



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΔΟΜΟΣΤΑΤΙΚΗΣ

**ΑΝΑΛΥΣΗ ΔΙΑΔΟΣΗΣ ΚΥΜΑΤΟΣ ΜΕ ΚΑΘΟΔΗΓΟΥΜΕΝΑ ΑΠΟ
ΤΗ ΦΥΣΙΚΗ ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ**



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αδαμαντία Βλαχοπούλου

Επιβλέπων: Σάββας Τριανταφύλλου, Επίκουρος Καθηγητής ΕΜΠ

Αθήνα, Νοέμβριος 2021

Copyright ©– Αδαμαντία Γ. Βλαχοπούλου, 2021.

Με επιφύλαξη παντός δικαιώματος

Η έγκριση της διπλωματικής εργασίας από τη Σχολή Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου δεν υποδηλώνει αποδοχή των απόψεων του συγγραφέα (Ν. 5343/1932, Άρθρο 202).

Copyright © Adamantia G. Vlachopoulou 2021

All Rights Reserved

Approval of this diploma thesis by the School of Civil Engineering of the National Technical University of Athens (NTUA) does not constitute in any way an acceptance of the views of the author contained herein by the said academic organization (L. 5343/1932, art. 202).

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή της εργασίας αυτής κ. Σάββα Τριανταφύλλου, που μου έδωσε την ευκαιρία να εκπονήσω την παρούσα διπλωματική και να ασχοληθώ με το συγκεκριμένο θέμα. Η καθοδήγηση, η εμπύχωση, η υπομονή και η έμπρακτη συνεισφορά του σε όλα τα στάδια της εργασίας ήταν ανεκτίμητη. Δεδομένου ότι η παρούσα εργασία σηματοδοτεί και το τέλος των φοιτητικών μου χρόνων, θα ήθελα να ευχαριστήσω όλους εκείνους που βρίσκονταν δίπλα μου, η συμπαράσταση και η βοήθεια των οποίων ήταν πολύτιμη. Συγκεκριμένα ευχαριστώ θερμά τους Νάσια, Άννα, Λυδία, Κωνσταντίνα, Φλώρα, Μάνο, Γιάννη, Ράνια, Έλενα. Τέλος θα ήθελα να εκφράσω την ευγνωμοσύνη μου στην οικογένειά μου, τους γονείς μου και τα αδέρφια μου, Χρήστο και Νίκο, για την υποστήριξη και την προσφορά τους.

ΠΕΡΙΛΗΨΗ

Η επίλυση των μερικών διαφορικών εξισώσεων με τις κλασικές αριθμητικές μεθόδους είναι μια επίπονη διαδικασία μεγάλου υπολογιστικού κόστους και πολυπλοκότητας. Για αυτό τον λόγο εξετάζεται η συνεισφορά της μηχανικής μάθησης και των τεχνητών νευρωνικών δικτύων στην επίλυση τέτοιων εξισώσεων. Πρόσφατα, εισήχθη στη διεθνή βιβλιογραφία η μεθοδολογία επίλυσης σύνθετων προβλημάτων με τη χρήση τεχνητών νευρωνικών δικτύων τα οποία υπακούν νόμους της φυσικής, οι οποίοι περιγράφονται από μη γραμμικές μερικές διαφορικές εξισώσεις (PINNs - Physics Informed Neural Networks). Ανάλογα με τη φύση του προβλήματος και τα διαθέσιμα δεδομένα, διακρίνονται δύο κύριες κατηγορίες προβλημάτων και συγκεκριμένα η επίλυση της εξίσωσης του προβλήματος και η εύρεση των παραμέτρων της εξίσωσης για γνωστή τη λύση.

Στα πλαίσια της παρούσας διπλωματικής εργασίας, ελέγχεται η αξιοπιστία της μεθόδου PINNs, τόσο ως προς το πρόβλημα εξαγωγής, όσο και ως προς το πρόβλημα ταυτοποίησης παραμέτρων για την περίπτωση της διάδοσης κύματος σε ελαστικές δοκούς. Στην παρούσα εργασία, η κυματική εξίσωση επιλύεται με την εφαρμογή της μεθόδου PINNs και τα αποτελέσματα συγκρίνονται με την αναλυτική λύση, όπως αυτή προκύπτει από την ιδιομορφική ανάλυση. Παρατηρείται πως η μέθοδος PINNs μπορεί να προσεγγίσει με ακρίβεια τη λύση της διαφορικής εξίσωσης και τις παραμέτρους αυτής με γνωστή τη λύση της εξίσωσης, εντός της χωροχρονικής κλίμακας στην οποία έχει εκπαιδευτεί. Παρατηρήθηκε απόκλιση στις τιμές των συνοριακών συνθηκών δοκού.

Λέξεις κλειδιά: τεχνητά νευρωνικά δίκτυα, τεχνητά νευρωνικά δίκτυα που υπακούν νόμους της φυσικής (PINNs), μη γραμμικές μερικές διαφορικές εξισώσεις, ιδιομορφική ανάλυση, εξίσωση Euler-Bernoulli για δοκούς

ABSTRACT

The solution of nonlinear partial differential equations using numerical methods is an arduous, in cases complex, and computationally taxing procedure. In this thesis, the efficiency of machine learning as a means to rapidly tackle the solution of nonlinear partial differential equations is examined.

To this end, a new methodology for solving complex problems has recently emerged in the literature, i.e., Physics Informed Neural Networks (PINNs). This involves the utilization of artificial neural networks that obey a physical law and hence are described by nonlinear partial differential equations. Based on the nature of the problem and the available data, two classes of problems may be determined, i.e., those that aim at finding the solution of the equation that describes the physical problem and those that aim at estimating the parameters of the equation for a given known solution.

In this thesis, the robustness and accuracy of PINNs to tackle partial differential equations is examined. As a first step, the case of the nonlinear Burgers equation is investigated. Next, the case of elastic waves propagating in beams is examined. To this end, both the inference and the identification problem were addressed for the beam wave equation using PINNs and the results were compared with the analytical solution obtained through modal analysis. It has been observed that PINNs can accurately estimate the solution of this differential equation and estimate its parameters, while within the spatial and temporal scale that neural network has been trained. It established that there was a difference between the exact and predicted values of the beam boundary conditions.

Keywords: artificial neural networks, Physics Informed Neural Networks (PINNs), nonlinear partial differential equations, modal analysis, Euler-Bernoulli equation of beams

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΚΕΦΑΛΑΙΟ 1. Εισαγωγή	20
1.1 Γενική περιγραφή του προβλήματος	20
1.2 Αντικείμενο και στόχος της εργασίας	21
1.3 Μεθοδολογία	22
1.4 Διάρθρωση της εργασίας	23
ΚΕΦΑΛΑΙΟ 2. Εισαγωγή στην τεχνητή νοημοσύνη	25
2.1 Τεχνητή νοημοσύνη (Artificial Intelligence-AI)	25
2.2 Μηχανική μάθηση (Machine Learning-ML)	26
2.2.1 Επιβλεπόμενη μάθηση	28
2.2.1.1 Ταξινόμηση (Classification)	29
2.2.1.2 Παλινδρόμηση (Regression)	30
2.2.2 Μη επιβλεπόμενη μάθηση	31
2.2.3 Ενισχυτική μάθηση	32
2.2.4 Μερικώς επιβλεπόμενη μάθηση	33
2.2.5 Βαθιά μάθηση (Deep learning)	33
ΚΕΦΑΛΑΙΟ 3. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ	35
3.1 Βιολογικοί και τεχνητοί νευρώνες	35
3.2 Ο αλγόριθμος – Perceptron	36
3.3 Λειτουργίες τεχνητών νευρωνικών δικτύων	38
3.4 Τεχνητοί νευρώνες	38
3.5 Νευρωνικά δίκτυα πρόσθιας τροφοδότησης	40
3.6 Μεμονωμένο επίπεδο (single layer)	41
3.7 Πολύ-επίπεδα Νευρωνικά Δίκτυα	42
3.8 Συνάρτηση Ενεργοποίησης	42

3.9	Υπερπαράμετροι	44
3.10	Αλγόριθμοι βελτιστοποίησης (Optimization Algorithms)	46
3.10.1	Πρώτης τάξης μέθοδοι βελτιστοποίησης	46
3.10.1.1	Αλγόριθμος κατάβασης κλίσης (Gradient Descent)	46
3.10.1.2	Στοχαστική κατάβαση κλίσης (Stochastic gradient descent).....	47
3.10.1.3	Αλγόριθμος Adam.....	48
3.10.2	Μέθοδοι βελτιστοποίησης δευτέρας τάξης	49
3.10.2.1	Αλγόριθμος L-BFGS.....	49
3.10.3	Οπισθοδιάδοση (backpropagation)	50
3.10.4	Εκπαίδευση νευρωνικών δικτύων Neural network training.....	52
3.10.5	Βαθιά Νευρωνικά Δίκτυα -Deep neural networks	52
3.10.6	TensorFlow.....	52
3.10.7	Προβλήματα κατά την εκπαίδευση ενός νευρωνικού δικτύου	53
3.10.7.1	Υπερπροσαρμογή.....	53
3.10.7.2	Υποπροσαρμογή	53
3.10.8	Βασική ορολογία των ΝΔ	54
3.10.8.1	Εποχή(Epoch)	54
3.10.8.2	Μέγεθος ομάδας (Batch size)	55
3.10.8.3	Επανάληψη (Iteration)	55
3.10.8.4	Αριθμός υπολογισμού συναρτήσεων	56
3.10.8.5	Συνάρτηση σφάλματος (Loss function).....	56
3.10.8.6	Συνάρτηση κόστους (Cost function or Objective Function).....	57
3.10.8.7	Μέσο τετραγωνικό σφάλμα (Mean Squared Error).....	57
ΚΕΦΑΛΑΙΟ 4.	Νευρωνικά δίκτυα που υπακουν στους νόμους της φυσικής -Physics Informed Neural Networks (PINNS).....	59
4.1	Εισαγωγή στα PINNs.....	59

4.2	Ορισμός του προβλήματος	60
4.3	Εκπαίδευση των PINNs.....	62
4.4	Ταυτοποίηση παραμέτρων διαφορικής εξίσωσης (parameter identification)	63
4.5	Μέθοδος του λατινικού υπερκύβου - Latin Hypercube Sampling	64
ΚΕΦΑΛΑΙΟ 5. Δυναμική ανάλυση		65
5.1	Δυναμική ανάλυση των κατασκευών	65
5.2	Ιδιομορφική ανάλυση	65
5.3	Εξίσωση δυναμικής ισορροπίας	66
5.4	Ιδιομορφική ανάλυση γενικευμένων ελαστικών συστημάτων.....	69
5.4.1	Αναλυτική λύση ελεύθερης ταλάντωσης για αμφιέρειστη δοκό	70
5.4.2	Αναλυτική λύση ελεύθερης ταλάντωσης για δοκό πρόβολο	71
5.5	Εξαναγκασμένες ταλαντώσεις δοκών.....	72
ΚΕΦΑΛΑΙΟ 6. Εφαρμογή της μεθόδου PINNs στην εξίσωση Burgers		74
6.1	Εφαρμογή 1: Πρόβλημα επαγωγής εξίσωσης Burgers.....	74
6.2	Εφαρμογή 2: Ταυτοποίηση παραμέτρων εξίσωσης Burgers.....	84
6.2.1	Περίπτωση δεδομένων χωρίς θόρυβο	84
6.2.2	Περίπτωση δεδομένων με θόρυβο	86
ΚΕΦΑΛΑΙΟ 7. Εφαρμογή της μεθόδου PINN στην ανάλυση διάδοσης κύματος σε δοκό.....		89
7.1	Περιπτώσεις προβλημάτων επαγωγής (inference)	89
7.1.1	Εφαρμογή 1: Εξίσωση ταλάντωσης αμφιέρειστης δοκού υπό ημιτονοειδές φορτίο.....	89
7.1.2	Εφαρμογή 2: Εξίσωση ταλάντωσης δοκού- πρόβολου.....	103
7.1.3	Εφαρμογή 3: Εξίσωση ταλάντωσης αμφιέρειστης δοκού υπό σύνθεσης φορτίων.....	112
7.2	Περιπτώσεις προβλημάτων ταυτοποίησης παραμέτρων (identification)	120
7.2.1	Εφαρμογή 1: Εξίσωση ταλάντωσης αμφιέρειστης δοκού υπό ημιτονοειδές φορτίο (χωρίς θόρυβο) 120	
ΚΕΦΑΛΑΙΟ 8. Εφαρμογή της μεθόδου PINN για την πρόβλεψη της συμπεριφοράς της ΜΔΕ....		123

8.1	Εφαρμογή 1: Εξίσωση Burgers	123
8.2	Εφαρμογή 2: Εξίσωση ανάλυσης διάδοσης κύματος σε αμφιέριστη δοκό	125
ΚΕΦΑΛΑΙΟ 9. Συμπεράσματα και προτάσεις για μελλοντική έρευνα		127
9.1	Συμπεράσματα	127
9.2	Προτάσεις για μελλοντική έρευνα.....	128
ΚΕΦΑΛΑΙΟ 10. Αναφορές		130

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Εικόνα 2–1: Η βαθιά μάθηση ως υποσύνολο της μηχανικής μάθησης και οι δύο σαν υποσύνολα της τεχνητής νοημοσύνης (πηγή: www.intel.la.com)	26
Εικόνα 2–2: Η ροή της διαδικασίας της επιβλεπόμενης μάθησης (πηγή: www.tutorialandexample.com)	29
Εικόνα 2–3: Διαφορά μεταξύ ταξινόμησης και παλινδρόμησης (πηγή: www.javatpoint.com).....	30
Εικόνα 2–4: Η ροή της διαδικασίας της μη επιβλεπόμενης μηχανικής μάθησης (πηγή: www.educba.com)	32
Εικόνα 2–5: Τυπική ροή της διαδικασίας της ενισχυτικής μηχανικής μάθησης (πηγή: www.wikipedia.com).....	33
Εικόνα 3–1: Από τους βιολογικούς στους τεχνητούς νευρώνες (πηγή: www.ΜηχανικήΜάθηση(kallipos.gr)).....	35
Εικόνα 3–2: Στην εικόνα αναπαριστάται ο τρόπος με τον οποίο ο αλγόριθμος Perceptron διαμορφώνει τη γραμμή εκπαίδευσης όσο μεγαλώνει το δείγμα εκπαίδευσης (πηγή: www.wikipedia.com).....	38
Εικόνα 3–3: Τα βασικά συστατικά στοιχεία ενός τεχνητού νευρώνα (πηγή: www.openlearninglibrary.mit.edu).....	40
Εικόνα 3–4: Αναπαράσταση μεμονομένου επιπέδου με $x \in \mathbb{R}^m$ στοιχεία εισόδου σε κάθε νευρώνα και n πλήθος εξόδων του επιπέδου (πηγή: www.openlearninglibrary.mit.edu)	41
Εικόνα 3–5: Πολυεπίπεδο ΝΔ με 3 εισόδους, 2 κρυφά επίπεδα και μία έξοδο (πηγή: www.GitHub.com)	42
Εικόνα 3–6: Γραφικές παραστάσεις κυριότερων συναρτήσεων ενεργοποίησης (πηγή: www.openlearninglibrary.mit.edu).....	44
Εικόνα 3–7: Σχηματική απεικόνιση του τρόπου λειτουργίας του αλγόριθμου κατάβασης κλίσης, όπου στον καθeto άξονα αναπαριστώνται οι τιμές της συνάρτησης κόστους και στον οριζόντιο οι τιμές της παραμέτρου ως προς την οποία παραγωγίζεται η συνάρτηση κόστους (πηγή Gradient descent algorithm and its three types Clairvoyant Blog (clairvoyantsoft.com))	47
Εικόνα 3–8: οι καμπύλες στην περίπτωση της υπερπροσαρμογής, υποπροσαρμογής σε σχέση με την ιδανική λύση (πηγή: www.TowardsDataScience.com).	54
Εικόνα 5–1: Ισορροπία ενός στοιχείου της δοκού με μήκος dx αριστερά και το σχήμα της διατομής αριστερά (πηγή: Δυναμική ανάλυση των κατασκευών-Ι.Κατσικαδέλης, 2020)	67

Εικόνα 5–2: Ιδιομορφές ταλαντώσεως και φυσικές συχνότητες αμφιέρειστης δοκού σταθερής διατομής αριστερά και προβόλου σταθερής διατομής δεξιά (πηγή: Δυναμική ανάλυση των κατασκευών- I. Κατσικαδέλης, 2020).....	72
Εικόνα 6–1 Τα 100 σημεία των αρχικών και συνοριακών συνθηκών μαζί με 10000 εσωτερικά σημεία του χωρίου, όπως προέκυψαν από τη μέθοδο του λατινικού υπερκύβου LHS	76
Εικόνα 6–2: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Burgers.....	78
Εικόνα 6–3: Τρισδιάστατη απεικόνιση της λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Burgers.	79
Εικόνα 6–4: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $t=0.25$ sec	80
Εικόνα 6–5: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $t=0.50$ sec	80
Εικόνα 6–6: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $t=1$ sec	81
Εικόνα 6–7: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $x=-0.50$ m.....	81
Εικόνα 6–8: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $x=0.25$ m	82
Εικόνα 6–9: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $x=0.95$ m	82
Εικόνα 6–10: Σύγκριση του σφάλματος μεταξύ αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για τιμές νευρώνων και κρυφών επιπέδων. Τα συνοριακά και σημεία αρχικών συνθηκών είναι σταθερά ίσα με $Nu=100$, και τα εσωτερικά σημεία $Nf=10000$	83
Εικόνα 6–11: Σύγκριση του σφάλματος μεταξύ αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για διαφορετικά πλήθη συνοριακών και σημείων αρχικών συνθηκών. Οι νευρώνες και τα κρυφά επίπεδα είναι σταθερά ίσα με 8 και 20, αντίστοιχα.	83
Εικόνα 6–12: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου u της εξίσωσης Burgers.	85
Εικόνα 6–13: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου κινηματικού ιξώδες ν της εξίσωσης Burgers.	86
Εικόνα 6–14: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου u της εξίσωσης Burgers για δεδομένα με θόρυβο.....	87
Εικόνα 6–15: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου κινηματικού ιξώδες ν της εξίσωσης Burgers για δεδομένα με θόρυβο.	87
Εικόνα 7–1: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x,t) = \sin(2\pi t)$	91

Εικόνα 7–2: Σχηματική απεικόνιση των δεδομένων εκπαίδευσης του PINN για την επίλυση της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$	93
Εικόνα 7–3: Τρισδιάστατη απεικόνιση λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$	94
Εικόνα 7–4: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.25\text{sec}$	95
Εικόνα 7–5: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.7\text{sec}$	96
Εικόνα 7–6: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=1\text{sec}$	96
Εικόνα 7–7: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=0\text{ m}$	97
Εικόνα 7–8: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=1.52\text{ m}$	97
Εικόνα 7–9: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=3\text{ m}$	98
Εικόνα 7–10: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=5\text{ m}$	98
Εικόνα 7–11: Τρισδιάστατη απεικόνιση της λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ με διαφορετική εξίσωση σφάλματος.	100
Εικόνα 7–12: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=0\text{m}$	100
Εικόνα 7–13: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=1.46\text{ m}$	101
Εικόνα 7–14: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=4\text{ m}$	101

Εικόνα 7–15: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=5$ m	102
Εικόνα 7–16: Σύγκριση του σφάλματος μεταξύ αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για διαφορετικά πλήθη συνοριακών σημείων και σημείων αρχικών συνθηκών. Οι νευρώνες και τα κρυφά επίπεδα είναι σταθερά και ίσα με 8 και 20, αντίστοιχα.....	102
Εικόνα 7–17: Σύγκριση του σφάλματος μεταξύ αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για διαφορετικές τιμές νευρώνων και κρυφών επιπέδων. Τα συνοριακά σημεία και σημεία αρχικών συνθηκών είναι σταθερά και ίσα με $Nu = 2100$, και τα εσωτερικά σημεία $Nf=10000$	103
Εικόνα 7–18: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$	105
Εικόνα 7–19: Τρισδιάστατη απεικόνιση λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$	107
Εικόνα 7–20: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.25$ sec	108
Εικόνα 7–21: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.75$ sec	109
Εικόνα 7–22: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.99$ sec	109
Εικόνα 7–23: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=0$ m	110
Εικόνα 7–24: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=3.50$ m	110
Εικόνα 7–25: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=5$ m	111
Εικόνα 7–26 Σχηματική απεικόνιση των δεδομένων εκπαίδευσης του PINN για την επίλυση της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$	111

Εικόνα 7–27: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	113
Εικόνα 7–28: Τρισδιάστατη απεικόνιση λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	115
Εικόνα 7–29 Σχηματική απεικόνιση των δεδομένων εκπαίδευσης του PINN για την επίλυση της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	116
Εικόνα 7–30: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$ για $t=1$ sec	116
Εικόνα 7–31: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό για $t=3.50$ sec υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	117
Εικόνα 7–32: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $t=4.70$ sec για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	117
Εικόνα 7–33: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $x=1.50$ m για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	118
Εικόνα 7–34: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $x=3$ m για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	118
Εικόνα 7–35: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $x=4.50$ m για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	119
Εικόνα 7–36: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $x=4.95$ m για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin 2\pi t + \cos \pi 2t + \sin \pi 9t + \sin(\pi t + \pi 3)$	119
Εικόνα 7–37: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου του γινομένου του EI της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$	122
Εικόνα 8–1: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Burgers για χρόνο $1 \leq t \leq 5$ sec	123
Εικόνα 8–2: Τρισδιάστατη απεικόνιση της λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Burgers για χρόνο $1 \leq t \leq 5$ sec για τον οποίο το δίκτυο δεν έχει εκπαιδευτεί.	124
Εικόνα 8–3: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για χρόνο $0 \leq t \leq 2$ sec	125

Εικόνα 8–4: Τρισδιάστατη απεικόνιση της λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για χρόνο $1 \leq t \leq 2$ sec για τον οποίο το δίκτυο δεν έχει εκπαιδευτεί. 126

ΚΕΦΑΛΑΙΟ 1. Εισαγωγή

1.1 Γενική περιγραφή του προβλήματος

Υπάρχουν φαινόμενα με γνωστούς τους φυσικούς νόμους που τα διέπουν αλλά που η επίλυση των μαθηματικών εξισώσεων, που περιγράφουν αυτούς τους νόμους με τις υπάρχουσες αριθμητικές μεθόδους (πχ πεπερασμένα στοιχεία), αποτελεί μια επίπονη διαδικασία μεγάλου υπολογιστικού κόστους. Για αυτόν τον λόγο και ιδιαίτερα για την περίπτωση των μη γραμμικών μερικών διαφορικών εξισώσεων τα τελευταία χρόνια υπάρχει η τάση να χρησιμοποιούνται μέθοδοι μηχανικής μάθησης για την επίλυσή τους.

Στην παρούσα εργασία εξετάζεται η συνεισφορά της μηχανικής μάθησης και των τεχνητών νευρωνικών δικτύων στην επίλυση προβλημάτων μηχανικού και πιο συγκεκριμένα στην ανάλυση διάδοσης κύματος σε δοκό. Η μέθοδος που εφαρμόστηκε καλείται PINNs (Physics Informed Neural Networks). Πρόκειται για τεχνητά νευρωνικά δίκτυα τα οποία υπακούν νόμους της φυσικής. Με τη μέθοδο αυτή επιλύονται δύο ειδών προβλήματα ανάλογα με τα διαθέσιμα δεδομένα και τη φύση του προβλήματος. Το πρώτο είναι η επίλυση της εξίσωσης που περιγράφει το πρόβλημα και το δεύτερο είναι η εύρεση των παραμέτρων της εξίσωσης με γνωστή τη λύση της εξίσωσης. Η αξιοπιστία της μεθόδου ελέγχεται επιλύοντας το πρόβλημα της επαγωγής και της ταυτοποίησης παραμέτρων της εξίσωσης Burgers για την οποία υπάρχουν διαθέσιμα δεδομένα εκπαίδευσης από την αναλυτική της λύση. Το πρόβλημα της ταυτοποίησης των παραμέτρων επιλύεται για την περίπτωση δεδομένων χωρίς θόρυβο και με θόρυβο.

Όπως πολλά προβλήματα πολιτικού μηχανικού έτσι και η εξίσωση ανάλυσης κύματος σε δοκό περιγράφεται από μερική διαφορική εξίσωση που είναι γνωστή ως Euler-Bernoulli. Η αναλυτική λύση της εξίσωσης είναι γνωστή και προκύπτει από την ιδιομορφική ανάλυση. Στην εξίσωση αυτή για τις περιπτώσεις ανάλυσης κύματος σε αμφιέριστη και δοκό πρόβολο εφαρμόζεται η μεθοδολογία PINNs. Επιπλέον για την περίπτωση της αμφιέριστης δοκού εφαρμόζεται η μέθοδος PINNs για την ταυτοποίηση των παραμέτρων της εξίσωσης. Συμπεραίνεται πως η εν λόγω μέθοδος δύναται να εκτιμήσει ικανοποιητικά τη λύση της διαφορικής εξίσωσης εντός του χωροχρονικού πεδίου για το οποίο έχει εκπαιδευτεί, αν και παρατηρηθήκαν αποκλίσεις μεταξύ των τιμών της αναλυτικής λύσης και εκείνων που υπολογίστηκαν από το δίκτυο στις συνοριακές θέσεις των δοκών. Με ακρίβεια

υπολογίζονται και οι παράμετροι των εξισώσεων με εφαρμογή της μεθόδου PINNs για το πρόβλημα της ταυτοποίησης παραμέτρων.

Τέλος εξετάζεται η δυνατότητα των PINNs να προβλέψουν τη συμπεριφορά των εξισώσεων εκτός του πεδίου για το οποίο το δίκτυο έχει εκπαιδευτεί. Τα αποτελέσματα δεν είναι ιδιαίτερα ικανοποιητικά ως προς αυτήν την κατεύθυνση.

1.2 Αντικείμενο και στόχος της εργασίας

Η παρούσα εργασία πραγματεύεται τη χρήση της μεθόδου PINNs, η οποία εφαρμόζει Τεχνητά Νευρωνικά Δίκτυα-TND- που εκπαιδεύονται για την επίλυση προβλημάτων μη γραμμικών μερικών διαφορικών εξισώσεων-MDE- οι οποίες περιγράφουν κάποιο νόμο της φυσικής, για την ανάλυση διάδοσης κύματος σε δοκό. Αρχικά, περιγράφεται η φιλοσοφία που διέπει τη μεθοδολογία αυτή και τον τρόπο με τον οποίο μπορεί να εφαρμοστεί σε προβλήματα πολιτικού μηχανικού. Στη συνέχεια παρουσιάζεται η δυναμική των PINNs μέσω της εφαρμογής τους στην εξίσωση Burgers και κατόπιν εξετάζεται η επιτυχία της μεθόδου σε περιπτώσεις αμφιέριστης δοκού, δοκού πρόβολου υπό φόρτιση, για τις οποίες υπάρχουν δεδομένα για την παραμορφωσιακή τους κατάσταση από τις αναλυτικές τους λύσεις με αριθμητικές μεθόδους. Συγκεκριμένα στην παρούσα εργασία τα δίκτυα αυτά εφαρμόζονται για την επίλυση της γραμμικής MDE που στη θεωρία της δυναμικής ανάλυσης είναι γνωστή ως εξίσωση Euler-Bernoulli. Επιπρόσθετα παρουσιάζεται και η δυνατότητα των PINNs για εύρεση των παραμέτρων της διαφορικής εξίσωσης, έχοντας γνωστή τη λύση της. Τέλος, εξετάζεται το ενδεχόμενο η μέθοδος PINNs να προβλέπει τη συμπεριφορά του εκάστοτε φαινομένου εκτός του χωροχρονικού πεδίου για το οποίο έχει εκπαιδευτεί το δίκτυο.

Στην προκειμένη εργασία στόχος είναι να εξεταστούν οι δυνατότητες της μεθόδου PINNs, και συγκεκριμένα της μηχανικής μάθησης, στην επίλυση μερικών διαφορικών εξισώσεων των οποίων η επίλυση με τις κλασικές αριθμητικές μεθόδους είναι ιδιαίτερα επίπονη διαδικασία με μεγάλο υπολογιστικό κόστος. Αρχικά η μέθοδος εφαρμόζεται για την επίλυση της εξίσωσης Burgers, τόσο για το πρόβλημα της επαγωγής όσο και για εκείνο της ταυτοποίησης των παραμέτρων της. Κατόπιν μελετάται η συνεισφορά της μεθόδου στην ανάλυση διάδοσης κύματος σε δοκό και συγκεκριμένα στην επίλυση της εξίσωσης Euler- Bernoulli. Η μέθοδος εφαρμόστηκε σε περίπτωση αμφιέριστης και δοκού πρόβολου. Για την περίπτωση της αμφιέριστης δοκού εφαρμόζεται η μέθοδος PINNs και για την εύρεση παραμέτρων. Σκοπός της μεθοδολογίας PINNs δεν είναι η αντικατάσταση των αριθμητικών μεθόδων που ήδη υπάρχουν στη βιβλιογραφία αλλά να συνυπάρξουν με τους αλγορίθμους επίλυσης μερικών

διαφορικών εξισώσεων και να αποτελέσουν χρήσιμο εργαλείο στο πεδίο των εφαρμοσμένων μαθηματικών και της μηχανικής στις περιπτώσεις που τα διαθέσιμα δεδομένα για την επίλυση της εξίσωσης δεν είναι αρκετά ή οι τιμές των παραμέτρων της είναι άγνωστες.

1.3 Μεθοδολογία

Στα πλαίσια της εργασίας, η εφαρμογή της μεθόδου PINNs αποδομείται στα εξής μέρη:

1. Συλλογή Δεδομένων Εκπαίδευσης: Αναλυτική επίλυση της παραμόρφωσης του υπό μελέτη στοιχείου υπό δεδομένο φορτίο και σε ορισμένο χωροχρονικό πεδίο με χρήση της ιδιομορφικής ανάλυσης σε προγραμματιστικό περιβάλλον MATLAB. Αποθήκευση των δεδομένων σε μητρώα.
2. Εφαρμογή του PINN: Ανακατασκευή του PINN βάσει της υπο μελέτη περίπτωσης. Τροφοδότηση του PINN με τα δεδομένα εκπαίδευσης όπως προέκυψαν από την αναλυτική λύση. Εκπαίδευση του PINN, πραγματοποίηση πρόβλεψης σε προγραμματιστικό περιβάλλον PYTHON. Αποθήκευση των δεδομένων σε μητρώα.
3. Επεξεργασία δεδομένων: Προβολή και σύγκριση των αποτελεσμάτων της αναλυτικής επίλυσης και της πρόβλεψης του PINN μέσω γραφημάτων σε προγραμματιστικό περιβάλλον MATLAB. Εξαγωγή συμπερασμάτων.

Αρχικά, πραγματοποιήθηκε η αναλυτική επίλυση της εξίσωσης του προβλήματος με χρήση των εξισώσεων της δυναμικής και συγκεκριμένα της ιδιομορφικής ανάλυσης. Έπειτα ένα τυχαίο πλήθος των δεδομένων αυτών χρησιμοποιούνται ως δεδομένα εκπαίδευσης του PINN, αφού το δίκτυο έχει προηγουμένως ανακατασκευαστεί ως προς την εξίσωση σφάλματος εκπαίδευσης καθώς και την εκάστοτε εξίσωση. Στη συνέχεια πραγματοποιείται πρόβλεψη της συμπεριφοράς της εξίσωσης από το PINN και σύγκριση των αποτελεσμάτων που προέκυψαν με εκείνα της αναλυτικής λύσης.

Για την επίτευξη των παραπάνω συντάχθηκαν προγράμματα σε γλώσσα MATLAB με δεδομένα εισόδου τη γεωμετρία του στοιχείου που μελετάται, τις μηχανικές ιδιότητες του υλικού του, τον αριθμό των ιδιομορφών, το χωροχρονικό πεδίο και τη διακριτοποίηση στον χώρο και τον χρόνο. Ως έξοδος δίδεται το μητρώο των μετατοπίσεων εντός του χωροχρονικού πεδίου που έχει οριστεί. Ακολούθως το μητρώο των μετατοπίσεων τροφοδοτεί το PINN προκειμένου αυτό να εκπαιδευτεί χρησιμοποιώντας τυχαίο πλήθος στοιχείων του μητρώου. Παράλληλα τα PINNs τροφοδοτούνται από τα μητρώα των θέσεων και των χρονικών στιγμών που έχουν οριστεί για το εκάστοτε πρόβλημα μαζί με τις τιμές των παραμέτρων της εξίσωσης στην περίπτωση των προβλημάτων επαγωγής. Στην περίπτωση των προβλημάτων ταυτοποίησης παραμέτρων οι παράμετροι του προβλήματος λογίζονται ως μεταβλητές και επιχειρείται η εύρεση της τιμής τους ώστε να περιγράψουν ακριβέστερα τα δεδομένα. Τα PINNs

είναι κατασκευασμένα σε γλώσσα προγραμματισμού PYTHON. Αφού πραγματοποιηθεί η ανακατασκευή της εξίσωσης που εφαρμόζεται καθώς και της εξίσωσης σφάλματος, επιχειρούνται δοκιμές ώστε να βρεθούν οι κατάλληλοι συντελεστές βαρύτητας κάθε όρου της εξίσωσης σφάλματος, ώστε το σφάλμα του δικτύου να προκύπτει το μικρότερο δυνατό. Στην περίπτωση των προβλημάτων επαγωγής τα PINNs δίνουν ως έξοδο το μητρώο των μετατοπίσεων εντός του χωροχρονικού πεδίου που έχει οριστεί, το σφάλμα εκπαίδευσης και το σφάλμα της πρόβλεψης. Στην περίπτωση των προβλημάτων εύρεσης παραμέτρων δίδεται επιπλέον ως έξοδος η τιμή της παραμέτρου και το σφάλμα της τιμής συγκριτικά με την τιμή που χρησιμοποιήθηκε στην αναλυτική επίλυση του προβλήματος. Τέλος τα αποτελέσματα της μεθόδου PINN συγκρίνονται με εκείνα της αναλυτικής λύσης.

1.4 Διάρθρωση της εργασίας

Στο πρώτο κεφάλαιο γίνεται εισαγωγή στο θέμα της παρούσας εργασίας. Αναφέρονται ο σκοπός και το αντικείμενο της, συνοπτικά κάποια βασικά στοιχεία, καθώς και εργαλεία και εφαρμογές που πραγματοποιήθηκαν. Επίσης περιγράφεται η μεθοδολογική προσέγγιση και η συλλογιστική πορεία των βημάτων της εργασίας.

Στο δεύτερο κεφάλαιο αναπτύσσεται συνοπτικά το θεωρητικό υπόβαθρο της τεχνητής νοημοσύνης, της μηχανικής μάθησης και των τεχνητών νευρωνικών δικτύων.

Στο τρίτο κεφάλαιο παρουσιάζονται διεξοδικότερα τα τεχνητά νευρωνικά δίκτυα, αναλύοντας τη δομή και τη λειτουργία τους.

Στο τέταρτο κεφάλαιο γίνεται παρουσίαση της μεθόδου που εξετάζεται στην παρούσα εργασία, συγκεκριμένα των νευρωνικών δικτύων που υπακούν στους νόμους της φυσικής (Physics Informed Neural Networks-PINNs). Δίδεται ο γενικός ορισμός του προβλήματος, ο τρόπος με τον οποίο εφαρμόζονται τα PINNs καθώς και η δυνατότητα εφαρμογής τους για την εύρεση των παραμέτρων της εξίσωσης του εκάστοτε φαινομένου. Τέλος αναπτύσσεται συνοπτικά η θεωρία του λατινικού υπερκύβου.

Στο πέμπτο κεφάλαιο παρατίθενται βασικά θεωρητικά στοιχεία της δυναμικής ανάλυσης των κατασκευών και ορίζεται η ιδιομορφική ανάλυση. Περιγράφεται το βασικό μαθηματικό υπόβαθρο της εξίσωσης της δυναμικής ισορροπίας, της ιδιομορφικής ανάλυσης γενικευμένων ελαστικών συστημάτων, της αναλυτικής λύσης της ελεύθερης ταλάντωσης για αμφιέρειστη δοκό και δοκο πρόβολο και τέλος των εξαναγκασμένων ταλαντώσεων δοκών.

Στο έκτο κεφάλαιο παρατίθενται οι εφαρμογές των δικτύων PINNs για την επίλυση της εξίσωσης Burgers και για την εύρεση της παραμέτρου της εξίσωσης στην περίπτωση δεδομένων με θόρυβο και δεδομένων χωρίς θόρυβο

Στο έβδομο κεφάλαιο παρουσιάζονται οι εφαρμογές των δικτύων PINNs στα προβλήματα επαγωγής της εξίσωσης διάδοσης κύματος σε αμφιέριστη δοκό υπό ημιτονοειδές φορτίο και σε σύνθεση φορτίων και σε δοκό πρόβολο υπό ημιτονοειδές φορτίο. Ακόμα επιλύεται και το πρόβλημα για την εύρεση της παραμέτρου του γινομένου του μέτρου ελαστικότητας και της δυσκαμψίας της εξίσωσης διάδοσης κύματος σε αμφιέριστη δοκό υπό ημιτονοειδές φορτίο.

Στο όγδοο κεφάλαιο εξετάζεται η δυνατότητα της μεθόδου PINNs να προβλέπει τη συμπεριφορά των ΜΔΕ που επιλύονται σε χωροχρονικό πεδίο εκτός του πεδίου για το οποίο έχει εκπαιδευτεί. Οι εφαρμογές έγιναν στην εξίσωση Burgers και στην εξίσωσης διάδοσης κύματος σε αμφιέριστη δοκό υπό ημιτονοειδές φορτίο.

Στο ένατο κεφάλαιο συνοψίζονται τα συμπεράσματα που προκύπτουν από την παρούσα εργασία και δίνονται ενδεικτικές προτάσεις για μελλοντική έρευνα

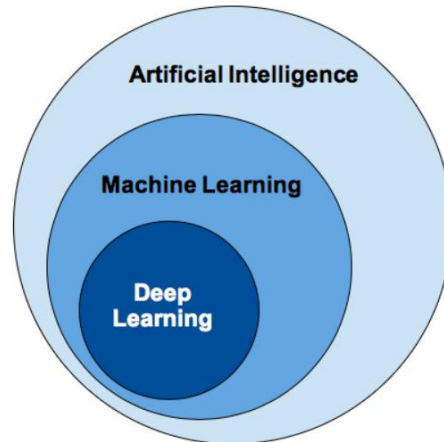
ΚΕΦΑΛΑΙΟ 2. Εισαγωγή στην τεχνητή νοημοσύνη

2.1 Τεχνητή νοημοσύνη (Artificial Intelligence-AI)

Η ιδέα και ο όρος τεχνητή νοημοσύνη επινοήθηκε για πρώτη φορά το 1956, αλλά κατά τα τελευταία χρόνια η δημοφιλία της είναι ολοένα και μεγαλύτερη. Ο κυριότερος λόγος ήταν η έλλειψη επαρκών δεδομένων, τα οποία είναι απαραίτητα για ακριβείς προβλέψεις. Στη σύγχρονη εποχή το πλήθος των διαθέσιμων δεδομένων έχει αυξηθεί δραματικά και συνεχίζει να αυξάνεται. Ταυτόχρονα, η υπολογιστική δύναμη και η μνήμη των υπολογιστών έχει αυξηθεί, οι αλγόριθμοι έχουν εξελιχθεί και άρα δύνανται να διαχειριστούν το τεράστιο πλήθος δεδομένων

Η τεχνητή νοημοσύνη αποτελεί ένα από τα ταχύτερα εξελισσόμενα πεδία της σύγχρονης επιστήμης και αναμένεται να έχει καθοριστικό ρόλο ως τεχνολογία του μέλλοντος. Με τον όρο τεχνητή νοημοσύνη νοείται ο κλάδος της πληροφορικής που έχει ως αντικείμενο την έρευνα, τη σχεδίαση και την υλοποίηση υπολογιστικών συστημάτων που έχουν ικανότητα αναπαραγωγής των ανθρώπινων γνωστικών λειτουργιών όπως είναι η μάθηση, ο σχεδιασμός, η δημιουργικότητα, η επίλυση προβλημάτων. Ουσιαστικά, πρόκειται για τεχνικές οι οποίες προσπαθούν να μιμηθούν την ανθρώπινη συμπεριφορά. Με την ΤΝ είναι δυνατό σε μια μηχανή να «μάθει» από την εμπειρία. Οι μηχανές ανταποκρίνονται σε νέα δεδομένα και ερεθίσματα, εκτελώντας ενέργειες αντίστοιχες με εκείνες που πραγματοποιεί ο ανθρώπινος εγκέφαλος. Η ΤΝ μπορεί να εκπαιδευτεί έτσι ώστε να επιτελεί εργασίες με κατάλληλη επεξεργασία μεγάλου πλήθους δεδομένων και αναγνωρίζοντας πρότυπα σε αυτά.

Η τεχνητή νοημοσύνη όπως φαίνεται και στην Εικόνα 2-1 χωρίζεται σε δύο υποσύνολα τη Μηχανική Μάθηση (Machine Learning) και την Βαθιά Μάθηση (Deep Learning) με τη βαθιά μάθηση να είναι ένα υποσύνολο της ΜΜ η οποία με τη σειρά της, είναι ένα υποσύνολο της Τεχνητής Νοημοσύνης



Εικόνα 2–1: Η βαθιά μάθηση ως υποσύνολο της μηχανικής μάθησης και οι δύο σαν υποσύνολα της τεχνητής νοημοσύνης (πηγή:www.intel.la.com)

Παρότι οι πρώτες εμφανίσεις της τεχνητής νοημοσύνης έγιναν περί τα 60 χρόνια πριν, η ανάπτυξή της είναι αλματώδης ως αποτέλεσμα της εξέλιξης των ηλεκτρονικών υπολογιστών, του πλήθους των δεδομένων και νέων αλγορίθμων. Ήδη οι εφαρμογές της επηρεάζουν σημαντικά τη σύγχρονη καθημερινότητα και έχουν συντελέσει στην ψηφιακή μεταμόρφωση της κοινωνίας. Αναμένεται να μεταμορφώσει όλες τις πτυχές της οικονομίας και της καθημερινότητας με τους κλάδους της υγείας, της γεωργίας, των μεταφορών, της δημόσιας διοίκησης και υπηρεσιών και του μεταποιητικού κλάδου, ήδη να αξιοποιούν εφαρμογές της.

2.2 Μηχανική μάθηση (Machine Learning-ML)

Η μηχανική μάθηση (MM) πρωτοεμφανίστηκε στα τέλη της δεκαετίας του '80 με αρχές του '90, και εξελίχθηκε προκειμένου να επιλυθούν κάποια βασικά προβλήματα διαφόρων επιστημονικών πεδίων. Στον τομέα της στατιστικής το πρόβλημα ήταν πώς να εκπαιδευτεί αποτελεσματικά ένα μεγάλο και περίπλοκο μοντέλο, στον τομέα της επιστήμης υπολογιστών και της τεχνητής νοημοσύνης το πρόβλημα ήταν πώς να εκπαιδευτεί ένα πιο ισχυρό σύστημα τεχνητής νοημοσύνης, ενώ στον τομέα της νευροεπιστήμης οι ερευνητές προσπαθούσαν να σχεδιάσουν ένα μοντέλο λειτουργίας του εγκεφάλου.

Η MM αποτελεί υποσύνολο αλλά και κινητήρια δύναμη της τεχνητής νοημοσύνης. Πλέον η MM δεν πρόκειται για μια συμβολική προσέγγιση, αλλά έχει υιοθετήσει αρκετά στατιστικά στοιχεία και μεθόδους ανάλυσης της θεωρίας πιθανοτήτων. Έχει εξελιχθεί ως η επιστήμη ανάλυσης δεδομένων που αυτοματοποιεί τη δημιουργία αναλυτικών μοντέλων. Τα αναλυτικά μοντέλα και οι υπολογιστικοί

αλγόριθμοι εκπαιδεύονται και βελτιώνουν την επίδοση τους βάσει δεδομένων. Κύριος στόχος τους είναι η λήψη αποφάσεων, ο προσδιορισμός προτύπων και η δημιουργία προβλέψεων, βάσει της στατιστικής, της θεωρίας των πιθανοτήτων και της βελτιστοποίησης. Επομένως, η MM είναι ένα υποσύνολο της TN το οποίο χρησιμοποιώντας στατιστικές μεθόδους δίνει τη δυνατότητα στα υπολογιστικά συστήματα, με ελάχιστη ανθρώπινη παρέμβαση, να βελτιώνουν την απόδοσή τους όσο δέχονται νέα ερεθίσματα και άρα όσο αποκτούν περισσότερη εμπειρία.

Από μια αφαιρετική σκοπιά, με την τεχνητή νοημοσύνη να έχει ως στόχο τη μίμηση των ανθρώπινων ικανοτήτων, η MM είναι ο κλάδος της που εκπαιδεύει μια μηχανή πώς να μαθαίνει.

Σε αντίθεση με τα στατιστικά μοντέλα τα οποία βασίζονται σε μια θεωρία που είναι μαθηματικά αποδεδειγμένη και απαιτούν από τα δεδομένα να ικανοποιούν ορισμένες προϋποθέσεις, η MM χρησιμοποιεί τους υπολογιστές ώστε να ανιχνεύσουν τα δεδομένα για την εύρεση μοτίβων και χαρακτηριστικών σε τεράστιο όγκο δεδομένων, ακόμα και αν δεν υπάρχει θεωρία για τα μοτίβα αυτά. Αυτό επιτυγχάνεται μέσω επαναλήψεων, ώστε να βρεθεί το βέλτιστο πρότυπο-μοντέλο που θα δίνει το καλύτερο αποτέλεσμα.

Ένας επίσημος ορισμός της μηχανικής μάθησης αναφέρει :

Ένα υπολογιστικό πρόγραμμα θεωρείται ότι αποκτά εμπειρία E (experience) για δεδομένο εγχείρημα (task) T με απόδοση P (performance), αν η απόδοσή του P αυξάνεται όσο μεγαλώνει η εμπειρία του E [et al., Tom.M.Mitchell 1997].

Συνήθως η εμπειρία E είναι ένα δείγμα δεδομένων στο οποίο ο αλγόριθμος MM καλείται να εκπαιδευτεί προκειμένου να πραγματοποιήσει το εγχείρημα T . Υπάρχει πληθώρα αλγορίθμων MM, αναλόγως του προβλήματος. Τα ΝΔ, χρησιμοποιούνται σαν το υπολογιστικό εργαλείο των προβλημάτων MM. Η διαδικασία της εκπαίδευσης των ΝΔ είναι άμεσα εξαρτώμενη από την βελτιστοποίηση του αντικειμενικού σφάλματος και άρα η επιλογή της μεθόδου βελτιστοποίησης είναι εξαιρετικής σημασίας κατά τον σχεδιασμό του ΝΔ.

Για τις βιομηχανίες όπου η αποδοτικότητα τους εξαρτάται από την ανάλυση μεγάλου όγκου δεδομένων, η τεχνολογία της MM είναι ένα απίστευτα χρήσιμο εργαλείο. Το εργαλείο αυτό αξιοποιείται ήδη σε βιομηχανίες υγείας, λιανεμπορίου, μεταφορών χρηματοπιστωτικών και κυβερνητικών υπηρεσιών, για τη εύρεση νέων πηγών ενέργειας πετρελαίου και φυσικού αερίου.

Ορισμένα παραδείγματα MM την τελευταία δεκαετία, είναι τα αυτο-οδηγούμενα αυτοκίνητα (με χαρακτηριστικό παράδειγμα τα μοντέλα της Tesla), έμπρακτη αναγνώριση ομιλίας, τα φίλτρα ανεπιθύμητης αλληλογραφίας, η αποτελεσματική αναζήτηση στο διαδίκτυο (με χαρακτηριστικά

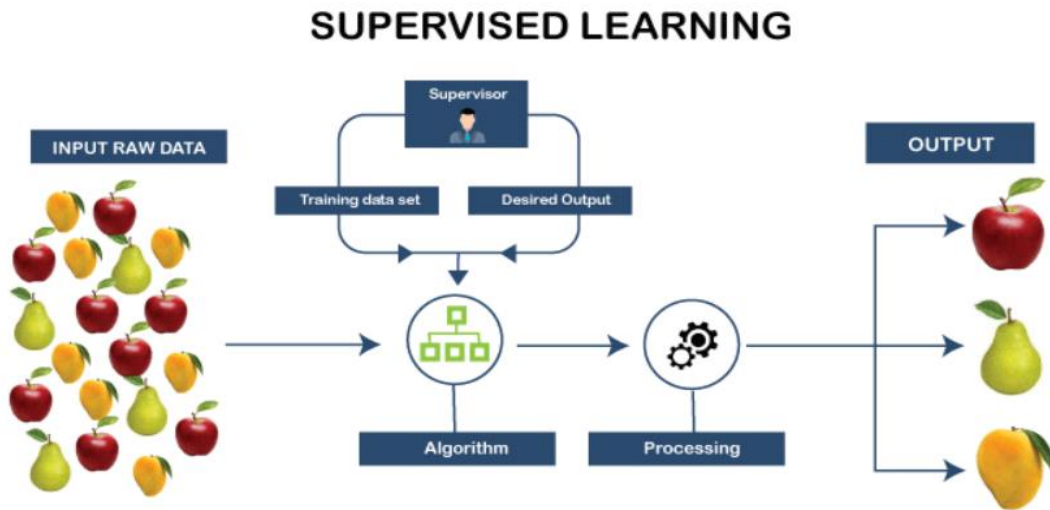
παραδείγματα από την καθημερινή ζωή τις διαδικτυακές πλατφόρμες Google, Amazon, Netflix όπου με χρήση της MM παρέχουν αποτελέσματα αναζήτησης βάσει αλγορίθμων που αναλύουν το ιστορικό αναζήτησης, αγορών και προβολής ενός χρήστη ώστε να γίνονται όσο το δυνατόν οι βέλτιστες προβλέψεις σχετικά με τις επιθυμίες του χρήστη) και μια σημαντικά βελτιωμένη κατανόηση του ανθρώπινου γονιδιώματος. Ωστόσο, οι εφαρμογές της είναι πολυάριθμες και διάχυτες στη σύγχρονη καθημερινότητα.

Όπως αποφαίνεται και από τα παραπάνω στη σύγχρονη εποχή υπάρχει μεγάλη ποικιλία από διαφορετικά προβλήματα MM, τα οποία και κατηγοριοποιούνται σε 3 βασικές κατηγορίες: την επιβλεπόμενη μάθηση (supervised learning), τη μη-επιβλεπόμενη μάθηση (unsupervised learning) και την ενισχυτική μάθηση (reinforcement learning). Σε αυτές τις βασικές κατηγορίες προστίθεται και η μερικώς επιβλεπόμενη μάθηση (semi-supervised learning).

2.2.1 Επιβλεπόμενη μάθηση

Κατά την επιβλεπόμενη μηχανική μάθηση, στο σύστημα δίδεται ένα σύνολο δεδομένων εισόδου (inputs), καθώς και η ετικέτα ή αποτέλεσμα (label or output) με το οποίο κάθε δείγμα (sample) πρέπει να αντιστοιχηθεί στη φάση της εκπαίδευσης. Το στοιχείο εξόδου κάθε δείγματος μπορεί να είναι αριθμητική ή ποιοτική τιμή.

Αναλυτικότερα, κατά την εκπαίδευση το σύστημα τροφοδοτείται με ένα σύνολο από ζεύγη στοιχείων εισόδου και στοιχείων εξόδου/ετικετών, που αποτελούν τα δεδομένα εκπαίδευσης (training data). Για κάθε δείγμα το δίκτυο υπολογίζει μια εκτιμώμενη έξοδο την οποία στη συνέχεια συγκρίνει με την επιθυμητή, με τη διαφορά τους να ορίζεται ως σφάλμα. Το σφάλμα, ανάλογα με τον αλγόριθμο εκπαίδευσης που χρησιμοποιεί το δίκτυο, αξιοποιείται έτσι ώστε η νέα έξοδος να συγκλίνει όσο το δυνατόν περισσότερο με την επιθυμητή. Κατά αυτόν τον τρόπο, το δίκτυο προσαρμόζεται στα δεδομένα εισόδου και εκπαιδεύεται να παράγει τιμές ίσες με την ζητούμενη έξοδο, όπως φαίνεται και στην Εικόνα 2-2.



Εικόνα 2–2: Η ροή της διαδικασίας της επιβλεπόμενης μάθησης (πηγή: www.tutorialandexample.com)

Η επιβλεπόμενη μηχανική μάθηση χρησιμοποιείται συνήθως σε εφαρμογές όπου τα ιστορικά δεδομένα προβλέπουν πιθανά μελλοντικά γεγονότα/αποτελέσματα. Ωστόσο, στην περίπτωση που το εκπαιδευμένο σύστημα τροφοδοτηθεί με δεδομένα που παρουσιάζουν σημαντική απόκλιση από τα δεδομένα εκπαίδευσης, οι προβλέψεις ενδέχεται να μην ανταποκρίνονται στην πραγματικότητα.

Ανάλογα με το αν τα στοιχεία εξόδου επιλέγονται από ένα μικρό πεπερασμένο σύνολο ή από ένα μεγάλο ή μικρό σύνολο συνεχών στοιχείων, η επιβλεπόμενη μηχανική μάθηση χωρίζεται στα προβλήματα ταξινόμησης και παλινδρόμησης.

2.2.1.1 Ταξινόμηση (Classification)

Τα δεδομένα εκπαίδευσης D_n είναι σε μορφή ζευγών $\{(x^{(1)}, y^{(1)}), \dots, (x^{(n)}, y^{(n)})\}$, όπου το $x^{(i)}$ συνήθως είναι ένα διάνυσμα d-διαστάσεων πραγματικών και/ή διακριτών τιμών και αναπαριστά ένα στοιχείο που θα ταξινομηθεί, και το $y^{(i)}$ είναι ένα στοιχείο από ένα διακριτό σετ τιμών. Στόχος του προβλήματος ταξινόμησης είναι για ένα νέο στοιχείο εισόδου $x^{(n+1)}$ το μοντέλο να προβλέψει την έξοδο $y^{(n+1)}$.

Ένα πρόβλημα ταξινόμησης είναι δυαδικό (binary) όταν το $y^{(i)}$ επιλέγεται από ένα σύνολο δύο μόνο πιθανών τιμών.

Πιο συγκεκριμένα, κατά την ταξινόμηση εκτελείται από το σύστημα μια χαρτογράφηση από ένα χώρο χαρακτηριστικών X σε ένα σύνολο ετικετών Y . Αξίζει να αναφερθεί η διαφορά με την συσταδοποίηση

(clustering) που ανήκει στη μη επιβλεπόμενη μάθηση και που ο αλγόριθμος χωρίζει τα δεδομένα αυτόνομα σε συστάδες (clusters) έτσι ώστε τα δεδομένα κάθε συστάδας να έχουν κοινά στο χώρο χαρακτηριστικών.

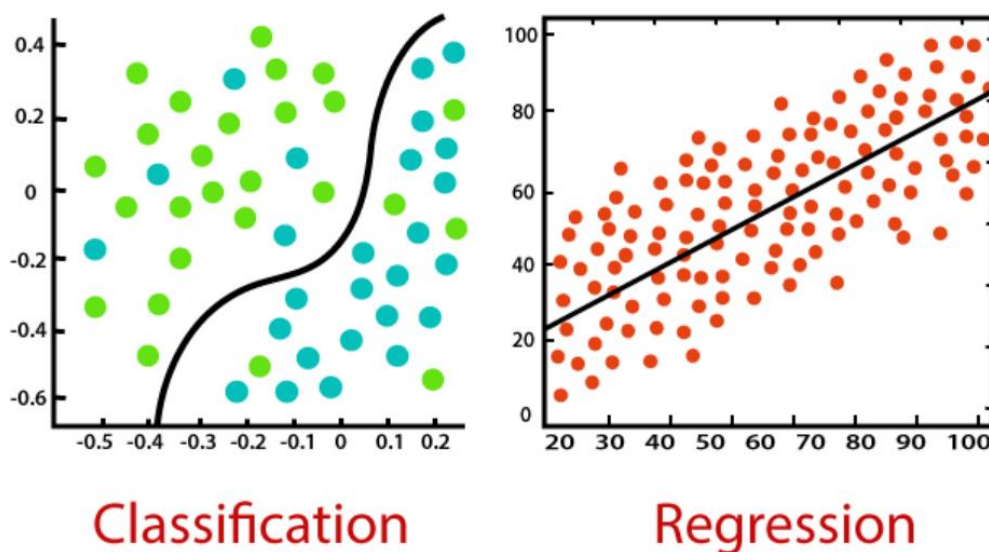
2.2.1.2 Παλινδρόμηση (Regression)

Το πρόβλημα παλινδρόμησης αποτελεί ένα είδος επιβλεπόμενης μηχανικής μάθησης και μοιάζει με το πρόβλημα ταξινόμησης, με τη διαφορά ότι $y^{(i)} \in \mathbb{R}^k$.

Η παλινδρόμηση χωρίζεται σε γραμμική (linear Regression) και λογιστική (logistic Regression). Η σημαντικότερη διαφοροποίηση μεταξύ λογιστικής και γραμμικής παλινδρόμησης βασίζεται στη φύση της επιλεγμένης μεταβλητής απόκρισης, η οποία στην μεν πρώτη μπορεί να είναι κατηγορική στη δευτέρα αποκλειστικά ποσοτική.

Σκοπός της γραμμικής παλινδρόμησης είναι η πρόβλεψη της τιμής του y με βάση κάποιο διάνυσμα – είσοδο $x \in \mathbb{R}^n$. Το αποτέλεσμα εξόδου είναι μία συνεχής ποσοτική μεταβλητή, για παράδειγμα μεταβλητές όπως το «βάρος» ή «ευρώ», για παράδειγμα η πρόβλεψη της τιμής y σε ευρώ ενός σπιτιού, βάσει των χαρακτηριστικών του x (το μέγεθος, παλαιότητα κτλ).

Στην περίπτωση που η μεταβλητή y είναι διακριτή και όχι συνεχής, όπως για παράδειγμα ένα πλέγμα από pixel με τιμές ή , το πρόβλημα τάσσεται στην κατηγορία της ταξινόμησης (classification). Η διαφορά μεταξύ ταξινόμησης και παλινδρόμησης διαφαίνεται και στην Εικόνα 2–3.



Εικόνα 2–3: Διαφορά μεταξύ ταξινόμησης και παλινδρόμησης (πηγή: www.javatpoint.com)

Η Λογιστική Παλινδρόμηση είναι ένας απλός αλγόριθμος εκμάθησης για τη λήψη τέτοιων αποφάσεων. Οι κατηγορικές μεταβλητές απόκρισης μπορούν να είναι είτε τακτικές, δηλαδή οι τιμές τους ανήκουν σε κατηγορίες που μπορούν να μπουν σε μία διάταξη-σειρά, είτε ονομαστικές, όπου δεν μπορεί να γίνει κάποια σύγκριση, αριθμητική πράξη ή υπολογισμός στατιστικών μέτρων, είτε τέλος τη να έχουν δυο μόνο κατηγορίες (διωνυμική παλινδρόμηση).

Η λογιστική παλινδρόμηση ερευνά το μη γραμμικό αποτέλεσμα μίας εξαρτημένης κατηγορικής μεταβλητής y συναρτήσει πολλών ανεξάρτητων μεταβλητών. Χαρακτηρίζεται, αναλόγως της φύσης των κατηγοριών της εξαρτημένης μεταβλητής, από τρεις κατηγορίες μοντέλων, την τακτική (οι κατηγορίες διατάσσονται με αυξητική τάση) και την ονομαστική (ποιοτικές κατηγορίες) .

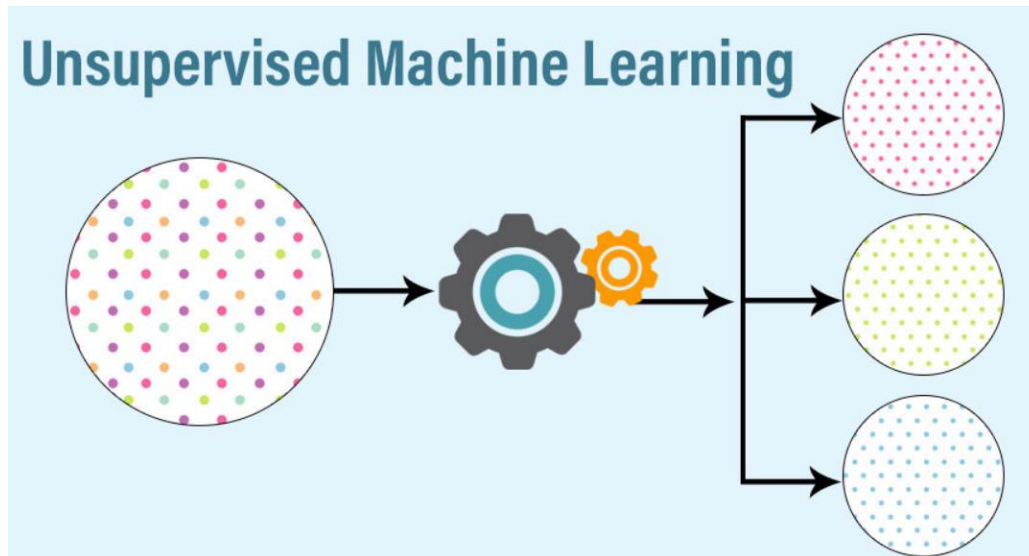
2.2.2 Μη επιβλεπόμενη μάθηση

Στη μη-επιβλεπόμενη μάθηση, στόχος είναι η ομαδοποίηση και η αναζήτηση συσχετίσεων και κοινών χαρακτηριστικών σε ένα σύνολο δεδομένων εκπαίδευσης για τα δείγματα του οποίου δεν υπάρχουν ετικέτες.

Αυτό επιτυγχάνεται με αλγορίθμους οι οποίοι μοντελοποιούν συναρτήσεις πυκνότητας πιθανότητας στα άγνωστα δεδομένα. Η μη-επιβλεπόμενη μάθηση χρησιμοποιείται κυρίως σε προβλήματα:

- Ανάλυσης συσχετίσεων (Association analysis): Αναζήτηση και εξαγωγή συσχετίσεων μεταξύ διαφορετικών αντικειμένων.
- Ομαδοποίησης (Clustering): Οργάνωση των διαφορετικών στοιχείων ενός συνόλου δεδομένων σε διακριτές ομάδες, βάσει κάποιου κοινού χαρακτηριστικού τους. Το χαρακτηριστικό ως προς το οποίο γίνεται η ομαδοποίηση συχνά αποτελεί επίσης ζητούμενο του προβλήματος.

Σχηματική αναπαράσταση της λειτουργίας της μη-επιβλεπόμενης μάθησης δίδεται στην Εικόνα 2–4.

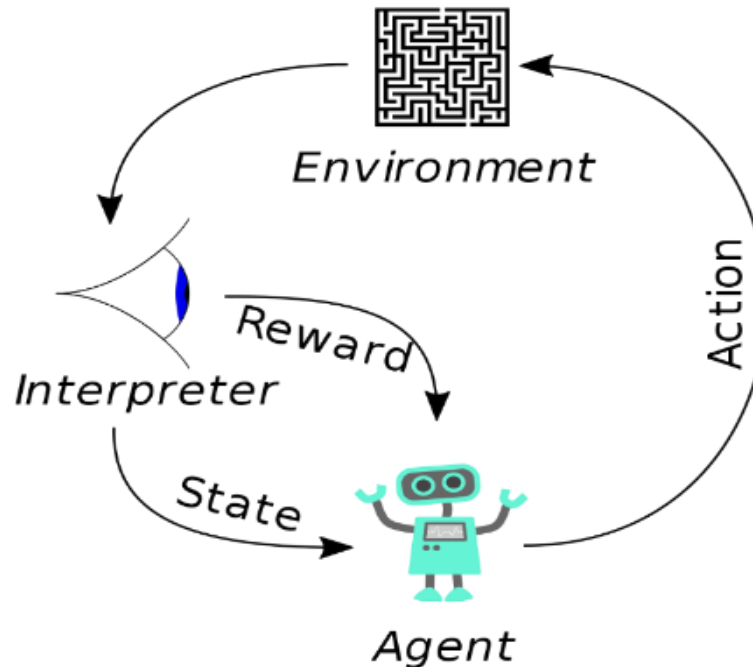


Εικόνα 2–4: Η ροή της διαδικασίας της μη επιβλεπόμενης μηχανικής μάθησης (πηγή: www.educba.com)

Η μη-επιβλεπόμενη μάθηση λειτουργεί αποτελεσματικά σε δεδομένα συναλλαγών. Για παράδειγμα, μπορεί να προσδιορίζει ομάδες πελατών με παρόμοια χαρακτηριστικά, που μπορούν να προσεγγίζονται με παρόμοιο τρόπο από καμπάνιες μάρκετινγκ.

2.2.3 Ενισχυτική μάθηση

Στόχος της ενισχυτικής μάθησης, είναι ο αλγόριθμος να ανακαλύψει μέσω δοκιμής και σφάλματος την καλύτερη «τακτική», δηλαδή ποιες ενέργειες αντιστοιχούν στα βέλτιστα αποτελέσματα, χωρίς να έχει προηγηθεί καθοδήγηση. Η ουσία αυτού του τύπου μάθησης είναι πώς ο παράγοντας (ο εκπαιδευόμενος ή ο υπεύθυνος λήψης αποφάσεων) αλληλοεπιδρά με το περιβάλλον του. Αρχικά παρατηρεί το περιβάλλον του (οτιδήποτε τον επηρεάζει), εν συνεχεία επιλέγει μια ενέργεια για την οποία λαμβάνει από έναν διερμηνέα ένα «βραβείο» και κατόπιν ενημερώνει την κατάστασή του, όπως αναπαρίσταται στην Εικόνα 2–5.



Εικόνα 2–5: Τυπική ροή της διαδικασίας της ενισχυτικής μηχανικής μάθησης (πηγή: www.wikipedia.com)

Στόχος είναι ο παράγοντας να επιλέξει την καλύτερη δυνατή πολιτική αποφάσεων έτσι ώστε οι ενέργειες του να έχουν το βέλτιστο αποτέλεσμα για ένα δεδομένο χρονικό διάστημα. Χρησιμοποιείται στη ρομποτική, σε παιχνίδια και στην πλοήγηση.

2.2.4 Μερικώς επιβλεπόμενη μάθηση

Μια ακόμη κατηγορία επιβλεπόμενης μάθησης που για την εκπαίδευση του συστήματος συνδυάζει τόσο δεδομένα που έχουν επισημανθεί με ετικέτες, όσο και άγνωστα δεδομένα. Χρησιμοποιείται στις περιπτώσεις εκείνες όπου το κόστος για την απόκτηση δεδομένων με ετικέτα είναι μεγάλο. Παράδειγμα αποτελεί η αναγνώριση προσώπου σε μια web κάμερα.

2.2.5 Βαθιά μάθηση (Deep learning)

Η βαθιά μάθηση συνδυάζει την υπολογιστική ισχύ τους ειδικούς τύπους και τις αρχιτεκτονικές των νευρωνικών δικτύων, ώστε να μαθαίνει αρκετά περίπλοκα πρότυπα από σύνθετα δεδομένα. Οι τεχνικές βαθιάς μάθησης για την αναγνώριση αντικειμένων σε εικόνες και λέξεων σε ήχους είναι αρκετά

πρόσφατες και γίνεται προσπάθεια να αξιοποιηθούν σε πιο περίπλοκες εφαρμογές όπως αυτόματη μετάφραση γλώσσας, ιατρικές διαγνώσεις αλλά και σε άλλα κοινωνικο-οικονομικά ζητήματα

Η Βαθιά Μάθηση είναι μια εξειδικευμένη μορφή μηχανικής μάθησης. Αξίζει να σημειωθεί ότι οι διαφορές μεταξύ της MM (machine learning) και των νευρωνικών δικτύων βαθιάς μάθησης (deep learned neural networks) είναι ότι στην πρώτη χρησιμοποιούνται αλγόριθμοι οι οποίοι αναλύουν δεδομένα, εκπαιδεύονται από αυτά και πραγματοποιούν προβλέψεις βάσει της εκπαίδευσής τους. Η MM είναι διαδικασία που εκτελείται βήμα-βήμα, η εξαγωγή των χαρακτηριστικών γίνεται από τον άνθρωπο, η ταξινόμηση των στοιχείων είναι αυτόματη και λειτουργεί με λιγότερα δεδομένα συγκριτικά με τη βαθιά μάθηση. Η βαθιά μάθηση είναι εξ ολοκλήρου αυτόματη διαδικασία όπου τα δεδομένα εκπαιδεύουν τους νευρώνες και τα πιθανά σφάλματα κατά την εκπαίδευση βοηθούν στην προσαρμογή του μοντέλου. Άρα η βαθιά μάθηση δομεί αλγόριθμους σε πολλά επίπεδα δημιουργώντας ένα ΤΝΔ που έχει τη δυνατότητα μάθησης και λήψης αποφάσεων από μόνο του.

ΚΕΦΑΛΑΙΟ 3. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

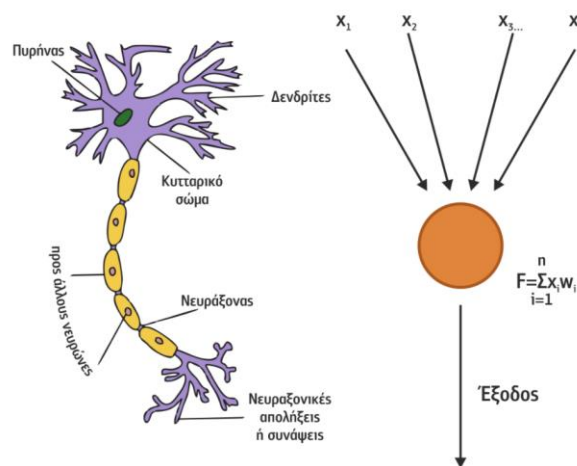
3.1 Βιολογικοί και τεχνητοί νευρώνες

Τα τεχνητά νευρωνικά δίκτυα είναι εμπνευσμένα και επιχειρούν να μιμηθούν τα βιολογικά νευρωνικά δίκτυα που απαντώνται στο κεντρικό νευρικό σύστημα των θηλαστικών και κυρίως στην περιοχή του εγκεφάλου τους. Ο βιολογικός ιστός, οι χημικές ουσίες και τα ηλεκτρικά σήματα των ΒΝΔ, αντιστοιχίζονται στα ΤΝΔ με ένα σύνολο ηλεκτρονικών και μηχανικών συστημάτων συνοδευόμενων από ευφυείς αλγορίθμους η αρχιτεκτονική των οποίων επιχειρεί να προσομοιάσει εκείνη του εγκεφάλου. Ουσιαστικά, ένα ΤΝΔ είναι ένα υπολογιστικό σύστημα που αποτελείται από ένα συνδυασμό συναρτήσεων $f^{(i)}$ που έχουν ως είσοδο κάποια δεδομένα x και δίνουν ως έξοδο τα δεδομένα y .

Τα ΤΝΔ είναι μια τεχνολογία την οποία η ερευνητική κοινότητα διαρκώς εξελίσσει καθότι για την ώρα οι μηχανισμοί και οι λειτουργίες του ανθρώπινου εγκεφάλου είναι αρκετά πιο σύνθετες και πιο περίπλοκες.

Τα κοινά χαρακτηριστικά των δυο κατηγοριών νοημοσύνης συνοψίζονται στην Εικόνα 3–1 και είναι τα εξής:

- Η εκπαίδευσή τους γίνεται μέσα από ένα πλήθος παραδειγμάτων, έτσι ώστε να κατανοήσουν το περιβάλλον τους και να βελτιώσουν την απόδοσή τους
- Στα σημεία συνάψεων των ΒΝΔ, αποθηκεύεται η αποκτηθείσα γνώση, ανάλογα με την ισχύ των δεσμών τους. Κάτι αντίστοιχο συμβαίνει με τα συναπτικά βάρη στα ΤΝΔ.



Εικόνα 3–1: Από τους βιολογικούς στους τεχνητούς νευρώνες (πηγή: [www.Μηχανική Μάθηση \(kallipos.gr\)](http://www.Μηχανική Μάθηση (kallipos.gr)))

Συμπερασματικά, τα τεχνητά και τα βιολογικά ΝΔ είναι ένα σύνολο νευρώνων, όπου κάθε νευρώνας αποτελεί μια μονάδα αποθήκευσης και επεξεργασίας πληροφορίας. Τα βασικά στοιχεία του νευρώνα είναι ένα σύνολο συνάψεων, ένας αθροιστής και μια συνάρτηση ενεργοποίησης. Μπορεί να έχει μεγάλο αριθμό εισόδων αλλά μία έξοδο, η οποία στην περίπτωση του νευρωνικού δικτύου, ενδέχεται να είναι είσοδος άλλου νευρώνα. Οι συνάψεις (βάρη) διαφέρουν μεταξύ των νευρώνων, ενώ η συνάρτηση ενεργοποίησης καθορίζει την έξοδο του νευρώνα.

3.2 Ο αλγόριθμος – Perceptron

Βασισμένος στο μοντέλο νευρώνα των McCulloch-Pitts, ο αλγόριθμος – Perceptron αναπτύχθηκε το 1962 από τον ψυχολόγο Rosenblatt και αποτέλεσε τη βάση για την περαιτέρω εξέλιξη των νευρωνικών δικτύων. Αποτελεί την απλούστερη δυνατή μορφή ενός νευρωνικού δικτύου και ταξινομεί γραμμικά διαχωρίσιμα πρότυπα. Αποτελείται από ένα μεμονωμένο νευρώνα με συναπτικά βάρη και μία προδιάθεση (πόλωση) που προσαρμόζονται κατά την εκπαίδευση. Για δύο γραμμικά διαχωρίσιμες κλάσεις ο αλγόριθμος Perceptron, προσαρμόζει τις ελεύθερες παραμέτρους, συγκλίνει και τοποθετεί την επιφάνεια απόφασης μεταξύ των δύο επιπέδων με τη μορφή ενός υπερεπίπεδου.

Στην περίπτωση ταξινόμησης προτύπων που ανήκουν σε περισσότερες από δύο κλάσεις, στο επίπεδο εξόδου του perceptron προστίθενται επιπλέον νευρώνες. Για την αποτελεσματική λειτουργία του Perceptron, οι κλάσεις πρέπει να είναι γραμμικά διαχωρίσιμες.

Το perceptron βασίζεται στον μη γραμμικό νευρώνα McCulloch-Pitts. Στο συγκεκριμένο μοντέλο νευρώνα, τα συναπτικά βάρη και οι εισοδοί συνδυάζονται γραμμικά στον κόμβο άθροισης στα οποία ενσωματώνεται και μια εξωτερικά παραγόμενη πόλωση (ή ‘προδιάθεση’). Στη συνέχεια, στο άθροισμα που προκύπτει εφαρμόζεται η συνάρτηση πρόσημου ή κάποια σιγμοειδής συνάρτηση από έναν απότομο περιοριστή. Ως απόκριση ο νευρώνας παράγει έξοδο ίση με +1 εαν η είσοδος του απότομου περιοριστή είναι θετική και -1 εάν είναι αρνητική.

Με δύο δυνατές περιοχές απόφασης το υπερεπίπεδο που τις χωρίζει ορίζεται από την εξίσωση :

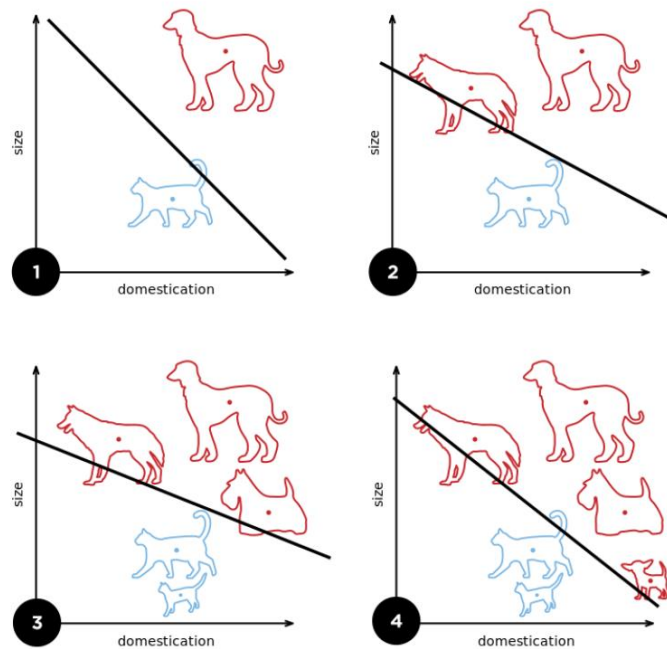
$$\sum_{i=1}^d (\theta^T x) + \theta_0 = 0$$

Όπου, $x \in \mathbb{R}^d$ είναι το διάνυσμα εισόδου, $\theta \in \mathbb{R}^d$ είναι το διάνυσμα των συναπτικών βαρών με διαστάσεις $(d \times 1)$ και $\theta_0 \in \mathbb{R}$ είναι το βαθμωτό μέγεθος της πόλωσης το οποίο εκφράζει την απόσταση του υπερεπιπέδου απόφασης από την αρχή των αξόνων .

Για πλήθος δεδομένων εισόδου D_n με $x \in \mathbb{R}^d$ και για έξοδο $y \in \{-1, +1\}$ ο Perceptron εκπαιδεύει επαναληπτικά τον γραμμικό ταξινομητή για την εύρεση των παραμέτρων θ, θ_0 κάνοντας τ αριθμό επαναλήψεων.

```
PERCEPTRON( $\tau, \mathcal{D}_n$ )
1  $\theta = [0 \ 0 \ \dots \ 0]^T$ 
2  $\theta_0 = 0$ 
3 for t = 1 to  $\tau$ 
4   for i = 1 to n
5     if  $y^{(i)} (\theta^T x^{(i)} + \theta_0) \leq 0$ 
6        $\theta = \theta + y^{(i)} x^{(i)}$ 
7        $\theta_0 = \theta_0 + y^{(i)}$ 
8 return  $\theta, \theta_0$ 
```

Σε κάθε επαναληπτικό βήμα, αν οι παράμετροι ως έχουν ταξινομούν ορθά το δείγμα $x^{(i)}$, δεν πραγματοποιούνται αλλαγές στις τιμές τους. Στην αντίθετη περίπτωση, τα θ, θ_0 προσαρμόζονται έτσι ώστε η ταξινόμηση του δείγματος να είναι η αναμενόμενη. Μια σχηματική αναπράσταση της διαδικασίας είναι η Εικόνα 3-2



Εικόνα 3–2: Στην εικόνα αναπαριστάται ο τρόπος με τον οποίο ο αλγόριθμος Perceptron διαμορφώνει τη γραμμή εκπαίδευσης όσο μεγαλώνει το δείγμα εκπαίδευσης (πηγή: www.wikipedia.com)

3.3 Λειτουργίες τεχνητών νευρωνικών δικτύων

Συνοπτικά, τα ΤΝΔ επιτελούν δύο βασικές λειτουργίες:

- Εκπαίδευση (Training): Κατά την εκπαίδευση, το δίκτυο τροφοδοτείται από μία συλλογή δεδομένων εκπαίδευσης (training dataset). Για το σύνολο των δεδομένων αυτών, επαναληπτικά, το δίκτυο ενημερώνει την εσωτερική του κατάσταση ούτως ώστε αυτή να ανταποκρίνεται όσο το δυνατόν πιο πιστά στα δεδομένα.
- Συμπερασματολογία (Inference): Το δίκτυο δέχεται ένα ερέθισμα εισόδου και παράγει μία έξοδο, αξιοποιώντας την γνώση που έχει συσσωρεύσει κατά την εκπαίδευση.

3.4 Τεχνητοί νευρώνες

Στοιχειώδης δομική μονάδα των τεχνητών νευρωνικών δικτύων είναι οι τεχνητοί νευρώνες ή μονάδες (artificial neurons ή units). Πρόκειται για εννοιολογικές δομές η μορφή και η λειτουργία των οποίων είναι εμπνευσμένη από τη μορφή και την λειτουργία των νευρώνων του ανθρώπινου εγκεφάλου, όπως αυτοί είναι γνωστοί από την βιολογία.

Πρόκειται για μια μη γραμμική συνάρτηση με είσοδο ένα διάνυσμα m στοιχείων, $x \in \mathbb{R}^m$, και έξοδο το βαθμωτό μέγεθος $a \in \mathbb{R}$. Η συνάρτηση έχει παραμέτρους το διάνυσμα των συναπτικών βαρών (weights) με m στοιχεία, $(w_1, \dots, w_m) \in \mathbb{R}^m$ και την σταθερή τιμή της πόλωσης (bias) $w_0 \in \mathbb{R}$. Επιπλέον, ορίζεται μία συνάρτηση ενεργοποίησης (activation function) $f : \mathbb{R} \rightarrow \mathbb{R}$, υποχρεωτικά παραγωγίσιμη, η οποία εξασφαλίζει τη μη γραμμικότητα του νευρώνα.

Επομένως, τα βασικά συστατικά στοιχεία ενός τεχνητού νευρώνα είναι:

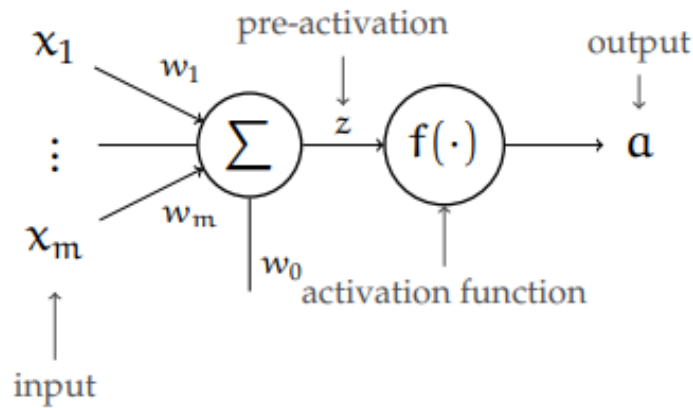
- Ένα σύνολο συνάψεων εισόδου, κάθε μία από τις οποίες αντιστοιχίζεται με ένα βάρος w_i , το οποίο αντιστοιχεί στην σημαντικότητα, ή στην συνεισφορά της εκάστοτε εισόδου στην ενεργοποίηση του νευρώνα.
- Ένας κεντρικός κόμβος υπολογισμού, όπου για κάθε σύναψη υπολογίζεται το γινόμενο της εισόδου x_i με το αντίστοιχο βάρος w_i . Τα συναπτικά βάρη του μοντέλου αρχικά λαμβάνουν κάποιες μικρές τυχαίες τιμές, οι οποίες μέσω του αλγόριθμου βελτιστοποίησης ενημερώνονται και προσαρμόζονται ώστε να ελαχιστοποιείται το σφάλμα της εκπαίδευσης. Τα επιμέρους γινόμενα αθροίζονται, και στο αποτέλεσμα προστίθεται ο σταθερός όρος b , ο οποίος ονομάζεται πόλωση ή κατώφλι. Το άθροισμα που προκύπτει είναι η προ-ενεργοποίηση του νευρώνα, z , και φαίνεται στη Σχέση (3-1)

$$z = \sum_{j=1}^m (x_j w_j + w_0) \quad (3-1)$$

- Στην ενεργοποίηση του νευρώνα εφαρμόζεται η συνάρτηση ενεργοποίησης $f(z)$ και το αποτέλεσμα που προκύπτει είναι η έξοδος του νευρώνα. Η Σχέση(3-2) αναπαριστά τον νευρώνα:

$$a_0 = f(z) = f \sum_{j=1}^m (x_j w_j + w_0) = f (w^T x + w_0) \quad (3-2)$$

Σχηματικά τα βασικά συστατικά στοιχεία ενός τεχνητού νευρώνα απεικονίζονται στην Εικόνα 3–3.



Εικόνα 3-3: Τα βασικά συστατικά στοιχεία ενός τεχνητού νευρώνα
(πηγή: www.openlearninglibrary.mit.edu)

3.5 Νευρωνικά δίκτυα πρόσθιας τροφοδότησης

Με την προσθήκη πολλών νευρώνων δημιουργείται ένα νευρωνικό δίκτυο. Η σύνδεση των νευρώνων μεταξύ τους και η οργάνωσή τους δύναται να γίνει με διαφορετικούς τρόπους σε κάθε δίκτυο. Είσοδοι ενός νευρώνα μπορούν να είναι τα στοιχεία x_j και/ή οι έξοδοι άλλων νευρώνων. Σε γενικές γραμμές ένα νευρωνικό δίκτυο λαμβάνει ως είσοδο ένα διάνυσμα $x \in \mathbb{R}^m$, και παράγει ως έξοδο το βαθμωτό μέγεθος $a \in \mathbb{R}$.

Σε ένα δίκτυο πρόσθιας τροφοδότησης (feed-forward network) το δίκτυο ορίζεται σαν ένας κατευθυνόμενος άκυκλος γράφος, όπου η είσοδος ενός νευρώνα δεν μπορεί να εξαρτάται από την έξοδό του. Η πληροφορία κατευθύνεται προς μία κατεύθυνση, από την είσοδο του δικτύου προς την έξοδο. Η συνάρτηση που πραγματοποιείται από το δίκτυο είναι η σύνθεση των συναρτήσεων που εκτελούν οι επιμέρους μονάδες-νευρώνες. Παρόλο που το γράφημα ενός ΝΔ πρόσθιας τροφοδότησης μπορεί να είναι οποιοδήποτε ικανοποιεί αυτή τη συνθήκη, χάριν απλότητας οι νευρώνες οργανώνονται σε μία σειρά από διαδοχικά επίπεδα (*layers*). Ένα επίπεδο είναι μια ομάδα από «παράλληλα» τοποθετημένους νευρώνες, όπου τα στοιχεία εισόδου x_i ενός νευρώνα j ενός επιπέδου N είναι τα στοιχεία εξόδου των νευρώνων του προηγούμενου επιπέδου $N-1$. Ομοίως, τα στοιχεία εξόδου a_j του δεδομένου επιπέδου θα αποτελέσουν είσοδοι των νευρώνων του επόμενου επιπέδου $N+1$.

Η διάταξη πρόσθιας τροφοδότησης δεν είναι η μοναδική. Υπάρχει πληθώρα αρχιτεκτονικών ανάλογα το πρόβλημα, με τη διάταξη πρόσθιας τροφοδότησης να είναι η πιο συνηθισμένη. Σε άλλες

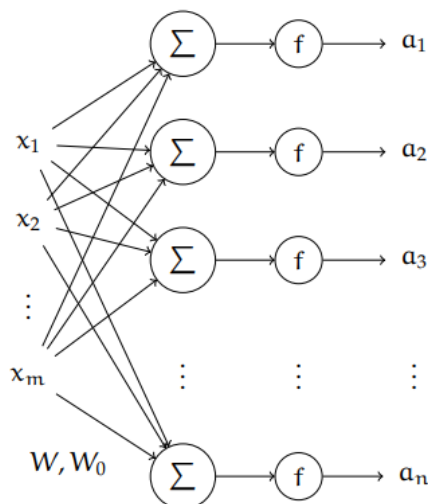
αρχιτεκτονικές πολλοί νευρώνες ενός επιπέδου μπορούν να συνδέονται με έναν μόνο νευρώνα του επόμενου, ή να υπάρχουν νευρώνες οι οποίοι προωθούνται απευθείας σε κάποιο άλλο από τα επόμενα επίπεδα του δικτύου

3.6 Μεμονωμένο επίπεδο (single layer)

Οι νευρώνες που ανήκουν στο ίδιο επίπεδο, N , δεν συνδέονται μεταξύ τους. Ένα επίπεδο ονομάζεται πλήρως συνδεδεμένο (*fully connected*) όταν καθένας από τους νευρώνες του επιπέδου $N-1$ συνδέεται με όλους τους νευρώνες του επιπέδου N , δηλαδή κάθε νευρώνας του επιπέδου N έχει το ίδιο διάνυσμα εισόδου, $x \in \mathbb{R}^m$

Για μεμονωμένο επίπεδο N όπως απεικονίζεται στην Εικόνα 3–4 με $x \in \mathbb{R}^m$ στοιχεία εισόδου σε κάθε νευρώνα, δηλαδή X είναι διάνυσμα ($m \times 1$), n πλήθος νευρώνων και n πλήθος εξόδων του επιπέδου και δεδομένου ότι κάθε νευρώνας j του επιπέδου N έχει ένα διάνυσμα συναπτικών βαρών και ένα μοναδικό κατώφλι, τα συναπτικά βάρη όλου του επιπέδου ορίζονται

- Τα συναπτικά βάρη σαν ένας πίνακας W ($m \times n$)
- το κατώφλι σαν διάνυσμα W_0 ($n \times 1$).
- Η προ- ενεργοποίηση κάθε νευρώνα $Z = w^T x + w_0$, προκύπτει διάνυσμα Z ($n \times 1$), και
- η ενεργοποίηση - έξοδος του νευρώνα $A = f(z) = f(w^T x + w_0)$, σαν διάνυσμα A ($n \times 1$)

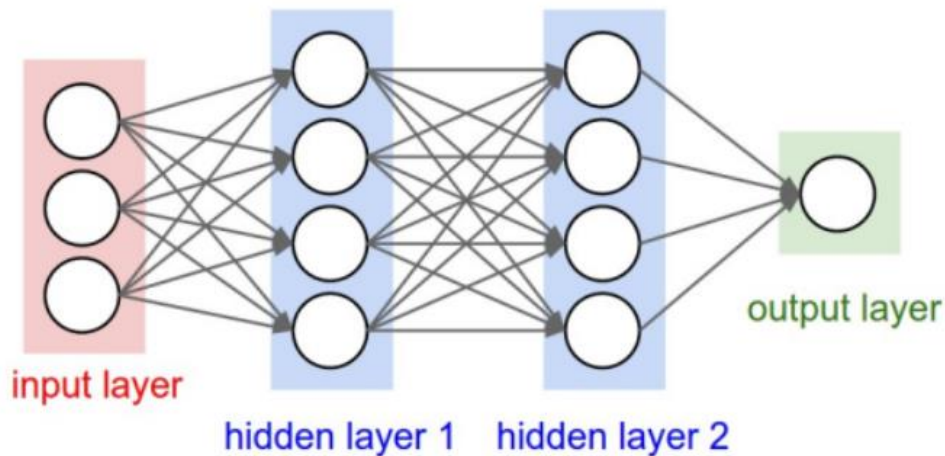


Εικόνα 3–4: Αναπαράσταση μεμονωμένου επιπέδου με $x \in \mathbb{R}^m$ στοιχεία εισόδου σε κάθε νευρώνα και n πλήθος εξόδων του επιπέδου (πηγή: www.openlearninglibrary.mit.edu)

3.7 Πολύ-επίπεδα Νευρωνικά Δίκτυα

Για να εξασφαλιστεί η επίλυση πολύπλοκων προβλημάτων προστίθενται επιπλέον επίπεδα και άρα η συνηθέστερη μορφή οργάνωσης των ΝΔ είναι μια σειρά διαδοχικών επιπέδων, όπου οι έξοδοι ενός επιπέδου είναι τα στοιχεία εισόδου του επόμενου (πολυεπίπεδα ΝΔ πρόσθιας τροφοδότησης). Το πρώτο επίπεδο, λαμβάνει ως ερέθισμα τα δεδομένα εισόδου από το εξωτερικό περιβάλλον του δικτύου, δεν εκτελεί υπολογισμούς απλά παρεμβάλλεται ανάμεσα στις περιβαλλοντικές εισόδους του δικτύου και τους υπολογιστικούς νευρώνες και καλείται επίπεδο εισόδου (input layer). Αντίστοιχα, το τελευταίο επίπεδο είναι αυτό που υπολογίζει και διοχετεύει στο περιβάλλον την έξοδο του δικτύου και ονομάζεται επίπεδο εξόδου (output layer). Τα ενδιάμεσα επίπεδα ονομάζονται υπολογιστικά ή κρυφά επίπεδα (hidden layers) επειδή οι τιμές τους δεν εμφανίζονται στα δεδομένα εκπαίδευσης. Οι νευρώνες των διαφόρων επιπέδων συνδέονται μεταξύ τους μέσω των συνάψεων με τον τρόπο που αναφέρθηκε στην παραπάνω ενότητα.

Στην Εικόνα 3–5 παρουσιάζεται παράδειγμα πολυεπίπεδου ΝΔ πρόσθιας τροφοδότησης τριών εισόδων, μιας εξόδου και δύο κρυφών επιπέδων.



Εικόνα 3–5: Πολυεπίπεδο ΝΔ με 3 εισόδους, 2 κρυφά επίπεδα και μία έξοδο (πηγή: www.GitHub.com)

3.8 Συνάρτηση Ενεργοποίησης

Η συνάρτηση ενεργοποίησης $\varphi(z)$ του νευρώνα είναι μια οποιαδήποτε μη γραμμική συνάρτηση. Συνηθέστερα χρησιμοποιούνται σιγμοειδείς συναρτήσεις, καθότι είναι συνεχείς σε όλο το πεδίο ορισμού τους, φραγμένες, γνησίως μονότονες (συνήθως αύξουσες) και παραγωγίσιμες, εξυπηρετώντας περιπτώσεις που εφαρμόζεται οπισθοδιάδοση του σφάλματος με την εφαρμογή του κανόνα της

αλυσίδας στη σχέση 1.2. Τυπικές συναρτήσεις ενεργοποιήσεις που χρησιμοποιούνται στην πλειοψηφία των εφαρμογών είναι οι εξής .

- Ημιγραμμική (Rectified Linear Unit - ReLU)

$$\varphi(z) = z^+ = \max(0, z) = \begin{cases} 0, & u < 0 \\ z, & u \geq 0 \end{cases}$$

- Υπερβολική Εφαπτομένη (Hyperbolic tangent)

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

- Λογιστική ή Σιγμοειδής (Logistic or Sigmoid function), μπορεί να ερμηνευτεί ως πιθανότητα καθότι για κάθε τιμή του z το σύνολο τιμών της συνάρτησης είναι $[0,1]$

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

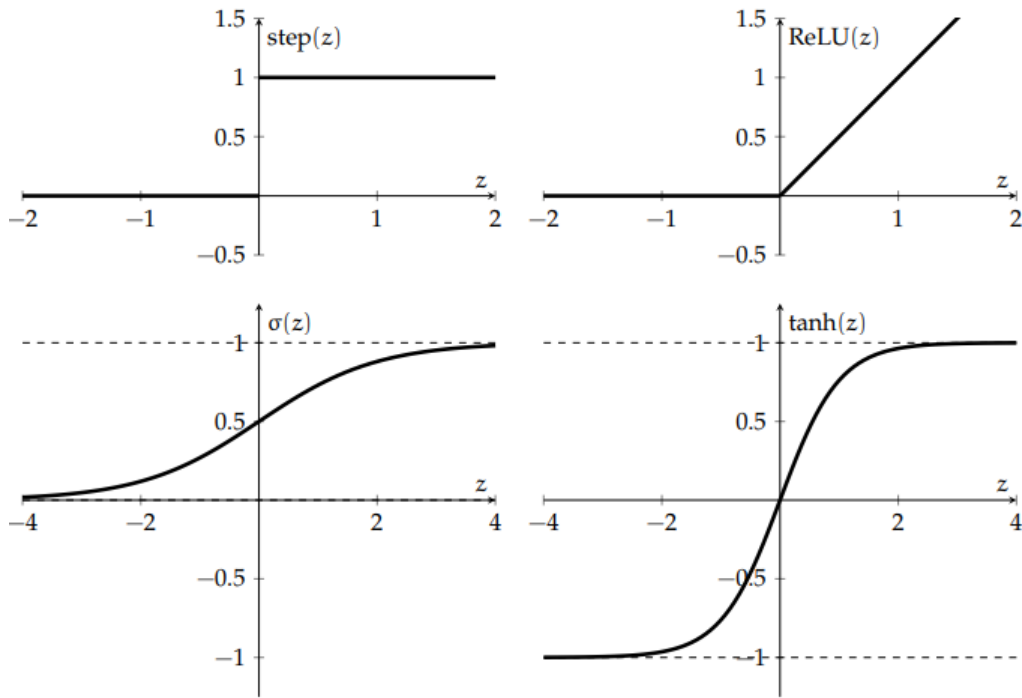
- Βηματική (step)

$$\text{Step}(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases}$$

- Softmax function, χρησιμοποιείται συχνά σαν η τελευταία συνάρτηση ενεργοποίησης σε ένα ΝΔ. Αποτελεί γενίκευση της λογιστικής συνάρτησης και έχει περισσότερες διαστάσεις. Λαμβάνει ως είσοδο ένα διάνυσμα $Z \in \mathbb{R}^n$ και δίνει ως έξοδο διάνυσμα $A \in [0,1]^n$ με την ιδιότητα $\sum_{i=1}^n A_i = 1$. Αρα μπορεί να ερμηνευτεί σαν μια κατανομή πιθανότητας για n στοιχεία.

$$\text{Softmax}(z) = \begin{matrix} \exp(z_1) / \sum_i \exp(z_i) \\ \vdots \\ \exp(z_n) / \sum_i \exp(z_i) \end{matrix}$$

Στην Εικόνα 3–6 παριστάνονται οι κυριότερες μορφές συναρτήσεων ενεργοποίησης. Η διαφορά των συναρτήσεων παρατηρείται στη διαφορετική κλίση τους γύρω από το 0. Η βηματική συνάρτηση έχει παντού μηδενική κλίση εκτός από το 0, ενώ στην ημιγραμμική συνάρτηση η κλίση είναι μηδενική στις αρνητικές τιμές της. Με τη χρήση συνάρτησης με μηδενική κλίση, οι τιμές εξόδου παρουσιάζουν μεγάλη διαφορά για μικρές μεταβολές των τιμών εισόδου. Αντίθετα η λογιστική συνάρτηση και η συνάρτηση υπερβολικής εφαπτομένης έχουν ήπιες μεταβολές στην έξοδο. Μεγάλες αρνητικές τιμές στη λογιστική συνάρτηση δίνουν τιμές πολύ κοντά στο 0, ενώ μεγάλες θετικές τιμές δίνουν τιμές πολύ κοντά στο 1.



Εικόνα 3–6: Γραφικές παραστάσεις κυριότερων συναρτήσεων ενεργοποίησης (πηγή: www.openlearninglibrary.mit.edu)

3.9 Υπερπαράμετροι

Οι υπερπαράμετροι ενός ΝΔ είναι οι σταθερές που καθορίζονται πριν την εκπαίδευση του δικτύου εμπειρικά είτε μέσω δοκιμών. Συμβάλουν στην ορθή λειτουργία του δικτύου και στην ικανότητα εκπαίδευσής του, δηλαδή στις ποιοτικές και ποσοτικές μεταβολές της συμπεριφοράς του.

Οι συνηθέστερες υπερπαράμετροι ενός ΝΔ είναι οι παρακάτω:

- Ο αριθμός των νευρώνων στα επίπεδα εισόδου και εξόδου του δικτύου. Συνήθως ο καθορισμός τους γίνεται από την μορφή των διαθέσιμων δεδομένων εισόδου και εξόδου.
- Ο αριθμός των κρυφών επιπέδων του νευρωνικού δικτύου. Η προσθήκη κρυφών επιπέδων σε ένα νευρωνικό δίκτυο αυξάνει την πολυπλοκότητά του και την δυσκολία της εκπαίδευσής του. Ωστόσο, NN με μεγάλο αριθμό κρυφών επιπέδων αναγνωρίζουν αντικείμενα, πρότυπα, μοτίβα ή να πραγματοποιούν συσχετίσεις υψηλότερου επιπέδου.

- Ο αριθμός των νευρώνων σε κάθε ένα από τα κρυφά επίπεδα του δικτύου. Αντίστοιχα, αύξηση του του αριθμού των νευρώνων συνεπάγεται αύξηση της χωρητικότητας του δικτύου, οπότε δύναται το δίκτυο να αναγνωρίζει πιο σύνθετες μορφές δεδομένων.
- Η συνάρτηση ενεργοποίησης που χρησιμοποιείται στους νευρώνες κάθε επιπέδου του δικτύου.
- Οι αρχικές τιμές των παραμέτρων του δικτύου, δηλαδή οι αρχικές τιμές των βαρών των συνάψεων μέσω των οποίων συνδέονται οι νευρώνες του δικτύου και οι πολώσεις αυτών. Τα βάρη κατά την εκκίνηση αρχικοποιούνται σε μικρές τιμές, συνήθως αυτές είναι τυχαίες και ακολουθούν την κανονική κατανομή με μέση τιμή 0 και τυπική απόκλιση 1. Ο σκοπός της αρχικοποίησης των βαρών είναι να εμποδίσει τις εξόδους ενός επιπέδου από το να γίνουν υπερβολικά μικρές ή υπερβολικά μεγάλες. Αν συμβεί κάτι από τα δύο κατά τη διάρκεια μιας πρόσθιας τροφοδότησης, οι παράγωγοι της συνάρτησης σφάλματος ως προς τα βάρη θα είναι πολύ μεγάλες ή πολύ μικρές, αντίστοιχα (exploding and vanishing gradients, respectively). Καθότι τα βάρη αναβαθμίζονται και προσαρμόζονται αναλογικά με τις παραγώγους της συνάρτησης σφάλματος, στην πρώτη περίπτωση, τα βάρη θα γίνουν υπερβολικά μεγάλα, ενώ στη δεύτερη υπερβολικά μικρά. Άρα η οπισθοδιάδοση του σφάλματος δεν θα είναι αποδοτική και το Δ δεν θα συγκλίνει.
- Ο αλγόριθμος εκπαίδευσης. Πιο συνηθισμένοι είναι οι αλγόριθμοι βελτιστοποίησης κατάβασης κλίσης (gradient descent), ενώ σε περιπτώσεις επιβλεπόμενης μάθησης, χρησιμοποιείται ο αλγόριθμος οπισθοδιάδοσης σφάλματος ή κάποια παραλλαγή αυτού.
- Ο ρυθμός εκπαίδευσης. Μεγαλύτερος ρυθμός εκπαίδευσης οδηγεί γρηγορότερα σε σύγκλιση του δικτύου, με μικρότερη όμως ακρίβεια αποτελεσμάτων. θυσιάζοντας μέρος της ακριβειάς του. Αντίθετα, μικρότερος ρυθμό εκπαίδευσης απαιτεί μεγαλύτερο υπολογιστικό χρόνο άλλα τα αποτελέσματα είναι ακριβέστερα.
- Η συνάρτηση κόστους (cost function) που χρησιμοποιείται κατά την εκπαίδευση. Η συνάρτηση κόστους είναι η μέση τιμή των τιμών της συνάρτησης σφάλματος για κάθε δείγμα της εκπαίδευσης. Η εκπαίδευση ενός νευρωνικού δικτύου αποτελεί κατά βάση ένα πρόβλημα βελτιστοποίησης και συγκεκριμένα στην ελαχιστοποίηση ενός αντικειμενικού σφάλματος. Σε προβλήματα επιβλεπόμενης μάθησης, η συνάρτηση κόστους είναι μία μετρική απόκλισης από την γνωστή επιθυμητή έξοδο, όπως το μέσο τετραγωνικό σφάλμα (mean squared error) ή κόστος διασταυρούμενης εντροπίας (crossentropy loss). Σε προβλήματα μη-επιβλεπόμενης μάθησης, όπου η επιθυμητή έξοδος είναι άγνωστη ως συνάρτηση κόστους χρησιμοποιείται κάποια μετρική της στατιστικής διασποράς των εξόδων του δικτύου.

3.10 Αλγόριθμοι βελτιστοποίησης (Optimization Algorithms)

3.10.1 Πρώτης τάξης μέθοδοι βελτιστοποίησης

Πρώτης τάξης μέθοδοι βελτιστοποίησης είναι μέχρι στιγμής οι πιο δημοφιλείς μέθοδοι βελτιστοποίησης. Αυτές χρησιμοποιούν την κλίση της συνάρτησης σφάλματος και με τη βαθμιαία μείωσή της συγκλίνουν στο επιθυμητό αποτέλεσμα. Χρησιμοποιούν δηλαδή την πρώτη παράγωγο της συνάρτησης κόστους για την εύρεση του βέλτιστου. Ο πιο βασικός τέτοιος αλγόριθμος που λειτουργεί αποτελεσματικά σε εφαρμογές μηχανικής μάθησης είναι ο αλγόριθμος Στοχαστικής Κατάβασης Κλίσης (Stochastic Gradient Descent), ενώ ένας πιο πολύπλοκος αλγόριθμος πρώτης τάξης είναι ο ADAM.

3.10.1.1 Αλγόριθμος κατάβασης κλίσης (Gradient Descent)

Είναι ο αλγόριθμος πάνω στον οποίο βασίζονται αρκετοί αλγόριθμοι βελτιστοποίησης και στόχος του είναι η προσαρμογή των παραμέτρων/βαρών της εξίσωσης κόστους έτσι ώστε αυτή να ελαχιστοποιηθεί. Ειδικότερα, ο αλγόριθμος κατάβασης κλίσης υπολογίζει την αρνητική μερική παράγωγο της συνάρτησης κόστους ως προς το διάνυσμα των παραμέτρων της και προχωρά προς την κατεύθυνση της φθίνουσας κλίσης, άρα ο υπολογισμός της αρνητικής μερικής παραγώγου επαναλαμβάνεται εως ότου βρεθεί το ολικό ή ένα τοπικό ελάχιστο της συνάρτησης. Βασίζεται δηλαδή στο γεγονός ότι η κλίση μιας συνάρτησης πάντα δείχνει προς την κατεύθυνση της μέγιστης αύξησης, οπότε με τη μετακίνηση προς την αντίθετη κατεύθυνση της κλίσης, μπορεί να επιτευχθεί μείωση τη τιμής της συνάρτησης κόστους. Το βήμα με το οποίο ο αλγόριθμος μετακινείται ονομάζεται ρυθμός μάθησης ϵ (learning rate). Ως μετακίνηση, νοείται η αφαίρεση από μια τιμή της παραμέτρου, το γινόμενο αυτής της τιμής της παραμέτρου με τον ρυθμό μάθησης. Ισχύει ότι μικρός ρυθμός μάθησης συνεπάγεται περισσότερες επαναλήψεις και άρα περισσότερος χρόνος. Από την άλλη, για μεγάλους ρυθμούς μάθησης ελλοχύνει ο κίνδυνος κίνδυνος το χαμηλότερο σημείο να προσπεραστεί αφού η κλίση αλλάζει συνεχώς. Αυτό μπορεί να οδηγήσει τον αλγόριθμο σε αδυναμία σύγκλισης απόκλιση και εύρεσης καλής λύσης.

Η μαθηματική περιγραφή του αλγόριθμου είναι:

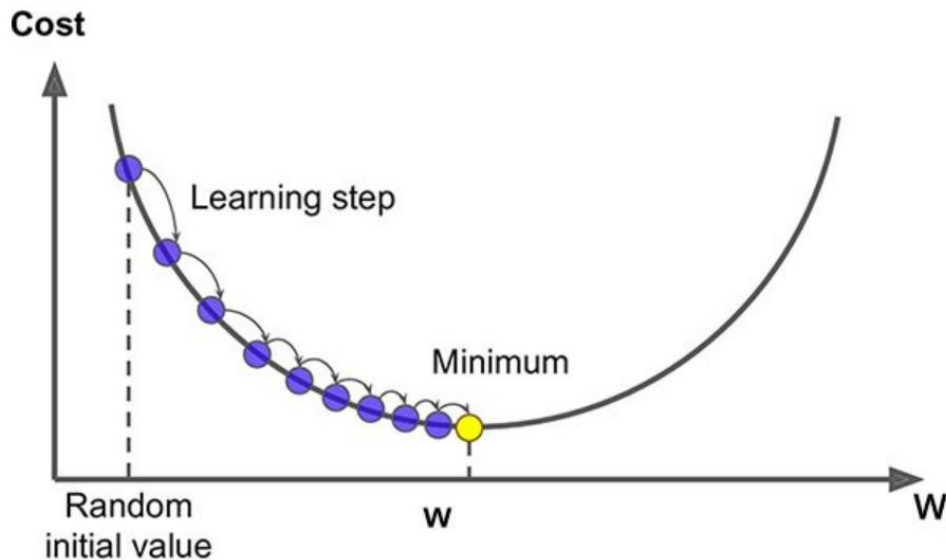
$$g = \nabla w \sum_i C(f(x^{(i)}; w), y^{(i)}) \quad (3-3)$$

Με

$$w = w - \epsilon \times w \quad (3-4)$$

Όπου g (gradient) είναι η παράγωγος, $x^{(i)}, y^{(i)}$ τα δεδομένα εκπαίδευσης, $f(x^{(i)}; w)$ είναι το ΤΝΔ με w (weights) τις παραμέτρους/βάρη ως προς τις οποίες υπολογίζεται η παράγωγος της συνάρτησης κόστους C και ϵ ο ρυθμός μάθησης του αλγορίθμου.

Ο αλγόριθμος κατάβασης κλίσης παριστάνεται σχηματικά στην Εικόνα 3–7.



Εικόνα 3–7: Σχηματική απεικόνιση του τρόπου λειτουργίας του αλγορίθμου κατάβασης κλίσης, όπου στον καθετο άξονα αναπαριστώνται οι τιμές της συνάρτησης κόστους και στον οριζόντιο οι τιμές της παραμέτρου ως προς την οποία παραγωγίζεται η συνάρτηση κόστους (πηγή Gradient descent algorithm and its three types | Clairvoyant Blog (clairvoyantsoft.com))

3.10.1.2 Στοχαστική κατάβαση κλίσης (Stochastic gradient descent)

Πρακτικά ο αλγόριθμος κατάβασης κλίσης υπολογίζει τις παραγώγους της συνάρτησης κόστους όπως αυτή προκύπτει σε κάθε βήμα για κάθε στοιχείο εκπαίδευσης. Συνέπεια της μεθόδου είναι ο τεράστιος υπολογιστικός φόρτος όταν το σύνολο δεδομένων εκπαίδευσης είναι μεγάλο.

Λύση στο πρόβλημα αυτό δίνει ο αλγόριθμος στοχαστικής κατάβασης κλίσης (SGD algorithm) με την τυχαία επιλογή ενός μόνο στοιχείου εκπαίδευσης για τον υπολογισμό των παραγώγων της συνάρτησης κόστους. Παρόλη την απλότητα της μεθόδου, έχει αποδειχθεί αρκετά αποτελεσματική σε αρκετά προβλήματα μηχανικής μάθησης. Εξαιτίας όμως της επιλογής ενός μόνο στοιχείου από το σύνολο των δεδομένων εκπαίδευσης, ενδέχεται οι μεταβολές της συνάρτησης κόστους να μην είναι σταθερές και ο αλγόριθμος να μη συγκλίνει στο βέλτιστο αποτέλεσμα.. Για αυτόν τον λόγο, συχνά επιλέγεται μια υποομάδα των δεδομένων (minibatch-SGD algorithm).

Ακόμη ένα πρόβλημα του SGD αλγόριθμου είναι η διατήρηση του ίδιου ρυθμού μάθησης για όλες τις αναβαθμίσεις των παραμέτρων χωρίς να αλλάζει κατά τη διάρκεια της εκπαίδευσης. Έτσι υπάρχει η πιθανότητα ο αλγόριθμος να παγιδευτεί σε περιοχές μικρής κλίσης όταν το βήμα της εκπαίδευσης είναι επίσης μικρό ή να προσπεράσει περιοχές με μεγάλη κλίση όταν το βήμα της εκπαίδευσης είναι μεγάλο. Αυτά τα προβλήματα καλείται να επιλύσει ο αλγόριθμος Adam.

Η μαθηματική περιγραφή του SGD αλγόριθμου είναι:

$$g = \frac{1}{m} \nabla_w \sum_i C(f(x^{(i)}; w), y^{(i)}) \quad (3-5)$$

Με

$$w = w - \epsilon \times g \quad (3-6)$$

Όπου g (gradient) είναι η παράγωγος, $x^{(i)}, y^{(i)}$ τα δεδομένα εκπαίδευσης, $f(x^{(i)}; w)$ είναι το ΤΝΔ με w τις παραμέτρους ως προς τις οποίες υπολογίζεται η παράγωγος της συνάρτησης κόστους C , m η υποομάδα των δεδομένων και ϵ ο ρυθμός μάθησης του αλγόριθμου.

3.10.1.3 Αλγόριθμος Adam

Ο αλγόριθμος Adam (Adaptive moments) είναι ένας αρκετά δημοφιλής αλγόριθμος για την εκπαίδευση ΝΔ, ο οποίος είναι μια προέκταση του SGD-αλγόριθμου αλλά συνδυάζει και διάφορους αλγόριθμους βελτιστοποίησης. Έχει αποδειχτεί ότι δίνει ικανοποιητικά αποτελέσματα σε αρκετές περιπτώσεις προβλημάτων.

Η μέθοδος υπολογίζει για κάθε παράμετρο έναν διαφορετικό ρυθμό εκπαίδευσης, ο οποίος προσαρμόζεται σε κάθε βήμα, όσο εξελίσσεται η εκπαίδευση, βάσει των ροπών 1^{ης} και 2^{ης} τάξης των παραγώγων, δηλαδή των μέσων όρων και των διακυμάνσεων των παραγώγων αντίστοιχα.

Κατα βάση ο αλγόριθμος Adam επιταχύνει τη διαδικασία εύρεσης των κατάλληλων παραμέτρων που ελχιστοποιούν το σφάλμα, δημιουργώντας έναν εκθετικά εξασθενημένο μέσο όρο των προηγούμενων παραγώγων και των διακυμάνσεων των παραγώγων.

Μαθηματικά ο αλγόριθμος περιγράφεται

$$g = \frac{1}{m'} \nabla_w \sum_i C(f(x^{(i)}; w), y^{(i)}) \quad (3-7)$$

Με

$$m = \beta_1 m + (1 - \beta_1)g, \quad (3-8)$$

$$s = \beta_2 s + (1 - \beta_2)g^T g, \quad (3-9)$$

$$\hat{m} = \frac{m}{1 - \beta_1^t}, \quad (3-10)$$

$$\hat{s} = \frac{s}{1 - \beta_2^t} \quad (3-11)$$

$$w = w - \epsilon \times \frac{\hat{m}}{\sqrt{\hat{s} + \epsilon}} \quad (3-12)$$

Όπου g (gradient) είναι η παράγωγος, $x^{(i)}, y^{(i)}$ τα δεδομένα εκπαίδευσης, $f(x^{(i)}; w)$ είναι το ΤΝΔ με w τις παραμέτρους ως προς τις οποίες υπολογίζεται η παράγωγος της συνάρτησης κόστους C , m η υποομάδα των δεδομένων, ϵ ο ρυθμός μάθησης του αλγορίθμου, t ο αριθμός της επανάληψης ($t=1,2,3,\dots$) και β_1, β_2 είναι σταθερές με συνηθέστερες τιμές $\beta_1 = 0.9, \beta_2 = 0.999$ και $\epsilon = 10^{-8}$.

3.10.2 Μέθοδοι βελτιστοποίησης δευτέρας τάξης

Παράλληλα με τη χρήση των παραγώγων για την εύρεση των παραμέτρων που δίνουν το μικρότερο σφάλμα, οι δευτέρας τάξης μέθοδοι χρησιμοποιούν και τον Εσσιανό πίνακα, ή παραλλαγές αυτού, για την εύρεση της καμπυλότητας της συνάρτησης κόστους. Σε γενικές γραμμές οι δευτέρας τάξης μέθοδοι βελτιστοποίησης, δεν έχουν ακόμα αποδειχθεί το ίδιο αποτελεσματικοί, όσο και οι πρώτης τάξης. Ο κύριος λόγος είναι ο μεγάλος φόρτος υπολογισμού τους. Μία από αυτές τις μεθόδους είναι ο αλγόριθμος L-BFGS.

3.10.2.1 Αλγόριθμος L-BFGS

Ο BFGS αλγόριθμος αποτελεί έναν από τους αλγορίθμους βελτιστοποίησης που έχουν δημιουργηθεί έτσι ώστε να εκμεταλλεύονται τα πλεονεκτήματα της μέθοδου του Νεύτωνα (Newton-Raphson method) χωρίς την υπολογιστική πολυπλοκότητα που απαιτείται για τη δημιουργία του Εσσιανού πίνακα.

Με την μέθοδο του Νεύτωνα υπολογίζεται ο Εσσιανός πίνακας της συνάρτησης σφάλματος, που έχει σαν στοιχεία του τις δευτέρες μερικές παραγώγους της συνάρτησης, και απαιτείται τόσο ο υπολόγισμός

του πίνακα όσο και η αντιστροφή του για κάθε επανάληψη, ενέργειες που είναι χρονοβόρες και υπολογιστικά πολύπλοκες.

Με χρήση του BFGS αλγορίθμου δεν υπολογίζεται σε κάθε επανάληψη ο Εσσιανός πίνακας της συνάρτησης κόστους, αλλά επιχειρείται η προσέγγιση του, αποφεύγοντας τον αναλυτικό υπολογισμό όλων των δευτέρων παραγώγων της συνάρτησης κόστους. Δημιουργεί επομένως εσσιανού πίνακες με στοιχεία τόσα όσες και οι παράμετροι του ΝΔ

Ωστόσο, μειονέκτημα του BFGS αλγορίθμου είναι η ανάγκη αποθήκευσης του πίνακα της προηγούμενης επανάληψης για τον υπολογισμό του επόμενου. Γίνεται σαφές ότι για την εκπαίδευση ενός ΝΔ όπου υπάρχει μεγάλο πλήθος από παραμέτρους (βάρη και κατώφλια) είναι αδύνατη η αποθήκευση του πίνακα μεταξύ των επαναλήψεων. Λύση σε αυτό το πρόβλημα δίνει ο L-BFGS αλγόριθμος, ο οποίος προσομοιάζει τον BFGS χρησιμοποιώντας όμως περιορισμένη υπολογιστική μνήμη (Limited-memory). Ειδικότερα, ο L-BFGS αποθηκεύει περιορισμένο αριθμό διανυσμάτων, τα οποία είναι αντιπροσωπευτικά της προσέγγισης του Εσσιανού και καθίσταται κατάλληλος για προβλήματα βελτιστοποίησης με τεράστιο πλήθος μεταβλητών

3.10.3 Οπισθοδιάδοση (backpropagation)

Ο αλγόριθμος οπισθοδιάδοσης σφάλματος (back-propagation) είναι ο πλέον διαδεδομένος αλγόριθμος εκπαίδευσης για τα ΤΝΔ πρόσθιας διάδοσης. Η εφαρμογή στην οποία βασίζεται η λειτουργία του αλγορίθμου είναι ο κανόνας της αλυσίδας. Ειδικότερα, το σφάλμα που προκύπτει για κάθε εποχή εκπαίδευσης του δικτύου αξιοποιείται από τις επόμενες εποχές, προσαρμόζοντας κατάλληλα νέες τιμές στις παραμέτρους του δικτύου (βάρη και κατώφλια). Σε κάθε εποχή εκπαίδευσης, το σφάλμα υπολογίζεται από το επίπεδο εξόδου του δικτύου. Για τον λόγο αυτό, χρησιμοποιούνται υποχρεωτικά παραγωγίσιμες συναρτήσεις ενεργοποίησης των νευρώνων.

Αρχικά, το δίκτυο τροφοδοτείται με διάνυσμα εισόδου $\bar{X} = (x^{(1)}, x^{(2)}, \dots, x^{(i)})$ βάσει του οποίου υπολογίζεται το διάνυσμα εξόδου \hat{Y} με στοιχεία $y^{(i)} = f(x^{(i)}; w)$, όπου, $x^{(i)}$ τα δεδομένα εισόδου της εκπαίδευσης, $f(x^{(i)}; w)$ είναι το ΤΝΔ με w τις παραμέτρους του. Στα δεδομένα εκπαίδευσης περιλαμβάνονται και οι επιθυμητές εξόδοι, $\bar{y}^{(i)}$, κάθε εισόδου. Συνεπώς, για κάθε νευρώνα i του επιπέδου εξόδου υπολογίζεται το αντίστοιχο σφάλμα, όπως φαίνεται στη Σχέση (3-13):

$$e_i = |y_i - \bar{y}_i| \quad (3-13)$$

Έπειτα, δεδομένης της ενεργοποίησης z_{ij} και της συνάρτησης ενεργοποίησης $\phi(z)$, υπολογίζονται τα σταθμισμένα σφάλματα d_{ij} για κάθε μία από τις εισόδους j του νευρώνα i στο επίπεδο εξόδου, σύμφωνα με τη Σχέση

$$d_{ij} = e_i * \frac{d\phi}{dz}(z_{ij}) \quad (3-14)$$

Έτσι, το συνολικό σφάλμα e_i αποδίδεται σε κάθε μία από τις εισόδους του νευρώνα του επιπέδου j . Προκειμένου να εκτελεστεί η οπισθοδιάδοση του σφάλματος στο προηγούμενο επίπεδο του δικτύου $j-1$, υπολογίζεται το άθροισμα των σταθμισμένων σφαλμάτων των k νευρώνων του j επιπέδου που οφείλονται σε αυτό, πολλαπλασιασμένα με το βάρος κάθε σύναψης $w_{i(j-1),kj}$. Η ενέργειες αυτές περιγράφονται μαθηματικά με την παρακάτω σχέση:

$$d_{i(j-1)} = \frac{d\phi}{du}(u_{i(j-1)}) * \sum_k d_{kj} * w_{i(j-1),kj} \quad (3-15)$$

Εφαρμόζοντας επαναληπτικά την Σχέση (3-15) το σφάλμα διαδίδεται από το επίπεδο της εξόδου σε όλα τα προηγούμενα, με δυνατότητα απόδοσης διαφορετικού ποσοστού ευθύνης σε κάθε βάρος του ΝΔ, ανάλογα με το πόσο το καθένα έχει συνεισφέρει στο σφάλμα. Τέλος, τα βάρη του δικτύου τροποποιούνται βάσει των σταθμισμένων σφαλμάτων d_{ij} του κάθε νευρώνα. Για την σύναψη που συνδέει τον νευρώνα i του επιπέδου j με τον νευρώνα k του επιπέδου $j-1$ με βάρος $w_{ij,k(j-1)}$, η μεταβολή του βάρους περιγράφεται από την Σχέση (3-16):

$$\Delta w_{ij,k(j-1)} = -\epsilon * d_{ij} * y_{k(j-1)} \quad (3-16)$$

Ο όρος $y_{k(j-1)}$ είναι η τιμή εξόδου του νευρώνα k στο επίπεδο $j-1$. Ο όρος ϵ είναι ο ρυθμός εκπαίδευσης (learning rate) του δικτύου και μπορεί να είναι σταθερός, ή να μεταβάλλεται ως φθίνουσα συνάρτηση του αριθμού εποχών εκπαίδευσης που έχουν ολοκληρωθεί, έτσι ώστε να πραγματοποιείται μεγάλη μεταβολή των παραμέτρων του δικτύου στις αρχικές εποχές (αρα γρηγορότερη προσαρμογή στα δεδομένα εκπαίδευσης) και μικρότερη στις τελευταίες ώστε να επιτευχθεί η τελική προσαρμογή (fine-tuning) στα δεδομένα.

3.10.4 Εκπαίδευση νευρωνικών δικτύων **Neural network training**

Όπως έχει προαναφερθεί η εκπαίδευση (training) ενός ΤΝΔ είναι η διαδικασία προσαρμογής των παραμέτρων του (βάρη και πολώσεις), έτσι ώστε το αποτέλεσμα, δηλαδή η έξοδος του δικτύου να συγκλίνει όσο το δυνατόν περισσότερο με την επιθυμητή.

Προκειμένου η έξοδος, ή αλλιώς πρόβλεψη, ενός ΝΔ να είναι ακριβής, το δίκτυο πρέπει να εκπαιδευτεί. Πιο συγκεκριμένα πρέπει οι παράμετροι του δικτύου (βάρη και πολώσεις) να προσαρμοστούν. Υπάρχουν διάφοροι εναλλακτικοί τρόποι με τους οποίους μπορεί να επιτευχθεί αυτό με τα βασικά βήματα να παραμένουν ίδια για κάθε εναλλακτική. Καταρχήν τα διαθέσιμα δεδομένα χωρίζονται σε δεδομένα εκπαίδευσης (training data) και σε δεδομένα εξέτασης (testing data). Ακολουθεί αρχικοποίηση των παραμέτρων του δικτύου με τυχαίες τιμές και στη συνέχεια το δίκτυο τροφοδοτείται με τα δεδομένα εκπαίδευσης, τα οποία αποτελούνται από επιμέρους δείγματα δεδομένων. Για κάθε ένα δείγμα το δίκτυο υπολογίζει την έξοδο/πρόβλεψη και κατόπιν, συγκρίνοντας την με την επιθυμητή έξοδο του δείγματος, υπολογίζεται το σφάλμα βάσει της επιλεχθείσας συνάρτησης κόστους. Πρακτικά το σφάλμα είναι μία τιμή του πόσο η πρόβλεψη του ΝΔ αποκλίνει από την επιθυμητή έξοδο του εκάστοτε δείγματος και χρησιμοποιείται για την αναβάθμιση και προσαρμογή των παραμέτρων του δικτύου. Κατά τη διάρκεια της εκπαίδευσης, το δίκτυο χρησιμοποιεί επαναληπτικά το σύνολο των δεδομένων εκπαίδευσης. Κάθε μία επανάληψη ονομάζεται μία εποχή (epoch) εκπαίδευσης. Στο τέλος κάθε εποχής υπολογίζεται το συνολικό σφάλμα του δικτύου. Η διαδικασία της εκπαίδευσης τερματίζει όταν το συνολικό σφάλμα εκπαίδευσης λαμβάνει μια συγκεκριμένη τιμή ή όταν το δίκτυο αποκτήσει μια συγκεκριμένη τιμή ακρίβειας ή με την ολοκλήρωση προκαθορισμένου αριθμού εποχών εκπαίδευσης ή με το πέρας ενός προκαθορισμένου χρονικού ορίου.

3.10.5 Βαθιά Νευρωνικά Δίκτυα -**Deep neural networks**

Ένα ΤΝΔ με ένα μόνο επίπεδο μπορεί να χρησιμοποιηθεί για την μοντελοποίηση μίας γραμμικής δομής, επομένως απαιτούνται ΤΝΔ με παραπάνω επίπεδα για τη μοντελοποίηση μη γραμμικών προβλημάτων. Τα ΤΝΔ με παραπάνω από ένα κρυφά επίπεδα, καλούνται βαθιά νευρωνικά δίκτυα.

3.10.6 TensorFlow

Το TensorFlow είναι μια βιβλιοθήκη νευρωνικών δικτύων ανοιχτού κώδικα (open-source) της Google, που αναπτύχθηκε από την ομάδα του Google Brain για πολλαπλές χρήσεις. Ουσιαστικά το TensorFlow καταργεί την ανάγκη δημιουργίας ενός νευρικού δικτύου από την αρχή και δίνει τη δυνατότητα στον

χρήστη να εκπαιδεύσει το Tensorflow με τα δικά του δεδομένα, χρησιμοποιώντας τα αποτελέσματα που επιθυμεί.

3.10.7 Προβλήματα κατά την εκπαίδευση ενός νευρωνικού δικτύου

Στην Ενότητα 3.10.4 περιγράφηκε ο τρόπος με τον οποίο εκπαιδεύεται ένα νευρωνικό δίκτυο, συχνά είναι όμως και τα προβλήματα που προκύπτουν κατά τη διάρκεια της εκπαίδευσης. Σε αυτήν την ενότητα θα γίνει αναφορά σε αυτά τα προβλήματα.

3.10.7.1 Υπερπροσαρμογή

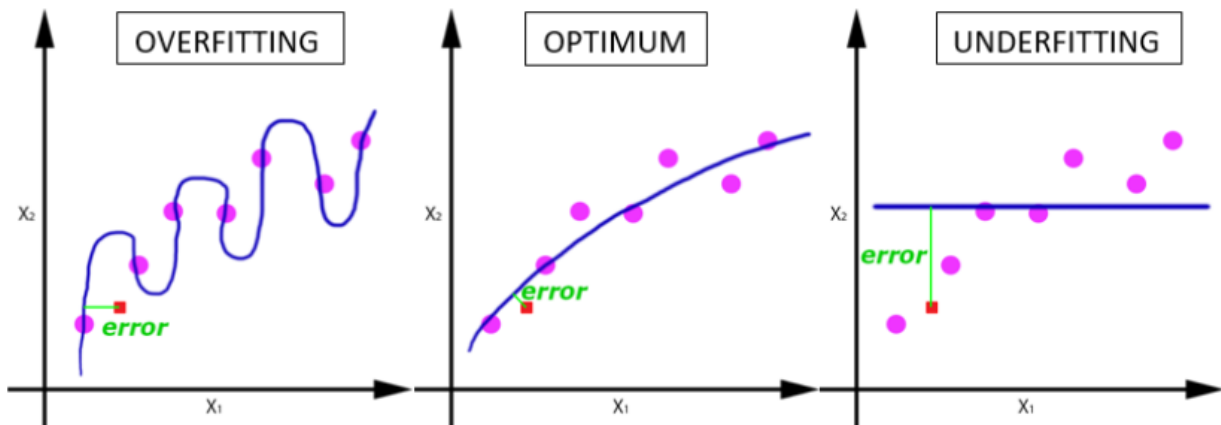
Στην περίπτωση που ένα στοχαστικό μοντέλο εκπαιδεύεται έτσι ώστε να προσαρμόζεται στα εκάστοτε δεδομένα με αρκετά μεγάλη ακρίβεια και αδυνατεί να γενικευτεί ικανοποιητικά σε καινούρια και άγνωστα δεδομένα, υπάρχει το πρόβλημα της υπερπροσαρμογής (overfitting). Ειδικότερα, ο θόρυβος που υπάρχει σε ένα υποσύνολο των δεδομένων και αποτελεί βασικό χαρακτηριστικό του συνόλου τους, εξαλείφεται παντελώς. Συνέπεια, το μοντέλο αποκτά εξειδίκευση στο συγκεκριμένο υποσύνολο, παρεκκλίνοντας από τον πρωτεύοντα στόχο του μοντέλου, ο οποίος είναι ο εντοπισμός των συσχετίσεων που διέπει το σύνολο των δεδομένων. Η υπερεκπαίδευση (overtraining) του δικτύου είναι η αιτία της υπερπροσαρμογής του και συγκεκριμένα όταν τα κριτήρια επιλογής (selection criteria) ενός μοντέλου (συνήθως επιλέγεται αυτό που ελαχιστοποιεί το σφάλμα πάνω στα δεδομένα εκπαίδευσης) διαφέρουν από τα κριτήρια καταλληλότητάς (fitness criteria) του (η ικανότητά του να γενικεύει σε άγνωστα δεδομένα τα οποία δεν είχε συναντήσει κατά την εκπαίδευση). Επισημαίνεται ότι στην υπερεκπαίδευση κριτήριο επιλογής του μοντέλου αποτελεί η μεγάλη ελαχιστοποίηση του σφάλματος πάνω στα δεδομένα εκπαίδευσης, εις βάρος του κριτηρίου καταλληλότητάς του μοντέλου, δηλαδή της ικανότητάς του να γενικεύεται σε δεδομένα για τα οποία δεν έχει εκπαιδευτεί.

3.10.7.2 Υποπροσαρμογή

Αντίθετο πρόβλημα της υπερπροσαρμογής αποτελεί η υποπροσαρμογή (underfitting), η αδυναμία δηλαδή του μοντέλου να αναγνωρίσει και να αφομοιώσει ικανοποιητικά τα πρότυπα που διέπουν τη δομή των δεδομένων εκπαίδευσης. Πιθανή αιτία της υποπροσαρμογής ενός μοντέλου MM είναι ο τερματισμός της εκπαίδευσης προτού πραγματοποιηθεί ο κατάλληλος αριθμός εποχών εκπαίδευσης, με συνέπεια οι παράμετροι του δικτύου να μην έχουν προλάβει να προσαρμοστούν κατάλληλα. Άλλη πιθανή αιτία είναι το ανεπαρκές πλήθος δειγμάτων για τα δεδομένα εκπαίδευσης, με συνέπεια το δίκτυο

να μην προσαρμόζεται κατάλληλα στην κατανομή τους. Επίσης, η λανθασμένη επιλογή αρχιτεκτονικής και υπερπαραμέτρων δικτύου είναι μια ακόμη αιτία, με την κάθε αρχιτεκτονική να μπορεί να αντιμετωπίζει ικανοποιητικά συγκεκριμένες μόνο κατηγορίες προβλημάτων MM, ενώ οι υπερπαραμέτροι του δικτύου συντελούν στην πολυπλοκότητα, την ευρωστία και την χωρητικότητα του δικτύου.

Στην Εικόνα 3–8 απεικονίζονται οι καμπύλες στην περίπτωση της υπερπροσαρμογής, υποπροσαρμογής σε σχέση με την ιδανική λύση καθώς και το πόσο μπορεί να αποκλίνει μια τιμή σε κάθε περίπτωση.



Εικόνα 3–8: οι καμπύλες στην περίπτωση της υπερπροσαρμογής, υποπροσαρμογής σε σχέση με την ιδανική λύση (πηγή: www.TowardsDataScience.com).

3.10.8 Βασική ορολογία των ΝΔ

Παρακάτω γίνεται αναφορά σε βασικούς όρους που χρησιμοποιούνται στον τομέα των νευρωνικών δικτύων.

3.10.8.1 Εποχή(Epoch)

Το ΝΔ «μαθαίνει» τα μοτίβα των δεδομένων εκπαίδευσης, πραγματοποιώντας διάφορους υπολογισμούς σε αυτά. Με το πρώτο «πέρασμα» των δεδομένων το TN δεν είναι σε θέση να εκπαιδευτεί πλήρως και ορθά. Απαιτούνται αρκετές επαναλήψεις με είσοδο τα ίδια δεδομένα, ώστε μαθαίνοντας από τα λάθη του το δίκτυο να προσαρμοστεί καλύτερα στα δεδομένα.

Κάθε προσπάθεια εκπαίδευσης του δικτύου από τα ίδια δεδομένα καλείται εποχή. Συγκεκριμένα ένα ολοκληρωμένο πρόσθιο πέρασμα του συνόλου των δεδομένων στο δίκτυο και ένα όπισθεν πέρασμά τους, αποτελούν μία εποχή.

Επομένως μία εποχή αναφέρεται σε έναν κύκλο εκπαίδευσης του δικτύου όπου ολόκληρο το σύνολο των δεδομένων εισέρχεται μία φορά στο δίκτυο. Συνήθως η εκπαίδευση ενός ΝΔ απαιτεί πάνω από μία εποχές για να επιτευχθεί. Πάραυτα, αύξηση του αριθμού των εποχών δεν συνεπάγεται καλύτερα αποτελέσματα του δικτύου. Επομένως, εμπειρικά διαλέγεται ο αριθμός των εποχών για τον οποίο τα αποτελέσματα παραμένουν ίδια έπειτα από μερικούς κύκλους.

3.10.8.2 Μέγεθος ομάδας (Batch size)

Το σύνολο των δεδομένων εκπαίδευσης χωρίζεται σε ομάδες με συγκεκριμένο αριθμό δεδομένων. Καθεμία από αυτές τις ομάδες εισέρχονται στο ΝΔ όπου πραγματοποιείται ένα πρόσθιο και ένα όπισθεν πέρασμά τους. Πρόσθιο και όπισθεν πέρασμα θεωρείται μία επανάληψη, ενώ με το πέρασμα όλων των ομάδων έχει ολοκληρωθεί μία εποχή. Όσο περισσότερες οι ομάδες που χωρίζεται το σύνολο των δεδομένων, τόσο μεγαλύτερη η μνήμη που απαιτείται.

3.10.8.3 Επανάληψη (Iteration)

Για κάθε ολοκληρωμένη εποχή, πραγματοποιούνται αρκετές επαναλήψεις (iterations). Ο αριθμός των επαναλήψεων είναι ο αριθμός των ομάδων (batches) ή βημάτων όπου όταν εκτελεστούν αποτελούν μία εποχή. Στην ορολογία της ΜΜ μία επανάληψη ορίζεται ως το πρόσθιο και όπισθεν πέρασμα μίας ομάδας δεδομένων στο δίκτυο. Επιπλέον, με κάθε επανάληψη οι παράμετροι του ΝΔ αναβαθμίζονται. Η εκπαίδευση ενός ΝΔ απαιτεί αρκετές επαναλήψεις.

Ένα τυπικό παράδειγμα μίας επανάληψης κατά το στάδιο εκπαίδευσης ενός ΝΔ αποτελείται από τα παρακάτω βήματα:

- Επεξεργασία της ομάδας του συνόλου των δεδομένων εκπαίδευσης από τον αλγόριθμο του δικτύου
- Υπολογισμός της συνάρτησης κόστους
- Οπισθοδιάδοση του σφάλματος και προσαρμογή των παραμέτρων του δικτύου

Για παράδειγμα: Αν υπάρχουν 1000 δεδομένα εκπαίδευσης και ανά 500 χωριστούν σε ομάδες, απαιτούνται 2 επαναλήψεις για να ολοκληρωθεί μία εποχή

3.10.8.4 Αριθμός υπολογισμού συναρτήσεων

Ο αριθμός υπολογισμού συναρτήσεων του ΝΔ είναι από τα μεγαλύτερης σημασίας κριτήρια για την επίδοση ενός αλγορίθμου βελτιστοποίησης. Παρότι ο αριθμός των συνολικών επαναλήψεων για διαφορετικούς αλγορίθμους βελτιστοποίησης σε κάποιο ΝΔ μπορεί να είναι ίδιος, δεν αναπαριστά το πραγματικό υπολογιστικό κόστος του αλγορίθμου, δηλαδή πόση «προσπάθεια» απαιτείται από τον αλγόριθμο. Συνηθέστερα, για το εκάστοτε πρόβλημα επιλέγονται εκείνοι οι αλγόριθμοι βελτιστοποίησης με το μικρότερο υπολογιστικό κόστος.

3.10.8.5 Συνάρτηση σφάλματος (Loss function)

Η συνάρτηση σφάλματος είναι η μέθοδος που περιγράφει «πόσο σωστά ο αλγόριθμος περιγράφει τα δεδομένα που του δόθηκαν». Αν οι προβλέψεις που δίνει ως απάντηση ο αλγόριθμος είναι τελείως διαφορετικές από τις πραγματικές, τότε η συνάρτηση σφάλματος θα έχει ως έξοδο έναν μεγάλο αριθμό. Στην αντίθετη περίπτωση, η τιμή της θα είναι πολύ κοντά στο 0. Επομένως ο δείκτης για το πόσο ορθά ο αλγόριθμος ρυθμίζεται έτσι ώστε να αναπαριστά με μεγαλύτερη ακρίβεια τα δεδομένα, είναι η τιμή της συνάρτησης σφάλματος. Το σφάλμα (error) είναι η μετρική που δείχνει πόσο η τιμή μιας πρόβλεψης διαφέρει από την πραγματική.

Τα είδη της συνάρτησης σφάλματος είναι τα εξής:

- Παλινδρομική συνάρτηση σφάλματος (*Regression Loss Function*)

Τα μοντέλα παλινδρόμησης χρησιμοποιούνται για την πρόβλεψη μιας συνεχούς τιμής όπως για παράδειγμα την τιμή ενός σπιτιού ή τις διαστάσεις ενός δωματίου. Αυτό είναι και το είδος σφάλματος που χρησιμοποιείται στην περίπτωση των PINNs.

- Συνάρτηση σφάλματος δυαδικής ταξινόμησης (*Binary Classification Loss Functions*)

Η δυαδική ταξινόμηση πρόκειται για έναν αλγόριθμο πρόβλεψης όπου η τιμή εξόδου θα είναι μία εκ των δύο τιμών είτε 0 είτε 1. Η έξοδος μιας δυαδικής ταξινόμησης είναι ένα σκορ πρόβλεψης. Η ταξινόμηση γίνεται αφού έχει οριστεί ένα κατώφλι, συνήθως με την τιμή 0.5. Αν το σκορ πρόβλεψης είναι μεγαλύτερο από το κατώφλι τότε η έξοδος λαμβάνει την τιμή 1, αλλιώς με 0.

- Πολυβάθμια ταξινόμηση συνάρτησης σφάλματος (*Multi-class Classification Loss Functions*).

Πρόκειται για την προέκταση του προβλήματος της δυαδικής ταξινόμησης και αφορά μοντέλα πρόβλεψης με περισσότερες μεταβλητές.

3.10.8.6 Συνάρτηση κόστους (Cost function or Objective Function)

Η συνάρτηση σφάλματος υπολογίζεται για ένα μεμονωμένο ζεύγος δεδομένων εκπαίδευσης. Αν τα ζεύγη εκπαίδευσης είναι m τότε ο μέσος όρος της συνάρτησης σφάλματος από όλα τα ζεύγη των δεδομένων εκπαίδευσης ονομάζεται συνάρτηση κόστους (cost function).

$$\text{Cost function (J)} = \frac{1}{m} (\sum_{i=1}^m \text{LossFunction}_i)$$

Η γραφική παράσταση της συνάρτησης κόστους ως προς τις παραμέτρους του ΝΔ W, b είναι μια παραβολή με μία ελάχιστη τιμή που καλείται «τοπικό βέλτιστο».

Η συνάρτηση κόστους είναι ο δείκτης επίδοσης του μοντέλου μηχανικής μάθησης που χρησιμοποιείται για δοθέν δείγμα δεδομένων εκπαίδευσης. Ποσοτικοποιεί με έναν πραγματικό αριθμό το σφάλμα μεταξύ των προβλέψεων που κάνει ο αλγόριθμος και των αναμενόμενων τιμών από τα δεδομένα. Ανάλογα με το πρόβλημα, η συνάρτηση κόστους μπορεί να σχεδιαστεί έτσι ώστε ο σκοπός της να είναι να βρεθεί είτε η ελάχιστη, είτε η μέγιστη τιμή της.

Στην περίπτωση ελαχιστοποίησης της συνάρτησης κόστους, στόχος είναι να βρεθούν οι τιμές των παραμέτρων του μοντέλου για τις οποίες η συνάρτηση κόστους έχει τη μικρότερη δυνατή τιμή. Αντίθετα στην περίπτωση μεγιστοποίησης της συνάρτησης, η τιμή που επιστρέφει δίνει ονομάζεται βραβείο (reward) και στόχος είναι η μέγιστη δυνατή τιμή της.

Σε προβλήματα επιβλεπόμενης μάθησης η συνάρτηση κόστους είναι συνήθως είτε το μέσο τετραγωνικό σφάλμα ή κόστος διασταυρούμενης εντροπίας. Για τη μη επιβλεπόμενη μάθηση, δεδομένου ότι δεν υπάρχει γνωστή εκ των προτέρων έξοδος χρησιμοποιείται κάποιος δείκτης στατιστικής διασποράς των εξόδων του δικτύου.

3.10.8.7 Μέσο τετραγωνικό σφάλμα (Mean Squared Error)

Οι αλγόριθμοι των PINNs χρησιμοποιούν ως συνάρτηση κόστους το μέσο τετραγωνικό σφάλμα. Πρόκειται για το μέσο όρο των διαφορών μεταξύ των προβλέψεων του δικτύου και των αναμενόμενων τιμών υψωμένων στο τετράγωνο.

Στην συνάρτηση κόστους όπου χρησιμοποιείται το μέσο απόλυτο σφάλμα, κάθε μερική τιμή σφάλματος ισούται με την απόσταση μεταξύ των σημείων στο σύστημα συντεταγμένων. Αντίστοιχα, στην περίπτωση του ΜΤΣ, κάθε μερική τιμή σφάλματος ισούται με το τετράγωνο της απόστασης μεταξύ των σημείων. Όλες οι τετραγωνικές αποστάσεις προστίθενται και λαμβάνεται η μέση τιμή τους. Αυτή η μορφή της συνάρτησης κόστους λειτουργεί πολύ καλύτερα σε αλγορίθμους που χρησιμοποιούν τον

αλγόριθμο Gradient Descent, καθότι έχει μαθηματικές ιδιότητες που καθιστούν ευκολότερη την παραγωγή της. Η μορφή του ΜΤΣ είναι:

$$MTΣ = \frac{1}{n} \sum_{i=1}^n (y_i - \widehat{y}_i)^2 \quad (3-17)$$

Όπου i ο δείγματος (index of sample), \widehat{y} η πρόβλεψη, y η αναμενόμενη τιμή, n πλήθος δειγμάτων των δεδομένων εκπαίδευσης

ΚΕΦΑΛΑΙΟ 4. Νευρωνικά δίκτυα που υπακούουν στους νόμους της φυσικής -Physics Informed Neural Networks (PINNs)

4.1 Εισαγωγή στα PINNs

Η εκρηκτική αύξηση των διαθέσιμων δεδομένων, η ανάπτυξη των υπολογιστικών μεθόδων, παράλληλα με τις πρόσφατες εξελίξεις της μηχανικής μάθησης και της ανάλυσης δεδομένων έχουν επιφέρει πρωτοφανείς αλλαγές σε ποικίλα επιστημονικά πεδία της σύγχρονης εποχής όπως αναγνώριση εικόνων, γνωσιακή φυσική και γονιοδιωματική. Στις περιπτώσεις όμως επιστημονικών πεδίων, που η απόκτηση των απαραίτητων δεδομένων είναι μια διαδικασία ιδιαίτερα κοστοβόρα, η τεχνολογία αυτή δεν μπορεί να αξιοποιηθεί. Τα εργαλεία της MM που για την ώρα χρησιμοποιούνται, δηλαδή τα βαθιά, τα συνελκτικά και τα επαναληπτικά νευρωνικά δίκτυα, δεν δύνανται να συγκλίνουν επιτυχώς σε προβλήματα με ελάχιστα ή μη επαρκή δεδομένα εκπαίδευσης.

Παρόλαυτα, σε αρκετές από αυτές τις περιπτώσεις, όπου επιχειρείται η μοντελοποίηση βιολογικών και φυσικών φαινομένων με λίγα δεδομένα, υπάρχει πληθώρα θεωρητικής γνώσης στη βιβλιογραφία σχετικά με τους φυσικούς νόμους, τις εμπειρικές σχέσεις ή άλλες πληροφορίες-νόμους που διέπουν αυτά τα φαινόμενα, η οποία δεν αξιοποιούνταν από τις σύγχρονες μεθόδους MM. Η χρήση των φυσικών νόμων όπου ορίζουν την χρονικά εξαρτώμενη δυναμική ενός φαινομένου ή κάποιων άλλων εμπειρικά αποδεκτών κανόνων ή άλλης μορφής εξειδικευμένων βασικών θεωρητικών αρχών για τον κάθε τομέα επιστήμης, λειτουργούν ως ρυθμιστές των αποδεκτών λύσεων ενός ΝΔ (πχ σε ρευστοδυναμικά προβλήματα ασυμπίεστης ροής το σύστημα MM που αξιοποιεί τις βασικές αρχές που διέπουν αυτό το φαινόμενο απαλλάσσεται από τις λύσεις που δεν τηρούν την αρχή διατήρησης της μάζας). Παρέχοντας σε έναν αλγόριθμο εκπαίδευσης τέτοιου είδους πληροφορία, η οποία θα εκφραστεί μέσω μιας μη γραμμικής διαφορικής εξίσωσης προκύπτουν επιπλέον δεδομένα εκπαίδευσης και έτσι ο αλγόριθμος κατευθύνεται γρηγορότερα προς την επιθυμητή κατεύθυνση επίλυσης του προβλήματος, ενώ λειτουργεί ικανοποιητικά σε άγνωστες τιμές εισόδου.

Τα ΝΔ που εκπαιδεύονται για την επίλυση προβλημάτων επιβλεπόμενης μηχανικής μάθησης και ταυτόχρονα τηρούν τους φυσικούς νόμους του εκάστοτε προβλήματος καλούνται PINNs.

Ο στόχος των ΝΔ που υπακούουν στους νόμους της φυσικής (PINNs) είναι να επιλύσουν πραγματικά υπάρχοντα προβλήματα τριών κατηγοριών. Στην πιο συνηθισμένη κατηγορία, ορισμένοι από τους φυσικούς νόμους που διέπουν το εκάστοτε φαινόμενο είναι γνωστοί και μπορούν να περιγραφούν, ενώ

ταυτόχρονα είναι διαθέσιμα και κάποια δεδομένα με τα οποία μπορεί το ΝΔ να εκπαιδευτεί, χωρίς να είναι δεσμευτικό τα δεδομένα αυτά να είναι οι αρχικές ή συνοριακές συνθήκες του προβλήματος. Τα κενά στη γνώση όλων των φυσικών νόμων οφείλονται είτε στην πολυπλοκότητα του φαινομένου που εξετάζεται, είτε στη δυσκολία επίλυσης τους (πχ για τα φαινόμενα των γεωφυσικών αναταράξεων παρόλο που οι νόμοι και οι αρχές που τα διέπουν είναι γνωστοί, η επίλυσή τους καθίσταται δύσκολη ακόμα και με χρήση τελευταίας τεχνολογίας υπολογιστών). Επομένως, με γνωστό ένα μέρος των αρχών που διέπουν το φαινόμενο γίνεται προσπάθεια για την εύρεση των υπολοίπων, με χρήση των δεδομένων που διατίθενται.

Οι άλλες δύο κατηγορίες που εφαρμόζονται τα PINNs αποτελούν ακραίες περιπτώσεις, όπου στη μια είναι γνωστοί επακριβώς όλοι οι φυσικοί νόμοι και οι βασικές αρχές του φαινομένου, με την αρωγή των εφαρμοσμένων μαθηματικών, και επιδιώκεται η εύρεση αλγορίθμων για την ταχύτερη επίλυσή τους. Στην τελευταία περίπτωση υπάρχουν μόνο δεδομένα για το φαινόμενο και δεν είναι γνωστοί οι φυσικοί νόμοι που το περιγράφουν.

4.2 Ορισμός του προβλήματος

Όπως διαπιστώνεται η χρήση των PINNs καθιστά εφικτή τη διαχείριση μη γραμμικών προβλημάτων, χωρίς να υπάρχει η ανάγκη δημιουργίας νέων υποθέσεων, γραμμικοποίησης ή εφαρμογή αριθμητικών μεθόδων, όπως τα πεπεραμένα στοιχεία. Τα PINNs είναι κατασκευασμένα έτσι ώστε να τηρούν τις όποιες συμμετρίες, σταθερές τιμές ή αρχές διατήρησης που προκύπτουν από τους φυσικούς νόμους κάθε φαινομένου ή γεγονότος, και άρα διέπουν και τα δεδομένα που χρησιμοποιούνται. Το εκάστοτε φαινόμενο αναπαρίσταται με γενικές χρονο-εξαρτώμενες και μη γραμμικές διαφορικές εξισώσεις, και δύναται να αναπαριστά μεγάλο εύρος προβλημάτων υπολογιστικής επιστήμης. Ταυτόχρονα, εισάγει μια τεχνολογία που αναπτύσσει νέες μεθόδους αξιοποίησης δεδομένων, νέες τάξεις και προσεγγίσεις αριθμητικών επιλύσεων των ΜΔΕ.

Απώτερος σκοπός των PINNs είναι να τεθούν οι βάσεις για ένα νέο πρότυπο μοντελοποίησης και υπολογισμού που θα εμπλουτίζει τη βαθιά μάθηση με τις όποιες εξελίξεις συντελούνται στον τομέα των εφαρμοσμένων μαθηματικών. Τα PINNs χρησιμοποιούν σχετικά απλές, πρόσθιας διάδοσης αρχιτεκτονικές ΝΔ με την συνάρτηση ενεργοποίησης να είναι η συνάρτηση υπερβολικής εφαπτομένης, χωρίς πρόσθετη κανονικοποίηση.

Δύο είναι οι κατηγορίες προβλημάτων στις οποίες έχουν δοκιμαστεί τα PINNs, τα προβλήματα επίλυσης βάσει των δεδομένων (data-driven solution) και τα προβλήματα ανακάλυψης βάσει των

δεδομένων (data-driven discovery). Τα πρώτα αναφέρονται σε προβλήματα όπου χρησιμοποιούνται νευρωνικά δίκτυα που έχουν εκπαιδευτεί με ελάχιστα δεδομένα. Τα δεδομένα αυτά ικανοποιούν κάποιο γνωστό φυσικό νόμο ο οποίος περιγράφεται από ένα σύστημα ΜΔΕ. Τέτοιες περιπτώσεις απαντώνται σε πληθώρα φαινομένων φυσικής, βιολογίας και μηχανικής όπου χάρη στις εξελίξεις στον τομέα των εφαρμοσμένων μαθηματικών είναι γνωστός ο τρόπος με τον οποίο τα φαινόμενα αυτά δομούνται, αλληλοεπιδρούν και εξελίσσονται δυναμικά στον χρόνο. Επομένως η γνώση ενός φυσικού νόμου στον οποίο υπακούν τα δεδομένα ενός φαινομένου, αποτελεί επιπλέον πληροφορία για την εκπαίδευση του δικτύου και συμβάλλει στην ομαλοποίηση του δικτύου χωρίς προβλήματα υπερ-προσαρμογής, ακόμα και αν τα διαθέσιμα δεδομένα εκπαίδευσης είναι λιγοστά.

Έστω η παραμετροποιημένη, μη γραμμική ΜΔΕ με γενική μορφή (Karniadakis., et al 2017) :

$$u_t + N[u; \lambda] = 0, x \in \Omega, t \in [0, T] \quad (4-1)$$

Όπου η $u(t, x)$ είναι η λύση της εξίσωσης, $N[u; \lambda]$ είναι ο μη γραμμικός τελεστής της εξίσωσης (ως τελεστής στα μαθηματικά ορίζεται γενικά μία συνάρτηση που δρα πάνω σε κάποια άλλη συνάρτηση, μετασχηματίζοντάς την κατά ένα καθορισμένο τρόπο) με παράμετρο λ , και Ω είναι ένα υποσύνολο του R^d . Κατά αυτόν τον τρόπο, περιλαμβάνεται μεγάλο εύρος προβλημάτων του τομέα των εφαρμοσμένων μαθηματικών, όπως αρχές διατήρησης, διαδικασίες διάχυσης, συστήματα μεταγωγών, αντιδράσεων και εξισώσεις κινηματικής (conservation laws, diffusion processes, advection diffusion reaction systems, and kinetic equations).

Προκειμένου το ΤΝΔ να αξιοποιήσει την πληροφορία που προκύπτει από τη γνώση του φυσικού νόμου του προβλήματος, ο αλγόριθμος τροποποιείται και ορίζεται η συνάρτηση $f(t, x)$ η οποία θα δίνεται από το αριστερό μέλος της εξίσωσης (4-1) δηλαδή

$$f := u_t + N[u] \quad (4-2)$$

Με αυτόν τον τρόπο το ΤΝΔ δύναται να συγκλίνει στη λύση του προβλήματος αξιοποιώντας τη ΜΔΕ που ορίζει το φαινόμενο και τις αρχικές και συνοριακές συνθήκες που την περιγράφουν.

Για παράδειγμα, η μιας διάστασης εξίσωση του Burgers αντιστοιχεί στην περίπτωση όπου

$$N[u; \lambda] = \lambda_1 u u_x - \lambda_2 u_{xx} \quad \text{και} \quad \lambda = (\lambda_1, \lambda_2) \quad (4-3)$$

με τους δείκτες να δηλώνουν είτε την χρονική, είτε τη χωρική μεταβλητή ως προς την οποία γίνεται η μερική παραγωγή της u . Δεδομένων των μετρήσεων με θόρυβο του συστήματος, ενδιαφέρει η επίλυση δύο ξεχωριστών προβλημάτων. Το πρώτο πρόβλημα είναι αυτό της επαγωγής (inference) ή το πρόβλημα για το οποίο οι επιλύσεις των ΜΔΕ είναι καθοδηγούμενες από τα δεδομένα, το οποίο με δεδομένη τιμή της παραμέτρου λ αναζητείται η λύση της άγνωστης συνάρτησης του συστήματος $u(t, x)$ Το δεύτερο

πρόβλημα είναι αυτό της ταυτοποίησης, βάσει των δεδομένων που έχουν παρατηρηθεί για την ΜΔΕ, η εύρεση δηλαδή της βέλτιστης τιμής της παραμέτρου λ της εξίσωσης, με την οποία περιγράφονται πιστότερα τα δεδομένα.

4.3 Εκπαίδευση των PINNs

Τα δεδομένα εκπαίδευσης του ΤΝΔ είναι κάποιο πλήθος σημείων αρχικών και συνοριακών συνθηκών του προβλήματος, $N_u=(t_u, x_u)$, και ορισμένα σημεία του χωροχρονικού πεδίου του προβλήματος $N_f=(t_f, x_f)$ καθώς και η λύση της $u(t, x)$ σε αυτά τα σημεία. Ο τρόπος που επιλέγονται είναι τυχαίος και βάσει του προηγούμενου ορισμού, η f θα πρέπει να μηδενίζεται σε αυτά τα σημεία. Οι παράμετροι του ΤΝΔ θα βρεθούν από την ελαχιστοποίηση της συνάρτησης κόστους, η οποία αποτελείται από δύο όρους, με τον ένα να αναφέρεται στα σημεία N_u και τον άλλο στα N_f . Συνεπώς λαμβάνει τη μορφή:

$$MSE = MSE_u + MSE_f \quad (4-4)$$

Με

$$MSE_u = \frac{1}{N_u} \sum_{i=1}^{N_u} |u(t_u^i, x_u^i) - u^i|^2 \quad (4-5)$$

$$MSE_f = \frac{1}{N_f} \sum_{i=1}^{N_f} |f(t_f^i, x_f^i)|^2 \quad (4-6)$$

Όπου οι όροι t_u^i, x_u^i, u^i αποτελούν τα στοιχεία εκπαίδευσης με τον όρο $u(t_u^i, x_u^i)$ να αναφέρεται στην πρόβλεψη του ΤΝΔ ενώ ο όρος u^i δηλώνει την πραγματική τιμή της u στα συγκεκριμένα σημεία. Αντίστοιχα για τα σημεία N_f το ΤΝΔ δίνει τα αποτελέσματα $f(t_f^i, x_f^i)$, τα οποία θα πρέπει να συγκλίνουν στο μηδέν προκειμένου να εξασφαλιστεί ότι ικανοποιείται ο φυσικός νόμος που διέπει το φαινόμενο. Και στις δύο περιπτώσεις η μορφή της συνάρτησης κόστους παραμένει ίδια.

Επισημαίνεται ότι βιβλιογραφικά δεν υπάρχει κάποια επίσημη θεωρία που να αποδεικνύει ότι η παραπάνω διαδικασία καταλήγει σε ολικό ελάχιστο. Εντούτοις μέσω εμπειρικών δοκιμών άγεται το συμπέρασμα ότι τα PINNs δύνανται να δώσουν αρκετά ακριβή πρόβλεψη της συμπεριφοράς του φαινομένου εντός του χωροχρονικού πεδίου στο οποίο έχει οριστεί το εκάστοτε φαινόμενο χρησιμοποιώντας ελάχιστα δεδομένα εκπαίδευσης. Αυτό προϋποθέτει ότι τα νευρωνικά δίκτυα των PINNs θα πρέπει να έχουν επαρκώς ορισμένη αρχιτεκτονική, δηλαδή αριθμό επιπέδων (layers) και νευρώνων σε κάθε κρυμμένο επίπεδο (Karniadakis et al., 2017). Η εύρεση του ολικού ελαχίστου

επιτυγχάνεται δοκιμάζοντας τον κατάλληλο αλγόριθμο βελτιστοποίησης και την καταλληλότερη αρχιτεκτονική των δικτύων.

Για τη βέλτιστη προσέγγιση της λύσης από το PINN ερευνάται η αρχιτεκτονική του δικτύου και ο αριθμός των δεδομένων εκπαίδευσης με τα οποία τροφοδοτείται. Η ορθή επιλογή των παραπάνω χαρακτηριστικών εξαρτάται από την εμπειρία του μελετητή στην εκπαίδευση αλγορίθμων. Η γενική ιδέα είναι πώς αυξάνοντας τα δεδομένα εισόδου με τα οποία τροφοδοτείται το δίκτυο, Nu και Nf , βελτιώνονται οι προβλέψεις που δίνει.

Σε αντίθεση με τις κλασικές αριθμητικές μεθόδους η επίλυση μιας ΜΔΕ με τη μέθοδο PINNs έχει επιτευχθεί χωρίς καμία διακριτοποίηση στο χώρο και στο χρόνο, ενώ δεν χρειάζεται ειδική μεταχείριση το κάθε ξεχωριστό φαινόμενο που μελετάται όσον αφορά την αρχιτεκτονική του ΤΝΔ. Το αποτέλεσμα που προκύπτει είναι πλήρως παραγωγίσιμο σε όλο το πεδίο ορισμού της εξίσωσης. Η έννοια της παραγώγου που εισάγεται στο ΤΝΔ είναι ιδιαίτερα σημαντική πληροφορία για την ανάπτυξη λύσης στο εκάστοτε πρόβλημα. Αυτό σημαίνει ότι είναι μια πολύ καλή προσέγγιση της αναλυτικής και ακριβής λύσης. Τα πλεονεκτήματα αυτά, μαζί με την ευκολία στην υλοποίηση, καθιστούν τα PINNs ένα ιδιαίτερα ισχυρό εργαλείο.

4.4 Ταυτοποίηση παραμέτρων διαφορικής εξίσωσης (parameter identification)

Στις περιπτώσεις όπου το δίκτυο προβλέπει την εξέλιξη του φαινομένου στο πεδίο ορισμού που έχει οριστεί και η διαφορική εξίσωση που διέπει το φαινόμενο είναι γνωστή στην πλήρη μορφή της με τις παραμέτρους της γνωστές και σταθερές κατά την εκπαίδευση του δικτύου, τα προβλήματα αυτά καλούνται προβλήματα επαγωγής (inference).

Ενδέχεται όμως η λύση της ΜΔΕ να είναι γνωστή από μετρήσεις πεδίου, αναλυτικές επιλύσεις, ή προσομοιώσεις του φαινομένου μέσω αριθμητικών μεθόδων, αλλά να ερευνάται η τιμή των παραμέτρων. Αυτά καλούνται προβλήματα ταυτοποίησης (identification), και το δίκτυο εκπαιδεύεται έτσι ώστε οι προβλέψεις που δίνει να προσεγγίζουν τα δεδομένα, αλλά με τις παραμέτρους λ να αντιμετωπίζονται σαν μεταβλητές του δικτύου και άρα να μεταβάλλεται συνεχώς η τιμή τους κατά τη φάση εκπαίδευσης.

Η εκπαίδευση των προβλημάτων ταυτοποίησης μπορεί να πραγματοποιηθεί με μία εποχή πάνω σε μεγάλο αριθμό δεδομένων, με πολλές εποχές αν τα δεδομένα δεν είναι αρκετά ή με συνδυασμό των δύο αυτών μεθόδων.

4.5 Μέθοδος του λατινικού υπερκύβου - Latin Hypercube Sampling

Η μέθοδος του λατινικού υπερκύβου χρησιμοποιείται στη μέθοδο PINNs. Ουσιαστικά εφαρμόζεται σε προβλήματα στατιστικής για την παραγωγή ενός σχεδόν τυχαίου δείγματος παραμετρικών τιμών μιας πολυδιάστατης κατανομής. Η συγκεκριμένη μέθοδος δειγματοληψίας χρησιμοποιείται στην κατασκευή πειραμάτων με H/Y. Κατα βάση, η LHS διαμερίζει κάθε διάσταση, η οποία αναπαριστά μια μεταβλητή, σε N τμήματα, όπου N είναι ο αριθμός των σημείων των δειγμάτων, και λαμβάνεται ένα σημείο για κάθε τμήμα.

Κατά τη δειγματοληψία μιας συνάρτησης μεταβλητών, το εύρος κάθε μεταβλητής χωρίζεται σε ίσα πιθανά διαστήματα. Έπειτα επιλέγεται από κάθε διάστημα ένα δείγμα, έτσι ώστε να ικανοποιούνται οι απαιτήσεις του Λατινικού Υπερκύβου. Έτσι εξασφαλίζεται ότι ο αριθμός των διαμερίσεων θα είναι ίσος για κάθε μεταβλητή. Αυτή η μέθοδος δειγματοληψίας δεν απαιτεί περισσότερα δείγματα για περισσότερες διαστάσεις και αυτή η ανεξαρτησία είναι το κύριο πλεονέκτημα αυτής της μεθόδου. Το άλλο πλεονέκτημα είναι ότι ο αλγόριθμος διαθέτει μνήμη, δηλαδή «θυμάται» ποια δείγματα έχουν ως τώρα επιλεγεί.

ΚΕΦΑΛΑΙΟ 5. Δυναμική ανάλυση

5.1 Δυναμική ανάλυση των κατασκευών

Οι διεγέρσεις και τα φορτία που αυτές ασκούν σε μία κατασκευή χωρίζονται σε δύο κατηγορίες. Η πρώτη είναι τα στατικά φορτία, τα οποία είναι σταθερά και άρα η ένταση και οι παραμορφώσεις που αναπτύσσονται είναι μοναδικές (κατά τη γραμμική θεωρία). Η δεύτερη κατηγορία είναι τα δυναμικά φορτία, δηλαδή φορτία ή διεγέρσεις των οποίων το μέγεθος, η διεύθυνση ή και η θέση μεταβάλλονται συναρτήσει του χρόνου, με συνέπεια η ένταση και η παραμόρφωση του φορέα να αλλάζουν ανά χρονική στιγμή. Αναπτύσσεται δηλαδή στο φορέα διαφορετική παραμορφωσιακή και εντατική κατάσταση συναρτήσει χρόνου. Επομένως, τα σημεία του φορέα, η μάζα των οποίων δεν αμελείται, αποκτούν ταχύτητα και επιτάχυνση, καθώς αλλάζουν θέση ανα χρονική στιγμή, με συνέπεια να αναπτύσσονται αδρανειακές δυνάμεις. Ο προσδιορισμός της παραμορφώσεως και εντάσεως σε μία κατασκευή, όταν αυτή υποβάλλεται σε δυναμική φόρτιση, αποτελεί το αντικείμενο της Δυναμικής Αναλύσεως των Κατασκευών.

Όταν ο τρόπος με τον οποίο τα φορτία μεταβάλλονται με το χρόνο είναι γνωστός και σταθερός τα δυναμικά φορτία χαρακτηρίζονται αιτιοκρατικά. Διαφορετικά καλούνται τυχαία ή στοχαστικά και ο ακριβής προσδιορισμός της δυναμικής φόρτισης γίνεται ιδιαίτερα απαιτητικός για τον μελετητή. Τέτοιου είδους φορτία είναι τα φορτία που προκαλούνται εξαιτίας του ανέμου, των θαλάσσιων κυματισμών, της διέλευσης οχημάτων σε γέφυρες, του σεισμού και άλλα.

5.2 Ιδιομορφική ανάλυση

Η ιδιομορφική ανάλυση είναι μία από τις μεθόδους δυναμικής ανάλυσης και πρόκειται για τη μελέτη και τον προσδιορισμό των δυναμικών χαρακτηριστικών ενός προβλήματος μηχανικής. Η μέθοδος συνοψίζεται στην εύρεση των ιδιοσυχνοτήτων και ιδιομορφών του συστήματος και τα αποτελέσματά της έχουν μεγάλο εύρος πρακτικών εφαρμογών τόσο στη μελέτη και κατασκευή νέων κατασκευών, όσο και στην επίβλεψη υφιστάμενων.

Ως συχνότητα έχει οριστεί ο αριθμός των επαναλήψεων ενός συμβάντος στη μονάδα του χρόνου, με τη συχνότητα να αποτελεί ιδιότητα του αιτίου που προκαλεί το επαναλαμβανόμενο γεγονός.

Οι ιδιοσυχνότητες ή φυσικές συχνότητες είναι φυσικά χαρακτηριστικά της απόκρισης των συστημάτων και λόγω του ότι δεν εξαρτώνται από κανέναν άλλο εξωτερικό παράγοντα θεωρούνται εσωτερικές ιδιότητες. Οι έννοιες της ιδιοσυχνότητας και ιδιομορφής γίνονται αντιληπτές από τον ορισμό της ελεύθερης ταλάντωσης. Συγκεκριμένα ελεύθερη καλείται η ταλάντωση που εκτελεί ένα μηχανικό σύστημα χωρίς καμία εξωτερική επιρροή εκτός από την αρχική ώθηση που την έθεσε σε κίνηση. Η ταλάντωση που εκτελεί το σύστημα εξαρτάται από τον τρόπο με τον οποίο αυτή αρχικά τέθηκε σε ισχύ, καθώς επίσης από τη μάζα και τη δυσκαμψία του συστήματος. Κάθε ταλάντωση πρόκειται για επαλληλία προδιαγεγραμμένων φυσικών μετατοπίσεων, με κάθε μετατόπιση να είναι ανεξάρτητη με τη δική της μορφή και συχνότητα οι οποίες καλούνται ιδιομορφές και ιδιοσυχνότητες, αντίστοιχα. Όλες οι ιδιομορφές δεν συμβάλλουν το ίδιο στη συνολική μετατόπιση. Αυτές με τη μεγαλύτερη επιρροή ονομάζονται κύριες. Όσο η τάξη των ιδιομορφών αυξάνεται οι ιδιομορφές συνεισφέρουν όλο και λιγότερο και η αξιοπιστία τους μειώνεται. Ο βαθμός συνεισφοράς κάθε ιδιομορφής στην ταλάντωση εξαρτάται από την πηγή που διεγείρει το σύστημα. Στην ιδιομορφική ανάλυση η απόκριση ενός δυναμικού συστήματος εκφράζεται ως υπέρθεση των ιδιομορφών.

Τα βήματα της ιδιομορφικής ανάλυσης είναι:

- Καθορισμός της γεωμετρίας και των ιδιοτήτων της κατασκευής (μάζα, δυσκαμψία, σταθερά απόσβεσης)
- Προσδιορισμός ιδιοσυχνοτήτων και ιδιομορφών
- Απόκριση κατασκευής για κάθε ιδιομορφή
- Σχηματισμός και επίλυση των ιδιομορφικών εξισώσεων για κάθε ιδιομορφή
- Υπολογισμός επικόμβιων μετατοπίσεων για κάθε ιδιομορφή
- Υπολογισμός εντατικών μεγεθών των στοιχείων που σχετίζονται με επικόμβιες μετατοπίσεις για κάθε ιδιομορφή
- Υπέρθεση των συνεισφορών όλων των ιδιομορφών για τον προσδιορισμό της ολικής δυναμικής απόκρισης της κατασκευής.

5.3 Εξίσωση δυναμικής ισορροπίας

Η δοκός είναι ένας επιμήκης ευθύγραμμος φορέας, δηλαδή ένα στοιχείο με μεγάλο μήκος σε σχέση με το πλάτος και το ύψος του, που μπορεί και παραλαμβάνει ή/και μεταφέρει δυνάμεις. Στη δοκό εφαρμόζονται ροπές και δυνάμεις που βρίσκονται στο ίδιο επίπεδο που βρίσκεται και ο διαμήκης άξονάς

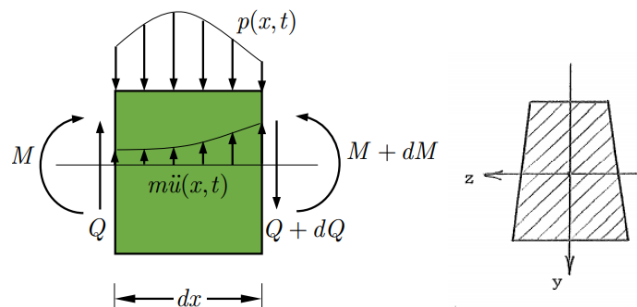
της. Υπόκειται δηλαδή και/ή μόνο σε κάμψη, επομένως ο άξονάς της παραμορφώνεται σε μια καμπύλη. Για τη μοντελοποίηση και επίλυση της συμπεριφοράς της δοκού οι βασικές θεωρίες που έχουν αναπτυχθεί είναι η κλασική Τεχνική Θεωρία και η γενική Θεωρία της Ελαστικότητας, με τις παραλλαγές τους.

Η συνήθως χρησιμοποιούμενη είναι η κλασική θεωρία και ειδικότερα αυτές που αναπτύχθηκαν από τους Euler-Bernoulli, de Saint-Venant και Timoshenko. Οι πρώτες δύο αγνοούν τις διατμητικές παραμορφώσεις, ενώ το μοντέλο του Timoshenko θεωρεί ότι υπάρχει ομοιόμορφη κατανομή διατμητικής παραμόρφωσης κατά μήκος της διατομής της δοκού ενώ δεν ισχύει ότι οι αρχικά επίπεδες και κάθετες διατομές στο άξονα της δοκού παραμένουν επίπεδες και κάθετες στον παραμορφωμένο ουδέτερο άξονα της δοκού, όπως θεωρεί η παραδοχή Bernoulli.

Οι βασικές παραδοχές που πρέπει να ικανοποιούνται για την εφαρμογή της θεωρίας της δοκού των Euler-Bernoulli, είναι γνωστές από την αντοχή των υλικών κατά την κάμψη και είναι:

Η δοκός είναι πρισματική και ευθύγραμμη από ελαστικό, ιστροπικό και ομογενές υλικό. Ο κεντροβαρικός της άξονας είναι μία ευθεία που ορίζεται ως άξονας x και ο κάθετος σε αυτόν άξονας είναι ο y , ο οποίος είναι και άξονας συμμετρίας της διατομής. Όλες οι εγκάρσιες φορτίσεις της δοκού ασκούνται στο επίπεδο συμμετρίας (επίπεδο $x-y$), το οποίο καλείται και επίπεδο φόρτισης. Η δοκός υπο την επίδραση των φορτίσεων θα παραμορφωθεί αλλά θα παραμείνει συμμετρική ως προς το αρχικό επίπεδο συμμετρίας. Κάθε διατομή επίπεδη και κάθετη στον άξονα πρισματικής δοκού, παραμένει έτσι και στον καμπυλωμένο άξονα. Το επίπεδο δηλαδή της διατομής διέρχεται από το επίπεδο καμπυλότητας. Επίσης, οι εγκάρσιες παραμορφώσεις θεωρούνται μικρές.

Για την διατύπωση της εξίσωσης της δυναμικής ισορροπίας της δοκού εξετάζεται η ισορροπία ενός στοιχείου της δοκού με μήκος dx , όπως φαίνεται στην Εικόνα 5-1.



Εικόνα 5-1: Ισορροπία ενός στοιχείου της δοκού με μήκος dx αριστερά και το σχήμα της διατομής αριστερά (πηγή: Δυναμική ανάλυση των κατασκευών-Ι.Κατσικαδέλης, 2020)

Όπου η M αναπαριστά την εσωτερική καμπτική ροπή και η Q την εσωτερική εγκάρσια δύναμη, με το $p = p(x, t)$ να είναι το εξωτερικά επιβαλλόμενο φορτίο, το οποίο δύναται να μεταβάλλεται συναρτήσει της θέσης και του χρόνου.

Λαμβάνοντας υπόψη τη δυναμική ισορροπία στο στοιχείο μήκους dx , και χρησιμοποιώντας το Νόμο του Νεύτωνα πρέπει να οριστούν κάποιες επιπλέον παραδοχές για αυτό το τυπικό στοιχείο της δοκού λαμβάνοντας υπόψη και την επιτάχυνσή του. Αυτές οι παραδοχές καθορίζουν το εύρος εφαρμογής αυτής της θεωρίας και ορίζουν πως το υλικό ακολουθεί το νόμο του Hooke, η αξονική παραμόρφωση είναι αμελητέα σε σχέση με την παραμόρφωση κάμψης και η περιστροφή του στοιχείου είναι αμελητέα συγκριτικά με την κάθετη και πλάγια παραμόρφωση.

Με αυτές τις παραδοχές και θέτοντας τις κάθετες δυνάμεις στο στοιχείο της δοκού ίσες με τη μάζα επί την επιτάχυνση, από την εξίσωση δυναμικής ισορροπίας κατά τον κατακόρυφο άξονα y λαμβάνεται

$$\frac{\partial Q}{\partial x} = -p(x, t) + m \frac{\partial^2 u}{\partial t^2} \quad (5-1)$$

Ταυτόχρονα, αθροίζοντας τις ροπές M κάθε στοιχείου μήκους dx της δοκού, ισχύει:

$$\frac{\partial M}{\partial x} = Q \quad (5-2)$$

Όπου Q η τέμνουσα δύναμη στο στοιχείο μήκους dx , ρ η πυκνότητα του υλικού της δοκού, A το εμβαδό της διατομής του στοιχείου, $u = u(x, t)$ η εγκάρσια μετατόπιση και άρα $\frac{\partial^2 u}{\partial t^2}$ είναι η επιτάχυνση της δοκού. Συνδυάζοντας τις (5.1),(5.2) προκύπτει:

$$\frac{\partial^2 M}{\partial x^2} = -p(x, t) + \rho A \frac{\partial^2 u}{\partial t^2} \quad (5-3)$$

Από το σχήμα της παραμόρφωσης, δεδομένου ότι ικανοποιούνται οι υποθέσεις που τέθηκαν και χρησιμοποιώντας τον νόμο του Hooke προκύπτει:

$$\frac{\partial^2 u}{\partial x^2} = - \frac{M}{EI} \quad (5-4)$$

Όπου E το μέτρο ελαστικότητας, I η ροπή αδράνειας ως προς τον άξονα z . Το μέτρο δυσκαμψίας $EI(x)$ και η μάζα $m(x)$ μπορούν να μεταβάλλονται με τη θέση x , ενώ όπως έχει αναφερθεί το φορτίο $p = p(x, t)$ μπορεί να μεταβάλλεται ανά χρονική στιγμή και θέση.

Τελικά από τις (5-3) και (5-4) προκύπτει:

$$EI \frac{\partial^4 u}{\partial x^4} + m \frac{\partial^2 u}{\partial t^2} = p(x, t) \quad (5-5)$$

η οποία είναι μία μερική διαφορική εξίσωση υπερβολικού τύπου και αναπαριστά τη γνωστή εξίσωση δυναμικής ισορροπίας ή εξίσωση κινήσεως της ταλαντευόμενης δοκού.

5.4 Ιδιομορφική ανάλυση γενικευμένων ελαστικών συστημάτων

Όπως προαναφέρθηκε οι εγκάρσιες καμπτικές ταλαντώσεις μίας δοκού περιγράφονται από την εξίσωση (5-5), όπου οι αρχικές και συνοριακές συνθήκες του εκάστοτε συστήματος είναι γνωστές.

Αν στην (5-5) θεωρηθεί $p(x, t) = 0$, δηλαδή μηδενική φόρτιση, προκύπτει η ομογενής εξίσωση:

$$EI \frac{\partial^4 u}{\partial x^4} + m \frac{\partial^2 u}{\partial t^2} = 0 \quad (5-6)$$

Όπου \bar{m} = σταθερό, δηλαδή η μάζα ανα μονάδα μήκους θεωρείται ότι διατηρείται σταθερή.

Η (5-6) περιγράφει τις ελεύθερες ταλαντώσεις της δοκού, τις οποίες έχει προκαλέσει κάποια αρχική διέγερση και αποτελεί μια μερική διαφορική εξίσωση. Για την επίλυσή της θα χρησιμοποιηθεί η μέθοδος των χωριζομένων μεταβλητών, όπου η λύση θα είναι το γινόμενο μιας χωρικής και μιας χρονικής μεταβλητής με μορφή:

$$u(x, t) = \phi(x)Y(t) \quad (5-7)$$

Με αντικατάσταση της (5-7) στην (5-6) και διαιρώντας με $\bar{m}\phi(x)Y(t)$ προκύπτει:

$$\frac{EI}{\bar{m}} \frac{\phi^{(iv)}(x)}{\phi(x)} = - \frac{\ddot{Y}(t)}{Y(t)} \quad (5-8)$$

με την παραδοχή ότι οι ρίζες του παρανομαστή στην εξίσωση (5-8) συμπίπτουν με αυτές του αριθμητού επιτρέποντας έτσι την διαίρεση (Cain, G.L and Meyer et al., 2006).

Δεδομένου ότι το αριστερό μέλος της εξισώσεως (5-8) εξαρτάται αποκλειστικά από το x και το δεξί από το t , προκύπτει ότι τα δύο μέλη είναι ίσα με μια σταθερά λ :

$$\frac{EI}{\bar{m}} \frac{\phi^{(iv)}(x)}{\phi(x)} = - \frac{\ddot{Y}(t)}{Y(t)} = \lambda \quad (5-9)$$

$$\ddot{Y}(t) + \lambda Y(t) = 0 \quad (5-10)$$

$$\phi^{(iv)}(x) - \lambda \frac{\bar{m}}{EI} \phi(x) = 0 \quad (5-11)$$

Για να παριστάνει η (5-10) ταλάντωση, το λ πρέπει να είναι θετική σταθερά, $\lambda = \omega^2$. Έτσι, η λύση της (5-10) είναι:

$$Y(t) = A \cos \omega t + B \sin \omega t \quad (5-12)$$

όπου A,B είναι αυθαίρετες σταθερές εξαρτώμενες από τις αρχικές συνθήκες $Y(0)$ και $\dot{Y}(0)$. Έτσι προκύπτει:

$$Y(t) = Y(0) \cos \omega t + \frac{\dot{Y}(0)}{\omega} \sin \omega t \quad (5-13)$$

όπου ω είναι η αρχικά άγνωστη φυσική συχνότητα.

Η (5-11) λαμβάνει τη μορφή:

$$\phi^{(iv)}(x) - \beta^4 \phi(x) = 0 \quad (5-14)$$

Έχοντας θέσει:

$$\beta^4 = \omega^2 \frac{\bar{m}}{EI} \quad (5-15)$$

Όπου με την κατάλληλη μαθηματική επεξεργασία και κάνοντας χρήση του τύπου του Euler προκύπτει η λύση της (5-11) ως:

$$\phi(x) = C_1 \cos \beta x + C_2 \sin \beta x + C_3 \cosh \beta x + C_4 \sinh \beta x \quad (5-16)$$

στην οποία οι $C_i (i = 1,2,3,4)$ είναι αυθαίρετες σταθερές που μπορούν να προσδιορισθούν από τις συνοριακές συνθήκες, δηλαδή τις συνθήκες που επικρατούν στις στηρίξεις της δοκού ενός ανοίγματος.

5.4.1 Αναλυτική λύση ελεύθερης ταλάντωσης για αμφιέρειστη δοκό

Έστω αμφιέρειστη δοκός μήκους L . Οι συνοριακές συνθήκες που ισχύουν συναρτήσει της μετατόπισης u , αφορούν την κάθετη μετατόπιση της δοκού στα άκρα της $x = 0$ και $x = L$, και την καμπυλότητα επίσης στα άκρα της $x = 0$ και $x = L$. Οι συνοριακές αυτές συνθήκες ισχύουν για κάθε τιμή t και είναι όλες μηδενικές:

$$u(0,t) = 0, \quad M(0,t) = -EIu''(0,t) = 0 \quad (5-17)$$

$$u(L,t) = 0, \quad M(L,t) = -EIu''(L,t) = 0 \quad (5-18)$$

Αντικαθιστώντας τις (5-17) και (5-18) στην (5-7) και επειδή οι σχέσεις ισχύουν για κάθε χρονική στιγμή t μπορεί να γραφτεί:

$$\phi(0) = 0, \quad \phi''(0) = 0 \quad (5-19)$$

$$\phi(L) = 0, \quad \phi''(L) = 0 \quad (5-20)$$

Όπου με αντικατάσταση των εξισώσεων (5-19) στην (5-16) προκύπτει $C_1 = C_3 = 0$, ενώ αντικαθιστώντας τις εξισώσεις (5-20) στην (5-16) προκύπτει:

$$C_2 \sin \beta L + C_4 \sinh \beta L = 0 \quad (5-21)$$

$$\beta^2 [-C_2 \sin \beta L + C_4 \sinh \beta L] = 0 \quad (5-22)$$

Επιλύοντας το σύστημα των εξισώσεων (5-21) και (5-22) για την εύρεση των C_2, C_4 , έτσι ώστε το σύστημα να έχει μη τετριμμένη λύση, θα πρέπει να ισχύει $\sin \beta L = 0$, και άρα:

$$\beta_n = \frac{n\pi}{L}, \quad n = 1, 2, \dots \quad (5-23)$$

Αντικαθιστώντας την (5-23) στην (5-15) προκύπτουν οι αντίστοιχες φυσικές συχνότητες:

$$\omega_n = n^2 \pi^2 \sqrt{\frac{EI}{mL^4}} \quad (5-24)$$

Ενώ από την επεξεργασία του συστήματος των εξισώσεων (5-21) και (5-22) λαμβάνεται $C_4 = 0, C_2 =$ αυθαίρετο, άρα τελικά:

$$\phi(x) = C_2 \sin \beta x \quad (5-25)$$

Όπου αντικαθιστώντας στην (5-25) την (5-15) προκύπτουν οι ιδιομορφές ταλάντωσης:

$$\phi_n(x) = C_2 \sin\left(\frac{n\pi}{L} x\right), \quad n = 1, 2, \dots \quad (5-26)$$

5.4.2 Αναλυτική λύση ελεύθερης ταλάντωσης για δοκό πρόβολο

Σε δοκό πρόβολο ισχύουν οι συνοριακές συνθήκες στα $x=0$ και $x=L$ για κάθε t :

$$u(0, t) = 0, \quad u'(0, t) = 0 \quad (5-27)$$

$$M(L, t) = -EIu''(L, t) = 0, \quad Q(L, t) = -EIu'''(L, t) = 0 \quad (5-28)$$

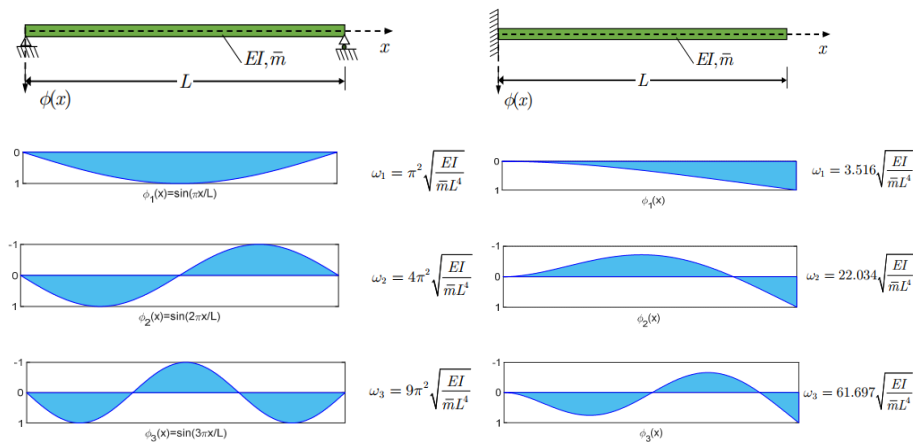
Όπως φαίνεται για $x=0$ σε κάθε χρονική στιγμή t η μετατόπιση και η στροφή είναι μηδενικές, ενώ για $x=L$ σε κάθε χρονική στιγμή t η καμπυλότητα και η τέμνουσα είναι μηδενικές.

Ακολουθώντας παρόμοια διαδικασία με αυτήν που περιγράφηκε στην περίπτωση της αμφιέριστης δοκού προκύπτουν οι ιδιομορφές με τις οποίες ταλαντώνεται ο πρόβολος όπως παρουσιάζονται στην Εικόνα 5–2 και η μαθηματική περιγραφή των οποίων είναι:

$$\phi_n(x) = C_1 \left[\cos \beta_n x - \cosh \beta_n x - \frac{\cos \beta_n L + \cosh \beta_n L}{\sin \beta_n L + \sinh \beta_n L} (\sin \beta_n x - \sinh \beta_n x) \right] \quad (5-29)$$

Σημειώνεται ότι για την περίπτωση δοκού πρόβολου οι τιμές των ριζών $\beta_n L$ λαμβάνονται προσεγγιστικά:

$$\beta_n L \simeq \frac{\pi}{2} (2n - 1) \quad (5-30)$$



Εικόνα 5–2: Ιδιομορφές ταλαντώσεως και φυσικές συχνότητες αμφιέριστης δοκού σταθερής διατομής αριστερά και πρόβολου σταθερής διατομής δεξιά (πηγή: Δυναμική ανάλυση των κατασκευών-Ι. Κατσιακαδέλης, 2020)

5.5 Εξαναγκασμένες ταλαντώσεις δοκών

Στην περίπτωση των εξαναγκασμένων ταλαντώσεων στη δοκό υποβάλλονται οι αρχικές συνθήκες

$$u(x, 0) = f(x), \quad \dot{u}(x, 0) = g(x) \quad (5-31)$$

Η μέθοδος χωριζομένων μεταβλητών εφαρμόζεται και για την επίλυση της εξίσωσης (5-5) που περιγράφει το φαινόμενο των εξαναγκασμένων ταλαντώσεων σε δοκό.

Η ολική καμπτική μετατόπιση των σημείων της δοκού προκύπτει ως η επαλληλία όλων των ιδιομορφικών συνεισφορών και είναι:

$$u(x, t) = \sum_{n=1}^{\infty} \phi_n(x) Y_n(t) \quad (5-32)$$

Με χρήση της σχέσης (5-15) και της συνθήκης ορθογωνικότητας, η οποία ισχύει για το άπειρο σύνολο των ιδιομορφών της ελεύθερης ταλαντώσεως λαμβάνεται:

$$M_n \ddot{Y}_n(t) + K_n Y_n(t) = P_n(t) \quad (5-33)$$

Όπου η ιδιομορφική μάζα, η ιδιομορφική ακαμψία και η ιδιομορφική εξωτερική δύναμη (φορτίο) είναι αντίστοιχα:

$$M_n = \bar{m} \int_0^L \phi_n^2(x) dx, \quad K_n = \omega_n^2 M_n, \quad P_n = \int_0^L \phi_n(x) p(x, t) dx \quad (5-34)$$

Ενώ η λύση της (5-33) είναι η:

$$Y_n(t) = \frac{\dot{Y}_n(0)}{\omega_n} \sin \omega_n t + Y_n(0) \cos \omega_n t + \frac{1}{M_n \omega_n} \int_0^t P_n(\tau) \sin[\omega_n(t - \tau)] d\tau \quad (5-35)$$

Οι αρχικές συνθήκες $Y_n(0), \dot{Y}_n(0)$ για τη χρονική συνάρτηση προκύπτουν από τη σχέση (5-32) ως:

$$u(x, 0) = \sum_{n=1}^{\infty} \phi_n(x) Y_n(0) = f(x) \quad (5-36)$$

Πολλαπλασιάζοντας τη σχέση (5-36) με $\phi_n(x)$, χρησιμοποιώντας τις συνθήκες ορθογωνικότητας και τελικά ολοκληρώνοντας στο διάστημα $[0, L]$ λαμβάνονται οι:

$$Y_n(0) = \frac{\int_0^L f(x) \phi_n(x) dx}{\int_0^L \phi_n^2(x) dx} \quad (5-37)$$

Και

$$\dot{Y}_n(0) = \frac{\int_0^L g(x) \phi_n(x) dx}{\int_0^L \phi_n^2(x) dx} \quad (5-38)$$

ΚΕΦΑΛΑΙΟ 6. Εφαρμογή της μεθόδου PINNs στην εξίσωση Burgers

6.1 Εφαρμογή 1: Πρόβλημα επαγωγής εξίσωσης Burgers

Η εξίσωση Burgers αποτελεί την απλούστερη ΜΔΕ που περιέχει τον μη γραμμικό όρο μεταφοράς και τον όρο διάχυσης. Εφαρμόζεται σε αρκετούς τομείς των εφαρμοσμένων μαθηματικών με ορισμένους από αυτούς να είναι η μηχανική των ρευστών, η θερμοδυναμική και η θεωρία πιθανοτήτων. Προκύπτει από τις εξισώσεις Navier-Stokes για μονοδιάστατη ροή ασυμπίεστου ρευστού με αμελητέα τη βαρύτητα και την κλίση πίεσης. Είναι παραβολική ΜΔΕ με γενική της μορφή:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2} \quad (6-1)$$

Όπου $u(x,t)$ ένα δοσμένο πεδίο και ν γνωστό κινηματικό ιξώδες.

Για μία χωρική διάσταση, η εξίσωση Burgers μαζί με τις οριακές συνθήκες του Dirichlet (Dirichlet boundary conditions), λαμβάνει τη μορφή:

$$\begin{aligned} u_t + uu_x - \left(\frac{0.01}{\pi}\right)u_{xx} &= 0, \quad x \in [-1,1], \quad t \in [0,1] \\ u(0,x) &= -\sin(\pi x), \\ u(t,-1) &= u(t,1) = 0 \end{aligned} \quad (6-2)$$

Όπου ο δείκτης t δηλώνει τη μερική παράγωγο ως προς το χρόνο, ο δείκτης x τη μερική παράγωγο ως προς το χώρο και ο δείκτης xx τη δεύτερη μερική παράγωγο ως προς τον χώρο αντίστοιχα.

Έστω τώρα η συνάρτηση $f(t, x)$:

$$f := u_t + uu_x - \left(\frac{0.01}{\pi}\right)u_{xx} \quad (6-3)$$

Με τη λύση $u(x,t)$ της ΜΔΕ να προσεγγίζεται από ένα βαθύ ΤΝΔ (deep neural network), η $f(t, x)$ αποτελεί ένα δίκτυο PINN. Οι κοινές παράμετροι των δικτύων $u(x,t)$ και $f(t,x)$ προσδιορίζονται από την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος, σχέση (4-4) με τον τρόπο που περιγράφηκε στην παράγραφο 4.3, με τον όρο MSE_u να αναφέρεται στα αρχικά και συνοριακά σημεία του προβλήματος και η μορφή του δίδεται από τη σχέση (4-5), ενώ ο όρος MSE_f αφορά έναν πεπερασμένο αριθμό εσωτερικών σημείων, δίδεται από τη σχέση (4-6) και θα πρέπει να έχει μηδενική τελική τιμή ορίζοντας έτσι την δομή της εξίσωσης που περιγράφει τον φυσικό νόμο του φαινομένου.

Παρακάτω δίδεται τμήμα του κώδικα στο οποίο ορίζεται η ΜΔΕ καθώς και τα νευρωνικά δίκτυα *net_u* και *net_f*. Ο κώδικας γράφτηκε στην γλώσσα προγραμματισμού Python, ενώ ταυτόχρονα χρησιμοποιήθηκε το TensorFlow1, μια από τις δημοφιλέστερες και πιο πλήρεις ελεύθερες βιβλιοθήκες που χρησιμοποιείται για την επίλυση προβλημάτων μηχανικής μάθησης.

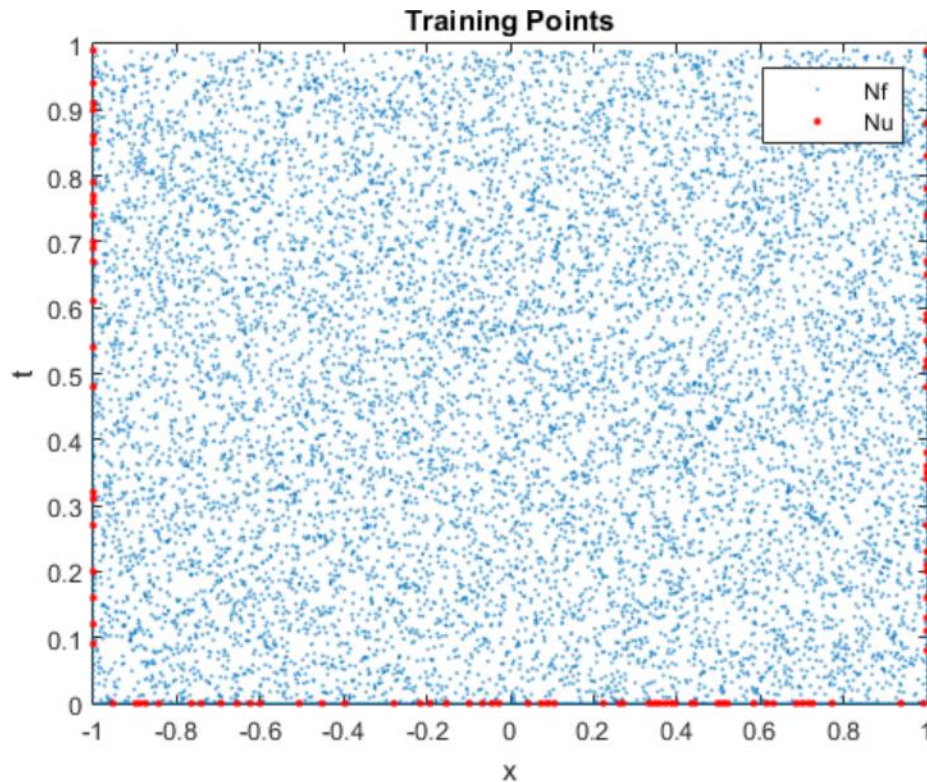
```
def neural_net(self, X, weights, biases):
    num_layers = len(weights) + 1

    H = 2.0*(X - self.lb)/(self.ub - self.lb) - 1.0
    for l in range(0,num_layers-2):
        W = weights[l]
        b = biases[l]
        H = tf.tanh(tf.add(tf.matmul(H, W), b))
    W = weights[-1]
    b = biases[-1]
    Y = tf.add(tf.matmul(H, W), b)
    return Y

def net_u(self, x, t):
    u = self.neural_net(tf.concat([x,t],1), self.weights, self.biases)
    return u

def net_f(self, x,t):
    u = self.net_u(x,t)
    u_t = tf.gradients(u, t)[0]
    u_x = tf.gradients(u, x)[0]
    u_xx = tf.gradients(u_x, x)[0]
    f = u_t + u*u_x - self.nu*u_xx
    return f
```

Στη συγκεκριμένη εφαρμογή για την επίλυση της εξίσωσης Burgers χρησιμοποιήθηκε ένα ΤΝΔ με 8 κρυμμένα επίπεδα, με 20 νευρώνες σε κάθε ένα από αυτά. Το επίπεδο εισόδου έχει 2 νευρώνες καθότι 2 είναι οι είσοδοι στο δίκτυο ενώ το επίπεδο εξόδου έχει 1 νευρώνα, αφού δίνει μία έξοδο. Τα διαθέσιμα δεδομένα αφορούν 100 χρονικές στιγμές, με $t \in [0,0.99]$ με $\Delta t=0.01\text{sec}$, 256 χωρικές θέσεις, $x \in [-1,1]$ με $\Delta x= 0.0078\text{m}$ και τις παρατηρηθείσες μετατοπίσεις κάθε σημείου ανά χρονική στιγμή. Άρα συνολικά υπάρχουν 25600 δείγματα. Ως δεδομένα εκπαίδευσης των αρχικών και συνοριακών συνθηκών λαμβάνονται τυχαία, από σύνολο 456 σημείων (ζεύγη (x, t)) τα 100 από αυτά ($Nu = 100$), με τα οποία τροφοδοτείται το νευρωνικό δίκτυο *u*. Από τα 25144 ζεύγη που αποτελούν τα εσωτερικά σημεία επιλέγονται με τη μέθοδο LHS 10000 από αυτά ($Nf= 10000$). Τα ζεύγη αυτά τροφοδοτούν το νευρωνικό δίκτυο *f* και παρουσιάζονται στην Εικόνα 6–1.



Εικόνα 6–1 Τα 100 σημεία των αρχικών και συνοριακών συνθηκών μαζί με 10000 εσωτερικά σημεία του χωρίου, όπως προέκυψαν από τη μέθοδο του λατινικού υπερκύβου LHS

Αφού το δίκτυο εκπαιδευτεί με τα δείγματα που επιλέχθηκαν, θα δώσει λύση για ολο το χωρίο του προβλήματος. Κάτι ανάλογο είναι σαφώς πιο δύσκολο να επιτευχθεί με τις κλασικές αριθμητικές μεθόδους, στις οποίες θα απαιτηθεί μια επίμονη και σχολαστική διακριτοποίηση στο χώρο και στο χρόνο.

Για κάθε νευρώνα η συνάρτηση ενεργοποίησης που εφαρμόζεται είναι η υπερβολική εφαπτομένη \tanh . Η προσέγγιση της λύσης $u(x,t)$ προκύπτει από το μέσο τετραγωνικό σφάλμα, σχέση (4-4), και ο αλγόριθμος βελτιστοποίησης που χρησιμοποιείται είναι ο L-BFGS-B. Δεν δίνεται συγκεκριμένος αριθμός εποχών εκπαίδευσης και άρα η εκπαίδευση επαναλαμβάνεται έως ότου υπερβεί κάποιον από τους περιορισμούς του L-BFGS-B. Οι περιορισμοί που τίθενται είναι:

- αριθμός των επαναλήψεων να μην υπερβεί τις 50000 ($\text{maxiter}=50000$),
- ο αριθμός των υπολογισμών των συναρτήσεων (loss function) να μην ξεπερνά επίσης τις 50000 ($\text{maxfun}=50000$), Ως Function evaluations δηλαδή υπολογισμοί συνάρτησης λογίζονται οι φορές που υπολογίζεται η συνάρτηση σφάλματος (loss function) σε κάθε επανάληψη (iteration) για τον υπολογισμό της παραγώγου στις gradient based μεθόδους βελτιστοποίησης. Στη συγκεκριμένη

περίπτωση ο αριθμός των επαναλήψεων και ο αριθμός των υπολογισμών της loss function είναι ίδιος. Ωστόσο, σε περιπλοκότερους αλγορίθμους βελτιστοποίησης, η loss function ενδεχομένως να υπολογίζεται παραπάνω από 1 φορές για την ακριβή εύρεση των παραγώγων

- οι μετρήσεις διορθώσεων των μεταβλητών του δικτύου να μην ξεπερνούν τις 50 ($maxcor=50$). Αυτές είναι που καθορίζουν και την περιορισμένη μνήμη του πίνακα L-BFG-S, καθότι δεν αποθηκεύει στη μνήμη ολόκληρο τον Εσσιανό πίνακα, αλλά μόνο όσους όρους ορίζονται στη μεταβλητή $maxcor$, ως μια προσέγγιση του Εσσιανού πίνακα. Η διαδικασία βελτιστοποίησης με χρήση παραγώγων δεν είναι τίποτα άλλο από μια προσπάθεια επίλυσης μη γραμμικής αλγεβρικής εξίσωσης, που προκύπτει από το μηδενισμό της αντικειμενικής συνάρτησης (εδώ το loss function). Αυτό που κάνει ο αλγόριθμος είναι να δοκιμάσει διάφορες τιμές των μεταβλητών που ορίζουν τη συνάρτηση και με αυτές τις διάφορες τιμές υπολογίζει παραγώγους της συνάρτησης. Αυτή είναι η εσσιανή του πίνακα. Επίσης είναι γνωστό ότι η κατεύθυνση προς την οποία μειώνεται η παράγωγος δίνει την κατεύθυνση που ελαχιστοποιείται η συνάρτηση, αρα δίνει μια κατεύθυνση στην οποία θα προβλέψω κάποια νέα x για να υπολογίσω μια πιθανή ελάχιστη τιμή της συναρτησης
- ο μέγιστος αριθμός των βημάτων line search ανα επανάληψη ($maxls=50$) Line search είναι μια μέθοδος η οποία βοηθάει στην αποφυγή εγκλωβισμού της συνάρτησης βελτιστοποίησης (ειδικά στις τις gradient based μεθόδους βελτιστοποίησης) σε τοπικά ελάχιστα για την εύρεση του ολικού ελαχίστου. Αρα η line search βοηθάει τον αλγόριθμο να ξεπεράσει πιθανά τοπικά ελάχιστα. Πραγματοποιεί αναζήτηση στην περιοχή γύρω από το σημείο που αρχικά έχει βρεθεί ελάχιστο από τις παραγώγους, σημείου που αποτελεί άλλη θέση τοπικού ελαχίστου.
- Η εκάστοτε επανάληψη ορίζεται να σταματά όταν ικανοποιείται η συνθήκη, η οποία ορίζει τον μικρότερο δυνατό θετικό αριθμό που μπορεί να αναπαραστήσει το υπολογιστικό σύστημα που χρησιμοποιείται. Πρόκειται δηλαδή για μια ασυμπτωτική μικρή πιθανότητα που χρησιμοποιείται για την αποφυγή της λύσης της μαθηματικής παράστασης $\log 0$ καθότι αυτή απειρίζεται.

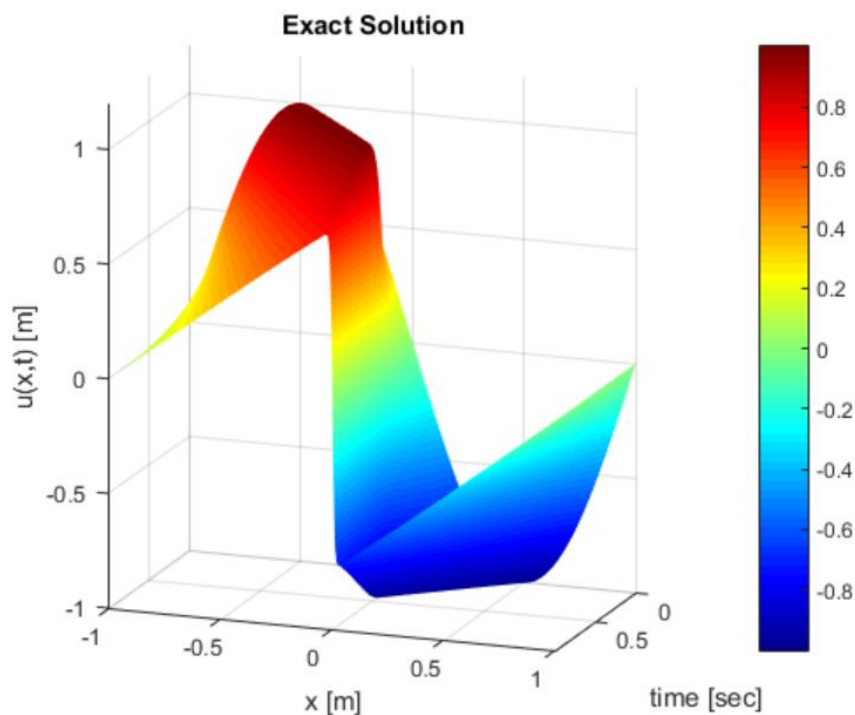
Η απόκλιση της πρόβλεψης του δικτύου από τις πραγματικές τιμές, δηλαδή το σφάλμα του δικτύου, υπολογίστηκε με χρήση της σχετικής L2 ή ευκλείδειας νόρμας, σχέση (6-4), και προέκυψε ίση με $2.57 \times 10^{-3} \%$.

Η εξίσωση που την περιγράφει την ευκλείδια νόρμα είναι η εξής:

$$L_2 \text{ error} = \frac{\sqrt{\sum_{i=1}^N (u_{\text{exact}}(i) - u_{\text{PINN}}(i))^2}}{\sqrt{\sum_{i=1}^N u_{\text{exact}}(i)^2}} \quad (6-4)$$

όπου στον αριθμητή υπολογίζεται η νόρμα της διαφοράς της εκτιμώμενης τιμής του i στοιχείου $u_{\text{PINN}(i)}$ από την αναλυτική λύση $u_{\text{exact}(i)}$ και διαιρείται με τη νόρμα της αναλυτικής λύσης.

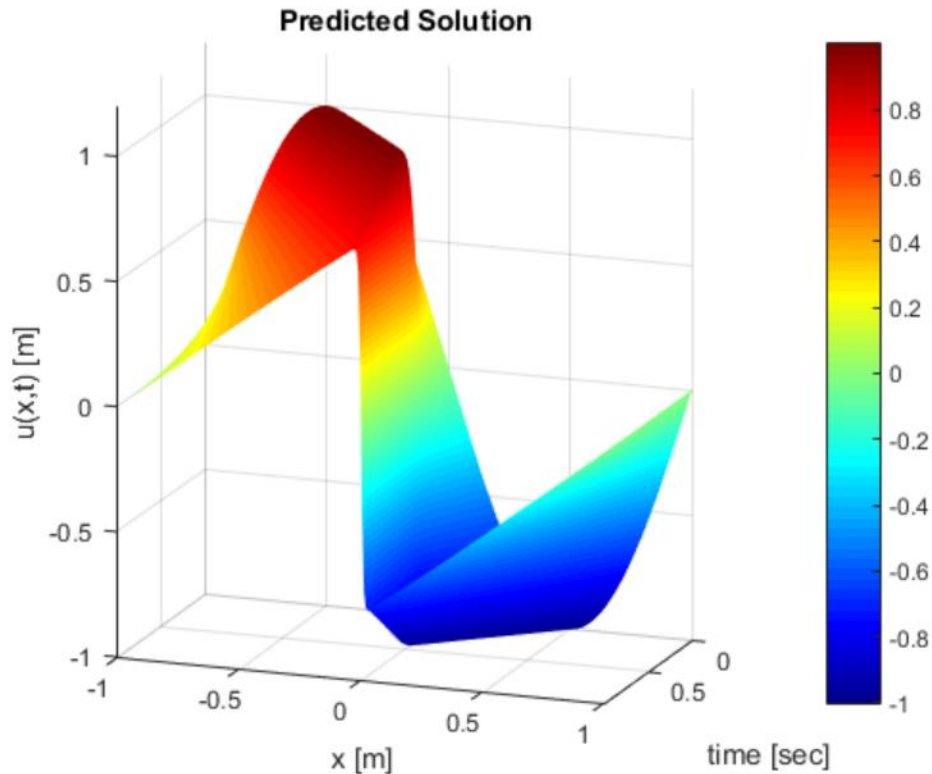
Η αναλυτική λύση της εξίσωσης Burgers απεικονίζεται στην Εικόνα 6–2, ενώ η Εικόνα 6–3 απεικονίζει τη λύση που προέκυψε από την επίλυση του PINN.



Εικόνα 6–2: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Burgers

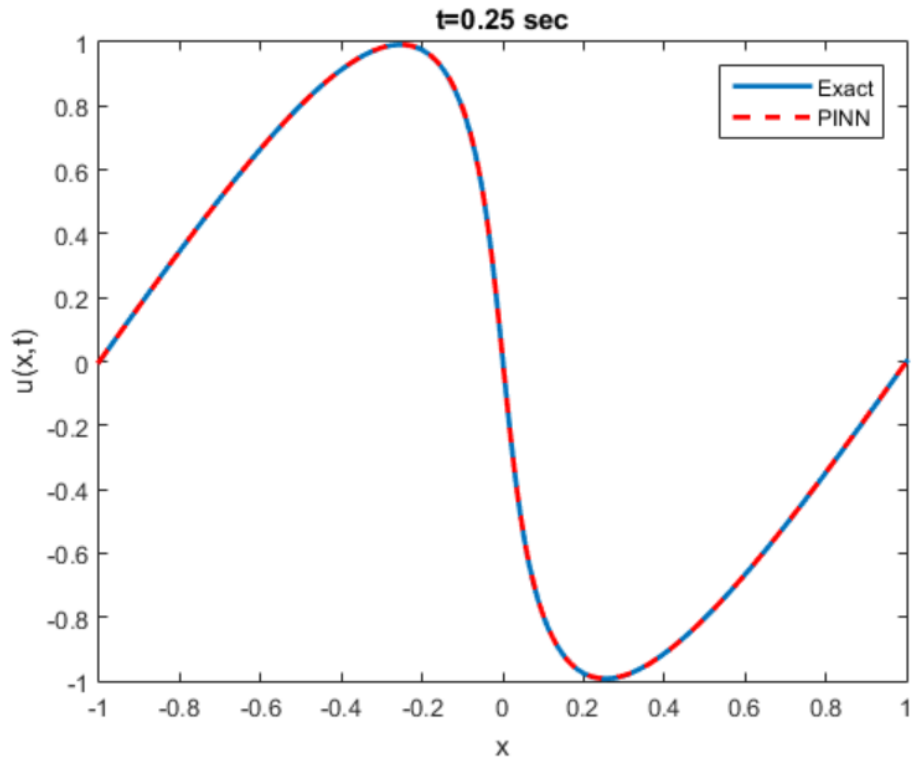
Τα γραφήματα της Εικόνα 6–4, της Εικόνα 6–5, της Εικόνα 6–6 αποτυπώνουν και συγκρίνουν τη γραφική παράσταση της αναλυτικής λύσης της ΜΔΕ και της πρόβλεψης του PINN για τις χρονικές στιγμές $t= 0.25, 0.50, 1 \text{ sec}$, αντίστοιχα. Επιπλέον στα γραφήματα της Εικόνα 6–7, της Εικόνα 6–8 και της Εικόνα 6–9 αποτυπώνονται και συγκρίνονται η γραφική παράσταση της αναλυτικής λύσης της ΜΔΕ και της πρόβλεψης του δικτύου στις θέσεις $x= -0.50, 0.25, 0.95 \text{ m}$, αντίστοιχα. Από τα διαγράμματα αυτά συμπεραίνεται ότι το δίκτυο αξιοποιώντας ένα μέρος των αρχικών συνθηκών, των συνοριακών συνθηκών και των εσωτερικών σημείων προσέγγισε με αρκετά μεγάλη ακρίβεια την πολυπλοκότητα και

τη μη γραμμικότητα της εξίσωσης Burgers αποφεύγοντας τη διακριτοποίηση στον χώρο και τον χρόνο που θα απαιτούνταν για την εφαρμογή των κλασικών αριθμητικών μεθόδων.

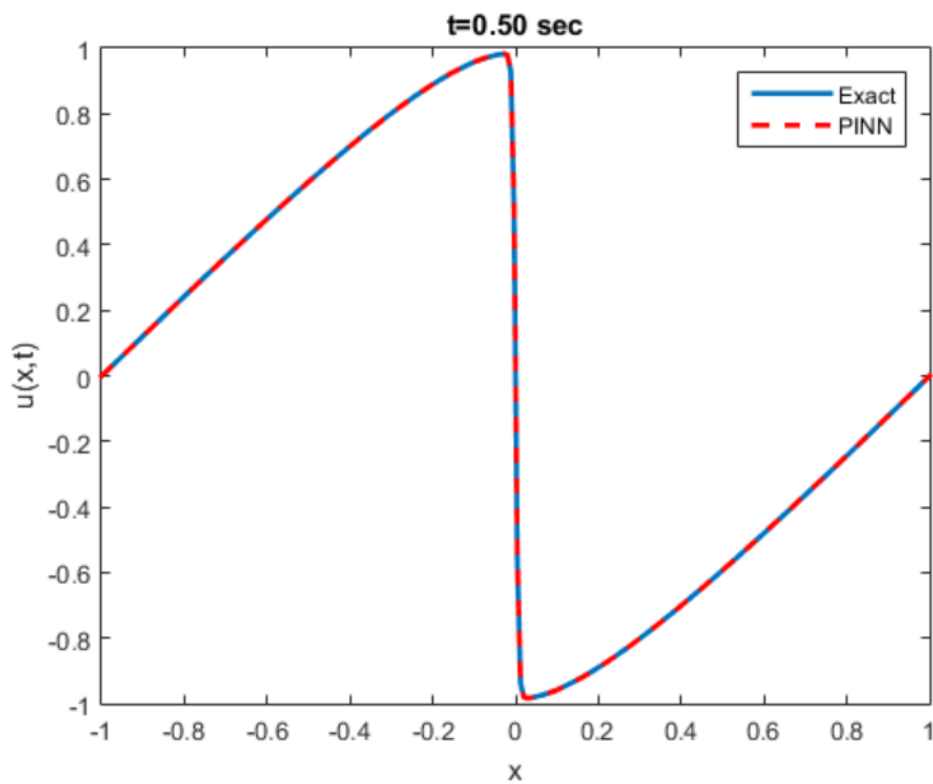


Εικόνα 6–3: Τρισδιάστατη απεικόνιση της λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Burgers.

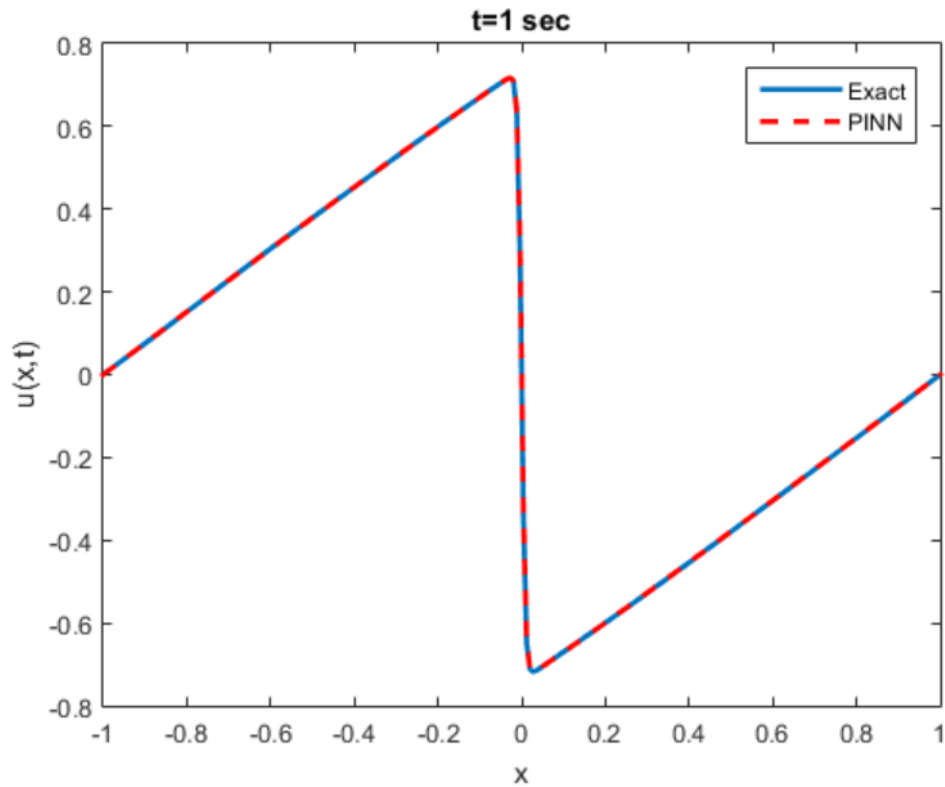
Η αρχιτεκτονική του δικτύου, το πλήθος των σημείων που επιλέγονται και οι συντελεστές βαρύτητας κάθε όρου στη συνάρτηση σφάλματος πρέπει να ερευνώνται προκειμένου να επιτευχθεί μεγαλύτερη ακρίβεια στην πρόβλεψη. Συγκεκριμένη μεθοδολογία δεν υπάρχει και η λύση με τη μεγαλύτερη προγνωστική ακρίβεια προκύπτει εμπειρικά. Η γενικότερη κατεύθυνση που συστήνεται από τους συγγραφείς των PINNs για τη βελτίωση της πρόβλεψης είναι αύξηση των δεδομένων εισόδου Nu και Nf , και / ή αύξηση των κρυμμένων επιπέδων και νευρώνων. Τα συμπεράσματα αυτά απεικονίζονται στην Εικόνα 6–10, όπου για αριθμό νευρώνων μικρότερο του 4 και πλήθος επιπέδων μικρότερο του 40 το σφάλμα είναι μεγαλύτερο. Στα παραδείγματα αυτά χρησιμοποιήθηκαν συνοριακά σημεία και σημεία αρχικών συνθηκών, $Nu = 100$, και εσωτερικά σημεία $Nf=10000$. Επίσης, στην Εικόνα 6–11 για 8 κρυφά επίπεδα με 20 νευρώνες το καθένα προκύπτει ότι αυξάνοντας το πλήθος των εσωτερικών σημείων πάνω από 5000 το σφάλμα πρόβλεψης μικραίνει ανεξαρτήτως του πλήθους των Nu .



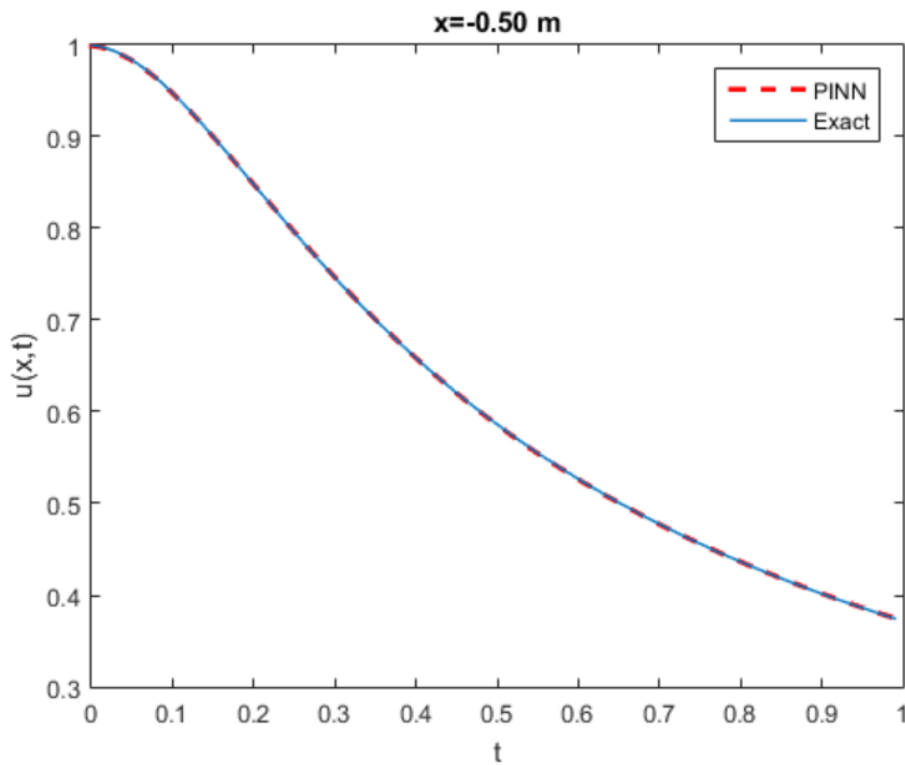
Εικόνα 6-4: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $t= 0.25$ sec



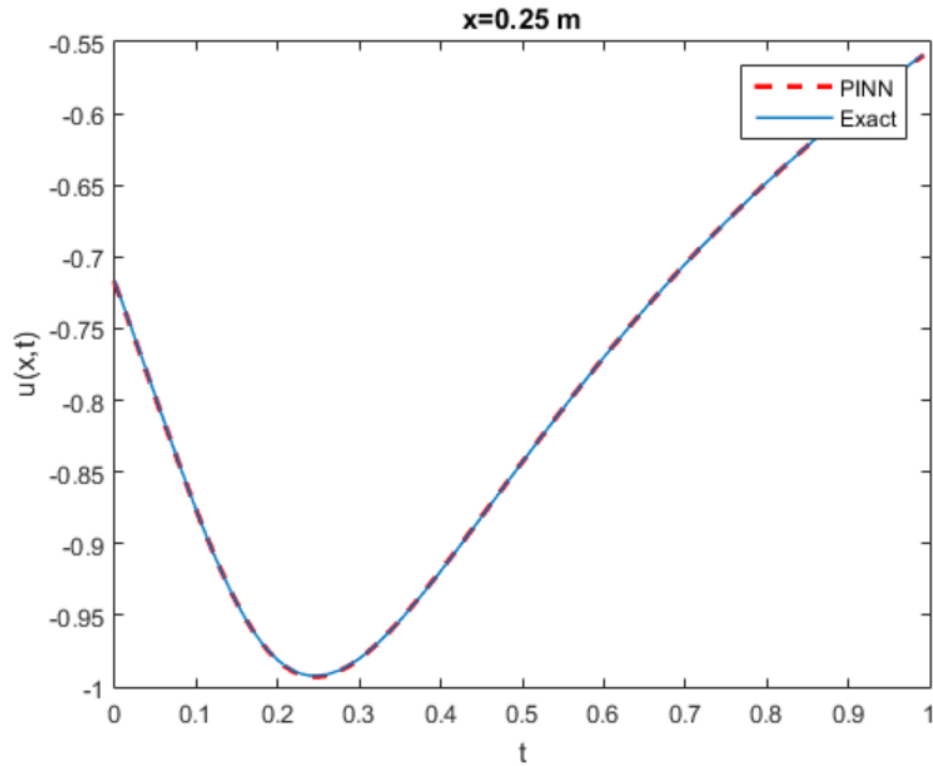
Εικόνα 6-5: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $t= 0.50$ sec



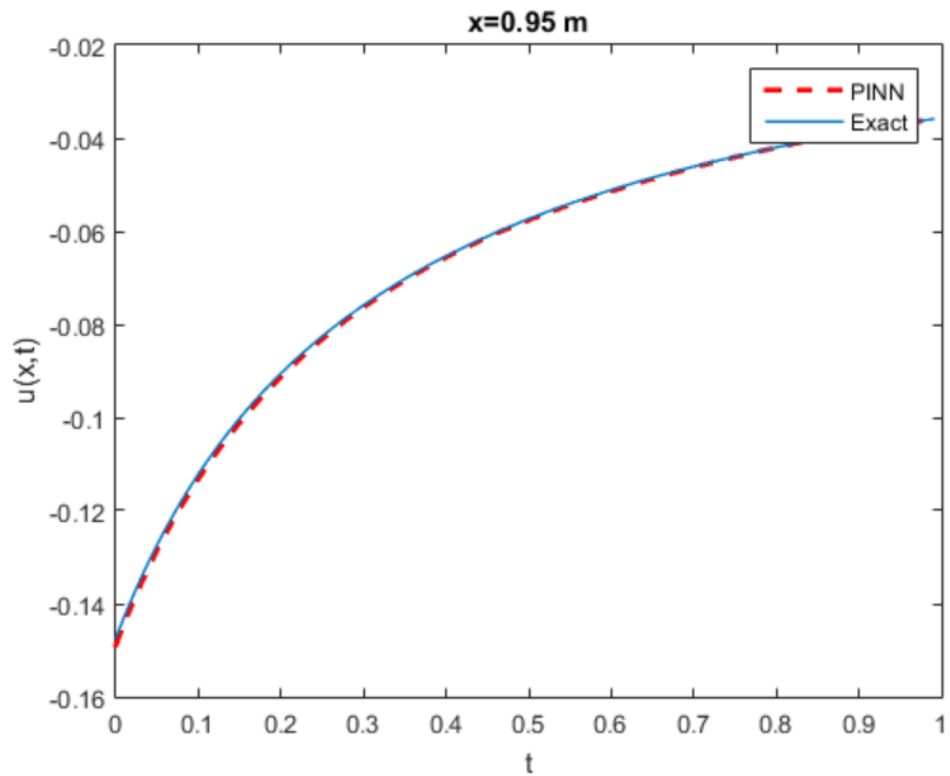
Εικόνα 6–6: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $t = 1 \text{ sec}$



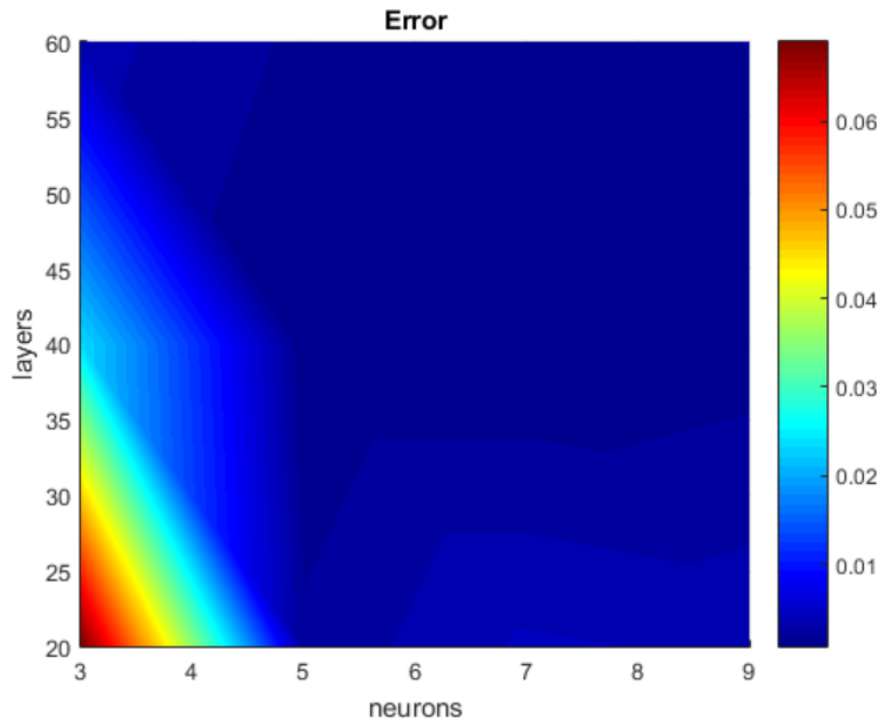
Εικόνα 6–7: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $x = -0.50 \text{ m}$



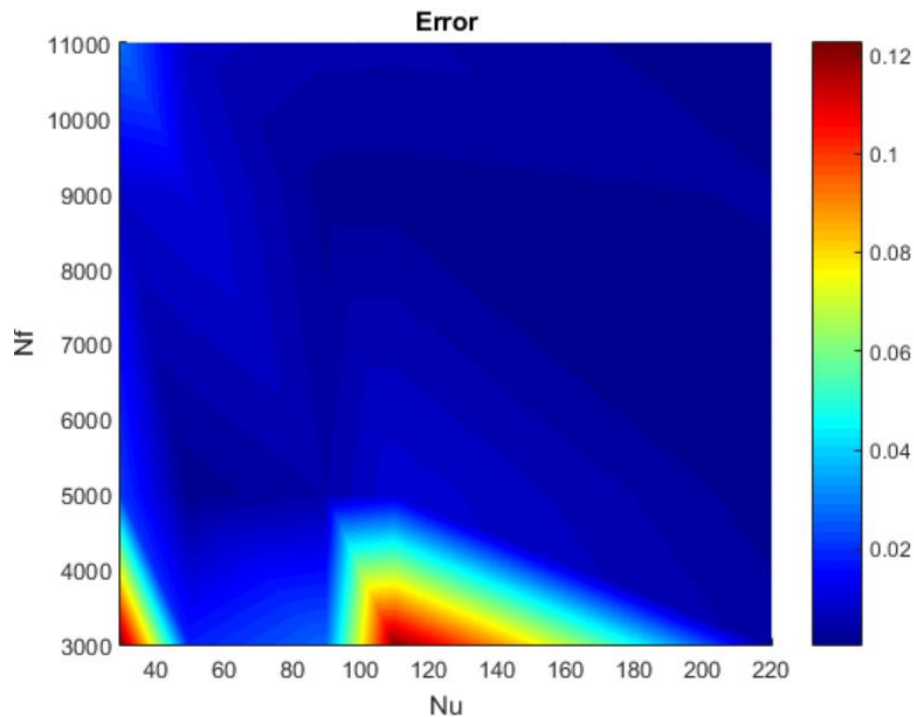
Εικόνα 6–8: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $x=0.25m$



Εικόνα 6–9: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για $x=0.95m$



Εικόνα 6–10: Σύγκριση του σφάλματος μεταξύ αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για τιμές νευρώνων και κρυφών επιπέδων. Τα συνοριακά και σημεία αρχικών συνθηκών είναι σταθερά ίσα με $Nu = 100$, και τα εσωτερικά σημεία $Nf=10000$.



Εικόνα 6–11: Σύγκριση του σφάλματος μεταξύ αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Burgers για διαφορετικά πλήθη συνοριακών και σημείων αρχικών συνθηκών. Οι νευρώνες και τα κρυφά επίπεδα είναι σταθερά ίσα με 8 και 20, αντίστοιχα.

6.2 Εφαρμογή 2: Ταυτοποίηση παραμέτρων εξίσωσης Burgers

6.2.1 Περίπτωση δεδομένων χωρίς θόρυβο

Για την ταυτοποίηση των παραμέτρων της εξίσωσης στα δεδομένα εισόδου του PINN (identification πρόβλημα) χρησιμοποιείται η ίδια αρχιτεκτονική δικτύου με το πρόβλημα της επαγωγής, καθώς επίσης και τα ίδια δεδομένα. Στο συγκεκριμένο πρόβλημα κατά την εκπαίδευση του Ν.Δ. δεν γίνεται διαφοροποίηση μεταξύ των δεδομένων που περιγράφουν τις αρχικές και συνοριακές συνθήκες και μεταξύ των εσωτερικών σημείων, αλλά επιλέγονται τυχαία 2000 ζεύγη. Το κινηματικό ιξώδες ν αντιμετωπίζεται σαν παράμετρος του προβλήματος και έτσι ορίζεται στον κώδικα, όπως επίσης ως παράμετρος λαμβάνεται και η σταθερά έμπροσθεν του δεύτερου όρου της εξίσωσης (6-1). Ταυτόχρονα ορίζεται και η u ως μεταβλητή για την οποία επιχειρείται ταυτοποίηση.

Αρχικά στις παραμέτρους δίδονται οι τιμές όπως ορίζονται στο παρακάτω τμήμα του κώδικα

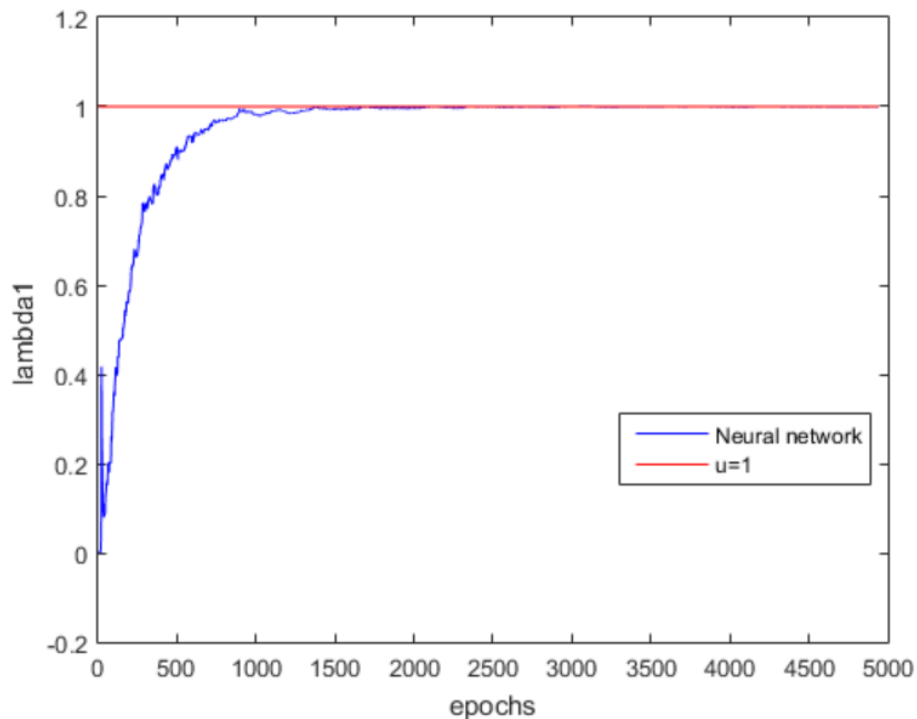
```
self.lambda_1 = tf.Variable([0.0], dtype=tf.float32)
self.lambda_2 = tf.Variable([-6.0], dtype=tf.float32)
```

Η παράμετρος της σταθεράς του δεύτερου όρου της εξίσωσης Burgers που αποτελεί την $lambda_1$ στον κώδικα αρχικά λαμβάνει την τιμή 0, ενώ η παράμετρος ν , που αποτελεί την $lambda_2$ στον κώδικα, αρχικά πήρε την τιμή -0.6. Μέσω της διαδικασίας της βελτιστοποίησης οι τιμές αυτές μεταβάλλονται κατά την διάρκεια των εποχών (epochs) έτσι ώστε τελικά να βρεθούν οι τιμές των η παραμέτρων που θα συντελέσουν σε ορθή πρόβλεψη. Με τη χρήση πολλών εποχών επιχειρείται η βέλτιστη λύση του προβλήματος και η ακριβέστερη εύρεση των παραμέτρων, αφού κάθε εποχή αποτελεί ένα γύρο εκπαίδευσης πάνω στα δεδομένα Αρχικά χρησιμοποιήθηκε ο αλγόριθμος βελτιστοποίησης Adam για 10000 εποχές, ώστε να επιταχυνθεί η διαδικασία της βελτιστοποίησης και ο αλγόριθμος να συγκλίνει στην τιμή του σφάλματος εκπαίδευσης και στη συνέχεια χρησιμοποιήθηκε ο αλγόριθμος L-BFGS. Η μεταβολή της παραμέτρου του κινηματικού ιξώδες ν και της παραμέτρου u κατά την διάρκεια των εποχών φαίνονται στο Σχήμα. Το σφάλμα του PINN προκύπτει $1.47 * 10^{-3}$, η τιμή σφάλματος στην πρόβλεψη του I1 είναι $lambda_1 = 0.046\%$ και τιμή σφάλματος στην πρόβλεψη του I2 είναι $lambda_2 = 0.06\%$. Η μεταβολή του σφάλματος της εκτίμησης της παραμέτρου του κινηματικού ιξώδες ν και της παραμέτρου της σταθεράς του δεύτερου όρου της εξίσωσης Burgers κατά την διάρκεια

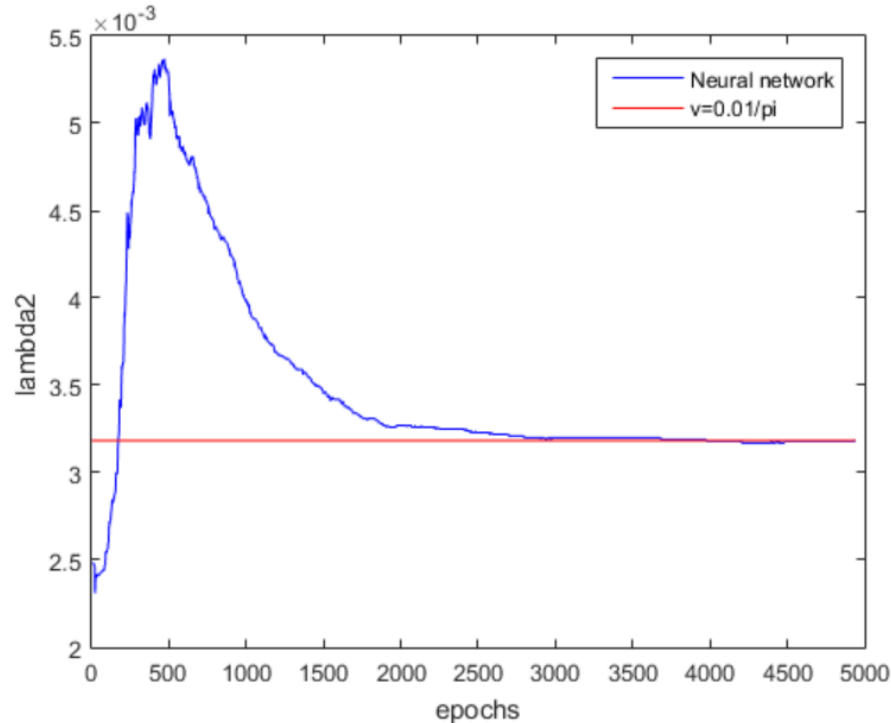
των εποχών φαίνονται στις Εικόνες Εικόνα 6–12Εικόνα 6–13. Ο κατακόρυφος άξονας κάθε διαγράμματος υποδεικνύει τη μεταβολή της παραμέτρου και ο οριζόντιος αριθμεί τις εποχές. Η πραγματική τιμή της παραμέτρου, που αποτελεί και την τιμή-στόχο απεικονίζεται με την σταθερή ευθεία. Κατα αυτόν τον τρόπο διαφαίνεται πώς το πλήθος των εποχών εκπαίδευσης επηρεάζει την ταυτοποίηση των παραμέτρων, καθώς και την ταχύτητα με την οποία επιτυγχάνεται η προσέγγιση της επιθυμητής τιμής.

Όπως φαίνεται, αναφορικά με την παράμετρο u , ξεκινώντας από την τιμή 0, η οποία τέθηκε ως αρχική τιμή, το ΤΝΔ μόλις από τις πρώτες 1000 εποχές συγκλίνει στην πραγματική τιμή. Η τελική τιμή της παραμέτρου u λίγο πριν τις 4000 εποχές παίρνει την τιμή $u = 0.99867$, όπου σε σύγκριση με την πραγματική τιμή $u=1$ συμπεραίνεται ότι το ΤΝΔ προσεγγίζει τη λύση με μεγάλη ακρίβεια.

Όμοια είναι και τα αποτελέσματα της τιμής του κινηματικού ιξώδους, όπου το ΤΝΔ μόλις από τις πρώτες 2000 εποχές συγκλίνει στην πραγματική τιμή. Η τελική τιμή της παραμέτρου ν λίγο πριν τις 4000 εποχές παίρνει την τιμή $\nu = 0.00325$, όπου σε σύγκριση με την πραγματική τιμή $\nu = 0.01/\pi = 0.0032$ συμπεραίνεται ότι το ΤΝΔ προσεγγίζει τη λύση με μεγάλη ακρίβεια, όπως φαίνεται και στα σχήματα Εικόνα 6–12Εικόνα 6–13.



Εικόνα 6–12: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου u της εξίσωσης Burgers.



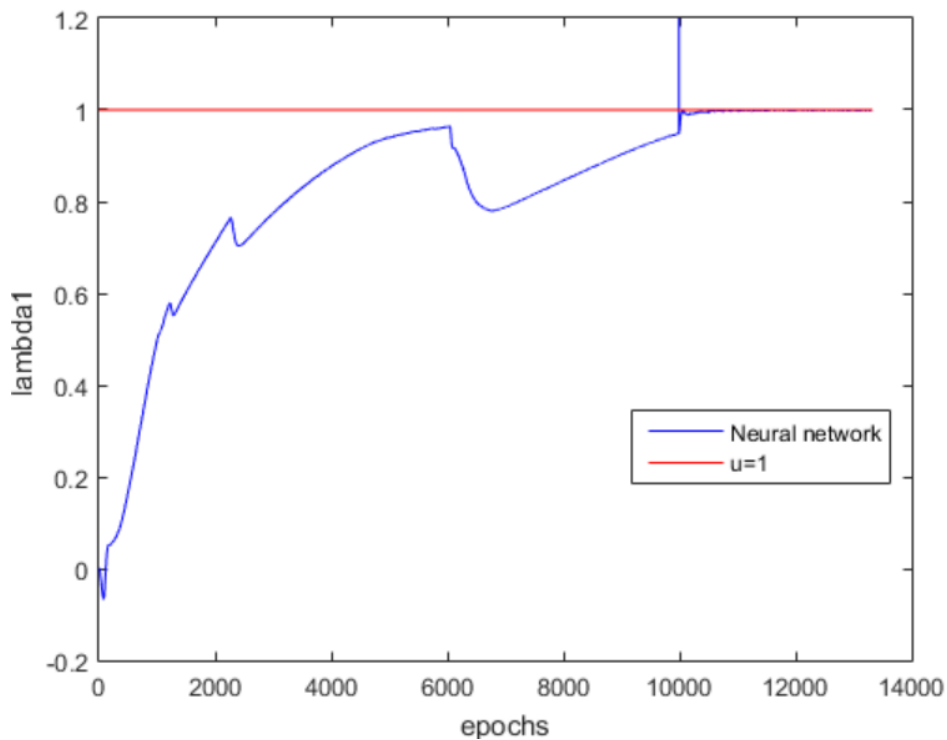
Εικόνα 6–13: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου κινηματικού ιξώδες ν της εξίσωσης Burgers.

6.2.2 Περίπτωση δεδομένων με θόρυβο

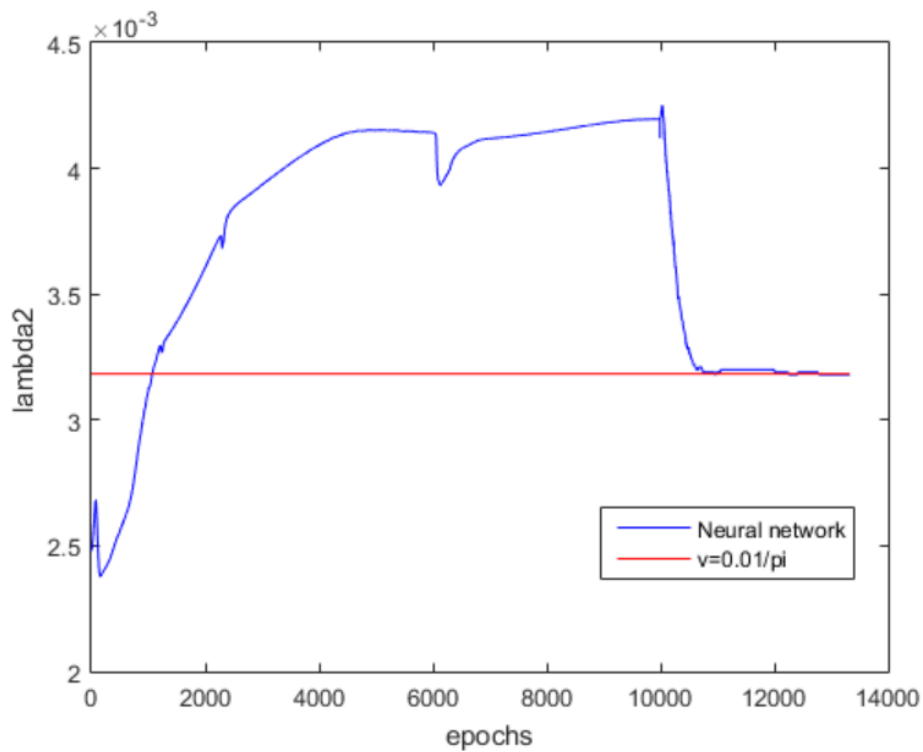
Πραγματοποιήθηκε επίσης προσπάθεια ταυτοποίησης των ίδιων παραμέτρων και χρησιμοποιώντας το ίδιο ΝΔ προσθέτοντας θόρυβο στις μετρήσεις που τροφοδοτούν το δίκτυο. Συγκεκριμένα ο θόρυβος λήφθηκε ίσος με 0.01. Χρησιμοποιήθηκαν δύο αλγόριθμοι βελτιστοποίησης ο αλγόριθμος Adam αρχικά και στη συνέχεια ο L-BFGS. Αυτό διότι με αυτόν τον τρόπο επιταχύνεται η διαδικασία. Ο Adam φέρνει γρήγορα το ΝΔ κοντά σε μία καλή εκτίμηση. Ο L-BFGS ξεκινάει από εκεί που σταμάτησε ο Adam και τον φέρνει αργά στη βέλτιστη λύση. Προφανώς η αποκλειστική χρήση του Adam απαιτεί πολλά iteration για την εύρεση λύσης, ενώ μόνο με τον L-BFGS-B απαιτούνται πολλές επαναλήψεις και αρκετός χρόνος.

Τα αποτελέσματα που προκύπτουν είναι $\text{error } \lambda_{\text{lambda1}}=0.156\%$ και $\text{error } \lambda_{\text{lambda2}}= 0.541\%$

Η μεταβολή του σφάλματος της εκτίμησης της παραμέτρου του κινηματικού ιξώδες ν και της παραμέτρου u κατά την διάρκεια των εποχών φαίνονται στα Σχήματα Εικόνα 6–14 Εικόνα 6–15.



Εικόνα 6–14: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου u της εξίσωσης Burgers για δεδομένα με θόρυβο.



Εικόνα 6–15: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου κινηματικού ιξώδους ν της εξίσωσης Burgers για δεδομένα με θόρυβο.

Για την παράμετρο u ξεκινώντας από την τιμή 0, η οποία τέθηκε ως αρχική τιμή, το ΤΝΔ από τις πρώτες 10000 εποχές συγκλίνει στην πραγματική τιμή. Η τελική τιμή της παραμέτρου u λίγο πριν τις 14000 εποχές παίρνει την τιμή $u = 0.99844$, και άρα το ΤΝΔ προσεγγίζει τη λύση με μεγάλη ακρίβεια με περισσότερες όμως εποχές από τα δεδομένα χωρίς θόρυβο.

Όμοια είναι και τα αποτελέσματα της τιμής του κινηματικού ιξώδους, όπου το ΤΝΔ από τις 10000 εποχές και έπειτα συγκλίνει στην πραγματική τιμή. Η τελική τιμή της παραμέτρου u λίγο πριν τις 14000 εποχές παίρνει την τιμή $\nu = 0.00317$, αρκετά κοντά στην πραγματική τιμή $\nu = 0.01/\pi = 0.0032$.

Συμπερασματικά, το ΤΝΔ μπορεί να δώσει πολύ ικανοποιητικές προβλέψεις των παραμέτρων της εξίσωσης και για δεδομένα εκπαίδευσης με θόρυβο, πραγματοποιώντας όμως 3.5 φορές περισσότερες εποχές εκπαίδευσης.

ΚΕΦΑΛΑΙΟ 7. Εφαρμογή της μεθόδου PINN στην ανάλυση διάδοσης κύματος σε δοκό

7.1 Περιπτώσεις προβλημάτων επαγωγής (inference)

7.1.1 Εφαρμογή 1: Εξίσωση ταλάντωσης αμφιέρειστης δοκού υπό ημιτονοειδές φορτίο

Για την σύγκριση της μεθόδου των ΤΝΔ, αρχικά πραγματοποιήθηκε επίλυση με αριθμητική μέθοδο. Η λύση της εξίσωσης (5.5) πραγματοποιήθηκε με αριθμητική ολοκλήρωση μέσω γραμμικών πεπερασμένων διαφορών, σύμφωνα με τις εξισώσεις ταλαντούμενης αμφιέρειστης δοκού. όπως περιγράφηκε στο προηγούμενο κεφάλαιο, στο προγραμματιστικό περιβάλλον matlab.

Η εφαρμογή έγινε για περίπτωση χαλύβδινης δοκού με τα παρακάτω χαρακτηριστικά:

- Μήκος αμφιέρειστης δοκού $L = 5$
- Μέτρο ελαστικότητας δοκού $E = 210 * 10^6 \text{ KPa}$
- Διατομή δοκού $IPE 80$
- Εμβαδόν διατομής δοκού $A = 7.64 * 10^{-4} \text{ m}^2$
- Ροπή αδράνειας διατομής $I = 80.1 * 10^{-8} \text{ m}^4$
- Πυκνότητα διατομής $\rho = 7500 \text{ kg/m}^3$

Επιλέχθηκε ημιτονοειδές φορτίο μορφής $p(x, t) = \sin(2\pi t)$ και η επίλυση πραγματοποιήθηκε για μια ολόκληρη περίοδο ($T = 1 \text{ sec}$), δηλαδή για χρόνο $0 \leq t \leq 1 \text{ sec}$ για χρονική και χωρική διακριτοποίηση, $\Delta t = 0.001 \text{ sec}$ και $\Delta x = 0.05 \text{ m}$ αντίστοιχα. Επιλέχθηκε μεγαλύτερη διακριτοποίηση στον χρόνο σε σχέση με το χώρο καθότι διαπιστώθηκε ότι με αυτόν τον τρόπο το δίκτυο προσέγγιζε καλύτερα την επιθυμητή λύση αναφορικά με τη συμπεριφορά της δοκού κατά τις αρχικές συνθήκες. Συγκεκριμένα, ο τρόπος με τον οποίο έχει οριστεί ο αλγόριθμος PINN που χρησιμοποιήθηκε στην προκειμένη περίπτωση, λαμβάνει έναν τυχαίο αριθμό από το σύνολο των αρχικών και συνοριακών συνθηκών. Μικρότερη διακριτοποίηση στον χρόνο θα έχει ως αποτέλεσμα λιγότερα δεδομένα εκπαίδευσης αρχικών συνθηκών και άρα μη ακριβής πρόβλεψη της συμπεριφοράς της δοκού για $t = 0$.

Η εξίσωση που περιγράφει τη συμπεριφορά της δοκού είναι η (5-32), η οποία προκύπτει από την επαλληλία όλων των ιδιομορφικών συνεισφορών ταλάντωσης. Επομένως, επιχειρείται αρχικά επίλυση της (5-33), όπου για τις ιδιομορφές ταλάντωσης σύμφωνα με τη σχέση (5-26) για αυθαίρετη σταθερά $C_2 = 1$ και με πλήθος ιδιομορφών $n=20$, ισχύει:

$$\varphi_n = \sin \frac{n\pi x}{L},$$

Όπου $n=1,2,3,\dots,20$ ο δείκτης της ιδιομορφής της ταλάντωσης.

Επιπλέον, από την σχέση (5-24) ισχύει:

$$\omega_n = n^2 \pi^2 \sqrt{\frac{EI}{mL^4}}$$

Άρα με γνωστά τα M_n, K_n από τις σχέσεις (5-34) απομένει η επίλυση της P_n , όπου αντικαθιστώντας κατάλληλα προκύπτει:

$$P_n = \sin(2\pi t) \int_0^L \varphi_n(x) dx$$

Για $a = \int_0^L \varphi_n(x) dx$ και $\Omega = 2\pi$ η σχέση(5-33) με αντικατάσταση λαμβάνει τη μορφή:

$$M_n Y_n''(t) + K_n Y_n(t) = a * \sin(\Omega t)$$

Η λύση Y_n είναι της μορφής:

$$Y_n(t) = A \cos(\omega_n t) + B \sin(\omega_n t) + \frac{1}{1 - \beta_n^2} \sin(\omega_n t)$$

Όπου

$$A = Y_n(0), \quad B = \frac{Y_n'(0)}{\omega_n}, \quad \beta_n = \frac{\Omega}{\omega_n}$$

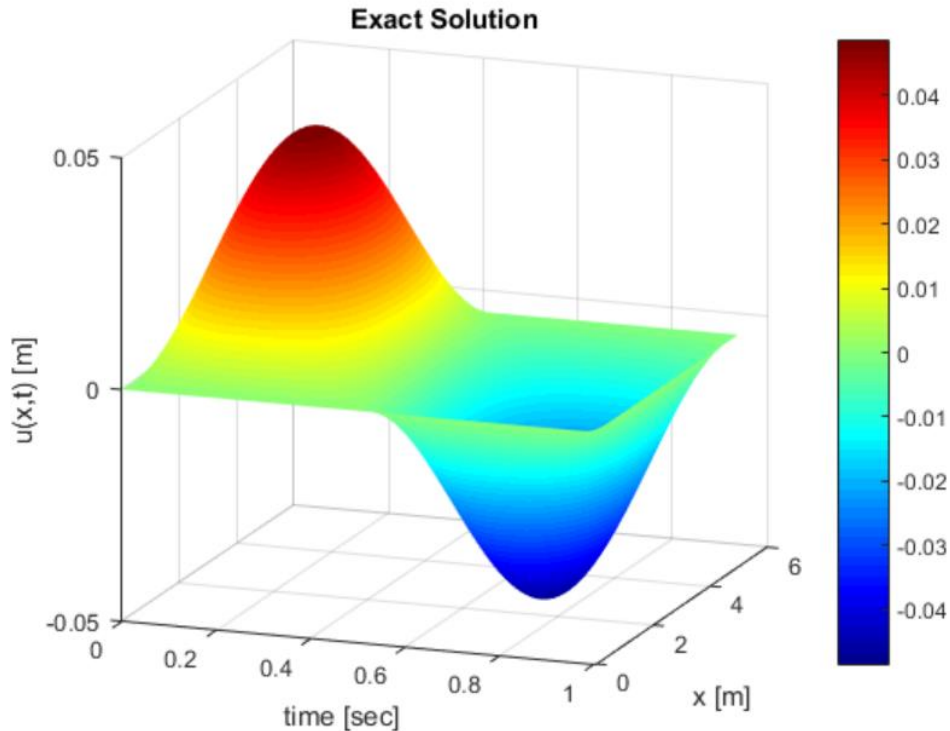
Με τα A,B να είναι αυθαίρετες σταθερές που καθορίζονται από τις αρχικές συνθήκες. Στην προκειμένη περίπτωση εξετάστηκαν οι μηδενικές αρχικές συνθήκες της μετατόπισης και της ταχύτητας κάθε σημείου της δοκού σε κάθε ιδιομορφή, $Y_n(0) = 0$ και $Y_n'(0) = 0$.

Από την αναλυτική επίλυση (Exact solution) προέκυψε η τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου, όπως παρουσιάζεται στην Εικόνα 7-1.

Είναι εμφανές ότι οι αρχικές συνθήκες ($t = 0$) είναι μηδενικές σε όλη τη δοκό και ότι στα δύο άκρα της δοκού οι μετατοπίσεις παραμένουν μηδενικές όλο το διάστημα. Η μέγιστη και η ελάχιστη μετατόπιση συμβαίνουν σε χρόνους $\frac{T}{4} = 0.24 \text{ s}$ και $\frac{3T}{4} = 0.75 \text{ s}$, αντίστοιχα και ισούνται με $u_{min} = -0.0484 \text{ m}$ και $u_{max} = 0.0484 \text{ m}$.

Οι τιμές των κάθετων μετατοπίσεων που προκύπτουν από την αναλυτική επίλυση, καθώς και οι καμπυλότυπες των σημείων της δοκού (δηλαδή οι τιμές της 2^{ης} παραγώγου της κατακόρυφης

μετατόπισης) χρησιμοποιούνται ως δεδομένα εκπαίδευσης στο PINN, αφού με αυτές τις τιμές μπορεί να καθοριστεί επαρκώς η συμπεριφορά της εκάστοτε ταλαντούμενης δοκού.



Εικόνα 7–1: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$

Το ΤΝΔ που χρησιμοποιήθηκε για την επίλυση της ΜΔΕ της ταλάντωσης αμφιέρειστης δοκού δημιουργήθηκε στο προγραμματιστικό περιβάλλον της Python, με χρήση της βιβλιοθήκης Tensorflow.

Ορίζεται το PINN ως η συνάρτηση $f(t, x)$:

$$f := EI \frac{\partial^4 u}{\partial x^4} + m \frac{\partial^2 u}{\partial t^2} - p(x, t)$$

Όπου $p(x, t) = \sin(2\pi t)$

Το δίκτυο τροφοδοτείται και εκπαιδεύεται με τα δεδομένα εισόδου x_i, t_i, u_i όπως αυτά έχουν προκύψει από την αναλυτική επίλυση του προβλήματος.

Έχει συμπεριληφθεί το κομμάτι του κώδικα στο οποίο ορίζεται το PINN, δηλαδή η υπό μελέτη ΜΔΕ:

```
def net_f(self, x,t):

    u,u_xx = self.net_u(x,t)
    u_t = tf.gradients(u, t)[0]
    u_tt = tf.gradients(u_t, t)[0]
    u_xxx = tf.gradients(u_xx, x)[0]
    u_xxxx = tf.gradients(u_xxx, x)[0]

    pi_2=np.pi*2.0
    f = self.nu2*u_tt + self.nu1*u_xxxx-tf.sin(pi_2*t, name ='sin')

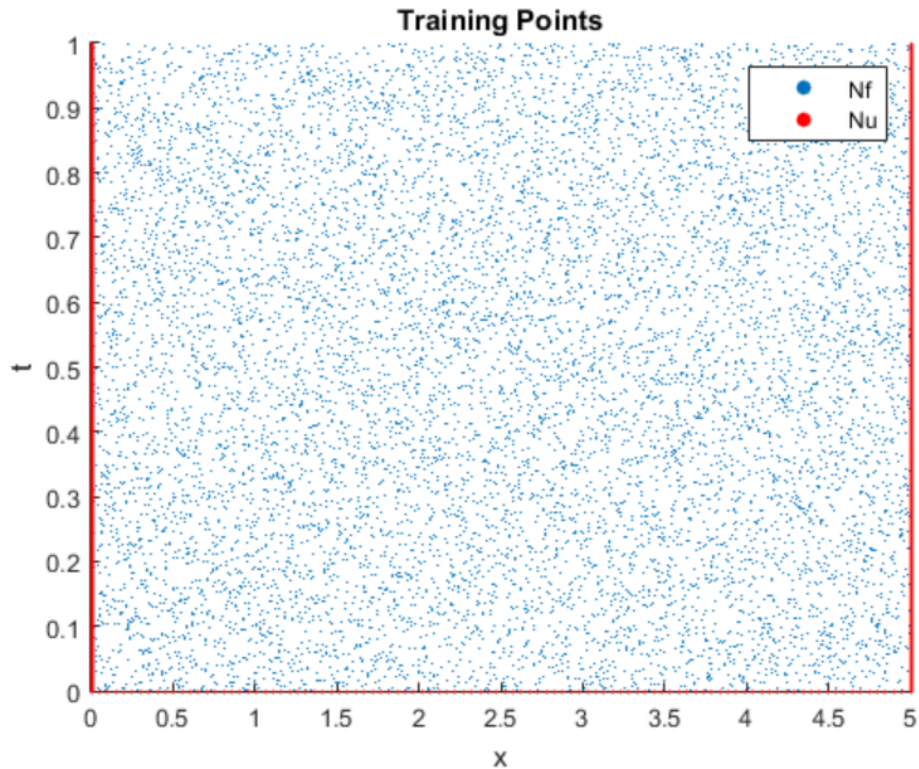
    return f
```

Για την εκπαίδευση του ΤΝΔ πραγματοποιήθηκαν αρκετές δοκιμές προκειμένου να το δίκτυο να δώσει την καλύτερη πρόβλεψη. Τελικά επιλέχθηκαν τα σημεία όλων των αρχικών ($t=0$) και συνοριακών συνθηκών ($x=0$ και $x=L=5$), $N_u=2102$, ενώ από τα εσωτερικά επιλέχθηκαν $N_f=10000$ τυχαία καταναμημένα σημεία στο πεδίο ορισμού της εξίσωσης, από σύνολο 100100 σημείων (το 10% των συνολικών σημείων).

Στην Εικόνα 7–2 παριστάνονται τα σημεία που τυχαία επιλέχθηκαν για την εκπαίδευση του δικτύου. Ο χρόνος t λαμβάνει τιμές στο διάστημα $[0,1]$, ενώ ο χώρος x στο διάστημα $[0,5]$.

Η αρχιτεκτονική του ΤΝΔ αποτελείται από 8 επίπεδα νευρώνων, με κάθε κρυφό επίπεδο να περιέχει 20 νευρώνες και με συνάρτηση ενεργοποίησης την υπερβολική εφαπτομένη \tanh .

Η εκπαίδευση έγινε χρησιμοποιώντας την συνάρτηση του μέσου τετραγωνικού σφάλματος. Συγκεκριμένα, υπολογίζεται από το δίκτυο το άθροισμα των σφαλμάτων των μετατοπίσεων u των συνοριακών σημείων, των σφαλμάτων των τιμών της καμπυλότητας των συνοριακών σημείων u_{xx} (οι οποίες είναι μηδενικές καθότι αμφιέρειστη δοκός) και του σφάλματος της συνάρτησης f , η οποία θα πρέπει να έχει μηδενική τιμή. Θεωρήθηκε ότι κάθε μία από τις προαναφερθείσες συνιστώσες της συνάρτησης σφάλματος δεν συνεισφέρει εξίσου στο τελικό αποτέλεσμα. Συγκεκριμένα, δόθηκε μεγαλύτερη βαρύτητα στο σφάλμα της μετατόπισης και την καμπυλότητας των συνοριακών σημείων της δοκού, προκειμένου το δίκτυο να εκπαιδευτεί κατάλληλα για τη συγκεκριμένη δοκό. Το σφάλμα της συνάρτησης θεωρείται ότι έχει μικρότερη βαρύτητα στο σφάλμα της εκπαίδευσης του δικτύου, σε ποσοστό τάξης 1 προς 1000.



Εικόνα 7–2: Σχηματική απεικόνιση των δεδομένων εκπαίδευσης του PINN για την επίλυση της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$

Το τμήμα του κώδικα στο οποίο ορίζεται η συνάρτηση σφάλματος της εκπαίδευσης είναι το παρακάτω:

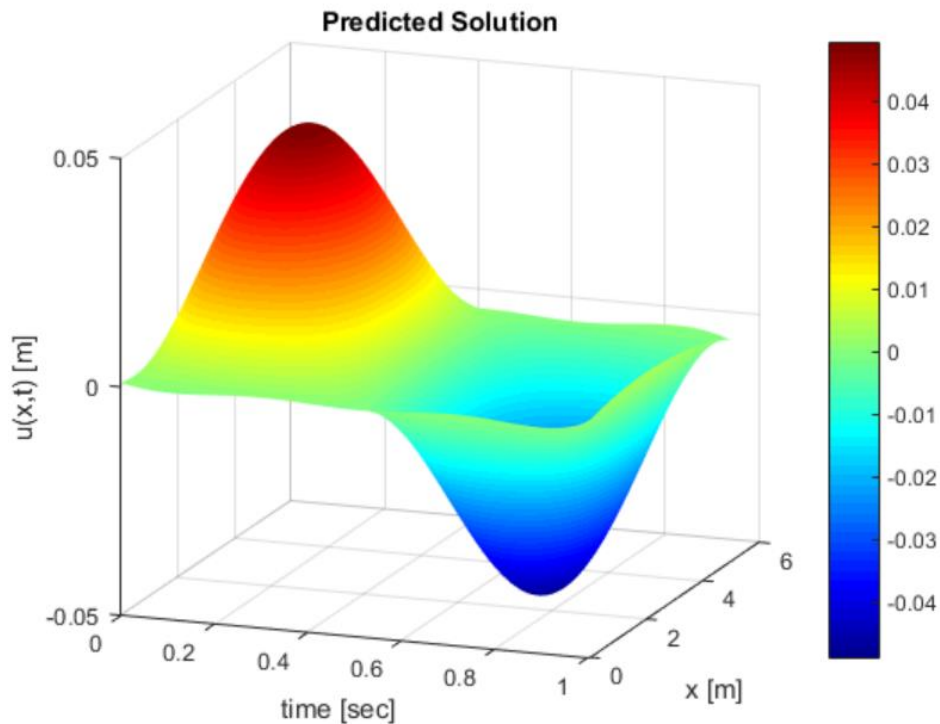
```
self.loss = tf.reduce_mean(tf.square(self.u_tf - self.u_pred)) + \
            tf.reduce_mean(tf.square(self.u_xx_tf - self.u_xx_pred)) + \
            0.001*tf.reduce_mean(tf.square(self.f_pred))
```

Ο αλγόριθμος βελτιστοποίησης που χρησιμοποιείται είναι ο L-BFGS-B καθώς το δίκτυο συγκλίνει γρήγορα στις βέλτιστες παραμέτρους που περιγράφουν το πρόβλημα. Επισημαίνεται ότι θεωρητικά δεν αποδεικνύεται ότι η μέθοδος αυτή οδηγεί σε ολικό ελάχιστο του σφάλματος. Ωστόσο, μια πλήρως ορισμένη ΜΔΕ έχει μοναδική λύση και επομένως χρησιμοποιώντας δίκτυο κατάλληλης αρχιτεκτονικής με επαρκή αριθμό δεδομένων για την εκπαίδευση της f , τότε η πρόβλεψη του δικτύου θα είναι ικανοποιητική.

Το σφάλμα του δικτύου υπολογίστηκε $\text{error} = 3.77 \cdot 10^{-2} \%$ με τις αρχικές συνθήκες ($t=0$) να είναι μηδενικές σε όλη τη δοκό και μηδενικές μετατοπίσεις στα δύο άκρα της δοκού καθόλο το διάστημα. Η

μέγιστη και η ελάχιστη μετατόπιση συμβαίνουν σε χρόνους $t_{min} = 0.24 \text{ s} \approx \frac{T}{4}$ και $t_{max} = 0.75 \approx \frac{3T}{4}$ αντίστοιχα και ισούνται με $u_{min} = -0.0492\text{m}$ και $u_{max} = 0.0488\text{m}$

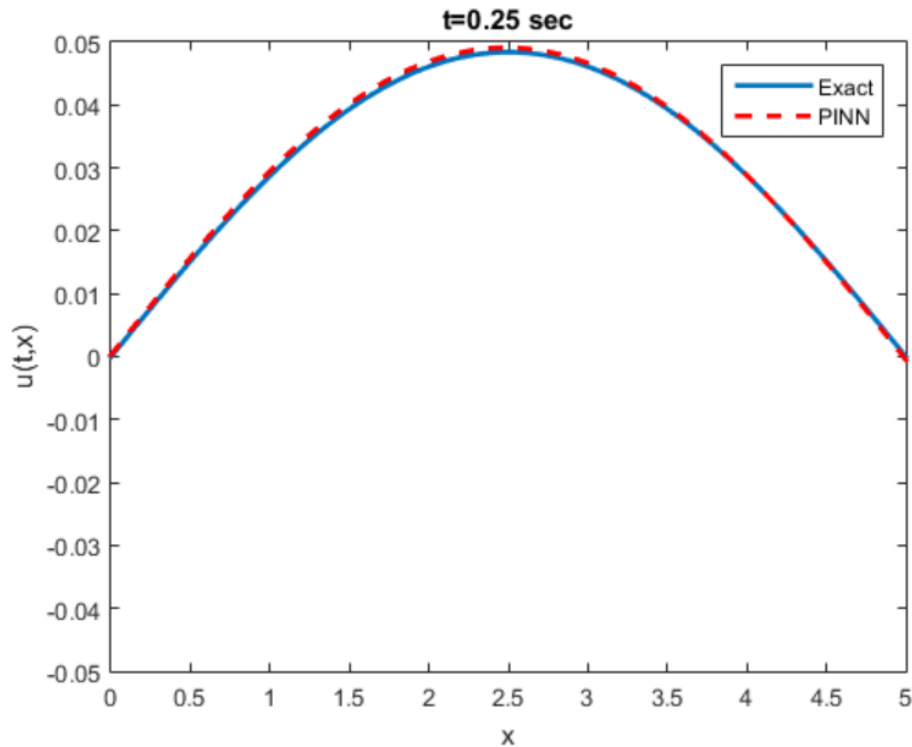
Στη συνέχεια στην Εικόνα 7–3 απεικονίζεται η τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου, όπως προέκυψε από το ΤΝΔ.



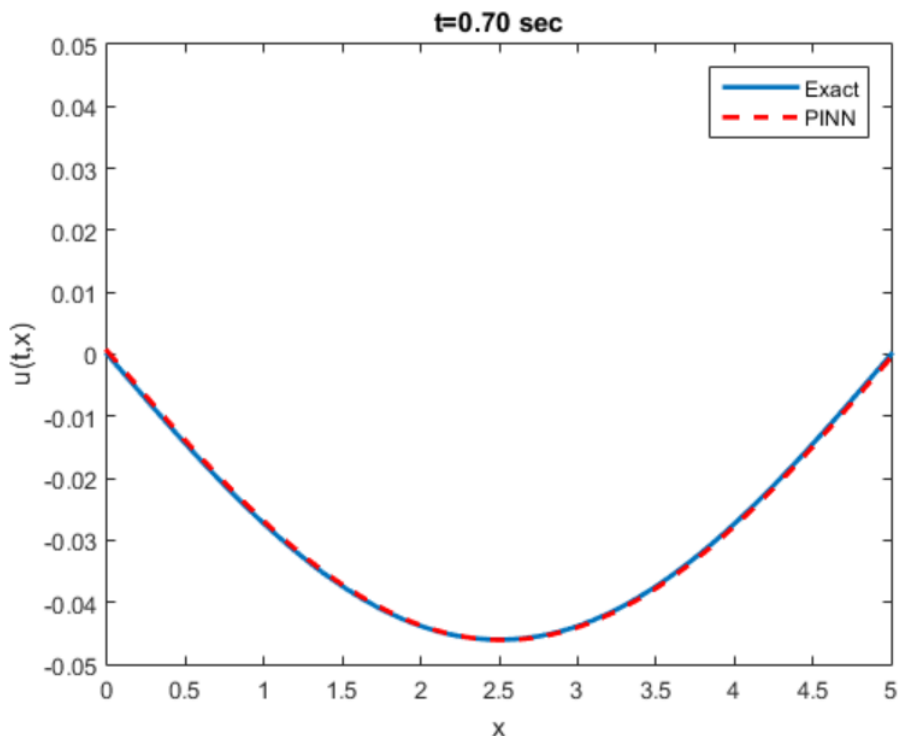
Εικόνα 7–3: Τρισδιάστατη απεικόνιση λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$

Στην Εικόνα 7–4, Εικόνα 7–5 και Εικόνα 7–6 απεικονίζονται η γραφική παράσταση της αναλυτικής λύσης της παραμόρφωσης της δοκού και η πρόβλεψη του PINN για τις χρονικές στιγμές $t=0.25 \text{ sec}$, $t=0.70\text{sec}$ και $t=1.00 \text{ sec}$, ενώ στην Εικόνα 7–7, στην Εικόνα 7–8, στην Εικόνα 7–9 και στην Εικόνα 7–10 η πρόβλεψη του PINN για τις θέσεις $x=0$, $x=1.52$, $x=3$ και $x=5\text{m}$. Όπως παρατηρείται, το δίκτυο έχοντας τροφοδοτηθεί με τις αρχικές και συνοριακές συνθήκες καθώς επίσης και με μικρό ποσοστό (10%) των εσωτερικών σημείων του χωροχρονικού πεδίου που ορίστηκε, προσεγγίζει με αρκετή ακρίβεια την αναλυτική λύση εντός του χωροχρονικού πεδίου στο οποίο έχει οριστεί με τη μη γραμμικότητα της ΜΔΕ που περιγράφει το φαινόμενο της διάδοσης κύματος σε αμφιέρειστη δοκό έχει προβλεφθεί αρκετά επιτυχώς από το δίκτυο. Όπως συμπεραίνεται όμως από την Εικόνα 7–10 το PINN φαίνεται να μην έχει «αντιληφθεί» με ακρίβεια τις συνοριακές συνθήκες στα στηριγμένα άκρα της

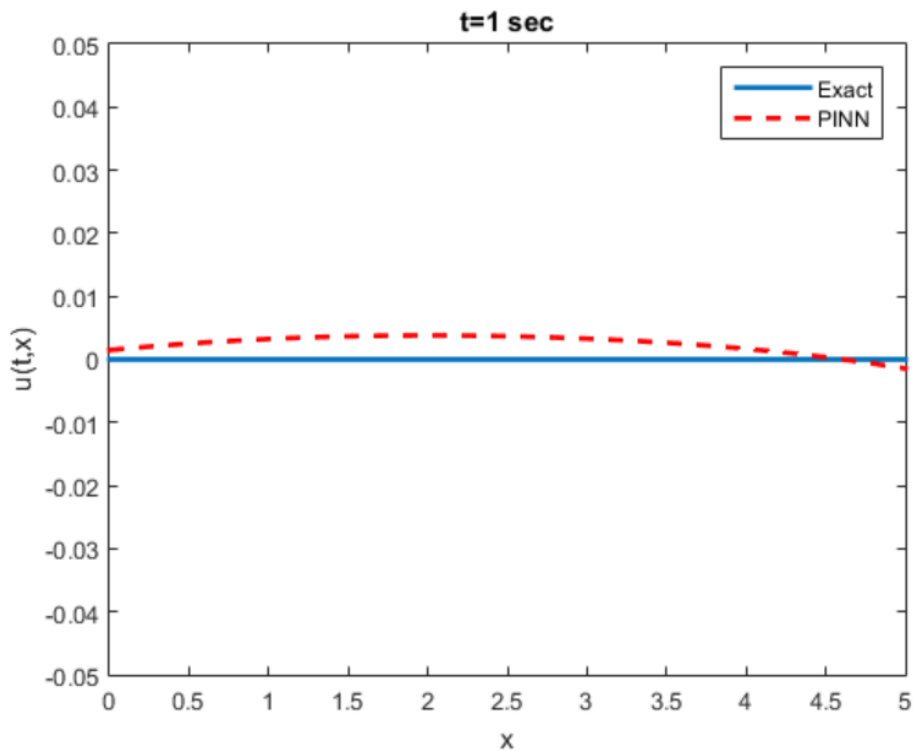
δοκού, και ιδιαίτερα για $x=5\text{m}$. Για την αντιμετώπιση αυτού του προβλήματος δοκιμάστηκε η εκπαίδευση του PINN με διαφορετικούς συντελεστές βαρύτητας σε κάθε όρο της συνάρτησης σφάλματος και ίδιο το πλήθος των σημείων των αρχικών και συνοριακών συνθηκών καθώς επίσης και των εσωτερικών σημείων.



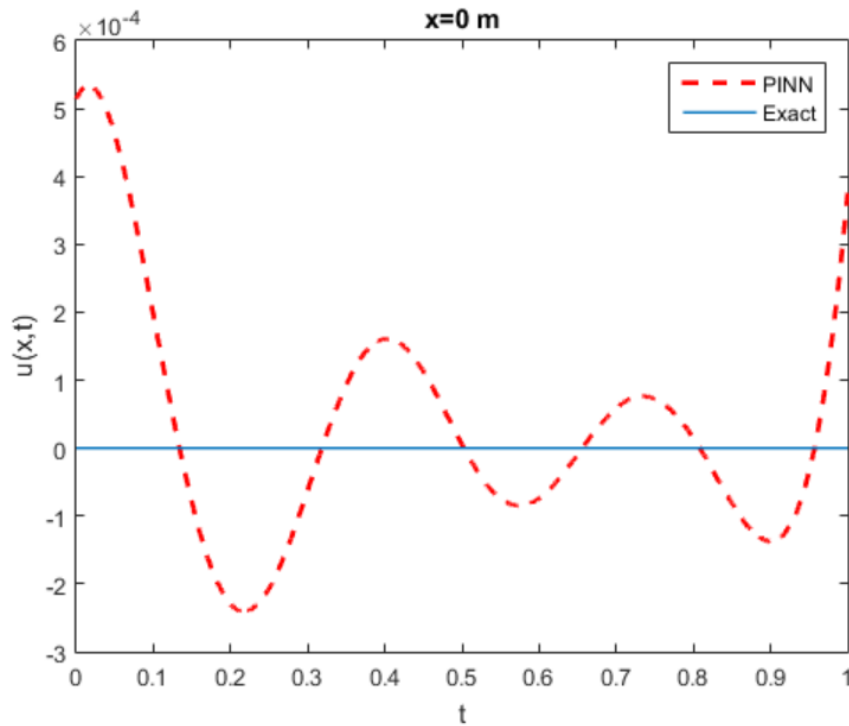
Εικόνα 7-4: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.25\text{sec}$



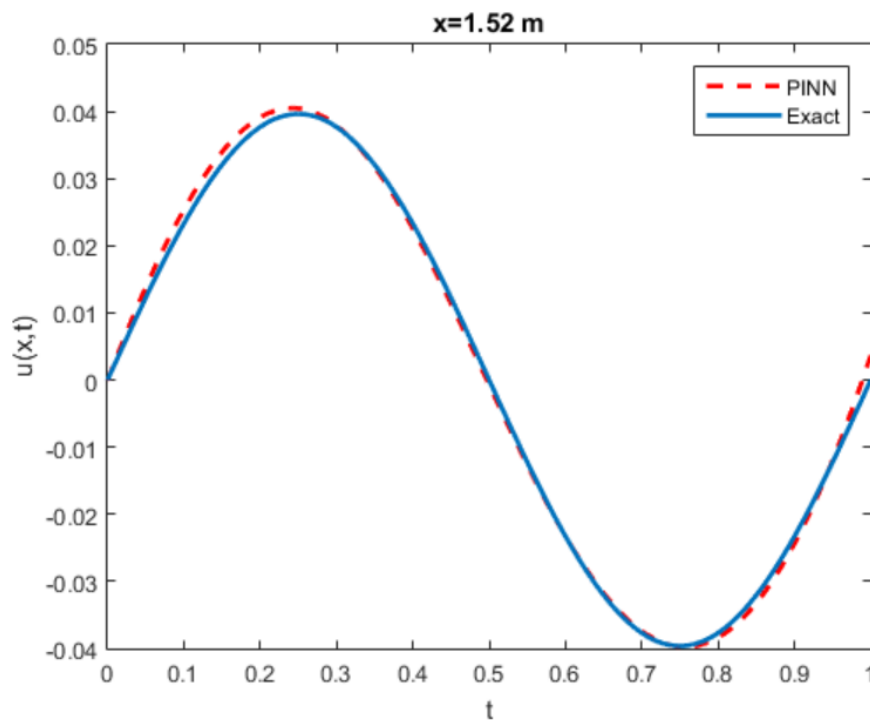
Εικόνα 7-5: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.7\text{sec}$



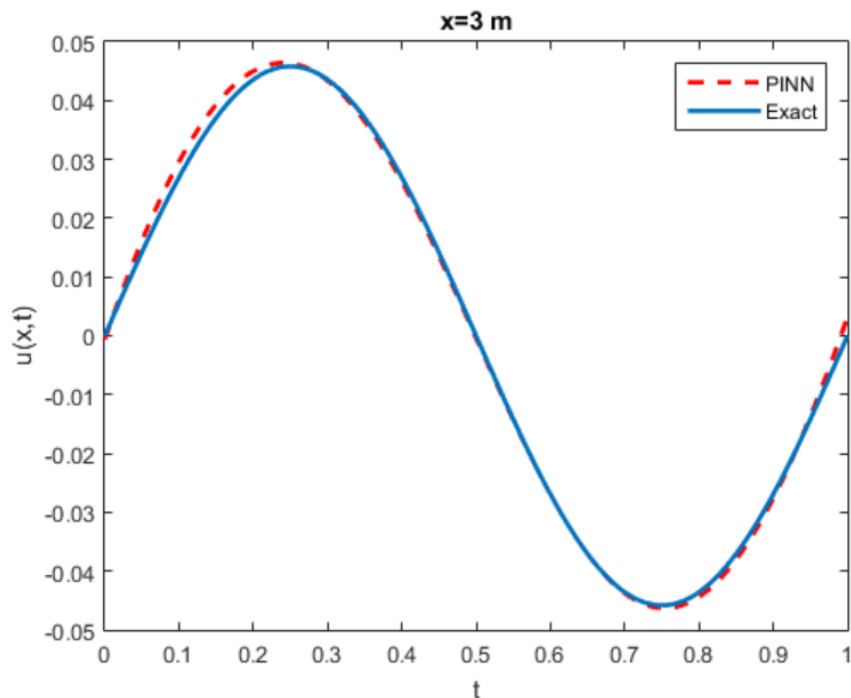
Εικόνα 7-6: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=1\text{sec}$



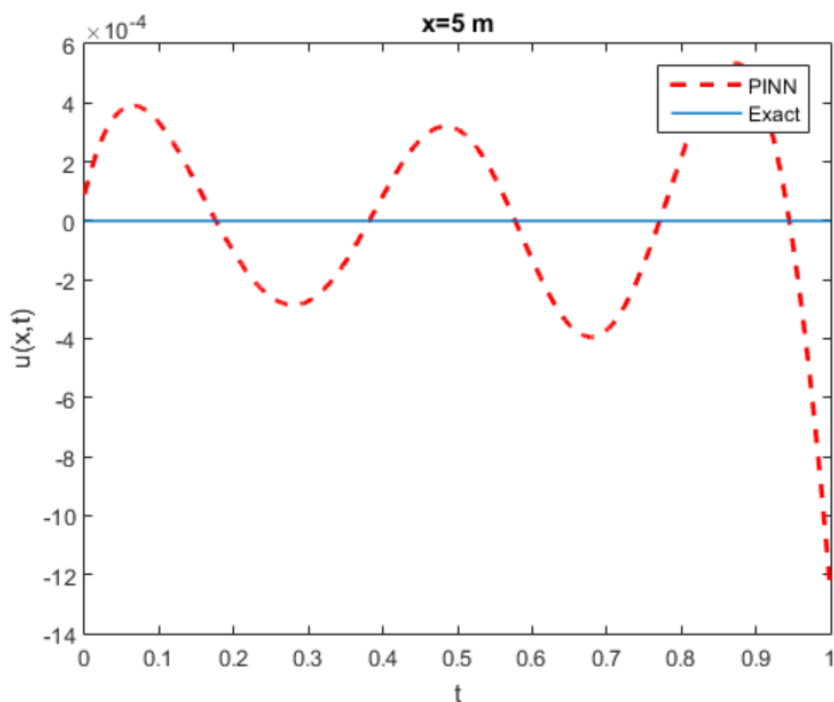
Εικόνα 7-7: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=0$ m



Εικόνα 7-8: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=1.52$ m



Εικόνα 7–9: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=3$ m



Εικόνα 7–10: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=5$ m

Ενδεικτικά παρουσιάζεται μία προσπάθεια επίλυσης του δικτύου ώστε να προβλέπει με μεγαλύτερη ακρίβεια τις μετατοπίσεις στο άκρο $x=5\text{m}$ της δοκού, με τη συνάρτηση σφάλματος να είναι της μορφής:

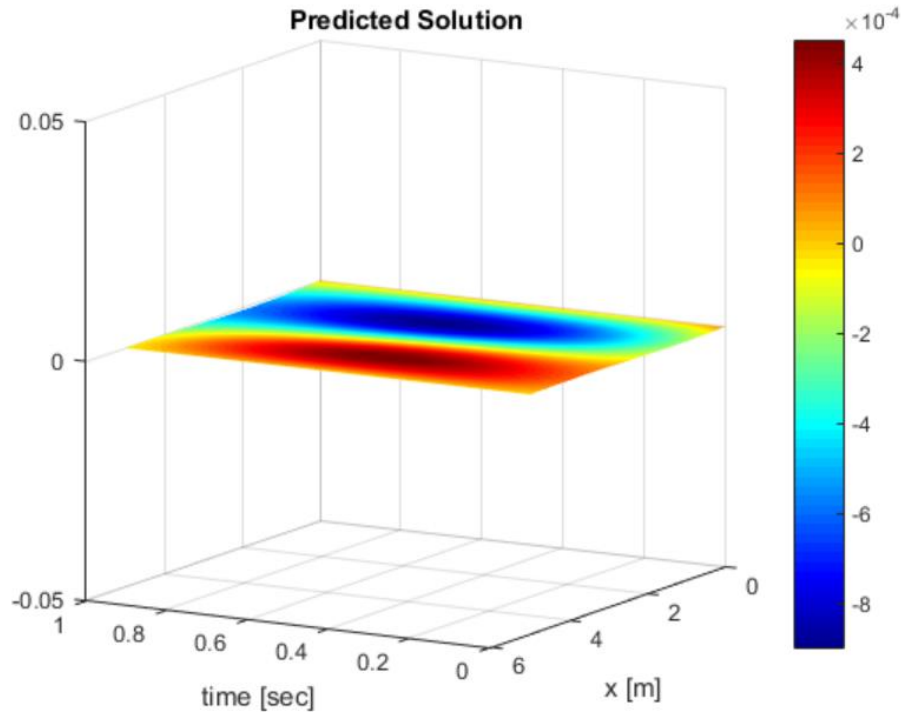
```
self.loss = 100000*tf.reduce_mean(tf.square(self.u_tf - self.u_pred)) + \
            100000*tf.reduce_mean(tf.square(self.u_xx_tf )) + \
            0.001*tf.reduce_mean(tf.square(self.f_pred))
```

Πιο συγκεκριμένα στην συγκεκριμένη περίπτωση στη συνάρτηση του σφάλματος εκπαίδευσης του δικτύου, δόθηκε 10^5 φορές μεγαλύτερη έμφαση στον όρο των μετατοπίσεων καθώς και στον όρο της καμπυλότητας των συνοριακών σημείων. Παράλληλα, η βαρύτητα του όρου του σφάλματος της συνάρτησης f μειώθηκε 10^{-3} φορές.

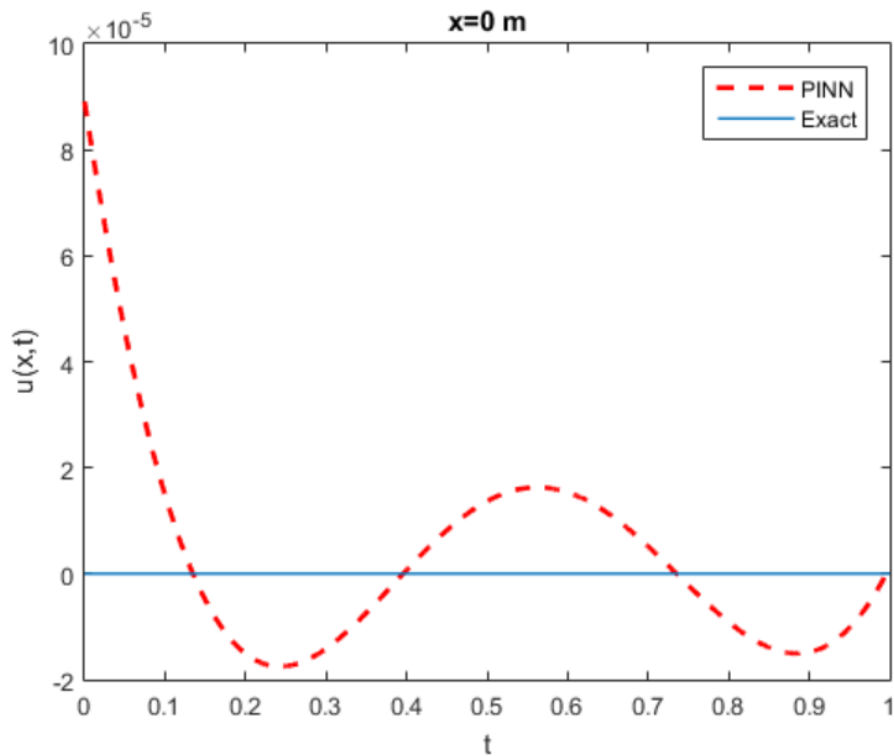
Στην Εικόνα 7–11 απεικονίζεται η τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου, όπως προέκυψε από το PINN με τη συνάρτηση σφάλματος όπως ορίστηκε προηγουμένως. Το σφάλμα του δικτύου υπολογίστηκε $\text{error} = 0.998\%$ έναντι $\text{error} = 3.77 \cdot 10^{-2}\%$ που προέκυψε στην αρχική περίπτωση.

Παρακάτω στην Εικόνα 7–13, στην Εικόνα 7–14, στην Εικόνα 7–15 και στην Εικόνα 7–15 απεικονίζονται η γραφική παράσταση της αναλυτικής λύσης της παραμόρφωσης της δοκού και η πρόβλεψη του PINN στις θέσεις $x=0$, $x=1.46$, $x=4$, $x=5\text{m}$. Όπως διαπιστώνεται παρότι σε αυτήν την περίπτωση το PINN δίνει καλή προσέγγιση της συμπεριφοράς της δοκού στα άκρα της, αδυνατεί να προβλέψει ικανοποιητικά τη συμπεριφορά της στο υπόλοιπο χωρίο.

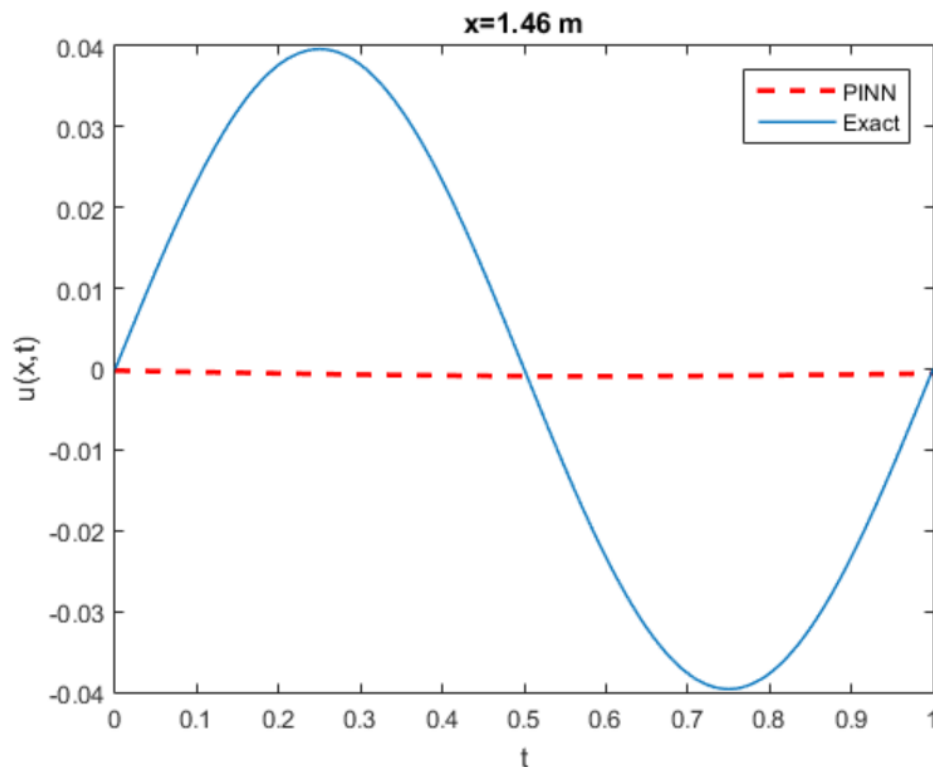
Όπως αναφέρθηκε και προηγουμένως η γενικότερη κατεύθυνση που συστήνεται από τους συγγραφείς των PINNs για τη βελτίωση της πρόβλεψης είναι αύξηση των δεδομένων εισόδου Nu και Nf , και / ή αύξηση των κρυμμένων επιπέδων και νευρώνων. Όπως φαίνεται και από τις εικόνες Εικόνα 7–16 και Εικόνα 7–17 τα συμπεράσματα αυτά επιβεβαιώνονται μόνο στην Εικόνα 7–16 όπου για πλήθος εσωτερικών σημείων πάνω από 5000 το σφάλμα πρόβλεψης μικραίνει ανεξαρτήτως του πλήθους των Nu , σε δίκτυο με 8 κρυφά επίπεδα των 20 νευρώνων το καθένα. Αντίθετα στην, Εικόνα 7–17, όπου $Nu=2100$ και $Nf=10000$ φαίνεται να υπάρχει το πρόβλημα της υπερπροσαρμογής (overfitting). Ειδικότερα, για δίκτυο με 12 κρυφά επίπεδα των 60 νευρώνων το σφάλμα είναι μεγαλύτερο από δίκτυο με λιγότερα από 6 κρυφά επίπεδα.



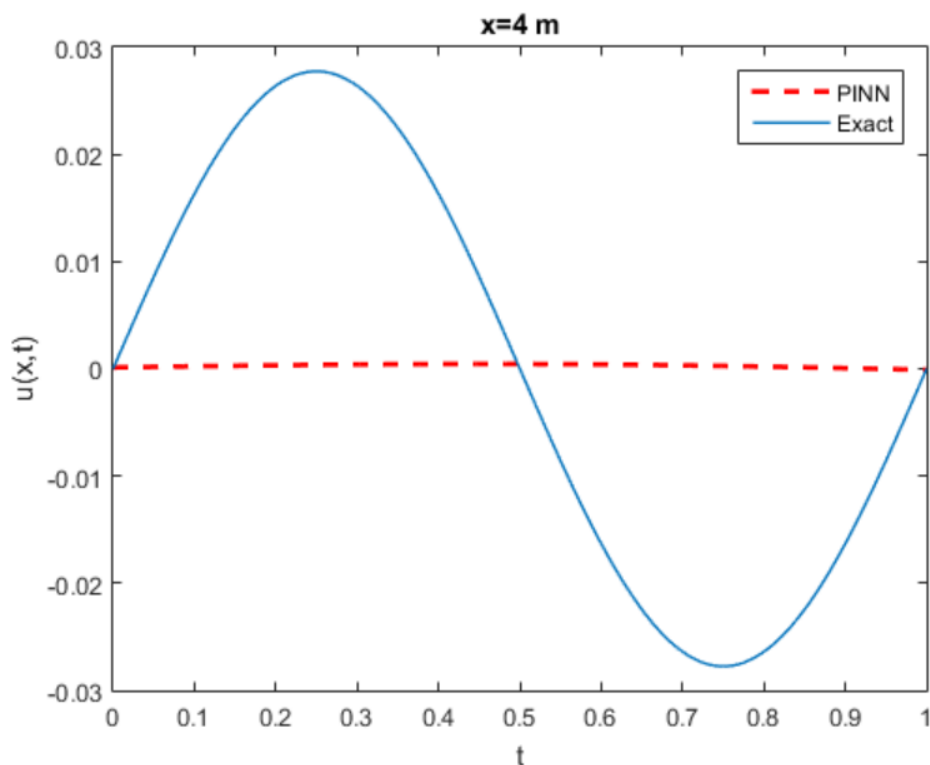
Εικόνα 7–11: Τρισδιάστατη απεικόνιση της λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ με διαφορετική εξίσωση σφάλματος.



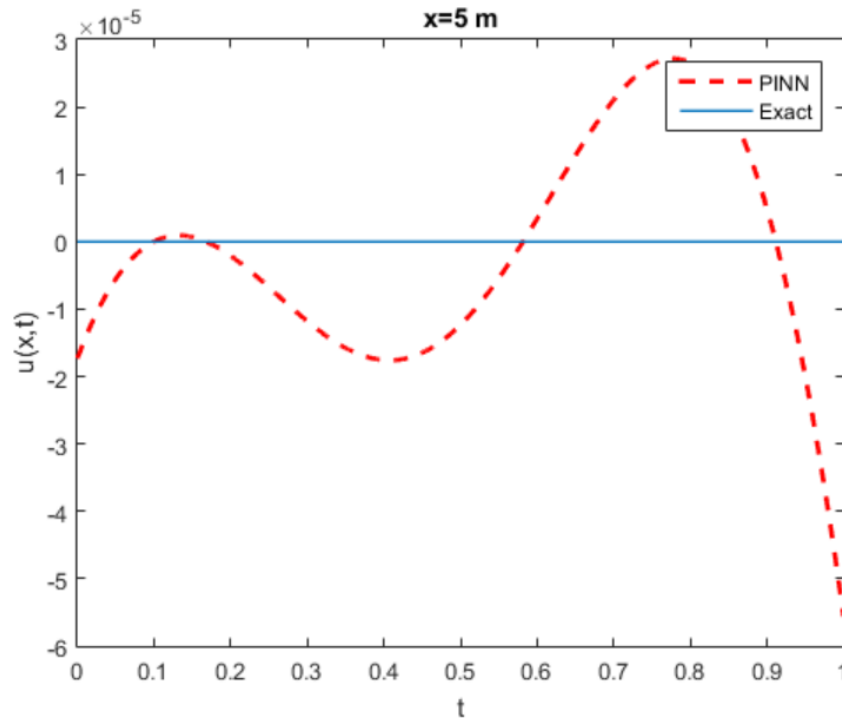
Εικόνα 7–12: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t = \sin(2\pi t)$ για $x=0m$



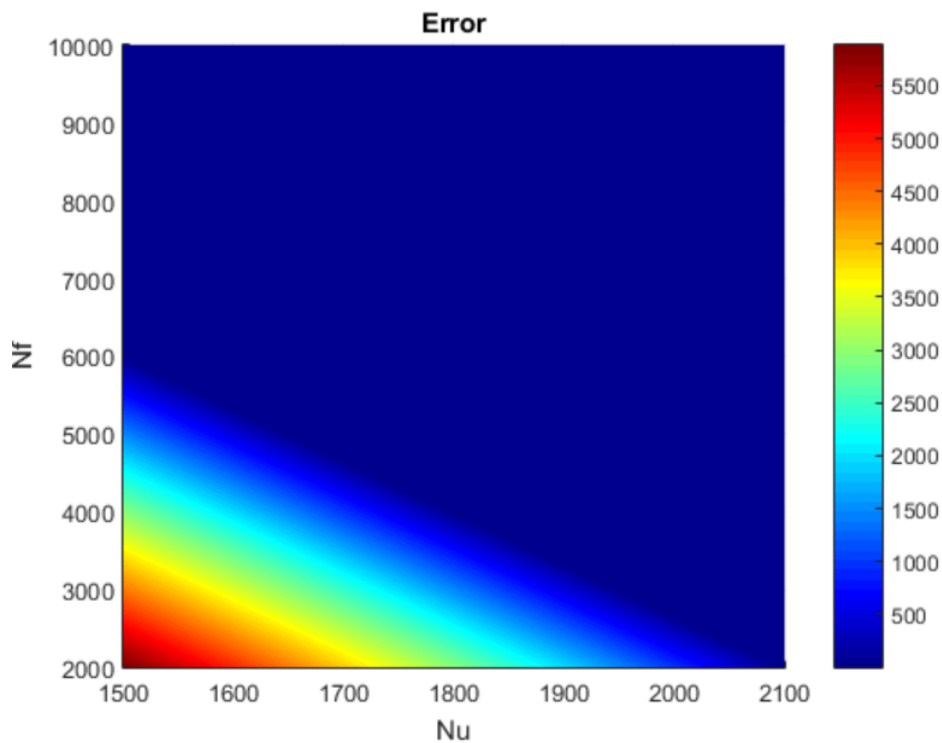
Εικόνα 7–13: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=1.46$ m



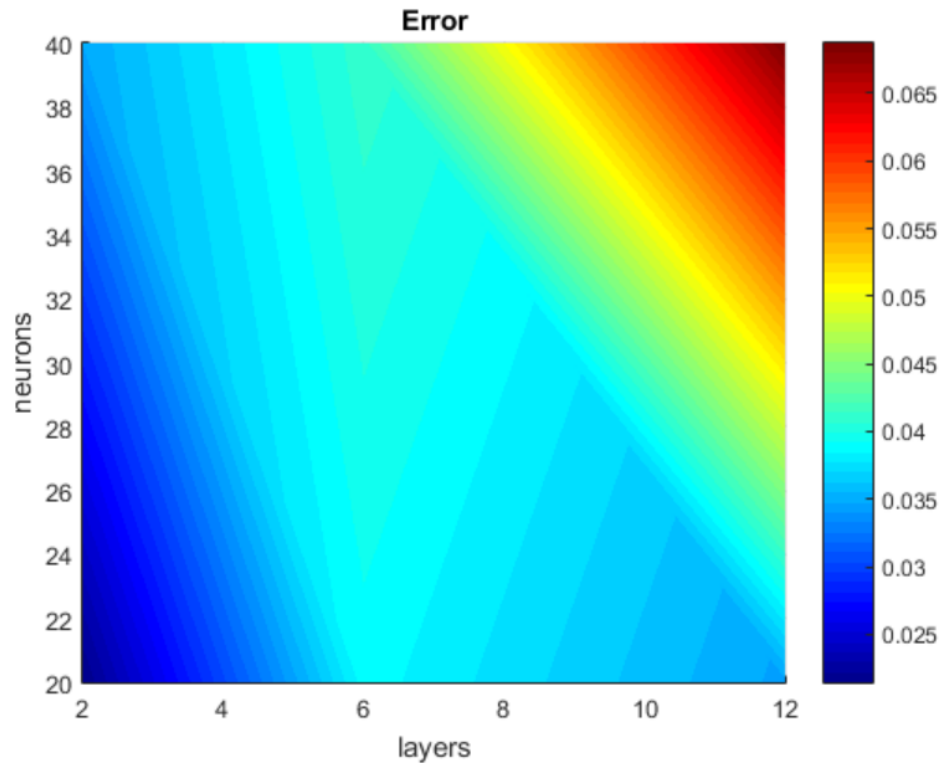
Εικόνα 7–14: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=4$ m



Εικόνα 7–15: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=5$ m



Εικόνα 7–16: Σύγκριση του σφάλματος μεταξύ αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για διαφορετικά πλήθη συνοριακών σημείων και σημείων αρχικών συνθηκών. Οι νευρώνες και τα κρυφά επίπεδα είναι σταθερά και ίσα με 8 και 20, αντίστοιχα.



Εικόνα 7–17: Σύγκριση του σφάλματος μεταξύ αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli με διαφοροποιημένη εξίσωση σφάλματος για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για διαφορετικές τιμές νευρώνων και κρυφών επιπέδων. Τα συνοριακά σημεία και σημεία αρχικών συνθηκών είναι σταθερά και ίσα με $Nu = 2100$, και τα εσωτερικά σημεία $Nf=10000$.

7.1.2 Εφαρμογή 2: Εξίσωση ταλάντωσης δοκού- πρόβολου

Όμοια με την περίπτωση της αμφιέρειστης δοκού, για την περίπτωση της δοκού-προβόλου αρχικά πραγματοποιήθηκε η αναλυτική λύση της ταλάντωσης της δοκού με χρήσης της γλώσσας προγραμματισμού matlab.

Η εφαρμογή έγινε για περίπτωση χαλύβδινης δοκού με τα παρακάτω χαρακτηριστικά:

- Μήκος αμφιέρειστης δοκού $L = 5$
- Μέτρο ελαστικότητας δοκού $E = 210 * 10^6 \text{ KPa}$
- Διατομή δοκού $IPE 80$
- Εμβαδόν διατομής δοκού $A = 7.64 * 10^{-4} \text{ m}^2$
- Ροπή αδράνειας διατομής $I = 80.1 * 10^{-8} \text{ m}^4$
- Πυκνότητα διατομής $\rho = 7500 \text{ kg/m}^3$

Η επίλυση πραγματοποιήθηκε για χρόνο $0 \leq t \leq 1 \text{ sec}$ με χρονική και χωρική διακριτοποίηση, $\Delta t = 0.0099 \text{ sec}$ και $\Delta x = 0.00005 \text{ m}$ αντίστοιχα. Εξετάστηκε η περίπτωση μηδενικού επιβαλλόμενου

φορτίου με μοναδιαία αρχική μετακίνηση του ελεύθερου άκρου της δοκού στη θέση $x = L$, $u(L, 0) = 1$. Η εξίσωση που περιγράφει τη συμπεριφορά της δοκού προβόλου είναι είναι η (5-33), με τη λύση του $Y_n(t)$ να δίνεται από την (5-35). Εξετάστηκε η περίπτωση μηδενικής αρχικής ταχύτητας $Y_n(\dot{0})$ για κάθε ιδιομορφή ταλάντωσης και στη θέση $x=L$ μοναδιαία αρχική μετακίνηση. Οι συνοριακές συνθήκες που εφαρμόζονται στη δοκό πρόβολο είναι αυτές που περιγράφονται από τις σχέσεις (5-27). Επομένως, με τη θεώρηση ότι η εξίσωση που περιγράφει τη μετακίνηση της δοκού για $t=0$ συναρτήσει της θέσης είναι ένα κυβικό πολυώνυμο της μορφής:

$$u(x, 0) = ax^3 + bx^2 + cx + d$$

Όπου a, b, c, d σταθερές των οποίων οι τιμές λαμβάνονται αντικαθιστώντας κατάλληλα με τις συνοριακές συνθήκες που ισχύουν στην δοκό πρόβολο, προκύπτει η εξίσωση $f(x)$ της σχέσης (5-36), για την οποία ισχύει $f(x) = u(x, 0)$.

Επειτα από τις σχέσεις(5-24), (5-29), (5-30) με αυθαίρετη σταθερά $C_1 = 0$ υπολογίζονται τα $\omega_n, \varphi_n, \beta_n$ για κάθε ιδιομορφή ταλάντωσης και από τις σχέσεις (5-34) λαμβάνονται οι τιμές των M_n, K_n, P_n .

Με τις κατάλληλες αντικαταστάσεις των $Y_n(\dot{0}) = 0, P_n = 0$ η σχέση (5-35) λαμβάνει τη μορφή:

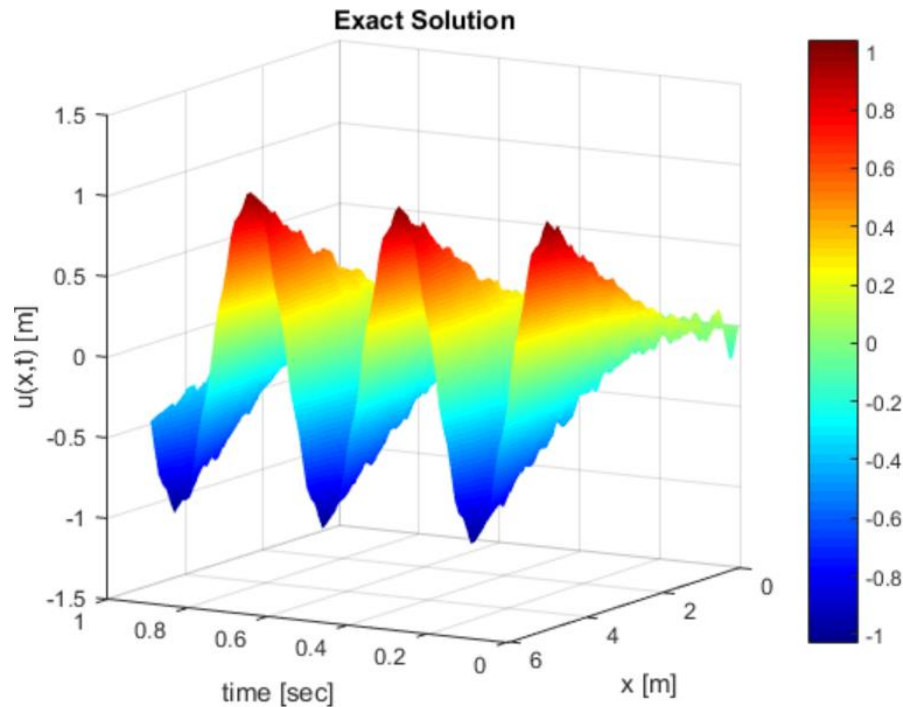
$$Y_n(t) = Y_n(0)\cos\omega_n t$$

Τελικά από τη σχέση (5-32) λαμβάνεται η αναλυτική επίλυση της ταλάντωσης ελαστικής δοκού προβόλου για μοναδιαία αρχική μετακίνηση στο ελεύθερο άκρο της.

Από την αναλυτική επίλυση (Exact solution) προέκυψε η παρακάτω τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου Εικόνα 7–18 όπου είναι εμφανές ότι $u(0,t)=0$ και ότι το ελεύθερο άκρο της δοκού ταλαντώνεται μεταξύ των τιμών ± 1 .

Οι τιμές των κάθετων μετατοπίσεων που προκύπτουν από την αναλυτική επίλυση, καθώς και οι καμπυλότυπες των σημείων της δοκού (δηλαδή οι τιμές της 2^{ης} παραγώγου της κατακόρυφης μετατόπισης) χρησιμοποιούνται ως δεδομένα εκπαίδευσης στο PINN, αφού με αυτές τις τιμές μπορεί να καθοριστεί επαρκώς η συμπεριφορά της εκάστωτε ταλαντούμενης δοκού.

Το τεχνητό νευρωνικό δίκτυο, το οποίο χρησιμοποιήθηκε για την επίλυση της ΜΔΕ της ταλάντωσης δοκού προβόλου δημιουργήθηκε επίσης στο προγραμματιστικό περιβάλλον της Python, κάνοντας χρήση της βιβλιοθήκης Tensorflow.



Εικόνα 7–18: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$

Ορίζεται το PINN ως η συνάρτηση $f(t, x)$:

$$f := EI \frac{\partial^4 u}{\partial x^4} + m \frac{\partial^2 u}{\partial t^2} - p(x, t)$$

Όπου $p(x, t) = 0$

Το δίκτυο τροφοδοτείται και εκπαιδεύεται με τα δεδομένα εισόδου x_i, t_i, u_i όπως αυτά έχουν προκύψει από την αναλυτική επίλυση του προβλήματος.

Έχει συμπεριληφθεί το κομμάτι του κώδικα στο οποίο ορίζεται το PINN, δηλαδή η υπό μελέτη ΜΔΕ:

```
def net_f(self, x,t):
    u = self.net_u(x,t)
    u_t = tf.gradients(u, t)[0]
    u_tt = tf.gradients(u_t, t)[0]
    u_x = tf.gradients(u, x)[0]
    u_xx = tf.gradients(u_x, x)[0]
    u_xxx = tf.gradients(u_xx, x)[0]
    u_xxxx = tf.gradients(u_xxx, x)[0]

    f = self.nu2*u_tt + self.nu1*u_xxxx
    return f
```

Για την εκπαίδευση του ΤΝΔ επιλέχθηκαν οι συνοριακές συνθήκες της δοκού. Συγκεκριμένα, ελήφθησαν όλο το πλήθος των σημείων της δοκού με τις μετατοπίσεις τους για $t = 0$, $uu1$, καθώς και των μετατοπίσεων του πακτωμένου άκρου $uu2$ και των μετατοπίσεων του ελεύθερου άκρου της δοκού $uu3$ για το χρονικό διάστημα $0 \leq t \leq 1 \text{ sec}$. Παράλληλα, για το πακτωμένο άκρο λήφθηκε η συνθήκη της μηδενικής στροφής ($\frac{du}{dx} = 0$) για τις μετατοπίσεις που εξετάζονται, $duu2dx$, ενώ για τις εξεταζόμενες μετατοπίσεις του ελεύθερου άκρου εφαρμόστηκαν οι συνθήκες της μηδενικής καμπυλότητας ($\frac{d^2u}{dx^2} = 0$), $d2uu3dx2$ και τέμνουσας ($\frac{d^3u}{dx^3} = 0$), $d3uu3dx3$.

Για το PINN επιλέχθηκε πλήθος 10000 τυχαίων κατανεμημένων σημείων του πεδίου ορισμού της εξίσωσης, $N_f = 10000$ από σύνολο 1000000 σημείων (το 1% των συνολικών σημείων). Η συνάρτηση σφάλματος είναι της μορφής:

```
self.loss = tf.reduce_mean(tf.square(self.uu1_tf - self.uu1_pred)) + \
            tf.reduce_mean(tf.square(self.uu2_tf - self.uu2_pred)) + \
            tf.reduce_mean(tf.square(self.uu3_tf - self.uu3_pred)) + \
            tf.reduce_mean(tf.square(self.duu2dx_tf - self.duu2dx_pred)) + \
            tf.reduce_mean(tf.square(self.d2uu3dx2_tf - self.d2uu3dx2_pred)) + \
            tf.reduce_mean(tf.square(self.d3uu3dx3_tf - self.d3uu3dx3_pred)) + \
            tf.reduce_mean(tf.square(self.f_pred))
```

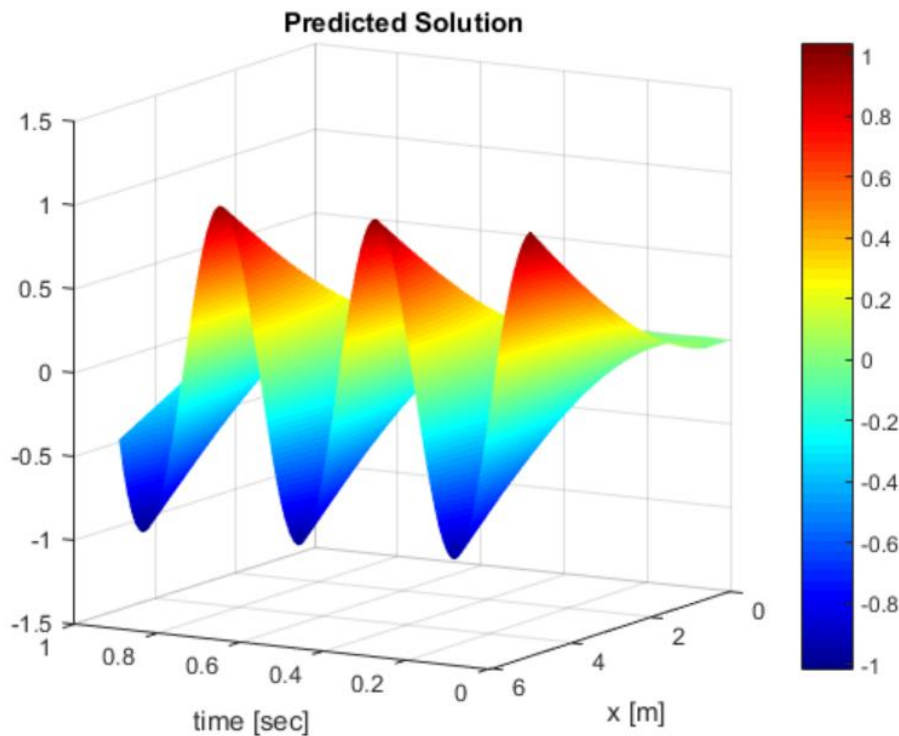
Η αρχιτεκτονική του ΤΝΔ αποτελείται από 8 επίπεδα νευρώνων, με κάθε κρυφό επίπεδο να περιέχει 20 νευρώνες και με συνάρτηση ενεργοποίησης την υπερβολική εφαπτομένη \tanh .

Η εκπαίδευση έγινε χρησιμοποιώντας την συνάρτηση του μέσου τετραγωνικού σφάλματος. Συγκεκριμένα, υπολογίζεται από το δίκτυο το άθροισμα των σφαλμάτων των μετατοπίσεων u των συνοριακών σημείων, των σφαλμάτων των τιμών της καμπυλότητας των συνοριακών σημείων u_{xx} (οι οποίες είναι μηδενικές καθότι αμφιέρειστη δοκός) και του σφάλματος της συνάρτησης f , η οποία θα πρέπει να έχει μηδενική τιμή. Θεωρήθηκε ότι κάθε μία από τις προαναφερθείσες συνιστώσες της συνάρτησης σφάλματος δεν συνεισφέρει εξίσου στο τελικό αποτέλεσμα. Συγκεκριμένα, δόθηκε μεγαλύτερη βαρύτητα στο σφάλμα της μετατόπισης και την καμπυλότητας των συνοριακών σημείων της δοκού, προκειμένου το δίκτυο να εκπαιδευτεί κατάλληλα για τη συγκεκριμένη δοκό. Το σφάλμα της συνάρτησης θεωρείται ότι έχει μικρότερη βαρύτητα στο σφάλμα της εκπαίδευσης του δικτύου, σε ποσοστό τάξης 1 προς τη μάζα της δοκού (1/1000).

Όσον αφορά τον αλγόριθμο βελτιστοποίησης, χρησιμοποιήθηκε ο αλγόριθμος Adam για 30000 εποχές εκπαίδευσης και κατόπιν ο L-BFGS-B. Μολονότι θεωρητικά δεν υπάρχει απόδειξη ότι η μέθοδος αυτή οδηγεί σε ολικό ελάχιστο του σφάλματος, είναι εμπειρικά γνωστό ότι αν η δοθείσα ΜΔΕ είναι πλήρως ορισμένη, έχει μοναδική λύση, χρησιμοποιείται επαρκής αριθμός σημείων N_f του πεδίου ορισμού της εξίσωσης τα οποία χρησιμοποιούνται για την εκπαίδευση της f και το ΤΝΔ που χρησιμοποιείται έχει την κατάλληλη αρχιτεκτονική, μπορεί το δίκτυο να δώσει μια αρκετά ικανοποιητική πρόβλεψη.

Το σφάλμα του δικτύου υπολογίστηκε $error = 9.796 \cdot 10^{-2} \%$ με τις αρχικές συνθήκες ($t=0$) να είναι μηδενικές σε όλη τη δοκό και μηδενικές μετατοπίσεις στα δύο άκρα της δοκού όλο το διάστημα.

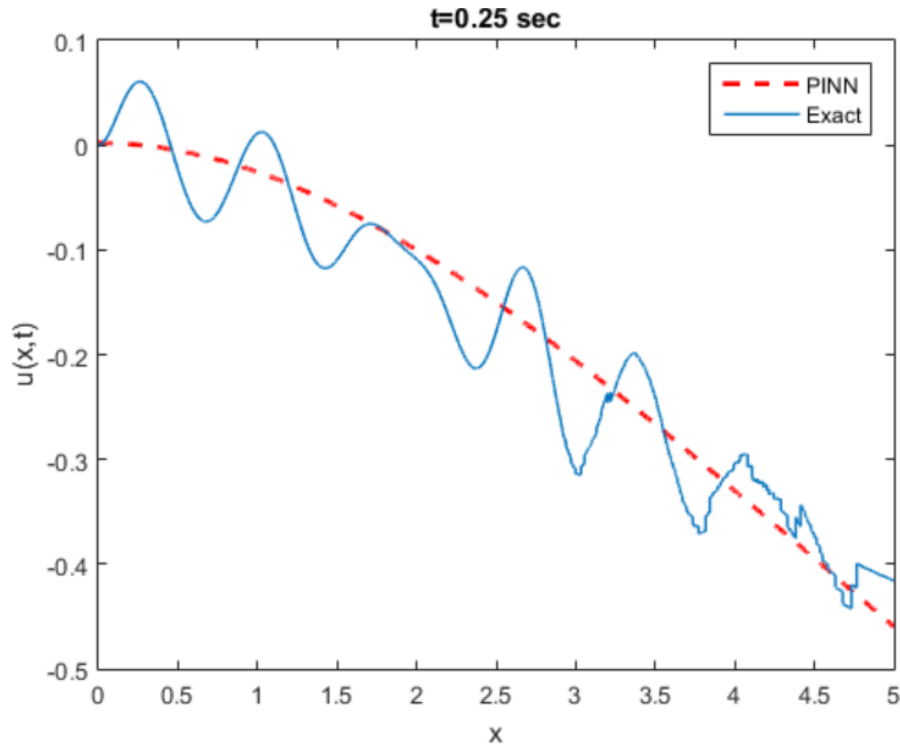
Στη συνέχεια στην Εικόνα 7–19 απεικονίζεται η τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου, όπως προέκυψε από το ΤΝΔ.



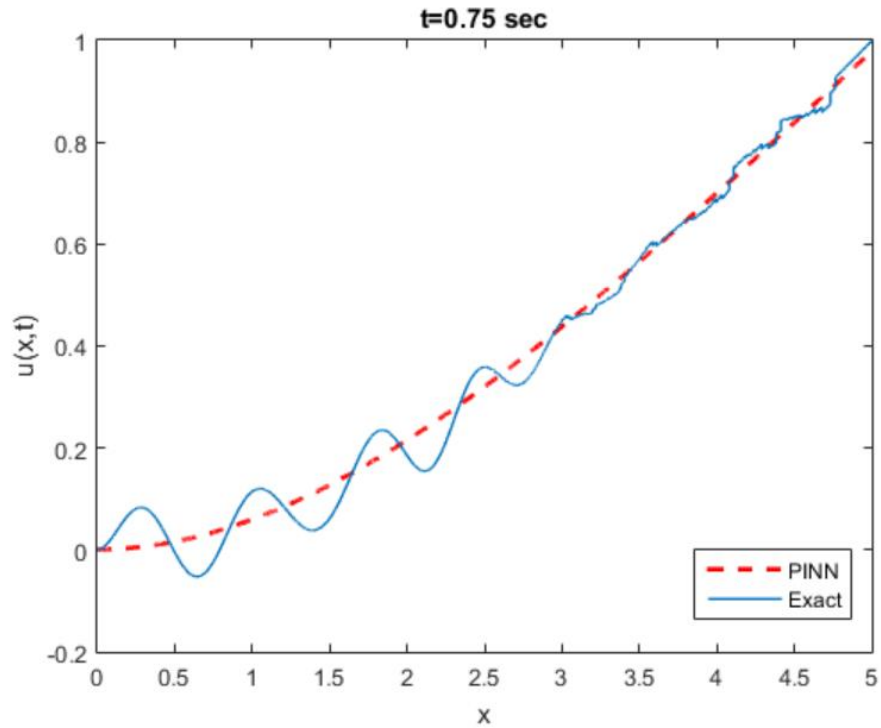
Εικόνα 7–19: Τρισδιάστατη απεικόνιση λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$

Τα γραφήματα Εικόνα 7–20, Εικόνα 7–21, Εικόνα 7–22 αποτυπώνουν και συγκρίνουν τη γραφική παράσταση της αναλυτικής λύσης της ΜΔΕ και της πρόβλεψης του PINN για τις χρονικές στιγμές $t = 0.25, 0.75, 0.99$ sec, αντίστοιχα. Επιπλέον στα γραφήματα Εικόνα 7–23, Εικόνα 7–24, Εικόνα 7–25 αποτυπώνονται και συγκρίνονται η γραφική παράσταση της αναλυτικής λύσης της ΜΔΕ και της

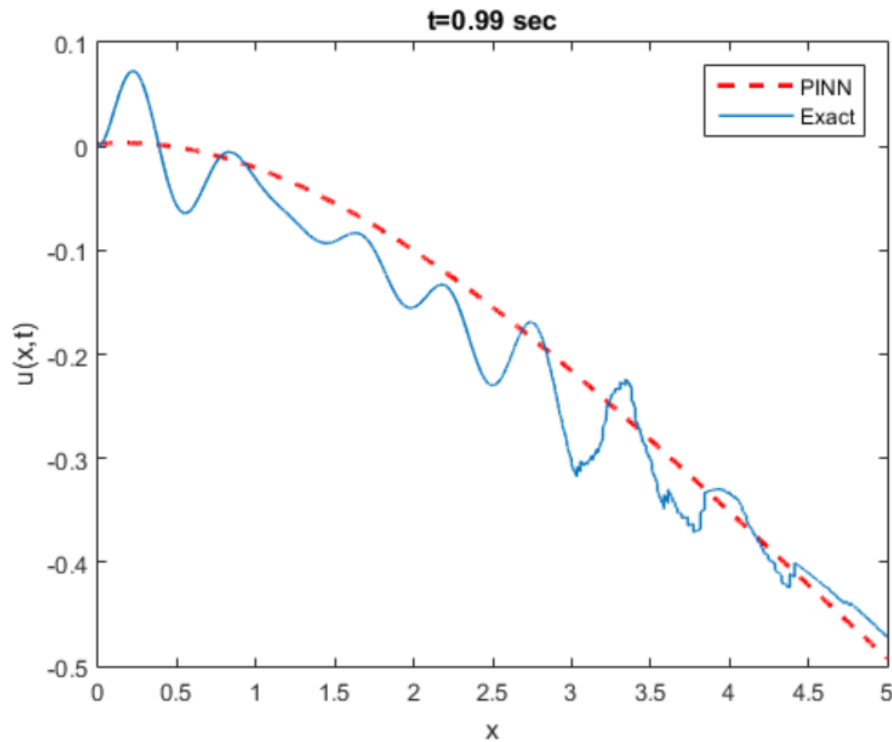
πρόβλεψης του δικτύου στις θέσεις $x=0,3,5,5$ m, αντίστοιχα. Από τα διαγράμματα αυτά συμπεραίνεται ότι το δίκτυο αξιοποιώντας ένα μέρος των αρχικών συνθηκών, των συνοριακών συνθηκών και των εσωτερικών σημείων προσέγγισε με αρκετά μεγάλη ακρίβεια την πολυπλοκότητα της εξίσωσης.



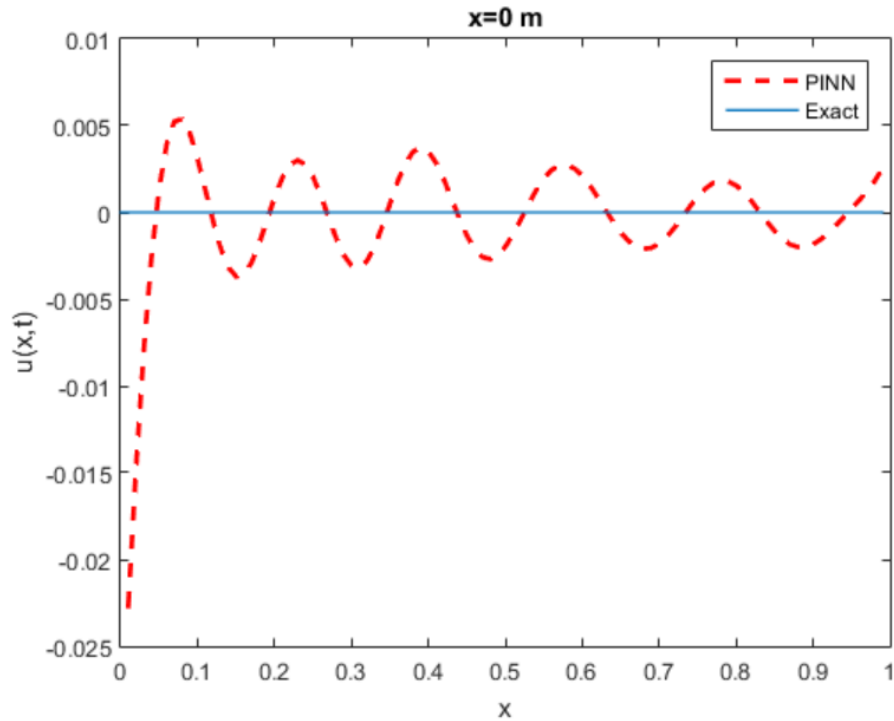
Εικόνα 7–20: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.25$ sec



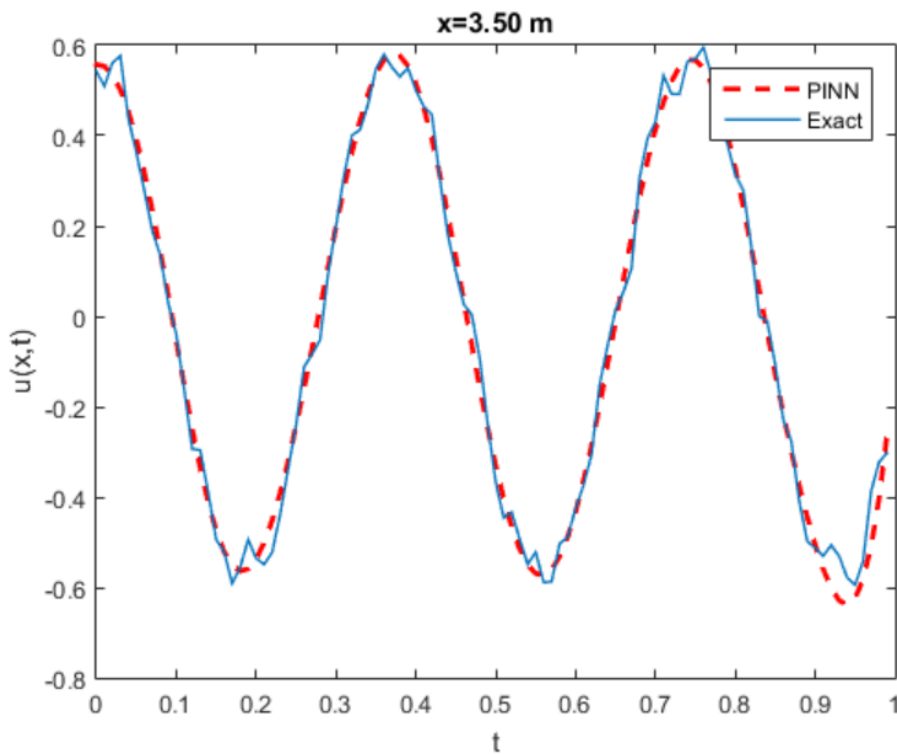
Εικόνα 7–21: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.75$ sec



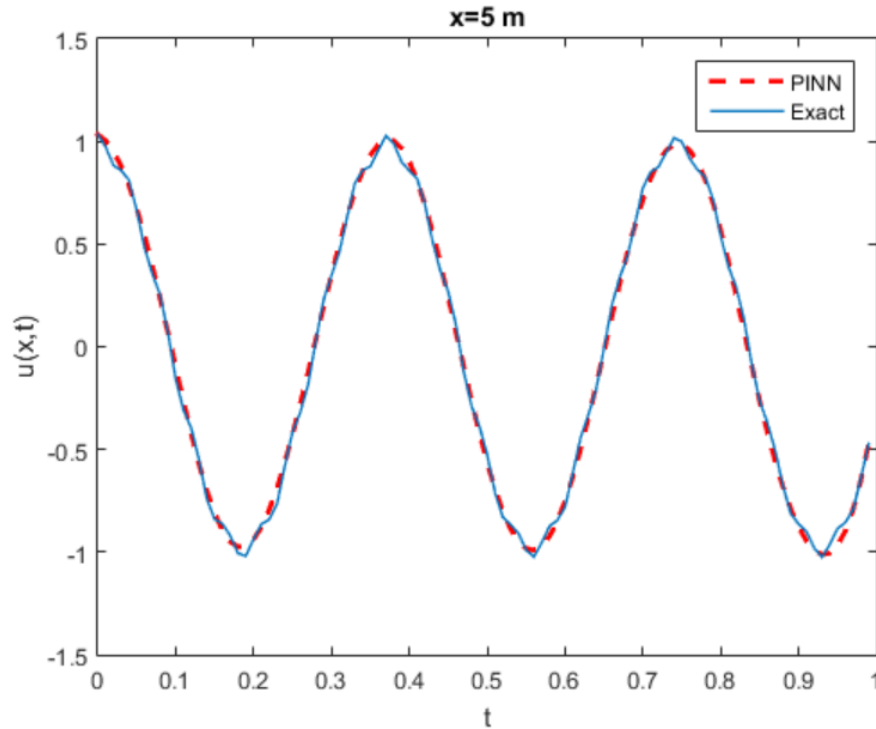
Εικόνα 7–22: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $t=0.99$ sec



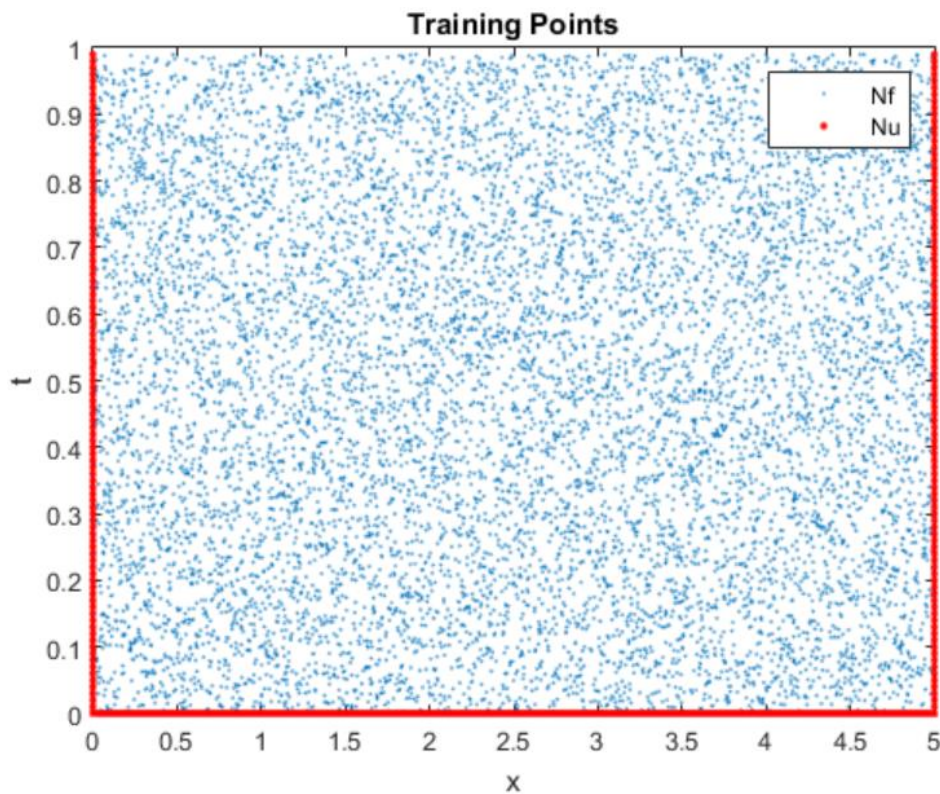
Εικόνα 7–23: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=0$ m



Εικόνα 7–24: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=3.50$ m



Εικόνα 7–25: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$ για $x=5$ m



Εικόνα 7–26 Σχηματική απεικόνιση των δεδομένων εκπαίδευσης του PINN για την επίλυση της εξίσωσης Euler-Bernoulli για δοκό πρόβολο υπό φορτίο $p(x, t) = \sin(2\pi t)$

7.1.3 Εφαρμογή 3: Εξίσωση ταλάντωσης αμφιέρειστης δοκού υπό σύνθεσης φορτίων

Όπως στην περίπτωση ταλάντωσης της αμφιέρειστης δοκού υπό ημιτονοειδές φορτίο, έτσι για την σύγκριση της μεθόδου των ΤΝΔ για την περίπτωση της ταλάντωσης αμφιέρειστης δοκού υπό σύνθεση τριγωνοειδών φορτίων, αρχικά πραγματοποιήθηκε επίλυση με αριθμητική μέθοδο. Η λύση της εξίσωσης (5.5) πραγματοποιήθηκε με αριθμητική ολοκλήρωση μέσω γραμμικών πεπερασμένων διαφορών, σύμφωνα με τις εξισώσεις ταλαντούμενης αμφιέρειστης δοκού. όπως περιγράφηκε στο προηγούμενο κεφάλαιο, στο προγραμματιστικό περιβάλλον matlab.

Η εφαρμογή έγινε για περίπτωση χαλύβδινης δοκού με τα παρακάτω χαρακτηριστικά:

- Μήκος αμφιέρειστης δοκού $L = 5$
- Μέτρο ελαστικότητας δοκού $E = 210 * 10^6 \text{ KPa}$
- Διατομή δοκού $IPE 80$
- Εμβαδόν διατομής δοκού $A = 7.64 * 10^{-4} \text{ m}^2$
- Ροπή αδράνειας διατομής $I = 80.1 * 10^{-8} \text{ m}^4$
- Πυκνότητα διατομής $\rho = 7500 \text{ kg/m}^3$

Επιλέχθηκε σύνθεση φορτίων της μορφής $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin(\pi t + \frac{\pi}{3})$ και η επίλυση πραγματοποιήθηκε για χρόνο $0 \leq t \leq 5 \text{ sec}$ για χρονική και χωρική διακριτοποίηση, $\Delta t = 0.01 \text{ sec}$ και $\Delta x = 0.005 \text{ m}$ αντίστοιχα. Η εξίσωση που περιγράφει τη συμπεριφορά της δοκού είναι η 8.3.59, η οποία προκύπτει από την επαλληλία όλων των ιδιομορφικών σευνεισφορών ταλάντωσης. Στη συνέχεια, όμοια με την περίπτωση ταλάντωσης της αμφιέρειστης δοκού υπό ημιτονοειδές φορτίο επιχειρείται σε αρχική φάση η επίλυση της 8.3.52, με τα A,B να είναι αυθαίρετες σταθερές που καθορίζονται από τις αρχικές συνθήκες. Στην προκειμένη περίπτωση εξετάστηκαν οι μηδενικές αρχικές συνθήκες της μετατόπισης και της ταχύτητας κάθε σημείου της δοκού σε κάθε ιδιομορφή, $Y_n(0) = 0$ και $\dot{Y}_n(0) = 0$.

Από την αναλυτική επίλυση (Exact solution) προέκυψε η τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου, Εικόνα 7–27. Είναι εμφανές ότι οι αρχικές συνθήκες ($t = 0$) είναι μηδενικές σε όλη τη δοκό και ότι στα δύο άκρα της δοκού οι μετατοπίσεις παραμένουν μηδενικές όλο το διάστημα. Οι υπόλοιπες κάθετες μετατοπίσεις της δοκού στο χωροχρονικό πεδίο που έχει επιλεχθεί προκύπτουν ως αποτέλεσμα σύνθεσης ταλαντώσεων. Οι τιμές αυτές όπως προκύπτουν από την αναλυτική επίλυση, καθώς και οι καμπυλότυπες των σημείων της δοκού (δηλαδή οι τιμές της

2^{ης} παραγώγου της κατακόρυφης μετατόπισης) χρησιμοποιούνται ως δεδομένα εκπαίδευσης στο PINN, αφού με αυτές τις τιμές μπορεί να καθοριστεί επαρκώς η συμπεριφορά της ταλαντούμενης δοκού.

Το ΤΝΔ το οποίο χρησιμοποιήθηκε για την επίλυση της ΜΔΕ της ταλάντωσης αμφιέρειστης δοκού δημιουργήθηκε σε προγραμματιστικό περιβάλλον Python με χρήση της βιβλιοθήκης Tensorflow.

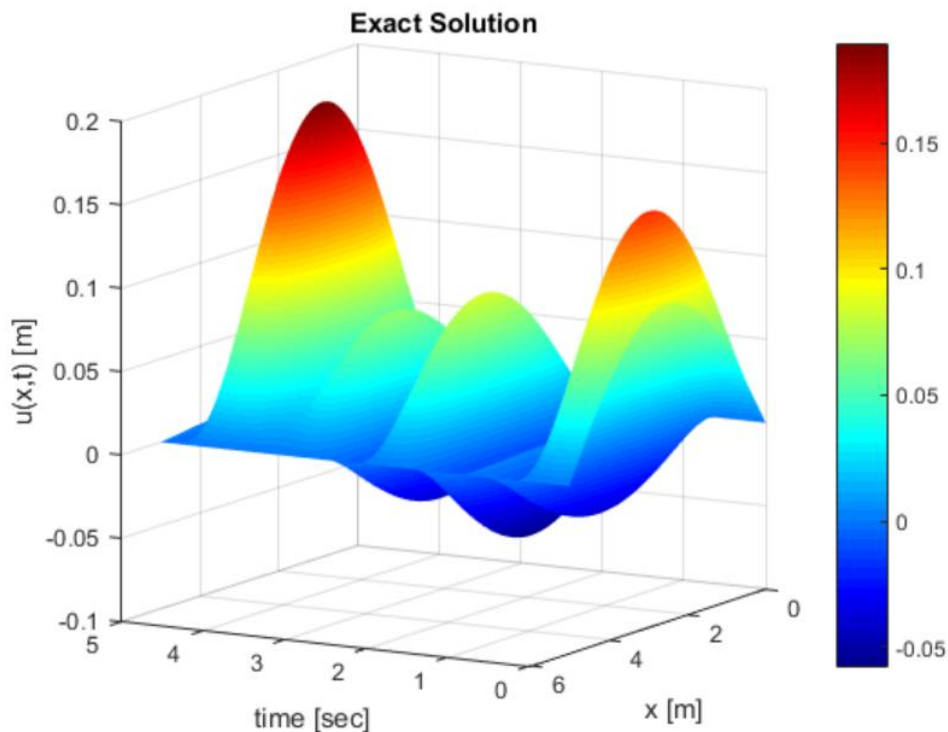
Ορίζεται το PINN ως η συνάρτηση $f(t, x)$:

$$f := EI \frac{\partial^4 u}{\partial x^4} + m \frac{\partial^2 u}{\partial t^2} - p(x, t)$$

Όπου $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$

Το δίκτυο τροφοδοτείται και εκπαιδεύεται με τα δεδομένα εισόδου x_i, t_i, u_i όπως αυτά έχουν προκύψει από την αναλυτική επίλυση του προβλήματος.

Το κομμάτι του κώδικα στο οποίο ορίζεται το PINN είναι παρόμοιο με την περίπτωση της ταλάντωσης της αμφιέρειστης δοκού υπό ημιτονοειδές φορτίο.



Εικόνα 7–27: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$

Για την εκπαίδευση του ΤΝΔ επιλέχθηκε το σύνολο όλων των συνοριακών σημείων, $N_u = 2000$ από σύνολο 2002 αρχικών και συνοριακών σημείων, ενώ από τα εσωτερικά επιλέχθηκαν $N_f = 100000$ τυχαία καταναμημένα σημεία στο πεδίο ορισμού της εξίσωσης, από σύνολο 501000 σημείων (το 19% των συνολικών σημείων).

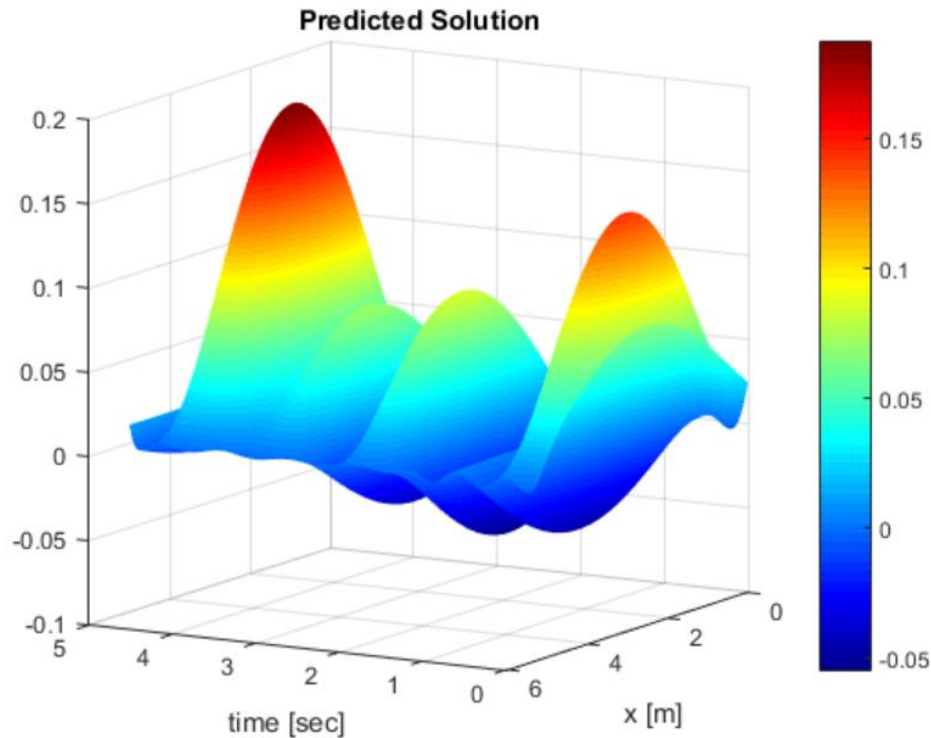
Η αρχιτεκτονική του ΤΝΔ αποτελείται από 8 επίπεδα νευρώνων, με κάθε κρυφό επίπεδο να περιέχει 20 νευρώνες και με συνάρτηση ενεργοποίησης την υπερβολική εφαπτομένη \tanh .

Η εκπαίδευση έγινε χρησιμοποιώντας την συνάρτηση του μέσου τετραγωνικού σφάλματος. Συγκεκριμένα, υπολογίζεται από το δίκτυο το άθροισμα των σφαλμάτων των μετατοπίσεων u των συνοριακών σημείων, των σφαλμάτων των τιμών της καμπυλότητας των συνοριακών σημείων u_{xx} (οι οποίες είναι μηδενικές καθότι αμφιέρειστη δοκός) και του σφάλματος της συνάρτησης f , η οποία θα πρέπει να έχει μηδενική τιμή. Θεωρήθηκε ότι κάθε μία από τις προαναφερθείσες συνιστώσες της συνάρτησης σφάλματος συνεισφέρει εξίσου στο τελικό αποτέλεσμα.

Όσον αφορά τον αλγόριθμο βελτιστοποίησης, χρησιμοποιήθηκε ο αλγόριθμος L-BFGS-B και ο αλγόριθμος Adam προκειμένου να υπάρξει γρηγορότερη σύγκλιση του δικτύου στις βέλτιστες παραμέτρους που περιγράφουν το πρόβλημα καθώς ο αριθμός των δεδομένων εκπαίδευσης N_u είναι μικρός αλλά ορίστηκαν 20000 εποχές εκπαίδευσης του. Θεωρείται ότι χρησιμοποιήθηκε επαρκής αριθμός σημείων N_f του πεδίου ορισμού της εξίσωσης για την εκπαίδευση της f και το ΤΝΔ που ορίστηκε έχει την κατάλληλη αρχιτεκτονική, ώστε να μπορεί το δίκτυο να δώσει μια αρκετά ικανοποιητική πρόβλεψη.

Το σφάλμα του δικτύου υπολογίστηκε $\text{error} = 7.08 * 10^{-2}\%$ με τις αρχικές συνθήκες ($t=0$) να είναι μηδενικές σε όλη τη δοκό και μηδενικές μετατοπίσεις στα δύο άκρα της δοκού όλο το διάστημα. Η μέγιστη μετατόπιση πραγματοποιείται στην θέση $x = 2.4975 \text{ m}$ τη χρονική στιγμή $t = 4.22 \text{ sec}$. Η ελάχιστη μετατόπιση είναι η -0.057 πραγματοποιείται στην θέση $x = 2.4975 \text{ m}$ τη χρονική στιγμή $t = 1.68 \text{ sec}$.

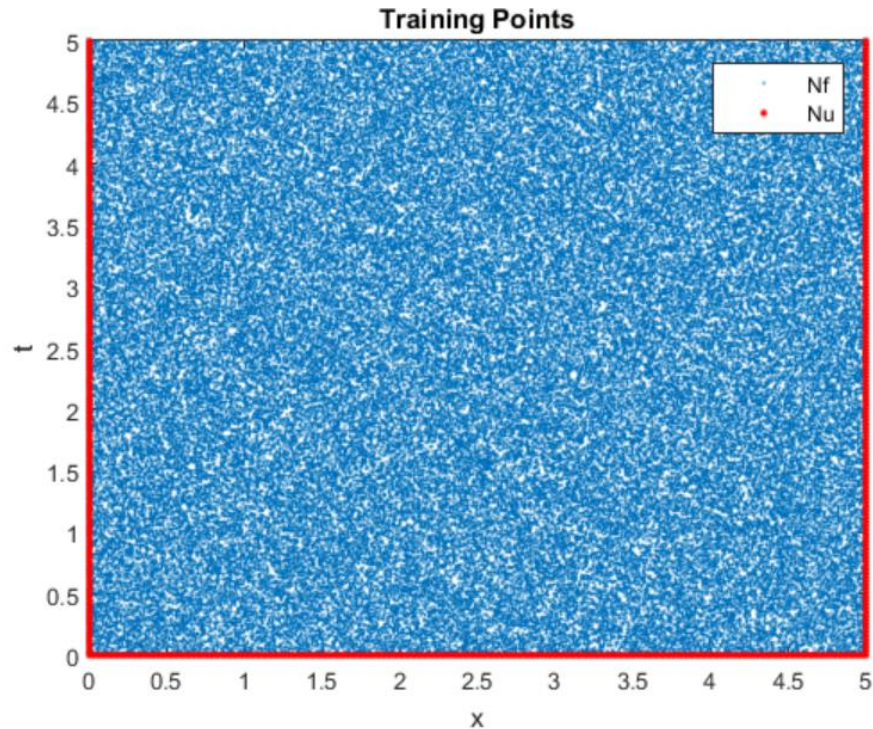
Στη συνέχεια στο Σχήμα απεικονίζεται η τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου, όπως προέκυψε από το ΤΝΔ.



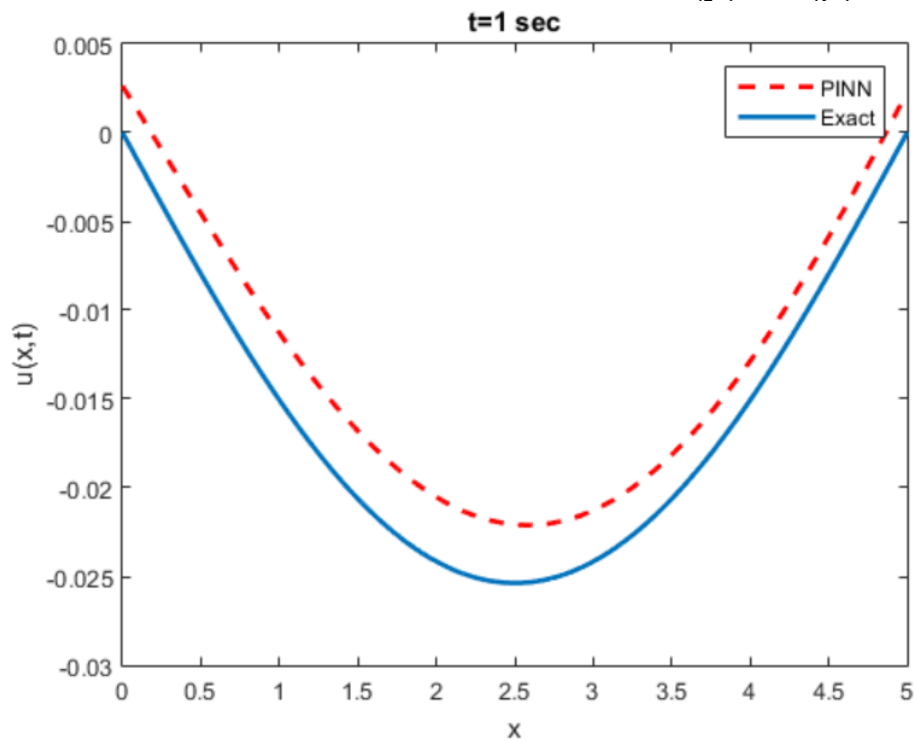
Εικόνα 7–28: Τρισδιάστατη απεικόνιση λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$

Η Εικόνα 7–30, Εικόνα 7–31,

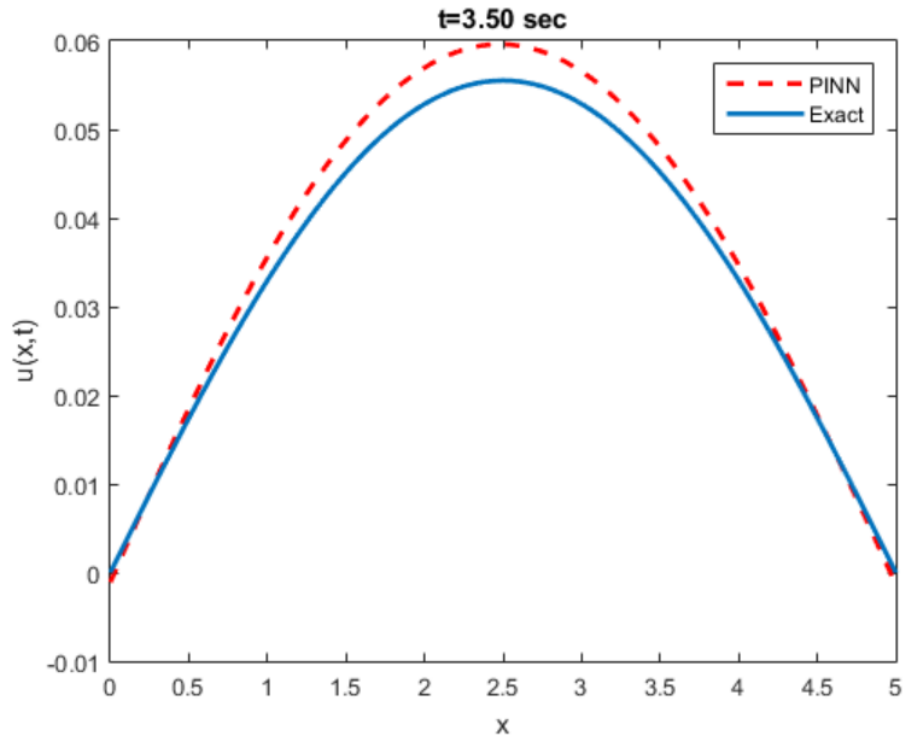
Εικόνα 7–32 αποτυπώνουν και συγκρίνουν τη γραφική παράσταση της αναλυτικής λύσης της ΜΔΕ και της πρόβλεψης του PINN για τις χρονικές στιγμές $t = 1, 3.50, 4.70$ sec, αντίστοιχα. Επιπλέον στα γραφήματα Εικόνα 7–33, Εικόνα 7–34, Εικόνα 7–35 αποτυπώνονται και συγκρίνονται η γραφική παράσταση της αναλυτικής λύσης της ΜΔΕ και της πρόβλεψης του δικτύου στις θέσεις $x = 1.50, 3, 4.50, 4.95$ m, αντίστοιχα. Από τα διαγράμματα αυτά συμπεραίνεται ότι το δίκτυο αξιοποιώντας ένα μέρος των αρχικών συνθηκών, των συνοριακών συνθηκών και των εσωτερικών σημείων προσέγγισε με αρκετά μεγάλη ακρίβεια την πολυπλοκότητα εξίσωσης.



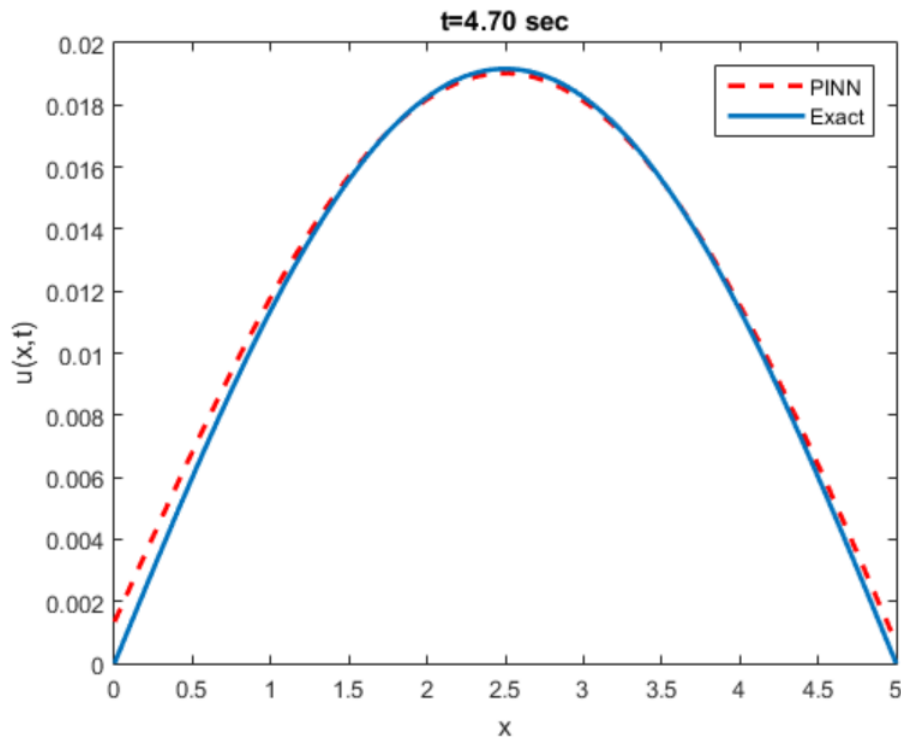
Εικόνα 7–29 Σχηματική απεικόνιση των δεδομένων εκπαίδευσης του PINN για την επίλυση της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$



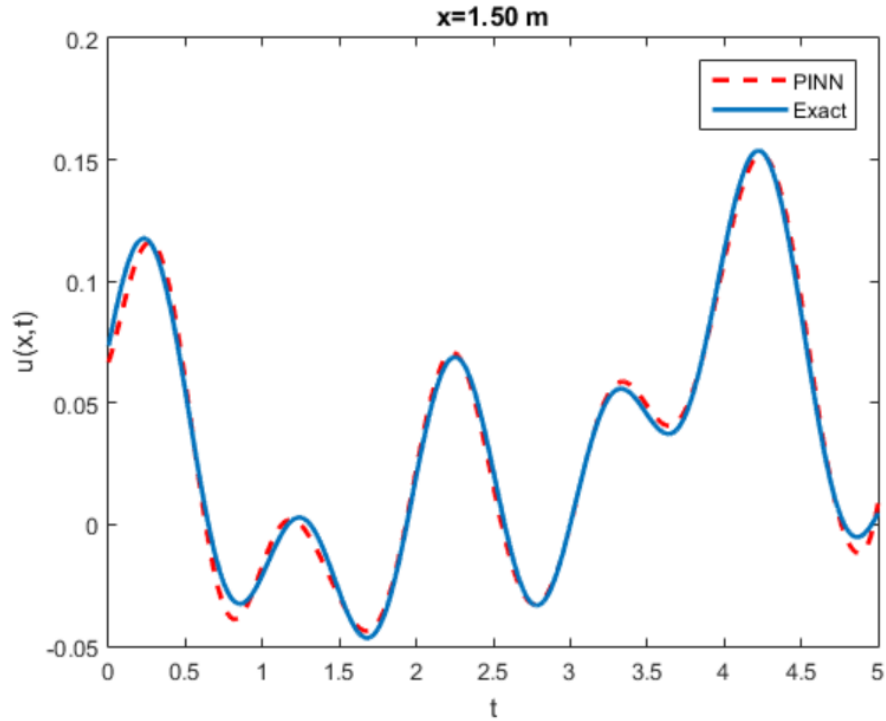
Εικόνα 7–30: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$ για $t=1$ se



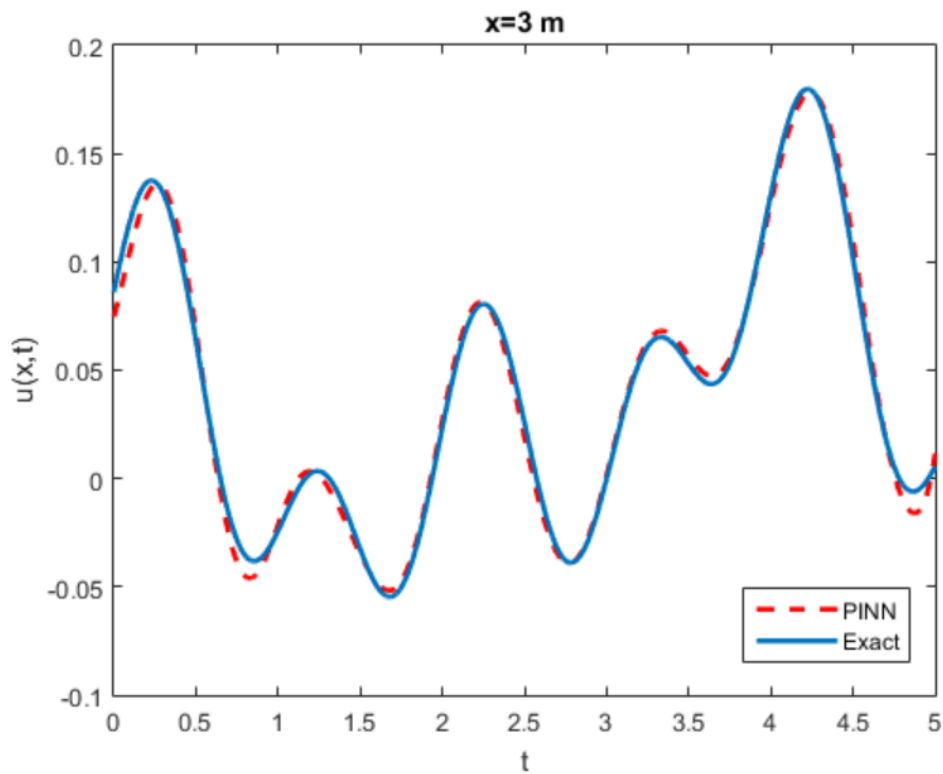
Εικόνα 7–31: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για για αμφιέριστη δοκό για $t=3.50$ sec υπό φορτίο $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$



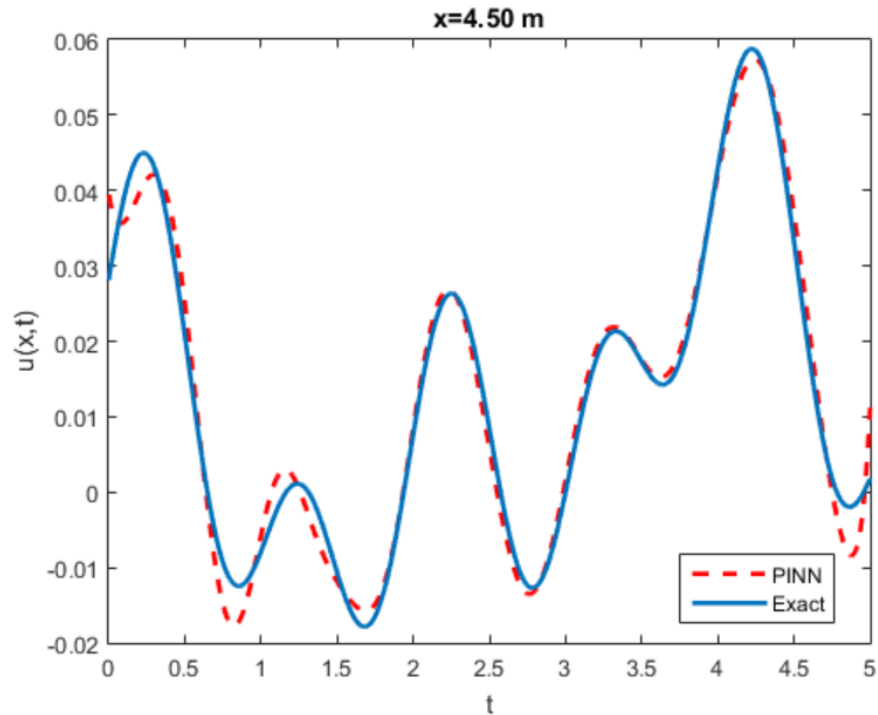
Εικόνα 7–32: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $t=4.70$ sec για αμφιέριστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$



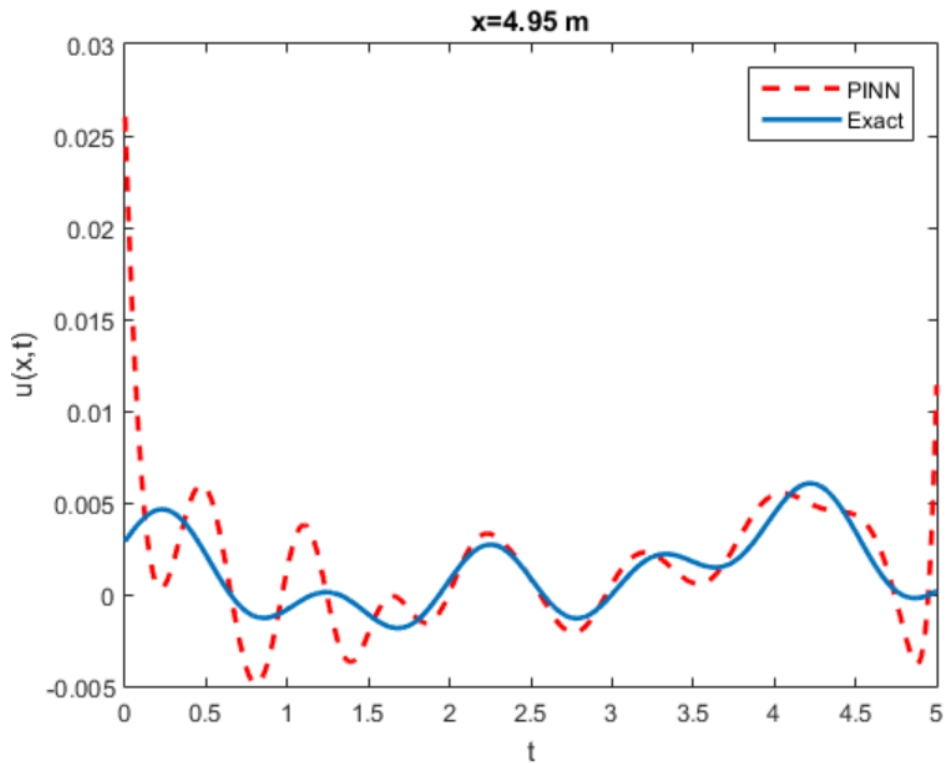
Εικόνα 7–33: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $x=1.50$ m για αμφιέρειστη δοκό υπό φορτίο $p(x,t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$



Εικόνα 7–34: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $x=3$ m για αμφιέρειστη δοκό υπό φορτίο $p(x,t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$



Εικόνα 7–35: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $x=4.50$ m για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$



Εικόνα 7–36: Σύγκριση αναλυτικής και εκτιμώμενης λύσης της εξίσωσης Euler-Bernoulli για $x=4.95$ m για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t) + \cos\left(\frac{\pi}{2}t\right) + \sin\left(\frac{\pi}{9}t\right) + \sin\left(\pi t + \frac{\pi}{3}\right)$

7.2 Περιπτώσεις προβλημάτων ταυτοποίησης παραμέτρων (identification)

7.2.1 Εφαρμογή 1: Εξίσωση ταλάντωσης αμφιέρειστης δοκού υπό ημιτονοειδές φορτίο (χωρίς θόρυβο)

Η εφαρμογή των ΤΝΔ τα οποία υπακούν νόμους της φυσικής (PINN) δίνει την δυνατότητα της εύρεσης των παραμέτρων της εξίσωσης με αρκετά ακριβή τρόπο, γεγονός που αποτελεί μεγάλο πλεονέκτημα της μεθόδου καθώς η συλλογή της αναλυτικής λύσης μέσω εργαστηριακών δοκιμών ή από μετρήσεις πεδίου είναι ιδιαίτερα δύσκολη και κοστοβόρα.

Όπως είναι φανερό, στο πρόβλημα ταλάντωσης της ελαστικής αμφιέρειστης δοκού, η οποία περιγράφεται από την εξίσωση (5.5), οι άγνωστες παράμετροι είναι η παράμετρος του γινομένου $E \cdot I$ και του γινομένου $\rho \cdot A$ της δοκού.

Στην παρούσα εργασία επιχειρήθηκε η ταυτοποίηση της παραμέτρου του γινομένου του μέτρου ελαστικότητας και της δυσκαμψίας, EI , με γνωστή τη δεύτερη παράμετρο της εξίσωσης διάδοσης κύματος σε δοκό, δηλαδή του γινομένου της πυκνότητας της δοκού με το εμβαδόν της. Η παράμετρος λ_{1} ορίζεται ως μεταβλητή στην γλώσσα προγραμματισμού Python, όπως φαίνεται παρακάτω με χρήση της βιβλιοθήκης Tensorflow (tf) και κατά την εκκίνηση του δικτύου λαμβάνει την τιμή 0.5 η οποία ανα εποχή αλλάζει έως ότου προσεγγίσει την τιμή εκείνη για την οποία περιγράφονται καλύτερα τα δεδομένα της εκπαίδευσης.

```
# Initialize parameter
self.lambda_1 = tf.Variable([0.5], dtype=tf.float32)
```

Για την εκπαίδευση του ΤΝΔ επιλέχθηκε το σύνολο όλων των συνοριακών σημείων $X_{u_train_b}$, τροφοδοτώντας το δίκτυο και με τις κατακόρυφες μετακινήσεις των συνοριακών σημείων u_{train_b} , αλλά και με τις καμπυλότητές τους u_{xx_train} . Από τα εσωτερικά επιλέχθηκαν $N_f = 10000$ τυχαία καταναμημένα σημεία στο πεδίο ορισμού της εξίσωσης, από σύνολο 100100 σημείων (περίπου το 10% των συνολικών σημείων).

Η αρχιτεκτονική του ΤΝΔ αποτελείται από 8 επίπεδα νευρώνων, με κάθε κρυφό επίπεδο να περιέχει 20 νευρώνες και με συνάρτηση ενεργοποίησης την υπερβολική εφαπτομένη \tanh .

Όμοια με το πρόβλημα της επαγωγής, η εκπαίδευση έγινε με την ελαχιστοποίηση της συνάρτησης κόστους όπου έγινε πάλι χρήση του μέσου τετραγωνικού σφάλματος και υπολογίστηκε το άθροισμα των

σφαλμάτων των μετατοπίσεων των σημείων u των συνοριακών σημείων, των σφαλμάτων των τιμών της καμπυλότητας των συνοριακών σημείων u_{xx} (οι οποίες είναι μηδενικές καθότι αμφιέρειστη δοκός), και των σφαλμάτων της τιμής της συνάρτησης f (η οποία θα πρέπει να έχει μηδενική τιμή). Το τμήμα του κώδικα που ορίζεται η συνάρτηση σφάλματος (loss function) του δικτύου είναι:

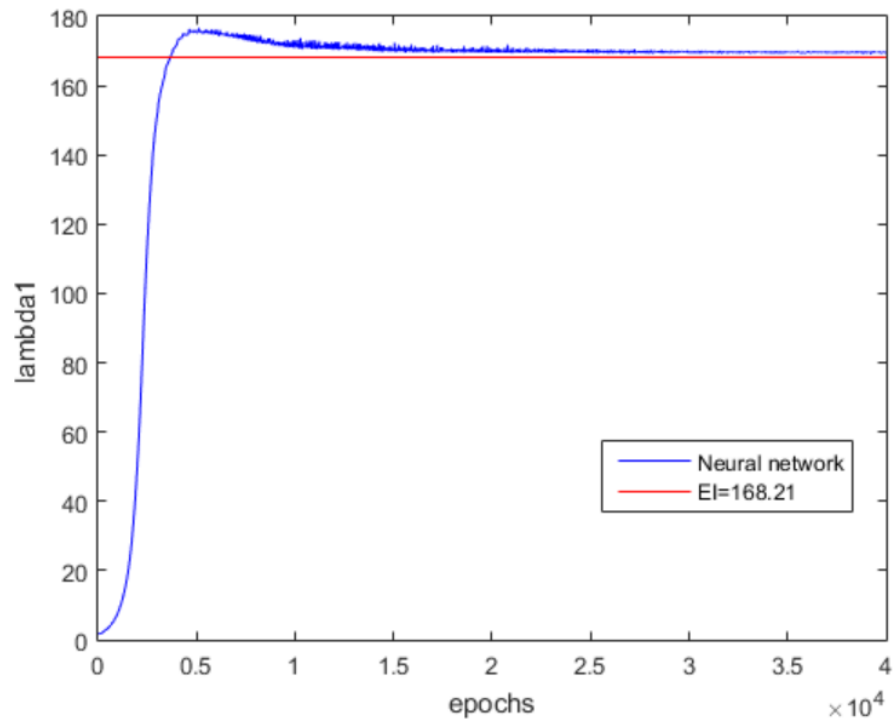
```
self.loss = tf.reduce_mean(tf.square(self.u_tf - self.u_pred)) + \  
             tf.reduce_mean(tf.square(self.u_xx_tf - self.u_xx_pred)) + \  
             0.001*tf.reduce_mean(tf.square(self.f_pred))
```

Όπως διαπιστώνεται από το παραπάνω σχήμα, κάθε μία από τις προαναφερθείσες συνιστώσες της συνάρτησης σφάλματος ορίστηκε να μην συνεισφέρει εξίσου στο τελικό αποτέλεσμα. Συγκεκριμένα, δόθηκε μεγαλύτερη βαρύτητα στο σφάλμα της μετατόπισης και της καμπυλότητας των συνοριακών σημείων της δοκού, προκειμένου το δίκτυο να εκπαιδευτεί κατάλληλα για τη συγκεκριμένη δοκό. Αντίθετα, το σφάλμα της συνάρτησης συνεισφέρει στο ολικό σφάλμα της εκπαίδευσης του δικτύου, σε ποσοστό τάξης 1 προς 1000. Αν δεν δοθεί στο δίκτυο επαρκής πληροφορία για τη συμπεριφορά της δοκού στα συνοριακά της σημεία κατά την ταλάντωση, το δίκτυο δεν δύναται να συγκλίνει στις βέλτιστες τιμές των παραμέτρων. Αυτό διότι, το δίκτυο αντιλαμβάνεται τη δοκό σαν ένα στερεό σώμα μη επαρκώς στηριγμένο. Στην περίπτωση αυτή, ναι μεν θα επιτυγχάνεται ικανοποιητική πρόβλεψη της παραμόρφωσης της δοκού στα εσωτερικά της σημεία, αλλά θα είναι αδύνατον να προβλεφθεί από το δίκτυο η συμπεριφορά στα άκρα. Επίσης, αν παρέχεται πληροφορία μόνο για τις συνοριακές συνθήκες και κανένα δεδομένο του υπόλοιπου χωρίου, τα ζεύγη των παραμέτρων που δίνει το ΤΝΔ ως λύση αποτελούν ένα τυχαίο ζεύγος από τα πολλά ζεύγη παραμέτρων που το δίκτυο βρίσκει ότι ικανοποιούν την εξίσωση f . Αυτό συμβαίνει επειδή για δεδομένες συνοριακές συνθήκες οποιαδήποτε αμφιέρειστη δοκός μπορεί να ικανοποιεί τη δεδομένη εξίσωση f .

Όσον αφορά τον αλγόριθμο βελτιστοποίησης, χρησιμοποιήθηκε αρχικά ο αλγόριθμος Adam, πραγματοποιώντας 40000 εποχές εκπαίδευσης. Με το πέρας των εποχών η βελτιστοποίηση συνεχίστηκε με τον αλγόριθμο L-BFGS-B. Έγινε διαδοχική χρήση των δύο αλγόριθμων βελτιστοποίησης καθώς το συγκεκριμένο πρόβλημα έχει μεγαλύτερο υπολογιστικό φόρτο συγκριτικά με το πρόβλημα της επαγωγής, αφού εκτιμά ταυτόχρονα και την τιμή της παραμέτρου λ_1 .

Το σφάλμα του δικτύου υπολογίστηκε $\text{error} = 1.62 \cdot 10^{-3} \%$ και το σφάλμα της παραμέτρου λ_1 είναι $\text{Error } \lambda_1 = 0.8\%$. Συγκεκριμένα η πρόβλεψη της λ_1 είναι $\lambda_1 = 169.538$, με την τιμή-στόχο να είναι $\lambda_1 = 168.21$.

Η σχηματική αναπαράσταση της ταυτοποίησης της παραμέτρου λ_1 από το ΤΝΔ είναι η Εικόνα 7–37.



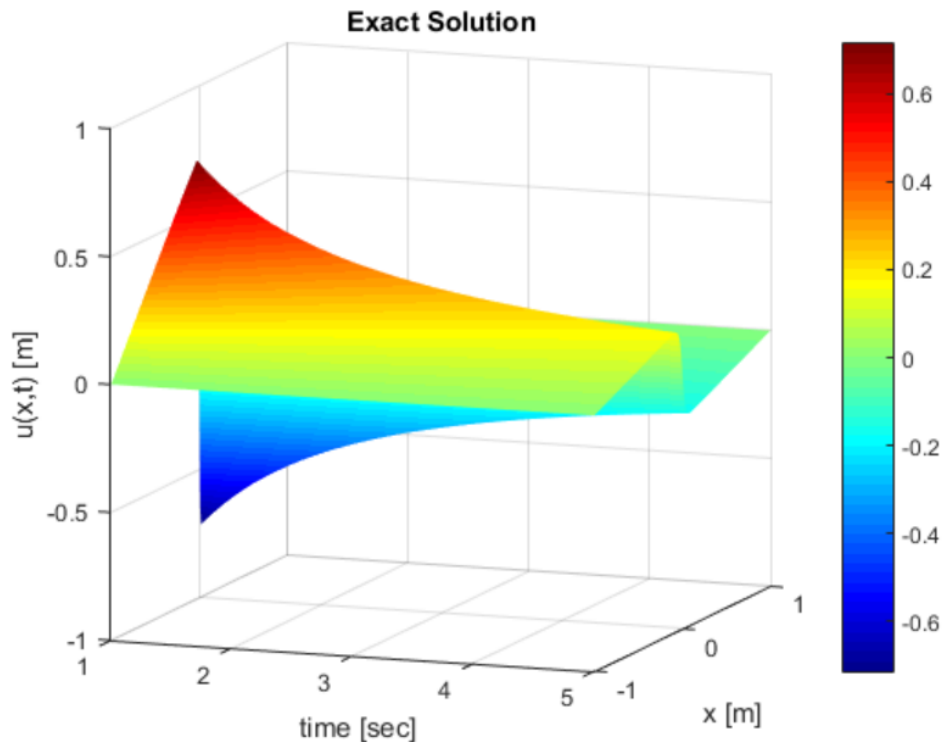
Εικόνα 7–37: Αποτέλεσμα του PINN για την ταυτοποίηση της παραμέτρου του γινομένου του EI της εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$

ΚΕΦΑΛΑΙΟ 8. Εφαρμογή της μεθόδου PINN για την πρόβλεψη της συμπεριφοράς της ΜΔΕ

8.1 Εφαρμογή 1: Εξίσωση Burgers

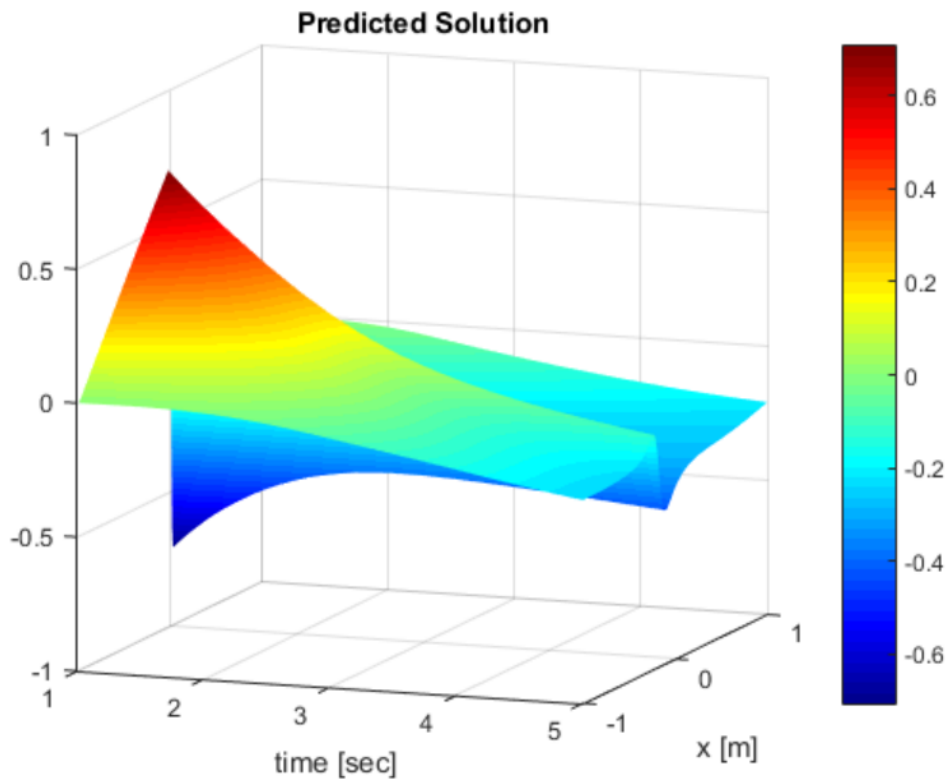
Αρχικά εξετάστηκε η δυνατότητα πρόβλεψης της συμπεριφοράς της εξίσωσης Burgers με εφαρμογή της μεθόδου PINN. Για την εύρεση της αναλυτικής λύσης (Exact solution) της εξίσωσης (6-1) κατασκευάστηκε κώδικας στην γλώσσα προγραμματισμού MATLAB. Οι οριακές συνθήκες που χρησιμοποιήθηκαν είναι του Dirichlet, εξισώσεις (6-2), και η επίλυση πραγματοποιήθηκε για χρόνο $0 \leq t \leq 5 \text{ sec}$ για χρονική και χωρική διακριτοποίηση, $\Delta t = 0.01 \text{ sec}$ και $\Delta x = 0.0078 \text{ m}$ αντίστοιχα. Το σφάλμα του δικτύου υπολογίστηκε $\text{error} = 0.795\%$.

Από την αναλυτική επίλυση προέκυψε η παρακάτω τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου με $0 \leq t \leq 5 \text{ sec}$ και στις θέσεις $-1 \leq x \leq 1 \text{ m}$.



Εικόνα 8–1: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης Burgers για χρόνο $1 \leq t \leq 5 \text{ sec}$

Στη συνέχεια στην Εικόνα 8–2 απεικονίζεται η τρισδιάστατη σχηματική αναπαράσταση της πρόβλεψης της ταλάντωσης της δοκού συναρτήσει του χρόνου για $0 \leq t \leq 5 \text{ sec}$, όπως προέκυψε από PINN ίδιας αρχιτεκτονικής και ίδιας εξίσωσης σφάλματος με την παράγραφο 6.1. το οποίο έχει εκπαιδευτεί για χρόνο $0 \leq t \leq 1 \text{ sec}$ για θέσεις $-1 \leq x \leq 1 \text{ m}$. Για την εκπαίδευση πραγματοποιήθηκαν 7000 εποχές εκπαίδευσης με τον αλγόριθμο βελτιστοποίησης Adam και κατόπιν με τον L-BFGS.



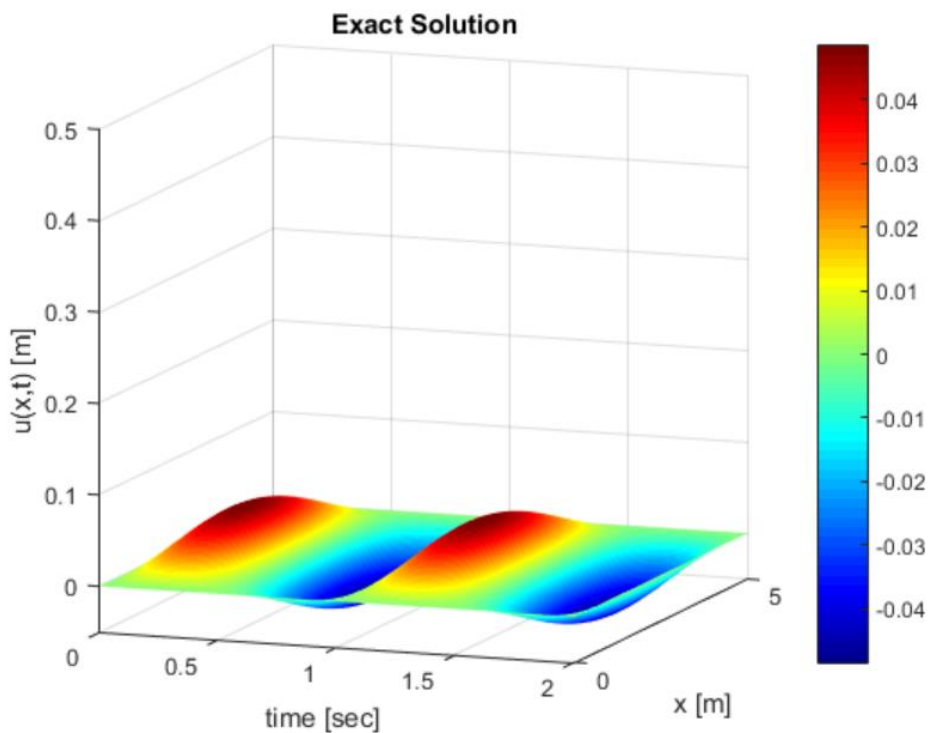
Εικόνα 8–2: Τρισδιάστατη απεικόνιση της λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Burgers για χρόνο $1 \leq t \leq 5 \text{ sec}$ για τον οποίο το δίκτυο δεν έχει εκπαιδευτεί.

Όπως φαίνεται και από την σύγκριση των εικόνων Εικόνα 8–1 Εικόνα 8–2 η μέθοδος PINN αδυνατεί να προβλέψει με ακρίβεια τη συμπεριφορά της εξίσωσης για πεδίο που δεν έχει εκπαιδευτεί προηγουμένως.

8.2 Εφαρμογή 2: Εξίσωση ανάλυσης διάδοσης κύματος σε αμφιέρειστη δοκό

Εξετάστηκε η δυνατότητα πρόβλεψης της συμπεριφοράς της εξίσωσης ταλάντωσης αμφιέρειστης δοκού υπό ημιτονοειδές φορτίο με εφαρμογή της μεθόδου PINN. Η διαδικασία που ακολουθήθηκε για την εύρεση της αναλυτικής λύσης, καθώς και το φορτίο $p(x, t)$ της δοκού, είναι όμοια με αυτήν που περιγράφηκε στην παράγραφο 7.1.1. Όμοια είναι επίσης και η αρχιτεκτονική του PINN που εφαρμόστηκε. Το δίκτυο εκπαιδεύτηκε για χρόνο $0 \leq t \leq 1 \text{ sec}$ για χρονική και χωρική διακριτοποίηση, $\Delta t = 0.001 \text{ sec}$ και $\Delta x = 0.05 \text{ m}$ αντίστοιχα και στη συνέχεια κλήθηκε να προβλέψει τη συμπεριφορά της δοκού για $1 \leq t \leq 2 \text{ sec}$ για τις ίδιες θέσεις. Το σφάλμα του δικτύου υπολογίστηκε $\text{error} = 1.067\%$

Από την αναλυτική επίλυση (Exact solution) προέκυψε η παρακάτω τρισδιάστατη σχηματική αναπαράσταση της ταλάντωσης της δοκού συναρτήσει του χρόνου για $0 \leq t \leq 2 \text{ sec}$.

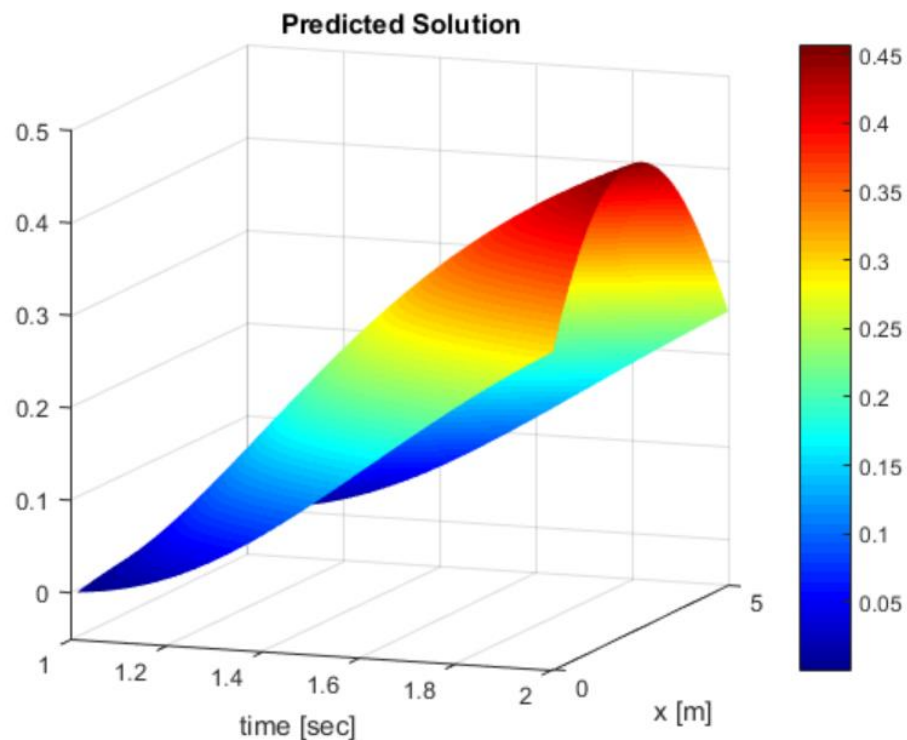


Εικόνα 8–3: Τρισδιάστατη απεικόνιση αναλυτικής λύσης εξίσωσης εξίσωσης Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για χρόνο $0 \leq t \leq 2 \text{ sec}$

Στη συνέχεια στην Εικόνα 8–4 απεικονίζεται η τρισδιάστατη σχηματική αναπαράσταση της πρόβλεψης της ταλάντωσης της δοκού συναρτήσει του χρόνου για $1 \leq t \leq 2 \text{ sec}$, όπως προέκυψε από

το PINN. Για την εκπαίδευση πραγματοποιήθηκαν 18000 εποχές εκπαίδευσης με τον αλγόριθμο βελτιστοποίησης Adam και κατόπιν με τον L-BFGS.

Όπως συμπεραίνεται με τον τρόπο που είναι ορισμένο το PINN στην παρούσα εργασία αδυνατεί να προβλέψει ορθά την συμπεριφορά της ταλάντωσης της δοκού για κάποιο χωροχρονικό πεδίο εκτός εκείνου για το οποίο έχει εκπαιδευτεί. Άρα τα συγκεκριμένα PINNs δεν ενδείκνυνται σε περιπτώσεις που ζητούμενο είναι η πρόβλεψη της συμπεριφοράς μιας ΜΔΕ σε ορισμένο χωροχρονικό πεδίο, το οποίο είναι άγνωστο για το δίκτυο. Με τον όρο «άγνωστο» νοείται το πεδίο εκείνο για το οποίο το PINN δεν έχει τροφοδοτηθεί με κάποιο δεδομένο προκειμένου να εκπαιδευτεί.



Εικόνα 8-4: Τρισδιάστατη απεικόνιση της λύσης όπως προέκυψε από την εφαρμογή της μεθόδου PINN για την εξίσωση Euler-Bernoulli για αμφιέρειστη δοκό υπό φορτίο $p(x, t) = \sin(2\pi t)$ για χρόνο $1 \leq t \leq 2 \text{ sec}$ για τον οποίο το δίκτυο δεν έχει εκπαιδευτεί.

ΚΕΦΑΛΑΙΟ 9. Συμπεράσματα και προτάσεις για μελλοντική έρευνα

9.1 Συμπεράσματα

Στόχος της παρούσας εργασίας είναι η παρουσίαση της μεθόδου επίλυσης ΜΔΕ με καθοδηγούμενα από τη φυσική ΤΝΔ (PINNs) καθώς και η διερεύνηση της συνεισφοράς της εν λόγω μεθόδου σε προβλήματα ανάλυσης διάδοσης κύματος σε αμφιέριστη και δοκό πρόβολο. Σκοπός των PINNs δεν είναι η εξ ολοκλήρου αντικατάσταση των αριθμητικών μεθόδων (πχ πεπερασμένα στοιχεία) που υπάρχουν στη βιβλιογραφία για την επίλυση τέτοιων προβλημάτων αλλά η συνύπαρξή τους με τους αλγορίθμους πρόβλεψης και επίλυσης ΜΔΕ. Παρακάτω παρατίθενται κάποια συμπεράσματα για τα PINNs, όπως προέκυψαν ύστερα από την επεξεργασία των εφαρμογών.

1. Τα PINNs, που αποτελούνται από απλά πολυεπίπεδα ΤΝΔ, είναι μία αποδοτική μέθοδος, η οποία μπορεί με σχετικά μικρό όγκο δεδομένων, να προσεγγίσει την πραγματική συμπεριφορά του φαινομένου με ακρίβεια. Επισημαίνεται πως η ποιότητα και η ποσότητα των δεδομένων είναι μεγάλης σημασίας για την επιτυχία των PINNs.
2. Η πρόβλεψη της ανάλυσης κύματος σε δοκό είναι αρκετά επιτυχής εντός του χωροχρονικού πεδίου για το οποίο ο αλγόριθμος έχει τροφοδοτηθεί με δεδομένα, ακόμα και με μικρό πλήθος αυτών. Ωστόσο στην περίπτωση των συνοριακών συνθηκών τα PINNs παρουσιάζουν αδυναμία ακριβούς προσέγγισης της αναλυτικής λύσης στα σημεία αυτά, ακόμα και αν το δίκτυο τροφοδοτηθεί με όλα τα δεδομένα αυτών των σημείων.
3. Τα PINNs λαμβάνοντας πληροφορία από τις παραγώγους προσεγγίζουν με ακρίβεια την αναλυτική λύση εντός του χωροχρονικού πεδίου στο οποίο έχει οριστεί το πρόβλημα εν αντιθέσει με τις αριθμητικές μεθόδους που απαιτείται διακριτοποίηση στο πεδίο του χώρου και του χρόνου.
4. Για την επιτυχή προσέγγιση της αναλυτικής λύσης απαιτείται η προσθήκη επιπλέον όρων και η τροποποίηση των παραμέτρων βαρύτητας κάθε όρου της εξίσωσης σφάλματος του PINN. Πλεονέκτημα όμως αποτελεί η ευκολία υλοποίησης της μεθόδου και η παρόμοια μεταχείριση του εκάστοτε προβλήματος.
5. Από την εφαρμογή της εξίσωσης Burgers και της εξίσωσης διάδοσης κύματος σε δοκό συμπεραίνεται ότι ορθώς ορισμένα PINNs προσεγγίζουν ικανοποιητικά την αναλυτική λύση πολύπλοκων ΜΔΕ εντός του χωροχρονικού πεδίου που έχει οριστεί το πρόβλημα. Μειονέκτημα

αποτελεί η αδυναμία πρόβλεψης της συμπεριφοράς των συνόρων σημείων στις εφαρμογές της εξίσωσης διάδοσης κύματος σε δοκό.

6. Όπως διαπιστώθηκε τα PINNs αδυνατούν να προβλέψουν την εξέλιξη της εξίσωσης για χωροχρονικό πεδίο για το οποίο δεν έχουν εκπαιδευτεί. Τα PINNs δεν μπορούν να χρησιμοποιηθούν για πρόβλεψη μελλοντικών καταστάσεων χωρίς καθόλου δεδομένα.
7. Στις εφαρμογές των PINNs που επιχειρήθηκε η εύρεση παραμέτρου αποτελέσματα ήταν αρκετά ικανοποιητικά. Συμπεραίνεται ότι κατάλληλα ορισμένα PINNs μπορούν να εφαρμόσουν με επιτυχία το αντίστροφο πρόβλημα της εύρεσης παραμέτρων της εξίσωσης.

Τα συμπεράσματα που προέκυψαν είναι ιδιαίτερα ικανοποιητικά για την προσφορά των ΤΝΔ στην επίλυση προβλημάτων μηχανικού, ενώ η δυνατότητα και η προοπτική που έχουν για περαιτέρω ανάπτυξη καθιστά ακόμα πιο ελκυστική και πιθανή την εφαρμογή τους στο μέλλον. Ωστόσο, δεν είναι εφικτή η γενίκευση των PINNs για πρόβλεψη της συμπεριφοράς του εκάστοτε προβλήματος εκτός του χωροχρονικού πεδίου στο οποίο έχει εκπαιδευτεί, σε σχέση με άλλες μεθόδους επίλυσης ΜΔΕ όπως Variational PINNs των E. Kharazmi, Z. Zhang, G. E. Karniadakis.

9.2 Προτάσεις για μελλοντική έρευνα

Ο σκοπός της παρούσας διπλωματικής εργασίας είναι η παρουσίαση μιας καινούργιας μεθόδου επίλυσης των εξισώσεων διόδευσης κύματος σε δοκούς, αυτής των ΤΝΔ που υπακούν στους νόμους της φυσικής. Τα αποτελέσματα, στην πλειονότητά τους, όσον αναφορά την προσέγγιση της αναλυτικής λύσης της εκάστοτε ΜΔΕ που μελετήθηκε είναι ενθαρρυντικά. Τα συμπεράσματα που προαναφέρθηκαν αποτελούν κίνητρο για περαιτέρω έρευνα στο αντικείμενο της επίλυσης ΜΔΕ για προβλήματα τόσο μηχανικού όσο και κάθε άλλου κλάδου των εφαρμοσμένων μαθηματικών. Παρά το μεγάλο εύρος εφαρμοσιμότητας και ευελιξίας που προσφέρουν τα PINNs, αναφορικά με τα προβλήματα ανάλυσης διόδευσης κύματος που εφαρμόστηκαν προέκυψαν το εξής βασικό μειονέκτημα. Η αδυναμία του δικτύου να «κατανοήσει» και να προσομοιάσει με ακρίβεια την συμπεριφορά στηριγμένου άκρου, παρά το πλήθος των δεδομένων με τα οποία τροφοδοτείται, αποτελεί γεγονός αρκετά περιοριστικό για την εφαρμογή της μεθόδου.

Με δεδομένο ότι η μέθοδος PINNs έχει αρκετά περιθώρια για βελτίωση, οι προτάσεις για μελλοντική έρευνα συνοψίζονται ως εξής:

1. Γενίκευση στον δισδιάστατο και κατ'επέκταση στον τρισδιάστατο χώρο με διαφοροποίηση του κώδικα αλλά όχι με σημαντική αλλαγή των κύριων σημείων του.

2. Εκμετάλλευση των PINNs για την πρόβλεψη της συμπεριφοράς του φαινομένου για χωροχρονικό πεδίο εκτός του πεδίου για το οποίο έχει εκπαιδευτεί.
3. Αξιοποίηση της δυνατότητας των PINNs για εύρεση της τιμής των παραμέτρων ΜΔΕ, σε περιπτώσεις με επαρκή δεδομένα.
4. Βελτιστοποίηση της αρχιτεκτονικής των ΤΝΔ που προορίζονται για προβλήματα δυναμικής ανάλυσης μηχανικού, κυρίως στο κομμάτι της «κατανόησης» από το δίκτυο της συμπεριφοράς των συνοριακών συνθηκών στην ανάλυση διόδευσης κύματος.

ΚΕΦΑΛΑΙΟ 10. Αναφορές

Βουθούνης, Π. Α., Σταματίου, Α. Ε., & Βουθούνη, Π. Π. (2015). Μηχανική παραμορφώσιμου στερεού.

Κατσικαδέλης, Ι. Θ. (2012). Δυναμική ανάλυση των κατασκευών. Εκδόσεις Συμμετρία, Ζωγράφου, Αθήνα.

Bathe, Klaus-Jürgen: *Finite element procedures*, 2nd edn, Prentice Hall, (2014)

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of machine learning*. MIT press.

Basdevant, C., Deville, M., Haldenwang, P., Lacroix, J. M., Ouazzani, J., Peyret, R., ... & Patera, A. T. (1986). Spectral and finite difference solutions of the Burgers equation. *Computers & fluids*, 14(1), 23-41.

Ian Goodfellow, Yoshua Bengio and Aaron Courville. Deep Learning. MIT Press, 2016-11-18.

Neural Networks and Learning Machines Third Edition Simon Haykin McMaster University Hamilton, Ontario, Canada

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707.

Jones, R. P. N. (1956). THE REFLECTION OF TRANSVERSE WAVES IN BEAMS. *The Quarterly Journal of Mechanics and Applied Mathematics*, 9(4), 499-507.

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge university press.

Sammut, C., & Webb, G. I. (Eds.). (2011). *Encyclopedia of machine learning*. Springer Science & Business Media.

- Alpaydin, E. (2020). Introduction to machine learning. MIT press.
- Doyle, J. F. (1989). Wave propagation in structures. In Wave propagation in structures (pp. 126-156). Springer, New York, NY.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Yam, J. Y., & Chow, T. W. (2000). A weight initialization method for improving training speed in feedforward neural network. *Neurocomputing*, 30(1-4), 219-232.
- Karlsson, L., & Bonde, O. (2020). A Comparison of Selected Optimization Methods for Neural Networks.
- Adlers, J., & Pihl, G. (2018). Prediction of training time for deep neural networks in TensorFlow.
- Dai, Y. H. (2002). Convergence properties of the BFGS algorithm. *SIAM Journal on Optimization*, 13(3), 693-701.
- Weaver Jr, W., Timoshenko, S. P., & Young, D. H. (1990). Vibration problems in engineering. John Wiley & Sons.
- Katsikadelis, J. T., & Tsiatas, G. C. (2004). Non-linear dynamic analysis of beams with variable stiffness. *Journal of Sound and Vibration*, 270(4-5), 847-863.
- Ames, W. F. (2014). Numerical methods for partial differential equations. Academic press.
- Haghighat, E., Raissi, M., Moure, A., Gomez, H., & Juanes, R. (2020). A deep learning framework for solution and discovery in solid mechanics. *arXiv preprint arXiv:2003.02751*.
- Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.
- Ian Goodfellow, Yoshua Bengio and Aaron Courville. Deep Learning. MIT Press, 2016-11-18.

Neural Networks and Learning Machines Third Edition Simon Haykin McMaster University Hamilton,
Ontario, Canada

Helton, J. C., & Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1), 23-69.

Τεπετίδης, Ν. (2020). Διόδευση πλημμυρών με χρήση τεχνητών νευρωνικών δικτύων.

Ιωάννης, Φ. (2020). Βελτιστοποίηση αποτίμηση σεισμικής συμπεριφοράς υφισταμένων κατασκευών με σύγκλιση αποτελεσμάτων ιδιομορφικής ανάλυσης.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.

Bauchau, O. A., & Craig, J. I. (2009). Euler-Bernoulli beam theory. In *Structural analysis* (pp. 173-221). Springer, Dordrecht.

Heylen, W., Lammens, S., & Sas, P. (1997). *Modal analysis theory and testing* (Vol. 200, No. 7). Leuven, Belgium: Katholieke Universiteit Leuven.

