



**ΠΟΛΥΤΕΧΝΕΙΟ  
ΚΡΗΤΗΣ**  
TECHNICAL  
UNIVERSITY  
OF CRETE

Diploma Thesis

---

# Neural network on the assessment of danger of abdominal aortic aneurysms (AAA)

Νευρωνικό δίκτυο πρόβλεψης  
επικινδυνότητας ανευρύσματος  
κοιλιακής αορτής (ΑΚΑ)

---

*Author:*  
Kyriakos P. Siaterlis

*Committee:*  
Dr. Michalis Zervakis  
Dr. Georgios E. Stavroulakis  
Dr. Evangelos Kalogerakis

*A thesis submitted in fulfillment of the requirements  
for the integrated master's degree of Electrical and Computer Engineering  
in the*

[Department of Electrical and Computer Engineering](#)

August 30, 2025

“ σε εμάς,  
που μοιραστήκαμε ήλιους και φεγγάρια,  
άνθη και φύλλα μαραμένα,  
εισπνοές και εκπνοές,  
μαζί ”

Ιούλιος

# *Abstract*

Department of Electrical and Computer Engineering

## **Neural network on the assessment of danger of abdominal aortic aneurysms (AAA)**

Νευρωνικό δίκτυο πρόβλεψης επικινδυνότητας ανευρύσματος κοιλιακής αορτής (AKA)

by Kyriakos P. Siaterlis

In this thesis, we study the effectiveness of using neural networks to assess abdominal aortic aneurysms (AAA) risk by predicting maximum von Mises stress and maximum displacement at the walls of the aorta. The datasets, derived from finite-element method (FEM) simulations on geometric models reconstructed from MRI scans of real patients, come from two independent studies.

We develop and evaluate two multi-output regression architectures: (a) an architecture with specialized neurons in the hidden layers (separate pathways toward each output), and (b) a fully connected feedforward network with shared neurons in the hidden-layers. In addition, we examine output normalization due to the different orders of magnitude between the two outputs, aiming to counter the training bias toward the higher scale output, as well as feature-selection strategies.

Evaluation is performed with exhaustive leave-2-out cross validation, i.e., for  $n$  samples we train  $\binom{n}{2}$  independent models. Experiments are run using the AdamW and Levenberg-Marquardt (LM) optimizers.

Overall, AdamW exhibits more stable generalization on small/heterogeneous datasets like ours, whereas LM can offer faster convergence but is more prone to overfitting. Output normalization achieves its goal, and the architecture with specialized neurons is beneficial when it reduces interference between targets, meaning that its effectiveness depends on how the targets are produced by the FEM. Moreover, using targeted feature subsets yields significant improvements compared to using all available features, underscoring the importance of feature selection.

Results on both datasets show good target prediction given the number of samples relative to the spread of the target values we aim to predict. The results highlight the potential for accessible medical applications for AAA risk assessment, but more work is needed before it can be used for clinical decision-making.

## Περίληψη

Στην παρούσα διπλωματική εργασία μελετάται η αποτελεσματικότητα χρήσης νευρωνικών δικτύων για την εκτίμηση επικινδυνότητας ανευρύσματος κοιλιακής αορτής (ΑΚΑ), προβλέποντας μέγιστη τάση von Mises και μέγιστη μετατόπιση τοιχώματος. Τα σύνολα δεδομένων, που προκύπτουν από προσομοιώσεις πεπερασμένων στοιχείων (FEM) σε γεωμετρικά μοντέλα τα οποία έχουν ανακατασκευαστεί από MRI πραγματικών ασθενών, προέρχονται από δύο ανεξάρτητες εργασίες.

Αναπτύσσονται και αξιολογούνται δυο αρχιτεκτονικές πολλαπλών εξόδων (multi-output regression): (α) αρχιτεκτονική με εξειδικευμένους νευρώνες στα κρυφά επίπεδα (διαχωρισμός καναλιών προς κάθε έξοδο) και (β) πλήρως συνδεδεμένο feedforward δίκτυο με κοινούς νευρώνες στα κρυφά επίπεδα. Επιπρόσθετα, εξετάζεται η κανονικοποίηση τιμών στόχων (output normalization) λόγω διαφορετικών τάξεων μεγέθους μεταξύ των δύο τιμών εξόδου, με στόχο την αντιμετώπιση της εύνοιας της εξόδου με τη μεγαλύτερη κλίμακα κατά την εκπαίδευση, καθώς και στρατηγικές επιλογής χαρακτηριστικών.

Η αξιολόγηση γίνεται με εξαντλητική leave-2-out cross validation, δηλαδή για  $n$  δείγματα εκπαιδεύονται  $\binom{n}{2}$  ανεξάρτητα μοντέλα. Τα πειράματα τρέχουν με βελτιστοποιητές AdamW και Levenberg-Marquardt (LM).

Συνολικά, ο AdamW δείχνει πιο σταθερή γενίκευση σε μικρά/ετερογενή σύνολα όπως τα δικά μας, ενώ ο LM μπορεί να προσφέρει ταχύτερη σύγκλιση, αλλά είναι πιο επιρρεπής σε υπερπροσαρμογή. Η κανονικοποίηση τιμών στόχων πετυχαίνει το στόχο της και η αρχιτεκτονική με εξειδικευμένους νευρώνες ωφελεί όταν μειώνει τις παρεμβολές μεταξύ των στόχων, με την αποτελεσματικότητα της να εξαρτάται από τον τρόπο παραγωγής των στόχων από τη FEM. Επίσης, η χρήση στοχευμένων χαρακτηριστικών παρουσιάζει σημαντική βελτίωση σε σύγκριση με τη χρήση όλων των διαθέσιμων χαρακτηριστικών, αναδεικνύοντας τη σημασία της.

Τα αποτελέσματα και των δύο συνόλων δεδομένων παρουσιάζουν καλές προβλέψεις των στόχων, για το πλήθος των δειγμάτων σε σχέση με τη διασπορά στις τιμές στόχους που θέλουμε να προβλέψουμε. Τα αποτελέσματα αναδεικνύουν τη δυνατότητα για προσιτές ιατρικές εφαρμογές για εκτίμηση κινδύνου ΑΚΑ, αλλά απαιτείται ακόμη περισσότερη δουλειά ώστε να μπορεί να χρησιμοποιείται για τη λήψη κλινικών αποφάσεων.

## *Ευχαριστίες*

Όλα αυτά τα χρόνια που υπήρξα φοιτητής δεν υπάρχει λόγος και χώρος να εκφραστούν εδώ. Έχω περάσει πολύ γεμάτα χρόνια, πολύ γρήγορα χρόνια και πολύ αργά χρόνια. Έχω μάθει πολλά πράγματα για τον εαυτό μου, άλλους ανθρώπους και τον κόσμο και είμαι ευγνώμων που μου δόθηκε η ευκαιρία να το κάνω στο χρόνο μου.

Πάνω απ' όλα νιώθω ευγνώμων για τους όμορφους ανθρώπους που γνώρισα και αγαπώ και που μου στάθηκαν σαν δεύτερη οικογένεια στα φοιτητικά μου χρόνια και την οικογένεια μου που με στήριξε σε όλη τη διάρκεια αυτού του ταξιδιού.

Τέλος θέλω να ευχαριστήσω τους καθηγητές που έκαναν ενδιαφέρουσα τη φοίτηση στη σχολή που επέλεξα να σπουδάσω και με βοήθησαν καθ' όλη τη διάρκεια.

## Περιεχόμενα

Abstract.....	iii
Περίληψη .....	iv
Ευχαριστίες .....	v
Λίστα πινάκων .....	viii
Λίστα γραφημάτων.....	x
Λίστα εικόνων .....	xi
Εισαγωγή .....	1
Προσομοίωση .....	2
1.1 Διαδικασία προσομοίωσης .....	2
1.2 Μειονεκτήματα και παραδοχές προσομοίωσης .....	6
Νευρωνικά δίκτυα .....	7
2.1 Εισαγωγή.....	7
2.2 Συναρτήσεις ενεργοποίησης .....	8
2.2.1 Βηματικές συναρτήσεις.....	8
2.2.2 Γραμμικές συναρτήσεις .....	8
2.2.3 Σιγμοειδείς συναρτήσεις .....	9
2.2.4 Συνάρτηση ReLU .....	9
2.2.5 Συνάρτηση LReLU .....	9
2.2.6 Συνάρτηση PReLU .....	10
Νευρωνικό δίκτυο πρόβλεψης επικινδυνότητας ανευρύσματος .....	11
3.1 Δομή .....	11
3.2 Επεξεργασία δεδομένων .....	11
3.3 Υπερπροσαρμογή.....	12
3.4 Υπερπαράμετροι.....	12
3.4.1 Epoch .....	12
3.4.2 Learning Rate.....	12
3.4.3 Batch size.....	12
3.4.4 Patience .....	12
3.4.5 weight decay (L2 regularization) .....	12
3.5 Τρόπος εκπαίδευσης .....	13
3.6 Αλγόριθμοι βελτιστοποίησης(Optimizers) .....	15
Διαδικασία Πειραμάτων.....	17
4.1 Γενικά.....	17
4.2 Παραμετροποιήσεις .....	18

4.3 Μεθοδολογία αξιολόγησης.....	19
Αποτελέσματα μοντέλου με το πρώτο σύνολο δεδομένων .....	20
5.1 Εισαγωγή.....	20
5.2 Πρώτα πειραματικά αποτελέσματα μοντέλου 15 νευρώνων ανά κρυφό επίπεδο (κοινοί νευρώνες, με κανονικοποίηση στόχων) .....	21
5.2.1 Παρατηρήσεις αποτελεσμάτων .....	22
5.3 Περαιτέρω πειραματικά αποτελέσματα πάνω σε επιλογές χαρακτηριστικών που παρουσιάζουν καλές επιδόσεις .....	23
5.3.1 Παρατηρήσεις αποτελεσμάτων και γραφήματα .....	25
5.3.2 Αξιοσημείωτες περιπτώσεις .....	26
5.4 Πρώτα πειραματικά αποτελέσματα μοντέλου 25 νευρώνων ανά κρυφό επίπεδο (κοινοί νευρώνες, με κανονικοποίηση στόχων) .....	28
5.4.1 Παρατηρήσεις αποτελεσμάτων .....	29
5.5 Περαιτέρω πειραματικά αποτελέσματα πάνω σε επιλογές χαρακτηριστικών που παρουσιάζουν καλές επιδόσεις .....	30
5.5.1 Παρατηρήσεις αποτελεσμάτων και γραφήματα .....	32
5.5.2 Αξιοσημείωτες περιπτώσεις .....	34
Αποτελέσματα μοντέλου με το δεύτερο σύνολο δεδομένων .....	35
6.1 Εισαγωγή.....	35
6.2 Πρώτα πειραματικά αποτελέσματα μοντέλου 15 νευρώνων ανά κρυφό επίπεδο (κοινοί νευρώνες, με κανονικοποίηση στόχων) .....	36
6.2.1 Παρατηρήσεις αποτελεσμάτων .....	37
6.3 Περαιτέρω πειραματικά αποτελέσματα πάνω σε επιλογές χαρακτηριστικών που παρουσιάζουν καλές επιδόσεις .....	38
6.3.1 Παρατηρήσεις αποτελεσμάτων και γραφήματα .....	39
6.3.2 Αξιοσημείωτες περιπτώσεις .....	40
6.4 Πρώτα πειραματικά αποτελέσματα μοντέλου 25 νευρώνων ανά κρυφό επίπεδο (κοινοί νευρώνες, με κανονικοποίηση στόχων) .....	42
6.4.1 Παρατηρήσεις αποτελεσμάτων .....	43
6.5 Περαιτέρω πειραματικά αποτελέσματα πάνω σε επιλογές χαρακτηριστικών που παρουσιάζουν καλές επιδόσεις .....	44
6.5.1 Παρατηρήσεις αποτελεσμάτων και γραφήματα .....	46
6.5.2 Αξιοσημείωτες περιπτώσεις .....	47
Ανάλυση αποτελεσμάτων .....	49
Συμπεράσματα.....	51
Μελλοντική έρευνα.....	52
Βιβλιογραφία.....	53

## Λίστα πινάκων

Πίνακας 1 Διαφορών των υλοποιήσεων των προσομοιώσεων Ξυδιάς/Κυπαρίσσης .....	3
Πίνακας 2 Δεδομένων ασθενών .....	4
Πίνακας 3 Μέγιστες τιμές τάσης και μετατόπισης κατά Κυπαρίσση .....	5
Πίνακας 4 Μέγιστες τιμές τάσης και μετατόπισης κατά Ξυδιά .....	5
Πίνακας 5 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής $R^2$ ( $R^2$ score) για μοντέλα με 15 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή AdamW. Σύνολο δεδομένων Κυπαρίσση. ....	21
Πίνακας 6 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής $R^2$ ( $R^2$ score) για μοντέλα με 15 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή Levenberg-Marquardt. Σύνολο δεδομένων Κυπαρίσση. ....	21
Πίνακας 9 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση. ....	23
Πίνακας 10 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση. ....	24
Πίνακας 11 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Αιματοκρίτης, μήκος αυχένα, μέγιστη διάμετρος ανευρύσματος, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου, διάμετρος δεξιάς λαγονίου, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση. ....	24
Πίνακας 7 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής $R^2$ ( $R^2$ score) για μοντέλα με 25 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή AdamW. Σύνολο δεδομένων Κυπαρίσση. ....	28
Πίνακας 8 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής $R^2$ ( $R^2$ score) για μοντέλα με 25 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή Levenberg-Marquardt. Σύνολο δεδομένων Κυπαρίσση. ....	28
Πίνακας 12 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση. ....	30
Πίνακας 13 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση. ....	31
Πίνακας 14 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Μήκος αυχένα, Μήκος ανευρύσματος, Διάμετρος αριστερής λαγονίου και Διάμετρος δεξιάς λαγονίου, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση. ....	31
Πίνακας 15 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής $R^2$ ( $R^2$ score) για μοντέλα με 15 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή AdamW. Σύνολο δεδομένων Ξυδιά. ....	36



Πίνακας 16 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής $R^2$ ( $R^2$ score) για μοντέλα με 15 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή Levenberg-Marquardt. Σύνολο δεδομένων Ξυδιά. ....	36
Πίνακας 19 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.....	38
Πίνακας 20 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά. ....	39
Πίνακας 17 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής $R^2$ ( $R^2$ score) για μοντέλα με 25 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή AdamW. Σύνολο δεδομένων Ξυδιά.....	42
Πίνακας 18 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής $R^2$ ( $R^2$ score) για μοντέλα με 25 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή Levenberg-Marquardt. Σύνολο δεδομένων Ξυδιά. ....	42
Πίνακας 21 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: : Όλα τα χαρακτηριστικά, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.....	44
Πίνακας 22 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: : Διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.....	45
Πίνακας 23 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: : Ηλικία, διαγνωστική διάμετρος ΑΚΑ, διάμετρος αυχένα και διάμετρος δεξιάς λαγονίου, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.....	45

## Λίστα γραφημάτων

Γράφημα 1 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση. .	25
Γράφημα 2 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση. ....	26
Γράφημα 3 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, χωρίς κανονικοποίηση στόχων και με 16 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.....	26
Γράφημα 4 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, χωρίς κανονικοποίηση στόχων και με 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση. ....	27
Γράφημα 5 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση. .	32
Γράφημα 6 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση. ....	33
Γράφημα 7 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: μήκος αυχένα, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.....	33
Γράφημα 8 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, χωρίς κανονικοποίηση στόχων και με 26 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.....	34
Γράφημα 9 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, χωρίς κανονικοποίηση στόχων και με 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση. ....	34
Γράφημα 10 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW και με κανονικοποίηση στόχων και 16 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά. ....	40
Γράφημα 11 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: LM και με	

κανονικοποίηση στόχων και 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά. ....	40
Γράφημα 12 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW και χωρίς κανονικοποίηση στόχων και με 16 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά. ....	41
Γράφημα 13 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: LM και χωρίς κανονικοποίηση στόχων και με 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά. ....	41
Γράφημα 14 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW και με κανονικοποίηση στόχων και 26 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά. ....	46
Γράφημα 15 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: LM και με κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά. ....	47
Γράφημα 16 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW και χωρίς κανονικοποίηση στόχων και 26 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά. ....	47
Γράφημα 17 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: LM και χωρίς κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά. ....	48

## Λίστα εικόνων

Figure 1 Λειτουργία νευρώνα.....	8
Figure 2 Διάγραμμα ροής εκπαίδευσης .....	14

## Εισαγωγή

Τα τελευταία χρόνια παρατηρείται ραγδαία αύξηση στη χρήση του υπολογιστικού μοντέλου των νευρωνικών δικτύων, λόγω της ικανότητάς τους να προσεγγίζουν πολύπλοκες μη γραμμικές συναρτήσεις πίσω από δεδομένα, ακόμη και σε περιπτώσεις που δεν γνωρίζουμε κάποια μαθηματική περιγραφή [ [1] Ch. 1]. Τα νευρωνικά δίκτυα έχουν τη δυνατότητα να μαθαίνουν αφηρημένα μοτίβα από πληθώρα δεδομένων όπως εικόνες, σήματα, πίνακες τιμών κ.α.. Η εκπαίδευσή τους περιλαμβάνει δύο βασικά στάδια: το “forward pass”, στο οποίο το δίκτυο κάνει πρόβλεψη με βάση τα δεδομένα εισόδου και το “backward pass”, όπου ρυθμίζει κατάλληλα τις παραμέτρους του προς τη σωστή κατεύθυνση για να διορθώσει τα σφάλματα πρόβλεψης, βάσει του πραγματικού αποτελέσματος [ [1] Ch. 2].

Ένας από τους κλάδους που έχει να επωφεληθεί ιδιαίτερα από τα νευρωνικά δίκτυα είναι η ιατρική, καθώς μεγάλο μέρος της διάγνωσης και της λήψης αποφάσεων γίνεται με βάση την ανάλυση αποτελεσμάτων εξετάσεων. Τα δεδομένα αυτά μπορούν με σχετική ευκολία να χρησιμοποιηθούν για την εκπαίδευση νευρωνικών δικτύων, τα οποία στη συνέχεια θα λειτουργήσουν ως πρόσθετα εργαλεία για ειδικούς, που θα τους παρέχουν την ικανότητα να δώσουν μια πιο ακριβή διάγνωση ιδιαίτερα σε περιπτώσεις αβεβαιότητας [2]. Επιπλέον, λόγω της συστηματικής καταγραφής μη προσωποποιημένων δεδομένων ασθενών από τα νοσοκομεία, υπάρχει μεγάλος όγκος διαθέσιμων δεδομένων που μπορεί να αξιοποιηθεί για την δημιουργία μοντέλων νευρωνικών δικτύων με υψηλή ακρίβεια [3].

Ένα πολύ σημαντικό ιατρικό πρόβλημα με το οποίο θα ασχοληθούμε είναι τα ανευρύσματα και πιο συγκεκριμένα τα κοιλιακά ανευρύσματα της αορτής, τα οποία μπορούν να έρθουν σε ρήξη και να προκαλέσουν θάνατο. Μέχρι σήμερα δεν υπάρχει αξιόπιστη μη επεμβατική μέθοδος που να καθορίζει με σχετική ακρίβεια την επικινδυνότητα ενός ανευρύσματος να οδηγήσει σε ρήξη της αορτής.

Η εργασία μας μελετά τη δυνατότητα χρήσης νευρωνικών δικτύων ως εργαλείο εκτίμησης της επικινδυνότητας ανευρυσμάτων, βασισμένα σε δεδομένα προσομοίωσης αορτής για υπολογισμό μέγιστων τάσεων και μετατοπίσεων στα τοιχώματα. Βασίζεται σε ήδη υπάρχουσες μελέτες [2] [4] με προσομοίωση αορτών από πραγματικά ανώνυμα ιατρικά δεδομένα με στόχο τη βελτιστοποίηση της διαδικασίας, παρακάμπτοντας το ενδιάμεσο τμήμα της προσομοίωσης, έτσι ώστε να προβλέπουμε από δεδομένα που αποσπούμε από εικόνες κατευθείαν τις τάσεις και τις μετατοπίσεις στα τοιχώματα της αορτής, μέσω κατάλληλα εκπαιδευμένου μοντέλου.

## Κεφάλαιο 1

### Προσομοίωση

#### 1.1 Διαδικασία προσομοίωσης

Τα δεδομένα με τα οποία εκπαιδεύσαμε το νευρωνικό δίκτυο προέρχονται από τις διπλωματικές εργασίες με τίτλους “Προσομοίωση και ανάλυση αορτών με ανεύρυσμα με χρήση μεθόδου πεπερασμένων στοιχείων” του Κ. Κυπαρίσσης [4] και “Μελέτη μη-γραμμικών μοντέλων ανευρυσματικών αορτών με χρήση μεθόδου πεπερασμένων στοιχείων” του Χ. Ξυδιά [2], στις οποίες πραγματοποιήθηκαν προσομοιώσεις με δεδομένα από MRI απεικονίσεις ασθενών με ανεύρυσμα κοιλιακής αορτής (ΑΚΑ) (AAA in English). Οι εικόνες χρησιμοποιήθηκαν για την κατασκευή τρισδιάστατων γεωμετρικών μοντέλων της αορτής με σκοπό την ανάλυση των εσωτερικών καταπονήσεων του τοιχώματος της αορτής με χρήση της μεθόδου πεπερασμένων στοιχείων (FEM).

Σκοπός της ανάλυσης των προσομοιώσεων ήταν η εκτίμηση των μηχανικών μεγεθών της μέγιστης μετατόπισης και της μέγιστης τάσης, δηλαδή της τάσης στο σημείο της μέγιστης μετατόπισης, για την εκτίμηση του κινδύνου ρήξης της αορτής. Η μέγιστη τάση χρησιμοποιείται συχνά ως δείκτης επικινδυνότητας, καθώς η ρήξη θεωρείται ότι προκύπτει στις περιοχές όπου η καταπόνηση υπερβαίνει την αντοχή του ιστού [5].

Η διαδικασία της προσομοίωσης ακολουθεί την τυπική ροή της υπολογιστικής μηχανικής.

Αρχικά γίνεται **κατασκευή της γεωμετρίας της αορτής** από τις ιατρικές εικόνες MRI· δημιουργείται **πλέγμα(mesh)** και καθορίζονται οι περιοχές ενδιαφέροντος· **αναθέτονται οι ιδιότητες του υλικού**, στη περίπτωση αυτή του τοιχώματος της αορτής· **ορίζονται οι συνοριακές συνθήκες**, δηλαδή καθορίζονται οι σταθερές περιοχές το μοντέλου και εφαρμόζονται φορτία ή πιέσεις και τέλος **επιλύεται το πρόβλημα** με χρήση πεπερασμένων στοιχείων και εξάγονται τα μεγέθη ενδιαφέροντος [2]. Οι δύο εργασίες διαφέρουν στο τρόπο που αντιμετωπίζουν την μοντελοποίηση της αορτής. Ειδικότερα η εργασία του Ξυδιά [2] αντιμετωπίζει το πρόβλημα με τέσσερις διαφορετικούς τρόπους, αλλά στην εργασία μας θα εστιάσουμε στη μη γραμμική ανάλυση με πολυστρωματικά μοντέλα της εργασίας αυτής. Οι πιο σημαντικές ομοιότητες και διαφορές των δύο εργασιών παρουσιάζονται στον παρακάτω πίνακα:

	ΞΥΔΙΑΣ	ΚΥΠΑΡΙΣΣΗΣ
ΤΥΠΟΣ ΜΟΝΤΕΛΟΥ	Μέθοδος πεπερασμένων στοιχείων(FEM) με στοιχεία κελύφους	Μέθοδος πεπερασμένων στοιχείων(FEM) με στοιχεία κελύφους
ΣΤΟΙΧΕΙΑ ΓΕΩΜΕΤΡΙΑΣ	Πλέγμα τρισδιάστατων τριγωνικών στοιχείων στο κέλυφος. Αριθμός τριγώνων:10924	Πλέγμα τρισδιάστατων τριγωνικών στοιχείων στο κέλυφος. Αριθμός τριγώνων:12385
ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΟΙΧΩΜΑΤΟΣ	Δυο στρώματα ένα εσωτερικό(Media) και ένα εξωτερικό(Adventitia),	Μονοστρωματικό τοίχωμα
ΠΑΧΟΣ ΤΟΙΧΩΜΑΤΟΣ	Media:1.4 mm, Adventitia:0.7 mm	2.7mm
POISSON'S RATIO	Media:0.45, Adventitia:0.45	0.45
ΥΛΙΚΟ	Μη γραμμικό υπερελαστικό μοντέλο(Υεoh) με τιμές (Pa) Media: C1:192821333 C2:-10000 C3:3500 Adventitia: C1:64273833 C2:-5000 C3:2000	Γραμμικό ελαστικό ιστροπικό υλικό με Young Modulus: 700000 Pa
ΣΥΝΟΡΙΑΚΕΣ ΣΥΝΘΗΚΕΣ	Πάκτωση στα άνω και κάτω άκρα, ομοιόμορφη αρτηριακή πίεση	Πάκτωση στα άνω και κάτω άκρα, ομοιόμορφη αρτηριακή πίεση
ΣΥΣΤΟΛΙΚΗ ΠΙΕΣΗ(SAP)	15 kPa	20 kPa
ΠΥΚΝΟΤΗΤΑ ΤΟΙΧΩΜΑΤΟΣ	1060 kg/m <sup>3</sup>	1095 kg/m <sup>3</sup>

Πίνακας 1 Διαφορών των υλοποιήσεων των προσομοιώσεων Ξυδιάς/Κυπαρίσσης

Τα δεδομένα των ασθενών που προκύπτουν από τις ιατρικές εικόνες MRI φαίνονται στον παρακάτω πίνακα:

Αρίθμηση	Ασθενής	Φύλο	Ηλικία	Διαγνωστική διάμετρος ΑΚΑ (mm)	Αιματοκρίτης (Ht)	Διάμετρος αυχένα (mm)	Μήκος αυχένα (mm)	Μέγιστη διάμετρος ανευρύσματος (mm)	Μήκος ανευρύσματος (mm)	Διάμετρος αριστερής λαγονίου (mm)	Διάμετρος δεξιάς λαγονίου (mm)
1	T1-P2	M	81	48	46,1	21	30	33	69	11	10
2	T1-P4	M	89	75	37,4	24	24	73	78	12	14
3	T1-P5	M	72	51	41,8	22	46	48	74	9	11
4	T1-P6	M	66	58	39,9	30	47	44	68	22	20
5	T1-P8	M	63	48	45,2	25	36	43	80	13	13
6	T1-P9	M	57	68	44,5	21	46	45	83	16	20
7	T1-P10	M	57	48	42,5	21	18	33	101	20	15
8	T1-P11	M	70	50	43,5	24	33	34	64	10	10
9	T1-P12	M	75	52	42,7	19	10	52	55	11	11
10	T1-P14	M	81	50	41,4	20	55	36	67	15	12
11	T1-P16	M	76	55	48,9	26	35	55	81	13	15
12	T1-P18	M	65	52	41,9	24	15	39	135	13	12
13	T1-P20	F	69	62	40,5	21	30	33	69	11	10
14	T2-P3	M	78	50	44,7	24	24	73	78	12	14
15	T2-P4	M	66	61	49,7	29	46	48	74	9	11
16	T2-P13	M	62	51	51,7	22	47	44	68	22	20
17	T2-P17	M	67	50	42,4	25	36	43	80	13	13

Πίνακας 2 Δεδομένων ασθενών

Μετά την ολοκλήρωση της στατικής ανάλυσης της προσομοίωσης και σε συνδυασμό με τα φυσικά χαρακτηριστικά των ασθενών καταλήγουμε στις ακόλουθες τιμές στόχους από τις εργασίες του Κυπαρίσση και του Ξυδιά

Αρίθμηση	Μέγιστη τιμή τάσης (VM-MPa)	Μέγιστη μετατόπιση επιφάνειας (mm)
1	0,35	13
2		
3	0,37	9
4	0,6	24
5		
6	0,35	18
7	0,55	35
8	0,25	7
9		
10	0,45	21
11	0,4	18
12		
13		
14	0,7	30
15	0,3	14
16	0,3	14
17	0,4	14

αντίστοιχα:

Πίνακας 3 Μέγιστες τιμές τάσης και μετατόπισης κατά Κυπαρίσση

Αρίθμηση	Μέγιστη τιμή τάσης (VM-MPa)	Μέγιστη μετατόπιση επιφάνειας (mm)
1	0.97	0.143
2	1.4	1.271
3	0.5	0.075
4	4.63	0.271
5	0.96	0.596
6	0.55	0.192
7	26.14	0.452
8	26.57	0.069
9	8.63	0.192
10	0.89	0.241
11	0.8	0.167
12	2.54	0.428
13	37.28	0.21
14	0.59	0.176
15	1.16	0.882
16	0.64	0.169
17	74.66	0.20

Πίνακας 4 Μέγιστες τιμές τάσης και μετατόπισης κατά Ξυδιά



## 1.2 Μειονεκτήματα και παραδοχές προσομοίωσης

Η μέθοδος προσομοίωσης με χρήση πεπερασμένων στοιχείων προσφέρει μεγάλη ακρίβεια, αλλά υπάρχουν και μερικά μειονεκτήματα που προέρχονται από τις υποθέσεις που γίνονται.

Στην εργασία του Κυπαρίσση [4], το τοίχωμα προσεγγίζεται ως ιστροπικό και ομοιογενές, ενώ στην πραγματικότητα έχει σύνθετη δομή με τρία στρώματα και ανομοιογενή συμπεριφορά. Επίσης, δε λαμβάνεται υπόψη η παρουσία θρόμβων, που επηρεάζει τη κατανομή των τάσεων. Τέλος η πίεση που είναι σημαντικός παράγοντας προσεγγίζεται στατικά και όχι παλμικά.

Αντίθετα στην εργασία του Ξυδιά [2], το τοίχωμα προσεγγίζεται πολυστρωματικά, με διαφορετικά μηχανικά χαρακτηριστικά για κάθε στρώμα και υπερελαστική, μη γραμμική συμπεριφορά. Ωστόσο και σε αυτή την εργασία δεν λαμβάνεται υπόψη η παρουσία θρόμβων και η πίεση προσομοιώνεται ως στατική αντί για παλμική.

Παρόλα αυτά, οι προσομοιώσεις και των δυο εργασιών αποτελούν πολύ καλές μη επεμβατικές μεθόδους πρόβλεψης των τιμών ενδιαφέροντος οι οποίες θα χρησιμοποιηθούν για την εκπαίδευση του νευρωνικού μας δικτύου.

## Κεφάλαιο 2

### Νευρωνικά δίκτυα

#### 2.1 Εισαγωγή

Το νευρωνικό δίκτυο είναι ένα υπολογιστικό μοντέλο, του οποίου το όνομα είναι εμπνευσμένο από τη βιολογία [ [6], Ch.1] , γιατί θυμίζει τη λειτουργία των νευρώνων του εγκεφάλου [ [6] Ch.6]. Το νευρωνικό δίκτυο που υλοποιούμε αποτελείται από τρία στρώματα, το στρώμα εισόδου, ένα ή περισσότερα κρυφά στρώματα και το στρώμα εξόδου, στην οποία περίπτωση χαρακτηρίζεται “βαθύ νευρωνικό δίκτυο” [ [6] Ch.6].

Στο πρώτο στρώμα, το στρώμα εισόδου, έχουμε τους νευρώνες εισόδου, οι οποίοι βρίσκονται στην αρχή του μοντέλου όπου εισέρχεται η πληροφορία από τα δεδομένα μας για την εκπαίδευση και αργότερα για τα δεδομένα που θέλουμε να υπολογίσουμε μετά την εκπαίδευση. Το στρώμα εισόδου ακολουθείται από ένα ή περισσότερα κρυφά στρώματα, όπου γίνονται οι υπολογισμοί των δεδομένων εισόδου. Στο τελευταίο στρώμα, το στρώμα εξόδου, έχουμε τους νευρώνες εξόδου, οι οποίοι βρίσκονται στο τέλος του μοντέλου απ’ όπου παίρνουμε το τελικό αποτέλεσμα μετά τη διαδικασία υπολογισμών.

Οι νευρώνες του κάθε επιπέδου είναι πυκνά συνδεδεμένοι με τους νευρώνες των γειτονικών επιπέδων, με τρόπο ώστε κάθε νευρώνας του επόμενου επιπέδου παίρνει ως είσοδο την έξοδο όλων των νευρώνων των προηγούμενων επιπέδων [ [6] Ch.6]. Επίσης κάθε μια τέτοια σύνδεση μεταξύ νευρώνων εμπεριέχει και έναν αριθμό βαρύτητας(συναπτικός συντελεστής), ο οποίος ρυθμίζεται κατά την εκπαίδευση και πολλαπλασιάζεται με την αντίστοιχη είσοδο σε κάθε νευρώνα. Πέρα από τους συναπτικούς συντελεστές βαρύτητας υπάρχει και άλλος ένας αριθμός, που ονομάζεται πόλωση και προκύπτει μέσω της εκπαίδευσης, ο οποίος προστίθεται σε κάθε νευρώνα αφού όλες οι είσοδοι πολλαπλασιαστούν με τους συντελεστές τους και προστεθούν μεταξύ τους. Τέλος οι νευρώνες χαρακτηρίζονται και από τη συνάρτηση ενεργοποίησης τους, με την οποία θα ασχοληθούμε παρακάτω [ [6] Ch.6].

Η έξοδος ενός νευρώνα προκύπτει από τη συνάρτηση ενεργοποίησής του, της οποίας η είσοδος είναι το άθροισμα της πόλωσης με το σύνολο εισόδων του νευρώνα πολλαπλασιασμένων με τους αντίστοιχους συναπτικούς συντελεστές.

$$Output = f \times \left[ b + \sum_{i=0}^n (x_i \times w_i) \right]$$

Όπου *output* η έξοδος του νευρώνα, *f* η συνάρτηση ενεργοποίησης, *x* η είσοδος, *w* ο συναπτικός συντελεστής και *b* η πόλωση

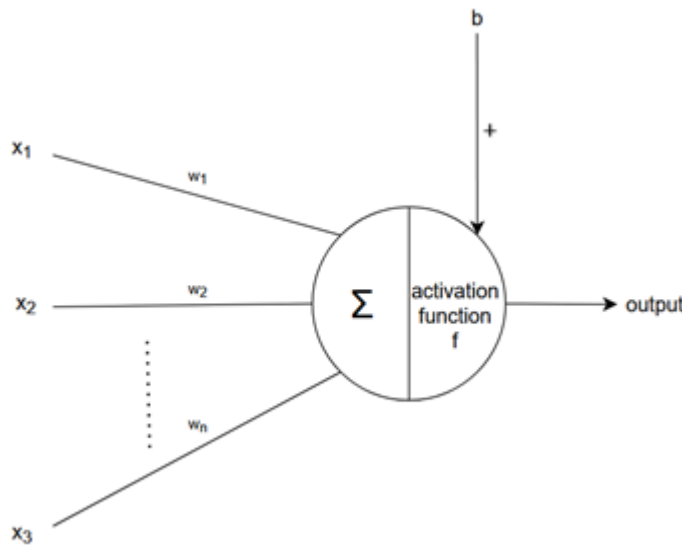


Figure 1 Λειτουργία νευρώνα

## 2.2 Συναρτήσεις ενεργοποίησης

### 2.2.1 Βηματικές συναρτήσεις

Ο σκοπός της βηματικής συνάρτησης είναι να προσομοιώσει την πυροδότηση ή μη πυροδότηση των νευρώνων με βάση τα δεδομένα που εισέρχονται στον νευρώνα. Η πιο απλή είναι η

$$f(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Ο νευρώνας θα πυροδοτήσει, εφόσον  $\sum_{i=0}^n (x_i \times w_i)$  έχει αποτέλεσμα μεγαλύτερο του 0. Οι συγκεκριμένες συναρτήσεις ενεργοποίησης σπάνια θα αποτελέσουν πια επιλογή [ [1] Ch.1].

### 2.2.2 Γραμμικές συναρτήσεις

Αποτυπώνονται σαν μια ευθεία γραμμή π.χ.  $F(x) = x$  όπου η έξοδος ισούται με την τιμή εισόδου. Το βασικό μειονέκτημά τους κρύβεται στην γραμμικότητα, καθώς όσα γραμμικά επίπεδα και αν έχει ένα νευρωνικό δίκτυο στο τέλος θα

ισοδυναμεί με γραμμικό μετασχηματισμό της εισόδου [ [6] Ch.6]. Εφαρμόζονται ακόμα, συνήθως στους νευρώνες του επιπέδου εξόδου.

### 2.2.3 Σιγμοειδείς συναρτήσεις

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

Αποτελεί την εναλλακτική που προτιμάται αντί για τις βηματικές συναρτήσεις επειδή είναι πιο πληροφοριακές. Αντί για μόνο 0 ή 1, επιστέφουν όλο το εύρος με 0 για είσοδο που πλησιάζει το αρνητικό άπειρο και 1 για είσοδο που πλησιάζει το θετικό άπειρο. Το εύρος τιμών της σιγμοειδούς συνάρτησης την καθιστά κατάλληλη σε περιπτώσεις δυαδικής κατηγοριοποίησης, καθώς η έξοδος μεταφράζεται σαν πιθανότητα. Το πρόβλημα με τις σιγμοειδείς συναρτήσεις εμφανίζεται για οριακές τιμές εισόδου. Σε αυτές τις περιοχές η παράγωγος τείνει στο 0 (vanishing gradient problem), με αποτέλεσμα κατά την εκπαίδευση πολλοί νευρώνες να ρυθμίζονται ελάχιστα ως καθόλου, οπότε η διαδικασία σύγκλισης του μοντέλου καθυστερεί. [ [1] Ch.1/2] [ [6] Ch.6/7] [ [7] Ch.5].

### 2.2.4 Συνάρτηση ReLU

Η Rectified Linear Unit (ReLU) ορίζεται ως:

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

Χάρη στη μη κορεσμένη, τμηματικά γραμμική μορφή της για θετικές εισόδους, μετριάζει το φαινόμενο vanishing gradient και διευκολύνει τη σύγκλιση σε βαθιά δίκτυα [8] [9].

Πολύ αποτελεσματική λόγω της ευκολίας στον υπολογισμό της και της μη γραμμικότητας της. Η μη γραμμικότητα της RELU είναι πολύ σημαντική καθώς επιτρέπει στο δίκτυο, με κατάλληλες παραμέτρους, να προσεγγίσει πολύπλοκες μη γραμμικές συναρτήσεις και η υπολογιστική απλότητα της την καθιστά πολύ αποδοτική [9], αλλά από την άλλη παρουσιάζει ένα σημαντικό μειονέκτημα. Κατά την εκπαίδευση για αρνητικές εισόδους, η έξοδος θα είναι 0 και η παράγωγος θα μηδενίζεται επίσης, με αποτέλεσμα πολλοί νευρώνες να μην εκπαιδεύονται (dying ReLU problem).

### 2.2.5 Συνάρτηση LReLU

Για την αντιμετώπιση του dying-ReLU προβλήματος, προτάθηκε η Leaky ReLU [10], η οποία ορίζεται ως:

$$f(x) = \begin{cases} x, & x > 0 \\ \alpha x, & x \leq 0 \end{cases}, \alpha \approx 0.01$$

Η LReLU επιτρέπει πολύ μικρή κλίση στο αρνητικό τμήμα, εξασφαλίζοντας ότι οι παράγωγοι δεν μηδενίζονται πλήρως και ότι οι νευρώνες παραμένουν ενεργοί κατά την εκπαίδευση.

### 2.2.6 Συνάρτηση PReLU

Μια ακόμα παραλλαγή της ReLU που προτάθηκε είναι η PReLU (Parametric ReLU) [11], η οποία γενικεύει τη Leaky ReLU και ορίζεται ως:

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

Σε αντίθεση με τη Leaky ReLU, όπου η κλίση για τις αρνητικές τιμές είναι σταθερή, στη PReLU η κλίση αυτή είναι εκπαιδεύσιμη παράμετρος (ανά νευρώνα) που ενημερώνεται μέσω του αλγορίθμου backpropagation. Αυτό δίνει μεγαλύτερη ευελιξία χωρίς σημαντικό υπολογιστικό κόστος και είναι ο λόγος που την επιλέξαμε.

## Κεφάλαιο 3

# Νευρωνικό δίκτυο πρόβλεψης επικινδυνότητας ανευρύσματος

### 3.1 Δομή

Η κατάλληλη επιλογή της δομής και των υπερπαραμέτρων ενός νευρωνικού δικτύου εξαρτάται από τα διαθέσιμα δεδομένα, τα χαρακτηριστικά εισόδου και τους στόχους εκπαίδευσης, και είναι μια πειραματική διαδικασία, καθώς δεν υπάρχει τρόπος να γνωρίζουμε προκαταβολικά τις κατάλληλες παραμέτρους. Στο νευρωνικό δίκτυο πρόβλεψης επικινδυνότητας ανευρύσματος που έχουμε εκπαιδεύσει πειραματιστήκαμε με διάφορες δομές και υπερπαραμέτρους μέχρι να καταλήξουμε σε αποδεκτά αποτελέσματα και να συλλέξουμε μετρήσεις. Οι διαφορετικές δομές και τα αποτελέσματά τους θα παρουσιαστούν στο κεφάλαιο μετρήσεων.

### 3.2 Επεξεργασία δεδομένων

Πριν ξεκινήσει η εκπαίδευση του δικτύου μας πρέπει να γίνει επεξεργασία των αρχείων με τα σετ δεδομένων. Η διαδικασία επιτυγχάνεται μέσω του προγράμματος που δημιουργήσαμε διαβάζοντας το αρχείο και φιλτράροντας τα δεδομένα που δεν είναι πλήρη για αφαίρεση. Περαιτέρω φιλτράρουμε τα δεδομένα ανάλογα με τα χαρακτηριστικά εισόδου που θέλουμε να κρατήσουμε για την εκπαίδευση και τις τιμές στόχους που θέλουμε να προβλέψουμε. Τέλος κανονικοποιούμε τις τιμές των χαρακτηριστικών με τη χρήση του `StandardScaler` και της συνάρτησης `fit_transform`, ώστε οι τιμές να ακολουθούν κατανομή με μέση τιμή 0 και τυπική απόκλιση 1 ( $N(0,1)$ ), ώστε να πετύχουμε ταχύτερη και πιο σταθερή σύγκλιση του δικτύου [ [6] Ch.6].

Μετά από τη διαδικασία επεξεργασίας και επιλογής δεδομένων, χωρίζουμε τα δεδομένα μας σε αυτά που θα χρησιμοποιηθούν για την εκπαίδευση του προγράμματος και τον έλεγχο του αντίστοιχα (train/test split). Επειδή έχουμε πολύ μικρό πλήθος δεδομένων, για να μην έχουμε μεγάλες αποκλίσεις σε διαφορετικούς κύκλους εκπαίδευσης, θα χρησιμοποιήσουμε επαναληπτική σχεδίαση (cross-validation) για αξιοπιστία, όπως θα εξηγήσουμε σε επόμενη ενότητα.

### 3.3 Υπερπροσαρμογή

Πριν παρουσιάσουμε τις υπερπαραμέτρους και τον τρόπο εκπαίδευσης, ορίζουμε συνοπτικά την υπερπροσαρμογή. Υπερπροσαρμογή (overfitting) συμβαίνει όταν το μοντέλο ταιριάζει υπερβολικά τις ιδιομορφίες και τον θόρυβο του συνόλου δεδομένων εκπαίδευσης, επιτυγχάνοντας πολύ χαμηλό σφάλμα στην εκπαίδευση αλλά υψηλό σφάλμα γενίκευσης σε νέα δεδομένα (μεγάλο train–test/generalization gap). Το φαινόμενο εμφανίζεται συνήθως όταν η πολυπλοκότητα/χωρητικότητα του μοντέλου είναι μεγάλη σε σχέση με το μέγεθος ή/και τον θόρυβο των δεδομένων. Συνήθεις στρατηγικές μετρίας περιλαμβάνουν την κανονικοποίηση (π.χ. L2, dropout), την έγκαιρη διακοπή (early stopping) και τη διασταυρούμενη επικύρωση (cross-validation). [7] [6]

### 3.4 Υπερπαραμέτροι

#### 3.4.1 Epoch

Η εποχή (epoch) είναι μια πλήρης διέλευση της διαδικασίας Forward Pass, Backward Pass, update(optimization) ολόκληρου του συνόλου των δεδομένων εκπαίδευσης. Με την υπερπαραμέτρο epochs, ορίζουμε το μέγιστο αριθμό που επιτρέπουμε στο δίκτυο μας να επαναλάβει τη διαδικασία. [6]

#### 3.4.2 Learning Rate

Ο ρυθμός μάθησης (learning rate) ορίζει το μέγεθος των βημάτων στις ενημερώσεις των συναπτικών συντελεστών(weights) από το βελτιστοποιητή σε μια εποχή. Ένας πολύ υψηλός ρυθμός μάθησης οδηγεί σε ταλάντωση γύρω από σημεία σύγκλισης ή ακόμα και σε πλήρη απόκλιση, ενώ ένας πολύ μικρός οδηγεί σε αργή σύγκλιση και απαιτεί υψηλό αριθμό εποχών. [ [6] Ch.8]

#### 3.4.3 Batch size

Το μέγεθος παρτίδας(batch size) είναι ο αριθμός δειγμάτων, κατά την εκπαίδευση, που τροφοδοτούνται στο μοντέλο πριν ενημερωθούν οι συναπτικοί συντελεστές(weights). Με τη χρήση μικρού μεγέθους επιτυγχάνεται καλύτερη γενίκευση, ενώ με μεγαλύτερα μεγέθη επιτυγχάνεται επιτάχυνση της διαδικασίας μάθησης [12] .

#### 3.4.4 Patience

Η υπομονή(patience) είναι ο αριθμός εποχών το μοντέλο θα συνεχίσει να εκπαιδεύεται χωρίς επιτυχία μικρότερου σφάλματος, μέχρι να διακόψουμε την εκπαίδευση. Κάθε φορά που σημειώνεται βελτίωση, επανεκκινούμε τον μετρητή. Η διαδικασία αυτή ονομάζεται early stopping και με τη χρήση της αποφεύγουμε την υπερπροσαρμογή(overfitting) [13].

#### 3.4.5 weight decay (L2 regularization)

Ο όρος της κανονικοποίησης L2 είναι ο αριθμός με τον οποίο πολλαπλασιάζεται το άθροισμα των τετραγώνων των συναπτικών

συντελεστών(weights). Το άθροισμα προστίθεται στη συνάρτηση σφάλματος, με σκοπό τη σταδιακή συρρίκνωση των συναπτικών συντελεστών(weights), ώστε να αποθαρρύνουμε το μοντέλο από την εξάρτηση σε μεμονωμένους νευρώνες στη διαδικασία επεξεργασίας, που οδηγεί σε αποτυχία γενίκευσης και υπερπροσαρμογή στα δεδομένα εκπαίδευσης [6] [14] .

### 3.5 Τρόπος εκπαίδευσης

Ο τρόπος που εκπαιδεύεται το νευρωνικό δίκτυο χωρίζεται σε τμήματα που επαναλαμβάνονται. Τα δεδομένα εκπαίδευσης χωρίζονται σε ομάδες (batches), τα οποία περνάνε μέσα από το νευρωνικό δίκτυο και παράγονται οι έξοδοι. Η διαδικασία αυτή ονομάζεται Forward pass [6]. Στη συνέχεια υπολογίζεται το σφάλμα του κάθε batch μέσω της συνάρτησης σφάλματος (loss function), στην συγκεκριμένη περίπτωση Μέσο Τετραγωνικό Σφάλμα (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \text{ όπου } y_i \text{ οι προβλέψεις και } \hat{y}_i \text{ οι πραγματικές τιμές}$$

,που ουσιαστικά συγκρίνει την απόκλιση των εξόδων του batch με τις πραγματικές τιμές, υπολογίζει το τετράγωνο της και διαιρεί με το batch size. Τέλος ανανεώνονται οι συναπτικοί συντελεστές και η πόλωση, με σκοπό τη μείωση του σφάλματος, ξεκινώντας από το επίπεδο εξόδου προχωρώντας προς το επίπεδο εισόδου με τη χρήση του αλγορίθμου backpropagation. Ο αλγόριθμος υπολογίζει τις παραγώγους του σφάλματος ως προς τους συναπτικούς συντελεστές μέσω του κανόνα αλυσίδας (chain rule) και διαδίδει το σφάλμα προς τα πίσω, ώστε να ανανεωθούν τα βάρη προς τη σωστή κατεύθυνση [7]. Η διαδικασία αυτή ονομάζεται Backward pass. Αυτά τα βήματα επαναλαμβάνονται μέχρι να γίνει πλήρης διέλευση από όλα τα batches και συνολικά ονομάζεται εποχή (epoch) [6]. Το τέλος της διαδικασίας εκπαίδευσης το ορίζουμε με την υπερπαραμέτρο του μέγιστου αριθμού εποχών, είτε νωρίτερα με τη σχεδιαστική μέθοδο που ονομάζεται early stopping. [13]

Το early stopping είναι ουσιαστικά η διακοπή της εκπαίδευσης πριν φτάσουμε τον μέγιστο αριθμό επαναλήψεων, epochs, που έχουμε ορίσει. Κάθε φορά στο τέλος ενός epoch υπολογίζουμε το σφάλμα του νευρωνικού δικτύου με τα δεδομένα ελέγχου για να δούμε αν η συνέχιση της εκπαίδευσης παράγει καλύτερα αποτελέσματα. Αν δεν βλέπουμε καλύτερα αποτελέσματα μετά από έναν αριθμό epochs που έχουμε ορίσει ως ανοχή(patience), τότε σταματάμε την εκπαίδευση και αποφύγουμε υπερπροσαρμογή πριν το δίκτυο να αρχίσει να μαθαίνει θόρυβο [13].

Παράλληλα με αυτή τη λειτουργία χρησιμοποιούμε και τη σχεδιαστική μέθοδο των checkpoints. Όταν υπολογίζουμε μικρότερο σφάλμα από το προηγούμενο



μικρότερο, που συγκρατούμε για να μετρήσουμε αν ξεπερνάμε την ανοχή, αποθηκεύουμε την κατάσταση του νευρωνικού δικτύου που πέτυχε καλύτερη απόδοση, έτσι ώστε αν ενεργοποιηθεί το early stopping να το επαναφέρουμε στην καλύτερη απόδοσή του [15].

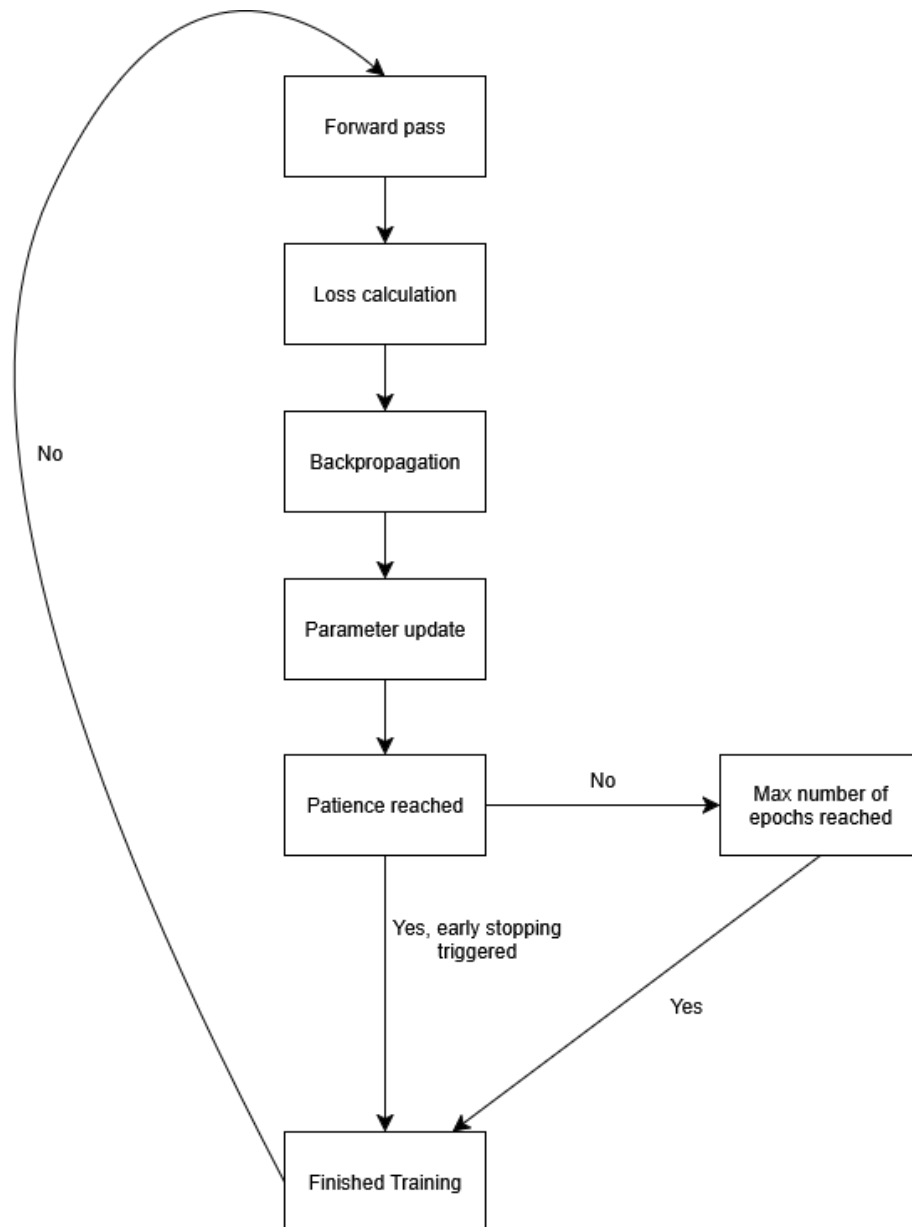


Figure 2 Διάγραμμα ροής εκπαίδευσης

### 3.6 Αλγόριθμοι βελτιστοποίησης(Optimizers)

Οι αλγόριθμοι βελτιστοποίησης (optimizers) είναι υπεύθυνοι για την ενημέρωση των παραμέτρων του νευρωνικού δικτύου κατά τη διάρκεια της εκπαίδευσης με στόχο την ελαχιστοποίηση της συνάρτησης σφάλματος (loss function) [6]. Στο νευρωνικό δίκτυο που ερευνούμε πειραματιστήκαμε και πήραμε μετρήσεις με δύο διαφορετικούς αλγόριθμους βελτιστοποίησης που καθένας έχει τα δικά του προτερήματα.

Ο πρώτος αλγόριθμος βελτιστοποίησης που χρησιμοποιήθηκε είναι ο AdamW (Adaptive Moment Estimation *with decoupled weight decay*) [16] [17], ο οποίος συνδυάζει τα πλεονεκτήματα των τεχνικών Momentum και RMS Propagation (RMSprop) για την ενημέρωση των παραμέτρων των νευρώνων. Ο Adam υπολογίζει για κάθε παράμετρο του δικτύου έναν προσαρμοστικό ρυθμό μάθησης (learning rate) σε κάθε βήμα ενημέρωσης. Ο υπολογισμός του ρυθμού μάθησης βασίζεται στην ταυτόχρονη εκτίμηση της πρώτης ροπής (first moment), που είναι ο μέσος όρος των προηγούμενων κλίσεων (gradients), και της δεύτερης ροπής (second moment), που είναι το μέσο τετράγωνο των κλίσεων (uncentered variance) . Με τη χρήση αυτών των μεγεθών προσαρμόζει δυναμικά το ρυθμό μάθησης για κάθε βάρος, επιτυγχάνοντας μείωση των ταλαντώσεων και της πιθανότητας παγίδευσης σε τοπικά ελάχιστα κατά την εκπαίδευση. Σύμφωνα με τους Kingma & Ba [16], ο Adam προσφέρει απλότητα στην εφαρμογή του, καθώς έχει λιγότερες υπερπαραμέτρους σε σύγκριση με άλλους αλγόριθμους βελτιστοποίησης, υπολογιστική αποδοτικότητα με μικρές απαιτήσεις μνήμης και καλή απόδοση σε προβλήματα με μεγάλα σύνολα δεδομένων ή πολλά χαρακτηριστικά. Η διαφορά του AdamW από τον Adam είναι ότι εφαρμόζει weight decay μετά την ενημέρωση των κλίσεων, ανεξάρτητα από την προσαρμογή του ρυθμού μάθησης, εξασφαλίζοντας σταθερότερη κανονικοποίηση και καλύτερη γενίκευση. Σύμφωνα με τους Loshchilov & Hutter [17], η προσέγγιση αυτή κάνει πιο προβλέψιμη την επίδραση του weight decay και την καθιστά ανεξάρτητη σε σχέση με τις άλλες υπερπαραμέτρους.

Ο δεύτερος αλγόριθμος βελτιστοποίησης που χρησιμοποιήθηκε είναι ο Levenberg-Marquardt (LM) [18] [19], ο οποίος επιλύει μη γραμμικά προβλήματα ελαχίστων τετραγώνων. Ο αλγόριθμος LM είναι ιδιαίτερα κατάλληλος για μικρά νευρωνικά δίκτυα με λίγα χαρακτηριστικά, προσφέροντας ταχύτητα στη σύγκλιση και συχνά καλύτερη προσαρμογή συγκριτικά με μεθόδους πρώτης τάξης [20]. Ο LM πετυχαίνει μια ισορροπία συνδυάζοντας τη μέθοδο Gauss-Newton, για πιο γρήγορη σύγκλιση, με μια προσαρμοστική απόσβεση για σταθερότητα, αποφεύγοντας τα προβλήματα που μπορούν να προκύψουν από την καθαρή μέθοδο Gauss-Newton. Η βασική του εξίσωση είναι:

$$(J^T J + \lambda I) \Delta w = -J^T e$$

όπου  $J$  είναι ο Jacobian πίνακας,  $\lambda$  ο συντελεστής απόσβεσης,  $I$  ο μοναδιαίος πίνακας,  $\Delta w$  οι ανανεώσεις στο διάνυσμα των συναπτικών συντελεστών και  $e$  το διάνυσμα σφαλμάτων.

Τον LM μπορούμε να τον φανταστούμε σαν Gradient Descent που χρησιμοποιεί την κλίση της συνάρτησης σφάλματος για την ενημέρωση των παραμέτρων, όμως, καθώς πλησιάζει στη λύση συμπεριφέρεται όλο και περισσότερο σαν Gauss-Newton χρησιμοποιώντας προσέγγιση δεύτερης τάξης για πιο ακριβή ενημέρωση παραμέτρων. Όσο μειώνεται το σφάλμα τόσο μειώνεται και το  $\lambda$ , οι ενημερώσεις γίνονται πιο έντονες και ο αλγόριθμος μοιάζει περισσότερο με Gauss-Newton. Όσο αυξάνεται το σφάλμα, αυξάνεται και το  $\lambda$ , οι ενημερώσεις γίνονται πιο προσεγμένες και ο αλγόριθμος μοιάζει περισσότερο με Gradient Descent. Επιπλέον, έχουμε προσθέσει early stopping για να περιορίσουμε την υπερπροσαρμογή, όπως περιγράψαμε παραπάνω.

## Κεφάλαιο 4

### Διαδικασία Πειραμάτων

#### 4.1 Γενικά

Στο πρόγραμμα μας εκπαιδεύουμε το μοντέλο με τους βελτιστοποιητές AdamW και Levenberg-Marquardt (LM), αλλά όχι μόνο μια φορά για τον καθένα. Η συνηθισμένη πρακτική για την αξιολόγηση της απόδοσης ενός μοντέλου είναι η χρήση k-fold cross validation [21] [6] κατά την οποία χωρίζουμε το σύνολο των δεδομένων σε k υποσύνολα (folds). Το μοντέλο εκπαιδεύεται με τα k-1 και ελέγχεται με το υποσύνολο που εξαιρείται από την εκπαίδευση. Αυτή η διαδικασία επαναλαμβάνεται k φορές, κάθε φορά χρησιμοποιώντας διαφορετικό υποσύνολο για τον έλεγχο, και υπολογίζεται η μέση απόδοση για την επικύρωση των αποτελεσμάτων.

Στη δική μας περίπτωση που έχουμε πολύ μικρό πλήθος συνόλου δεδομένων εφαρμόσαμε μια εξαντλητική leave-p-out cross validation [22] με p=2 (leave-2-out). Δηλαδή, τρέχουμε μια εξαντλητική επανάληψή πάνω στους συνδυασμούς που μπορούμε να δημιουργήσουμε αφήνοντας δύο (2) δείγματα για έλεγχο. Με αυτόν το τρόπο αξιολογούμε όλους τους πιθανούς συνδυασμούς, το πλήθος των οποίων δίνεται από τον τύπο:

$$C(n, r) = \binom{n}{r} = \frac{n!}{(r! (n - r)!)}$$

, όπου n αντικείμενα(objects) και r δείγμα(sample)

Για παράδειγμα στην περίπτωση που έχουμε 12 δεδομένα από την εργασία του Κυπαρίσση

$$\binom{12}{2} = 66$$

θα πραγματοποιήσουμε 66 επαναλήψεις για κάθε βελτιστοποιητή. Σε κάθε επανάληψη το μοντέλο εκπαιδεύεται πλήρως από την αρχή.

Για την αξιολόγηση της απόδοσης συλλέγουμε τις προβλεπόμενες και τις πραγματικές τιμές από όλες τις επαναλήψεις, οπότε μπορούμε να υπολογίσουμε το μέσο απόλυτο σφάλμα (MAE), δηλαδή το μέσο όρο της απόλυτης διαφοράς προβλεπόμενης-πραγματικής τιμής

$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$ , όπου  $y_i$  οι προβλέψεις και  $\hat{y}_i$  οι πραγματικές τιμές,

το μέσο τετραγωνικό σφάλμα (MSE), δηλαδή το μέσο όρο των τετραγώνων της διαφοράς προβλεπόμενης-πραγματικής τιμής

$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ , όπου  $y_i$  οι προβλέψεις και  $\hat{y}_i$  οι πραγματικές τιμές,

και το συντελεστή προσδιορισμού  $R^2$

$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ , όπου  $y_i$  οι προβλέψεις,  $\hat{y}_i$  οι πραγματικές τιμές και  $\bar{y}$  η συνολική μέση τιμή των πραγματικών τιμών,

δείκτης που μετρά πόσο καλά οι προβλεπόμενες τιμές προσεγγίζουν τις πραγματικές, καθώς παίρνει τιμή 1 για τέλειες προβλέψεις, 0 για προβλέψεις όσο καλές όσο η μέση τιμή, και αρνητικές τιμές για χειρότερες προβλέψεις από τη μέση τιμή. Επιπλέον σχεδιάζουμε διαγράμματα προβλέψεων έναντι πραγματικών τιμών, για να απεικονίσουμε γραφικά την ακρίβεια του μοντέλου.

## 4.2 Παραμετροποιήσεις

Το πρόγραμμα που αναπτύξαμε υποστηρίζει δυο διαφορετικές αρχιτεκτονικές εκπαίδευσης, οι οποίες επιλέγονται από το μενού έναρξης. Η πρώτη είναι ένα κλασικό fully connected feedforward νευρωνικό δίκτυο, στο οποίο κάθε νευρώνας συνδέεται με όλους τους νευρώνες των γειτονικών επιπέδων (fully connected) και η διάδοση της πληροφορίας γίνεται αποκλειστικά με κατεύθυνση από το επίπεδο εισόδου προς το επίπεδο εξόδου (feedforward). Το δίκτυο διαθέτει 1-9 νευρώνες εισόδου, ανάλογα με το πλήθος των χαρακτηριστικών που έχουν επιλεγεί, τους νευρώνες στα ενδιάμεσα επίπεδα, έχουμε πειραματιστεί με 15 και 25 νευρώνες ανά κρυφό επίπεδο και 2 νευρώνες εξόδου, έναν για κάθε τιμή πρόβλεψης. Η διάταξη αντιστοιχεί σε multi-output regression [ [6] Ch.6].

Στη δεύτερη αρχιτεκτονική το δίκτυο περιγράφεται από κοινή είσοδο, ενώ στα ενδιάμεσα επίπεδα οι νευρώνες χωρίζονται σε δύο ομάδες με ίσο πλήθος νευρώνων. Η πρώτη ομάδα εκπαιδεύεται για τη πρόβλεψη της μέγιστης τιμής τάσης, ενώ η δεύτερη για τη πρόβλεψη της μέγιστης τιμής μετατόπισης. Η λογική αυτής της αρχιτεκτονικής είναι η εκπαίδευση εξειδικευμένων νευρώνων για κάθε έξοδο, ώστε να αποφευχθεί η παρεμβολή όταν οι κλίσεις των δύο εξόδων κατευθύνουν τον βελτιστοποιητή προς διαφορετικές κατευθύνσεις [23].

Επιπλέον, έχει ενσωματωθεί δυνατότητα κανονικοποίησης των τιμών στόχων (*output normalization*) [ [6] Ch. 7] . Οι δύο στόχοι έχουν σημαντικά διαφορετικές τάξεις μεγέθους που μπορεί να οδηγήσει το βελτιστοποιητή να επικεντρωθεί στην ελαχιστοποίηση του σφάλματος της εξόδου με τη μεγαλύτερη κλίμακα.

Στη μελέτη μας συλλέξαμε δεδομένα και για τις δύο αρχιτεκτονικές, με και χωρίς κανονικοποίηση, ώστε να αξιολογηθεί η επίδραση της στην απόδοση του εκάστοτε μοντέλου.

### 4.3 Μεθοδολογία αξιολόγησης

Αρχικά, τα μοντέλα δοκιμάστηκαν με αρχιτεκτονική κοινών νευρώνων και κανονικοποίηση τιμών στόχων, χρησιμοποιώντας 15 και 25 νευρώνες ανά κρυφό επίπεδο. Η αξιολόγηση έγινε με δείκτες το Μέσο απόλυτο σφάλμα (MAE) και τον συντελεστή  $R^2$  με την εξής διαδικασία:

- χρήση όλων των διαθέσιμων χαρακτηριστικών ως είσοδοι,
- χρήση κάθε μεμονωμένου χαρακτηριστικού ως είσοδος,
- επιλογή των δύο χαρακτηριστικών εισόδου με την καλύτερη ατομική επίδοση, με βάση τους δείκτες αξιολόγησης, και σταδιακή προσθήκη των επόμενων δυο χαρακτηριστικών κάθε φορά -δηλαδή 2, 4, 6, 8 συνολικά χαρακτηριστικά- μέχρι και τα 8 από τα 9 διαθέσιμα.

Για το βήμα της σταδιακής προσθήκης χαρακτηριστικών, η επιλογή των χαρακτηριστικών έγινε ξεχωριστά για κάθε βελτιστοποιητή, AdamW και Levenberg-Marquardt (LM). Επιπλέον, κάθε σύνολο χαρακτηριστικών που προκύπτει για έναν βελτιστοποιητή εφαρμόστηκε και στον άλλον, ώστε να εξεταστεί τόσο η βελτιστοποίηση των χαρακτηριστικών εισόδου όσο και η απόδοση τους πάνω σε αυτά τα σύνολα χαρακτηριστικών. Συγκεκριμένα για κάθε πλήθος χαρακτηριστικών (2, 4, 6, 8), εκπαιδεύονται δύο μοντέλα, το ένα με τα χαρακτηριστικά που είχαν την καλύτερη ατομική απόδοση με AdamW, και το άλλο αντίστοιχα με LM.

## Κεφάλαιο 5

### Αποτελέσματα μοντέλου με το πρώτο σύνολο δεδομένων

#### 5.1 Εισαγωγή

Το πρώτο σύνολο δεδομένων που χρησιμοποιήθηκε προέρχεται από την εργασία του Κυπαρίσση και περιλαμβάνει 12 δείγματα. Η αξιολόγηση των μοντέλων πραγματοποιήθηκε με εξαντλητικό leave-2-out cross validation, όπως αναφέρθηκε στην προηγούμενη ενότητα, και συνολικά έγιναν 66 επαναλήψεις ανά μοντέλο - μια για κάθε δυνατό μοναδικό συνδυασμό.

Στους πίνακες που ακολουθούν παρουσιάζονται τα αποτελέσματα της παραπάνω διαδικασίας. Στην αριστερή στήλη κάθε πίνακα αναγράφονται ή αριθμούνται τα χαρακτηριστικά που αντιστοιχούν σε κάθε γραμμή αποτελεσμάτων. Σε περιπτώσεις συνδυασμών χαρακτηριστικών, για κάθε πλήθος χαρακτηριστικών εμφανίζονται δύο γραμμές. Η πρώτη γραμμή αντιστοιχεί στον συνδυασμό που επιλέχθηκε με AdamW και η δεύτερη στον συνδυασμό που επιλέχθηκε με LM.

## 5.2 Πρώτα πειραματικά αποτελέσματα μοντέλου 15 νευρώνων ανά κρυφό επίπεδο (κοινοί νευρώνες, με κανονικοποίηση στόχων)

	Μέγιστη VM τάση (MPa)		Μέγιστη μετατόπιση (mm)	
	MAE	$R^2$ score	MAE	$R^2$ score
Όλα τα χαρακτηριστικά	0,0621	0,567	3,2251	0,715
1.Ηλικία	0,0883	0,213	5,4186	0,2664
2.Διαγνωστική διάμετρος ΑΚΑ (mm)	0,0896	0,1904	5,8818	0,0891
3.Αιματοκρίτης (Ht)	0,0818	0,2567	5,6137	0,1487
4.Διάμετρος αυχένα (mm)	0,0925	0,1171	5,7916	0,1372
5.Μήκος αυχένα (mm)	0,0861	0,2291	4,824	0,3538
6.Μέγιστη διάμετρος ανευρύσματος (mm)	0,0872	0,2256	5,4461	0,2169
7.Μήκος ανευρύσματος (mm)	0,0848	0,2799	5,3267	0,2602
8.Διάμετρος αριστερής λαγονίου (mm)	0,0633	0,4676	3,0496	0,6974
9.Διάμετρος δεξιάς λαγονίου (mm)	0,0809	0,2845	4,5722	0,4062
8,9	0,0568	0,5899	3,0357	0,723
6,8	0,0631	0,4937	3,3924	0,6516
5,7,8,9	0,0644	0,5376	3,1286	0,6894
3,6,8,9	0,063	0,5004	3,1829	0,6737
1,3,5,7,8,9	0,0573	0,5803	3,4261	0,6405
3,5,6,7,8,9	0,0616	0,5949	3,3115	0,6984
1,2,3,5,6,7,8,9	0,0664	0,4907	3,4347	0,6286
1,2,3,5,6,7,8,9	0,0664	0,4907	3,4347	0,6286

Πίνακας 5 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής  $R^2$  ( $R^2$  score) για μοντέλα με 15 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή AdamW. Σύνολο δεδομένων Κυτταρίση.

	Μέγιστη VM τάση (MPa)		Μέγιστη μετατόπιση (mm)	
	MAE	$R^2$ score	MAE	$R^2$ score
Όλα τα χαρακτηριστικά	0,0793	0,2729	4,3487	0,4846
1.Ηλικία	0,1113	-0,1488	5,9742	0,0193
2.Διαγνωστική διάμετρος ΑΚΑ (mm)	0,1087	-0,1263	6,8104	-0,138
3.Αιματοκρίτης (Ht)	0,0941	0,0314	6,3573	-0,0989
4.Διάμετρος αυχένα (mm)	0,112	-0,1682	6,8634	-0,1582
5.Μήκος αυχένα (mm)	0,1086	-0,1031	6,2075	0,1194
6.Μέγιστη διάμετρος ανευρύσματος (mm)	0,0977	0,0935	5,8021	0,025
7.Μήκος ανευρύσματος (mm)	0,1035	-0,0774	6,2147	0,0818
8.Διάμετρος αριστερής λαγονίου (mm)	0,0766	0,3202	3,9808	0,527
9.Διάμετρος δεξιάς λαγονίου (mm)	0,0987	0,0045	5,5017	0,1966
8,9	0,0863	0,1963	3,7104	0,5965
6,8	0,0861	0,183	4,0559	0,5339
5,7,8,9	0,0902	0,1593	4,5043	0,4878
3,6,8,9	0,078	0,3071	4,2992	0,4972
1,3,5,7,8,9	0,0723	0,4546	4,6079	0,3833
3,5,6,7,8,9	0,0933	0,1311	4,6479	0,4602
1,2,3,5,6,7,8,9	0,0748	0,3915	4,1983	0,5447
1,2,3,5,6,7,8,9	0,0748	0,3915	4,1983	0,5447

Πίνακας 6 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής  $R^2$  ( $R^2$  score) για μοντέλα με 15 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή Levenberg-Marquardt. Σύνολο δεδομένων Κυτταρίση.



### 5.2.1 Παρατηρήσεις αποτελεσμάτων

- **Μέγιστη τάση von Mises (MPa)**
  - AdamW: **MAE 0.0568-0.0664** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
  - LM: **MAE 0.0723-0.0902** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
- **Μέγιστη μετατόπιση (mm)**
  - AdamW: **MAE 3.0357-3.4347** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
  - LM: **MAE 3.7104-4.6479** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
- **$R^2$** : Θετικό και γενικά υψηλότερο για τη μέγιστη μετατόπιση, σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
- **Συνδυασμοί χαρακτηριστικών (2/4/6/8)**
  - **AdamW**: καλύτερη επίδοση με τις επιλογές του για πλήθος χαρακτηριστικών 2 και 4, ενώ για 6 χαρακτηριστικά είχε καλύτερες επιδόσεις με το σύνολο που επιλέχθηκε μέσω LM· για πλήθος 8 χαρακτηριστικών συνέπεσαν οι επιλογές.
  - **LM**: καλύτερη επίδοση με τις επιλογές του για πλήθος χαρακτηριστικών 2 και 4, ενώ για 6 χαρακτηριστικά είχε καλύτερες επιδόσεις με το σύνολο που επιλέχθηκε μέσω AdamW για τη πρόβλεψη της μέγιστης τάσης VM, αλλά καλύτερες επιδόσεις με το σύνολο του για τη πρόβλεψη της μέγιστης μετατόπισης

### 5.3 Περαιτέρω πειραματικά αποτελέσματα πάνω σε επιλογές χαρακτηριστικών που παρουσιάζουν καλές επιδόσεις

Στη συνέχεια θα επιλέξουμε τα καλύτερα μεικτά αποτελέσματα, δηλαδή αυτά με καλή επίδοση και στις δύο τιμές στόχους, ώστε να τα αναλύσουμε περαιτέρω. Συγκεκριμένα θα συγκρίνουμε την αρχιτεκτονική με κοινούς νευρώνες με εκείνη των εξειδικευμένων νευρώνων, με και χωρίς κανονικοποίηση των τιμών στόχων.

Για το μοντέλο με τους **15 νευρώνες ανά κρυφό επίπεδο** οι καλύτερες επιδόσεις προέκυψαν με τις εξής επιλογές χαρακτηριστικών εισόδου:

1. **Διάμετρος αριστερής λαγονίου, διάμετρος δεξιάς λαγονίου.**
2. **Αιματοκρίτης, μήκος αυχένα, μέγιστη διάμετρος ανευρύσματος, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου, διάμετρος δεξιάς λαγονίου.**

Παραθέτουμε τα αποτελέσματα για τα παραπάνω χαρακτηριστικά μαζί με τα αντίστοιχα αποτελέσματα με επιλογή όλα τα χαρακτηριστικά για εισόδους ως σημείο αναφοράς, ώστε να φαίνεται καθαρά το κέρδος από την επιλογή χαρακτηριστικών.

		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
χωρίς κανονικοποίηση τιμών στόχων	εξειδικευμένοι νευρώνες	0,3536 <b>-30,0511</b>	4,0023 <b>0,5101</b>	0,5293 <b>-34,0024</b>	6,9498 <b>-0,3201</b>
	κοινοί νευρώνες	0,9303 <b>-114,9819</b>	3,918 <b>0,5395</b>	0,2393 <b>-11,6572</b>	4,5564 <b>0,4281</b>
με κανονικοποίηση τιμών στόχων	εξειδικευμένοι νευρώνες	0,0776 <b>0,3797</b>	4,7521 <b>0,4242</b>	0,1263 <b>-0,5145</b>	7,0519 <b>-0,2395</b>
	κοινοί νευρώνες	0,0621 <b>0,567</b>	3,2251 <b>0,715</b>	0,0793 <b>0,2729</b>	4,3487 <b>0,4846</b>

Πίνακας 7 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση.

		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
χωρίς κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	0,1574 <b>-1,8456</b>	2,2903 <b>0,7974</b>	0,2845 <b>-8,6156</b>	5,0587 <b>0,2266</b>
	κοινοί νευρώνες	0,3657 <b>-14,3304</b>	2,02 <b>0,8517</b>	0,3245 <b>-19,3358</b>	3,5322 <b>0,6623</b>
με κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	0,0711 <b>0,3872</b>	3,2187 <b>0,6677</b>	0,1064 <b>-0,1067</b>	5,0869 <b>0,2598</b>
	κοινοί νευρώνες	0,0568 <b>0,5899</b>	3,0357 <b>0,723</b>	0,0863 <b>0,1963</b>	3,7104 <b>0,5965</b>

Πίνακας 8 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση.

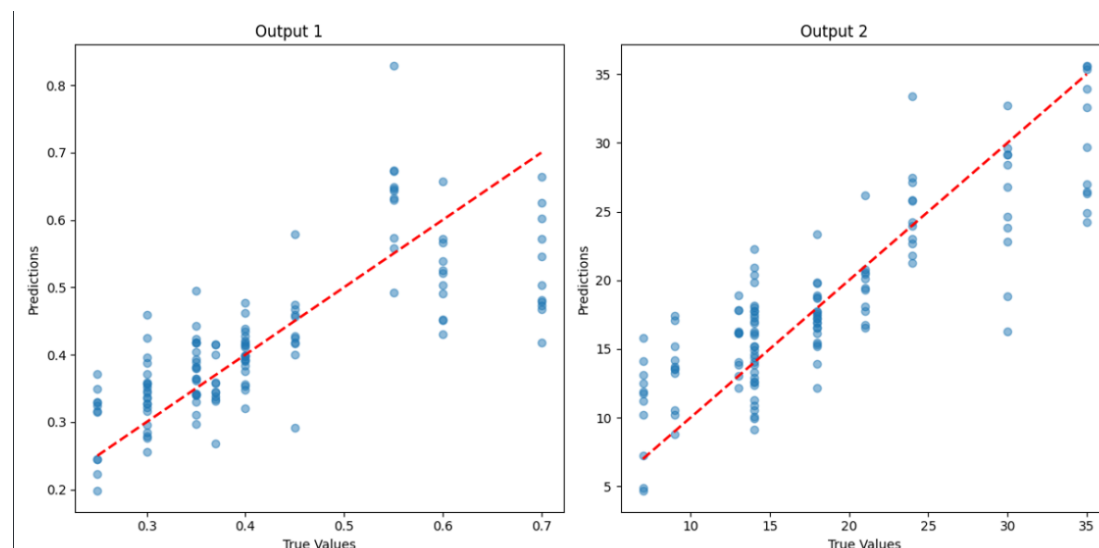
		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
χωρίς κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	0,2674 <b>-16,2395</b>	3,5155 <b>0,5695</b>	0,55 <b>-38,9656</b>	6,0955 <b>-0,02</b>
	κοινοί νευρώνες	0,7241 <b>-71,9118</b>	2,9589 <b>0,7037</b>	0,4189 <b>-31,3288</b>	5,0687 <b>0,286</b>
με κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	0,0697 <b>0,4342</b>	4,1703 <b>0,4502</b>	0,1 <b>0,018</b>	6,5829 <b>-0,2989</b>
	κοινοί νευρώνες	0,0616 <b>0,5949</b>	3,3115 <b>0,6984</b>	0,0933 <b>0,1311</b>	4,6479 <b>0,4602</b>

Πίνακας 9 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Αιματοκρίτης, μήκος αυχένα, μέγιστη διάμετρος ανευρύσματος, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου, διάμετρος δεξιάς λαγονίου, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση.

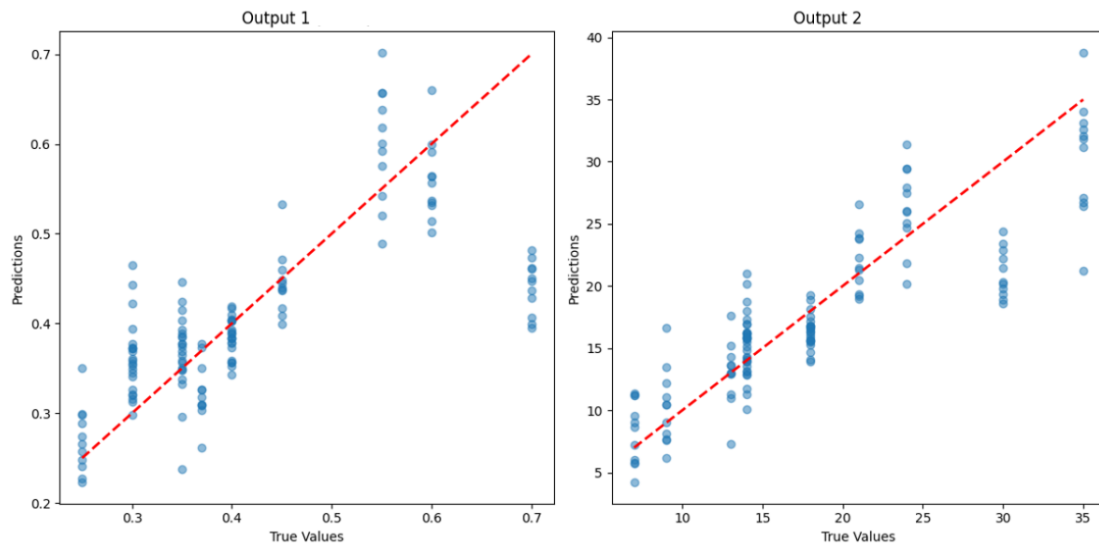
### 5.3.1 Παρατηρήσεις αποτελεσμάτων και γραφήματα

- **Καλύτερη** συνολική συμπεριφορά με **κοινούς νευρώνες** στα κρυφά επίπεδα και **κανονικοποίηση** τιμών στόχων.
- **Βέλτιστα** χαρακτηριστικά: **διάμετρος αριστερής και δεξιάς λαγονίου, καλύτερες προβλέψεις** και στις δύο τιμές στόχους σε σχέση με επιλογή **όλα τα χαρακτηριστικά**.
- Σε 6 χαρακτηριστικά: (+αιματοκρίτης, μήκος αυχένα, μέγιστη διάμετρος και μήκος ανευρύσματος): παρόμοια απόδοση σε πρόβλεψη **τιμών στόχων** σε σχέση με επιλογή **όλα τα χαρακτηριστικά**.
- **AdamW καλύτερη επίδοση από LM** σταθερά, για πρόβλεψη και των δύο τιμών στόχων, σε όλες τις ρυθμίσεις.

Παρακάτω παρουσιάζονται τα γραφήματα που σημείωσαν κάποια επιτυχία στη πρόβλεψη και των δύο τιμών στόχων.



Γράφημα 1 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυταρίσση.



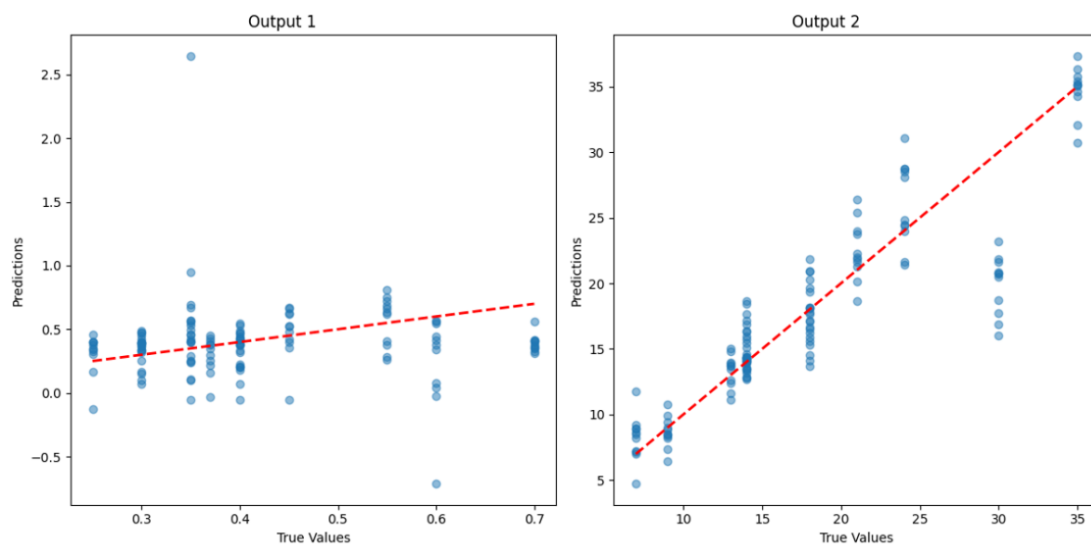
Γράφημα 2 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.

### 5.3.2 Αξιοσημείωτες περιπτώσεις

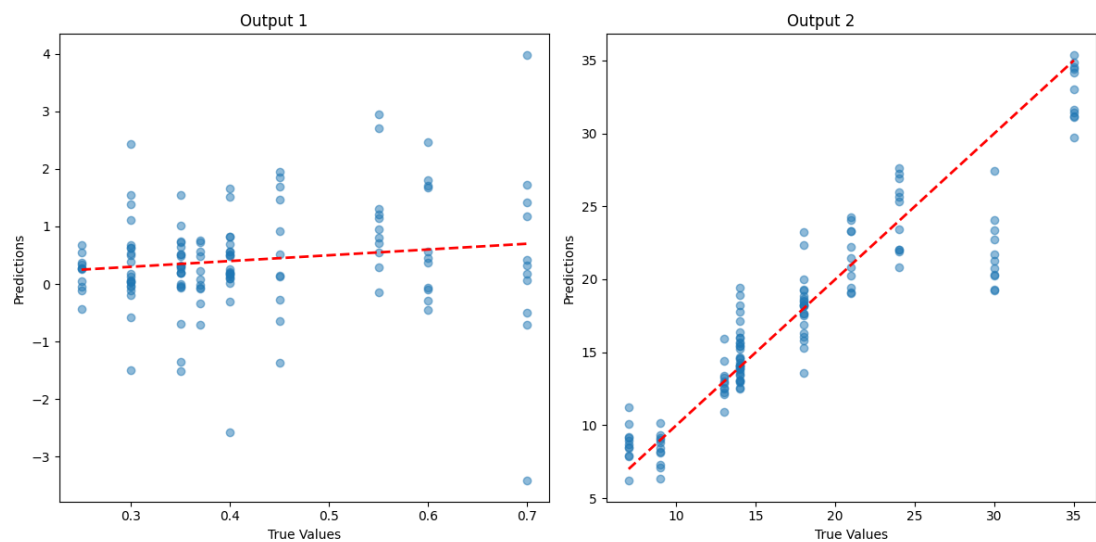
Τέλος, **χωρίς κανονικοποίηση** τιμών στόχων καταγράφηκε **ιδιαίτερα υψηλή ακρίβεια** στην **πρόβλεψη μέγιστης μετατόπισης**, με τις καλύτερες επιδόσεις να σημειώνονται με χαρακτηριστικά εισόδου τις **διαμέτρους δεξιάς και αριστερής λαγονίου** με ρυθμίσεις:

- AdamW+εξειδικευμένοι νευρώνες
- AdamW+κοινοί νευρώνες

Τα αντίστοιχα γραφήματα παρατίθενται παρακάτω.



Γράφημα 3 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, χωρίς κανονικοποίηση στόχων και με 16 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.



Γράφημα 4 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, χωρίς κανονικοποίηση στόχων και με 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.

## 5.4 Πρώτα πειραματικά αποτελέσματα μοντέλου 25 νευρώνων ανά κρυφό επίπεδο (κοινοί νευρώνες, με κανονικοποίηση στόχων)

	Μέγιστη VM τάση (MPa)		Μέγιστη μετατόπιση (mm)	
	MAE	$R^2$ score	MAE	$R^2$ score
Όλα τα χαρακτηριστικά	0,0683	0,4801	3,4263	0,4944
1.Ηλικία	0,0842	0,2602	4,9213	0,3495
2.Διαγνωστική διάμετρος ΑΚΑ (mm)	0,0846	0,2823	5,4158	0,2342
3.Αιματοκρίτης (Ht)	0,0762	0,366	5,2107	0,2848
4.Διάμετρος αυχένα (mm)	0,0835	0,2768	5,5736	0,2023
5.Μήκος αυχένα (mm)	0,0775	0,3463	4,5697	0,4471
6.Μέγιστη διάμετρος ανευρύσματος (mm)	0,0843	0,2722	4,9625	0,3585
7.Μήκος ανευρύσματος (mm)	0,0781	0,3769	4,6872	0,4306
8.Διάμετρος αριστερής λαγονίου (mm)	0,0593	0,5693	2,9854	0,7266
9.Διάμετρος δεξιάς λαγονίου (mm)	0,0715	0,452	4,581	0,3956
8,9	0,0594	0,5765	3,1456	0,6994
6,8	0,0634	0,5005	3,2521	0,7066
5,7,8,9	0,0641	0,5462	3,3288	0,6797
3,5,6,8	0,0687	0,4619	3,3768	0,6828
3,5,6,7,8,9	0,0701	0,5002	3,3978	0,6958
3,5,6,7,8,9	0,0701	0,5002	3,3978	0,6958
1,2,3,5,6,7,8,9	0,0646	0,5403	3,8285	0,6087
1,2,3,5,6,7,8,9	0,0646	0,5403	3,8285	0,6087

Πίνακας 10 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής  $R^2$  ( $R^2$  score) για μοντέλα με 25 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή AdamW. Σύνολο δεδομένων Κυπαρίσση.

	Μέγιστη VM τάση (MPa)		Μέγιστη μετατόπιση (mm)	
	MAE	$R^2$ score	MAE	$R^2$ score
Όλα τα χαρακτηριστικά	0,0801	0,3042	4,2569	0,5072
1.Ηλικία	0,1088	-0,147	5,9615	0,0482
2.Διαγνωστική διάμετρος ΑΚΑ (mm)	0,1086	-0,1334	6,8296	-0,1438
3.Αιματοκρίτης (Ht)	0,0952	0,0502	6,4193	-0,1233
4.Διάμετρος αυχένα (mm)	0,1142	-0,22	6,8336	-0,1671
5.Μήκος αυχένα (mm)	0,1053	-0,0859	6,1957	0,1335
6.Μέγιστη διάμετρος ανευρύσματος (mm)	0,0986	0,0775	5,9944	-0,0261
7.Μήκος ανευρύσματος (mm)	0,0993	-0,0153	6,251	0,1166
8.Διάμετρος αριστερής λαγονίου (mm)	0,0734	0,3573	4,154	0,5054
9.Διάμετρος δεξιάς λαγονίου (mm)	0,1011	-0,034	5,7058	0,132
8,9	0,0833	0,2435	3,9118	0,5866
6,8	0,0876	0,1872	4,193	0,4714
5,7,8,9	0,0923	0,1696	4,0738	0,5691
3,5,6,8	0,093	0,1278	4,8754	0,3967
3,5,6,7,8,9	0,0977	0,0569	4,7997	0,4347
3,5,6,7,8,9	0,0977	0,0569	4,7997	0,4347
1,2,3,5,6,7,8,9	0,0763	0,3497	4,5229	0,4456
1,2,3,5,6,7,8,9	0,0763	0,3497	4,5229	0,4456

Πίνακας 11 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής  $R^2$  ( $R^2$  score) για μοντέλα με 25 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή Levenberg-Marquardt. Σύνολο δεδομένων Κυπαρίσση.

#### 5.4.1 Παρατηρήσεις αποτελεσμάτων

- **Μέγιστη τάση von Mises (MPa)**
  - **AdamW: MAE 0.0594-0.0701** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
  - **LM: MAE 0.0763-0.0977** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
- **Μέγιστη μετατόπιση (mm)**
  - **AdamW: MAE 3.1456-3.8285** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
  - **LM: MAE 3.9118-4.8754** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
- **$R^2$ :** Θετικό και γενικά υψηλότερο για τη μέγιστη μετατόπιση, σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
- **Συνδυασμοί χαρακτηριστικών (2/4/6/8)**
  - **AdamW:** η πρόβλεψη της μέγιστης τάσης VM επιδόθηκε καλύτερα με τις επιλογές του, ενώ για πρόβλεψη της μέγιστης μετατόπισης είχε καλύτερη επίδοση με τις επιλογές μέσω LM· για πλήθος 6 και 8 χαρακτηριστικών συνέπεσαν οι επιλογές.
  - **LM:** απέδωσε καλύτερα με τις επιλογές χαρακτηριστικών μέσω AdamW



## 5.5 Περαιτέρω πειραματικά αποτελέσματα πάνω σε επιλογές χαρακτηριστικών που παρουσιάζουν καλές επιδόσεις

Για το μοντέλο με τους **25 νευρώνες ανά κρυφό επίπεδο** οι καλύτερες επιδόσεις προέκυψαν με τις εξής επιλογές χαρακτηριστικών εισόδου:

1. Διάμετρος αριστερής λαγονίου, διάμετρος δεξιάς λαγονίου.
2. Μήκος αυχένα, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου, διάμετρος δεξιάς λαγονίου.

Παραθέτουμε τα αποτελέσματα για τα δύο παραπάνω σύνολα χαρακτηριστικών μαζί με τα αποτελέσματα του σημείου αναφοράς “Όλα τα χαρακτηριστικά”.

		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
<i>χωρίς κανονικοποίηση τιμών στόχων</i>					
	εξειδικευμένοι νευρώνες	0,3958 <b>-23,848</b>	3,7521 <b>0,5296</b>	0,6016 <b>-52,8457</b>	7,072 <b>-0,2391</b>
	κοινοί νευρώνες	1,044	4,0716	0,2266	4,3095
		<b>-129,5525</b>	<b>0,4257</b>	<b>-7,0429</b>	<b>0,4539</b>
<i>με κανονικοποίηση τιμών στόχων</i>					
	εξειδικευμένοι νευρώνες	0,082 <b>0,3237</b>	4,749 <b>0,4346</b>	0,1456 <b>-0,903</b>	7,5582 <b>-0,3974</b>
	κοινοί νευρώνες	0,0683 <b>0,4801</b>	3,4263 <b>0,4944</b>	0,0801 <b>0,3042</b>	4,2569 <b>0,5072</b>

Πίνακας 12 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση.

		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
χωρίς κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	0,2604 <b>-12,6595</b>	2,2029 <b>0,7878</b>	0,4097 <b>-55,3429</b>	4,8315 <b>0,2897</b>
	κοινοί νευρώνες	0,6082 <b>-53,4487</b>	2,1732 <b>0,8223</b>	0,2386 <b>-9,6863</b>	3,6791 <b>0,6107</b>
	εξειδικευμένοι νευρώνες	0,0633 <b>0,4705</b>	3,6254 <b>0,5781</b>	0,1121 <b>-0,2023</b>	6,073 <b>-0,0552</b>
	κοινοί νευρώνες	0,0594 <b>0,5765</b>	3,1456 <b>0,6994</b>	0,0833 <b>0,2435</b>	3,9118 <b>0,5866</b>
με κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	0,0633 <b>0,4705</b>	3,6254 <b>0,5781</b>	0,1121 <b>-0,2023</b>	6,073 <b>-0,0552</b>
	κοινοί νευρώνες	0,0594 <b>0,5765</b>	3,1456 <b>0,6994</b>	0,0833 <b>0,2435</b>	3,9118 <b>0,5866</b>
	εξειδικευμένοι νευρώνες	0,0633 <b>0,4705</b>	3,6254 <b>0,5781</b>	0,1121 <b>-0,2023</b>	6,073 <b>-0,0552</b>
	κοινοί νευρώνες	0,0594 <b>0,5765</b>	3,1456 <b>0,6994</b>	0,0833 <b>0,2435</b>	3,9118 <b>0,5866</b>

Πίνακας 13 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση.

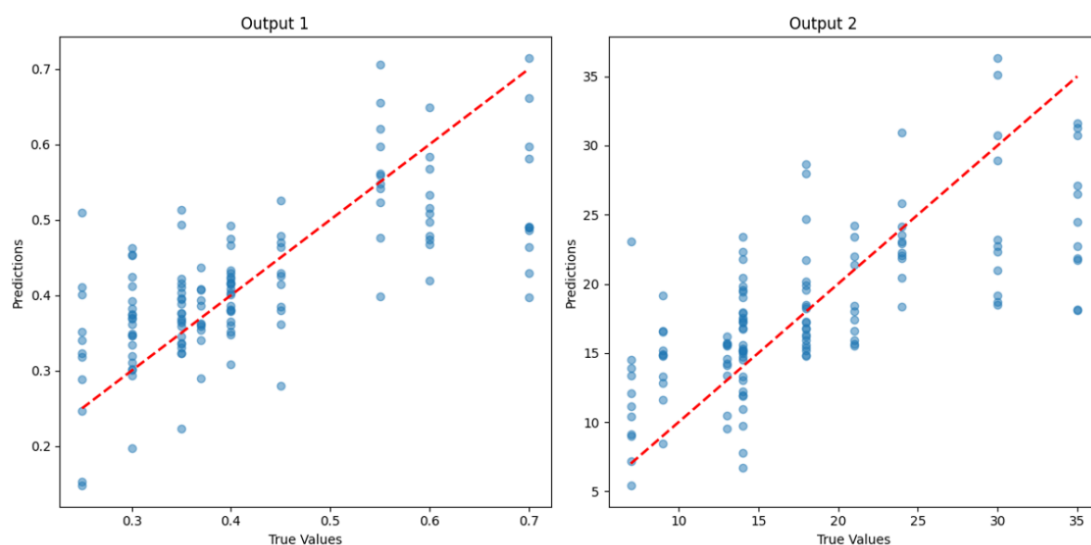
		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
χωρίς κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	0,2756 <b>-12,1908</b>	2,6966 <b>0,7365</b>	0,4711 <b>-26,6885</b>	6,2584 <b>0,0178</b>
	κοινοί νευρώνες	0,8005 <b>-75,7691</b>	2,4939 <b>0,8</b>	0,2803 <b>-11,8204</b>	4,3064 <b>0,4356</b>
	εξειδικευμένοι νευρώνες	0,0722 <b>0,4117</b>	3,673 <b>0,5982</b>	0,1415 <b>-1,0581</b>	6,3202 <b>0,0054</b>
	κοινοί νευρώνες	0,0641 <b>0,5462</b>	3,3288 <b>0,6797</b>	0,0923 <b>0,1696</b>	4,0738 <b>0,5691</b>
με κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	0,0722 <b>0,4117</b>	3,673 <b>0,5982</b>	0,1415 <b>-1,0581</b>	6,3202 <b>0,0054</b>
	κοινοί νευρώνες	0,0641 <b>0,5462</b>	3,3288 <b>0,6797</b>	0,0923 <b>0,1696</b>	4,0738 <b>0,5691</b>
	εξειδικευμένοι νευρώνες	0,0722 <b>0,4117</b>	3,673 <b>0,5982</b>	0,1415 <b>-1,0581</b>	6,3202 <b>0,0054</b>
	κοινοί νευρώνες	0,0641 <b>0,5462</b>	3,3288 <b>0,6797</b>	0,0923 <b>0,1696</b>	4,0738 <b>0,5691</b>

Πίνακας 14 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Μήκος αυχένα, Μήκος ανευρύσματος, Διάμετρος αριστερής λαγονίου και Διάμετρος δεξιάς λαγονίου, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Κυπαρίσση.

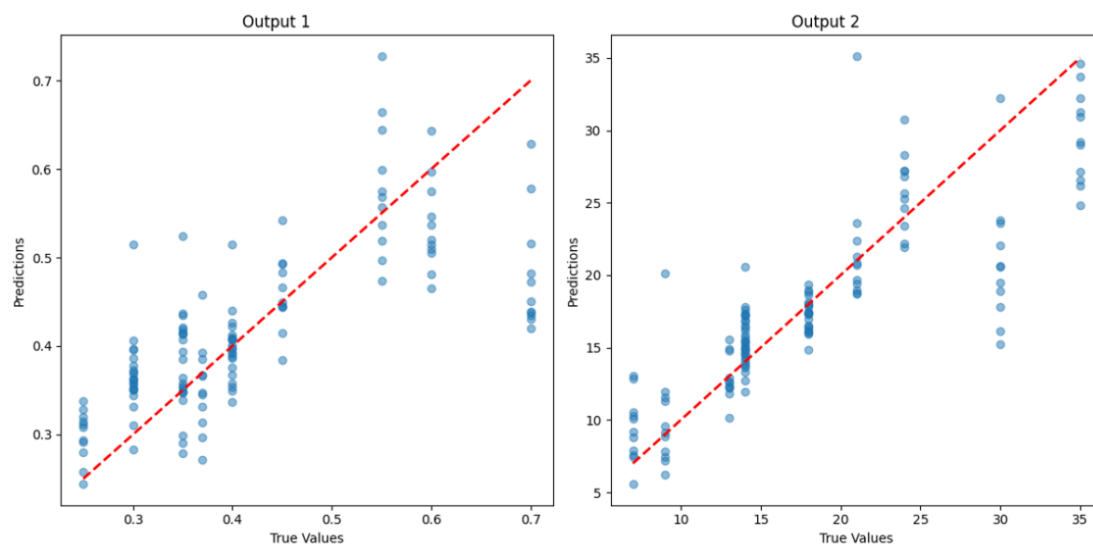
### 5.5.1 Παρατηρήσεις αποτελεσμάτων και γραφήματα

- **Καλύτερη** συνολική συμπεριφορά με **κοινούς νευρώνες** στα κρυφά επίπεδα και **κανονικοποίηση** τιμών στόχων.
- **Βέλτιστα** χαρακτηριστικά: **διάμετρος αριστερής και δεξιάς λαγονίου, καλύτερες προβλέψεις** και στις δύο τιμές στόχους σε σχέση με **“Όλα τα χαρακτηριστικά”**.
- Σε 4 χαρακτηριστικά: **(+μήκος αυχένα, μήκος ανευρύσματος )**: **καλύτερες προβλέψεις** και στις δύο τιμές στόχους σε σχέση με **“Όλα τα χαρακτηριστικά”**.
- **AdamW καλύτερη επίδοση από LM** σταθερά, για πρόβλεψη και των δύο τιμών στόχων, σε όλες τις ρυθμίσεις.

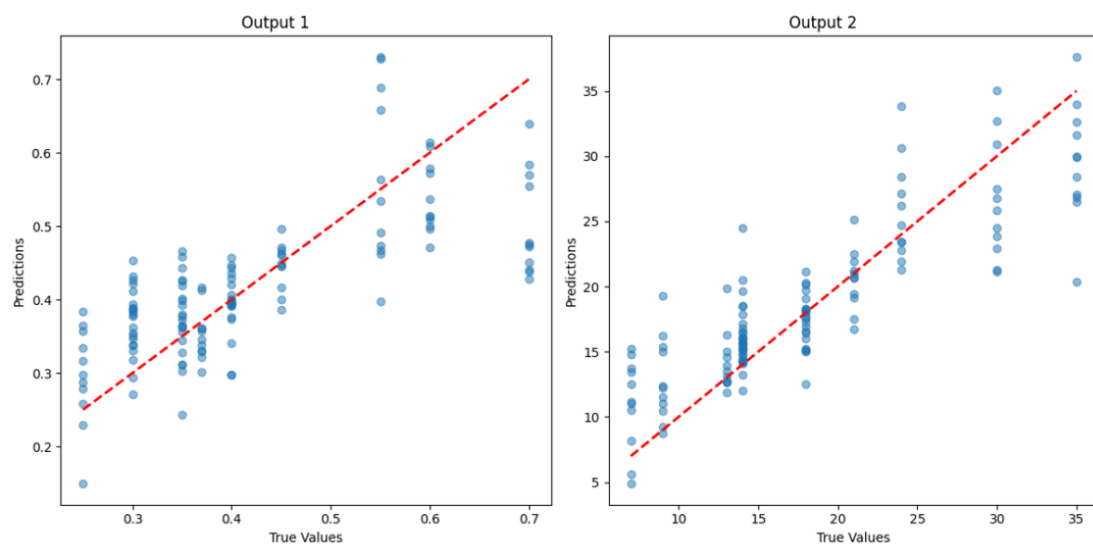
Παρακάτω παρουσιάζονται τα γραφήματα που σημείωσαν κάποια επιτυχία στη πρόβλεψη και των δύο τιμών στόχων.



Γράφημα 5 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυτταρίση.



Γράφημα 6 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.



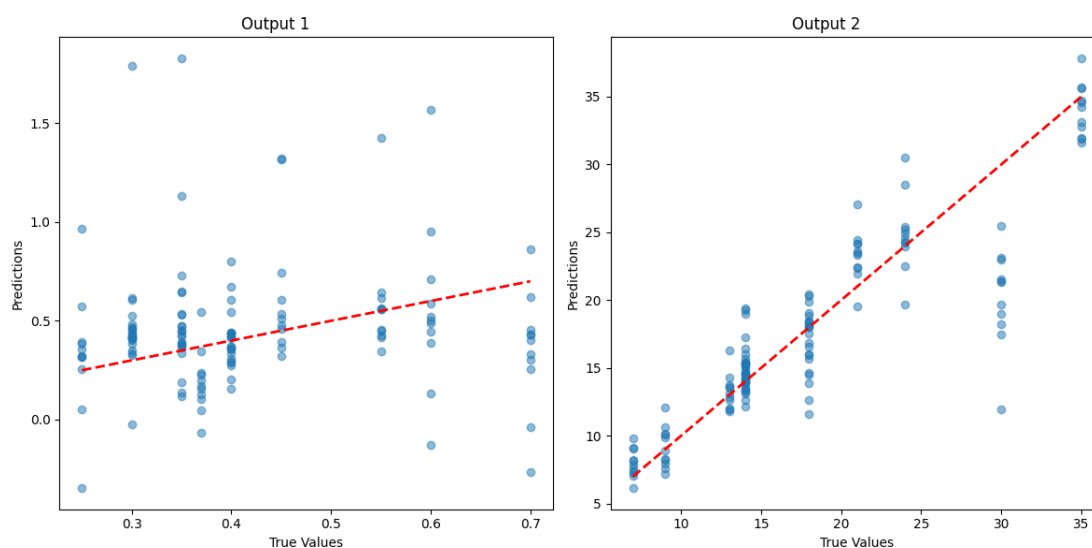
Γράφημα 7 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: μήκος αυχένα, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, με κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.

### 5.5.2 Αξιοσημείωτες περιπτώσεις

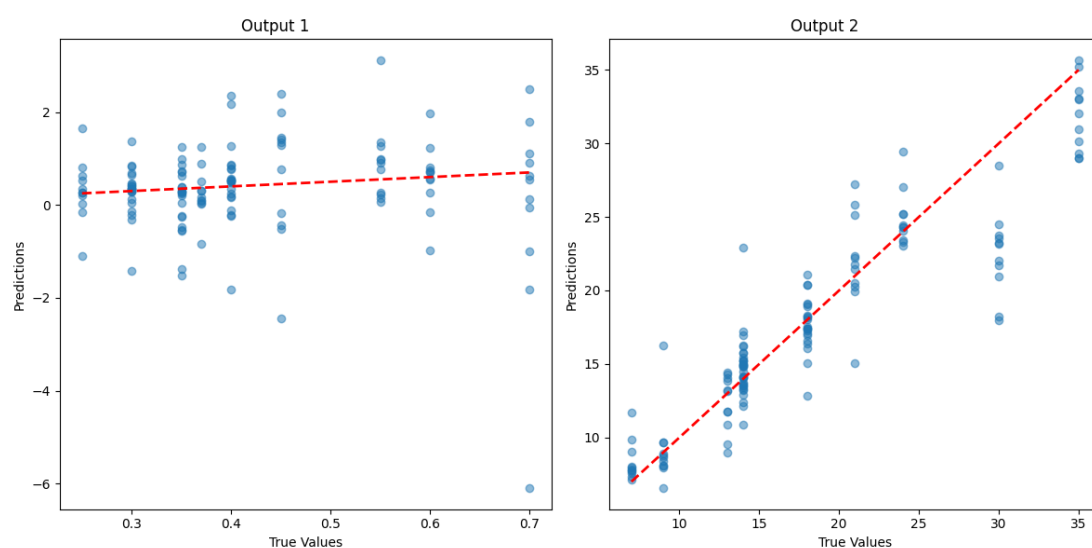
Τέλος, χωρίς κανονικοποίηση τιμών στόχων καταγράφηκε **ιδιαίτερα υψηλή ακρίβεια** στην **πρόβλεψη μέγιστης μετατόπισης**, με τις καλύτερες επιδόσεις να σημειώνονται με χαρακτηριστικά εισόδου τις **διαμέτρους δεξιάς και αριστερής λαγονίου** με ρυθμίσεις:

- AdamW+εξειδικευμένοι νευρώνες
- AdamW+κοινοί νευρώνες

Τα αντίστοιχα γραφήματα παρατίθενται παρακάτω.



Γράφημα 8 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, χωρίς κανονικοποίηση στόχων και με 26 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.



Γράφημα 9 προβλέψεων-πραγματικών τιμών με χαρακτηριστικά εισόδου: διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW, χωρίς κανονικοποίηση στόχων και με 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Κυπαρίσση.

## Κεφάλαιο 6

### Αποτελέσματα μοντέλου με το δεύτερο σύνολο δεδομένων

#### 6.1 Εισαγωγή

Το δεύτερο σύνολο δεδομένων που χρησιμοποιήθηκε προέρχεται από την εργασία του Ξυδιά και περιλαμβάνει 17 δείγματα. Η αξιολόγηση των μοντέλων πραγματοποιήθηκε με εξαντλητικό *leave-2-out cross validation*, όπως αναφέρθηκε στην ενότητα διαδικασία αξιολόγησης, και συνολικά έγιναν 136 επαναλήψεις ανά μοντέλο - μία για κάθε δυνατό μοναδικό συνδυασμό.

Στους πίνακες που ακολουθούν παρουσιάζονται τα αποτελέσματα της αξιολόγησης. Στην αριστερή στήλη κάθε πίνακα αναγράφονται ή αριθμούνται τα χαρακτηριστικά που αντιστοιχούν σε κάθε γραμμή αποτελεσμάτων. Σε περιπτώσεις συνδυασμών χαρακτηριστικών, για κάθε πλήθος χαρακτηριστικών εμφανίζονται δύο γραμμές. Η πρώτη γραμμή αντιστοιχεί στον συνδυασμό που επιλέχθηκε με AdamW και η δεύτερη στον συνδυασμό που επιλέχθηκε με LM.

## 6.2 Πρώτα πειραματικά αποτελέσματα μοντέλου 15 νευρώνων ανά κρυφό επίπεδο (κοινοί νευρώνες, με κανονικοποίηση στόχων)

	Μέγιστη VM τάση (MPa)		Μέγιστη μετατόπιση (mm)	
	MAE	$R^2$ score	MAE	$R^2$ score
Όλα τα χαρακτηριστικά	10,1858	0,3894	0,1676	0,415
1.Ηλικία	6,2592	0,755	0,164	0,4366
2.Διαγνωστική διάμετρος ΑΚΑ (mm)	6,3439	0,7355	0,1656	0,4237
3.Αιματοκρίτης (Ht)	6,3054	0,6421	0,1726	0,3748
4.Διάμετρος αυχένα (mm)	6,0919	0,7365	0,1715	0,3844
5.Μήκος αυχένα (mm)	4,884	0,8583	0,1829	0,265
6.Μέγιστη διάμετρος ανευρύσματος (mm)	5,5314	0,8247	0,1766	0,3209
7.Μήκος ανευρύσματος (mm)	5,158	0,8417	0,1842	0,2718
8.Διάμετρος αριστερής λαγονίου (mm)	5,1542	0,87	0,1905	0,2108
9.Διάμετρος δεξιάς λαγονίου (mm)	4,6753	0,8959	0,1691	0,3799
1,9	6,1952	0,7385	0,1788	0,3614
8,9	5,3465	0,8433	0,1742	0,3143
1,2,8,9	7,1861	0,6772	0,1679	0,3881
2,7,8,9	7,5836	0,653	0,16	0,4849
1,2,4,5,8,9	8,1467	0,6434	0,1577	0,4079
1,2,4,7,8,9	9,7682	0,5267	0,1642	0,4057
1,2,3,4,5,6,8,9	9,7729	0,4808	0,1617	0,4253
1,2,4,5,6,7,8,9	9,3406	0,504	0,1715	0,3653

Πίνακας 15 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής  $R^2$  ( $R^2$  score) για μοντέλα με 15 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή AdamW. Σύνολο δεδομένων Ξυδιά.

	Μέγιστη VM τάση (MPa)		Μέγιστη μετατόπιση (mm)	
	MAE	$R^2$ score	MAE	$R^2$ score
Όλα τα χαρακτηριστικά	7,5899	0,7254	0,2243	-0,0956
1.Ηλικία	3,8416	0,9113	0,1914	0,02
2.Διαγνωστική διάμετρος ΑΚΑ (mm)	4,4459	0,8831	0,2082	0,1036
3.Αιματοκρίτης (Ht)	4,412	0,8876	0,2355	-0,1599
4.Διάμετρος αυχένα (mm)	4,1561	0,906	0,2043	0,052
5.Μήκος αυχένα (mm)	4,1314	0,9014	0,2345	-0,1229
6.Μέγιστη διάμετρος ανευρύσματος (mm)	4,6367	0,8558	0,2134	-0,0466
7.Μήκος ανευρύσματος (mm)	3,6402	0,9212	0,2152	-0,0147
8.Διάμετρος αριστερής λαγονίου (mm)	3,4589	0,925	0,1918	0,1101
9.Διάμετρος δεξιάς λαγονίου (mm)	2,7709	0,9449	0,1677	0,3174
1,9	4,3761	0,8968	0,1858	0,0653
8,9	3,9765	0,9066	0,1861	0,2238
1,2,8,9	5,1749	0,8655	0,1791	0,3083
2,7,8,9	5,9315	0,81	0,1544	0,5745
1,2,4,5,8,9	6,0091	0,8382	0,2082	0,0221
1,2,4,7,8,9	6,4578	0,7862	0,1982	0,0772
1,2,3,4,5,6,8,9	6,6046	0,8026	0,2246	-0,1214
1,2,4,5,6,7,8,9	7,2283	0,7556	0,2161	-0,0525

Πίνακας 16 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής  $R^2$  ( $R^2$  score) για μοντέλα με 15 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή Levenberg-Marquardt. Σύνολο δεδομένων Ξυδιά.

### 6.2.1 Παρατηρήσεις αποτελεσμάτων

- **Μέγιστη τάση von Mises (MPa)**
  - **AdamW**
    - **MAE 5.3465-10.1858** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
    - **R<sup>2</sup> θετικό**
  - **LM**
    - **MAE 3.9765-7.5899** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
    - **R<sup>2</sup> θετικό**
- **Μέγιστη μετατόπιση (mm)**
  - **AdamW**
    - **MAE 0.1577-0.1788** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου
    - **R<sup>2</sup> θετικό**
  - **LM**
    - **MAE 0.1544-0.2246** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
    - **R<sup>2</sup> θετικό** έως και πλήθος 6 χαρακτηριστικών εισόδου, αρνητικό για περισσότερα χαρακτηριστικά εισόδου.
- **R<sup>2</sup>: Γενικά υψηλότερο για τη μέγιστη τάση VM σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.**
- **Συνδυασμοί χαρακτηριστικών (2/4/6/8)**
  - **AdamW**
    - Η πρόβλεψη της **μέγιστης τάσης VM** επιδόθηκε καλύτερα με τις επιλογές του για πλήθος 2, 4 και 6 χαρακτηριστικών, ενώ είχε καλύτερη επίδοση με τις επιλογές χαρακτηριστικών του **LM** για πλήθος 8 χαρακτηριστικών
    - Η πρόβλεψη της **μέγιστης μετατόπισης** επιδόθηκε καλύτερα με τις επιλογές του για πλήθος 2, 6 και 8 χαρακτηριστικών, ενώ είχε καλύτερη επίδοση με τις επιλογές χαρακτηριστικών του **LM** για πλήθος 4 χαρακτηριστικών
  - **LM**
    - Η πρόβλεψη της **μέγιστης τάσης VM** επιδόθηκε καλύτερα με τις επιλογές του για πλήθος 2 χαρακτηριστικών, ενώ είχε καλύτερη επίδοση με τις επιλογές χαρακτηριστικών του **AdamW** για πλήθος 4, 6 και 8 χαρακτηριστικών
    - Η πρόβλεψη της **μέγιστης μετατόπισης** επιδόθηκε παντού καλύτερα με τις επιλογές του



### 6.3 Περαιτέρω πειραματικά αποτελέσματα πάνω σε επιλογές χαρακτηριστικών που παρουσιάζουν καλές επιδόσεις

Στη συνέχεια θα επιλέξουμε τα καλύτερα μεικτά αποτελέσματα, δηλαδή αυτά με καλή επίδοση και στις δύο τιμές στόχους, ώστε να τα αναλύσουμε περαιτέρω. Συγκεκριμένα θα συγκρίνουμε την αρχιτεκτονική με κοινούς νευρώνες με εκείνη των εξειδικευμένων νευρώνων, με και χωρίς κανονικοποίηση των τιμών στόχων.

Για το μοντέλο με τους **15 νευρώνες ανά κρυφό επίπεδο** οι καλύτερες επιδόσεις προέκυψαν με τις εξής επιλογές χαρακτηριστικών εισόδου:

**Διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου.**

Παραθέτουμε τα αποτελέσματα για τα παραπάνω χαρακτηριστικά μαζί με τα αντίστοιχα αποτελέσματα με επιλογή όλα τα χαρακτηριστικά για εισόδους ως σημείο αναφοράς, ώστε να φαίνεται καθαρά το κέρδος από την επιλογή χαρακτηριστικών.

		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
<i>χωρίς κανονικοποίηση τιμών στόχων</i>					
	εξειδικευμένοι νευρώνες	2,2942 <b>0,94</b>	0,3822 <b>-3,6967</b>	4,0516 <b>0,8404</b>	0,5767 <b>-6,0225</b>
	κοινοί νευρώνες	2,8891 <b>0,9236</b>	1,213 <b>-55,332</b>	2,6278 <b>0,8992</b>	0,3863 <b>-2,9684</b>
<i>με κανονικοποίηση τιμών στόχων</i>					
	εξειδικευμένοι νευρώνες	7,5819 <b>0,6769</b>	0,1536 <b>0,514</b>	10,7105 <b>0,4041</b>	0,2818 <b>-0,6519</b>
	κοινοί νευρώνες	10,1858 <b>0,3894</b>	0,1676 <b>0,415</b>	7,5899 <b>0,7254</b>	0,2243 <b>-0,0956</b>

Πίνακας 17 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Όλα τα χαρακτηριστικά, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.

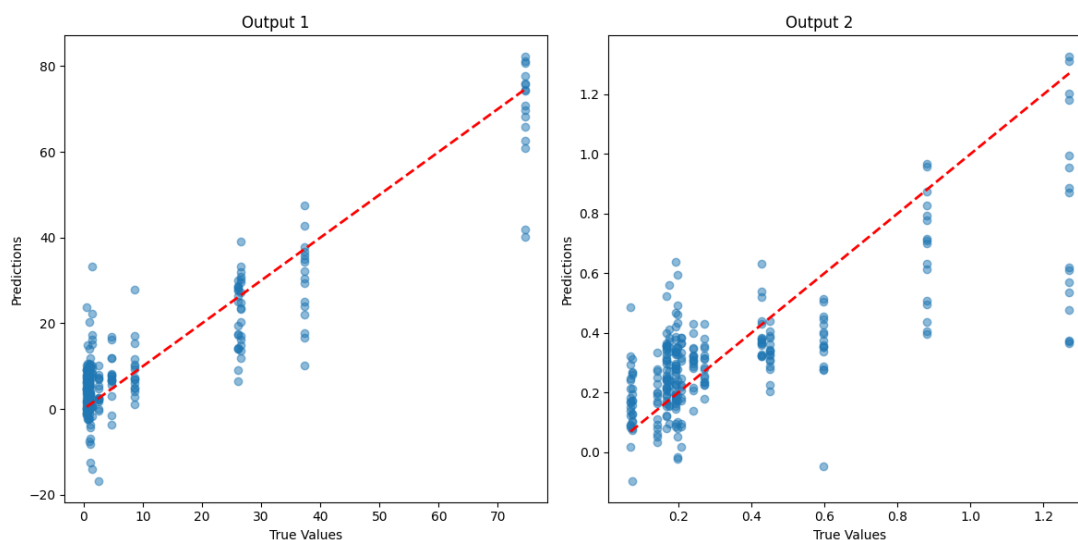
		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
χωρίς κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	1,4874	0,3256	2,3757	0,4766
		<b>0,9786</b>	<b>-1,4517</b>	<b>0,9303</b>	<b>-4,1806</b>
με κανονικοποίηση τιμών στόχων	κοινοί νευρώνες	1,4624	0,8657	1,1781	0,3233
		<b>0,9828</b>	<b>-23,0497</b>	<b>0,9814</b>	<b>-1,3927</b>
	εξειδικευμένοι νευρώνες	5,5326	0,1415	7,0059	0,2265
		<b>0,8337</b>	<b>0,5476</b>	<b>0,6756</b>	<b>0,0805</b>
	κοινοί νευρώνες	7,1861	0,1679	5,1749	0,1791
		<b>0,653</b>	<b>0,4849</b>	<b>0,81</b>	<b>0,5745</b>

Πίνακας 18 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: Διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, με 15 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.

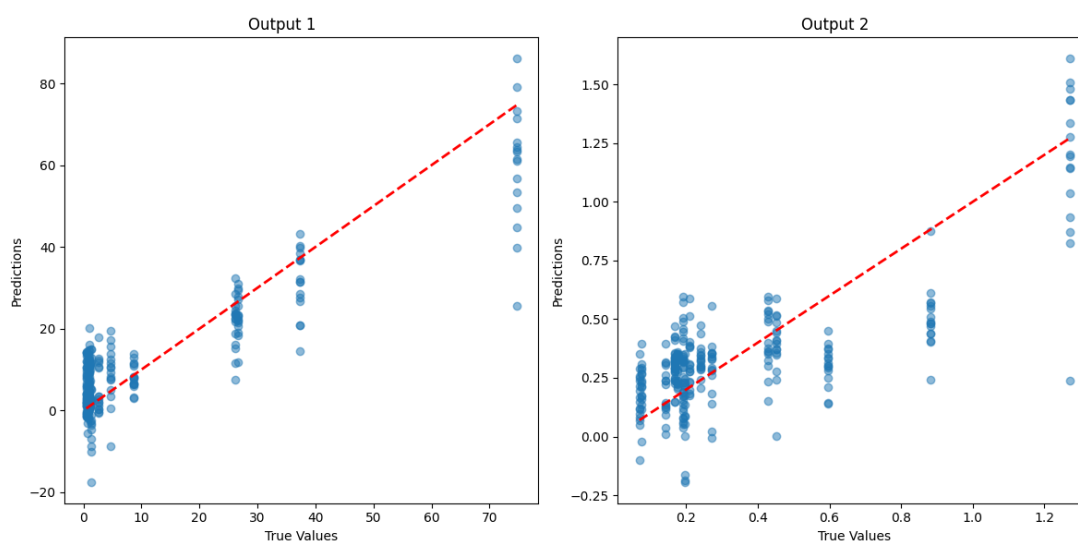
### 6.3.1 Παρατηρήσεις αποτελεσμάτων και γραφήματα

- Όλα τα χαρακτηριστικά:
  - καλή επίδοση μόνο με εξειδικευμένους νευρώνες + κανονικοποίηση στόχων + AdamW.
- Σύνολο 4 χαρακτηριστικών που ορίσαμε:
  - Συστηματικά καλύτερα αποτελέσματα σε όλο τον πίνακα, ανεξάρτητα από βελτιστοποιητή.
  - Ρυθμίσεις με ιδιαίτερη επιτυχία:
    - Εξειδικευμένοι νευρώνες + κανονικοποίηση στόχων + AdamW.
    - Κοινοί νευρώνες + κανονικοποίηση στόχων + Levenberg–Marquardt (LM).

Τα αντίστοιχα γραφήματα από τις ρυθμίσεις που σημείωσαν επιτυχία παρατίθενται παρακάτω.



Γράφημα 10 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW και με κανονικοποίηση στόχων και 16 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά.



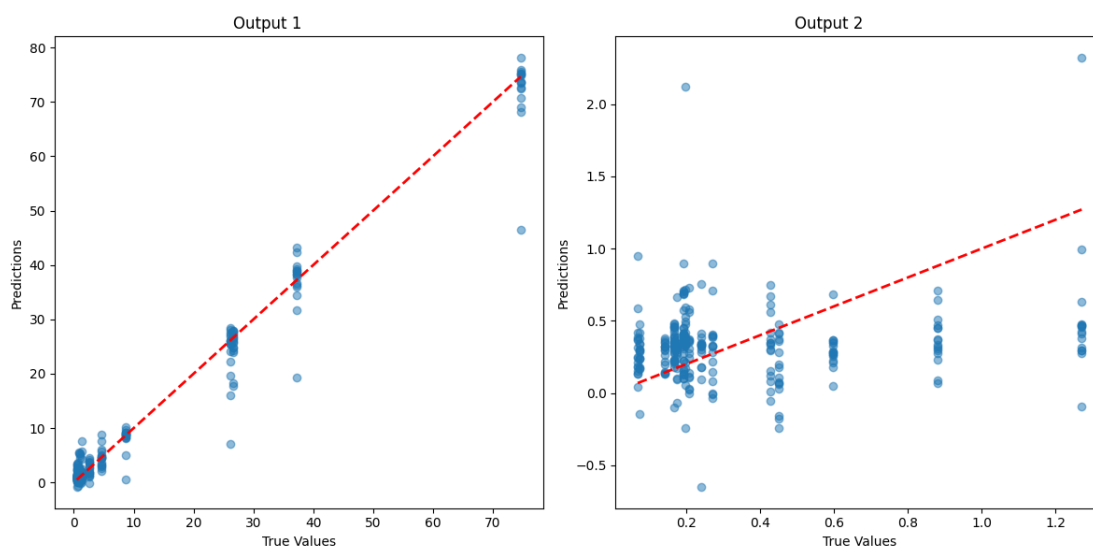
Γράφημα 11 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: LM και με κανονικοποίηση στόχων και 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά.

### 6.3.2 Αξιοσημείωτες περιπτώσεις

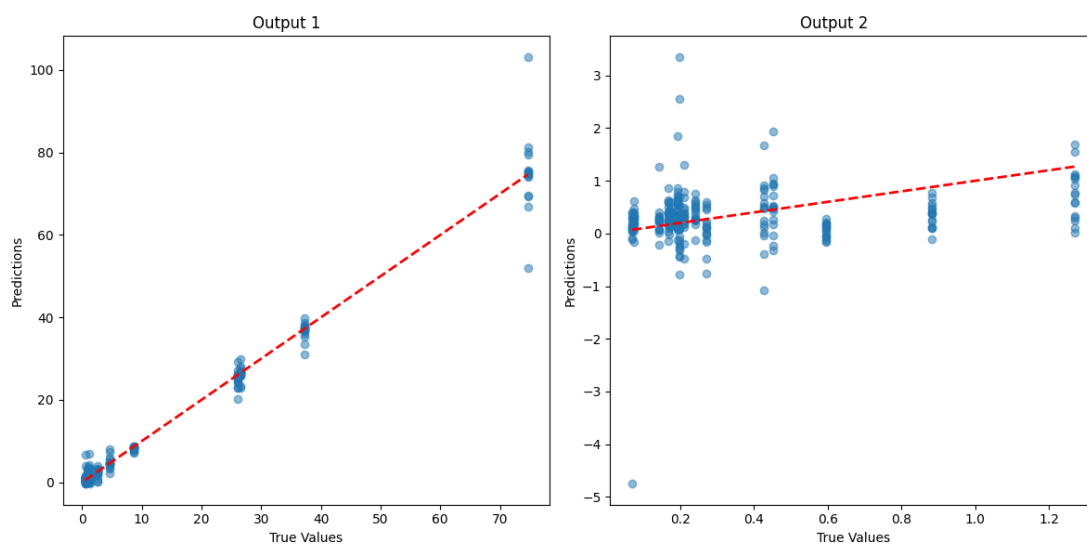
Τέλος, χωρίς κανονικοποίηση τιμών στόχων καταγράφηκε ιδιαίτερα υψηλή ακρίβεια στην πρόβλεψη μέγιστης τάσης VM, με τις καλύτερες επιδόσεις να σημειώνονται από:

- **AdamW+εξειδικευμένοι νευρώνες**
- **LM+κοινοί νευρώνες**

Τα αντίστοιχα γραφήματα παρατίθενται παρακάτω.



Γράφημα 12 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW και χωρίς κανονικοποίηση στόχων και με 16 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά.



Γράφημα 13 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ, μήκος ανευρύσματος, διάμετρος αριστερής λαγονίου και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: LM και χωρίς κανονικοποίηση στόχων και με 15 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά.

## 6.4 Πρώτα πειραματικά αποτελέσματα μοντέλου 25 νευρώνων ανά κρυφό επίπεδο (κοινοί νευρώνες, με κανονικοποίηση στόχων)

	Μέγιστη VM τάση (MPa)		Μέγιστη μετατόπιση (mm)	
	MAE	$R^2$ score	MAE	$R^2$ score
Όλα τα χαρακτηριστικά	10,6137	0,4285	0,176	0,3611
1.Ηλικία	6,3515	0,756	0,1709	0,3699
2.Διαγνωστική διάμετρος ΑΚΑ (mm)	6,8999	0,7115	0,1622	0,4615
3.Αιματοκρίτης (Ht)	6,2915	0,741	0,1787	0,3273
4.Διάμετρος αυχένα (mm)	5,908	0,804	0,1761	0,3127
5.Μήκος αυχένα (mm)	6,3817	0,7237	0,1832	0,2755
6.Μέγιστη διάμετρος ανευρύσματος (mm)	6,6575	0,7465	0,1843	0,222
7.Μήκος ανευρύσματος (mm)	6,7035	0,7679	0,1913	0,2285
8.Διάμετρος αριστερής λαγονίου (mm)	6,3839	0,7425	0,1795	0,3005
9.Διάμετρος δεξιάς λαγονίου (mm)	5,6358	0,8155	0,1728	0,3204
2,9	7,8338	0,5615	0,1611	0,4668
8,9	5,5441	0,8311	0,1793	0,3116
1,2,4,9	8,3339	0,6524	0,1635	0,4142
4,7,8,9	8,3225	0,6508	0,186	0,261
1,2,3,4,7,9	9,4125	0,4838	0,1771	0,3949
1,2,4,7,8,9	10,4357	0,4465	0,181	0,3252
1,2,3,4,6,7,8,9	10,9563	0,3623	0,1809	0,3502
1,2,4,5,6,7,8,9	9,9063	0,4515	0,1901	0,2884

Πίνακας 19 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής  $R^2$  ( $R^2$  score) για μοντέλα με 25 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή AdamW. Σύνολο δεδομένων Ξυδιά.

	Μέγιστη VM τάση (MPa)		Μέγιστη μετατόπιση (mm)	
	MAE	$R^2$ score	MAE	$R^2$ score
Όλα τα χαρακτηριστικά	8,7034	0,6617	0,2188	-0,0465
1.Ηλικία	3,3166	0,9185	0,1873	0,0386
2.Διαγνωστική διάμετρος ΑΚΑ (mm)	4,6495	0,8684	0,2146	0,0486
3.Αιματοκρίτης (Ht)	5,3207	0,849	0,2265	-0,0749
4.Διάμετρος αυχένα (mm)	3,5437	0,9197	0,2018	0,0614
5.Μήκος αυχένα (mm)	4,4403	0,8927	0,2315	-0,1047
6.Μέγιστη διάμετρος ανευρύσματος (mm)	4,4571	0,882	0,2152	-0,0398
7.Μήκος ανευρύσματος (mm)	3,6566	0,9269	0,2208	-0,0327
8.Διάμετρος αριστερής λαγονίου (mm)	2,7779	0,9478	0,1861	0,1345
9.Διάμετρος δεξιάς λαγονίου (mm)	2,9933	0,9413	0,1648	0,3736
2,9	4,7827	0,8694	0,1613	0,5406
8,9	3,724	0,9167	0,1815	0,2464
1,2,4,9	5,3925	0,845	0,1858	0,1836
4,7,8,9	6,4327	0,8054	0,206	0,049
1,2,3,4,7,9	7,2036	0,7451	0,2195	-0,0298
1,2,4,7,8,9	6,464	0,7649	0,1997	0,0737
1,2,3,4,6,7,8,9	8,2641	0,6939	0,2266	-0,1001
1,2,4,5,6,7,8,9	7,4215	0,7115	0,217	0,001

Πίνακας 20 Μέσο απόλυτο σφάλμα (MAE) και συντελεστής  $R^2$  ( $R^2$  score) για μοντέλα με 25 κοινούς νευρώνες ανά κρυφό επίπεδο και κανονικοποίηση με βελτιστοποιητή Levenberg-Marquardt. Σύνολο δεδομένων Ξυδιά.

#### 6.4.1 Παρατηρήσεις αποτελεσμάτων

- **Μέγιστη τάση von Mises (MPa)**
  - **AdamW**
    - **MAE 5.5441-10.9563** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
    - $R^2$  θετικό σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
  - **LM**
    - **MAE 3.724- 8.7034** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
    - $R^2$  θετικό σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
- **Μέγιστη μετατόπιση (mm)**
  - **AdamW**
    - **MAE 0.1611-0.1901** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
    - $R^2$  θετικό σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
  - **LM**
    - **MAE 0.1613-0.2266** σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
    - $R^2$  θετικό ως επί το πλείστον σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου, αρνητικό στις επιλογές του AdamW για πλήθος 6 και 8 χαρακτηριστικών εισόδου και για επιλογή εισόδων όλα τα χαρακτηριστικά
- $R^2$ : Γενικά υψηλότερο για τη μέγιστη τάση VM σε όλους τους συνδυασμούς χαρακτηριστικών εισόδου.
- **Συνδυασμοί χαρακτηριστικών (2/4/6/8)**
  - **AdamW**
    - Η πρόβλεψη της **μέγιστης τάσης VM** επιδόθηκε καλύτερα με τις επιλογές του για πλήθος 4 και 6 χαρακτηριστικών, ενώ είχε καλύτερη επίδοση με τις επιλογές χαρακτηριστικών του **LM** για πλήθος 2 και 8 χαρακτηριστικών
    - Η πρόβλεψη της **μέγιστης μετατόπισης** επιδόθηκε καλύτερα με τις επιλογές του
  - **LM**
    - Η πρόβλεψη της **μέγιστης τάσης VM** επιδόθηκε καλύτερα με τις επιλογές του για πλήθος 2, 6 και 8 χαρακτηριστικών, ενώ είχε καλύτερη επίδοση με τις επιλογές χαρακτηριστικών του **AdamW** για πλήθος 4 χαρακτηριστικών

- Η πρόβλεψη της **μέγιστης μετατόπισης** επιδόθηκε καλύτερα με τις επιλογές του για πλήθος 6 και 8 χαρακτηριστικών, ενώ είχε καλύτερη επίδοση με τις επιλογές χαρακτηριστικών του **AdamW** για πλήθος 2 και 4 χαρακτηριστικών

## 6.5 Περαιτέρω πειραματικά αποτελέσματα πάνω σε επιλογές χαρακτηριστικών που παρουσιάζουν καλές επιδόσεις

Για το μοντέλο με τους **25 νευρώνες ανά κρυφό επίπεδο** οι καλύτερες επιδόσεις προέκυψαν με τις εξής επιλογές χαρακτηριστικών εισόδου:

1. **Διαγνωστική διάμετρος ΑΚΑ, διάμετρος δεξιάς λαγονίου.**
2. **Ηλικία, διαγνωστική διάμετρος ΑΚΑ, διάμετρος αυχένα, διάμετρος δεξιάς λαγονίου.**

Παραθέτουμε τα αποτελέσματα για τα δύο παραπάνω σύνολα χαρακτηριστικών μαζί με τα αποτελέσματα του σημείου αναφοράς “Όλα τα χαρακτηριστικά”.

		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
<i>χωρίς κανονικοποίηση τιμών στόχων</i>					
	εξειδικευμένοι νευρώνες	2,5334 <b>0,9345</b>	0,497 <b>-7,7315</b>	4,9294 <b>0,7606</b>	0,6947 <b>-8,5473</b>
	κοινοί νευρώνες	3,2538 <b>0,9</b>	1,4358 <b>-59,3727</b>	2,5448 <b>0,9303</b>	0,3416 <b>-1,3141</b>
<i>με κανονικοποίηση τιμών στόχων</i>					
	εξειδικευμένοι νευρώνες	9,202 <b>0,561</b>	0,1589 <b>0,4654</b>	13,6547 <b>0,0072</b>	0,2967 <b>-0,7032</b>
	κοινοί νευρώνες	10,6137 <b>0,4285</b>	0,176 <b>0,3611</b>	8,7034 <b>0,6617</b>	0,2188 <b>-0,0465</b>

Πίνακας 21 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: : Όλα τα χαρακτηριστικά, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.

		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
χωρίς κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	1,024	0,3615	1,2925	0,4308
		<b>0,9927</b>	<b>-4,5014</b>	<b>0,973</b>	<b>-4,4491</b>
	κοινοί νευρώνες	1,8144	0,908	0,5414	0,2464
		<b>0,9731</b>	<b>-22,4553</b>	<b>0,996</b>	<b>-0,1111</b>
με κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	4,9448	0,1304	4,4426	0,2103
		<b>0,8277</b>	<b>0,6459</b>	<b>0,8478</b>	<b>0,088</b>
	κοινοί νευρώνες	7,8338	0,1611	4,7827	0,1613
		<b>0,5615</b>	<b>0,4668</b>	<b>0,8694</b>	<b>0,5406</b>

Πίνακας 22 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: : Διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.

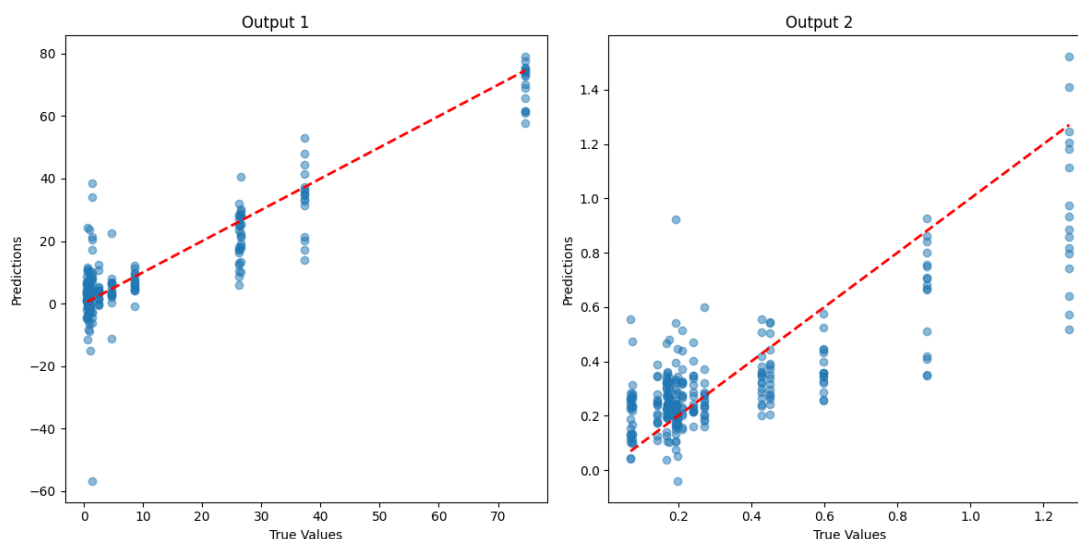
		MAE εξόδου 1 ADAMW	MAE εξόδου 2 ADAMW	MAE εξόδου 1 LM	MAE εξόδου 2 LM
		συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$	συντελεστής $R^2$
χωρίς κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	1.4855	0.3399	2.2933	0.447
		<b>0.9769</b>	<b>-2.6701</b>	<b>0.949</b>	<b>-4.2351</b>
	κοινοί νευρώνες	1.9371	1.1857	1.0013	0.3109
		<b>0.9713</b>	<b>-39.1575</b>	<b>0.9896</b>	<b>-1.0189</b>
με κανονικοποίηση τιμών στόχων					
	εξειδικευμένοι νευρώνες	6.6023	0.129	7.5408	0.2316
		<b>0.7064</b>	<b>0.6213</b>	<b>0.64</b>	<b>-0.2702</b>
	κοινοί νευρώνες	8.3339	0.1635	5.3925	0.1858
		<b>0.6524</b>	<b>0.4142</b>	<b>0.845</b>	<b>0.1836</b>

Πίνακας 23 Αποτελέσματα μοντέλου για Χαρακτηριστικά εισόδου: : Ηλικία, διαγνωστική διάμετρος ΑΚΑ, διάμετρος αυχένα και διάμετρος δεξιάς λαγονίου, με 25 νευρώνες ανά κρυφό επίπεδο. Σύνολο δεδομένων Ξυδιά.

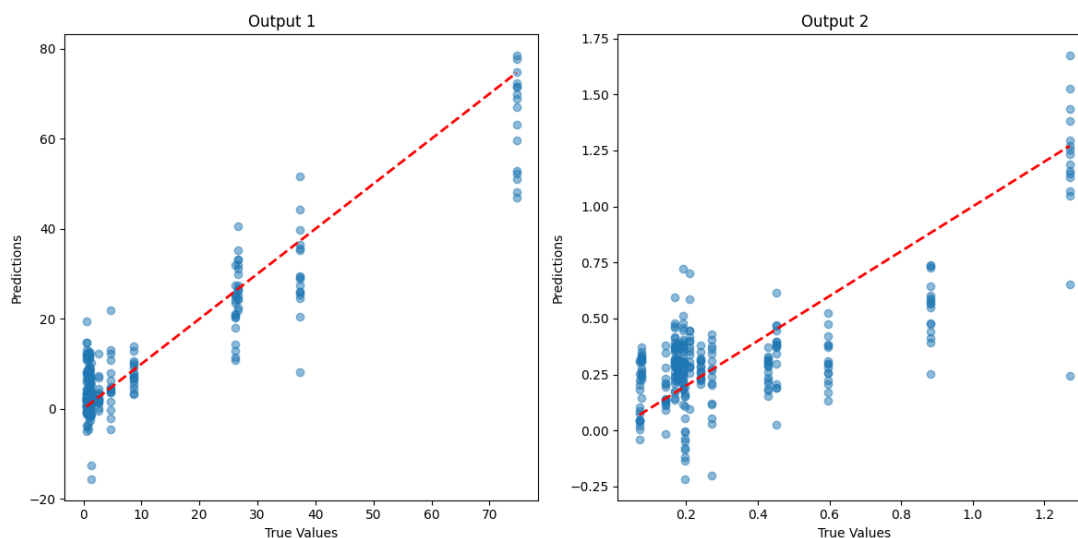


### 6.5.1 Παρατηρήσεις αποτελεσμάτων και γραφήματα

- Όλα τα χαρακτηριστικά:
  - καλή επίδοση μόνο με **εξειδικευμένους νευρώνες + κανονικοποίηση στόχων + AdamW**.
- **Σύνολο 2 χαρακτηριστικών** που ορίσαμε:
  - **Συστηματικά καλύτερα αποτελέσματα** σε όλο τον πίνακα, ανεξάρτητα από βελτιστοποιητή, σε σύγκριση με “**όλα τα χαρακτηριστικά**”.
  - Ρυθμίσεις με **ιδιαίτερη επιτυχία**:
    - **Εξειδικευμένοι νευρώνες + κανονικοποίηση στόχων + AdamW**.
    - **Κοινοί νευρώνες + κανονικοποίηση στόχων + Levenberg–Marquardt (LM)**.
- **Σύνολο 4 χαρακτηριστικών** που ορίσαμε:
  - **Συστηματικά καλύτερα αποτελέσματα** σε όλο τον πίνακα, ανεξάρτητα από βελτιστοποιητή, σε σύγκριση με “**όλα τα χαρακτηριστικά**”, αλλά χειρότερα αποτελέσματα σε σύγκριση με το σύνολο 2 χαρακτηριστικών.



Γράφημα 14 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW και με κανονικοποίηση στόχων και 26 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά.



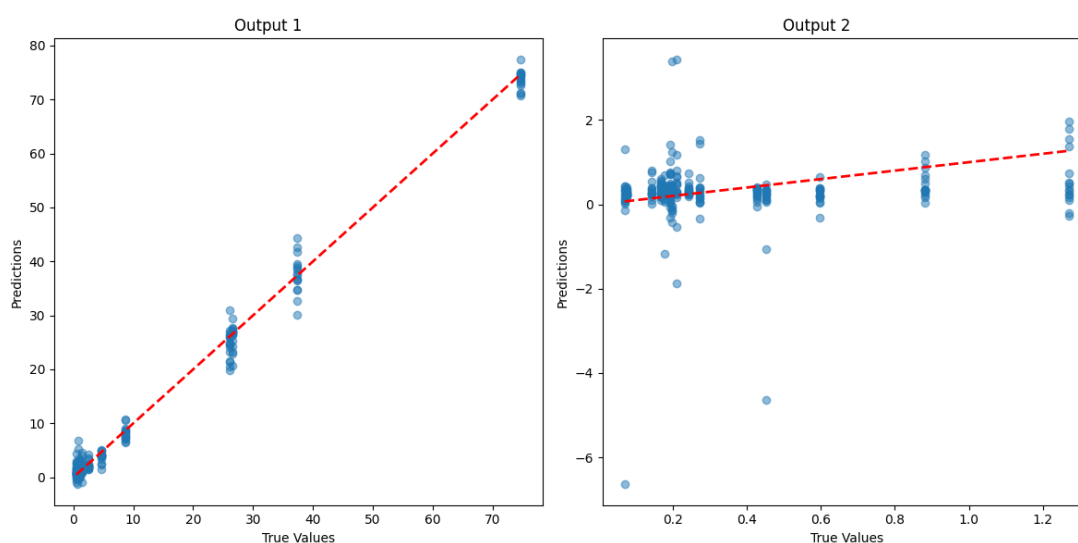
Γράφημα 15 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: LM και με κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά.

### 6.5.2 Αξιοσημείωτες περιπτώσεις

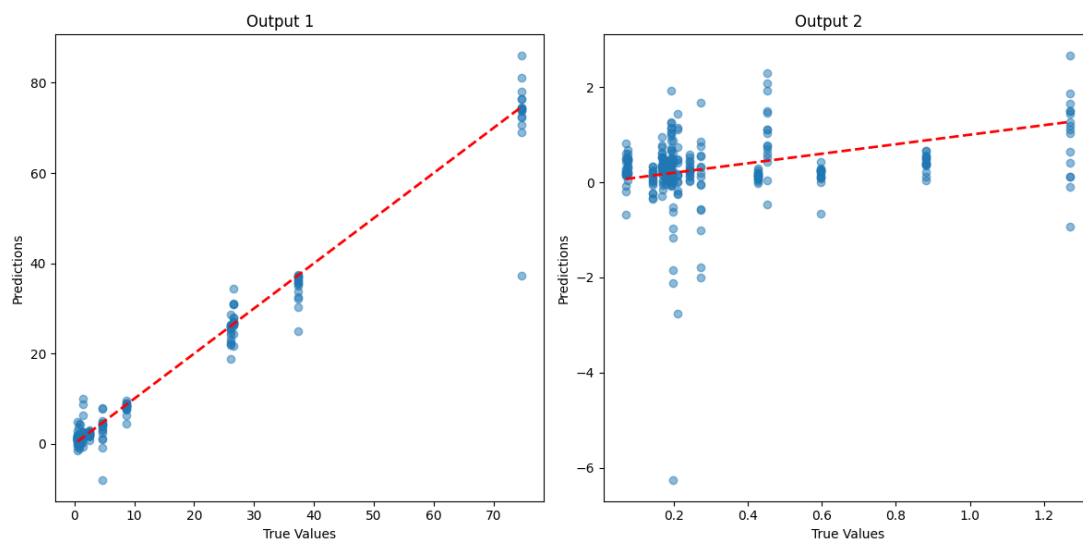
Τέλος, χωρίς κανονικοποίηση τιμών στόχων καταγράφηκε **ιδιαίτερα υψηλή ακρίβεια** στην **πρόβλεψη μέγιστης τάσης VM**, με τις καλύτερες επιδόσεις να σημειώνονται από :

- **AdamW+εξειδικευμένοι νευρώνες**
- **LM+κοινοί νευρώνες**

Τα αντίστοιχα γραφήματα παρατίθενται παρακάτω.



Γράφημα 16 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: AdamW και χωρίς κανονικοποίηση στόχων και 26 εξειδικευμένους νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά.



Γράφημα 17 προβλέψεων-πραγματικών τιμών για χαρακτηριστικά εισόδου: διαγνωστική διάμετρος ΑΚΑ και διάμετρος δεξιάς λαγονίου, βελτιστοποιητή: LM και χωρίς κανονικοποίηση στόχων και 25 κοινούς νευρώνες στα κρυφά επίπεδα. Σύνολο δεδομένων Ξυδιά.

## Κεφάλαιο 7

### Ανάλυση αποτελεσμάτων

Σκοπός αυτής της ενότητας είναι η ερμηνεία των ευρημάτων των πειραμάτων ως προς τη χρήση βελτιστοποιητή, αρχιτεκτονικής, κανονικοποίησης στόχων και επιλογής χαρακτηριστικών.

Αρχικά υπογραμμίζεται ότι οι απόλυτες μετρικές σφάλματος, όπως η MAE, επηρεάζονται από την κλίμακα των τιμών στόχων, ενώ ο δείκτης  $R^2$  αποτιμά τη σχετική επίδοση σε σχέση με τη διασπορά των πραγματικών τιμών.

Στο σύνολο δεδομένων Κυπαρίσση παρατηρήθηκε ότι η μέγιστη τάση VM εμφανίζει χαμηλό MAE αλλά και χαμηλότερο  $R^2$ , ενώ η μέγιστη μετατόπιση εμφανίζει υψηλότερο MAE αλλά και υψηλότερο  $R^2$ . Η συμπεριφορά αυτή εξηγείται από τα εύρη των τιμών στόχων:

- Μέγιστη τάση VM: 0.25-0.7 MPa (μικρό εύρος)
- Μέγιστη μετατόπιση: 7-35 mm (μεγαλύτερο εύρος)

Αντιστοίχως, στο σύνολο δεδομένων Ξυδιά με εύρη τιμών στόχων:

- Μέγιστη τάση VM: 0.497 έως 74.656 MPa (μεγάλο εύρος)
- Μέγιστη μετατόπιση: 0.0695 έως 1.2705 mm (μικρότερο εύρος)

καταγράφεται το αντίστροφο μοτίβο, όπου παρατηρούμε υψηλότερα MAE και  $R^2$  στη μέγιστη τάση VM και χαμηλότερα MAE και  $R^2$  στη μέγιστη μετατόπιση. Επισημαίνεται ότι, η ανάλυση γίνεται με συνδυασμό των μεγεθών σφάλματος και διασποράς.

Αν συγκρίνουμε την επίδοση των μοντέλων με βάση τους βελτιστοποιητές, παρατηρούμε ότι συνολικά ο AdamW υπερέχει συστηματικά σε MAE και  $R^2$  συχνά στις προβλέψεις και των δύο στόχων. Ο AdamW λόγω των χαρακτηριστικών του, που συζητήθηκαν στο κεφάλαιο 3.6 (προσαρμοστικός ρυθμός μάθησης, decoupled weight decay), παρουσιάζει σταθερότερη γενίκευση σε μικρά/θορυβώδη σύνολα. Ο LM είναι αποδοτικός σε πολύ μικρά δίκτυα, αλλά με λίγα δείγματα εμφανίζει υπερπροσαρμογή (overfitting).

Η κανονικοποίηση των τιμών στόχων επηρεάζει ουσιαστικά τη βελτιστοποίηση όταν χρησιμοποιείται κοινή συνάρτηση σφάλματος τύπου MSE.

Με κανονικοποίηση στο σύνολο δεδομένων Κυπαρίσση, βελτιώνεται αισθητά η επίδοση στη μέγιστη τάση VM, ενώ χωρίς κανονικοποίηση βελτιώνεται η επίδοση στη μέγιστη μετατόπιση.

Με κανονικοποίηση στο σύνολο δεδομένων Ξυδιά, βελτιώνεται αισθητά η επίδοση στη μέγιστη μετατόπιση, ενώ χωρίς κανονικοποίηση βελτιώνεται η επίδοση στη μέγιστη τάση VM.

Αυτό οφείλεται στις διαφορετικές κλίμακες στόχων σε συνδυασμό με την κοινή MSE για τον υπολογισμό σφάλματος κατά την εκπαίδευση. Έτσι δημιουργούνται ανισόρροπες κλίσεις και η εκπαίδευση ευνοεί τον στόχο με τη μεγαλύτερη κλίμακα, ώστε να πετύχει μεγαλύτερη/ταχύτερη μείωση σφάλματος.

Η ωφέλεια του διαχωρισμού των νευρώνων ανά τιμή στόχου εξαρτάται από τις στατιστικές σχέσεις μεταξύ των στόχων και το υποκείμενο μοντέλο προσομοίωσης

Στο σύνολο δεδομένων του **Κυπαρίσση**, οι **κοινοί νευρώνες** παρουσιάζουν συστηματικά **καλύτερες επιδόσεις**.

Στο σύνολο δεδομένων του **Ξυδιά**, οι **εξειδικευμένοι νευρώνες** παρουσιάζουν συστηματικά **καλύτερες επιδόσεις**.

Η επίδοση της αρχιτεκτονικής εξειδικευμένων νευρώνων εξαρτάται από τη διαδικασία προσομοίωσης της οποίας τα αποτελέσματα προσπαθούμε να προβλέψουμε. Με τη χρήση εξειδικευμένων νευρώνων μειώνεται η παρεμβολή μεταξύ στόχων, αλλά στη περίπτωση του Κυπαρίσση έχουμε θεμιτή (εποικοδομητική) παρεμβολή, ενώ στη περίπτωση του Ξυδιά έχουμε μη-θεμιτή (καταστροφική) παρεμβολή. Η επιλογή αρχιτεκτονικής πρέπει επομένως να λαμβάνει υπόψη την αλληλεξάρτηση των στόχων.

Η χρήση όλων των χαρακτηριστικών ως είσοδοι έδωσε χειρότερα αποτελέσματα από επιλογή υποσυνόλων και συγκεκριμένα ξεχώρισε η επίδοση των υποσυνόλων που περιλαμβάνουν τα χαρακτηριστικά διάμετρος αριστερής και διάμετρος δεξιάς λαγονίου. Με λιγότερα και στενευμένα χαρακτηριστικά πετυχαίνουμε μείωση θορύβου και κινδύνου υπερπροσαρμογής ειδικά σε μικρό σύνολο δεδομένων. Τα γεωμετρικά χαρακτηριστικά των διαμέτρων των λαγονίων πιθανώς κωδικοποιούν κρίσιμες πτυχές μορφολογίας που σχετίζονται με τις καταπονήσεις που προσομοιώνει η FEM.

## Κεφάλαιο 8

### Συμπεράσματα

Στην παρούσα εργασία μελετήσαμε τη δυνατότητα χρήσης νευρωνικών δικτύων ως υποκατάστατα μοντέλα για την πρόβλεψη μεγεθών ενδιαφέροντος από προσομοιώσεις πεπερασμένων στοιχείων (FEM) σε ανευρύσματα κοιλιακής αορτής.

Κύριο εύρημα αποτελεί ότι ο συνδυασμός AdamW, κανονικοποίησης τιμών στόχων και στοχευμένης επιλογής χαρακτηριστικών οδηγεί κατά κανόνα σε υψηλότερη ακρίβεια και χαμηλότερη διακύμανση. Η προτίμηση μεταξύ κοινών και εξειδικευμένων νευρώνων εξαρτάται από το εκάστοτε σύνολο δεδομένων και τις στατιστικές σχέσεις των στόχων.

Τα αποτελέσματα αναδεικνύουν τόσο τις δυνατότητες όσο και τους περιορισμούς της προσέγγισης. Με τα μικρά πλήθη δειγμάτων που έχουμε διαθέσιμα, οι αποδόσεις είναι ενθαρρυντικές αλλά μη γενικεύσιμες χωρίς περαιτέρω επικύρωση. Ως εκ τούτου, τα προτεινόμενα μοντέλα δεν είναι έτοιμα για κλινική χρήση ή λήψη αποφάσεων. Παρ' όλα αυτά, δείχνουν σαφή ικανότητα προσέγγισης των μεγεθών που παράγονται από FEM και υποδεικνύουν ότι με μεγαλύτερα και πιο ετερογενή σύνολα δεδομένων και αυστηρότερη μεθοδολογία επικύρωσης, τα νευρωνικά δίκτυα μπορούν να αποτελέσουν χρήσιμο συμπληρωματικό εργαλείο εκτίμησης κινδύνου.

## Κεφάλαιο 9

### Μελλοντική έρευνα

Τα μοντέλα που εκπαιδεύτηκαν περιορίζονται από το μικρό πλήθος των διαθέσιμων δεδομένων. Καθώς πρόκειται για σοβαρό ιατρικό πρόβλημα, προτεραιότητα αποτελεί η διεύρυνση και ετερογένεια των δειγμάτων και η δημιουργία ακριβέστερων συνόλων δεδομένων.

Παράλληλα κρίνεται χρήσιμος ο εμπλουτισμός των διαθέσιμων χαρακτηριστικών με μορφολογικούς δείκτες από την ανακατασκευή (π.χ. τοπική καμπυλότητα, δείκτες σχήματος ανευρύσματος) και πληροφορίες για ύπαρξη θρόμβων, όπου είναι διαθέσιμες, που σχετίζονται με τις κατανομές τάσεων.

Όσον αφορά τη συνάρτηση σφάλματος, μπορεί να εξεταστεί δυναμική στάθμιση των απωλειών ή και η χρήση robust συναρτήσεων σφάλματος π.χ. Huber, για μείωση της επίδρασης ακραίων τιμών.

Μια άλλη κατεύθυνση είναι η διερεύνηση PINN, όπου νόμοι της φυσικής ενσωματώνονται στη συνάρτηση κόστους, ώστε το δίκτυο να μαθαίνει από δεδομένα και εξισώσεις, ιδιαίτερα χρήσιμο σε καθεστώτα λίγων δειγμάτων.

Τέλος, μια decision-curve analysis θα επιτρέψει την αποτίμηση της κλινικής ωφέλειας για διαφορετικά κατώφλια των τιμών προβλέψεων, στοιχείο κρίσιμο για την αξιολόγηση του μοντέλου ως εργαλείου λήψης αποφάσεων.

## Βιβλιογραφία

- [1] M. A. Nielsen, *Neural Networks and Deep Learning*, Determination Press, 2015.
- [2] Χ.-Φ. Ξυδιάς, «Μελέτη μη-γραμμικών μοντέλων ανευρυσματικών αορτών με χρήση μεθόδου πεπερασμένων στοιχείων,» Χανιά, Ελλάς, 2025.
- [3] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov και M. DePristo, «A guide to deep learning in healthcare,» *Nat Med*, τόμ. 25, αρ. 1, pp. 24-29, 2019.
- [4] Κ. Κυπαρίσσης, «Προσομοίωση και ανάλυση αορτών με ανεύρυσμα με χρήση μεθόδου πεπερασμένων στοιχείων,» Χανιά, Ελλάς, 2024.
- [5] M. F. Fillinger, M. L. Raghavan, S. P. Marra, J. L. Cronenwett και F. E. Kennedy, «In vivo analysis of mechanical wall stress and abdominal aortic aneurysm rupture risk,» *JOURNAL OF VASCULAR SURGERY*, τόμ. 36, αρ. 3, pp. 589-597, 2002.
- [6] I. Goodfellow, Y. Bengio και A. Courville, *Deep Learning*, Cambridge, MA: MIT Press, 2016.
- [7] C. M. Bishop, *Pattern Recognition and Machine Learning*, New York, NY: Spring, 2006.
- [8] X. Glorot, A. Bordes και Y. Bengio, «Deep sparse rectifier neural networks,» σε *AISTATS*, Fort Lauderdale, FL, USA, 2011.
- [9] V. Nair και G. E. Hinton, «Rectified Linear Units Improve Restricted Boltzmann Machines,» σε *ICML*, Haifa, Israel,, 2010.
- [10] A. L. Maas, A. Y. Hannun και A. Y. Ng, «Rectifier Nonlinearities Improve Neural Network Acoustic Models,» σε *JMLR*, Atlanta, Georgia, USA, 2013.
- [11] K. He, X. Zhang, S. Ren και J. Sun, «Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification,» σε *ICCV*, 2015.
- [12] D. Masters και C. Luschi, «REVISITING SMALL BATCH TRAINING FOR,» *CoRR*, τόμ. abs/1804.07612, 2018.



- [13] L. Prechelt, «Early Stopping - But When?,» σε *Neural Networks: Tricks of the Trade*, τόμ. 1524, Berlin, Heidelberg, Springer, 1998, pp. 55-69.
- [14] A. Krogh και J. A. Hertz, «A Simple Weight Decay Can Improve Generalization,» σε *Advances in Neural Information Processing Systems 4 (NeurIPS)*, 1992.
- [15] I. Sutskever, M. James, G. Dahl και G. Dahl, «On the importance of initialization and momentum in deep learning,» σε *Proceedings of the 30th International Conference on Machine Learning, PMLR*, Atlanta, Georgia, USA, 2013.
- [16] D. P. Kingma και J. Ba, «Adam: A Method for Stochastic Optimization,» σε *ICLR*, 2015.
- [17] I. Loshchilov και F. Hutter, «DECOUPLED WEIGHT DECAY REGULARIZATION,» σε *ICLR*, 2019.
- [18] K. Levenberg, «A METHOD FOR THE SOLUTION OF CERTAIN NON-LINEAR PROBLEMS IN LEAST SQUARES,» *Quarterly of Applied Mathematics*, τόμ. 2, αρ. 2, pp. 164-168, 1944.
- [19] D. W. Marquardt, «An Algorithm for Least-Squares Estimation of Nonlinear Parameters,» *Society for Industrial and Applied Mathematics*, τόμ. 11, αρ. 2, pp. 431-441, 1963.
- [20] M. T. Hagan και M. B. Menhaj, «Training feedforward networks with the Levenberg–Marquardt algorithm,» *IEEE TRANSACTIONS ON NEURAL NETWORKS*, τόμ. 5, αρ. 6, p. 989–993, 1994.
- [21] R. Kohavi, «A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection,» σε *IJCAI*, 1995.
- [22] M. Stone, «Cross-Validatory Choice and Assessment of Statistical Predictions,» *Journal of the Royal Statistical Society: Series B (Methodological)*, τόμ. 36, αρ. 2, pp. 111-133, 1974.
- [23] R. Caruana, «Multitask Learning,» *Machine Learning*, τόμ. 28, p. 41–75, 1997.