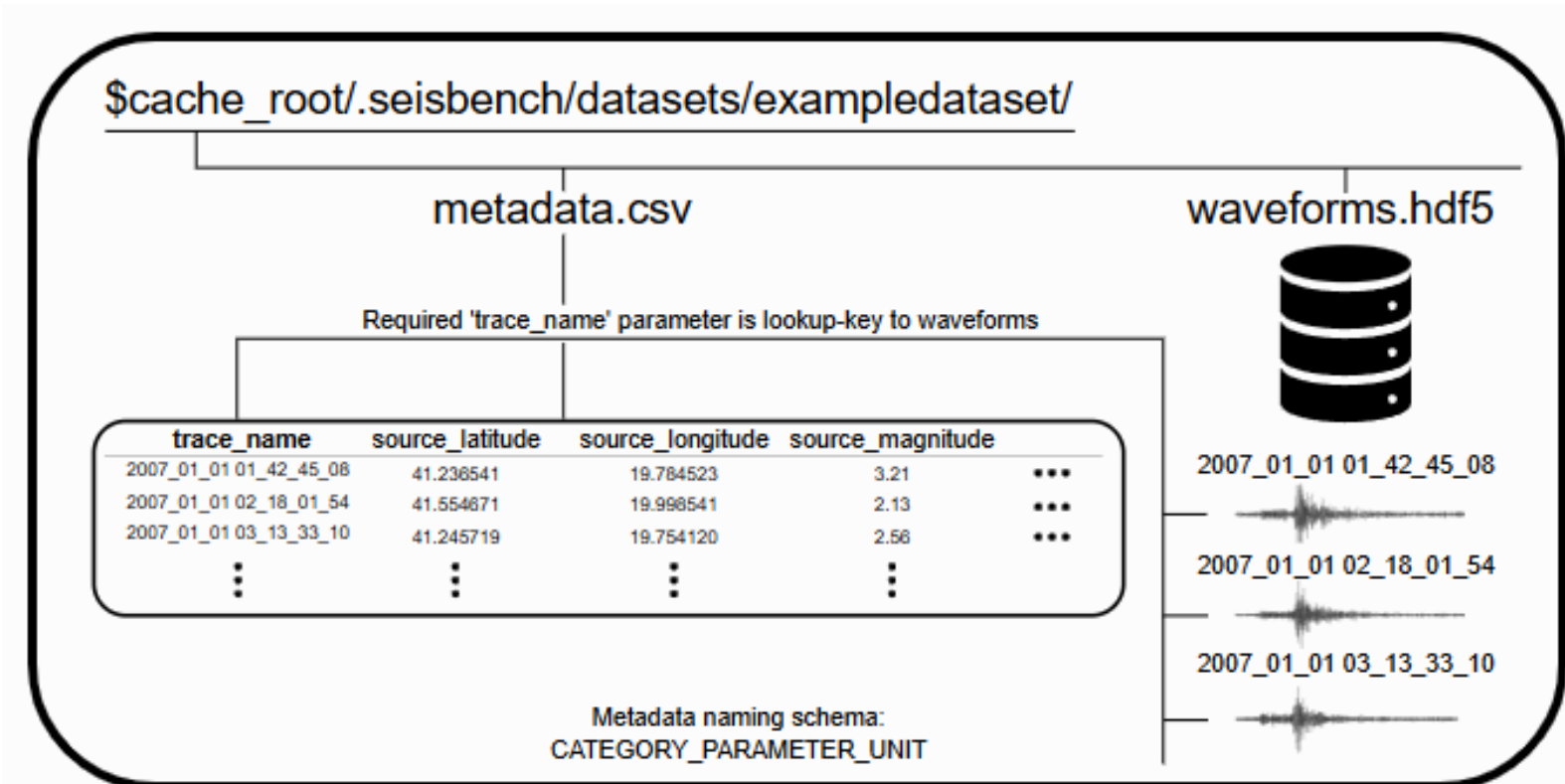
- Τα μοντέλα του SeisBench διατίθενται με προ-εκπαιδευμένα βάρη πάνω σε διαφορετικά σετ επισημασμένων δεδομένων, τα οποία μπορεί να επιλέξει ο χρήστης.



	Traces	Events	P picks	S picks	Noise	Region	Dist	Tr. length	f_s [Hz]
ETHZ	36,743	2,231	35,227	18,960	0	Switzerland	L/R	variable	100 - 500
INSTANCE	1,291,537	54,008	1,159,249	713,883	132,288	Italy	L/R	120 s	100
Iquique	13,400	409	13,327	11,361	0	N. Chile	L	variable	100
LenDB	1,244,942	303,902	629,095	0	615,847	various	L	27 s	20
SCEDC	8,111,060	378,528	7,571,970	4,364,155	0	S. California	L	variable	40 - 100
STEAD	1,265,657	441,705	1,030,231	1,030,231	235,426	various	L/R	60 s	100
GEOFON	275,274	2,270	284,240	2,847	0	global	R/T	variable	20 - 200
NEIC	1,354,789	137,424	1,025,000	329,789	0	global	R/T	60 s	40

Λογισμικό SeisBench

- Μέσω του SeisBench είναι επίσης δυνατή η λήψη ολόκληρων πρότυπων προσημασμένων σετ σεισμολογικών δεδομένων (π.χ., STEAD, INSTANCE, κτλ).
- Τα πρότυπα επισημασμένων δεδομένων (labeled datasets) περιλαμβάνουν τις κυματομορφές σε .hdf5 format και τα συνοδευτικά μεταδεδομένα (εστιακές παραμέτρους, χρόνους άφιξης P και S, κτλ.) σε .csv format.



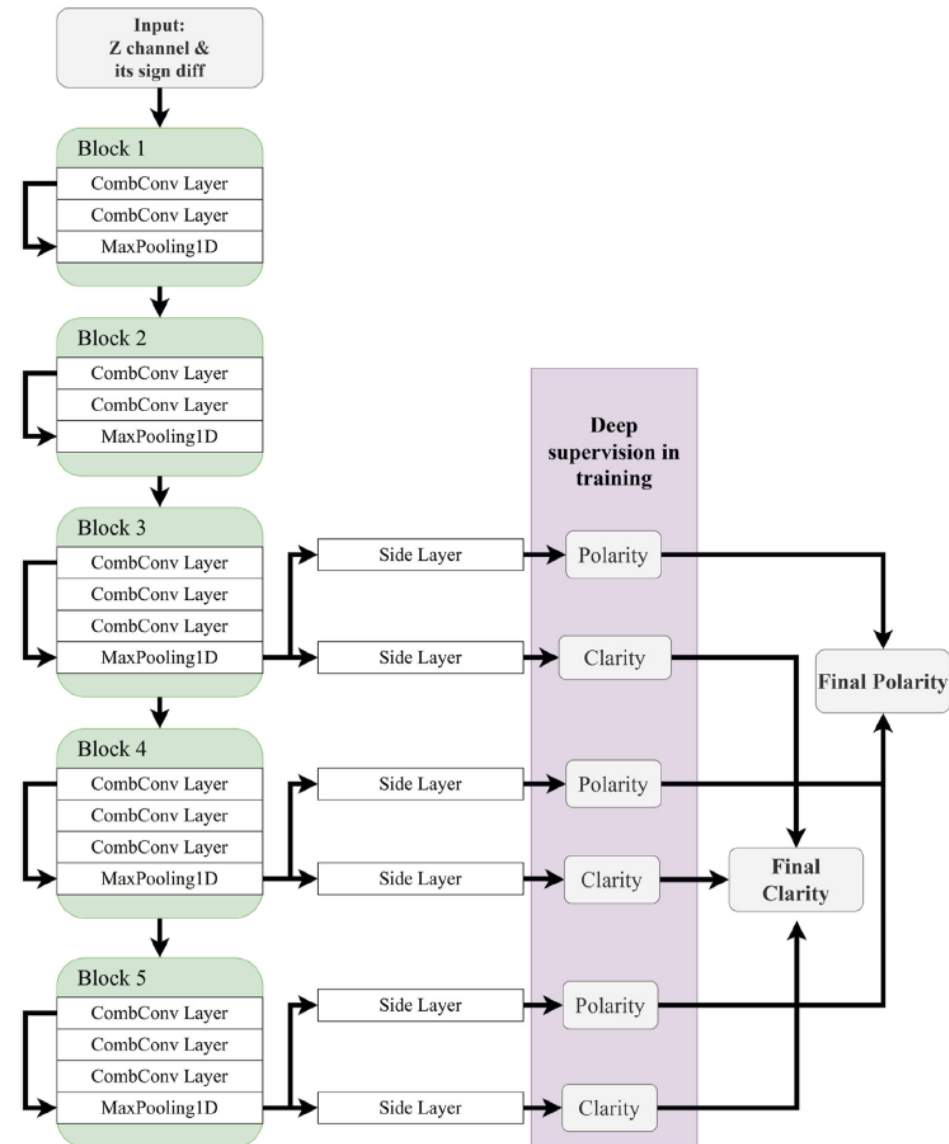
Μεταβιβαστική μάθηση

(transfer learning)

- Το SeisBench παρέχει επίσης δυνατότητα «μεταβιβαστικής μάθησης» για τα μοντέλα του.
- Προκειμένου να διευκολύνουμε τη γενίκευση ενός **ήδη** εκπαιδευμένου μοντέλου και να μπορέσουμε να το προσαρμόσουμε σε δεδομένα με τα οποία δεν έχει εκπαιδευτεί, μπορούμε να εφαρμόσουμε τη διαδικασία της **μεταβιβαστικής μάθησης** (transfer learning).
- Η διαδικασία αυτή είναι παρόμοια με την αρχική εκπαίδευση του μοντέλου, με τις παρακάτω διαφορές:
 - Δεν ξεκινάμε με τυχαία βάρη, αλλά χρησιμοποιούμε ως αρχικά τα βάρη του ήδη εκπαιδευμένου μοντέλου (με δεδομένα άλλης περιοχής).
 - Εκτελούμε την εκπαίδευση του μοντέλου (π.χ. με μέθοδο οπισθο-διάδοσης σφάλματος) χρησιμοποιώντας επισημασμένα δεδομένα από την συγκεκριμένη περιοχή που θέλουμε να μελετήσουμε.
 - Χρησιμοποιούμε διασταυρωμένη επικύρωση για να αποφύγουμε υπερπροσαρμογή.
- Η παραπάνω διαδικασία είναι χρήσιμη όταν δεν υπάρχουν επαρκή επισημασμένα δεδομένα για να εκπαιδεύσουμε από το μηδέν ένα μοντέλο, και ο σκοπός είναι απλώς να κάνουμε μικρο-ρυθμίσεις στα βάρη για να βελτιώσουμε τις επιδόσεις του προ-εκπαιδευμένου μοντέλου στη νέα περιοχή μελέτης.

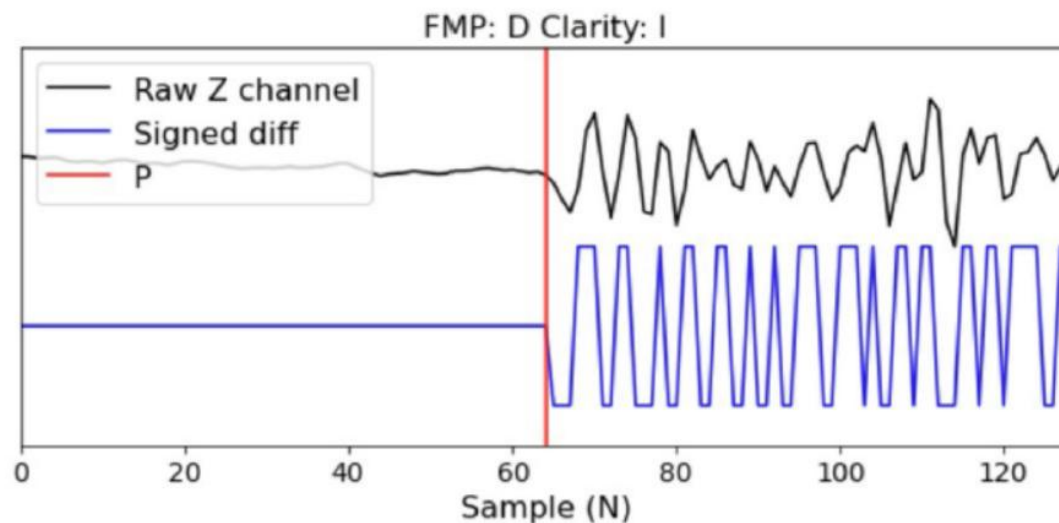
Αναγνώριση πολικότητας πρώτης απόκλισης επιμήκων κυμάτων: DiTingMotion (first motion polarity classification)

- Το μοντέλο βαθιάς μάθησης DiTingMotion είναι σχεδιασμένο για να αναγνωρίζει την πολικότητα της πρώτης απόκλισης των κυμάτων P στην κατακόρυφη συνιστώσα.
- Είναι εκπαιδευμένο με επισημασμένα σετ δεδομένων από τη Νότια Καλιφόρνια (SCSN-FMP σετ) και από την Κίνα (DiTing σετ).
- Η αρχιτεκτονική του στηρίζεται σε μια σειρά από επίπεδα συνέλιξης για την αναγνώριση του μοτίβου της πρώτης απόκλισης.
- Η έξοδος του ταξινομεί το σήμα ως προς την πολικότητα της πρώτης απόκλισης σε «**συμπίεση**» ή «**αραίωση**», καθώς και στον χαρακτηρισμό της ευκρίνειας του πρώτου παλμού σε «**κρουστικό**» (impulsive) ή «**αναδυόμενο**» (emergent).
- Οι μετρήσεις αυτές χρησιμοποιούνται για τον προσδιορισμό του μηχανισμού γένεσης σεισμών μικρού μεγέθους.

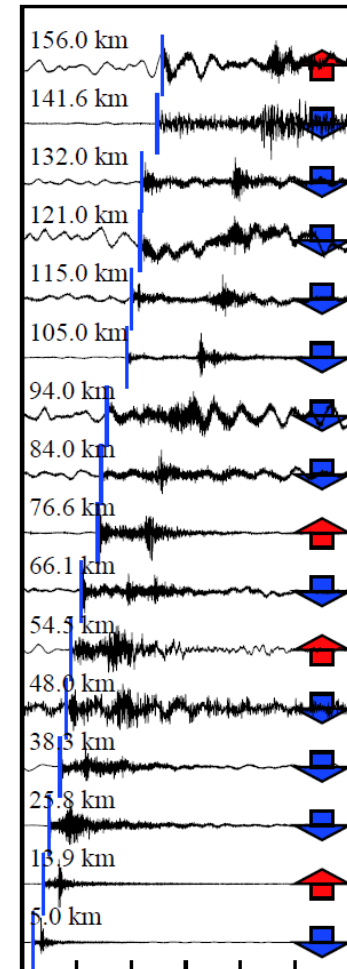


Αναγνώριση πολικότητας πρώτης απόκλισης επιμήκων κυμάτων: DiTingMotion (first motion polarity classification)

- Τα επισημασμένα σετ δεδομένων που χρησιμοποιούνται για την εκπαίδευση τέτοιων μοντέλων είναι απαραίτητο να διαθέτουν:
 - Σεισμικές καταγραφές στην κατακόρυφη συνιστώσα.
 - Χρόνο άφιξης των κυμάτων P, προσδιορισμένο με υψηλή ακρίβεια.
 - Χαρακτηρισμό της πρώτης απόκλισης των P ως προς την πολικότητα (C/D) και την ευκρίνεια (I/E).
 - Παραδείγματα καταγραφών στις οποίες **δεν** είναι δυνατός ο προσδιορισμός της πρώτης απόκλισης (ασαφής παλμός ή χαμηλός λόγος SNR), και πρέπει να **μην** χαρακτηριστούν ως C ή D.



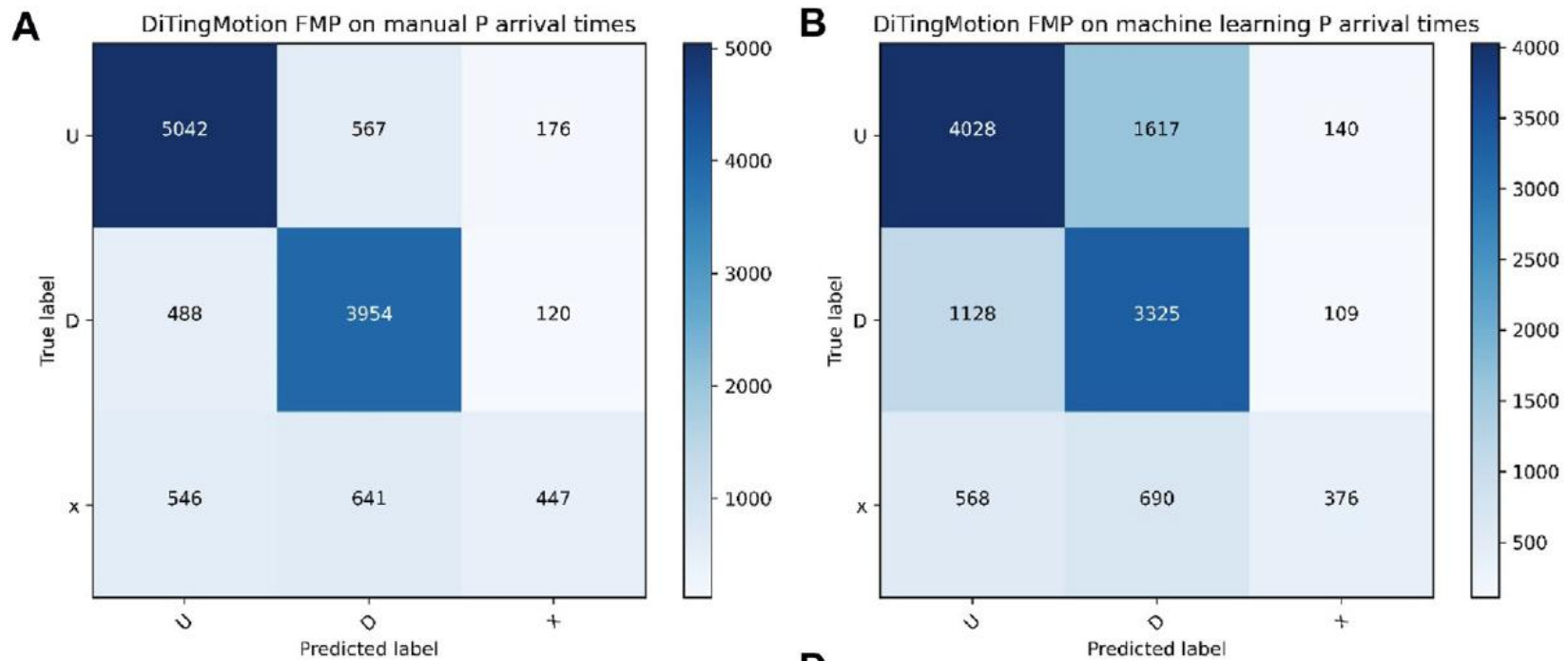
(a) Z channel



Time (s)

Αναγνώριση πολικότητας πρώτης απόκλισης επιμήκων κυμάτων: DiTingMotion (first motion polarity classification)

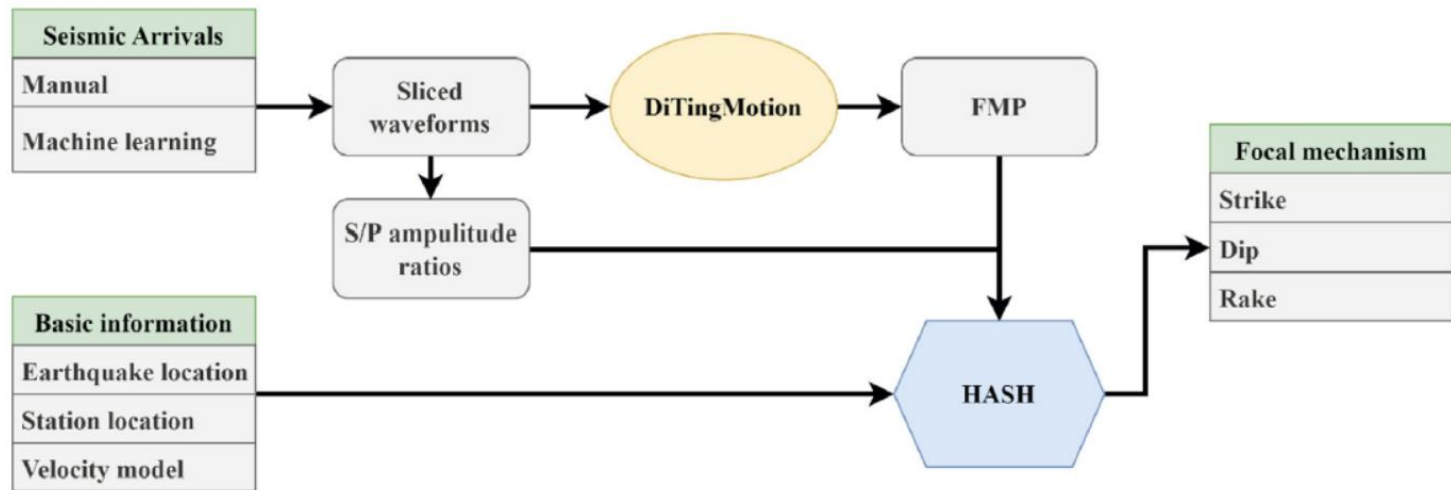
- Το μοντέλο αυτό δεν μετράει χρόνους άφιξης, παρά μόνο πραγματοποιεί ταξινόμηση στις προαναφερθείσες κλάσεις.
- Για τον έλεγχο των επιδόσεων, χρησιμοποιούνται πίνακες «σύγχυσης» (Confusion Matrix).



- Εξετάζονται τα ποσοστά **επιτυχημένων** ή **εσφαλμένων** προβλέψεων από το μοντέλο, συγκριτικά με τα επισημασμένα δεδομένα, τα ποσοστά **χαμένων** μετρήσεων (χωρίς χαρακτηρισμό) και τα ποσοστά **νέων** μετρήσεων (που δεν υπήρχαν στα επισημασμένα δεδομένα).

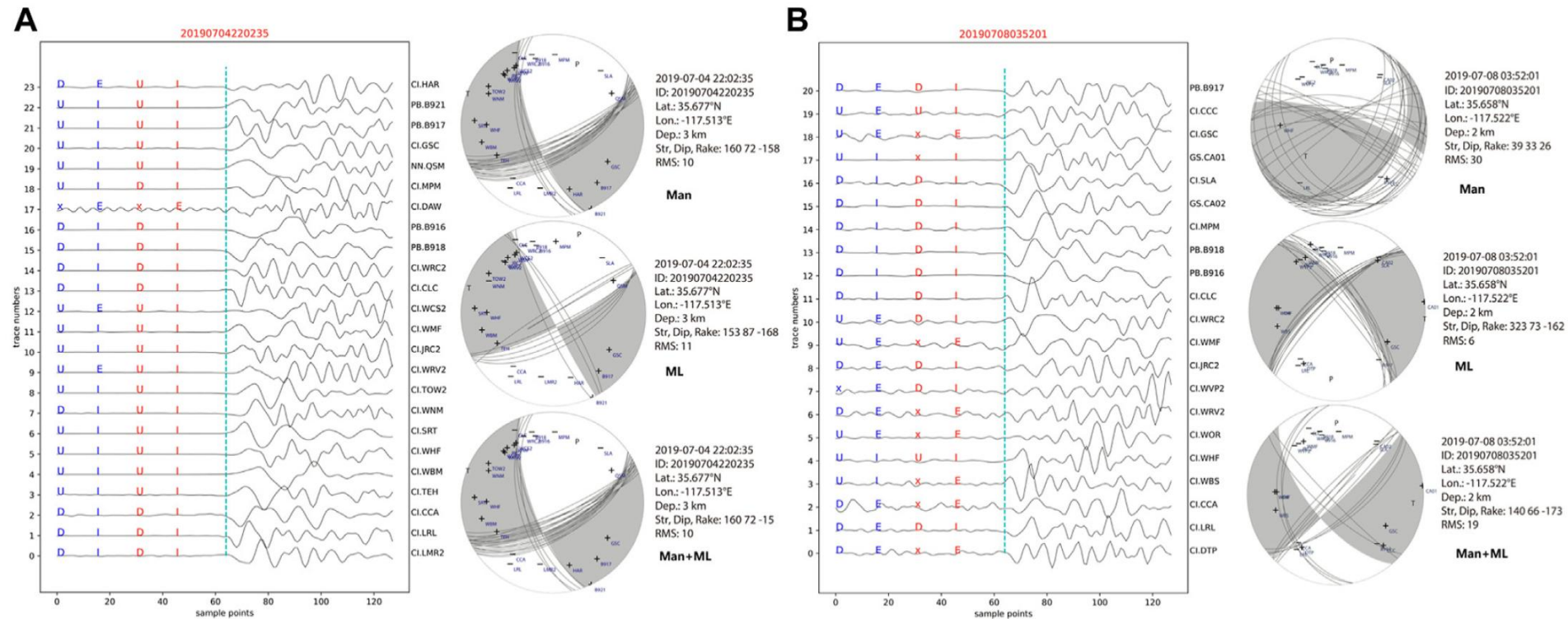
Αναγνώριση πολικότητας πρώτης απόκλισης επιμήκων κυμάτων: **DiTingMotion** (first motion polarity classification)

- Το μοντέλο αυτό μπορεί να μπει σε μια ροή εργασιών η οποία περιλαμβάνει τα εξής στάδια:
 - Μέτρηση χρόνων άφιξης (χειρακτική ή μέσω κάποιου μοντέλου βαθιάς μάθησης).
 - Αποκοπή κυματομορφών της κατακόρυφης συνιστώσας γύρω από την άφιξη των κυμάτων P σε κάθε σταθμό.
 - Εφαρμογή μοντέλου DiTingMotion για εξαγωγή χαρακτηρισμών της πρώτης απόκλισης των κυμάτων P (First Motion Polarities, ή FMP).
 - Συνδυασμός FMP με εστιακές παραμέτρους, θέσεις σταθμών, μοντέλο ταχυτήτων (και, προαιρετικά, λόγους πλατών S/P σε κάθε σταθμό).
 - Χρήση λογισμικού (π.χ., HASH) για αυτόματο προσδιορισμό μηχανισμού γένεσης (παράταξη, κλίση, γωνία ολίσθησης).



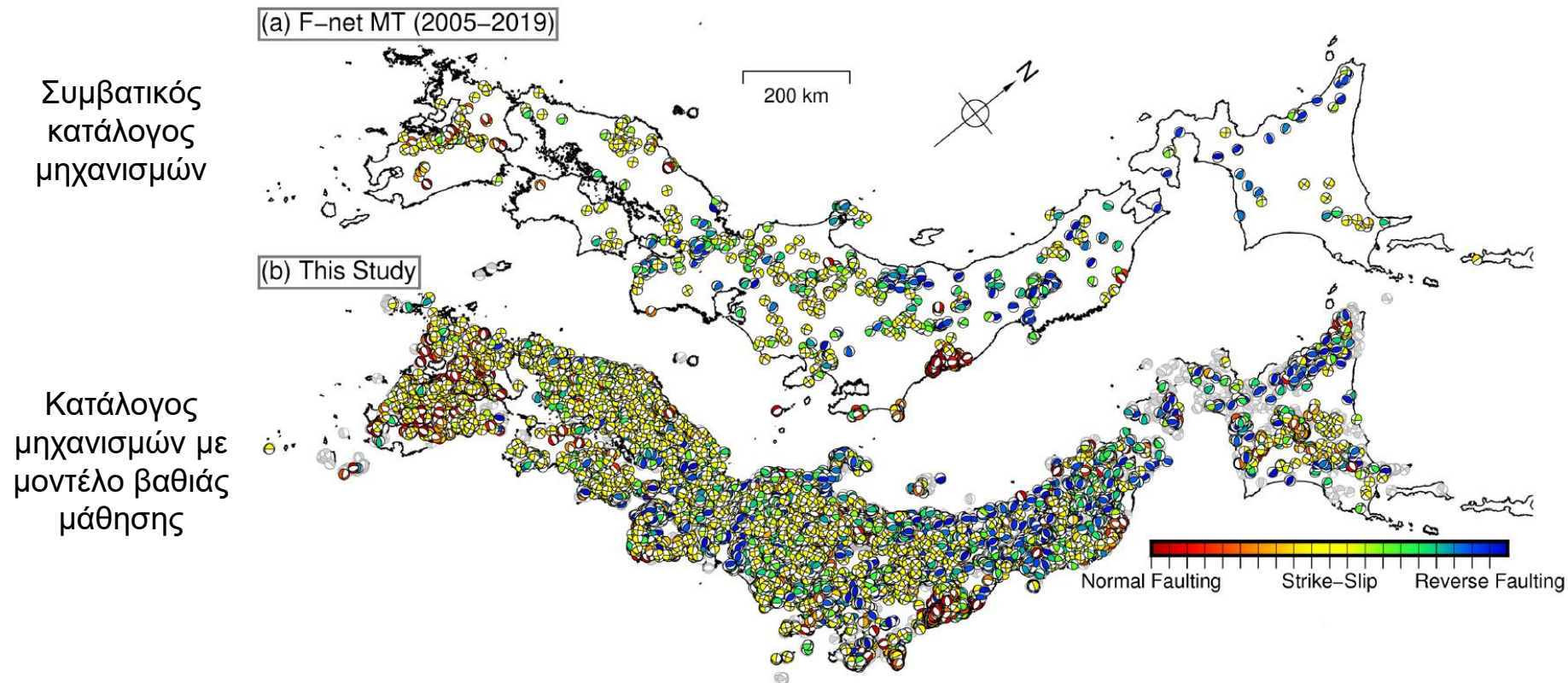
Αναγνώριση πολικότητας πρώτης απόκλισης επιμήκων κυμάτων: DiTingMotion (first motion polarity classification)

- Παράδειγμα εφαρμογής μοντέλου DiTingMotion σε ροή εργασιών η οποία καταλήγει σε αυτόματο υπολογισμό μηχανισμού γένεσης με το λογισμικό HASH.



Αναγνώριση πολικότητας πρώτης απόκλισης επιμήκων κυμάτων (first motion polarity classification)

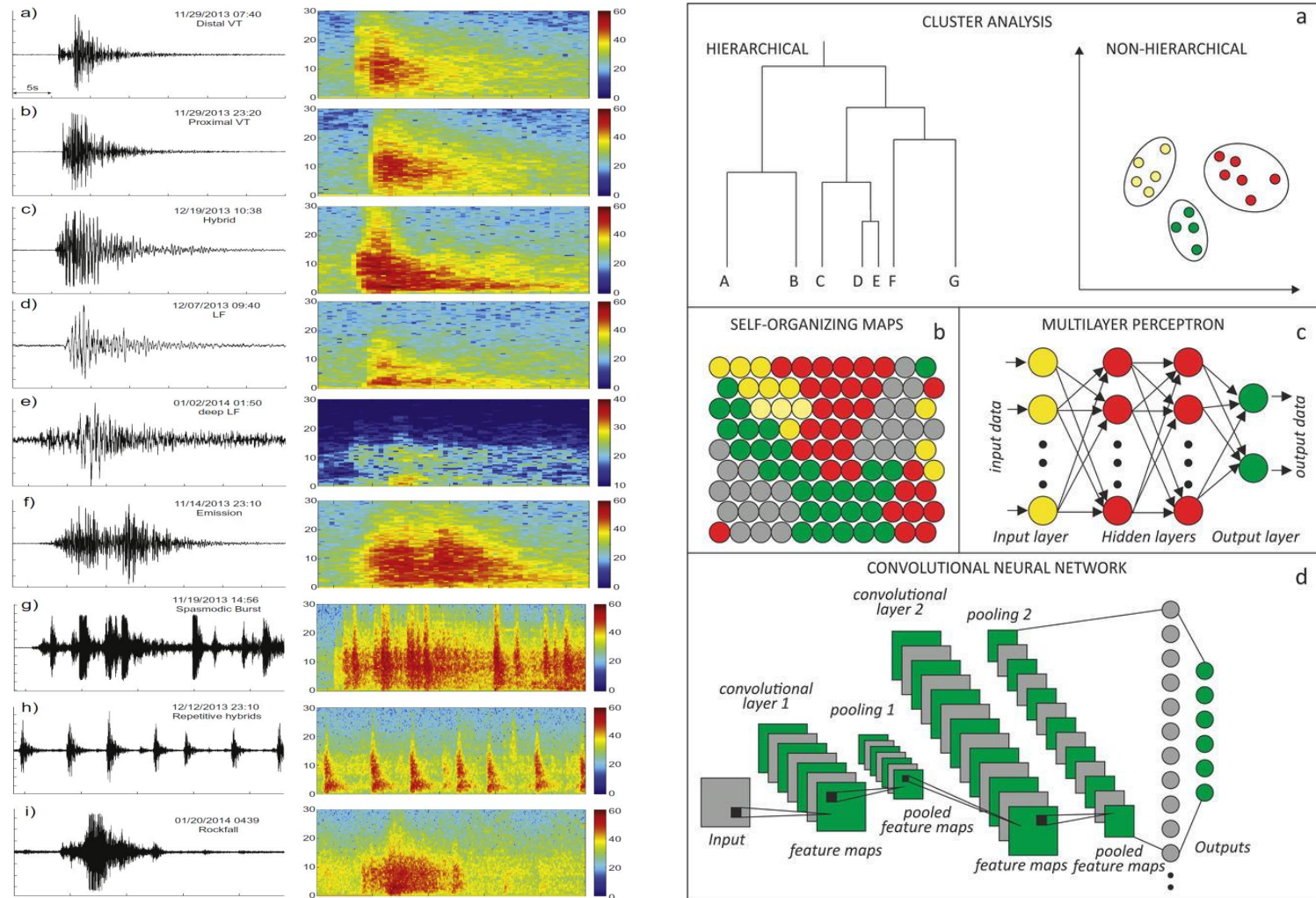
- Παράδειγμα εφαρμογής μοντέλου βαθιάς μάθησης για χαρακτηρισμό πρώτων αποκλίσεων και του λογισμικού HASH για προσδιορισμό μηχανισμών γένεσης σεισμών στην Ιαπωνία (2005-2019).



- Ο κατάλογος που δημιουργήθηκε με το μοντέλο βαθιάς μάθησης (κάτω) περιέχει ~110,000 μηχανισμούς γένεσης σεισμών, χρησιμοποιώντας ένα σύνολο ~1,800,000 πολικοτήτων πρώτων αποκλίσεων.

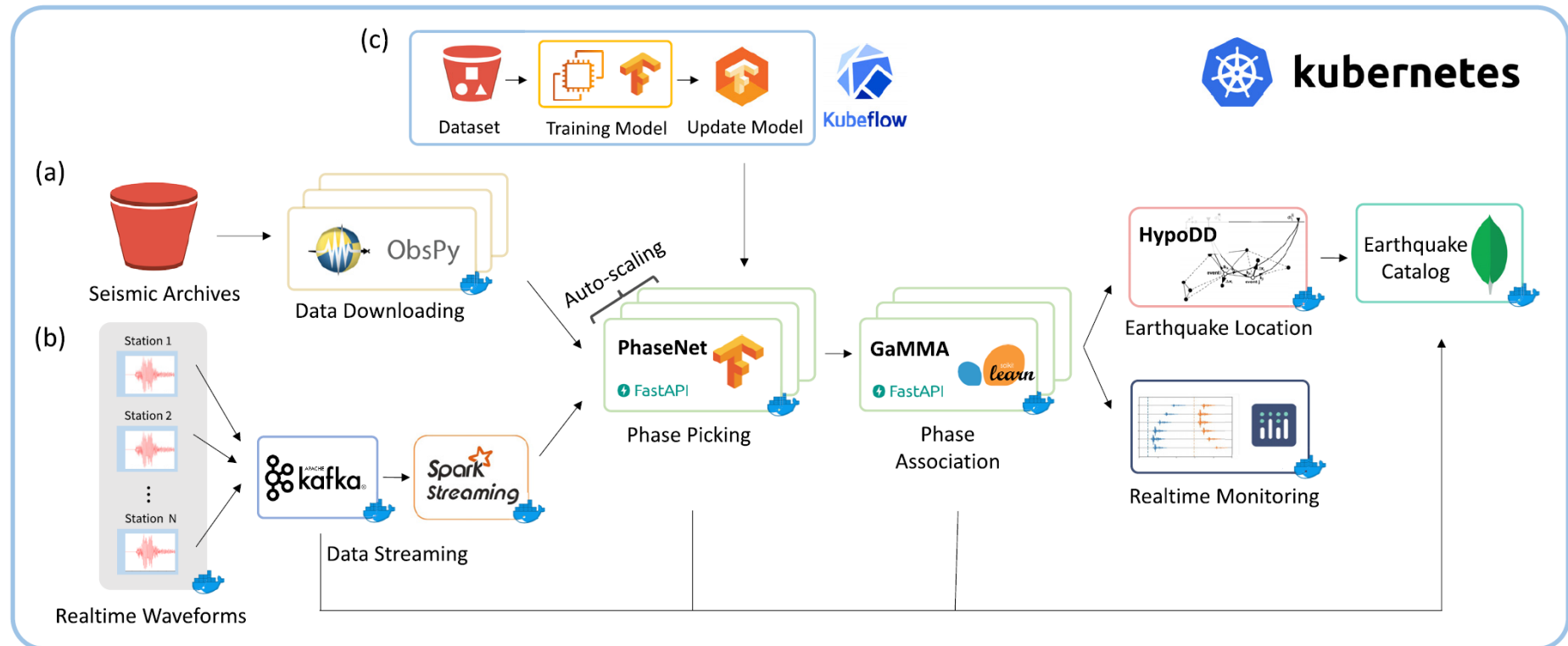
Αναγνώριση τύπου σεισμικών σημάτων (seismic signal classification)

- Μέθοδοι επιβλεπόμενης (και μη) μηχανικής μάθησης βρίσκουν εφαρμογή στον χαρακτηρισμό και την κατηγοριοποίηση σεισμικών σημάτων, ιδιαίτερα σε ηφαιστειακά περιβάλλοντα.



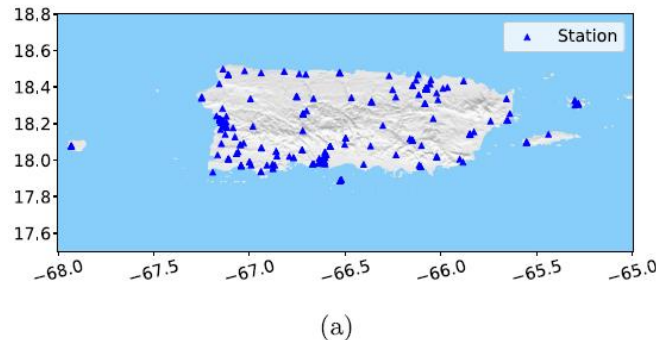
Πλήρως αυτοματοποιημένη ροή εργασιών: **QuakeFlow**

- Η εφαρμογή μοντέλων μηχανικής μάθησης μπορεί να συνδυαστεί με κλασικούς αλγορίθμους σε πλήρως αυτοματοποιημένες ροές εργασιών (workflows).
- Στη ροή εργασιών **QuakeFlow**, δίνονται στην είσοδο αρχειοθετημένες σεισμικές καταγραφές, ή καταγραφές που γίνονται σε πραγματικό χρόνο, και στην έξοδο λαμβάνεται ως τελικό προϊόν ο (επαναπροσδιορισμένος) σεισμικός κατάλογος.
- Η ροή εργασιών συνδυάζει τα μοντέλα **PhaseNet** και **GaMMA** για τη δημιουργία προκαταρκτικού σεισμικού καταλόγου, και το λογισμικό **HypoDD** για επαναπροσδιορισμό με τη μέθοδο των διπλών διαφρών.

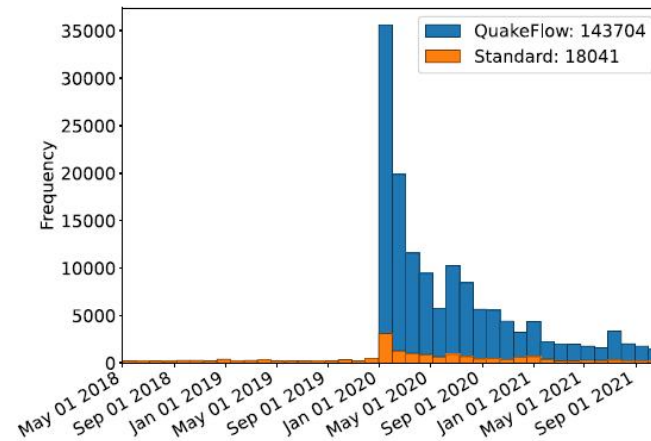


Πλήρως αυτοματοποιημένη ροή εργασιών: **QuakeFlow**

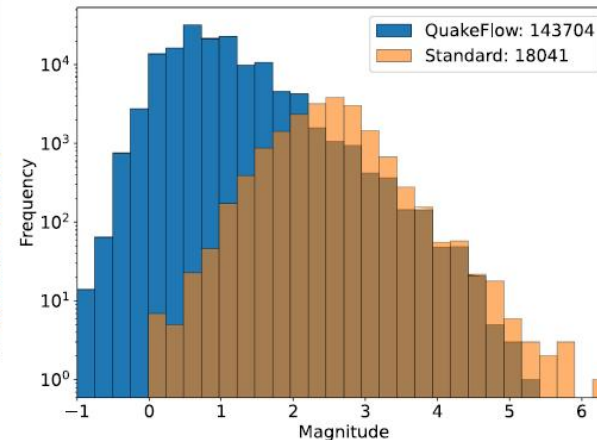
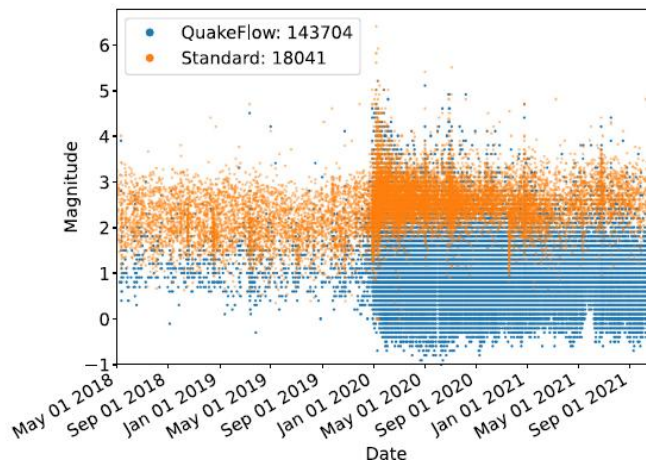
- Η εφαρμογή της ροής εργασιών **QuakeFlow** στη σεισμικότητα του Πουέρτο Ρίκο (1/5/2018 – 1/5/2021). Στις 7 Ιανουαρίου 2020 σημειώθηκε σεισμός M6.4.



(a)



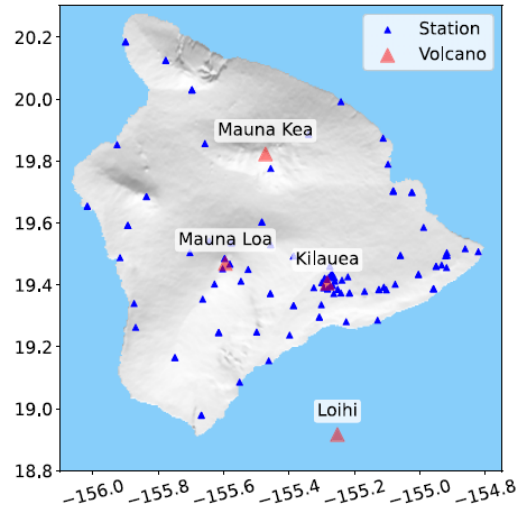
(b)



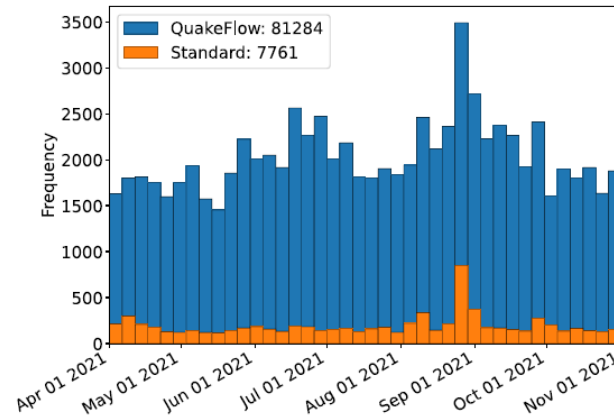
- Συγκριτικά με τον συμβατικό σεισμικό κατάλογο (πορτοκαλί), ο κατάλογος που δημιουργήθηκε μέσω QuakeFlow περιλαμβάνει ~8x αριθμό σεισμών, με το μέγεθος πληρότητας να έχει μειωθεί περισσότερο από μία τάξη μεγέθους.

Πλήρως αυτοματοποιημένη ροή εργασιών: **QuakeFlow**

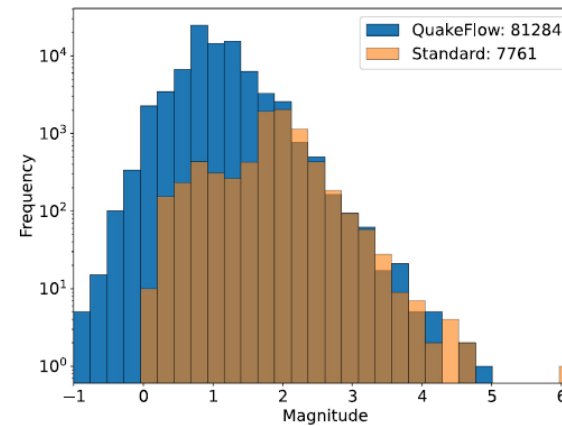
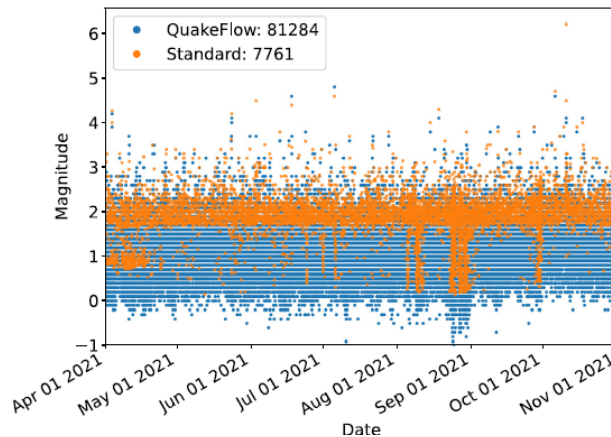
- Η εφαρμογή της ροής εργασιών **QuakeFlow** στη σεισμικότητα της Χαβάης (Big Island, 1/4 – 30/10 2021).



(a)



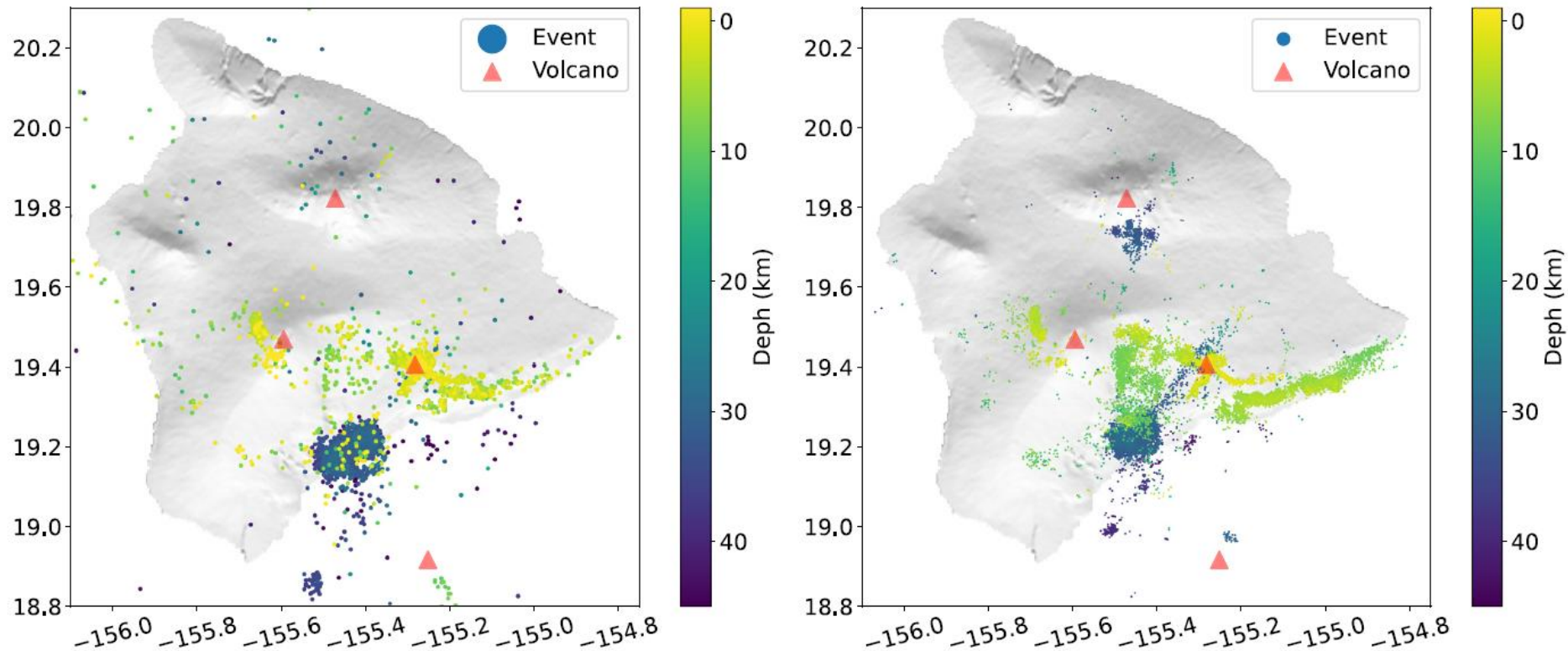
(b)



- Συγκριτικά με τον συμβατικό σεισμικό κατάλογο (πορτοκαλί), ο κατάλογος που δημιουργήθηκε μέσω QuakeFlow περιλαμβάνει **υπερδεκαπλάσιο** αριθμό σεισμών, με το μέγεθος πληρότητας να έχει μειωθεί ~1 τάξη μεγέθους.

Πλήρως αυτοματοποιημένη ροή εργασιών: **QuakeFlow**

- Η εφαρμογή της ροής εργασιών **QuakeFlow** στη σεισμικότητα της Χαβάης (Big Island, 1/4 – 30/10 2021).



- Σύγκριση κατανομής επικέντρων του συμβατικού σεισμικού καταλόγου του ηφαιστειακού παρατηρητηρίου HVO (αριστερά) με τον σεισμικό κατάλογο που παρήγαγε το QuakeFlow.

Συμπεράσματα

- Οι μέθοδοι μηχανικής μάθησης, και ιδιαίτερα τα μοντέλα βαθιάς μάθησης, είναι από τα σπουδαιότερα εργαλεία για σεισμολογικές εφαρμογές που βρίσκονται στην αιχμή της σημερινής τεχνολογίας.
- Συγκριτικά με συμβατικές μεθόδους, τα μοντέλα βαθιάς μάθησης είναι:
 - Περισσότερο αποδοτικά: λόγω κατασκευής τους επιδέχονται παράλληλη επεξεργασία (και με χρήση GPU)
 - Περισσότερο ευέλικτα: δυνατότητα εφαρμογής σε ποικίλα τεκτονικά και ηφαιστειακά περιβάλλοντα λόγω γενίκευσης
 - Ακριβέστερα στις εκτιμήσεις τους: η ακρίβειά τους προσεγγίζει εκείνη ανθρώπινων αναλυτών
 - Περισσότερο ευαίσθητα: είναι ικανά να διακρίνουν το σήμα σε συνθήκες πολύ χαμηλού SNR
 - Περισσότερο έγκυρα: μικρότερα ποσοστά εσφαλμένων εκτιμήσεων
 - Ικανά για επεξεργασία μεγάλου όγκου δεδομένων (big data)
 - Περισσότερο προσβάσιμα: πολλά από αυτά είναι δημοσίως διαθέσιμα σε ανοιχτό κώδικα (κυρίως σε γλώσσα Python)

- Carniel, R., Raquel Guzmán, S., 2021. Machine Learning in Volcanology: A Review, in: Updates in Volcanology - Transdisciplinary Nature of Volcano Science. IntechOpen. <https://doi.org/10.5772/intechopen.94217>
- Chaibeddra Tani, F., Derras, B., Theodoulidis, N., BARD, P.Y., 2025. Towards site-specific ground motion estimates in Greece using a partially non-ergodic, mixed-effects, neural network approach. Soil Dyn. Earthq. Eng. 194, 109343. <https://doi.org/10.1016/j.soildyn.2025.109343>
- Chollet, F., 2018. Deep learning with Python. Manning Publications. ISBN: 9781617294433
- Huang, S.C., Le, T.H., 2021. Principles and Labs for Deep Learning, in: Principles and Labs for Deep Learning. Elsevier, pp. 1–337. ISBN: 9780323901987. <https://doi.org/10.1016/B978-0-323-90198-7.09993-6>
- Kapetanidis, V., 2017. Spatiotemporal patterns of microseismicity for the identification of active fault structures using seismic waveform cross-correlation and double-difference relocation. PhD Thesis, University of Athens.
- Kapetanidis, V., Spingos, I., Kouskouna, V., Sakellariou, N., Galanos, N., Kaviris, G., 2023. Contribution of the local Paliki seismological stations to the monitoring of seismicity in the area of Cephalonia Island, Greece, assessed by machine-learning and probabilistic event detection methods. 8th International Colloquium on Historical Earthquakes, Palaeo-Macroseismology and Seismotectonics, 17-20 September 2023 - Lixouri, Greece, Book of Abstracts, Bulletin of the Geological Society of Greece, 2023, p137. <https://ejournals.epublishing.ekt.gr/index.php/geosociety/issue/view/1993/879>
- Kröse, B., van der Smagt, P., 1996. An introduction to Neural Networks. https://www.researchgate.net/publication/272832321_An_introduction_to_neural_networks
- Liu, S., Yang, W., Wang, J., Yuan, J., 2025. Visualization for Artificial Intelligence, Synthesis Lectures on Visualization. Springer Nature Switzerland, Cham. ISBN: 978-3-031-75339-8. <https://doi.org/10.1007/978-3-031-75340-4>
- Michellini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D., Lauciani, V., 2021. **INSTANCE** – the Italian seismic dataset for machine learning. Earth Syst. Sci. Data 13, 5509–5544. <https://doi.org/10.5194/essd-13-5509-2021>
- Mousavi, S.M., Zhu, W., Sheng, Y., Beroza, G.C., 2019a. **CRED**: A Deep Residual Network of Convolutional and Recurrent Units for Earthquake Signal Detection. Sci. Rep. 9, 10267. <https://doi.org/10.1038/s41598-019-45748-1>
- Mousavi, S.M., Sheng, Y., Zhu, W., Beroza, G.C., 2019b. STanford EArthquake Dataset (**STEAD**): A Global Data Set of Seismic Signals for AI. IEEE Access 7, 179464–179476. <https://doi.org/10.1109/ACCESS.2019.2947848>

- Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y., Beroza, G.C., 2020. **Earthquake transformer**—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* 11, 3952. <https://doi.org/10.1038/s41467-020-17591-w>
- Ross, Z.E., Meier, M., Hauksson, E., Heaton, T.H., 2018. Generalized Seismic Phase Detection with Deep Learning. *Bull. Seismol. Soc. Am.* 108, 2894–2901. <https://doi.org/10.1785/0120180080>
- Uchide, T., 2020. Focal mechanisms of small earthquakes beneath the Japanese islands based on first-motion polarities picked using deep learning. *Geophys. J. Int.* 223, 1658–1671. <https://doi.org/10.1093/gji/ggaa401>
- White, R.A., McCausland, W.A., 2019. A process-based model of pre-eruption seismicity patterns and its use for eruption forecasting at dormant stratovolcanoes. *J. Volcanol. Geotherm. Res.* 382, 267–297. <https://doi.org/10.1016/j.jvolgeores.2019.03.004>
- Woollam, J., Münchmeyer, J., Tilmann, F., Rietbrock, A., Lange, D., Bornstein, T., Diehl, T., Giunchi, C., Haslinger, F., Jozinović, D., Michelini, A., Saul, J., Soto, H., 2022. **SeisBench** - A Toolbox for Machine Learning in Seismology. *Seismol. Res. Lett.* 93, 1695–1709. <https://doi.org/10.1785/0220210324>
- Zhao, M., Xiao, Z., Zhang, M., Yang, Y., Tang, L., Chen, S., 2023a. **DiTingMotion**: A deep-learning first-motion-polarity classifier and its application to focal mechanism inversion. *Front. Earth Sci.* 11. <https://doi.org/10.3389/feart.2023.1103914>
- Zhao, M., Xiao, Z., Chen, S., Fang, L., 2023b. **DiTing**: A large-scale Chinese seismic benchmark dataset for artificial intelligence in seismology. *Earthq. Sci.* 36, 84–94. <https://doi.org/10.1016/j.eqs.2022.01.022>
- Zhu, W., Beroza, G.C., 2019. **PhaseNet**: A deep-neural-network-based seismic arrival-time picking method. *Geophys. J. Int.* 216, 261–273. <https://doi.org/10.1093/gji/ggy423>
- Zhu, W., McBrearty, I.W., Mousavi, S.M., Ellsworth, W.L., Beroza, G.C., 2022. Earthquake **Phase Association** Using a Bayesian Gaussian Mixture Model. *J. Geophys. Res. Solid Earth* 127, e2021JB023249. <https://doi.org/10.1029/2021JB023249>
- Zhu, W., Hou, A.B., Yang, R., Datta, A., Mousavi, S.M., Ellsworth, W.L., Beroza, G.C., 2023. **QuakeFlow**: a scalable machine-learning-based earthquake monitoring workflow with cloud computing. *Geophys. J. Int.* 232, 684–693. <https://doi.org/10.1093/gji/ggac355>

- Deep learning with Python: <https://github.com/fchollet/deep-learning-with-python-notebooks>

Λογισμικά:

- SeisBench: <https://github.com/seisbench/seisbench>
- EQTransformer: <https://github.com/smousavi05/EQTransformer>
- PhaseNet: <https://github.com/AI4EPS/PhaseNet>
- GPD: <https://github.com/interseismic/generalized-phase-detection>
- DeepDenoiser: <https://github.com/AI4EPS/DeepDenoiser>
- GaMMA: <https://github.com/AI4EPS/GaMMA>
- DiTingMotion: <https://github.com/mingzhaochina/DiTing-FOCALFLOW>
- QuakeFlow: <https://github.com/AI4EPS/QuakeFlow>

Labeled Datasets:

- STEAD: <https://github.com/AI4EPS/STEAD>
- INSTANCE: <https://github.com/INGV/instance>