Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών
Τομέας Ηλεκτρικών Βιομηχανικών Διατάξεων & Συστημάτων Αποφάσεων

Ανάπτυξη Προβλεπτικών Μοντέλων Χρονοσειρών με Εφαρμογή σε Συστήματα Boiler

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΓΕΩΡΓΙΟΣ ΚΑΛΑΡΑΣ

Επιβλέπων : Ευάγγελος Μαρινάκης
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2025

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών
Τομέας Ηλεκτρικών Βιομηχανικών Διατάξεων & Συστημάτων Αποφάσεων

Ανάπτυξη Προβλεπτικών Μοντέλων Χρονοσειρών με Εφαρμογή σε Συστήματα Boiler

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΓΕΩΡΓΙΟΣ ΚΑΛΑΡΑΣ

Επιβλέπων : Ευάγγελος Μαρινάκης
Επίκουρος Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 3η Ιουλίου 2025.

.....	.....	.....
Ευάγγελος Μαρινάκης	Ιωάννης Ψαρράς	Δημήτριος Ασκούνης
Επίκουρος Καθηγητής Ε.Μ.Π	Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π

Αθήνα, Ιούλιος 2025

.....
Γεώργιος Καλαράς

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Γεώργιος Καλαράς, 2025.
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Στη σύγχρονη εποχή, η ανάγκη περιορισμού των εκπομπών αερίων του θερμοκηπίου καθιστά επιτακτική τη μείωση της κατανάλωσης ορυκτών καυσίμων. Ένα κρίσιμο βήμα προς αυτή την κατεύθυνση είναι η βελτιστοποίηση της λειτουργίας των οικιακών καυστήρων φυσικού αερίου, κάτι που προϋποθέτει τη δυνατότητα πρόβλεψης του επιπέδου λειτουργίας τους. Η παρούσα διπλωματική εργασία εστιάζει στη χρήση μοντέλων τύπου Transformer για την πρόβλεψη της λειτουργίας των οικιακών boiler, καθώς και στη σύγκριση της ακρίβειας των προβλέψεών τους με αυτές μοντέλων μακράς βραχύχρονης μνήμης (LSTM). Το πρόβλημα θα μελετηθεί από δύο διαφορετικές πτυχές. Από τη μία, την πρόβλεψη της συμπεριφοράς του boiler ενός μεμονωμένου σπιτιού. Από την άλλη, την προσαρμογή των προβλεπτικών μοντέλων ώστε να εφαρμόζονται σε ένα σύνολο σπιτιών, επιτυγχάνοντας αξιόπιστες προβλέψεις σε συνολικό επίπεδο. Η μελέτη εστιάζει τόσο στην ακρίβεια των παραγόμενων προβλέψεων όσο και στον εντοπισμό των συνδυασμών κατηγοριών των δεδομένων εισόδου που οδηγούν στα καλύτερα αποτελέσματα για κάθε περίπτωση. Για κάθε μία από τις δύο πτυχές του προβλήματος, θα συγκριθούν τα μοντέλα με τις καλύτερες επιδόσεις. Αυτά θα προκύψουν από επιλογή χαρακτηριστικών που, μεμονωμένα οδηγούν σε καλές προβλέψεις, ενώ όταν συνδυάζονται ως είσοδος αποτυπώνουν αποτελεσματικά τη λειτουργία των boiler. Επιπλέον, με βάση αυτές τις επιλογές, θα υλοποιηθεί και ένας γενετικός αλγόριθμος για την αναζήτηση επιπλέον συνδυασμών χαρακτηριστικών που μπορούν να παράξουν βελτιωμένα αποτελέσματα πρόβλεψης. Τελικά, τα ευρήματα αποκαλύπτουν ότι τα μοντέλα Transformer μπορούν με κατάλληλα δεδομένα στην είσοδο να παράξουν αρκετά ακριβείς βραχύχρονες προβλέψεις και τις δύο προσεγγίσεις.

Λέξεις-Κλειδιά

οικιακοί λέβητες, λέβητες φυσικού αερίου, πρόβλεψη επιπέδου λειτουργίας, νευρωνικά δίκτυα, μετασχηματιστές, δίκτυα μακράς βραχύχρονης μνήμης, LSTM, γενετικός αλγόριθμος

Abstract

In the modern era, the need to limit greenhouse gas emissions makes the reduction of fossil fuel consumption imperative. A critical step in this direction is optimizing the operation of household natural gas boilers, which requires the ability to predict their modulation level. This thesis focuses on the use of Transformer-type models to predict the operation of residential boilers and compares the accuracy of their predictions with those of Long Short-Term Memory (LSTM) models. The problem will be examined from two different perspectives. On the one hand, predicting the behavior of the boiler in a single house. On the other hand, adapting the predictive models to be applicable to a group of houses, achieving reliable predictions at an aggregate level. The study focuses both on the accuracy of the generated predictions and on identifying the combinations of input data categories that lead to the best results in each case. For each of the two aspects of the problem, the best-performing models will be compared. These models will result from selecting features that individually lead to good predictions and, when combined as input, effectively represent the operation of boilers. Furthermore, based on these selections, a genetic algorithm will be implemented to search for additional combinations of features that can produce improved prediction results. Ultimately, the findings reveal that Transformer models, when provided with suitable input data, can produce fairly accurate short-term predictions in both approaches.

Key-Words

domestic boilers, natural gas boilers, modulation level forecast, neural networks, transformers, long short-term memory (LSTM), genetic algorithm

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή κ. Ευάγγελο Μαρινάκη για την ευκαιρία που μου έδωσε να εκπονήσω τη διπλωματική μου εργασία στο συγκεκριμένο θέμα. Επίσης, ευχαριστώ ιδιαίτερα τον διδάκτορα Βασίλη Μιχαλακόπουλο για την πολύτιμη βοήθεια και την άψογη καθοδήγηση που μου παρείχε καθ' όλη τη διάρκεια εκπόνησης της παρούσας εργασίας. Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου για την στήριξη και τη βοήθεια που μου προσέφεραν όλα αυτά τα χρόνια.

Περιεχόμενα

Περίληψη **6**
Abstract **8**
Ευχαριστίες **10**
Περιεχόμενα **12**
Κατάλογος Σχημάτων **14**
Κατάλογος Πινάκων **16**
1 Εισαγωγή **18**
1.1 Εισαγωγή στο Πρόβλημα 18
1.2 Σκοπός της Εργασίας 19
1.3 Βιβλιογραφική Ανασκόπηση 19
1.4 Συνεισφορά 21
1.5 Δομή Εργασίας 22
2 Θεωρητικό Υπόβαθρο **24**
2.1 Βαθιά Μάθηση 24
2.1.1 Νευρώνας 24
2.1.2 Αρχιτεκτονική Νευρωνικών Δικτύων 25
2.2 Recurrent Neural Networks (RNN) 26
2.2.1 Αρχιτεκτονική των RNN 26
2.3 Long Short-Term Memory (LSTM) 28
2.3.1 Αρχιτεκτονική των LSTM 28
2.4 Transformers 32
2.4.1 Αρχιτεκτονική των Transformer 33
2.5 Γενετικός Αλγόριθμος 36
2.5.1 Βασικές Έννοιες Γενετικού Αλγορίθμου 37
3 Πειραματικό Μέρος **39**
3.1 Εργαλεία 39
3.2 Δεδομένα 40
3.3 Μετρικές Σφαλμάτων 43
3.3.1 Mean Absolute Error (MAE) 43
3.3.2 Mean Squared Error (MSE) 43
3.3.3 Root Mean Squared Error (RMSE) 44
3.3.4 R-Squared (R2 - Coefficient of Determination) 44
3.4 Μετρικές Επίδοσης 45
3.4.1 Accuracy 45
3.4.2 Precision 46
3.4.3 Recall (True Positive Rate) 46
3.4.4 False Positive Rate 46
3.5 Πειραματικός Σχεδιασμός 47
3.5.1 Προεπεξεργασία Δεδομένων 49
3.5.2 Προετοιμασία Δεδομένων 50
3.5.3 Δομή Μοντέλου LSTM 52

3.5.4 Δομή Μοντέλου Transformer 52
3.5.5 Υλοποίηση του Γενετικού Αλγορίθμου 53
4 Αποτελέσματα 56
4.1 Επιλογή Βέλτιστης Διάρκειας Πρόβλεψης 56
4.2 Αποτελέσματα Πρόβλεψης σε Ένα Σπίτι 61
4.2.1 Απλοί συνδυασμοί κατηγοριών 61
4.2.2 Επιλογή Πιθανώς Βέλτιστων Μοντέλων 63
4.2.3 Γενετική Παραγωγή Μοντέλων 65
4.2.4 Σύγκριση Αποτελεσμάτων για Ένα Σπίτι 67
4.3 Αποτελέσματα Πρόβλεψης σε Σύνολο Σπιτιών 70
4.3.1 Επιλογή Πιθανώς Βέλτιστων Μοντέλων 70
4.3.2 Γενετική Παραγωγή Μοντέλων 72
4.3.3 Σύγκριση Αποτελεσμάτων για Σύνολο Σπιτιών 73
5 Συμπεράσματα και Μελλοντική Έρευνα 77
5.1 Συμπεράσματα 77
5.2 Μελλοντική Έρευνα 78
6 Βιβλιογραφία 80

Κατάλογος Σχημάτων

Σχήμα 2.1: Σχηματική αναπαράσταση της λειτουργίας ενός Νευρώνα 25
Σχήμα 2.2: Βαθύ νευρωνικό δίκτυο με 3 κρυφά επίπεδα 26
Σχήμα 2.3: RNN συμπιεσμένο (αριστερά) και ξεδιπλωμένο στο χρόνο (δεξιά) 27
Σχήμα 2.4: Εικόνα Δομής ενός LSTM 29
Σχήμα 2.5: Εικόνα Πύλης Λήθης LSTM 30
Σχήμα 2.6: Εικόνα Πύλης Εισόδου LSTM 31
Σχήμα 2.7: Εικόνα Πύλης Εξόδου LSTM 32
Σχήμα 2.9: Μηχανισμός κλιμακούμενης προσοχής με εσωτερικό γινόμενο (αριστερά) και Μηχανισμός πολλών κεφαλών προσοχής (δεξιά) 35
Σχήμα 2.10: Μεθοδολογία του Γενετικού Αλγορίθμου 36
Σχήμα 2.11: Διασταύρωση Γονιδίων Γονέων για παραγωγή Απογόνου 37
Σχήμα 2.12: Μετάλλαξη Γονιδίων Απογόνου 37
Σχήμα 3.1: Πίνακας Σφαλμάτων (Confusion Matrix) 45
Σχήμα 3.2: Σχηματική Απεικόνιση του Πειραματικού Σχεδιασμού 48
Σχήμα 3.3: Σχηματική Απεικόνιση της Προεπεξεργασίας των Δεδομένων 50
Σχήμα 3.4: Σχηματική Απεικόνιση της Προετοιμασίας των Δεδομένων 52
Σχήμα 3.5: Σχηματική Απεικόνιση της Λειτουργίας του Γενετικού Αλγορίθμου 54
Σχήμα 4.1: MAE και RMSE για Μοντέλα Αρχικής Δοκιμής Βέλτιστης Διάρκειας Πρόβλεψης 57
Σχήμα 4.2: Accuracy και Precision για Μοντέλα Αρχικής Δοκιμής Βέλτιστης Διάρκειας Πρόβλεψης 57
Σχήμα 4.3: Πρόβλεψη τυχαίας ημέρας από το test set για Transformer με είσοδο 'blr_mod_lvl' και πρόβλεψη για 1440 λεπτά 58
Σχήμα 4.4: MAE και RMSE για Μοντέλα Σύγκρισης Βέλτιστης Διάρκειας Πρόβλεψης 59
Σχήμα 4.5: Accuracy και Precision για Μοντέλα Σύγκρισης Βέλτιστης Διάρκειας Πρόβλεψης 59
Σχήμα 4.6: Πρόβλεψη τυχαίας ημέρας από το test set για Transformer με είσοδο 'blr_mod_lvl' και πρόβλεψη για 10 λεπτά 59
Σχήμα 4.7: MAE και RMSE για μοντέλα απλών συνδυασμών κατηγοριών σε ένα σπίτι 62
Σχήμα 4.8: Accuracy και Precision για μοντέλα απλών συνδυασμών κατηγοριών σε ένα σπίτι 62
Σχήμα 4.9: MAE και RMSE για μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε ένα σπίτι 64
Σχήμα 4.10: Accuracy και Precision για μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε ένα σπίτι 65
Σχήμα 4.11: Διαδικασία εκτέλεσης γενετικού αλγορίθμου σε περίπτωση για ένα σπίτι 66
Σχήμα 4.12: MAE και RMSE για συνολικά βέλτιστα μοντέλα στην περίπτωση ενός σπιτιού 68
Σχήμα 4.13: Accuracy και Precision για συνολικά βέλτιστα μοντέλα στην περίπτωση ενός σπιτιού 68

Σχήμα 4.14: MAE και RMSE για μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε σύνολο σπιτιών 71
Σχήμα 4.15: Accuracy και Precision για μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε σύνολο σπιτιών 71
Σχήμα 4.16: Διαδικασία εκτέλεσης γενετικού αλγορίθμου σε περίπτωση για σύνολο σπιτιών 72
Σχήμα 4.17: MAE και RMSE για συνολικά βέλτιστα μοντέλα στην περίπτωση συνόλου σπιτιών 74
Σχήμα 4.18: Accuracy και Precision για συνολικά βέλτιστα μοντέλα στην περίπτωση συνόλου σπιτιών 74

Κατάλογος Πινάκων

Πίνακας 3.1: Πίνακας Ανάλυσης και Ερμηνείας Μεταβλητών των Μεγεθών 42
Πίνακας 3.2: Πίνακας Ανάλυσης και Ερμηνείας των Νέων Μεγεθών 49
Πίνακας 4.1: Μοντέλα Αρχικής Δοκιμής Βέλτιστης Διάρκειας Πρόβλεψης 56
Πίνακας 4.2: Μοντέλα Σύγκρισης Βέλτιστης Διάρκειας Πρόβλεψης 58
Πίνακας 4.3: Μοντέλα απλών συνδυασμών κατηγοριών σε ένα σπίτι 61
Πίνακας 4.4: Μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε ένα σπίτι 64
Πίνακας 4.5: Βέλτιστο μοντέλο σε ένα σπίτι από γενετικό αλγόριθμο 66
Πίνακας 4.6: Συνολικά βέλτιστα μοντέλα στην περίπτωση ενός σπιτιού 67
Πίνακας 4.7: Μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε σύνολο σπιτιών 70
Πίνακας 4.8: Βέλτιστο μοντέλο σε σύνολο σπιτιών από γενετικό αλγόριθμο 73
Πίνακας 4.9: Συνολικά βέλτιστα μοντέλα στην περίπτωση συνόλου σπιτιών 74

1 Εισαγωγή

1.1 Εισαγωγή στο Πρόβλημα

Γενικά, όσον αφορά τους λέβητες (boilers), πρόκειται για συσκευές θέρμανσης νερού ή παραγωγής ατμού που αξιοποιούνται από διάφορες κατηγορίες κτιρίων, όπως εργοστάσια και κατοικίες. Οι χρήσεις τους είναι ποικίλες και περιλαμβάνουν την παραγωγή ενέργειας, την θέρμανση χώρων και την παραγωγή ζεστού νερού. Η βασική ιδέα λειτουργίας τους είναι η μεταφορά θερμότητας από μία διαδικασία καύσης (combustion), συνήθως πετρελαίου, φυσικού αερίου ή άνθρακα, σε σώμα νερού. Όμως, λόγω της προόδου της επιστήμης και της συνεχούς αύξησης της ζήτησης έχουν εξελιχθεί σε πιο περίπλοκες συσκευές. Ο κάθε κατασκευαστής ακολουθεί διαφορετική μεθοδολογία και έτσι έχουν προκύψει διαφορετικά είδη boiler, το κάθε ένα με τα δικά του ιδιαίτερα χαρακτηριστικά. [1] Δεδομένης της ευρύτητας της χρήσης των boiler για διάφορους σκοπούς, η βελτίωση της λειτουργίας τους θα έχει σημαντικά οφέλη για την εξοικονόμηση ορυκτών καυσίμων και την μείωση εκπομπών διοξειδίου του άνθρακα (CO2). [2]

Η σημερινή εποχή χαρακτηρίζεται από την στροφή προς ανανεώσιμες πηγές ενέργειας καθώς επίσης, από το διαρκώς αυξανόμενο ενδιαφέρον για την αξιοποίηση μεθόδων εξοικονόμησης ενέργειας και την χρήση “πράσινων” τεχνολογιών. [3] Η κύρια αιτία για την επιλογή αυτής της κατεύθυνσης είναι περιβαλλοντικοί λόγοι. Άλλωστε τα τελευταία χρόνια γίνονται έντονες προσπάθειες αποδέσμευσης από τη χρήση των ορυκτών καυσίμων, με στόχο την μείωση του αντίκτυπου που έχει η χρήση τους. Μάλιστα, είναι χαρακτηριστικό ότι η Ευρωπαϊκή Ένωση έχει θέσει ένα σύνολο στόχων για τις επόμενες δεκαετίες, το οποίο αποσκοπεί να την καταστήσει πιο φιλική προς το περιβάλλον. Μέρος αυτών των στόχων είναι αρχικά ο περιορισμός των εκπομπών αερίων του θερμοκηπίου στο 55% των επιπέδων του 1900 μέχρι και το 2030 και στη συνέχεια η πλήρης εξάλειψή τους μέχρι το 2050. [4]

Αυτό οδηγεί σε αλλαγές της προσέγγισης και των αξιοποιούμενων τεχνολογιών σε τομείς στους οποίους είναι έντονη η χρήση μη ανανεώσιμων πηγών ενέργειας. Χαρακτηριστικό παράδειγμα είναι, όπως αναφέρθηκε προηγουμένως, η θέρμανση, η οποία σε περιπτώσεις οικιακής χρήσης πραγματοποιείται σε μεγάλο βαθμό με τη χρήση καυστήρων (boiler) φυσικού αερίου. Μάλιστα, σύμφωνα με άρθρο της Ευρωπαϊκή Επιτροπής, τα κτίρια είναι υπεύθυνα για το 40% της συνολικής κατανάλωσης ενέργειας και το 36% των εκπομπών αερίου του θερμοκηπίου. [5] Επομένως, είναι ιδιαίτερα σημαντική η εύρεση τρόπων βελτίωσης της λειτουργίας των boiler. Κάτι τέτοιο θα επέτρεπε την διατήρηση των επιθυμητών επιπέδων θέρμανσης και άνεσης για τους χρήστες, μειώνοντας παράλληλα την ποσότητα καυσίμου που καταναλώνεται από την χρήση τους.

Μία λύση προς αυτή την κατεύθυνση αποτελούν τα boiler ρυθμιζόμενου επιπέδου λειτουργίας (boiler modulation level). Η απόδοση τους, άρα και η κατανάλωσή τους σε φυσικό αέριο, είναι ρυθμιζόμενη ανάλογα με το επιθυμητό επίπεδο θέρμανσης. Σε αντίθεση με τα πιο “παραδοσιακά” boiler, τα οποία λειτουργούν πάντα στο 100% των δυνατοτήτων τους, μπορούν να ρυθμίζουν την ένταση λειτουργίας τους και σε χαμηλότερα επίπεδα.
Για να γίνει αξιοποίηση αυτής της δυνατότητας τους στο μέγιστο δυνατό βαθμό, είναι αναγκαία η ανάπτυξη μεθόδων οι οποίες θα μπορούν να προβλέπουν την αναμενόμενη απαίτηση λειτουργίας των boiler ώστε να καθορίζεται ένα “βέλτιστο” πλάνο κατανάλωσης.

Με αυτό τον τρόπο θα μπορεί να επιλέγεται ο πιο αποδοτικός ρυθμός λειτουργίας τους ώστε να επιτυγχάνεται το επιθυμητό επίπεδο θέρμανσης των οικιακών χώρων, διατηρώντας παράλληλα την κατανάλωση του φυσικού αερίου σε όσο το δυνατόν πιο χαμηλά επίπεδα.

1.2 Σκοπός της Εργασίας

Αντικείμενο της παρούσας εργασίας είναι η ανάπτυξη έξυπνων μεθόδων με σκοπό την πρόβλεψη του αναμενόμενου επιπέδου έντασης λειτουργίας του boiler . Αυτό θα επιτρέψει την θέρμανση των οικιακών χώρων στον επιθυμητό βαθμό με την ελάχιστη δυνατή κατανάλωση καυσίμου. Πρόκειται για ένα πρόβλημα το οποίο εκ της φύσης του είναι δύσκολο, καθώς εξαρτάται από πολλούς παράγοντες. Για τον λόγο αυτό, όπως θα εξηγηθεί και στην ενότητα 4, έχουμε στην διάθεσή μας μετρήσεις πολλών μεγεθών τα οποία σχετίζονται με το πρόβλημα. Πιο συγκεκριμένα, αποσκοπούμε στην εύρεση των συνδυασμών των μεγεθών τα οποία όταν δοθούν ως είσοδος σε μοντέλα Transformer, οδηγούν στην παραγωγή βέλτιστων προβλέψεων. Μάλιστα, θα μελετηθεί η περίπτωση βελτιστοποίησης των προβλέψεων σε ένα boiler (ενός σπιτιού) καθώς και η περίπτωση προσαρμογής ενός μοντέλου στην πρόβλεψη ενός συνόλου boiler (ενός συνόλου σπιτιών).

1.3 Βιβλιογραφική Ανασκόπηση

Η δυνατότητα της πρόβλεψης της λειτουργίας των οικιακών boiler είναι ιδιαίτερης αξίας για την μείωση της κατανάλωσης του φυσικού αερίου. Η συγκεκριμένη πτυχή του προβλήματος είναι σημαντική εξαιτίας του πλήθους των οικιακών καυστήρων φυσικού αερίου που συνεπάγεται μεγάλο όγκο κατανάλωσης ως σύνολο. Ως συνέπεια έχει αρχίσει να στρέφεται μέρος μελετών και προς αυτή την κατεύθυνση τα τελευταία χρόνια. Βέβαια, υπάρχουν έρευνες σε παραπλήσιους τομείς όπως είναι η πρόβλεψη της κατανάλωσης φυσικού αερίου από οικιακούς χρήστες. Από την άλλη, η λειτουργία των οικιακών boiler καθορίζεται από την ανάγκη για θέρμανση, η οποία είναι με τη σειρά της σε άμεση εξάρτηση από τις συνθήκες εσωτερικού χώρου όπως η θερμοκρασία. Πράγματι το πεδίο της πρόβλεψης της θερμοκρασίας του οικιακού χώρου έχει σημαντική βιβλιογραφία.

Όσον αφορά την πρόβλεψη του επιπέδου λειτουργίας των οικιακών boiler οι Tsoumalis et al. [1] πραγματοποίησαν μία έρευνα η οποία δίνει τη σχετική πρόβλεψη για τα επόμενα λεπτά. Συγκεκριμένα, η έρευνά τους αφορούσε την βελτιστοποίηση της λειτουργίας του boiler με μείωση της κατανάλωσης του φυσικού αερίου. Για να το πετύχουν αυτό στηρίχθηκαν, ανάμεσα στα άλλα και σε βραχύχρονη πρόβλεψη του επιπέδου λειτουργίας του boiler. Η έρευνά τους αξιοποίησε ένα boiler ρυθμιζόμενου επιπέδου λειτουργίας συνδεδεμένο στο διαδίκτυο μέσω του οποίου είχαν πρόσβαση στα δεδομένα σε πραγματικό χρόνο. Τα δεδομένα αυτά συμπληρώνονταν από προβλέψεις με νευρωνικά δίκτυα της

αναμενόμενης θερμοκρασίας εσωτερικού χώρου και του εκτιμώμενου επιπέδου λειτουργίας του boiler για τα επόμενα λεπτά.

Οι Özmen et al. [6] πραγματοποίησαν έρευνα για βελτιστοποίηση της κατανάλωσης φυσικού αερίου από οικιακούς χρήστες. Στην περίπτωσή τους δεν περιορίστηκαν μόνο στην περίπτωση θέρμανσης αλλά αξιοποίησαν συνολικά στοιχεία από την κατανάλωση φυσικού αερίου (για παράδειγμα περιλαμβάνονταν και χρήσεις για μαγείρεμα). Στην μελέτη τους πραγματοποίησαν προβλέψεις για την κατανάλωση φυσικού αερίου για την επόμενη ημέρα. Τα αποτελέσματα παρήχθησαν από γραμμικά μοντέλα (Linear Models - LR), από απλά νευρωνικά δίκτυα (Neural Networks - NN) δύο κρυφών επιπέδων και εστίασαν κυρίως στα αποτελέσματα των MARS (Multivariate Adaptive Regression Splines) και CMARS (Conic Multivariate Adaptive Regression Splines) αλγορίθμων οι οποίοι είχαν καλύτερη επίδοση από τα απλούστερα μοντέλα.
Από την άλλη, οι Saloux et al. [7] εστίασαν στην ανάπτυξη (και εφαρμογή) μοντέλων τεχνητής νοημοσύνης για την ανάπτυξη προβλεπτικών μοντέλων ελέγχου της λειτουργίας των boiler για διάστημα μίας ημέρας. Εστίασαν την μελέτη τους σε boiler φυσικού αερίου για θέρμανση περιοχών (και όχι για μεμονωμένα σπίτια). Λαμβάνοντας υπόψη τους προηγούμενα και τρέχοντα δεδομένα δοκίμασαν διαφορετικά είδη αλγορίθμων μηχανικής μάθησης όπως τεχνητά νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης (feedforward ANN), CNN (Convolutional Neural Networks), LSTM, random forests. Τα αποτελέσματα από την πρακτική εφαρμογή των μοντέλων τους έδειξαν μείωση της κατανάλωσης φυσικού αερίου.

Στο κομμάτι της πρόβλεψης θερμοκρασίας εσωτερικού χώρου, οι Mateo et al. [8] συνέκριναν πολλές διαφορετικές μεθόδους για την πρόβλεψη της θερμοκρασίας ενός σπιτιού. Χρησιμοποίησαν μεθόδους όπως αυτο-παλινδρόμηση (Auto-Regression - AR) και Robust Multiple Linear Regression (RMLR) καθώς επίσης μη-γραμμικές μεθόδους μηχανικής μάθησης όπως Extreme Learning Machine (ELM) και το Non-linear Autoregressive Exogenous Multilayer Perceptron (MLP-NARX). Το MLP-NARX είχε τα καλύτερα αποτελέσματα από όλες τις μεθόδους που δοκίμασαν.
Μοντέλο NARX χρησιμοποιήθηκε επίσης από τους Afroz et al. [9]. Στηριζόμενοι σε δεδομένα πραγματικού χρόνου από παραπλήσια τμήματα ενός εμπορικού κτιρίου πραγματοποίησαν προβλέψεις για την θερμοκρασία του εσωτερικού χώρου.
Οι Shi et al. [10] έκαναν χρήση ενός νευρωνικού μοντέλου ανατροφοδότησης (Backpropagation Neural Network - BP) για την πρόβλεψη της θερμοκρασίας (αλλά και της υγρασίας) εσωτερικού χώρου σε ένα βιομηχανικό κτίριο για διαστήματα των 10 λεπτών. Τα παραγόμενα συμπεράσματα στηρίχθηκαν μόνο στα σχετικά δεδομένα (και όχι σε άλλα παρεμφερή στοιχεία) και είχαν αρκετά καλή ακρίβεια.
Οι Monteiro et al. [11] προσέγγισαν την πρόβλεψη θερμοκρασίας εσωτερικού χώρου από τη σκοπιά του IoT (Internet of Things), αξιοποιώντας δεδομένα που παρείχε ένα “έξυπνο” ψυγείο. Προέβλεψαν τη θερμοκρασία με μεγάλη ακρίβεια, ενώ από τις διαφορετικές προσεγγίσεις που μελέτησαν, ο σταθμισμένος συνδυασμός προηγούμενων παρατηρήσεων απέδωσε καλύτερα από πιο σύνθετες εναλλακτικές όπως μοντέλα ARIMA (Autoregressive Moving Average).

1.4 Συνεισφορά

Η συγκεκριμένη εργασία εστιάζει στην ανάπτυξη μοντέλων Μετασχηματιστών (Transformer) τα οποία στηρίζονται σε δομή κωδικοποιητή - αποκωδικοποιητή (encoder - decoder) για την πραγματοποίηση της πρόβλεψης του επιπέδου λειτουργίας του boiler. Θα μελετηθούν και θα συγκριθούν οι διαφοροποιήσεις των μοντέλων ως προς την επιτυχία των προβλέψεων ανάλογα με τα επιλεγμένα κάθε φορά στοιχεία εισόδου. Συγκεκριμένα, θα δοκιμαστούν διαφορετικοί συνδυασμοί κατηγοριών δεδομένων εισόδου ώστε να προκύψουν εκείνοι οι οποίοι "περιγράφουν" καλύτερα τις συνθήκες λειτουργίας του boiler και άρα επιτρέπουν στα μοντέλα να είναι πιο ακριβή στις προβλέψεις τους. Για κάθε μία από τις παραλλαγές των δεδομένων εισόδου θα προκύψει ένα μοντέλο Transformer. Παράλληλα, με αυτό θα εκπαιδευτεί και ένα αντίστοιχο μοντέλο LSTM (μακράς βραχύχρονης μνήμης) ώστε να συγκριθούν οι προσεγγίσεις των δύο διαφορετικών τύπων νευρωνικών δικτύων.

Η εκπαίδευση και αξιολόγηση των μοντέλων θα εστιάσει σε δύο παράλληλες κατευθύνσεις. Από τη μία, στην βελτιστοποίηση των μοντέλων για πρόβλεψη σε ένα σπίτι. Σε αυτό το τμήμα θα βρεθούν τα χαρακτηριστικά τα οποία επιτρέπουν στα μοντέλα να προσαρμόζονται σε μεγαλύτερο βαθμό στην συμπεριφορά ενός μόνο boiler ώστε να παράγονται ακριβείς προβλέψεις για την λειτουργία του.
Από την άλλη, τα μοντέλα θα βελτιστοποιηθούν ως προς την δυνατότητά τους να προβλέπουν το επίπεδο λειτουργίας των boiler για ένα σύνολο σπιτιών (μελετάμε την περίπτωση 28 σπιτιών). Σε αυτή την περίπτωση, θα βρεθούν οι συνδυασμοί χαρακτηριστικών τα οποία επιτρέπουν στα μοντέλα να έχουν καλύτερη δυνατότητα πρόβλεψης επί του συνόλου των σπιτιών. Δηλαδή, εκείνες οι κατηγορίες δεδομένων οι οποίες επιτρέπουν στα μοντέλα να "μάθουν" τη γενικότερη φύση του προβλήματος ώστε να είναι σε θέση να παράξουν ακριβείς προβλέψεις για διαφορετικά σπίτια.

Για κάθε μία από τις περιπτώσεις θα συγκριθούν διαφορετικοί συνδυασμοί κατηγοριών δεδομένων. Εκείνες οι οποίες σταθερά συνεισφέρουν σε μοντέλα με μεγάλη επιτυχία στις προβλέψεις θα χρησιμοποιηθούν για την δημιουργία των τελικών βέλτιστων μοντέλων. Ωστόσο, είναι πιθανό ορισμένες κατηγορίες να μπορούν να συνεισφέρουν στη παραγωγή καλύτερων προβλέψεων αλλά να μην έχουν εντοπιστεί. Για να καλυφθεί και αυτή η πιθανότητα θα γίνει χρήση γενετικού αλγορίθμου.
Για αυτό το βήμα θα αξιοποιηθούν οι μετρήσεις για την ακρίβεια των προβλέψεων των μοντέλων κάθε περίπτωσης. Με αυτό τον τρόπο ο γενετικός αλγόριθμος θα έχει μία "βάση" ώστε να εκτιμά την προβλεπτική ακρίβεια των μοντέλων που προτείνει.
Έτσι, για κάθε μία από τις δύο περιπτώσεις (τα μοντέλα που βελτιστοποιούνται ώστε να προβλέπουν τη λειτουργία του boiler σε ένα σπίτι και εκείνα που προβλέπουν σε σύνολο σπιτιών) θα μπορέσουμε να συγκρίνουμε τα καλύτερα μοντέλα, τα οποία προέκυψαν αφενός κατά την αρχική επιλογή των πιθανώς βέλτιστων συνδυασμών κατηγοριών και αφετέρου από την εκτέλεση του γενετικού αλγορίθμου.

1.5 Δομή Εργασίας

Η συγκεκριμένη εργασία αποτελείται από πέντε κεφάλαια. Στο κεφάλαιο 1 πραγματοποιήθηκε μία εισαγωγή στο αντικείμενο της εργασίας και στον σκοπό υλοποίησής της. Επίσης παρουσιάστηκε υπάρχουσα βιβλιογραφία σχετική με το θέμα. Στο κεφάλαιο 2 καλύπτεται το θεωρητικό υπόβαθρο σχετικά με τα μοντέλα (όπως μετασχηματιστές και LSTM) και τις διαδικασίες (γενετικός αλγόριθμος) που χρησιμοποιήθηκαν κατά την υλοποίηση της εργασίας. Στο κεφάλαιο 3 εξηγούνται αναλυτικά όλα τα βήματα της πειραματικής διαδικασία που ακολουθήθηκε. Στο κεφάλαιο 4 παρουσιάζονται και αξιολογούνται τα αποτελέσματα που προέκυψαν. Τέλος, στο κεφάλαιο 5 αναφέρονται τα τελικά συμπεράσματα τα οποία προέκυψαν καθώς και κατευθύνσεις στις οποίες μπορεί να υλοποιηθεί μελλοντική έρευνα.

2 Θεωρητικό Υπόβαθρο

2.1 Βαθιά Μάθηση

Η βαθιά μάθηση είναι υποσύνολο της μηχανικής μάθησης που εστιάζει στην χρήση μοντέλων γνωστών ως νευρωνικών δικτύων. Τα μοντέλα αυτά, εμπνευσμένα από τη νευροεπιστήμη, αποτελούνται από πολλά διαδοχικά επίπεδα. Το κάθε επίπεδο αποτελείται με τη σειρά του από πολλούς νευρώνες. Ο όρος "βαθιά" αναφέρεται στη χρήση μεγάλου αριθμού διαδοχικών επιπέδων. Κάθε επίπεδο επεξεργάζεται διαφορετικά χαρακτηριστικά των δεδομένων εισόδου, επιτρέποντας στο μοντέλο να μαθαίνει πιο αφηρημένες αναπαραστάσεις δεδομένων. [12]

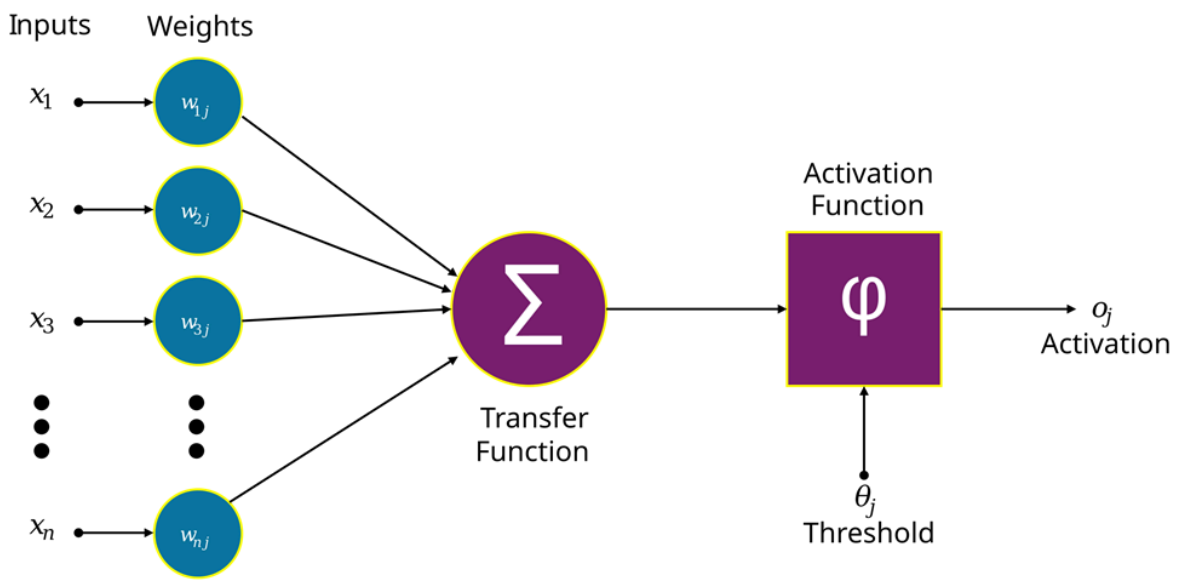
2.1.1 Νευρώνας

Ο νευρώνας αποτελεί τη δομική μονάδα των αποκαλούμενων νευρωνικών δικτύων. Πρόκειται για ένα μαθηματικό μοντέλο, η δομή και η λειτουργία του οποίου είναι αντίστοιχη με εκείνη των βιολογικών νευρώνων από τους οποίους και έχει εμπνευστεί. [13]
Ο τρόπος λειτουργίας των τεχνητών νευρώνων είναι ο υπολογισμός του σταθμισμένου αθροίσματος των n εισόδων και έπειτα εάν αυτό είναι μεγαλύτερο ενός ορίου θ δίνεται ως έξοδος 1. Αλλιώς η έξοδος είναι 0. [14]
Η λειτουργία αυτή για τον νευρώνα j περιγράφεται από την εξίσωση:

$$y_j = \varphi\left(\sum_{i=1}^{n} w_{ij} x_i\right)$$

όπου:

$\mathbf{y_j}$	: η έξοδος του νευρώνα j
$\boldsymbol{\varphi}$	: η συνάρτηση ενεργοποίησης
$\mathbf{w_{ij}}$	: το βάρος στάθμισης του νευρώνα j που αντιστοιχεί στην είσοδο i
$\mathbf{x_i}$	: η είσοδος i



Σχήμα 2.1: Σχηματική αναπαράσταση της λειτουργίας ενός Νευρώνα

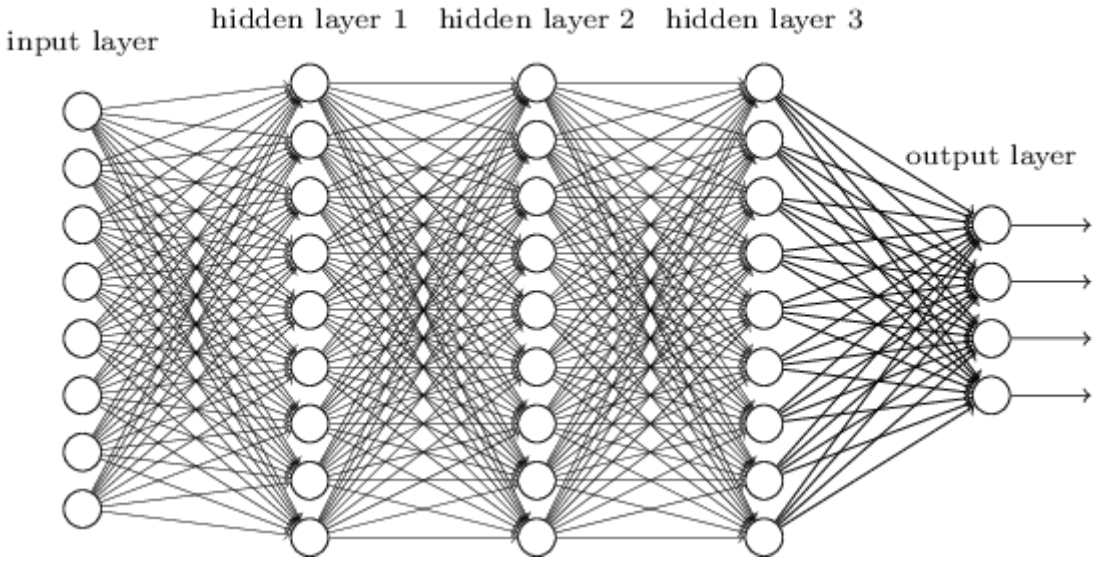
2.1.2 Αρχιτεκτονική Νευρωνικών Δικτύων

Τα νευρωνικά δίκτυα είναι υπολογιστικά μοντέλα εμπνευσμένα από τα βιολογικά νευρωνικά δίκτυα του εγκεφάλου. Βασική δομική μονάδα τους είναι οι τεχνητοί νευρώνες, σε αντιστοιχία με τους βιολογικούς νευρώνες. Αυτοί δομούνται σε επίπεδα, κάθε ένα από τα οποία εφαρμόζει διαφορετικό μετασχηματισμό στην είσοδο την οποία λαμβάνει από το προηγούμενο επίπεδο, παράγοντας έξοδο την οποία μεταδίδει στο επόμενο επίπεδο. Δηλαδή είναι “feed forward”, τα δεδομένα κινούνται μόνο προς μία κατεύθυνση. [15]

Τα επίπεδα των νευρωνικών δικτύων χωρίζονται στις εξής κατηγορίες:

1. **Input Layer (Επίπεδο Εισόδου):** Το πρώτο επίπεδο του δικτύου, το οποίο δέχεται τα δεδομένα εισόδου
2. **Hidden Layers (Κρυφά Επίπεδα):** Τα ενδιάμεσα επίπεδα, τα οποία μπορεί να μην υπάρχουν
3. **Output Layer (Επίπεδο Εξόδου):** Το τελευταίο επίπεδο, το οποίο παράγει το τελικό αποτέλεσμα

Τα πολυεπίπεδα δίκτυα μπορεί να είναι πλήρως συνδεδεμένα, οπότε κάθε νευρώνας ενός επιπέδου συνδέεται με κάθε νευρώνα του επόμενου, ή μερικώς συνδεδεμένα.
Τα βαθιά νευρωνικά δίκτυα είναι πολυεπίπεδα. Αυτό σημαίνει ότι μεταξύ των επιπέδων εισόδου και εξόδου έχουν πολλά κρυφά επίπεδα, τυπικά πρέπει να έχουν τουλάχιστον 2 κρυφά επίπεδα.



Σχήμα 2.2: Βαθύ νευρωνικό δίκτυο με 3 κρυφά επίπεδα

2.2 Recurrent Neural Networks (RNN)

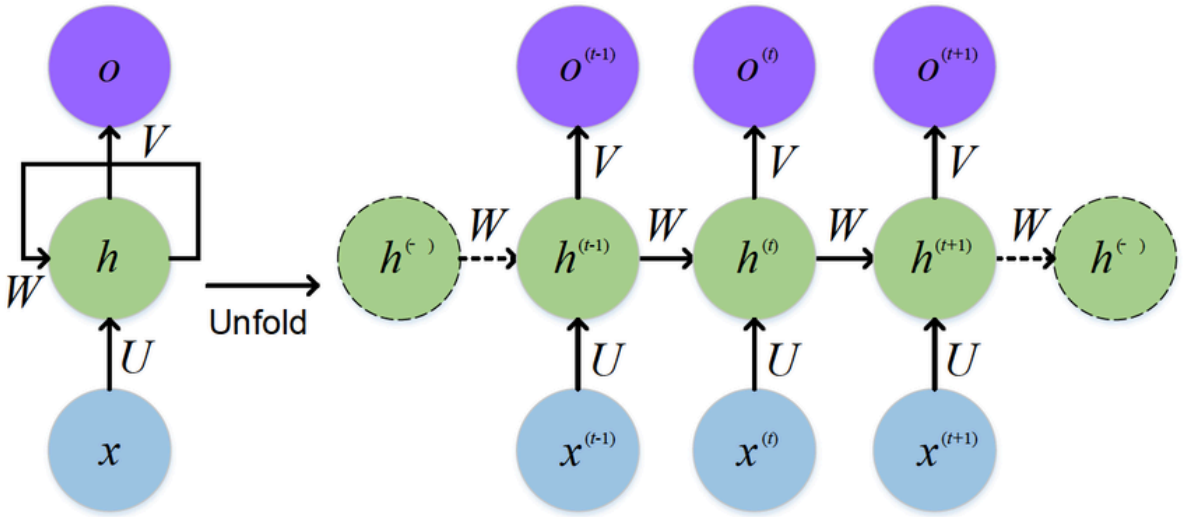
Τα αναδρομικά νευρωνικά δίκτυα (RNNs) είναι σχεδιασμένα για να επεξεργάζονται σειριακά σύνολα δεδομένων, στα οποία η σειρά των δεδομένων έχει σημασία. Αυτό το επιτυγχάνουν με τη δημιουργία κύκλων στο γράφο του δικτύου. Δηλαδή, σε αντίθεση με άλλα νευρωνικά δίκτυα στα οποία η ροή πληροφορίας είναι προς μία μόνο κατεύθυνση, στα RNNs πληροφορία μεταδίδεται και προς τα πίσω σε κάθε βήμα. Έτσι μπορούν να "θυμούνται" πληροφορία από προηγούμενα βήματα και να την χρησιμοποιούν για την εξαγωγή των μελλοντικών αποτελεσμάτων. [16], [17]
Βασικό μειονέκτημα των RNNs είναι η αδυναμία εκμάθησης μακροχρόνιων εξαρτήσεων [17] . Τέτοια προβλήματα επιλύονται με πιο σύγχρονες προσεγγίσεις όπως τα δίκτυα μακράς βραχύχρονης μνήμης (LSTM) [16] ή οι μετασχηματιστές κωδικοποιητή-αποκωδικοποιητή (encoder-decoder transformers) οι οποίοι βασίζονται σε μηχανισμό προσοχής (self-attention) αντί για επανάληψης (recurrence).

2.2.1 Αρχιτεκτονική των RNN

Ένα RNN χαρακτηρίζεται από 3 επίπεδα, την είσοδο, την μονάδα επανάληψης και την έξοδο. Η μονάδα επανάληψης είναι αυτή η οποία διατηρεί μία κρυφή κατάσταση στην οποία διατηρείται πληροφορία από προηγούμενα βήματα. Η μονάδα επανάληψης επανατροφοδοτεί την κρυφή της κατάσταση και έτσι μπορεί να θυμάται πληροφορία από προηγούμενα βήματα και να διακρίνει εξαρτήσεις στην πάροδο του χρόνου.

Βασικό βήμα για την κατανόηση της λειτουργίας των RNN είναι το ξεδίπλωμα (unfolding). Στη διαδικασία αυτή ξεδιπλώνεται η επαναλαμβανόμενη δομή στην πάροδο των χρονικών

βημάτων. Έτσι, όπως φαίνεται και στην παρακάτω εικόνα, κάθε χρονικό βήμα εμφανίζεται ως ένα ξεχωριστό "επίπεδο".



Σχήμα 2.3: RNN συμπιεσμένο (αριστερά) και ξεδιπλωμένο στο χρόνο (δεξιά)

Hidden State Calculation (Υπολογισμός Κρυφής Κατάστασης)
Στα RNN η κρυφή κατάσταση υπολογίζεται για κάθε είσοδο ώστε να διατηρούνται οι χρονικές εξαρτήσεις. Δηλαδή η κρυφή κατάσταση καθορίζει τη "μνήμη" του συστήματος και υπολογίζεται από τον τύπο:

$$h_t = f_h(Ux_t + Wh_{t-1} + b_h)$$

όπου:

$\mathbf{h_t}$	: η τρέχουσα κρυφή κατάσταση
$\mathbf{h_{t-1}}$	: η προηγούμενη κρυφή κατάσταση
$\mathbf{x_t}$	: η είσοδος στο τρέχον βήμα
$\mathbf{f_h}$	: η συνάρτηση ενεργοποίησης του κρυφού επιπέδου
U	: το βάρος που αντιστοιχεί στην είσοδο
W	: το βάρος που αντιστοιχεί στην προηγούμενη κρυφή κατάσταση
$\mathbf{b_h}$	: το διάνυσμα μεροληψίας (bias)

Output Calculation (Υπολογισμός της Εξόδου)
Η έξοδος υπολογίζεται από την εξίσωση:

$$y_t = f_O(Vh_t + b_O)$$

όπου:

$\mathbf{y_t}$	: η τρέχουσα έξοδος
$\mathbf{h_t}$	: η τρέχουσα κρυφή κατάσταση

$\mathbf{f_O}$: η συνάρτηση ενεργοποίησης της εξόδου
$\mathbf{V}$: το βάρος που αντιστοιχεί στην έξοδο
$\mathbf{b_O}$: το διάνυσμα μεροληψίας (bias)

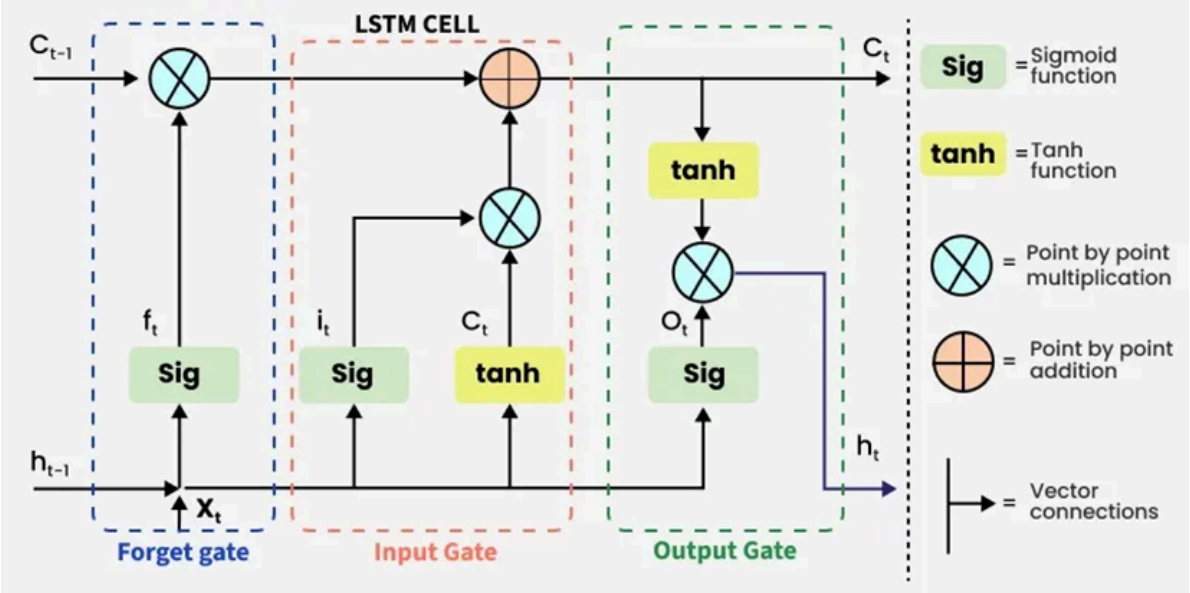
Η συμπεριφορά του RNN περιγράφεται από μη-γραμμικές εξισώσεις οι οποίες επαναλαμβάνονται στον χρόνο. Η κρυφή κατάσταση περιέχει πληροφορία για όλες τις προηγούμενες καταστάσεις στην πάροδο του χρόνου και σε κάθε βήμα παράγει για το επίπεδο εξόδου μία πρόβλεψη η οποία βασίζεται στην τρέχουσα είσοδο.

2.3 Long Short-Term Memory (LSTM)

Η κεντρική ιδέα στην οποία στηρίζονται τα δίκτυα μακράς βραχύχρονης μνήμης (LSTM) είναι η διατήρηση της κατάστασης για μεγαλύτερα χρονικά διαστήματα από ότι τα προηγούμενα RNNs [16] . Συγκεκριμένα, το πρόβλημα το οποίο καλούνται να λύσουν είναι η σταδιακή "εξαφάνιση" των παραγώγων που χρησιμοποιούνται, καθώς αυτές σταδιακά τείνουν προς το μηδέν.
Το πρόβλημα αυτό, επιλύεται στα LSTM με τη χρήση κελιών μνήμης (memory cells) και πυλών, τα οποία ρυθμίζουν τη ροή πληροφορίας. [18]

2.3.1 Αρχιτεκτονική των LSTM

Το LSTM αποτελείται από τρεις πύλες. Το κελί (cell) του LSTM παράγει την έξοδο (output) και διατηρεί πληροφορία, που αποκαλείται cell state (κατάσταση του κελιού), ενώ οι πύλες είναι υπεύθυνες για τη διαχείριση της μνήμης, δηλαδή ρυθμίζουν τη ροή πληροφοριών. Συγκεκριμένα, αυτές οι τρεις πύλες είναι η πύλη εισόδου (input gate), η πύλη εξόδου (output gate) και η πύλη λήθης (forget gate). [18], [19]



Σχήμα 2.4: Εικόνα Δομής ενός LSTM

Forget Gate (Πύλη Λήθης)

Η πύλη λήθης αποφασίζει ποια πληροφορία από το προηγούμενο κελί δεν είναι πλέον χρήσιμη και άρα πρέπει να "ξεχαστεί". Η πύλη λαμβάνει την τρέχουσα είσοδο και την προηγούμενη κατάσταση και παράγει ένα διάνυσμα τιμών μεταξύ 0 και 1. Συγκεκριμένα, εάν μια (στρογγυλοποιημένη) τιμή του διανύσματος είναι 1 τότε έχουμε διατήρηση πληροφορίας, ενώ εάν η τιμή είναι 0 η αντίστοιχη πληροφορία "ξεχνιέται". Η συμπεριφορά της πύλης λήθης περιγράφεται από την εξίσωση:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Όπου:

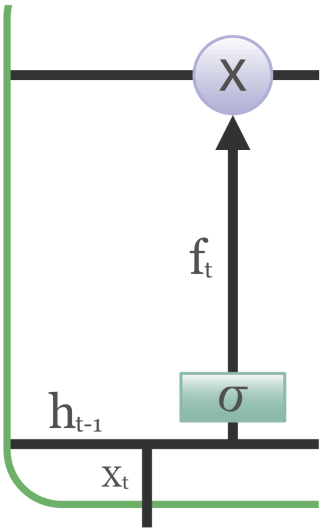
f_t : το διάνυσμα τιμών της πύλης λήθης ("αποφασίζει" ποιες τιμές διατηρούνται και ποιες θα ξεχαστούν)

σ : η σιγμοειδής συνάρτηση ενεργοποίησης

W_f : τα βάρη που αντιστοιχούν στην πύλη

$[h_{t-1}, x_t]$: η συνένωση του διανύσματος εξόδου της προηγούμενης κατάστασης h_{t-1} και του διανύσματος εισόδου της τρέχουσας κατάστασης x_t

b_f : το διάνυσμα μεροληψίας (bias) της πύλης

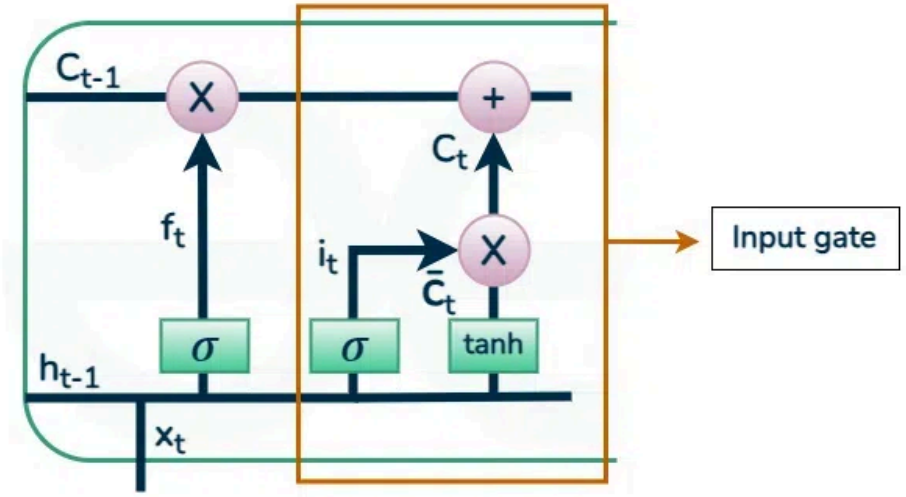


Σχήμα 2.5: Εικόνα Πύλης Λήθης LSTM

Input Gate (Πύλη Εισόδου)
Η πύλη εισόδου αποφασίζει ποιες από τις καινούργιες πληροφορίες θα αποθηκευτούν στην τρέχουσα κατάσταση. Λαμβάνει την προηγούμενη (κρυφή) κατάσταση και την τρέχουσα είσοδο και παράγει ένα διάνυσμα τιμών μεταξύ 0 και 1, το οποίο αντιπροσωπεύει πόση από την πληροφορία της τρέχουσας εισόδου θα διατηρηθεί. Οι εξισώσεις που περιγράφουν την συμπεριφορά της πύλης είναι:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$
$$\hat{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

Όπου:
i_t : το διάνυσμα τιμών της πύλης εισόδου (“αποφασίζει” ποιες υποψήφιες τιμές διατηρούνται και ποιες όχι)
$\hat{C}_t$: το διάνυσμα με τις υποψήφιες τιμές πληροφορίας που μπορούν να προστεθούν στην τρέχουσα κατάσταση
σ : η σιγμοειδής συνάρτηση ενεργοποίησης
W_i : τα βάρη για την πύλη εισόδου
$[h_{t-1}, x_t]$: η συνένωση του διανύσματος εξόδου της προηγούμενης κατάστασης h_{t-1} και του διανύσματος εισόδου της τρέχουσας κατάστασης x_t
b_i , b_c : τα διανύσματα μεροληψίας (bias) της πύλης
η συνάρτηση **tanh** αντιστοιχίζει τις τιμές στο διάστημα -1, +1



Σχήμα 2.6: Εικόνα Πύλης Εισόδου LSTM

Ενημέρωση Cell State (Κατάστασης Κελιού)
Με τα δεδομένα που προκύπτουν από την πύλη εισόδου ενημερώνεται η κατάσταση του κελιού μνήμης. Συγκεκριμένα, η τρέχουσα κατάσταση προκύπτει από τα δεδομένα της προηγούμενης κατάστασης που διατηρούνται, όπως αυτά επιλέγονται από την πύλη λήθης και από τα υποψήφια δεδομένα της τρέχουσας κατάστασης που τελικά θα διατηρηθούν. Η διαδικασία αυτή περιγράφεται από την εξίσωση:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \hat{C}_t$$

Όπου:
$\mathbf{C_t}$: η τρέχουσα (ενημερωμένη) κατάσταση του κελιού μνήμης
$\mathbf{C_{t-1}}$: η προηγούμενη κατάσταση του κελιού μνήμης
$\mathbf{\hat{C}_t}$: το διάνυσμα με τις υποψήφιες τιμές πληροφορίας που μπορούν να προστεθούν στην τρέχουσα κατάσταση
$\mathbf{f_t}$: το διάνυσμα τιμών της πύλης λήθης, "αποφασίζει" ποιες τιμές της προηγούμενης κατάστασης του κελιού μνήμης θα διατηρηθούν
$\mathbf{i_t}$: το διάνυσμα τιμών της πύλης εισόδου , "αποφασίζει" ποιες υποψήφιες τιμές θα προστεθούν τελικά στην τρέχουσα κατάσταση

Output Gate (Πύλη Εξόδου)
Η πύλη εξόδου ελέγχει ποιά πληροφορία από την τρέχουσα κατάσταση μνήμης θα χρησιμοποιηθεί για να παραχθεί η έξοδος. Για να παρθεί αυτή η απόφαση χρησιμοποιούνται η προηγούμενη κατάσταση και η τρέχουσα είσοδος. Η συμπεριφορά της πύλης εξόδου δίνεται από τις εξισώσεις:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \cdot tanh(C_t)$$

Όπου:

o_t : το διάνυσμα τιμών της πύλης εξόδου ("αποφασίζει" ποιες τιμές θα χρησιμοποιηθούν για την ενημέρωση της κρυφής κατάστασης στο τρέχον χρονικό βήμα)

h_t : η τρέχουσα κρυφή κατάσταση (είναι επίσης το διάνυσμα εξόδου της μονάδας LSTM)

σ : η σιγμοειδής συνάρτηση ενεργοποίησης

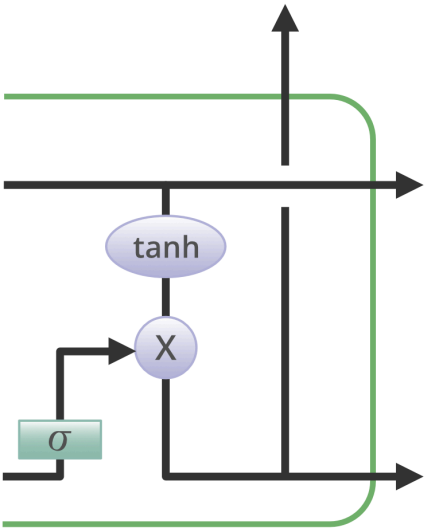
W_o : το διάνυσμα βαρών της πύλης εξόδου

$[h_{t-1}, x_t]$: η συνένωση του διανύσματος εξόδου της προηγούμενης κατάστασης h_{t-1} και του διανύσματος εισόδου της τρέχουσας κατάστασης x_t

b_o : το διάνυσμα μεροληψίας (bias) της πύλης

C_t : η τρέχουσα (ενημερωμένη) κατάσταση του κελιού μνήμης

η συνάρτηση **tanh** αντιστοιχίζει τις τιμές στο διάστημα -1, +1

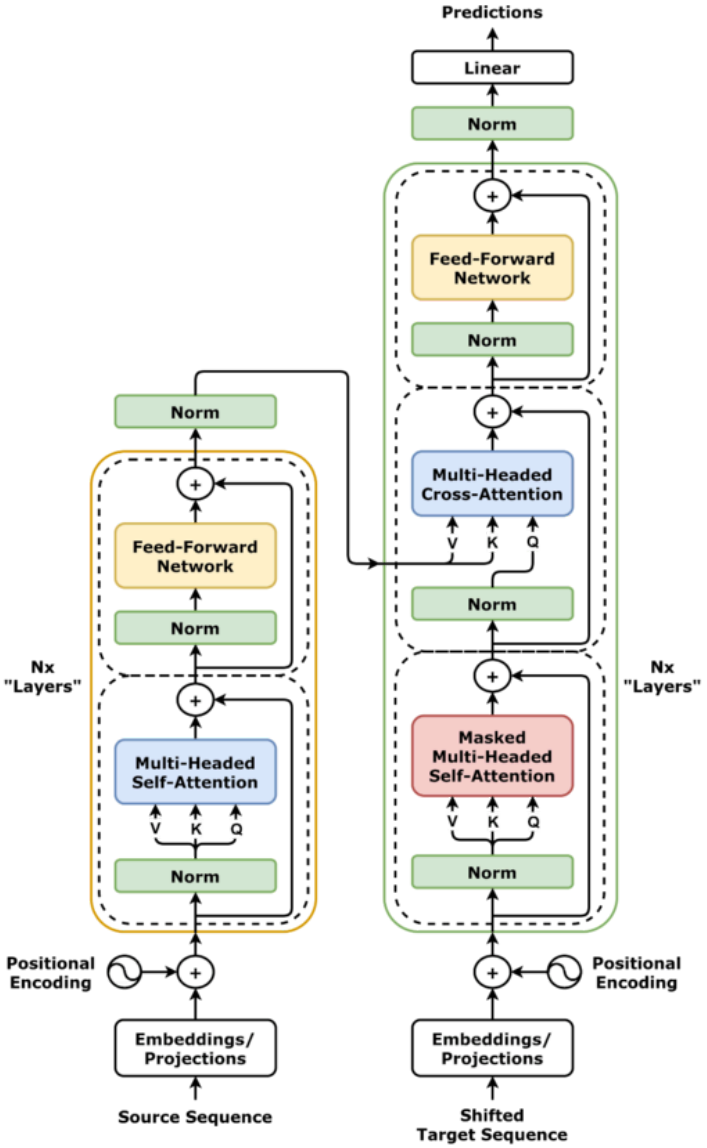


Σχήμα 2.7: Εικόνα Πύλης Εξόδου LSTM

2.4 Transformers

Οι μετασχηματιστές είναι μοντέλα βαθιάς μάθησης που αναπτύχθηκαν αρχικά ως μοντέλα επεξεργασίας φυσικής γλώσσας. Η βασική ιδέα στην οποία στηρίζονται είναι η αρχιτεκτονική κωδικοποιητή-αποκωδικοποιητή (encoder-decoder). Σε αντίθεση με προηγούμενα μοντέλα όπως τα αναδρομικά νευρωνικά δίκτυα (RNNs) δεν υποφέρουν από το πρόβλημα βραχύχρονης μνήμης (απώλεια μνήμης για μεγαλύτερα χρονικά διαστήματα). Από την άλλη, τα LSTM που έχουν επιλύσει αυτό το πρόβλημα, επεξεργάζονται τα δεδομένα σειριακά (βήμα-βήμα). Σε αντίθεση, ένα σημαντικό πλεονέκτημα των

μετασχηματιστών είναι η δυνατότητα παραλληλοποίησης, η οποία είναι εφικτή μέσω του μηχανισμού προσοχής. [20]
Συμπερασματικά, οι μετασχηματιστές αποτελούν βελτίωση των προγενέστερων καλύτερων μοντέλων. Αυτά ήταν σύνθετα RNNs τα οποία περιλάμβαναν έναν κωδικοποιητή (encoder) και έναν αποκωδικοποιητή (decoder) συνδεδεμένους με έναν μηχανισμό προσοχής (attention mechanism). [21]



Σχήμα 2.8: Εικόνα Δομής ενός Transformer

2.4.1 Αρχιτεκτονική των Transformer

Οι βασικές δομές του μετασχηματιστή είναι ο encoder (κωδικοποιητής) και ο decoder (αποκωδικοποιητής).

Encoder (Κωδικοποιητής)
Ο κωδικοποιητής αντιστοιχίζει την ακολουθία εισόδου που αποτελείται από διακριτά τμήματα σε συνεχές διάνυσμα. Αποτελείται από μία στοίβα πολλών (συνήθως 6) ίδιων επιπέδων. Κάθε επίπεδο έχει δύο βασικά τμήματα:

1. **Multi-head self-attention mechanism**: Βοηθάει στην κατανόηση των συσχετίσεων μεταξύ των διαφορετικών κομματιών της εισόδου
2. **Position-wise fully connected feed-forward network:** Μετασχηματίζει περαιτέρω την αναπαράσταση της εισόδου

Decoder (Αποκωδικοποιητής)
Ο αποκωδικοποιητής λαμβάνει το διάνυσμα από τον κωδικοποιητή και παράγει μια ακολουθία εξόδου ένα σύμβολο τη φορά. Αποτελείται επίσης από μία στοίβα (ίδιου αριθμού με τον κωδικοποιητή, συνήθως 6) ίδιων επιπέδων. Κάθε επίπεδο αποτελείται, όπως και στον κωδικοποιητή, από **Multi-head self-attention mechanism** και **Position-wise fully connected feed-forward network**. Όμως, επιπλέον αυτών υπάρχει επίσης και ένα τρίτο τμήμα το οποίο εκτελεί προσοχή πολλών σημείων (**multi-head attention**) στην έξοδο του κωδικοποιητή.

Όσον αφορά τα διάφορα τμήματα τα οποία χρησιμοποιούνται από τις παραπάνω δομές έχουμε:

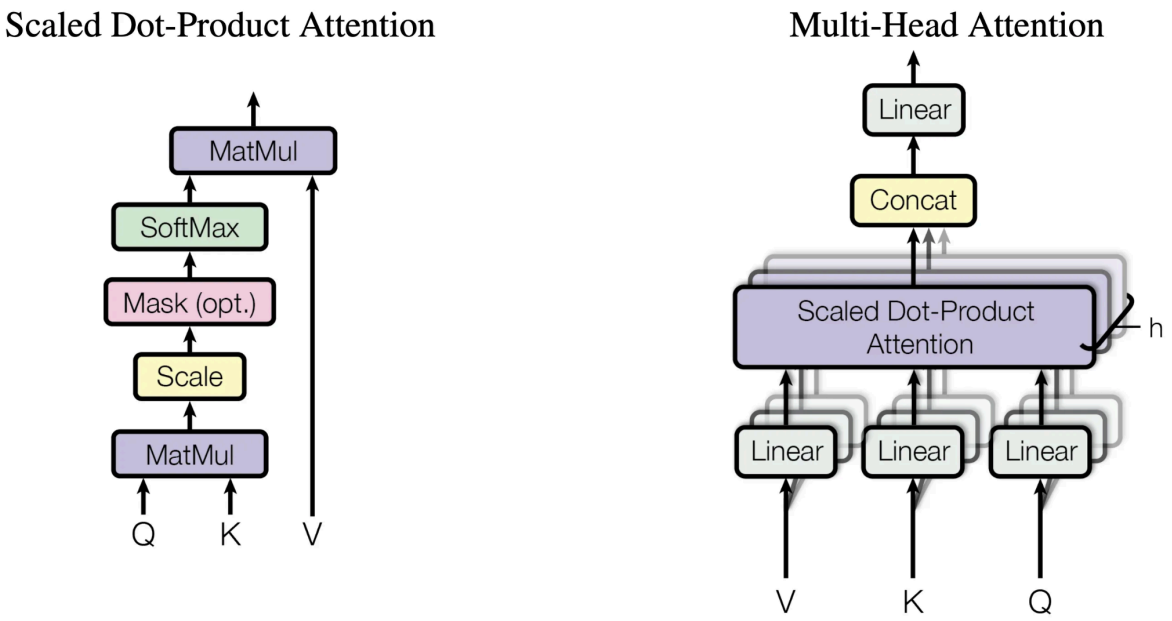
Multi-Head Self-Attention (Προσοχή Πολλών Κεφαλών)
Ο μηχανισμός αυτός επιτρέπει στο μοντέλο να καθορίσει ποια τμήματα της εισόδου έχουν μεγαλύτερη σχέση μεταξύ τους. Γενικά, ένας μηχανισμός προσοχής αντιστοιχίζει ένα ερώτημα (Q, query) και ένα σύνολο ζευγών κλειδιών-τιμής (K, key - V, value) σε μία έξοδο. Πιο συγκεκριμένα, όταν πρόκειται για παραλλαγή του μηχανισμού η οποία είναι πολλών "κεφαλών", αυτό σημαίνει ότι πολλά στρώματα προσοχής τρέχουν παράλληλα.

Στην περίπτωση των Μετασχηματιστών ο μηχανισμός προσοχής είναι συγκεκριμένα μηχανισμός κλιμακούμενης προσοχής με εσωτερικό γινόμενο (scaled dot-product attention) και υπολογίζεται από:

$$Attention(Q,\ K,\ V)\ =\ softmax(\ \frac{QK^T}{\sqrt{d_K}}\ )V$$

όπου:

Q	: το διάνυσμα query (ερωτήματος)
K	: το διάνυσμα key (κλειδιού)
V	: το διάνυσμα value (τιμής)
$\mathbf{d_K}$	: η διάσταση των διανυσμάτων query και key



Σχήμα 2.9: Μηχανισμός κλιμακούμενης προσοχής με εσωτερικό γινόμενο (αριστερά) και Μηχανισμός πολλών κεφαλών προσοχής (δεξιά)

Position-wise Feed-Forward Networks

Κάθε επίπεδο κωδικοποιητή και αποκωδικοποιητή περιέχει επίσης ένα πλήρως συνδεδεμένο δίκτυο εμπρόσθιας τροφοδότησης (feed-forward network) το οποίο εφαρμόζεται σε κάθε θέση ξεχωριστά και με πανομοιότυπο τρόπο. Αποτελείται από δύο γραμμικούς μετασχηματισμούς και μία συνάρτηση ενεργοποίησης ReLU ανάμεσά τους. Η διάσταση της εισόδου και της εξόδου είναι $d_{model} = 512$.

Περιγράφεται από την εξίσωση:

$$FFN(x) = max(0, xW_1 + b_1)W_2 + b_2$$

Positional Encoding (Κωδικοποίηση Θέσης)

Οι Μετασχηματιστές δεν διαθέτουν ανάδραση ή συνέλιξη και άρα για να αξιοποιήσουν τη σειρά της ακολουθίας χρειάζεται ένας διαφορετικός τρόπος για την εισαγωγή πληροφορίας που να αφορά τη σχετική θέση κάθε τμήματος στην ακολουθία. Για το σκοπό αυτό χρησιμοποιούνται οι κωδικοποιήσεις θέσεις (positional encodings) που έχουν επίσης διάσταση d_{model} .

Υπολογίζονται ως εξής:

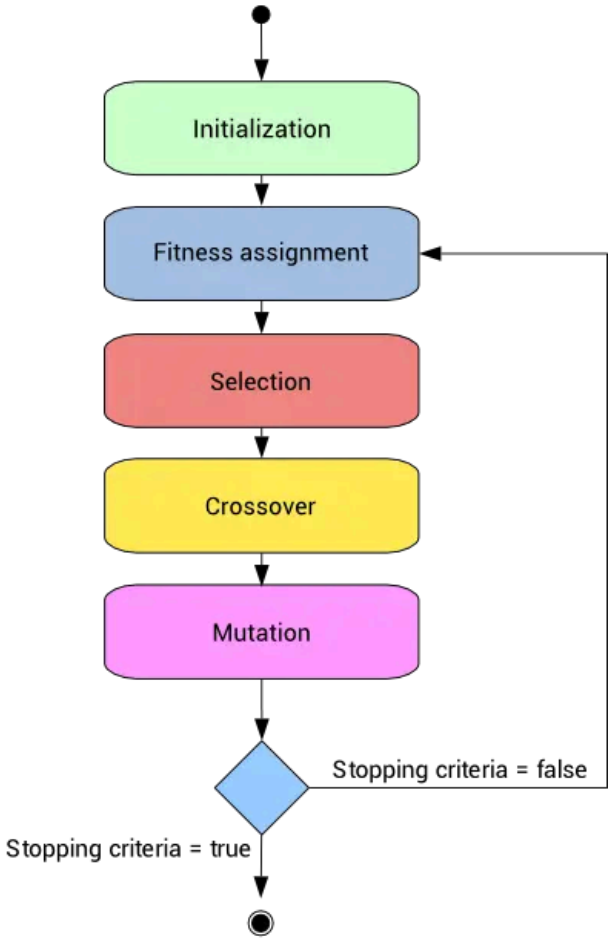
$$PE_{(pos,\,2i)} = sin(pos/10000^{2i/d_{model}})$$

$$PE_{(pos,\,2i+1)} = cos(pos/10000^{2i/d_{model}})$$

όπου:
pos : η θέση
i : η διάσταση

2.5 Γενετικός Αλγόριθμος

Ο γενετικός αλγόριθμος (genetic algorithm) είναι μία μέθοδος αναζήτησης και βελτιστοποίησης η οποία στηρίζεται στις ιδέες της φυσικής επιλογής και της γενετικής. Αντιστοιχίζει το χώρο επίλυσης του προβλήματος σε ένα σύνολο ατόμων. Σε αναλογία με τις βιολογικές διαδικασίες, ο πληθυσμός εξελίσσεται στην πάροδο γενεών παράγοντας νέα άτομα, κάθε ένα από τα οποία αντιστοιχεί σε μία πιθανή λύση του προβλήματος. Η επιτυχία ή η αποτυχία του κάθε ατόμου καθορίζεται από την προσαρμογή του (fitness) στις απαιτήσεις του προβλήματος. [22], [23]



Σχήμα 2.10: Μεθοδολογία του Γενετικού Αλγορίθμου

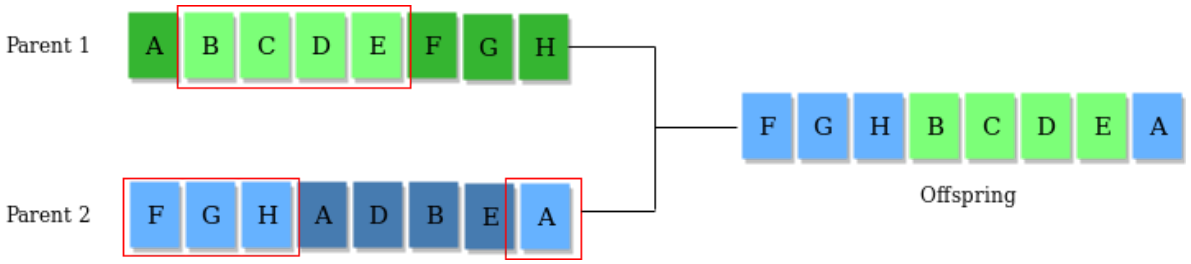
2.5.1 Βασικές Έννοιες Γενετικού Αλγορίθμου

Στον γενετικό αλγόριθμο κάθε πιθανή λύση είναι ένα άτομο εντός του χώρου λύσης του προβλήματος. Κάθε άτομο αντιπροσωπεύεται από ένα διάνυσμα τιμών (χρωμόσωμα), το οποίο απαρτίζεται από ένα σύνολο μεταβλητών χαρακτηριστικών, τα γονίδια.
Για κάθε άτομο υπολογίζεται ένας δείκτης προσαρμογής (fitness score) με τα άτομα τα οποία έχουν καλύτερη προσαρμογή να είναι επιθυμητά. Στην πάροδο των γενεών, τα άτομα με καλύτερη προσαρμογή είναι πιο πιθανό να χρησιμοποιηθούν στην παραγωγή απογόνων. Από την άλλη, επειδή το μέγεθος κάθε γενιάς είναι σταθερό, τα άτομα παλαιότερων γενεών με μικρή προσαρμοστικότητα απορρίπτονται ("πεθαίνουν"). Με αυτό τον τρόπο, κάθε διαδοχική γενιά έχει ως σύνολο, καλύτερα γονίδια από τις προηγούμενες.

Οι βασικές λειτουργίες οι οποίες εφαρμόζονται για την παραγωγή των ατόμων κάθε γενιάς είναι:

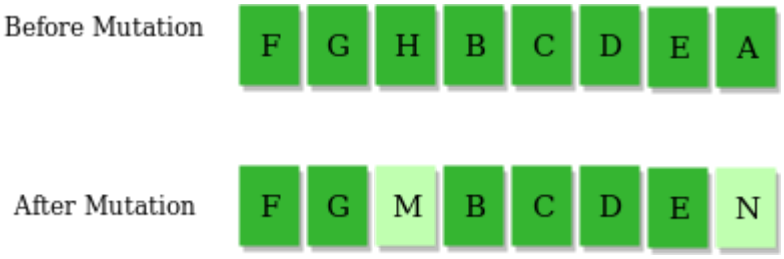
1. **Selection (Επιλογή) :** Δίνεται προτίμηση στα άτομα με καλύτερο δείκτη προσαρμογής (fitness score), κάνοντας πιο πιθανό το ενδεχόμενο να μεταδώσουν τα "καλά" τους γονίδια στις επόμενες γενιές

2. **Crossover (Διασταύρωση) :** Πραγματοποιείται σε αντιστοιχία με το ζευγάρωμα των βιολογικών ατόμων. Μεταξύ δύο ατόμων επιλέγονται τυχαία δύο σημεία διασταύρωσης μεταξύ των οποίων γίνεται αμοιβαία ανταλλαγή γονιδίων, παράγοντας έναν απόγονο



Σχήμα 2.11: Διασταύρωση Γονιδίων Γονέων για παραγωγή Απογόνου

3. **Mutation (Μετάλλαξη) :** Τυχαία γονίδια εισάγονται στους απογόνους για να διατηρηθεί η "ποικιλομορφία" του πληθυσμού και να αποφευχθούν φαινόμενα πρόωρης σύγκλισης και υπερ-προσαρμογής



Σχήμα 2.12: Μετάλλαξη Γονιδίων Απογόνου

3 Πειραματικό Μέρος

3.1 Εργαλεία

Για την υλοποίηση της πειραματικής διαδικασίας έγινε χρήση της γλώσσας προγραμματισμού python στο περιβάλλον google colaboratory. Η συγκεκριμένη γλώσσα προγραμματισμού επιλέχθηκε επειδή διαθέτει μεγάλο πλήθος βιβλιοθηκών που μπορούν να χρησιμοποιηθούν για την επεξεργασία μεγάλων συνόλων δεδομένων, την εκπαίδευση μοντέλων τεχνητής και την οπτικοποίηση των αποτελεσμάτων. Συγκεκριμένα, χρησιμοποιήθηκε η Pandas για την επεξεργασία των δεδομένων, η TensorFlow και το Keras για την εκπαίδευση των μοντέλων, ενώ έγινε και χρήση βιβλιοθηκών για πιο γενικούς σκοπούς όπως για παράδειγμα οι csv, matplotlib, numpy.
Συγκεκριμένα:

TensorFlow: Είναι μια από τις πιο διαδεδομένες βιβλιοθήκες, χρησιμοποιείται κυρίως για σκοπούς μηχανικής μάθησης και τεχνητής νοημοσύνης. Παρέχει τη δυνατότητα για ευελιξία στη διαχείριση δεδομένων και για χρήση σε διάφορες υπολογιστικές συσκευές όπως πολυπήρυνες CPU, GPU και TPU. Είναι ιδιαίτερα δημοφιλής για τον ορισμό, την κατασκευή και την εκπαίδευση νευρωνικών δικτύων. [24]

Keras: Είναι ένα υψηλού επιπέδου API της βιβλιοθήκης TensorFlow. Είναι ένα εύχρηστο και ευέλικτο εργαλείο το οποίο χρησιμοποιείται στην ανάπτυξη και κατασκευή νευρωνικών δικτύων. Περιέχει πολλά από τα δομικά στοιχεία που χρειάζονται συχνά στην διαδικασία αυτή, όπως επίπεδα και συναρτήσεις ενεργοποίησης. [25]

Pandas: Είναι μια βιβλιοθήκη η οποία χρησιμοποιείται για τη διαχείριση και ανάλυση δεδομένων. Συγκεκριμένα, προσφέρει γρήγορες, ευέλικτες και περιεκτικές δομές δεδομένων διευκολύνοντας την επεξεργασία σχεσιακών συνόλων δεδομένων. [26]

NumPy: Είναι η βασική βιβλιοθήκη της γλώσσας Python, η οποία χρησιμοποιείται για επιστημονικούς υπολογισμούς σε πολυδιάστατα διανύσματα και πίνακες. [27]

Matplotlib: Είναι βιβλιοθήκη η οποία χρησιμοποιείται για τη δημιουργία γραφημάτων. Η ιδέα πίσω από την συγκεκριμένη βιβλιοθήκη είναι η χρήση ενός εύκολου και γρήγορου τρόπου για την κατασκευή ευανάγνωστων γραφημάτων. [28]

3.2 Δεδομένα

Τα δεδομένα που χρησιμοποιήθηκαν για την εργασία αποτελούν μετρήσεις από αισθητήρες σε 28 διαφορετικά σπίτια. Συγκεκριμένα, τα στοιχεία αντιστοιχούν σε 19 κατηγορίες οι τιμές των οποίων λαμβάνονται σε χρονικά διαστήματα του ενός λεπτού και για συνολική διάρκεια μερικών μηνών. Πρόκειται για μετρήσεις του επιπέδου λειτουργίας του boiler (το μέγεθος που θέλουμε να προβλέψουμε), άλλων μεγεθών σχετικών με τη λειτουργία του boiler (όπως για παράδειγμα αν είναι ενεργοποιημένη η φλόγα του boiler ή αν γίνεται χρήση νερού), μετρήσεις θερμοκρασίας (όπως θερμοκρασίες εσωτερικού και εξωτερικού χώρου) και βοηθητικά δεδομένα (όπως ημερομηνία και ώρα). Αναλυτικότερα, τα μεγέθη που χρησιμοποιούνται και οι μεταβλητές που τα αναπαριστούν παρουσιάζονται στον ακόλουθο πίνακα:

Μεταβλητή	**Τύπος**	**Όνομα Μεγέθους**	**Περιγραφή**
blr_mod_lvl	*Float*	Boiler modulation level	Το ποσοστό του μέγιστου βαθμού ρύθμισης του boiler
blr_t	*Float*	Boiler temperature	Η τρέχουσα θερμοκρασία του νερού μέσα στο boiler
heat	*Boolean (0 or 1)*	Heat indicator	Δείχνει αν η κυκλοφορία του boiler είναι ενεργή
flame	*Boolean (0 or 1)*	Flame indicator	Δείχνει αν η φλόγα του boiler είναι αναμμένη
water	*Boolean (0 or 1)*	Domestic hot water indicator	Δείχνει αν έχει ζητηθεί ζεστό νερό στο σπίτι
t_out	*Float*	Outdoor temperature	Η τρέχουσα εξωτερική θερμοκρασία (θερμοκρασία εξωτερικού χώρου)

t_ret	*Float*	Inlet temperature	Η τρέχουσα θερμοκρασία του νερού που εισέρχεται στο boiler από τα καλοριφέρ
t_r	*Float*	Indoor temperature	Η τρέχουσα θερμοκρασία εσωτερικού χώρου στο σπίτι
t_r_set	*Float*	Indoor setpoint	Η επιθυμητή θερμοκρασία εσωτερικού χώρου που το σύστημα θέρμανσης πρέπει να φτάσει
otc_cur	*Float (0 - 10)*	Heating balance	Ρύθμιση χρήστη η οποία ελέγχει το ισοζύγιο οικονομίας-άνεσης. Ορίζει πόσο επίμονα ο προσαρμοστικός αλγόριθμος θα δώσει προτεραιότητα στην γρηγορότερη θέρμανση νερού ώστε να επιτευχθεί η επιθυμητή θερμοκρασία του boiler έναντι της λειτουργίας μειωμένης κατανάλωσης. Κυμαίνεται από 0 (πιο αποτελεσματικό) έως 10 (πιο επίμονο).
t_set	*Float*	Boiler setpoint	Η επιθυμητή θερμοκρασία που το boiler πρέπει να φτάσει
otc_maxt	*Float*	Max adaptive setpoint	Δείχνει το μέγιστο (MAX) επίπεδο που το boiler θα πρέπει να φτάσει στην προσαρμοστική λειτουργία βάσει του αλγόριθμου βρόχου ανάδρασης
bypass	*Enum (0, 1, 2, 4)*	Bypass indicator	Ελέγχει την πηγή του ελεγκτή θερμοκρασίας του boiler 0: προεπιλογή (65 ℃) [legacy], 1: αντιστάθμιση καιρού με καθοριζόμενο από τον χρήστη otc_cur [προσαρμοστικό],

			2: σταθερή θερμοκρασία boiler καθοριζόμενη από το cloud [legacy], 4: αντιστάθμιση καιρού με otc_cur καθοριζόμενο από το cloud [προσαρμοστικό]
nodata	*Boolean (0 or 1)*	Nodata indicator	Εάν 1, υποδηλώνει μη έγκυρα δεδομένα ή δεδομένα που λείπουν, οι τιμές θα είναι NAN

Πίνακας 3.1: Πίνακας Ανάλυσης και Ερμηνείας Μεταβλητών των Μεγεθών

3.3 Μετρικές Σφαλμάτων

Για την αξιολόγηση των διαφόρων μοντέλων που έχουν εκπαιδευτεί και την επιλογή κάποιων ως καλύτερων γίνεται χρήση του επιλεγμένου συνόλου αξιολόγησης (test dataset). Από τα δεδομένα εισόδου κάθε μοντέλο παράγει ως έξοδο μια πρόβλεψη η οποία συγκρίνεται ως προς την "ορθότητά" της με τα αναμενόμενα δεδομένα εξόδου. Για να προκύψει μια πλήρης εικόνα σχετικά με την συμπεριφορά κάθε μοντέλου γίνεται χρήση διαφορετικών μετρικών σφαλμάτων.

3.3.1 Mean Absolute Error (MAE)

Το μέσο απόλυτο σφάλμα ορίζεται ως η μέση τιμή των απόλυτων τιμών των σφαλμάτων. Είναι σχετικά εύκολο στην ερμηνεία καθώς εκφράζεται στις ίδιες μονάδες όπως και το προβλεπόμενο μέγεθος. Δίνει την ίδια "βαρύτητα" σε όλα τα σφάλματα και άρα δεν επηρεάζεται έντονα από ακραίες τιμές. Από την άλλη, αυτό σημαίνει ότι δεν "τιμωρεί" περισσότερο τις μεγάλες τιμές σφάλματος. [29]
Υπολογίζεται από τον τύπο:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - x_i\right| = \frac{1}{n}\sum_{i=1}^{n}\left|e_i\right|$$

όπου:

$\mathbf{y_i}$	: η προβλεπόμενη τιμή
$\mathbf{x_i}$	: η πραγματική τιμή
$\mathbf{\lvert e_i \rvert = \lvert y_i - x_i \rvert}$	: το απόλυτο σφάλμα

3.3.2 Mean Squared Error (MSE)

Το μέσο τετραγωνικό σφάλμα ορίζεται ως η μέση τιμή των τετραγώνων των σφαλμάτων. Βασικό του πλεονέκτημα είναι ότι, λόγω της ύψωσης στο τετράγωνο, επηρεάζεται περισσότερο από τα μεγάλα σφάλματα. Δηλαδή, "τιμωρεί" περισσότερο ένα μεγάλο σφάλμα από ότι πολλά μικρότερα. Από την άλλη, αυτό σημαίνει ότι είναι πιο ευαίσθητο σε λάθη από ακραίες τιμές. Επιπλέον, από την ύπαρξη εκθέτη στον υπολογισμό, οι μονάδες μέτρησης της μετρικής είναι διαφορετικές από του μετρούμενου μεγέθους και άρα μπορεί να μην είναι τόσο εύκολη η κατανόηση. [29]

Υπολογίζεται από τον τύπο:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(e_i)^2$$

όπου:

$\mathbf{y_i}$	: η προβλεπόμενη τιμή
$\mathbf{x_i}$	: η πραγματική τιμή

$|\mathbf{e_i}| = |\mathbf{y_i} - \mathbf{x_i}|$: το απόλυτο σφάλμα

3.3.3 Root Mean Squared Error (RMSE)

Η ρίζα του μέσου τετραγωνικού σφάλματος ορίζεται ως η τετραγωνική ρίζα του MSE. Όπως και το MSE επηρεάζεται σημαντικά από την ύπαρξη μεγάλων σφαλμάτων και άρα τα "τιμωρεί" περισσότερο από ότι το MAE. Σε αντίθεση με το MSE έχει τις ίδιες μονάδες μέτρησης με το αντίστοιχο μέγεθος και άρα είναι πιο εύκολο στην κατανόηση. Παραμένει ευαίσθητο σε ακραίες τιμές. [29]

Υπολογίζεται από τον τύπο:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i)^2}$$

όπου:

$\mathbf{y_i}$: η προβλεπόμενη τιμή
$\mathbf{x_i}$: η πραγματική τιμή
$|\mathbf{e_i}| = |\mathbf{y_i} - \mathbf{x_i}|$: το απόλυτο σφάλμα

3.3.4 R-Squared (R^2 - Coefficient of Determination)

Ο συντελεστής προσδιορισμού αντιπροσωπεύει το ποσοστό μεταβλητότητας το οποίο μπορεί να εξηγηθεί από το μοντέλο. Πρόκειται για ένα μέγεθος χωρίς μονάδα μέτρησης, με τιμές στο διάστημα -∞ έως 1. Βασικό του πλεονέκτημα είναι ότι είναι εύκολο στην κατανόηση, με τις τιμές κοντά στο 1 να αντιπροσωπεύουν καλύτερη προσαρμογή. [30]

Υπολογίζεται από τον τύπο:

$$SS_{res} = \sum_{i=1}^{n}(y_i - x_i)^2 = \sum_{i=1}^{n}(e_i)^2$$

$$SS_{tot} = \sum_{i=1}^{n}(y_i - \bar{x})^2$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

όπου:

$\mathbf{SS_{res}}$: το άθροισμα τετραγώνων υπολοίπων
$\mathbf{SS_{tot}}$: το συνολικό άθροισμα τετραγώνων
$\mathbf{y_i}$: η προβλεπόμενη τιμή

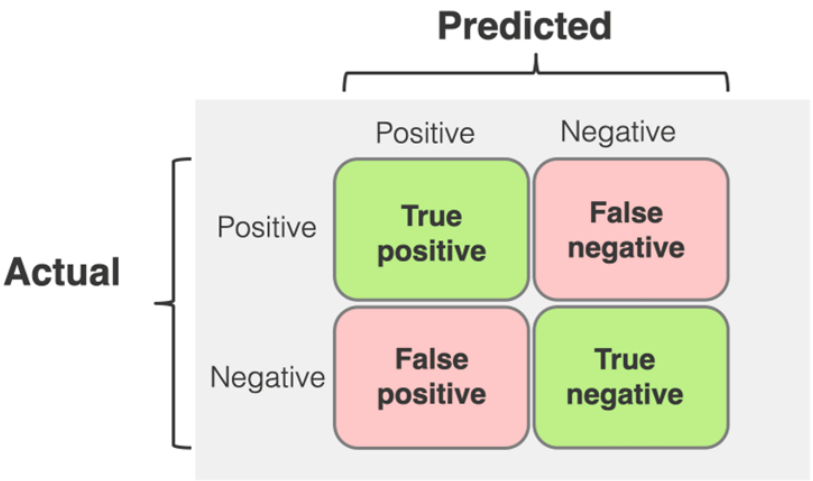
$\mathbf{x_i}$: η πραγματική τιμή
$|\mathbf{e_i}| = |\mathbf{y_i} - \mathbf{x_i}|$: το απόλυτο σφάλμα
$\bar{\mathbf{x}}$: η μέση τιμή των δεδομένων

3.4 Μετρικές Επίδοσης

Σε διάφορες διαδικασίες μηχανικής μάθησης (όπως ταξινόμηση και εύρεση αντικειμένων) για να αξιολογηθεί η ικανότητα των μοντέλων χρησιμοποιούνται οι μετρικές επίδοσης. [31] Για τον υπολογισμό αυτών χρειάζεται γνώση της προβλεπόμενης τιμής έναντι της πραγματικής για κάθε αποτέλεσμα.
Πράγματι όταν κάθε αποτέλεσμα μπορεί να πάρει δύο τιμές Αληθές (τιμή 1) ή Ψευδές (τιμή 0) και το ίδιο ισχύει και για τις τιμές της πρόβλεψης τότε προκύπτουν τέσσερα ενδεχόμενα:

1. **True Positive (Πραγματικά Αληθές)**
2. **False Positive (Ψευδώς Αληθές)**
3. **False Negative (Ψευδώς Αρνητικό)**
4. **True Negative (Πραγματικά Αρνητικό)**

Η ερμηνεία αυτών παρουσιάζεται διαγραμματικά στην ακόλουθη εικόνα:



Σχήμα 3.1: Πίνακας Σφαλμάτων (Confusion Matrix)

3.4.1 Accuracy

Εκφράζει τις σωστές προβλέψεις (θετικές και αρνητικές) στο σύνολο των προβλέψεων. Ένα ιδανικά τέλειο μοντέλο δεν θα έχει λανθασμένες προβλέψεις και άρα το accuracy παίρνει τιμή 1 (100%). Επειδή για τον υπολογισμό της αξιοποιεί όλα τα πιθανά ενδεχόμενα από τον πίνακα σφαλμάτων τείνει να είναι ένα καλό μέτρο της ποιότητας του μοντέλου σε

ισορροπημένα σύνολα δεδομένων. Όμως, όταν το σύνολο δεδομένων δεν είναι ισορροπημένο ή όταν οι λανθασμένες αξιολογήσεις του ενός ή του άλλου είδους "στοιχίζουν" περισσότερο, τότε μπορεί να οδηγήσει σε λανθασμένα συμπεράσματα.
Υπολογίζεται από την εξίσωση:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

3.4.2 Precision

Εκφράζει τις σωστές θετικές προβλέψεις στο σύνολο των θετικών προβλέψεων (σωστών και λανθασμένων). Ένα ιδανικά τέλειο μοντέλο δεν θα έχει λανθασμένα θετικά και άρα το precision παίρνει τιμή 1 (100%).
Υπολογίζεται από την εξίσωση:

$$Precision = \frac{TP}{TP + FP}$$

3.4.3 Recall (True Positive Rate)

Εκφράζει στο σύνολο των πραγματικών θετικών αποτελεσμάτων (προβλέφθηκαν ως θετικά και ως αρνητικά) πόσα προβλέφθηκαν σωστά ως θετικά. Ένα ιδανικά τέλειο μοντέλο δεν έχει λανθασμένα αρνητικά και άρα το recall παίρνει τιμή 1 (100%).
Υπολογίζεται ως εξής:

$$Recall = \frac{TP}{TP + FN}$$

3.4.4 False Positive Rate

Εκφράζει στο σύνολο των πραγματικών αρνητικών αποτελεσμάτων (προβλέφθηκαν ως θετικά και ως αρνητικά) πόσα προβλέφθηκαν λανθασμένα ως θετικά. Ένα ιδανικά τέλειο μοντέλο δεν έχει λανθασμένα θετικά και άρα το fpr παίρνει τιμή 0 (0%).
Υπολογίζεται:

$$FPR = \frac{FP}{FP + TN}$$

3.5 Πειραματικός Σχεδιασμός

Σε αυτό το κομμάτι παρουσιάζουμε και αναλύουμε τα βήματα της πειραματικής διαδικασιας που ακολουθήσαμε. Αρχικά θα δώσουμε μια συνοπτική αναπαράσταση των διαδοχικών σταδίων και σε επόμενες ενότητες θα ακολουθήσει πιο λεπτομερής περιγραφή. Πρόκειται για τα παρακάτω βήματα:

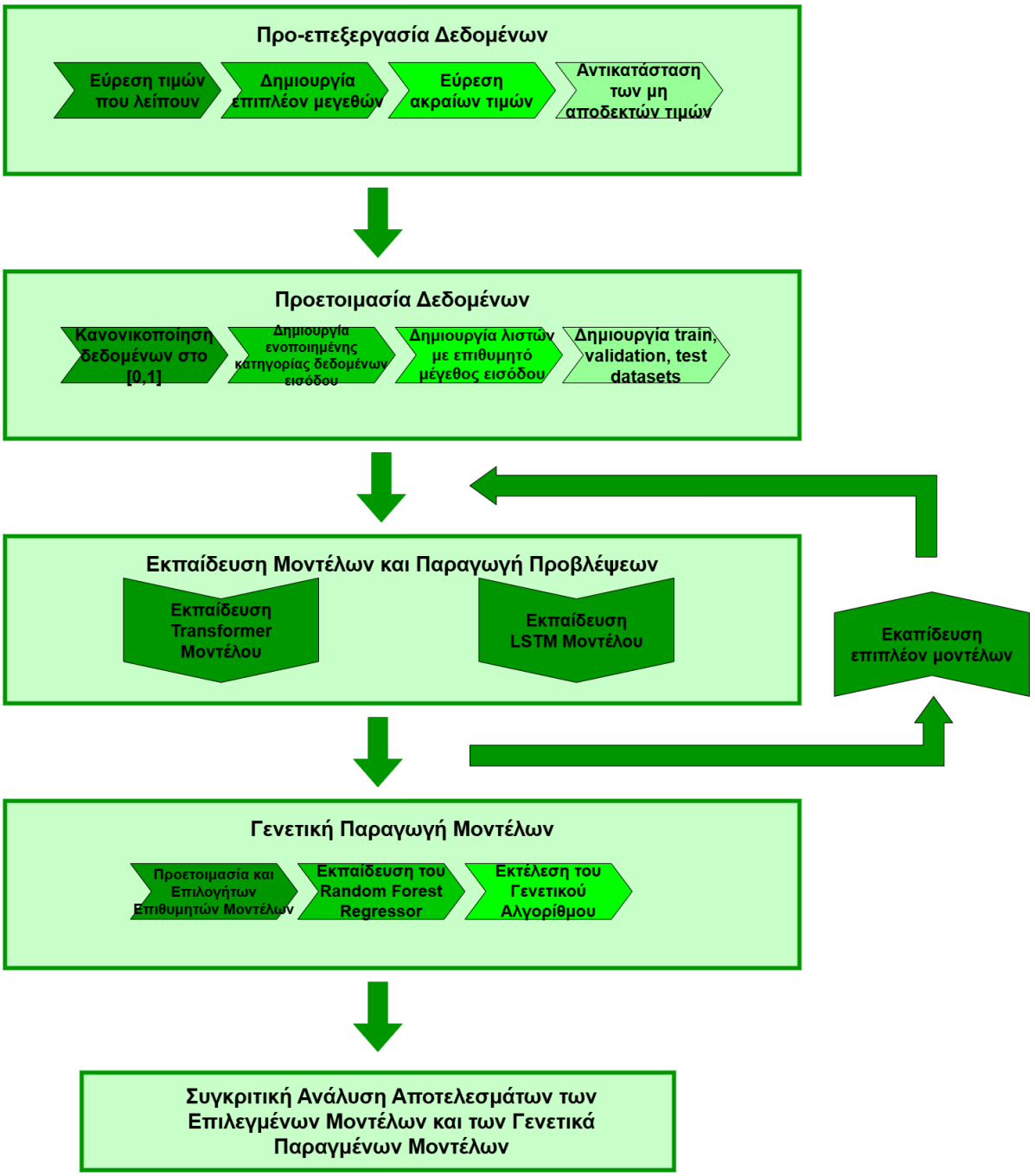
1. Την προεπεξεργασία των δεδομένων. Τα δεδομένα όπως είναι αρχικά έχουν προβλήματα, όπως είναι ελλιπείς τιμές και λάθη μετρήσεων. Αυτά πρέπει να επιλυθούν πριν μπορέσουμε να προχωρήσουμε στα επόμενα βήματα της επεξεργασίας.

2. Προετοιμασία των δεδομένων. Τα δεδομένα πρέπει να έρθουν σε μορφή κατάλληλη για την εκπαίδευση των μοντέλων και την παραγωγή προβλέψεων. Συγκεκριμένα, αυτό το βήμα εξαρτάται από χαρακτηριστικά όπως οι κατηγορίες και ο αριθμός σπιτιών που είναι μοναδικά του κάθε μοντέλου. Για αυτό θα επαναληφθεί ξεχωριστά (και με διαφορετικά χαρακτηριστικά) για κάθε μοντέλο.

3. Εκπαίδευση των μοντέλων. Για κάθε συνδυασμό χαρακτηριστικών (κατηγορίες, αριθμός σπιτιών) εκπαιδεύουμε ένα ξεχωριστό μοντέλο Transformer και ένα αντίστοιχο LSTM. Τα μοντέλα αυτά αποτελούνται από εκείνους του συνδυασμούς που θεωρούμε πιο πιθανό να αποδώσουν καλύτερα. Το βήμα αυτό υλοποιείται ξεχωριστά για μοντέλα βελτιστοποιημένα στην πρόβλεψη ενός σπιτιού και για μοντέλα που βελτιστοποιούνται για πρόβλεψη σε ένα σύνολο σπιτιών.

4. Παραγωγή βέλτιστων μοντέλων με χρήση γενετικού αλγορίθμου. Επειδή έχουν ήδη εκπαιδευτεί αρκετά διαφορετικά μοντέλα και για να επιταχυνθεί αυτό το βήμα, χρησιμοποιείται ένας Random Forest Regressor ο οποίος στηριζόμενος στα έως τώρα διαθέσιμα μοντέλα, μαθαίνει να εκτιμά τη μετρική precision, για δοσμένα χαρακτηριστικά εισόδου στο μοντέλο. Αυτό θα χρησιμοποιηθεί από τον γενετικό αλγόριθμο για την (αυτοματοποιημένη και πιο συστηματική) παραγωγή μοντέλων Transformer τα οποία έχουν καλή επίδοση στην πρόβλεψη του επιπέδου λειτουργίας του boiler (Αυτό το βήμα υλοποιείται δύο φορές ξεχωριστά. Μία για πρόβλεψη σε ένα σπίτι και μία για πρόβλεψη σε σύνολο σπιτιών)

5. Συγκριτική ανάλυση αποτελεσμάτων. Θα συγκρίνουμε ως προς την επίδοσή τους τα μοντέλα που έχουν προκύψει ως καλύτερα σε κάθε κατηγορία. Αυτά αποτελούνται από τα μοντέλα που επιλέξαμε αρχικά ως πιθανές βέλτιστες επιλογές και τα μοντέλα που προέκυψαν από τον γενετικό αλγόριθμο ως πιο ακριβή στις προβλέψεις τους. Έτσι, καταλήγουμε σε δύο διαδικασίες σύγκρισης. Από τη μία, για τα μοντέλα που έχουν προσαρμοστεί στην πρόβλεψη της συμπεριφοράς του boiler σε ένα σπίτι. Από την άλλη, τα μοντέλα που έχουν προβλεπτική δυνατότητα προσαρμοσμένη για ένα σύνολο σπιτιών.



Σχήμα 3.2: Σχηματική Απεικόνιση του Πειραματικού Σχεδιασμού

3.5.1 Προεπεξεργασία Δεδομένων

Η διαδικασία αυτή αποτελείται από τα ακόλουθα βασικά στάδια:

1. Την συμπλήρωση τιμών που λείπουν. Για τους σκοπούς της συγκεκριμένης εργασίας γίνεται εκπαίδευση των μοντέλων σε δεδομένα διαστήματος 1 λεπτού. Ωστόσο, τα αρχικά δεδομένα δεν είναι πλήρη καθώς σε μερικά από τα μεγέθη λείπουν ορισμένες μετρήσεις. Έτσι υπάρχουν "κενά" που αντιστοιχούν σε διαστήματα μερικών λεπτών το καθένα. Αφού πρώτα εντοπιστούν τα συγκεκριμένα "κενά", στη συνέχεια συμπληρώνονται με γραμμική παρεμβολή, δηλαδή με την μέση τιμή της προηγούμενης και της επόμενης υπάρχουσας μέτρησης.

2. Την δημιουργία επιπλέον μεγεθών. Για την καλύτερη αναπαράσταση των συνθηκών λειτουργίας του boiler και την εξαγωγή χρήσιμης πληροφορίας δημιουργούνται επιπλέον μεγέθη με πράξεις πάνω στα ήδη υπάρχοντα. Συγκεκριμένα:

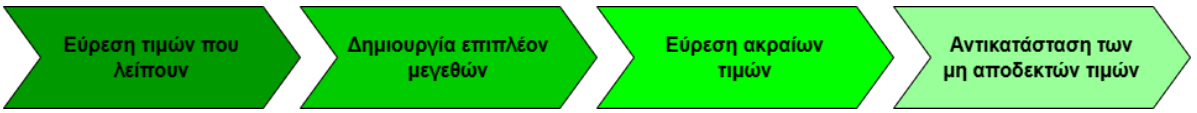
Μεταβλητή	**Υπολογισμός**	**Περιγραφή**
absorption	*blr_t - t_ret*	Η δυνατότητα του κάθε σπιτιού να απορροφήσει θερμότητα από τη λειτουργία του boiler, όπως αυτή προκύπτει από την μείωση της θερμοκρασίας του νερού μετά από έναν πλήρη κύκλο (από το boiler μέσω των σωλήνων σε όλα τα σώματα θέρμανσης στο σπίτι και τέλος πίσω στο boiler)
insulation	*t_r - t_out*	Η ικανότητα του κάθε σπιτιού να διατηρήσει την εσωτερική του θερμοκρασία σε σχέση με την εξωτερική θερμοκρασία. Δηλαδή, αντιπροσωπεύει την μονωτική δυνατότητα του σπιτιού.
day_of_year	*ο αριθμός της ημέρας εντός του έτους*	Ορίζεται ως ο αριθμός της ημέρας εντός του έτους, αντί για εντός του αντίστοιχου μήνα όπως με την μεταβλητή day. Δηλαδή, αντιστοιχεί στο συνδυασμό μεταβλητών day, month και υπολογίζεται ως μια εναλλακτική η οποία ίσως να αποδώσει καλύτερα αποτελέσματα.

Πίνακας 3.2: Πίνακας Ανάλυσης και Ερμηνείας των Νέων Μεγεθών

3. Την εύρεση εμφανώς ακραίων τιμών. Ορισμένες από τις κατηγορίες έχουν συγκεκριμένο εύρος αποδεκτών τιμών, οπότε μετρήσεις εκτός αυτού οφείλονται σε

σφάλμα μέτρησης και πρέπει να απορριφθούν. Για παράδειγμα, το μέγεθος blr_mod_lvl εκφράζει το επίπεδο λειτουργίας του boiler και παίρνει τιμές μόνο στο διάστημα 0% (καθόλου λειτουργία) έως 100% (μέγιστη λειτουργία).
Επίσης, γίνεται υπολογισμός του z-score ώστε να βρεθούν οι τιμές οι οποίες "απέχουν" πολύ από τη μέση τιμή του κάθε μεγέθους και άρα είναι ακραίες.

4. Την αντικατάσταση των τιμών οι οποίες εντοπίστηκαν στο προηγούμενο βήμα. Για να γίνει σωστή επεξεργασία στα δεδομένα, οι ακραίες τιμές που έχουν βρεθεί θα αφαιρεθούν. Στη θέση τους θα εισαχθούν τιμές οι οποίες με βάση τη συμπεριφορά του κάθε μεγέθους είναι πιο πιθανό να είναι "σωστές". Οι νέες, διορθωμένες τιμές θα υπολογιστούν με χρήση γραμμικής παρεμβολής, ως ο μέσος όρος της προηγούμενης και της επόμενης υπάρχουσας τιμής.



Σχήμα 3.3: Σχηματική Απεικόνιση της Προεπεξεργασίας των Δεδομένων

3.5.2 Προετοιμασία Δεδομένων

Έπειτα από τα στάδια προεπεξεργασίας των δεδομένων, ακολουθούν τα απαραίτητα βήματα ώστε να είναι έτοιμα να χρησιμοποιηθούν για την εκπαίδευση των μοντέλων. Για να γίνει αυτό χρειάζεται πρώτα να γίνουν κάποια επιπλέον βήματα.

1) Πρώτα να κανονικοποιηθούν τα δεδομένα, γιατί αφορούν διαφορετικές κατηγορίες, οι οποίες έχουν και διαφορετικά όρια τιμών. Για να μπορούν να εκπαιδευτούν τα μοντέλα με το βέλτιστο τρόπο, χωρίς να επηρεάζονται από σχετικά ανόμοιες τιμές διαφορετικών κατηγοριών, είναι επιθυμητό να κανονικοποιηθούν οι τιμές όλων στο διάστημα [0, 1]. Για να επιτευχθεί αυτό χρησιμοποιείται ένας κλιμακωτής (scaler) βελτιστοποιημένος για κάθε κατηγορία. Ο λόγος για τον οποίο δεν χρησιμοποιείται ο ίδιος scaler για όλες τις κατηγορίες είναι επειδή κάθε μία έχει διαφορετικό εύρος τιμών.
Επίσης, ο κλιμακωτής (scaler) είναι διαφορετικός και για κάθε σπίτι, αφού το κάθε ένα βρίσκεται σε διαφορετική περιοχή και άρα έχουν μεταξύ τους διαφορετικά χαρακτηριστικά. Συνεπώς, για μία συγκεκριμένη μεταβλητή παρατηρείται διαφορετική συμπεριφορά από σπίτι σε σπίτι και άρα ο κλιμακωτής που την κανονικοποιεί επιλέγεται να είναι ξεχωριστός ανά σπίτι.
Στη συνέχεια, οι τιμές της κάθε κατηγορίας αντιστοιχίζονται μέσω του κατάλληλου κλιμακωτή στο διάστημα από 0 έως 1. Με αυτό τον τρόπο οι τιμές οι οποίες θα χρησιμοποιηθούν για την εκπαίδευση των μοντέλων δεν θα τα επηρεάζουν υπέρμετρα προς μία κατεύθυνση αποφάσεων.

Έπειτα, πρέπει να γίνει και κατάλληλη μορφοποίηση των δεδομένων. Το μοντέλο θα λαμβάνει ως είσοδο μια λίστα με μήκος ίσο με την διάρκεια εισόδου (πόσα λεπτά "κοιτάει" πίσω για να παράξει την πρόβλεψη). Για κάθε λεπτό πρέπει να έχει τις τιμές όλων των

κατηγοριών στις οποίες θα στηριχθεί η πρόβλεψη. Μέχρι στιγμής, η μορφή των δεδομένων είναι ακόμα ως μία ξεχωριστή λίστα ανά κατηγορία. Κάθε λίστα απαρτίζεται από τιμές, κάθε μία από τις οποίες αντιπροσωπεύει διάρκεια ενός λεπτού.

2) Πρώτο βήμα σε αυτή τη διαδικασία είναι η συνένωση των δεδομένων των διαφορετικών κατηγοριών. Δημιουργείται μια συνάρτηση η οποία για κάθε επιθυμητό συνδυασμό κατηγοριών θα παράγει μία λίστα ως "συνένωση" των επιλεγμένων κατηγοριών. Συγκεκριμένα, για κάθε χρονική στιγμή, θα έχει αντί για μία τιμή, ένα σύνολο τιμών οι οποίες προέρχονται από τις διαφορετικές κατηγορίες. Ο λόγος για τον οποίο αυτό υλοποιείται μέσω συνάρτησης είναι επειδή θα χρειαστεί να επαναληφθεί για κάθε μοντέλο αφού για κάθε διαφορετικό αρχικό συνδυασμό κατηγοριών, η τελική λίστα θα περιέχει και διαφορετικό πλήθος τιμών για κάθε στιγμή.

3) Για να μπορεί να προκύψει ένα μοντέλο το οποίο θα προβλέπει επιλεγμένη χρονική διάρκεια, δημιουργείται μια νέα συνάρτηση η οποία λαμβάνει ως είσοδο τη λίστα των τιμών εισόδου και τη διασπά σε μικρότερες λίστες. Κάθε μία από αυτές έχει μήκος ίσο με την επιθυμητή διάρκεια πρόβλεψης. Για παράδειγμα, για 60 λεπτά (60 στιγμές), ως είσοδος θα είναι λίστα μήκους 60 με τις τιμές των επιλεγμένων κατηγοριών για κάθε στιγμή.

4) Τελευταίο βήμα είναι ο κατάλληλος διαχωρισμός των δεδομένων. Για να γίνει σωστά η εκπαίδευση μοντέλων είναι επιθυμητό να υπάρχει ένα σύνολο δεδομένων εκπαίδευσης (train dataset), ένα σύνολο δεδομένων επικύρωσης (validation dataset) και τέλος ένα δοκιμαστικό σύνολο δεδομένων (test dataset).

   A. Το δοκιμαστικό σύνολο (test dataset), το οποίο χρησιμοποιείται μετά την ολοκλήρωση της διαδικασίας εκπαίδευσης για να ελέγξει πόσο καλά το μοντέλο μπορεί να προβλέψει νέα δεδομένα. Αποτελείται από το τελευταίο 20% των συνολικών δεδομένων (δεν παίρνουμε τυχαίο 20% των δεδομένων, αφού πρόκειται για σειριακά δεδομένα στα οποία έχει σημασία η θέση της κάθε τιμής στην χρονοσειρά).

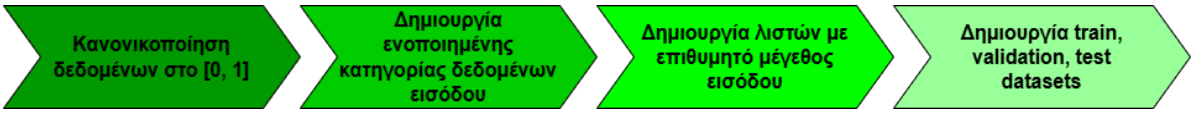
   Από τα υπόλοιπα δεδομένα, τα οποία αποτελούν το αρχικό 80% των συνολικών, διαχωρίζουμε τα άλλα δύο σύνολα:

   B. Το σύνολο εκπαίδευσης (train dataset), το οποίο όπως δηλώνεται από την ονομασία, γίνεται η εκπαίδευση των δεδομένων στις διαδοχικές εποχές της διαδικασίας.

   C. Το σύνολο επικύρωσης (validation dataset), το οποίο χρησιμοποιείται στο τέλος κάθε εποχής για να ελεγχθεί πόσο καλά το μοντέλο προβλέπει νέα δεδομένα, πέρα από αυτά στα οποία εκπαιδεύτηκε.

Έτσι, τώρα τα δεδομένα είναι έτοιμα να εισαχθούν στο κάθε μοντέλο, αρχικά για εκπαίδευση και επικύρωση στη διάρκεια των εποχών εκπαίδευσης και τέλος για παραγωγή των προβλέψεων από το δοκιμαστικό σύνολο. Αυτές οι προβλέψεις είναι που θα χρησιμοποιηθούν για την παραγωγή των μετρικών σφαλμάτων και ακριβείας για την αξιολόγηση και σύγκριση μεταξύ των διαφόρων μοντέλων.

Για να μπορέσουν να εξαχθούν αποτελέσματα και συμπεράσματα τα οποία να είναι συγκρίσιμα μεταξύ τους θα χρησιμοποιηθούν συγκεκριμένες δομές τόσο για τα μοντέλα τύπου LSTM όσο και για τα μοντέλα τύπου Transformer.



Σχήμα 3.4: Σχηματική Απεικόνιση της Προετοιμασίας των Δεδομένων

3.5.3 Δομή Μοντέλου LSTM

Τα μοντέλα τύπου LSTM τα οποία εκπαιδεύονται στη συγκεκριμένη μελέτη χρησιμοποιούνται ως βάση σύγκρισης των μοντέλων Transformer. Συγκεκριμένα για κάθε έναν Μετασχηματιστή που εκπαιδεύεται, θα έχουμε και ένα LSTM ως σχετικά πιο απλό μοντέλο τεχνητής νοημοσύνης, για να μπορούμε να συμπεράνουμε εάν οι Transformers (οι οποίοι τελικά μας ενδιαφέρουν) αποδίδουν συγκριτικά καλύτερα ή όχι.

Αναφορικά με τα μοντέλα LSTM, δημιουργούνται με χρήση της βιβλιοθήκης tensorflow. Χρησιμοποιούνται αρχικά 3 επίπεδα τύπου LSTM με σταδιακά μειούμενο μέγεθος. Ξεκινάμε από ένα με 120 μονάδες, έπειτα ένα με 60 και το τελευταίο με 32. Έπειτα ακολουθούν 3 επίπεδα Dense. Το πρώτο μεγέθους 60, το δεύτερο 30 ενώ το μέγεθος του τελευταίου εξαρτάται από το πλήθος των προβλεπόμενων τιμών. Συγκεκριμένα, εάν θέλουμε να προβλέψουμε μία ολόκληρη ημέρα, θα έχει μέγεθος 1440 (1440 λεπτά σε διάρκεια μιας ημέρας), ενώ για πρόβλεψη 10 λεπτών θα έχει μέγεθος 10.

3.5.4 Δομή Μοντέλου Transformer

Τα μοντέλα Transformer είναι αυτά για τα οποία εν τέλει ενδιαφερόμαστε και τα αποτελέσματα και τη συμπεριφορά των οποίων θα μελετήσουμε ιδιαίτερα. Όσον αφορά τη δομή την οποία χρησιμοποιούμε ισχύουν τα ακόλουθα στοιχεία για τις υπερπαραμέτρους:

1. Αριθμός των επιπέδων ορίζεται σε 4. Αφορά τον συνολικό αριθμό των επιπέδων (layers) στον κωδικοποιητή και στον αποκωδικοποιητή του μοντέλου. Περισσότερα επίπεδα επιτρέπουν την μάθηση πιο σύνθετων μοτίβων, αυξάνοντας όμως το υπολογιστικό κόστος

2. Διαστατικότητα ενσωματωσης (dimensionality of embeddings) είναι 64. Μεγαλύτερες τιμές αντιστοιχούν σε αναπαράσταση περισσότερης πληροφορίας ανά στοιχείο

3. Αριθμός κεφαλών προσοχής είναι 4. Επιτρέπει στο μοντέλο να εστιάζει ταυτόχρονα σε περισσότερα υποσύνολα πληροφορίας.

4. Εσωτερική διαστατικότητα του Feed Forward επιπέδου είναι 256. Συνήθως μεγαλύτερη της διαστατικότητας ενσωμάτωσης

5. Μέγεθος λεξιλογίου εισόδου, αντιστοιχεί στο μήκος των δεδομένων εισόδου (στη διάρκεια της εισόδου) που χρησιμοποιείται για τη παραγωγή της πρόβλεψης (για πρόβλεψη 1440 λεπτών, θα "κοιτάξει" στα προηγούμενα 1440 λεπτά)

6. Μέγεθος λεξιλογίου εξόδου, αντιστοιχεί στη διάρκεια του προβλεπόμενου διαστήματος

7. Ποσοστό απόρριψης είναι 0.1 . Χρησιμοποιείται για την αποφυγή υπερπροσαρμογής και ορίζει ποιο ποσοστό μονάδων θα απορριφθεί για αυτό το σκοπό. Το θέτουμε 10%

3.5.5 Υλοποίηση του Γενετικού Αλγορίθμου

Μέχρι αυτό το στάδιο έχουμε ήδη εκπαιδεύσει αρκετά μοντέλα. Τόσο απλών συνδυασμών κατηγοριών εισόδου (κυρίως για μια πρώτη εκτίμηση της κάθε κατηγορίας ως προς τη χρησιμότητά της για την παραγωγή των προβλέψεων) όσο και πιο σύνθετων συνδυασμών (τα οποία έχουμε εκτιμήσει ως πιθανά βέλτιστα για την περίπτωσή τους). Όμως, είναι πιθανό να μην έχουμε εντοπίσει κάποιους συνδυασμούς κατηγοριών δεδομένων οι οποίοι δίνουν μοντέλα με καλή προβλεπτική ικανότητα. Για αυτό θα υλοποιήσουμε έναν γενετικό αλγόριθμο:

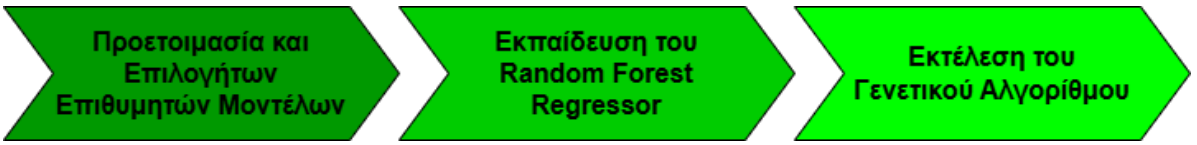
1. Προετοιμασία και επιλογή των επιθυμητών μοντέλων. Αρχικά επιλέγουμε τα κατάλληλα μοντέλα για να αξιοποιήσουμε ως βάση αυτής της διαδικασίας. Συγκεκριμένα, επειδή βελτιστοποιούμε δύο ξεχωριστές περιπτώσεις θα χρειαστεί σε κάθε μία από τις υλοποιήσεις να έχουμε κρατήσει μόνο τα σχετικά μοντέλα από αυτά που έχουμε ήδη εκπαιδεύσει.

2. Εκπαίδευση του Random Forest Regressor. Για να επιταχυνθεί η υλοποίηση του γενετικού αλγορίθμου και αφού έχουμε ήδη παράξει προβλέψεις από κάποια μοντέλα (για την περίπτωση του ενός σπιτιού και για την περίπτωση συνόλου σπιτιών) θα εκπαιδεύσουμε έναν Random Forest Regressor. Για δοσμένο σύνολο κατηγοριών εισόδου και πλήθος σπιτιών θα εκτιμά το αναμενόμενο precision.

Ο λόγος που προτιμάμε τη μετρική precision ως μέτρο βελτιστοποίησης οφείλεται στη δομή των δεδομένων και παρατηρήσεων επί των ήδη εκπαιδευμένων μοντέλων (όπως θα σχολιάσουμε στην επόμενη ενότητα που παρουσιάζουμε τα αποτελέσματα). Το boiler φαίνεται πως παραμένει απενεργοποιημένο για μεγάλο μέρος της ημέρας και άρα οι προβλέψεις είναι γενικά σωστές (επιτυχία στην πρόβλεψη της απενεργοποιημένης κατάστασης). Αυτό που μας ενδιαφέρει είναι να βρίσκουμε σωστά τις στιγμές που το boiler

ενεργοποιείται. Μάλιστα, από πλευράς εξοικονόμησης καυσίμου "κοστίζει" περισσότερο εάν προβλέψουμε λανθασμένη ενεργοποίηση, όταν το boiler δεν χρειάζεται στην πραγματικότητα να ανάψει (λειτουργεί και καίει καύσιμο χωρίς να χρειάζεται). Αυτά τα χαρακτηριστικά τα εκφράζει η μετρική precision και για αυτό την επιλέγουμε.

3. Εκτέλεση του γενετικού αλγορίθμου. Θα πραγματοποιήσουμε δύο εκτελέσεις. Μία για βελτιστοποίηση μοντέλων σε ένα σπίτι και μία για βελτιστοποίηση προβλέψεων σε σύνολο σπιτιών. Θα χρησιμοποιήσουμε διάρκεια 40 γενεών, με πιθανότητα διασταύρωσης (crossover) στο 0.5 (50%) και πιθανότητα μετάλλαξης (mutation) στο 0.2 (20%).



Σχήμα 3.5: Σχηματική Απεικόνιση της Λειτουργίας του Γενετικού Αλγορίθμου

4 Αποτελέσματα

Σε αυτή την ενότητα θα παρουσιάσουμε τα αποτελέσματα των μοντέλων για τις δύο περιπτώσεις. Από τη μία, θα αξιολογήσουμε τα μοντέλα τα οποία έχουν εκπαιδευτεί να προβλέπουν το αναμενόμενο επίπεδο λειτουργίας του boiler σε ένα σπίτι και άρα έχουν καλύτερη δυνατότητα προσαρμογής στο σύνολο δεδομένων που θα προβλέψουν. Από την άλλη, τα μοντέλα τα οποία έχουν εκπαιδευτεί σε ένα σύνολο σπιτιών και άρα έχουν βελτιστοποιηθεί ώστε να έχουν καλύτερη δυνατότητα πρόβλεψης επί του συνόλου. Δηλαδή, στη δεύτερη περίπτωση μας ενδιαφέρει η μέση επίδοση των μοντέλων και όχι οι επιμέρους προβλέψεις ανά σπίτι του συνόλου.
Τα μοντέλα Transformer κάθε περίπτωσης θα αξιολογηθούν τόσο μεταξύ τους όσο και με τα αντίστοιχα μοντέλα LSTM.

Επίσης σημειώνουμε ότι λόγω της ιδιομορφίας των δεδομένων θα κάνουμε και χρήση μετρικών επίδοσης, όπως το precision. Συγκεκριμένα, τα δεδομένα έχουν σε μεγάλο μηδέν επίπεδο λειτουργίας (το boiler είναι κλειστό) και μόνο σε κάποιες λίγες στιγμές εμφανίζονται κορυφές (peaks) που αντιστοιχούν στις περιόδους λειτουργίας του boiler. Επιπλέον, παρατηρούμε ότι οι ενεργοποιήσεις υπερβαίνουν το επίπεδο λειτουργίας 10%, οπότε θέτουμε αυτό ως όριό μας. Οπότε, για τιμές μικρότερες του 10% θεωρούμε ότι το boiler ειναι κλειστό (Ψευδές, τιμή 0) ενώ στην αντίθετη περίπτωση το boiler είναι ανοιχτό (Αληθές, τιμή 1).

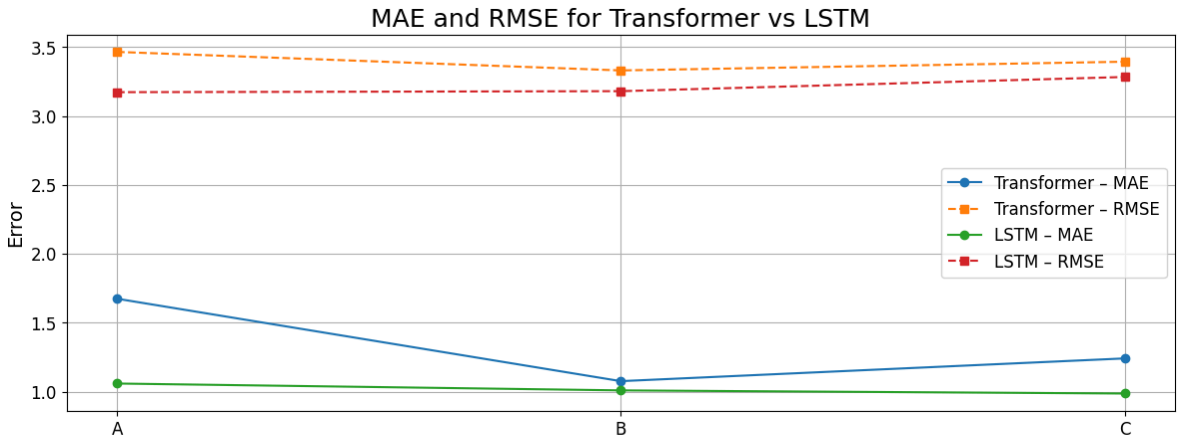
4.1 Επιλογή Βέλτιστης Διάρκειας Πρόβλεψης

Ως πρώτο βήμα της διαδικασίας πρέπει να επιλέξουμε τη βέλτιστη διάρκεια πρόβλεψης. Αυτό επιτεύχθηκε με την αξιολόγηση των ακόλουθων περιπτώσεων.
Αρχικά ξεκινήσαμε τις παρατηρήσεις από τις περιπτώσεις όπου προβλέπουμε είτε για μία ημέρα (1440 λεπτά) είτα για μία ώρα (60 λεπτά).
(Μάλιστα, λόγω των κακών επιδόσεων των μοντέλων A και C επιλέξαμε να εξετάσουμε και την πιο “σύνθετη” περίπτωση B σε περίπτωση που απέδιδε καλύτερα)

Μοντέλο	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**
A	60	blr_mod_lvl
B	60	blr_mod_lvl, day_of_year, hour_minute
C	1440	blr_mod_lvl

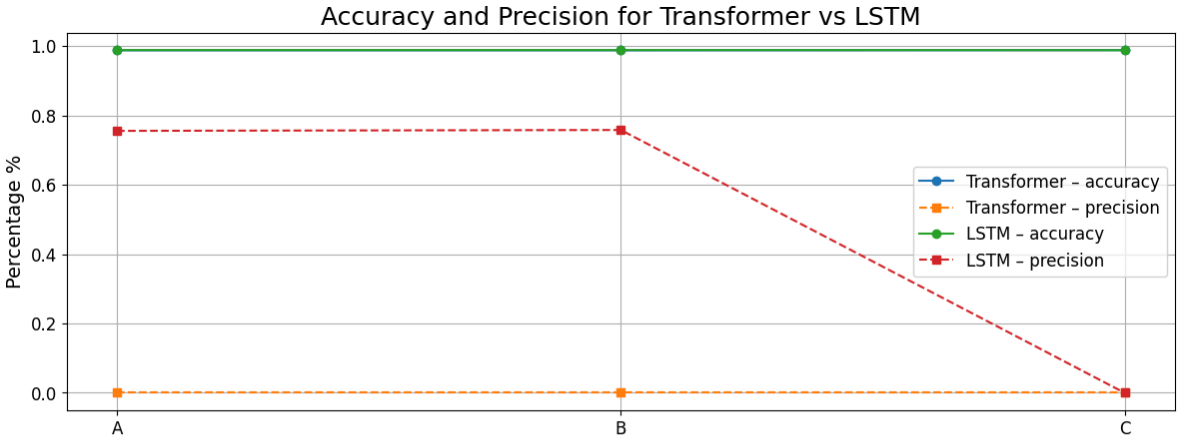
Πίνακας 4.1: Μοντέλα Αρχικής Δοκιμής Βέλτιστης Διάρκειας Πρόβλεψης

Βλέπουμε ότι όσον αφορά τις μετρήσεις του MAE και του RMSE, έχουμε μικρά σφάλματα:



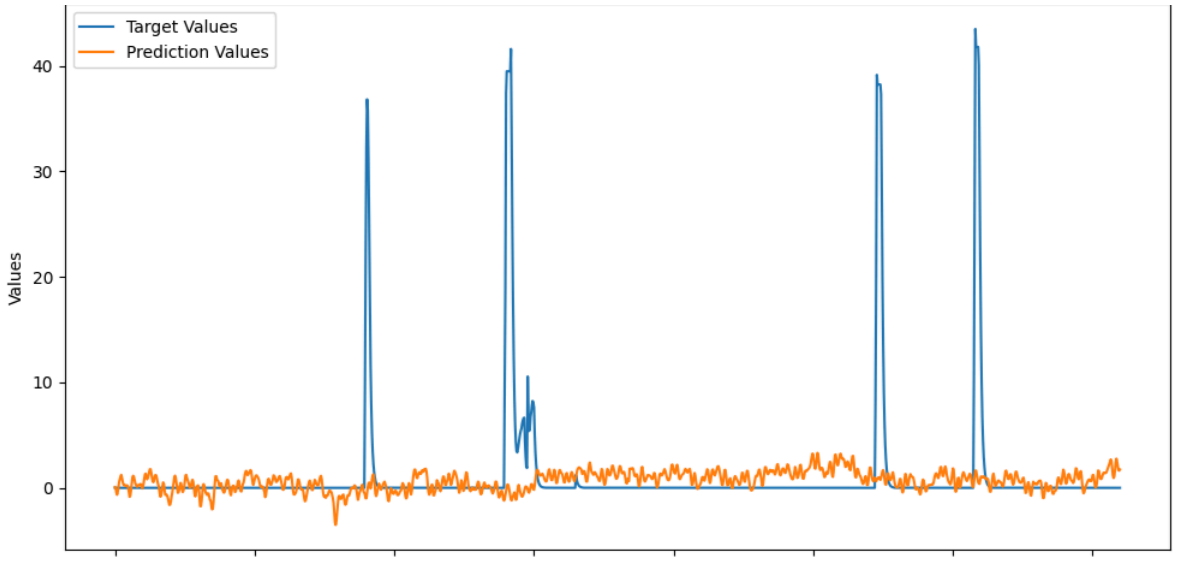
Σχήμα 4.1: MAE και RMSE για Μοντέλα Αρχικής Δοκιμής Βέλτιστης Διάρκειας Πρόβλεψης

Το accuracy είναι σχεδόν 1 για όλα τα μοντέλα, δηλαδή από όλες τις προβλέψεις, οι σωστές είναι σχεδόν στο 100% .
Όμως, μελετώντας το precision παρατηρούμε ότι κανένα μοντέλο Transformer δεν είναι εμφανώς μεγαλύτερο από το 0.0 (0%) . Δηλαδή, στο σύνολο των θετικών προβλέψεων δεν είναι σχεδόν καμία σωστή.



Σχήμα 4.2: Accuracy και Precision για Μοντέλα Αρχικής Δοκιμής Βέλτιστης Διάρκειας Πρόβλεψης

Αυτή η συμπεριφορά μπορεί να εξηγηθεί από το γεγονός ότι τα μοντέλα αυτά δεν προσαρμόζονται καλά στα δεδομένα που θέλουμε να προβλέψουν. Συγκεκριμένα για μοντέλο Transformer με είσοδο μόνο την κατηγορία “blr_mod_lvl” και διάρκεια πρόβλεψης 1440 λεπτά (1 ημέρα) έχουμε το ακόλουθο διάγραμμα για μία τυχαία ημέρα. Σε αυτό παρατηρούμε ότι η πρόβλεψη συνεχώς κυμαίνεται γύρω από το μηδέν. Αυτό είναι λάθος, αφού αν και οι περισσότερες τιμές είναι 0 (κλειστό boiler τον περισσότερο καιρό) σε κάποιες λίγες στιγμές εμφανίζονται κορυφές, όταν το boiler είναι ανοιχτό, τις οποίες και μας ενδιαφέρει να προβλέψουμε όσο το δυνατόν πιο σωστά.

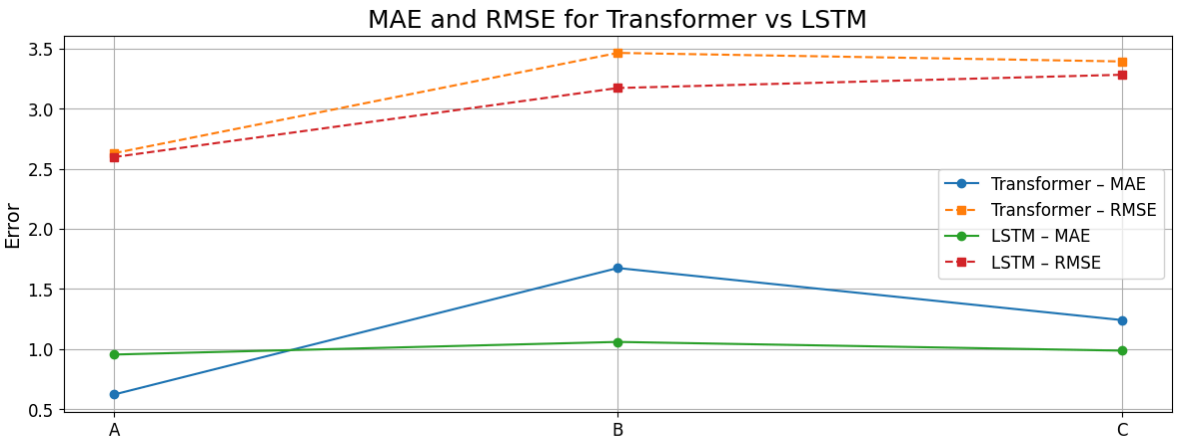


Σχήμα 4.3: Πρόβλεψη τυχαίας ημέρας από το test set για Transformer με είσοδο 'blr_mod_lvl' και πρόβλεψη για 1440 λεπτά

Επομένως, αξίζει να μελετήσουμε και την προβλεπτική συμπεριφορά των μοντέλων και για πιο μικρούς χρονικούς ορίζοντες, όπως τα 10 λεπτά. Για την επόμενη σύγκριση θα εστιάσουμε στην απλή περίπτωση όπου γίνεται πρόβλεψη με είσοδο μόνο την κατηγορία blr_mod_lvl (κατηγορία που προβλέπουμε είναι η blr_mod_lvl). Συγκεντρωτικά για τις τρεις περιπτώσεις χρονικής διάρκειας (10, 60, 1440 λεπτά) έχουμε τα ακόλουθα:

Μοντέλο	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**
A	10	blr_mod_lvl
B	60	blr_mod_lvl
C	1440	blr_mod_lvl

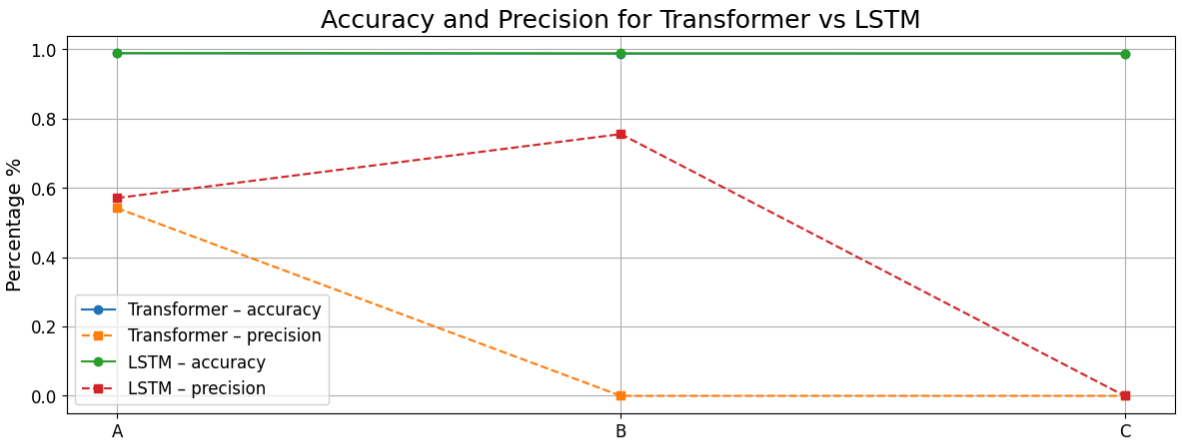
Πίνακας 4.2: Μοντέλα Σύγκρισης Βέλτιστης Διάρκειας Πρόβλεψης



Σχήμα 4.4: MAE και RMSE για Μοντέλα Σύγκρισης Βέλτιστης Διάρκειας Πρόβλεψης

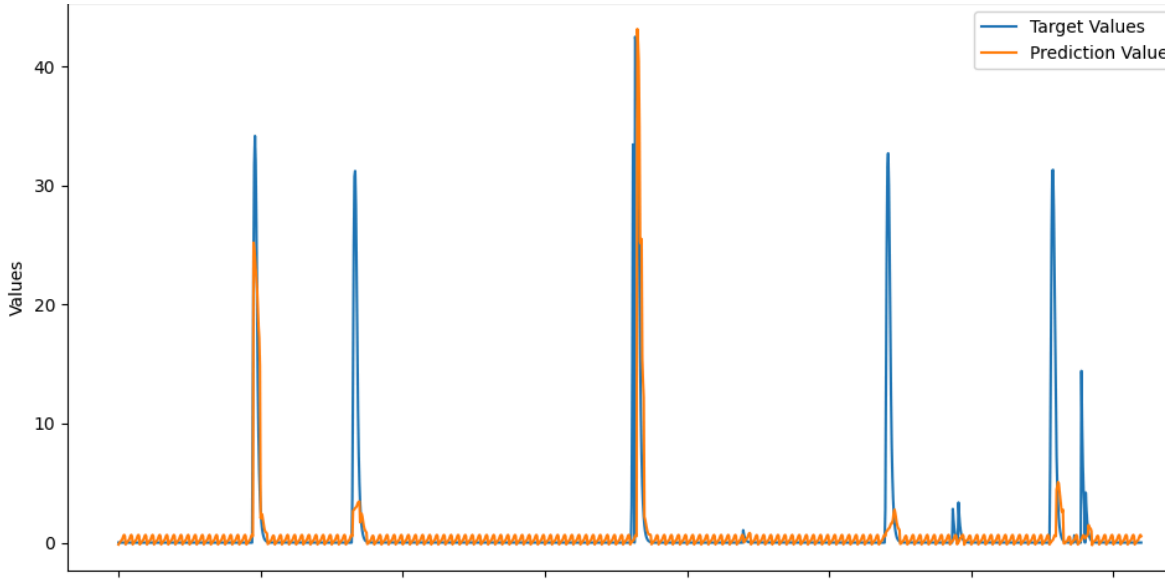
Όπως παρατηρούμε παραπάνω τα σφάλματα MAE και RMSE είναι εμφανώς μικρότερα στην περίπτωση των 10 λεπτών.

Και εδώ στα μοντέλα Transformer, για accuracy, η τιμή παραμένει σχεδόν στο 1 (100%), όμως για το recall στην περίπτωση των 10 λεπτών έχουμε τιμή περίπου 0.5 (50%), δηλαδή προβλέπονται σωστά σχεδόν τα μισά θετικά αποτελέσματα:



Σχήμα 4.5: Accuracy και Precision για Μοντέλα Σύγκρισης Βέλτιστης Διάρκειας Πρόβλεψης

Πράγματι, κοιτώντας μία τυχαία ημέρα του test set για τον Transformer που προβλέπει σε διάρκεια 10 λεπτών, παρατηρούμε καλή προσαρμογή τόσο στις "κορυφές" όταν το boiler είναι ενεργοποιημένο, όσο και στα υπόλοιπα διαστήματα που έχουμε 0 επίπεδο λειτουργίας:



Σχήμα 4.6: Πρόβλεψη τυχαίας ημέρας από το test set για Transformer με είσοδο 'blr_mod_lvl' και πρόβλεψη για 10 λεπτά

Οπότε, από εδώ και στο εξής θα προβλέπουμε για χρονικό διάστημα 10 λεπτών. Αυτό είναι επαρκώς μικρό ώστε να εντοπίζονται οι “κορυφές”, όταν το boiler είναι ενεργοποιημένο, πετυχαίνοντας επίσης καλή προσαρμογή στις υπόλοιπες στιγμές, όταν έχουμε τιμή 0 (απενεργοποιημένο boiler).

4.2 Αποτελέσματα Πρόβλεψης σε Ένα Σπίτι

Σε αυτό το τμήμα θα μελετήσουμε τα αποτελέσματα που αφορούν την πρόβλεψη του επιπέδου λειτουργίας του boiler σε ένα μόνο σπίτι. Σε αυτή την περίπτωση, τα δεδομένα παρουσιάζουν κάποια κοινά χαρακτηριστικά (σε αντίθεση με σύνολο σπιτιών, όπου το κάθε σπίτι έχει τα δικά του χαρακτηριστικά) και άρα είναι εφικτή μία αρκετά καλή προσαρμογή στο σύνολο δεδομένων. Αυτό θα έχει ως αποτέλεσμα καλύτερη επιτυχία στις προβλέψεις.

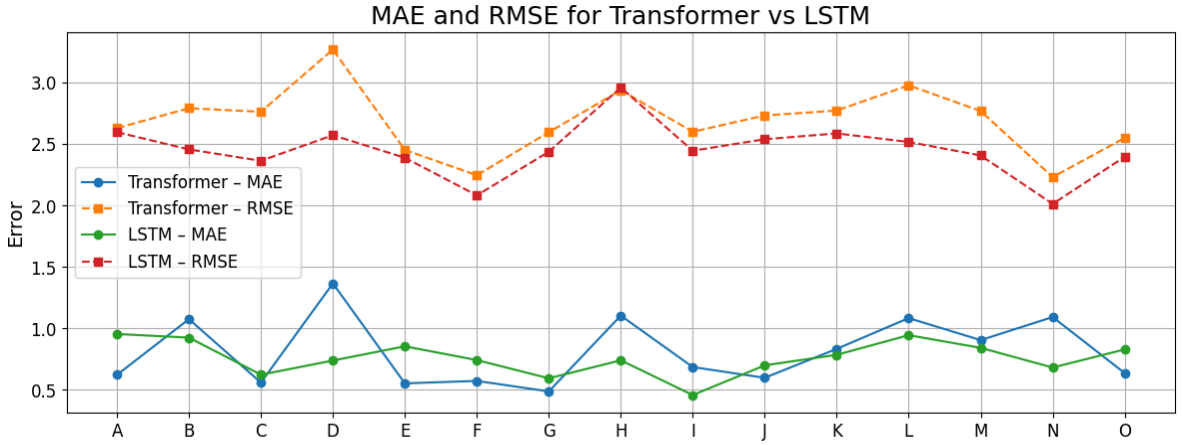
4.2.1 Απλοί συνδυασμοί κατηγοριών

Δεδομένου ότι η κατηγορία blr_mod_lvl είναι αυτή που θέλουμε να προβλέψουμε είναι εύλογο να υποθέσουμε ότι θα συμπεριλαμβάνεται ως είσοδος σε κάθε μοντέλο που εν τέλει θα έχει καλή συμπεριφορά.
Αρχικά, όμως, θέλουμε να αποκτήσουμε μία ιδέα για την καταλληλότητα στην εξαγωγή καλών προβλέψεων κάθε μίας εκ των υπόλοιπων κατηγοριών. Για το λόγο αυτό, θα δούμε αρχικά τα αποτελέσματα μοντέλων τα οποία έχουν ως είσοδο κάθε μία από τις κατηγορίες.

Μοντέλο	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**
A	10	blr_mod_lvl
B	10	blr_mod_lvl, absorption
C	10	blr_mod_lvl, blr_t
D	10	blr_mod_lvl, bypass
E	10	blr_mod_lvl, flame
F	10	blr_mod_lvl, heat
G	10	blr_mod_lvl, insulation
H	10	blr_mod_lvl, otc_cur
I	10	blr_mod_lvl, otc_maxt
J	10	blr_mod_lvl, t_out
K	10	blr_mod_lvl, t_r
L	10	blr_mod_lvl, t_r_set
M	10	blr_mod_lvl, t_ret
N	10	blr_mod_lvl, t_set
O	10	blr_mod_lvl, water

Πίνακας 4.3: Μοντέλα απλών συνδυασμών κατηγοριών σε ένα σπίτι

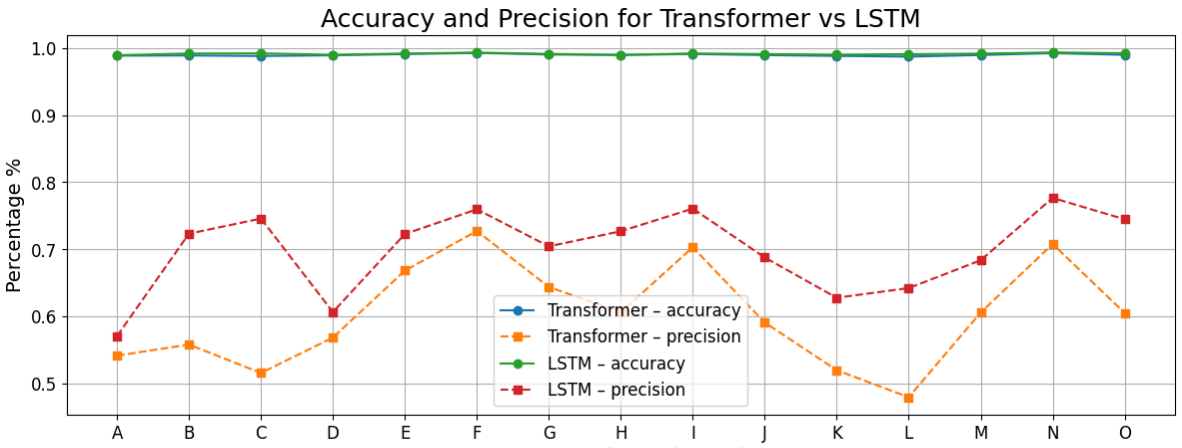
Όσον αφορά τις μετρικές σφαλμάτων MAE, RMSE έχουμε:



Σχήμα 4.7: MAE και RMSE για μοντέλα απλών συνδυασμών κατηγοριών σε ένα σπίτι

Πράγματι διαπιστώνουμε ότι γίνονται γενικά μικρά σφάλματα, αφού το MAE έχει, πλην λίγων εξαιρέσεων, τιμές που δεν ξεπερνούν τη 1 μονάδα. Από την άλλη όταν γίνονται σφάλματα, είναι εμφανή αλλά όχι ιδιαίτερα σημαντικά, αφού το RMSE κυμαίνεται γύρω στο 2.5 (και εν γένει δεν υπερβαίνει το 3.0). Βέβαια τα Transformer μοντέλα έχουν ελαφρά μεγαλύτερη τιμή RMSE από τα αντίστοιχα LSTM.
Αυτά τα αποτελέσματα είναι εν γένει αναμενόμενα, αφού το κάθε μοντέλο στηρίζεται σε δύο μόνο κατηγορίες (μία από τις διαθέσιμες κατηγορίες των δεδομένων και το επίπεδο λειτουργίας του boiler) και άρα υπάρχουν χαρακτηριστικά του προβλήματος που δεν μπορούν να εντοπίσουν ή να περιγράψουν.

Για τις μετρικές accuracy και precision έχουμε:



Σχήμα 4.8: Accuracy και Precision για μοντέλα απλών συνδυασμών κατηγοριών σε ένα σπίτι

Το accuracy είναι πρακτικά 1.0 , τόσο για Transformer όσο και για τα LSTM μοντέλα. Αυτό υποδηλώνει ότι οι σωστές προβλέψεις αποτελούν σχεδόν το σύνολο όλων των προβλέψεων, άρα δεν γίνονται πολλές λάθος προβλέψεις.
Βέβαια, από τις θετικές προβλέψεις, σωστές είναι ένα ποσοστό το οποίο ανάλογα την περίπτωση κυμαίνεται από περίπου 0.5 (ακριβώς οι μισές) έως και σχεδόν 0.8 . Εδώ αξίζει

να σημειωθεί πως τα Transformer μοντέλα δεν αποδίδουν τόσο καλά όσο τα αντίστοιχα LSTM. Με την τιμή του precision να είναι από 0.02 έως και 0.2 μονάδες μικρότερο, δηλαδή ακόμα και στις πιο ευνοϊκές περιπτώσεις έχουν 2% (περίπου) λιγότερη ακρίβεια στην σωστή πρόβλεψη ενεργοποιήσεων.

Από τα παραπάνω αποτελέσματα μπορούμε να συμπεράνουμε ότι τα Transformer μοντέλα βρίσκουν με αρκετά καλή ακρίβεια πότε ενεργοποιείται το boiler. Με πιθανώς κάποιες αστοχίες για το πόσο ακριβώς είναι το επίπεδο λειτουργίας σε ορισμένες περιπτώσεις, αφού όταν κάνουν λάθη, αυτά είναι πιο μεγάλα από ότι στα LSTM.

Όσον αφορά στην αξιολόγηση των "απλών" συνδυασμών κατηγοριών για τα Transformer μοντέλα με βάση τη συμπεριφορά που παρουσιάζουν ως προς το precision. Παρατηρούμε πως η πρόβλεψη μόνο με boiler modulation level δεν είναι από μόνη της αρκετά υψηλή (μόλις 55%). Ιδιαίτερα καλά αποτελέσματα έχουμε όταν συνδυάζεται αυτή η κατηγορία με μία από τις υπόλοιπες. Δηλαδή, για πιο σωστή εκτίμηση ενεργοποιήσεων, το precision να είναι τουλάχιστον 70%, το μοντέλο πρέπει να γνωρίζει αντίστοιχα αν έχει ενεργοποιηθεί η κυκλοφορία στο boiler (κατηγορία heat), ποιο είναι το μέγιστο επίπεδο που θα φτάσει το boiler (otc_maxt) ή την επιθυμητή θερμοκρασία που πρέπει να φτάσει το boiler (t_set). Συνεπώς, εκτιμούμε ότι αυτές οι κατηγορίες εκφράζουν σημαντικές πτυχές της λειτουργίας του boiler και άρα η συμπερίληψή τους στα δεδομένα εισόδου οδηγεί σε πιο ακριβείς προβλέψεις.

4.2.2 Επιλογή Πιθανώς Βέλτιστων Μοντέλων

Συνεχίζουμε να εξετάζουμε μοντέλα τα οποία εκπαιδεύονται σε 1 μόνο σπίτι. Όμως, τώρα θα εστιάσουμε σε πιο "σύνθετους" συνδυασμούς κατηγοριών. Συγκεκριμένα, έχουμε δει τα αποτελέσματα όταν χρησιμοποιείται ως είσοδος η προβλεπόμενη κατηγορία (blr_mod_lvl : επίπεδο λειτουργίας boiler) μόνη της ή μαζί με μόνο μία από τις υπόλοιπες κατηγορίες. Θα επικεντρωθούμε στην περίπτωση που χρησιμοποιούνται ως είσοδος συνδυασμοί πολλών κατηγοριών με στόχο την εύρεση αυτών που θα έχουν τελικά την βέλτιστη προβλεπτική συμπεριφορά. Επιλέγουμε τις κατηγορίες που έχουν προηγουμένως δώσει τα καλύτερα αποτελέσματα ή αυτές που εκφράζουν μεγέθη τα οποία λογικά εκτιμούμε ότι συνδυάζονται για να περιγράψουν πλήρως τη συμπεριφορά του boiler και άρα θα έχουν καλή προβλεπτική δυνατότητα.

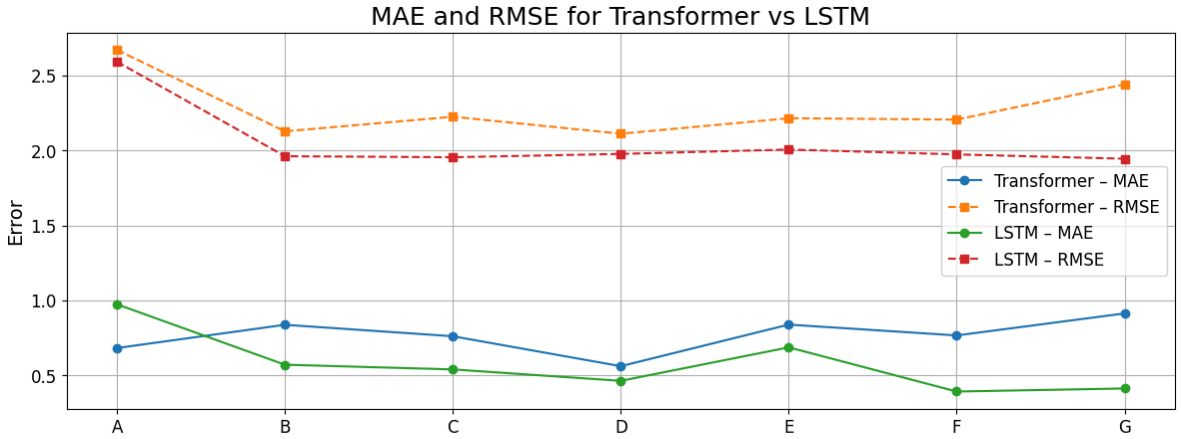
Θα δοκιμάσουμε τα μοντέλα που εκτιμούμε ότι θα έχουν καλύτερες επιδόσεις στην πρόβλεψη:

Μοντέλο	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**
A	10	blr_mod_lvl, day_of_year, hour, minute
B	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat

C	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, flame, water
D	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, day_of_year, hour, minute
E	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, flame, water, day_of_year, hour, minute
F	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, flame, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass
G	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, flame, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass, day_of_year, hour, minute

Πίνακας 4.4: Μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε ένα σπίτι

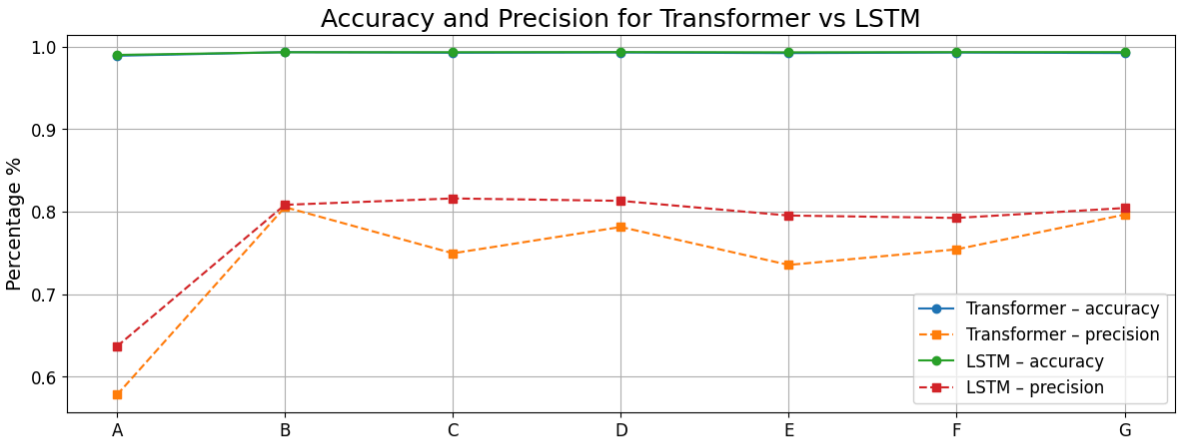
Όσον αφορά τις μετρήσεις των σφαλμάτων MAE, RMSE:



Σχήμα 4.9: MAE και RMSE για μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε ένα σπίτι

Παρατηρούμε ότι τα μοντέλα transformer εξακολουθούν να έχουν μεγαλύτερα λάθη από τα LSTM, αν και η διαφορά έχει μειωθεί σε λιγότερο από 0.3 . Αυτό ισχύει για το σύνολο των λαθών (MAE) όπου κανένα μοντέλο δεν υπερβαίνει το 1.0 . Ισχύει επίσης και όταν εστιάζουμε στα μεγάλα σφάλματα (RMSE) όπου όλες οι τιμές κυμαίνονται στο 2.3 .

Κοιτώντας τώρα τις αντίστοιχες μετρικές επίδοσης:



Σχήμα 4.10: Accuracy και Precision για μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε ένα σπίτι

Το accuracy, είναι για όλα τα μοντέλα κοντά στο 1 (100%), δηλαδή σχεδόν όλες οι προβλέψεις για την ενεργοποίηση ή όχι του boiler είναι σωστές.
Το precision έχει τιμές σχεδόν σταθερά στο 80% για τα LSTM και στο 75% για τα Transformer.

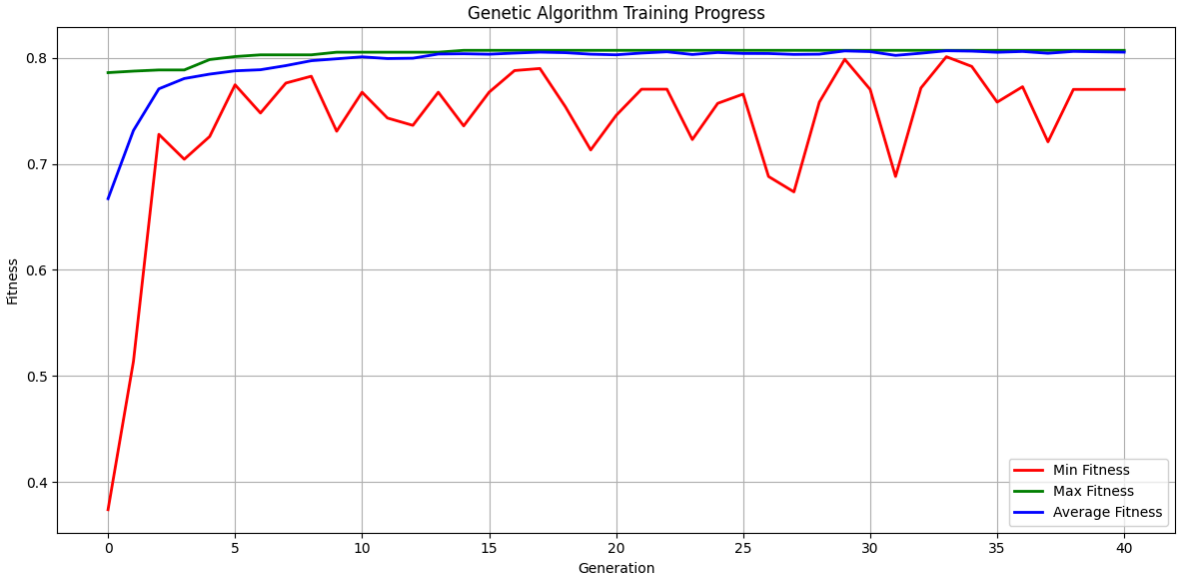
Παρατηρούμε δηλαδή ότι σε σχέση με τις περιπτώσεις λίγων κατηγοριών στην είσοδο έχουμε εμφανή βελτίωση από όλες τις απόψεις. Άλλωστε, το μοντέλο A είναι η απλή περίπτωση όπου ως είσοδος χρησιμοποιείται το επίπεδο λειτουργίας του boiler μαζί με την ημερομηνία. Όπως φαίνεται τα μοντέλα με πιο απλούς συνδυασμούς κατηγοριών εισόδου έχουν σημαντικά χαμηλότερες επιδόσεις.
Αναφορικά με τα υπόλοιπα μοντέλα, τα οποία έχουν είσοδο συνδυασμούς περισσοτέρων κατηγοριών, τα σφάλματα είναι σταθερά μικρότερα. Από την άλλη, οι σωστές θετικές προβλέψεις ως προς το σύνολο θετικών προβλέψεων (precision) έχουν επίσης αυξηθεί για όλα τα μοντέλα. Βέβαια, τα μοντέλα LSTM εξακολουθούν να αποδίδουν καλύτερα από τα μοντέλα Μετασχηματιστών.

4.2.3 Γενετική Παραγωγή Μοντέλων

Τώρα θα παράξουμε βέλτιστα μοντέλα για την πρόβλεψη του επιπέδου λειτουργίας σε ένα σπίτι με χρήση του γενετικού αλγορίθμου. Βέβαια, έχουμε ήδη δοκιμάσει κάποια μοντέλα που εκτιμούμε πως έχουν τη βέλτιστη δυνατή συμπεριφορά, ωστόσο θα χρησιμοποιήσουμε τον γενετικό αλγόριθμο για να βρούμε τυχόν μοντέλα που αγνοήσαμε. Δηλαδή, θα ελέγξουμε αν υπάρχουν κατηγορίες οι οποίες εκφράζουν σημαντικές πτυχές της λειτουργίας του boiler τις οποίες δεν έχουμε συμπεριλάβει στα βέλτιστα μοντέλα που έχουμε εξετάσει μέχρι στιγμής.

Για την επιτάχυνση της διαδικασίας χρησιμοποιείται ένας Random Forest Regressor ο οποίος έχει εκπαιδευτεί σε όλα τα μοντέλα που έχουμε ήδη παράξει.

Για τον γενετικό αλγόριθμο έχουμε την ακόλουθη συμπεριφορά:



Σχήμα 4.11: Διαδικασία εκτέλεσης γενετικού αλγορίθμου σε περίπτωση για ένα σπίτι

Παρατηρούμε ότι τα δοκιμαζόμενα μοντέλα γρήγορα συγκλίνουν σε ένα "βέλτιστο" precision σχεδόν στο 80% . Δηλαδή τα μοντέλα που προτείνει εκτιμούμε ότι έχουν πολύ καλή συμπεριφορά.

Το καλύτερο μοντέλο από την τελευταία γενιά έχει τα ακόλουθα χαρακτηριστικά:

Εκτιμώμενο precision	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**
0.807	10	blr_mod_lvl, insulation, t_out, blr_t, heat, t_set, otc_maxt, month

Πίνακας 4.5: Βέλτιστο μοντέλο σε ένα σπίτι από γενετικό αλγόριθμο

4.2.4 Σύγκριση Αποτελεσμάτων για Ένα Σπίτι

Τώρα μπορούμε να περάσουμε στην ολική σύγκριση των επιδόσεων των μοντέλων τα οποία αφορούν στην πρόβλεψη του επιπέδου λειτουργίας του boiler σε ένα σπίτι. Για το σκοπό θα συγκρίνουμε τα μοντέλα των "σύνθετων" συνδυασμών κατηγοριών τα οποία τελικά είχαν τις καλύτερες επιδόσεις και το μοντέλο με την καλύτερη αξιολόγηση από την διαδικασία του γενετικού αλγορίθμου.

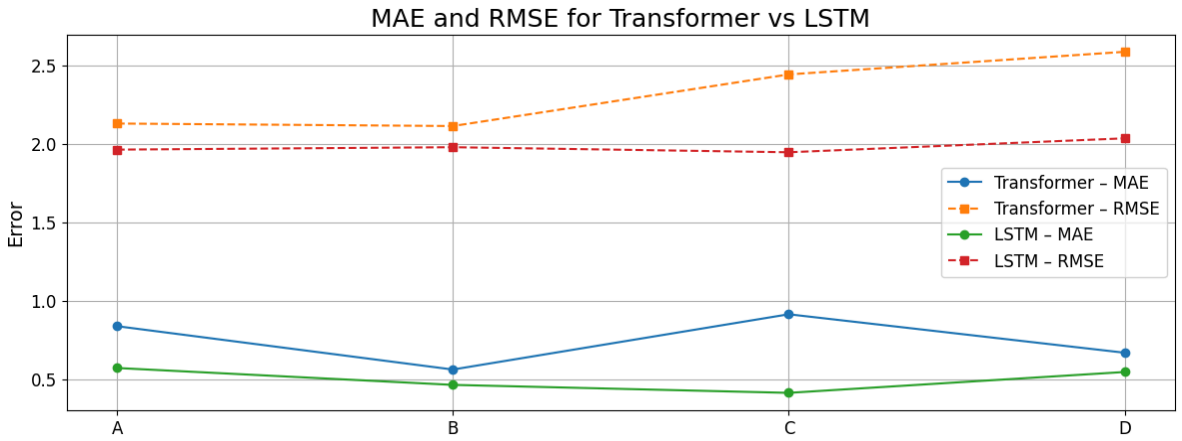
Γενικά, τα μοντέλα έχουν όλα επιδείξει χαμηλά σφάλματα ενώ η ουσιαστική τους διαφοροποίηση προκύπτει από την μετρική precision. Δηλαδή, από το πόσες φορές προβλέπουν ενεργοποίηση του boiler και αυτό όντως συμβαίνει.

Τα επιλεγμένα προς σύγκριση μοντέλα είναι:

Μοντέλο	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**	**Τρόπος Επιλογής**
A	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat	Αρχική επιλογή
B	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, day_of_year, hour, minute	Αρχική επιλογή
C	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, flame, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass, day_of_year, hour, minute	Αρχική επιλογή
D	10	blr_mod_lvl, insulation, t_out, blr_t, heat, t_set, otc_maxt, month	Γενετικός Αλγόριθμος

Πίνακας 4.6: Συνολικά βέλτιστα μοντέλα στην περίπτωση ενός σπιτιού

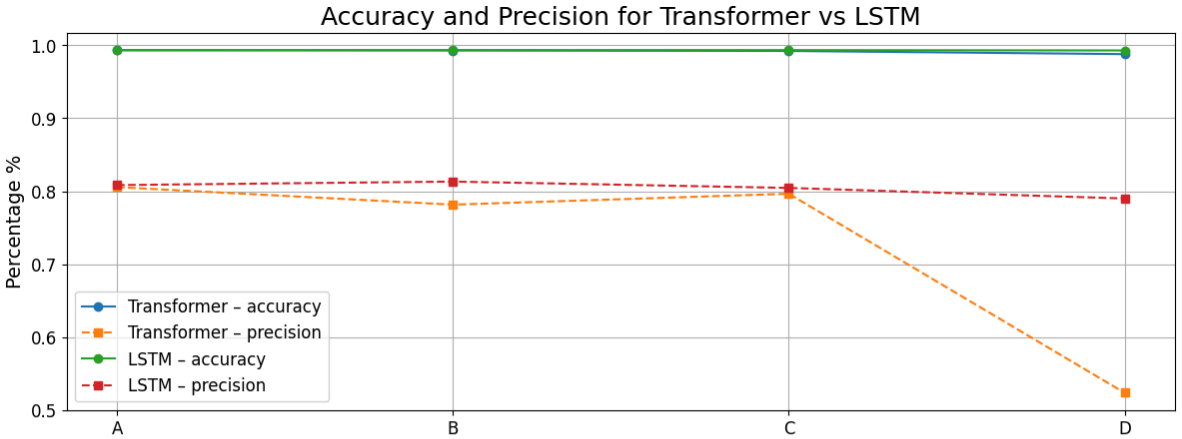
Όσον αφορά τις μετρήσεις σφαλμάτων MAE, RMSE έχουμε:



Σχήμα 4.12: MAE και RMSE για συνολικά βέλτιστα μοντέλα στην περίπτωση ενός σπιτιού

Διαπιστώνουμε ότι όλα τα μοντέλα έχουν αρκετά μικρά σφάλματα, με τις MAE να παραμένουν σταθερά μικρότερες από 1.0 . Όταν γίνονται σφάλματα, τότε δεν είναι ιδιαίτερα σημαντικά, αφού παρατηρούμε ότι το RMSE το οποίο επηρεάζεται περισσότερο από μεγάλα σφάλματα διατηρεί σταθερά μικρές τιμές μεταξύ 2.0 και 2.5 . Βέβαια, παραμένει χαρακτηριστικό ότι τα LSTM εξακολουθούν να αποδίδουν καλύτερα, έχοντας σε όλες τις περιπτώσεις μικρότερες τιμές σφαλμάτων από τα αντίστοιχα Transformer.

Όσον αφορά τις μετρικές επίδοσης:



Σχήμα 4.13: Accuracy και Precision για συνολικά βέλτιστα μοντέλα στην περίπτωση ενός σπιτιού

Το accuracy είναι σταθερά σχεδόν 1.0 , υποδηλώνοντας μία σχεδόν 100% επιτυχία στην πραγματοποίηση προβλέψεων από τα μοντέλα. Από την άλλη, το precision είναι περίπου στο 80% για τα μοντέλα τα οποία είχαμε επιλέξει αρχικά. Όμως, για το μοντέλο D το οποίο προέκυψε από την διαδικασία του γενετικού αλγορίθμου έχει τιμή λίγο μεγαλύτερη από 50%. Τα LSTM μοντέλα φαίνεται και από τις μετρικές επίδοσης ότι διατηρούν ένα οριακό προβάδισμα.

Συμπερασματικά, μπορούμε να αποφανθούμε ότι τα μοντέλα Transformer που εκπαιδεύονται σε ένα μόνο σπίτι μπορούν να προβλέψουν αρκετά καλά το επίπεδο λειτουργίας του boiler. Γενικά, τα λάθη είναι λίγα και μικρά όπως προκύπτει από τις τιμές

των MAE, RMSE. Επίσης όταν προβλέπουν ενεργοποίηση του boiler έχουν υψηλό ποσοστό επιτυχίας όπως φαίνεται από την μετρική precision.
Όσον αφορά τη σύγκρισή τους με τα LSTM μοντέλα, μπορούμε να αποφανθούμε ότι με πιο "σύνθετους" συνδυασμούς κατηγοριών ως είσοδο μπορούν να τα πλησιάσουν. Βέβαια, τα LSTM εξακολουθούν να έχουν καλύτερη δυνατότητα πρόβλεψης όταν χρησιμοποιούνται λίγες κατηγορίες στην είσοδο. Δηλαδή στην περίπτωση του ενός σπιτιού μπορούν να προσαρμοστούν πιο εύκολα και με λιγότερα δεδομένα από τους Transformers.
Σχετικά με το μοντέλο D, το οποίο προέκυψε ως το καλύτερο από την επιλογή του γενετικού αλγορίθμου, δεν έχει τόσο καλές επιδόσεις όσα τα υπόλοιπα. Παρατηρώντας τις κατηγορίες δεδομένων που λαμβάνει στην είσοδό του μπορούμε να δούμε ότι απαρτίζεται από κάποιες που έχουμε ήδη χρησιμοποιήσει στα άλλα 3 μοντέλα τα οποία αποδίδουν αρκετά καλά. Όμως δεν έχει την κατηγορία minute το οποίο είναι σημαντικό μειονέκτημα, αφού δεν μπορεί να γνωρίζει με ακρίβεια τις χρονικές στιγμές που πρέπει να προβλέψει, ειδικά αφού το παράθυρο πρόβλεψης είναι διάρκειας 10 λεπτών.

4.3 Αποτελέσματα Πρόβλεψης σε Σύνολο Σπιτιών

Τώρα θα εστιάσουμε σε μοντέλα τα οποία έχουν εκπαιδευτεί σε σύνολο των σπιτιών (28 σπίτια). Για κάθε ένα από αυτά τα μοντέλα οι μετρικές θα υπολογίζονται για κάθε ένα σπίτι ξεχωριστά και έπειτα θα υπολογίζεται η μέση τιμή τους. Με αυτό τον τρόπο, κάθε σπίτι του συνόλου θα έχει την ίδια βαρύτητα στην τελική αξιολόγηση του μοντέλου. Αυτό είναι ιδιαίτερα επιθυμητό, αφού στόχος του συγκεκριμένου τμήματος της εργασίας είναι να βρεθούν τα χαρακτηριστικά εκείνα τα οποία εν γένει θα επιτρέπουν σε ένα μοντέλο Transformer να προβλέπει το επίπεδο λειτουργίας του boiler για ένα σύνολο σπιτιών.

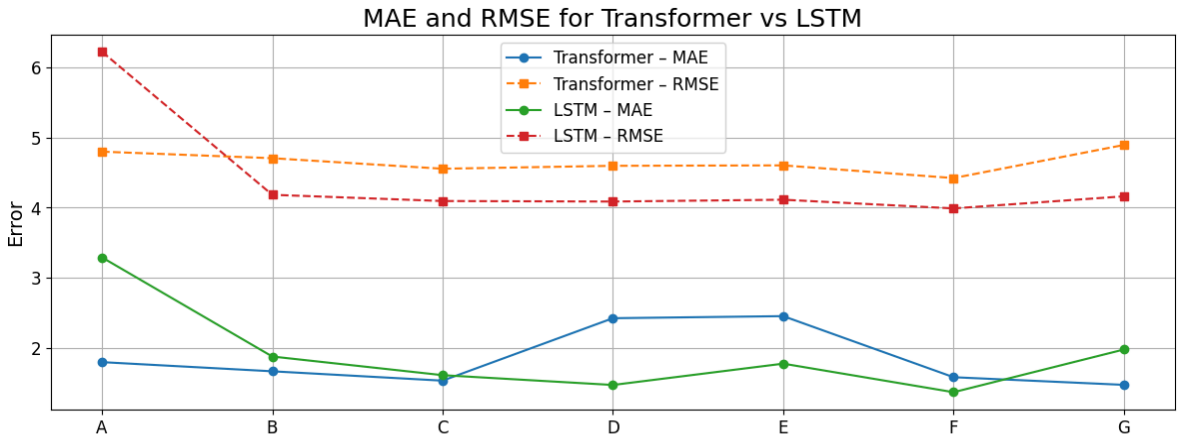
4.3.1 Επιλογή Πιθανώς Βέλτιστων Μοντέλων

Αρχικά, όπως και στην περίπτωση όπου εστιάζαμε σε ένα σπίτι θα επιλέξουμε να εκπαιδεύσουμε και να αξιολογήσουμε τα μοντέλα τα οποία ως είσοδο λαμβάνουν σύνολα κατηγοριών που είναι πιο πιθανό να έχουν καλή επίδοση. Η επιλογή αυτή στηρίζεται από τη μία, στις επιδόσεις διαφόρων κατηγοριών όπως φάνηκαν στην περίπτωση του ενός σπιτιού. Από την άλλη σε συνδυασμούς κατηγοριών οι οποίοι φαίνεται να περιγράφουν τη λειτουργία του boiler όσο το δυνατόν πληρέστερα.

Μοντέλο	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**
A	10	blr_mod_lvl, day_of_year, hour, minute
B	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat
C	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, flame, water
D	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, day_of_year, hour, minute
E	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, flame, water, day_of_year, hour, minute
F	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, flame, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass
G	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, flame, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass, day_of_year, hour, minute

Πίνακας 4.7: Μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε σύνολο σπιτιών

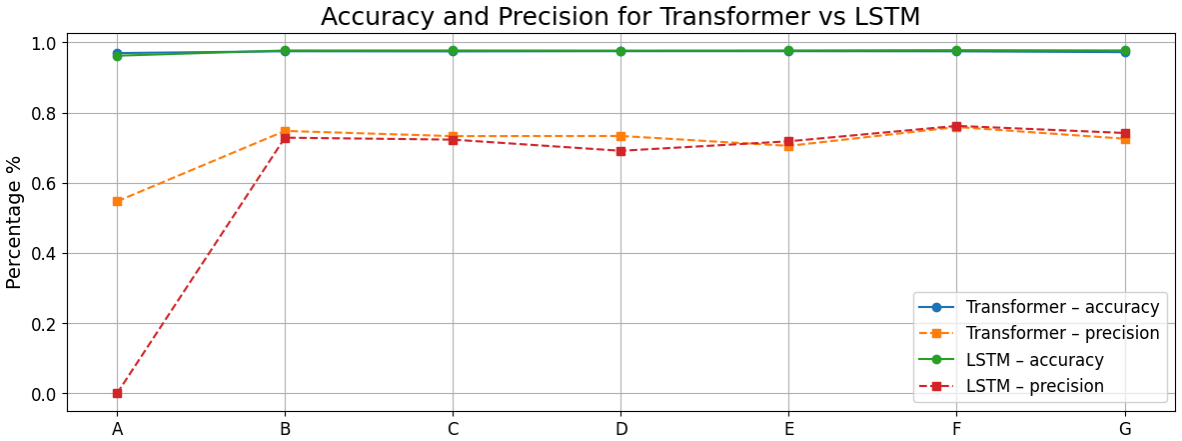
Ως προς τα σφάλματα:



Σχήμα 4.14: MAE και RMSE για μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε σύνολο σπιτιών

Παρατηρούμε ότι τα σφάλματα είναι εμφανώς μεγαλύτερα από ότι στην προηγούμενη περίπτωση όπου μελετούσαμε ένα μόνο σπίτι. Αυτό βέβαια ήταν αναμενόμενο, αφού κάθε σπίτι έχει διαφορετική συμπεριφορά και άρα τα μοντέλα δεν μπορούν να προσαρμοστούν τόσο καλά όσο στην περίπτωση του ενός σπιτιού.
Τα σφάλματα MAE είναι για όλες τις περιπτώσεις σχεδόν στο 2.5 . Από την άλλη, για το RMSE οι τιμές είναι κοντά στο 4.5 .
Μάλιστα, σε αντίθεση με την περίπτωση εκπαίδευσης σε ένα σπίτι, εδώ υπάρχουν περιπτώσεις όπου τα Transformer μοντέλα κάνουν μικρότερα σφάλματα από τα αντίστοιχα LSTM. Για το MAE ισχύει για τα A, B, C και G. Ενώ για το A ισχύει και το RMSE.

Σχετικά με τις μετρικές επίδοσης:



Σχήμα 4.15: Accuracy και Precision για μοντέλα πιθανώς βέλτιστων συνδυασμών κατηγοριών σε σύνολο σπιτιών

Το accuracy παραμένει σχεδόν σταθερά στο 1.0 για όλα τα μοντέλα. Δηλαδή η επιτυχία στις προβλέψεις είναι πρακτικά στο 100%. Για το precision οι τιμές είναι περίπου στο 75% για όλα τα μοντέλα. Αυτό σημαίνει ότι από τις προβλέψεις για ενεργοποίηση του boiler ένα πολύ μεγάλο ποσοστό είναι επιτυχείς.
Όσον αφορά τη σύγκριση με τα LSTM, έχουν επιδόσεις περίπου στα ίδια επίπεδα με τους Transformer. Εξαίρεση σε αυτό αποτελεί το μοντέλο A, για το οποίο αν και το precision

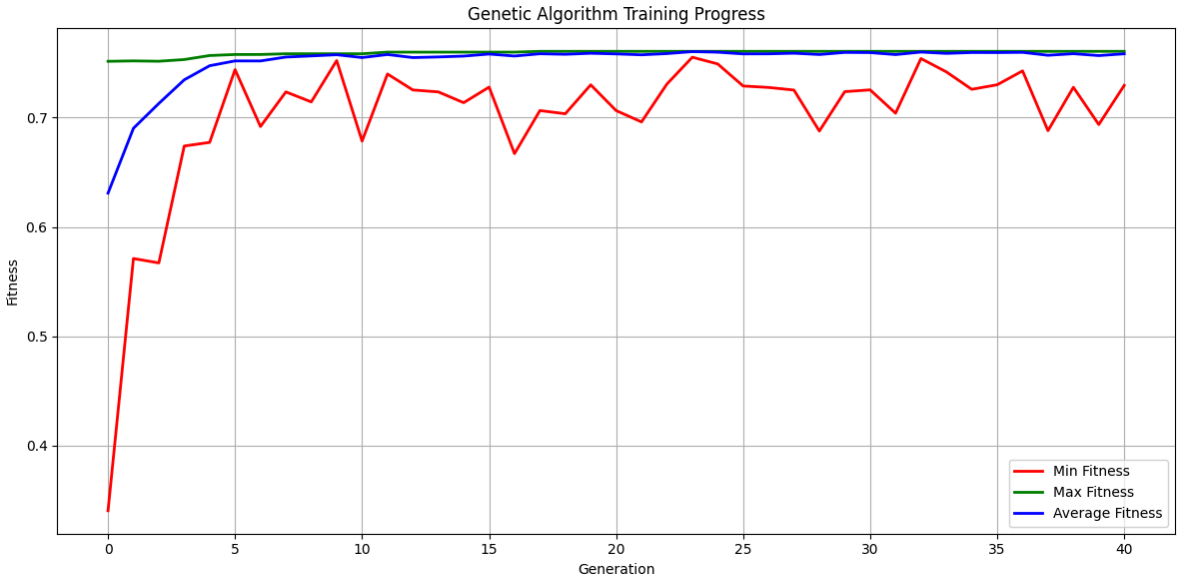
είναι χαμηλότερο του 60% (χειρότερο από τα υπόλοιπα αυτής της δοκιμής) , είναι υψηλότερο από του αντίστοιχου LSTM.

4.3.2 Γενετική Παραγωγή Μοντέλων

Σε αυτό το σημείο θα χρησιμοποιήσουμε τον γενετικό αλγόριθμο για να παράξουμε μοντέλα τα οποία θα έχουν καλή συμπεριφορά ως προς την πρόβλεψη του επιπέδου λειτουργίας του boiler σε ένα σύνολο σπιτιών (στην περίπτωσή μας 28 σπίτια).

Ο Random Forest Regressor εκπαιδεύεται και σε αυτή την περίπτωση σε όλα τα μέχρι στιγμής διαθέσιμα μοντέλα, ακόμα και στα μοντέλα του ενός σπιτιού. Ο λόγος για αυτή την επιλογή είναι ότι με την αξιοποίηση και των μοντέλων του ενός σπιτιού μπορεί να προκύψουν συσχετίσεις που δεν έχουμε διαπιστώσει μέχρι στιγμής.

Ο γενετικός αλγόριθμος κατά την εκτέλεσή του δίνει:



Σχήμα 4.16: Διαδικασία εκτέλεσης γενετικού αλγορίθμου σε περίπτωση για σύνολο σπιτιών

Παρατηρούμε ότι και εδώ τα μοντέλα συγκλίνουν σχετικά γρήγορα. Βέβαια, η εκτιμώμενη τιμή του precision εδώ εκτιμάται στο 75% . Πολύ καλή εκτιμώμενη ακρίβεια στις προβλέψεις, αν και δεν είναι τόσο καλή όσο για το 1 σπίτι.

Το καλύτερο μοντέλο από την τελευταία γενιά έχει τα ακόλουθα χαρακτηριστικά:

Εκτιμώμενο precision	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**

0.7606	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass

Πίνακας 4.8: Βέλτιστο μοντέλο σε σύνολο σπιτιών από γενετικό αλγόριθμο

4.3.3 Σύγκριση Αποτελεσμάτων για Σύνολο Σπιτιών

Τώρα θα συγκρίνουμε την ικανότητα των διαφόρων επιλεγμένων μοντέλων να προβλέψουν το επίπεδο λειτουργίας του boiler για ένα σύνολο σπιτιών (28 σπίτια). Θα αντιπαραθέσουμε τα μοντέλα που επιλέχθηκαν ως πιθανώς βέλτιστα όπως προκύπτει από τον συνδυασμό κατηγοριών της εισόδου τους με το μοντέλο που προκρίθηκε ως καλύτερο από την εκτέλεση του γενετικού αλγορίθμου.

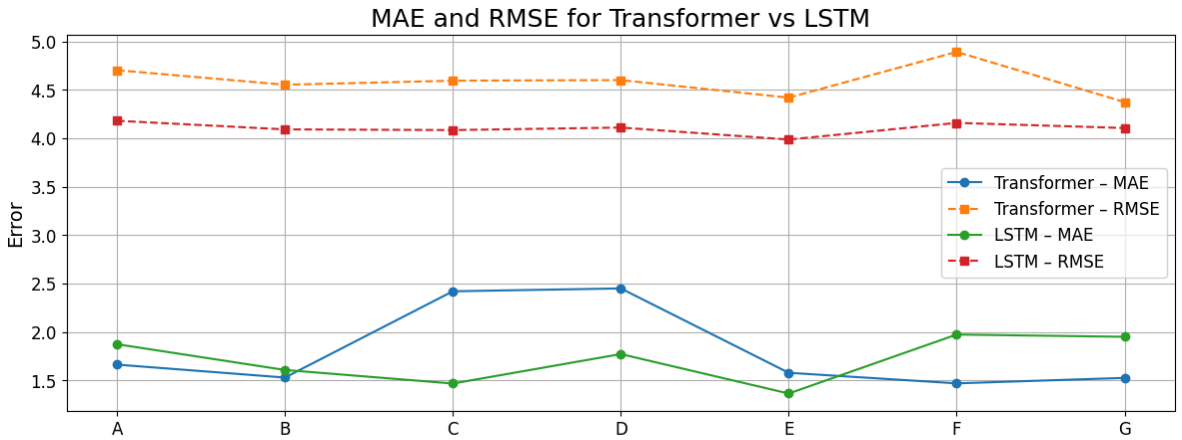
Στη συγκεκριμένη περίπτωση (σε αντίθεση με προηγουμένως που εξετάζαμε προσαρμογή σε ένα σπίτι) παρατηρούμε ότι σχεδόν όλα από τα αρχικά επιλεγμένα μοντέλα έχουν παρόμοια (και πολύ καλή) επίδοση. Αυτό μας οδηγεί στην διατήρηση πολλών μοντέλων για την τελική σύγκριση:

Μοντέλο	**Διάρκεια Πρόβλεψης**	**Κατηγορίες Εισόδου**	**Τρόπος Επιλογής**
A	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat	Αρχική επιλογή
B	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, flame, water	Αρχική επιλογή
C	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, day_of_year, hour, minute	Αρχική επιλογή
D	10	blr_mod_lvl, insulation, otc_maxt, blr_t, t_out, t_set, heat, flame, water, day_of_year, hour, minute	Αρχική επιλογή
E	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, flame, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass	Αρχική επιλογή
F	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, flame, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass, day_of_year, hour, minute	Αρχική επιλογή

G	10	blr_mod_lvl, absorption, insulation, t_r_set, t_out, blr_t, heat, water, t_ret, t_r, t_set, otc_cur, otc_maxt, bypass	Γενετικός Αλγόριθμος

Πίνακας 4.9: Συνολικά βέλτιστα μοντέλα στην περίπτωση συνόλου σπιτιών

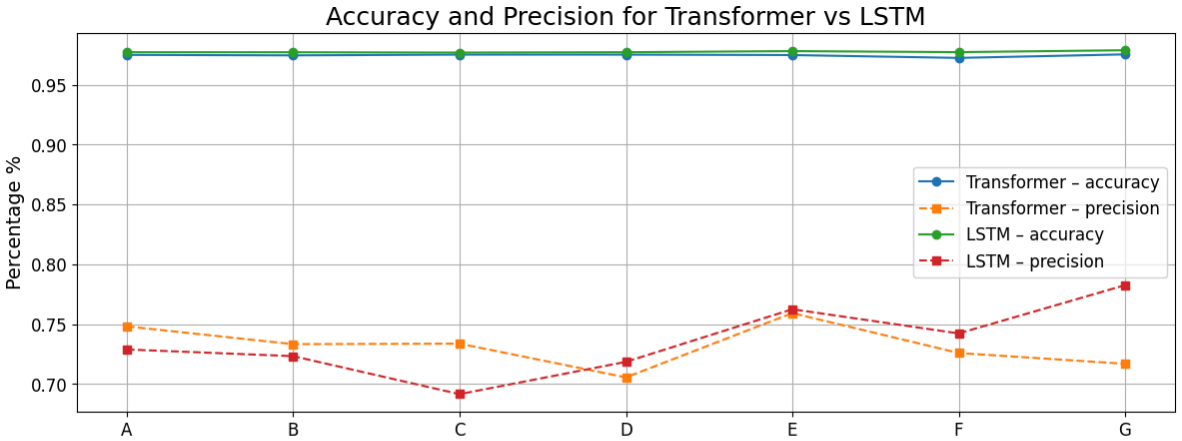
Για τις μετρικές MAE, RMSE έχουμε:



Σχήμα 4.17: MAE και RMSE για συνολικά βέλτιστα μοντέλα στην περίπτωση συνόλου σπιτιών

Παρατηρούμε ότι η τιμή του MAE είναι μικρή και κυμαίνεται εντός των ορίων 1.5 έως 2.5 . Μάλιστα, υπάρχουν μοντέλα για τα οποία η τιμή είναι μικρότερη από του αντίστοιχου LSTM. Το RMSE είναι για όλα σχεδόν στο 4.5 και μάλιστα διατηρείται σχεδόν σταθερά μεγαλύτερο από των αντίστοιχων LSTM κατά 0.5 .

Για τις μετρικές επίδοσης:



Σχήμα 4.18: Accuracy και Precision για συνολικά βέλτιστα μοντέλα στην περίπτωση συνόλου σπιτιών

Το accuracy για μία ακόμα περίπτωση σταθερά σχεδόν στο 1.0 (100%) υποδηλώνοντας ότι τα μοντέλα προβλέπουν σωστά τις περισσότερες φορές (βέβαια οι περισσότερες φορές

αντιστοιχούν στη μη ενεργοποίηση του boiler και άρα είναι εύκολο να προβλεφθούν σωστά). Το precision κυμαίνεται από 0.7 (70%) έως και 0.8 (80%).

Συμπερασματικά, για την περίπτωση της πρόβλεψης του επιπέδου λειτουργίας του boiler για ένα σύνολο σπιτιών, τα μοντέλα Transformer είναι αρκετά ακριβή. Τα σφάλματα όπως αυτά εκφράζονται από το MAE είναι γενικά σε χαμηλές τιμές. Για το RMSE, το οποίο είναι πιο ευαίσθητο σε μεγάλες τιμές σφαλμάτων, οι τιμές είναι μεγαλύτερες υποδηλώνοντας ότι συμβαίνουν λάθη στις περιπτώσεις που οι τιμές είναι μεγαλύτερες οι οποίες αντιστοιχούν σε ενεργοποιημένο boiler. Αυτό σε συνδυασμό με το γεγονός ότι το accuracy είναι σταθερά σχεδόν 1.0 (δηλαδή 100%) υποδεικνύει ότι αν και οι στιγμές μη λειτουργίας προβλέπονται σωστά, εντούτοις γίνονται λάθη στην πρόβλεψη του επιπέδου λειτουργίας όταν έχουμε ανοιχτό boiler. Η υψηλές τιμές του precision φανερώνουν περαιτέρω ότι οι στιγμές ενεργοποίησης προβλέπονται σωστά.
Επομένως, μπορούμε να συμπεράνουμε πως όταν ένα μοντέλο πρέπει να προσαρμοστεί στην πρόβλεψη για πολλά σπίτια ταυτόχρονα, αυτό μπορεί να γίνει με επιτυχία για κατάλληλο συνδυασμό κατηγοριών. Βέβαια, λαμβάνοντας υπόψη μας και τις τιμές όλων αυτών των μετρικών, εμφανίζονται κάποια σφάλματα (μεγαλύτερα από ότι για την περίπτωση ενός σπιτιού) τα οποία όμως δεν είναι επαρκώς μεγάλα ώστε να μειώσουν την επιτυχία των προβλέψεων. Εντούτοις, φανερώνουν ότι οι διαφοροποιήσεις μεταξύ των σπιτιών μπορεί να δυσκολέψουν την προσαρμογή των μοντέλων.

Όσον αφορά τη σύγκριση με τα μοντέλα LSTM φαίνεται ότι για σύνολο σπιτιών οι Transformers μπορούν να αποδώσουν καλύτερα. Δηλαδή, αν και οι τιμές των μετρικών δεν απέχουν πολύ μεταξύ τους, είναι φανερό ότι η πολυπλοκότητα του συνόλου των δεδομένων δυσκολεύει κάπως περισσότερο τα LSTMs.

Σχετικά με το μοντέλο G το οποίο προέκυψε από την υλοποίηση του γενετικού αλγορίθμου, μπορούμε να παρατηρήσουμε ότι προβλέπει με εξίσου καλή ακρίβεια όσο και τα υπόλοιπα μοντέλα. Αξίζει μάλιστα να σημειώσουμε ότι μαζί με τα μοντέλα A, B, F έχουν τιμές MAE μικρότερες από τα αντίστοιχα LSTM.

5 Συμπεράσματα και Μελλοντική Έρευνα

5.1 Συμπεράσματα

Στην παρούσα διπλωματική εργασία εστιάσαμε στην αξιοποίηση μοντέλων τύπου Transformer για την πρόβλεψη του επιπέδου λειτουργίας του boiler. Ταυτόχρονα χρησιμοποιήθηκαν μοντέλα LSTM για τις ίδιες λειτουργίες, ώστε τα αποτελέσματά τους να αξιοποιηθούν ως ένα μέτρο σύγκρισης. Μάλιστα, επικεντρωθήκαμε σε δύο πτυχές του προβλήματος. Από τη μία, στην παραγωγή μοντέλων που εστιάζουν στην πρόβλεψη της συμπεριφοράς ενός μόνο boiler (σε ένα μόνο σπίτι). Από άλλη, στην ανάπτυξη μοντέλων που θα έχουν καλή προβλεπτική ικανότητα για ένα σύνολο σπιτιών. Σε κάθε περίπτωση η προβλεπτική ικανότητα των μοντέλων διαφοροποιείται ανάλογα με τις κατηγορίες των δεδομένων οι οποίες αξιοποιούνται για την παραγωγή της πρόβλεψης.

Ένα γενικό συμπέρασμα το οποίο προκύπτει και από τα δύο σκέλη της μελέτης μας είναι ότι για να μπορέσει ένα μοντέλο να δώσει όσο το δυνατόν πιο σωστές προβλέψεις, πρέπει τα δεδομένα εισόδου να απαρτίζονται από σύνολο κατηγοριών ικανό να περιγράψει τις συνθήκες λειτουργίας του boiler. Είναι χαρακτηριστικό πως για συνδυασμούς πολλών κατηγοριών οι προβλέψεις είναι εμφανώς πιο ακριβείς.
Ένα δεύτερο γενικό συμπέρασμα είναι ότι τα μοντέλα LSTM μπορούν να προβλέπουν με τόσο καλή επιτυχία όσο τα αντίστοιχα μοντέλα Transformer. Σε πολλές περιπτώσεις οι προβλέψεις των LSTM είναι περισσότερο ακριβείς, ενώ ακόμα και όταν υπερισχύουν οι Transformers, η διαφορά όπως αποτυπώνεται στις μετρικές σφαλμάτων είναι αρκετά μικρή.

Όσον αφορά την πρόβλεψη σε ένα σπίτι, διαπιστώνουμε ότι οι Transformers μπορούν να "μάθουν" τη συμπεριφορά του boiler και να προβλέψουν με αρκετή επιτυχία με περίπου 10 κατηγορίες δεδομένων ως είσοδο. Ο ακριβής αριθμός των κατηγοριών εξαρτάται από το ποιές χρησιμοποιούνται σε κάθε παραλλαγή. Μάλιστα, υπερβολική αύξηση του αριθμού των κατηγοριών μπορεί να οδηγήσει σε μείωση της ακρίβειας των προβλέψεων. Τα βέλτιστα μοντέλα έχουν γενικά μικρές τιμές σφαλμάτων στην πρόβλεψη του επιπέδου λειτουργίας και μπορούν να προσδιορίσουν με μεγάλη ακρίβεια τις στιγμές κατά τις οποίες είναι ενεργοποιημένο το boiler.

Από την άλλη, οι Transformers οι οποίοι εκπαιδεύονται σε ένα σύνολο σύνολο σπιτιών, χρειάζονται περισσότερες κατηγορίες δεδομένων ως είσοδο. Επίσης, σε αυτή την περίπτωση ο ακριβής αριθμός εξαρτάται από το ποιές κατηγορίες έχουν κάθε φορά επιλεχθεί. Γενικά, τα σφάλματα και σε αυτή την περίπτωση είναι μικρά και η ακρίβεια των προβλέψεων για την ενεργοποίηση του boiler παραμένει υψηλή. Ωστόσο, σε σύγκριση με την περίπτωση πρόβλεψης για ένα σπίτι, τα σφάλματα είναι σχετικά μεγαλύτερα, αν και για τα καλύτερα μοντέλα οι τιμές των μετρικών πλησιάζουν σε αυτές της προηγούμενης περίπτωσης. Βέβαια, τέτοιες διαφοροποιήσεις ήταν αναμενόμενες εξ' αρχής, αφού για σύνολο σπιτιών τα μοντέλα πρέπει να "μάθουν" τη συμπεριφορά πολλών διαφορετικών boiler, κάποια από τα οποία είναι παρόμοια και κάποια που διαφέρουν.

Τέλος αξίζει να σημειώσουμε ότι επιλέγοντας συνδυασμούς των κατηγοριών που φαίνεται να καλύπτουν το σύνολο των ιδιαιτεροτήτων του προβλήματος, καταλήγουμε σε μοντέλα που αποδίδουν περίπου τόσο καλά όσο και αυτά που επιλέγονται με χρήση γενετικού

αλγορίθμου. Μάλιστα, και με τους δύο τρόπους καταλήγουμε σε μοντέλα που λαμβάνουν ως είσοδο παρόμοιους συνδυασμούς κατηγοριών δεδομένων. Από αυτά συμπεραίνουμε, ότι στο πρόβλημα για την πρόβλεψη του επιπέδου λειτουργίας του boiler δεν υπάρχουν "κρυφές" πτυχές ή εξαρτήσεις μεγεθών. Συνεπώς, μπορούμε σχετικά εύκολα να κατανοήσουμε γιατί οι επιλεγμένες σε κάθε περίπτωση κατηγορίες είναι απαραίτητες για την παραγωγή βέλτιστης πρόβλεψης.

5.2 Μελλοντική Έρευνα

Στην προσπάθεια για βελτιστοποίηση της λειτουργίας των οικιακών boiler φυσικού αερίου, η πρόβλεψη του επιπέδου λειτουργίας είναι απαραίτητη. Στηριζόμενοι στα συμπεράσματα αυτής της εργασίας είναι εμφανές ότι υπάρχουν αρκετές πτυχές του προβλήματος οι οποίες μπορούν να βελτιστοποιηθούν και να επεκταθούν περαιτέρω. Κατ' αρχάς μπορούν να αξιοποιηθούν διαφορετικές αρχιτεκτονικές μοντέλων για την παραγωγή των προβλέψεων. Συγκεκριμένα, αναφερόμαστε στις πολλές και διαφορετικές εκδοχές των Transformer καθώς επίσης και στα διαφορετικά είδη νευρωνικών δικτύων.
Μία άλλη κατεύθυνση που θα μπορούσε να ακολουθηθεί, είναι η επέκταση του χρονικού ορίζοντα πρόβλεψης. Για τα μοντέλα που μελετήσαμε, τόσο τους Transformers όσο και τα LSTMs, ιδανική διάρκεια ήταν τα 10 λεπτά. Ωστόσο, θα είχε ενδιαφέρον η χρήση άλλων ειδών μοντέλων ή υλοποιήσεων που θα επέτρεπαν την επέκταση αυτής της διάρκειας σε διάστημα ωρών ή ακόμα και ημερών, δίνοντας έτσι τη δυνατότητα για ακόμα μεγαλύτερη ευελιξία στην βελτιστοποίηση των boiler.
Μία άλλη εναλλακτική αποτελεί ο συνδυασμός στατιστικών μεθόδων πρόβλεψης χρονοσειρών με μοντέλα μηχανικής μάθησης. Αντί απλώς να εισαχθούν τα δεδομένα σε ένα νευρωνικό δίκτυο, όπως κάναμε στην παρούσα εργασία, θα μπορούσαν πρώτα να εξαχθούν χρήσιμες πληροφορίες για τη χρονοσειρά, υλοποιώντας κατάλληλη αρχική επεξεργασία, και έπειτα να προβλεφθεί η μελλοντική συμπεριφορά των boiler με κάποιου είδους νευρωνικό δίκτυο.

6 Βιβλιογραφία

[1] G. I. Tsoumalis, Z. N. Bampos, G. V. Chatzis, P. N. Biskas, και S. D. Keranidis, ‘Minimization of natural gas consumption of domestic boilers with convolutional, long-short term memory neural networks and genetic algorithm’, *Appl. Energy*, τ. 299, σελ. 117256, 2021, doi: https://doi.org/10.1016/j.apenergy.2021.117256.

[2] M. C. Barma, R. Saidur, S. M. A. Rahman, A. Allouhi, B. A. Akash, και S. M. Sait, ‘A review on boilers energy use, energy savings, and emissions reductions’, *Renew. Sustain. Energy Rev.*, τ. 79, σελ. 970–983, 2017, doi: https://doi.org/10.1016/j.rser.2017.05.187.

[3] V. Michalakopoulos *κ.ά.*, ‘A multi-dimensional unsupervised machine learning framework for clustering residential heat load profiles’. arXiv.org, 2024. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://arxiv.org/abs/2411.16701

[4] E. Commission, ‘Climate strategies & targets’. climate.ec.europa.eu, 2025. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://climate.ec.europa.eu/eu-action/climate-strategies-targets_en

[5] ‘In focus: Energy efficiency in buildings’. European Commission, Φεβρουάριος 2020. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://commission.europa.eu/news-and-media/news/focus-energy-efficiency-buildings-2020-02-17_en

[6] A. Özmen, Y. Yılmaz, και G.-W. Weber, ‘Natural gas consumption forecast with MARS and CMARS models for residential users’, *Energy Econ.*, τ. 70, σελ. 357–381, 2018, doi: https://doi.org/10.1016/j.eneco.2018.01.022.

[7] E. Saloux, J. Runge, και K. Zhang, ‘Operation optimization of multi-boiler district heating systems using artificial intelligence-based model predictive control: Field demonstrations’, *Energy*, τ. 285, σελ. 129524, 2023, doi: https://doi.org/10.1016/j.energy.2023.129524.

[8] F. Mateo, J. J. Carrasco, A. Sellami, M. Millán-Giraldo, M. Domínguez, και E. Soria-Olivas, ‘Machine learning methods to forecast temperature in buildings’, *Expert Syst. Appl.*, τ. 40, τχ. 4, σελ. 1061–1068, 2013, doi: https://doi.org/10.1016/j.eswa.2012.08.030.

[9] Z. Afroz, T. Urmee, G. M. Shafiullah, και G. Higgins, ‘Real-time prediction model for indoor temperature in a commercial building’, *Appl. Energy*, τ. 231, σελ. 29–53, 2018, doi: https://doi.org/10.1016/j.apenergy.2018.09.052.

[10] X. Shi, W. Lu, Y. Zhao, και P. Qin, ‘Prediction of Indoor Temperature and Relative Humidity Based on Cloud Database by Using an Improved BP Neural Network in Chongqing’, *IEEE Access*, τ. 6, σελ. 30559–30566, 2018, doi: 10.1109/ACCESS.2018.2844299.

[11] P. Monteiro, M. Zanin, E. Menasalvas Ruiz, J. Pimentão, και P. Alexandre da Costa Sousa, ‘Indoor Temperature Prediction in an IoT Scenario’, *Sensors*, τ. 18, σελ. 3610, Οκτωβρίου 2018, doi: 10.3390/s18113610.

[12] Y. LeCun, Y. Bengio, και G. Hinton, ‘Deep learning’, *Nature*, τ. 521, τχ. 7553, σελ. 436–444, Μαΐου 2015, doi: 10.1038/nature14539.

[13] A. Krenker, J. Bešter, και A. Kos, ‘Introduction to the Artificial Neural Networks’. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://cdn.intechopen.com/pdfs/14881/InTech-Introduction_to_the_artificial_neural_networks.pdf

[14] A. K. Jain, J. Mao, και K. M. Mohiuddin, ‘Artificial neural networks: a tutorial’, *Computer*, τ. 29, τχ. 3, σελ. 31–44, 1996, doi: 10.1109/2.485891.

[15] L. Hardesty, ‘Explained: Neural networks’. Massachusetts Institute of Technology, Απρίλιος 2017. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414
[16] Z. C. Lipton, J. Berkowitz, και C. Elkan, ‘A Critical Review of Recurrent Neural Networks for Sequence Learning’. arXiv.org, 2015. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://arxiv.org/abs/1506.00019
[17] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, και S. Valaee, ‘Recent Advances in Recurrent Neural Networks’, *ArXiv180101078 Cs*, Φεβρουαρίου 2018, [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://arxiv.org/abs/1801.01078
[18] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, και J. Schmidhuber, ‘LSTM: A Search Space Odyssey’, *IEEE Trans. Neural Netw. Learn. Syst.*, τ. 28, τχ. 10, σελ. 2222–2232, 2017, doi: 10.1109/TNNLS.2016.2582924.
[19] F. A. Gers, J. Schmidhuber, και F. Cummins, ‘Learning to Forget: Continual Prediction with LSTM’, *Neural Comput.*, τ. 12, σελ. 2451–2471, Οκτωβρίου 2000, doi: 10.1162/089976600300015015.
[20] Z. Ali, ‘A Comprehensive Overview and Comparative Analysis of CNN, RNN-LSTM and Transformer’, Ιανουαρίου 2025, doi: 10.2139/ssrn.5175090.
[21] A. Vaswani *κ.ά.*, ‘Attention Is All You Need’. Ιούνιος 2017. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://arxiv.org/pdf/1706.03762
[22] A. Lambora, K. Gupta, και K. Chopra, ‘Genetic Algorithm- A Literature Review’, *2019 Int. Conf. Mach. Learn. Big Data Cloud Parallel Comput. Com.*, σελ. 380–384, Φεβρουαρίου 2019, doi: 10.1109/comitcon.2019.8862255.
[23] L. Haldurai, T. Madhubala, και R. Rajalakshmi, ‘A Study on Genetic Algorithm and its Applications’, *Int. J. Comput. Sci. Eng.*, τ. 4, τχ. 10, σελ. 139–143, 2016.
[24] M. Abadi *κ.ά.*, ‘TensorFlow: A System for Large-Scale Machine Learning This paper is included in the Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI ’16). TensorFlow: A system for large-scale machine learning’. Νοέμβριος 2016. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf
[25] TensorFlow, ‘Keras | TensorFlow Core | TensorFlow’. TensorFlow, 2019. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://www.tensorflow.org/guide/keras
[26] ‘Package overview — pandas 1.0.3 documentation’. pandas.pydata.org. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://pandas.pydata.org/pandas-docs/stable/getting_started/overview.html
[27] NumPy, ‘What is NumPy? — NumPy v1.19 Manual’. numpy.org, 2022. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://numpy.org/doc/stable/user/whatisnumpy.html
[28] ‘History — Matplotlib 3.10.0 documentation’. Matplotlib.org, 2015. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: https://matplotlib.org/stable/project/history.html
[29] C. Willmott και K. Matsuura, ‘Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance’, *Clim. Res.*, τ. 30, σελ. 79–82, 2005.
[30] A. C. Cameron και F. A. G. Windmeijer, ‘An R-squared measure of goodness of fit for some common nonlinear regression models’, *J. Econom.*, τ. 77, τχ. 2, σελ. 329–342, 1997, doi: https://doi.org/10.1016/S0304-4076(96)01818-0.
[31] A. Tharwat, ‘Classification assessment methods’, *Appl. Comput. Inform.*, τ. 17, Αυγούστου 2018, doi: 10.1016/j.aci.2018.08.003.