

# Η απόδοση των μαθητών λαμβάνοντας υπόψη παράγοντες της καθημερινότητας τους

Ιωάννου Νικόλας - ge20718

Μπέκος Θωδωρής - ge

Δεκέμβριος 2024



Σχολή εφαρμοσμένων μαθηματικών και φυσικών επιστημών

Μάθημα: ΘΕΜΑ

Υπεύθυνος καθηγητής: Δρ. Στεφανέας Πέτρος

## **Ευχαριστίες:**

Θα θέλαμε να ευχαριστήσουμε τον κύριο Πέτρο Στεφανέα για το ενδιαφέρον και την εμπιστοσύνη που μας έδειξε κατά τη διάρκεια πραγματοποίησης της εν λόγω εργασίας. Επίσης, θα θέλαμε να ευχαριστήσουμε τις οικογένειες μας γιατί χωρίς αυτές δεν θα μπορούσαμε να βρισκόμασταν στη θέση που βρισκόμαστε τώρα.

# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>4</b>
<b>2</b>	<b>Ανάλυση της διαδικασίας επιλογής μοντέλου</b>	<b>6</b>
2.1	Ανάλυση του προβλήματος . . . . .	6
2.2	Επεξεργασία των δεδομένων . . . . .	8
2.3	Επιλογή χαρακτηριστικών . . . . .	11
2.4	Επιλογή υπερπαραμέτρων . . . . .	14
2.5	Επιλογή μοντέλου . . . . .	14
<b>3</b>	<b>Το μοντέλο παλινδρόμησης Lasso</b>	<b>17</b>
3.1	Έγκριση ειδικού πλαισίου χωροταξικού σχεδιασμού . . .	17
3.2	Καθορισμός περιοχών αιολικής προτεραιότητας . . . . .	17
3.3	Παράμετροι που καθορίζουν τις αποστάσεις . . . . .	17
<b>4</b>	<b>Σύγκριση με έρευνες</b>	<b>18</b>
<b>5</b>	<b>Βιβλιογραφία</b>	<b>19</b>

**1   Εισαγωγή**

km

# Introduction

gt

## 2 Ανάλυση της διαδικασίας επιλογής μοντέλου

### 2.1 Ανάλυση του προβλήματος

Το πρόβλημα το οποίο αντιμετωπίσαμε έχει να κάνει με την δημιουργία ενός μοντέλου μηχανικής μάθησης για την εκτίμηση των βαθμών που θα επιτύχουν οι μαθητές στην εξέταση ενός μαθήματος , μέσω διαφόρων παραγόντων που θα αναλύσουμε παρακάτω. Το πρόβλημα αυτό αποτελεί πρόβλημα παλινδρόμησης αφού ο τελικός βαθμός δεν αποτελεί κατηγορική αλλά ποσοτική μεταβλητή και παίρνει συνεχείς τιμές. Χρησιμοποιήσαμε το σύνολο δεδομένων Student Performance Factors που βρήκαμε στο Kaggle, όπου τα χαρακτηριστικά τα οποία το αποτελούν φαίνονται στον πίνακα παρακάτω:

Χαρακτηριστικά	Περιγραφή	Εύρος τιμών
Attendance	Ποσοστό μαθημάτων που παρευρέθηκε ο μαθητής	0-100
Previous Scores	Μέσος όρος βαθμών απο προηγούμενες εξετάσεις	0-100
Sleep Hours	Ώρες ύπνου ανά μέσο όρο κάθε βράδυ	4-10
Hours Studied	Αριθμός ωρών που διάβασε ο μαθητής μέσα στην βδομάδα	0-50
Tutoring Sessions	Αριθμός φροντιστηριακών μαθημάτων ανά μήνα	0-10
Family income	Οικογενειακό εισόδημα	Low/Medium/High
Teacher quality	Εκπαιδευτική ικανότητα καθηγητή	Low/Medium/High
Parental Involvement	Βαθμός εμπλοκής γονέων στην εκπαίδευση του μαθητή	Low/Medium/High

Access to Resources	Πρόσβαση μαθητή σε εκπαιδευτικό υλικό	Low/Medium/High
Motivation Level	Κίνητρο μαθητή	Low/Medium/High
Peer Influence	Επιρροή που δέχεται ο μαθητής απο ομήλικους του	Negative/Neutral/Positive
Distance from home	Απόσταση σχολείου απο το σπίτι	Near/Moderate/Far
School type	Τύπος σχολείου	Private/Public
Physical activity	Επίπεδο φυσικής δραστηριότητας μαθητή	Sedentary/Light/Moderate/Active/Very Active/Highly Active/Athlete
Parental education level	Επίπεδο εκπαίδευσης γονέα	High School/College/Postgraduate
Extracurricular activities	Εξωσχολικές δραστηριότητες	Yes/No
Learning disabilities	Μαθησιακές δυσκολίες	Yes/No
Internet Access	Πρόσβαση στο διαδίκτυο	Yes/No
Gender	Φύλο μαθητή	Male/Female
Exam Score	Βαθμός εξέτασης	0-100

Η μεταβλητή Exam Score αποτελεί την εξαρτημένη μεταβλητή(μεταβλητή στόχο) του μοντέλου ενώ οι υπόλοιπες μεταβλητές αποτελούν τις πιθανές ανεξάρτητες μεταβλητές.

## 2.2 Επεξεργασία των δεδομένων

Ρίχνοντας μια ματιά στα δεδομένα παρατηρήσαμε πως υπάρχουν παρατηρήσεις με ελλειπείς τιμές. Στην συγκεκριμένη περίπτωση έχουμε 3 επιλογές. Η πρώτη επιλογή ήταν να διαγράψουμε τις παρατηρήσεις αυτές, η δεύτερη να προβλέψουμε τις τιμές των παρατηρήσεων με κάποια μέθοδο (Παλινδρόμηση για τις ποσοτικές και ομαδοποίηση για της κατηγορικές μεταβλητές) και η τρίτη να αντικαταστήσουμε τις ελλειπείς τιμές με κάποιο λογικό τρόπο(π.χ. εάν έχουμε ποσοτικές τιμές με την μέση τιμή του χαρακτηριστικού ή εάν έχουμε κατηγορική με την τιμή, η οποία εμφανίζεται τις περισσότερες φορές). Επιλέξαμε να διαγράψουμε τις παρατηρήσεις οι οποίες έχουν ελλειπείς τιμές μιας και ο αριθμός των παρατηρήσεων είναι ικανοποιητικός και οι ελλειπείς τιμές αποτελούν ένα πολύ μικρό ποσοστό(<10%) των συνολικών παρατηρήσεων και δεν θα επιρεάσουν το αποτέλεσμα.

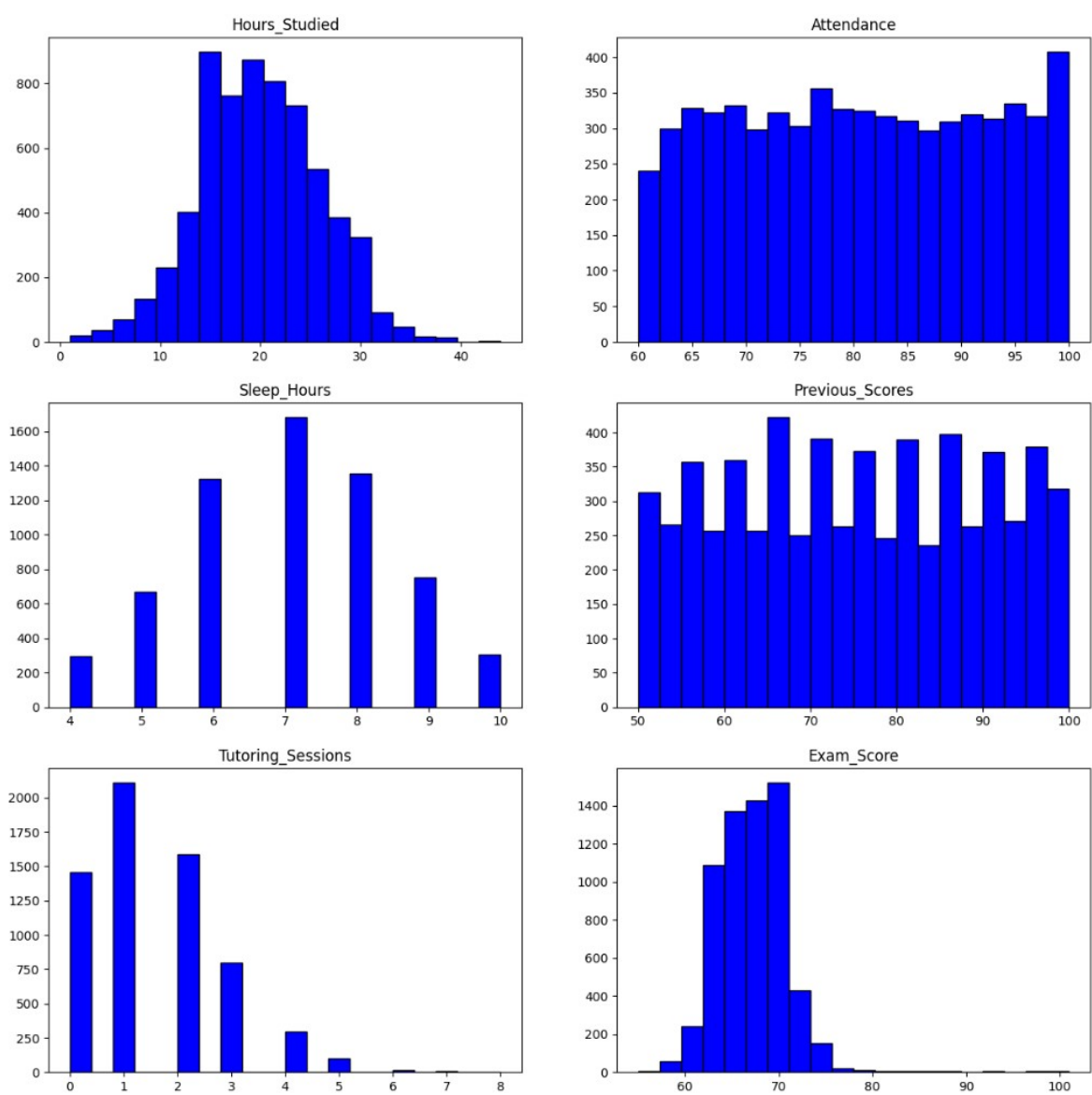
Η μεταβλητή Physical Activity έχει μια ιδιαιτερότητα. Παρόλο που περιλαμβάνει αριθμητικές τιμές(0-6), αυτές οι τιμές αποτελούν κατηγορίες. Για αυτό και εμείς αλλάξαμε τον τύπο της μεταβλητής σε αντικείμενο. Εύκολα μπορούμε να δούμε πως το dataset μας έχει 14 κατηγορικές μεταβλητές και 6 ποσοτικές.

Παρατηρώντας τα δεδομένα μπορούμε εύκολα να δούμε πως οι ποσοτικές μεταβλητές κυμαίνονται σε διαφορετικές κλίμακες. Όταν χρησιμοποιούμε αλγόριθμους μηχανικής μάθησης που βασίζονται σε αποστάσεις (όπως οι K-Neighbors Classifier, K-means, linear regression, polynomial regression κ.λπ.), είναι απαραίτητο να εφαρμόζουμε κάποια μορφή κλιμάκωσης στα δεδομένα μας. Η κλιμάκωση εξασφαλίζει ότι τα χαρακτηριστικά με μεγάλο εύρος τιμών δεν θα κυριαρχούν έναντι εκείνων με μικρό εύρος, αποτρέποντας έτσι την εισαγωγή σφαλμάτων στο μοντέλο. Ιδανικά,θέλουμε όλα τα αριθμητικά χαρακτηριστικά να έχουν την ίδια κλίμακα, έτσι ώστε να συνεισφέρουν ισότιμα στη διαδικασία εκπαίδευσης του μοντέλου. Η επιλογή της κατάλληλης κλιμάκωσης στα δεδομένα εξαρτάται από τα δομή του συνόλου δεδομένων. Για χαρακτηριστικά τα οποία προσεγγίζουν την κανονική κατανομή χρησιμοποιούμε Standard Scaler, για χαρακτηριστικά τα οποία περιέχουν ακραίες τιμές(outliers) χρησιμοποιούμε Robust Scaler



και για χαρακτηριστικά τα οποία προσεγγίζουν την ομοιόμορφη κατανομή τον MinMax Scaler.

Μπορούμε να εφαρμόσουμε κλιμάκωση ακόμα και αν οι ποσοτικές τιμές αποτελούν ακέραιες τιμές. Οι μέθοδοι κλιμάκωσης, όπως StandardScaler, MinMaxScaler και RobustScaler, λειτουργούν ανεξάρτητα από τον τύπο δεδομένων, εφόσον οι μεταβλητές είναι αριθμητικές και για κάθε τύπο κλιμάκωσης ισχύουν οι αντίστοιχες προϋποθέσεις. Παρακάτω δίνονται τα ιστογράμματα όλων των ποσοτικών μεταβλητών.



Εύκολα μπορούμε να παρατηρήσουμε πως εφαρμόζουμε Standard Scaler στα χαρακτηριστικά Hours Studied και Sleep Hours αφού όπως μπορούμε να δούμε ακολουθούν κατά προσέγγιση κανονική κατανομή, MinMax Scaler στα χαρακτηριστικά Attendance και Previous Scores αφού προσεγγίζουν την ομοιόμορφη κατανομή και Robust Scaler στο χαρακτηριστικό Tutoring Sessions αφού όπως μπορούμε να δούμε περιέχει ακραίες τιμές. Η μεταβλητή στόχος Exam Scores δεν χρειάζεται να κλιμακωθεί αφού αποτελεί την εξαρτημένη μεταβλητή του μοντέλου μας.

Οι σχέσεις για κάθε τύπο κλιμάκωσης δίνονται παρακάτω:

## Μέθοδοι Κλιμάκωσης

### StandardScaler

Η μετατροπή με τη μέθοδο StandardScaler ορίζεται ως:

$$X' = \frac{X - \mu}{\sigma}$$

όπου:

- $X$ : Η τιμή της παρατήρησης στο χαρακτηριστικό.
- $\mu$ : Ο μέσος όρος του χαρακτηριστικού.
- $\sigma$ : Η τυπική απόκλιση του χαρακτηριστικού.

### MinMaxScaler

Η μετατροπή με τη μέθοδο MinMaxScaler ορίζεται ως:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

όπου:

- $X$ : Η τιμή της παρατήρησης στο χαρακτηριστικό.
- $X_{\min}$ : Η ελάχιστη τιμή του χαρακτηριστικού.
- $X_{\max}$ : Η μέγιστη τιμή του χαρακτηριστικού.

### RobustScaler

Η μετατροπή με τη μέθοδο RobustScaler ορίζεται ως:

$$X' = \frac{X - Q_2}{Q_3 - Q_1}$$

όπου:

- $X$ : Η τιμή της παρατήρησης στο χαρακτηριστικό.
- $Q_2$ : Η διάμεσος (median) του χαρακτηριστικού.

- $Q_1$ : Το πρώτο τεταρτημόριο (25η εκατοστιαία θέση).
- $Q_3$ : Το τρίτο τεταρτημόριο (75η εκατοστιαία θέση).

Γνωρίζουμε πως τα μοντέλα μηχανικής μάθησης δεν μπορούν να διαχειριστούν μη αριθμητικά δεδομένα. Επομένως, μέσω των encoders θα τα μετατρέψουμε με κατάλληλο τρόπο σε αριθμητικά χωρίς όμως να αλλάζουμε την κατηγορική τους ιδιότητα. Θα χρησιμοποιήσουμε Ordinal encoder για τα χαρακτηριστικά τα οποία περιέχουν κατηγορίες, οι οποίες έχουν κάποια φυσική σειρά(π.χ. "Low", "Medium", "High") και OneHot encoder για τα χαρακτηριστικά τα οποία περιέχουν κατηγορίες, οι οποίες δεν έχουν κάποια φυσική σειρά(π.χ. "Yes", "No"). Επομένως τα κατηγορικά χαρακτηριστικά "Parental Involvement", "Access to Resources", "Motivation Level", "Family Income", "Teacher Quality", "Parental Education Level", "Distance from Home" και "Physical Activity" κωδικοποιούνται με ordinal encoding ενώ τα υπόλοιπα κατηγορικά χαρακτηριστικά με Onehot.

Η μέθοδος ordinal encoding κωδικοποιεί με αριθμητικές τιμές κάθε κατηγορία με βάση την σειρά ή την θέση τους ξεκινώντας από το 0. Για παράδειγμα εάν έχουμε τις κατηγορίες "Low", "Medium", "High" θα τους αναθέσει τις τιμές 0,1,2 αντίστοιχα.

Η μέθοδος onehot encoding διαγράφει την κατηγορική στήλη και δημιουργεί για κάθε κατηγορία μια νέα που για κάθε παρατήρηση περιέχει την τιμή 1 αν το δείγμα ανήκει σε αυτήν την κατηγορία και 0 διαφορετικά.

### 2.3 Επιλογή χαρακτηριστικών

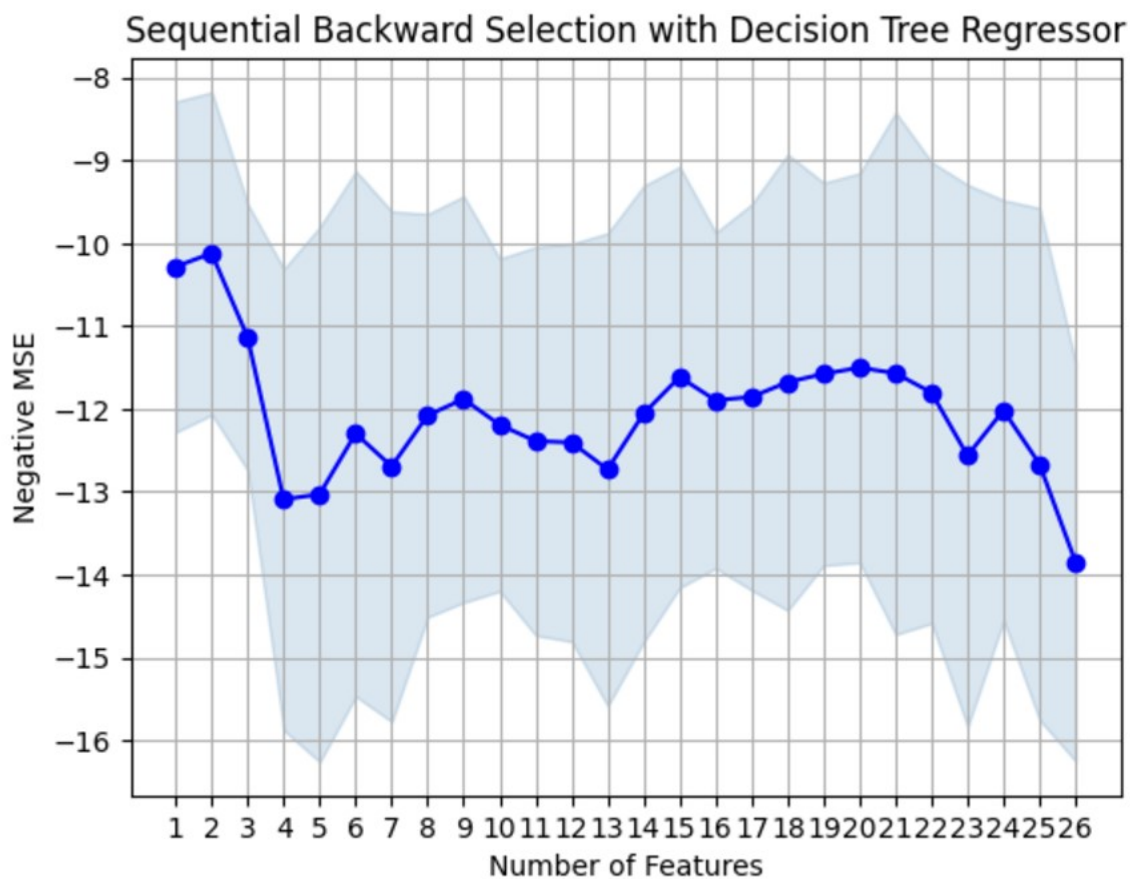
Η επιλογή των χαρακτηριστικών που θα χρησιμοποιήσουμε για την πρόβλεψη του στόχου αποτελεί σημαντικό βήμα. Από την μια προσπαθούμε να επιλέξουμε όλα τα σημαντικά χαρακτηριστικά έτσι ώστε να αποφύγουμε την υποπροσαρμογή, δηλαδή το μοντέλο μας να μην έχει τις απαραίτητες πληροφορίες για να αποτυπώσει την σχέση μεταξύ των δεδομένων και της μεταβλητής στόχου και από την άλλη προσπαθούμε να επιλέξουμε όσο το δυνατό λιγότερα μη σημαντικά χαρακτηριστικά έτσι ώστε να αποφύγουμε την υπερπροσαρμογή, δηλαδή το μοντέλο μας να αποδίδει πολύ καλά στα

δεδομένα εκπαίδευσης αλλά να αδυνατεί να γενικεύσει σε νέα δεδομένα. Αρχικά θα ορίσουμε 8 στο σύνολο μοντέλα παλινδρόμησης για τα οποία θα εξετάσουμε την απόδοση τους μέσω μιας μετρικής για κάθε συνδυασμό χαρακτηριστικών. Τα μοντέλα παλινδρόμησης τα οποία θα εξετάσουμε είναι τα εξής:

1. Πολυωνιμικό μοντέλο παλινδρόμησης
2. Γραμμική παλινδρόμηση
3. Decision Tree παλινδρόμηση
4. Random Forest παλινδρόμηση
5. Support Vector παλινδρόμηση
6. Gradient Boosting παλινδρόμηση
7. Παλινδρόμηση Ridge
8. Παλινδρόμηση Lasso

Για την πραγματοποίηση της παραπάνω διαδικασίας θα χρησιμοποιήσουμε τον αλγόριθμο Sequential feature selection (SFS). Θα μπορούσαμε να κάνουμε την διαδικασία αυτή και μέσω των πινάκων συσχέτισης Pearson, Cramer's V και ANOVA διαγράφοντας για κάθε ζεύγος μεταβλητών με ισχυρή συσχέτιση την μια από τις δύο μεταβλητές αλλά με την διαφορά πως δεν θα είχαμε τόσο καλά αποτελέσματα όσο με την μέθοδο SFS. Η μέθοδος SFS με επιλογή για backwards selection ξεκινά με το πλήρες σύνολο χαρακτηριστικών. Το σύνολο εκπαίδευσης χωρίζεται σε  $k$ =cross-validation ίσα μέρη(folds). Σε κάθε επανάληψη, αφαιρείται ένα χαρακτηριστικό από το τρέχον σύνολο και το μοντέλο εκπαιδεύεται χρησιμοποιώντας τα εναπομείναντα χαρακτηριστικά. Το μοντέλο εκπαιδεύεται  $k$  φορές, χρησιμοποιώντας κάθε φορά ένα διαφορετικό fold ως σύνολο δοκιμής (test set), ενώ τα υπόλοιπα  $k-1$  folds χρησιμοποιούνται ως σύνολο εκπαίδευσης (training set). Η απόδοση για κάθε αφαίρεση χαρακτηριστικού αξιολογείται με βάση τον μέσο όρο της απόδοσης για κάθε επιλογή fold ως test set χρησιμοποιώντας κάποιο προκαθορισμένο κριτήριο, όπως το μέσο τετραγωνικό

σφάλμα (MSE). Το χαρακτηριστικό που οδηγεί στη μικρότερη μείωση της απόδοσης αφαιρείται οριστικά από το σύνολο. Η διαδικασία συνεχίζεται μέχρι να επιτευχθεί ο επιθυμητός αριθμός χαρακτηριστικών. Για κάθε ένα από τα μοντέλα παλινδρόμησης που αναφέραμε παραπάνω εντοπίσαμε τα χαρακτηριστικά τα οποία οδηγούν στο μικρότερο μέσο τετραγωνικό σφάλμα. Για την καλύτερη κατανόηση της μεθόδου, παρουσιάζεται παρακάτω η γραφική παράσταση του βέλτιστου αρνητικού μέσου τετραγωνικού σφάλματος. Η γραφική αφορά το μοντέλο παλινδρόμησης Decision Tree και βασίζεται στα χαρακτηριστικά που ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα. Συγκεκριμένα, η γραφική απεικονίζει τη σχέση μεταξύ των χαρακτηριστικών όπως αυτά επιλέχθηκαν μέσω της συγκεκριμένης μεθόδου και του αρνητικού μέσου τετραγωνικού σφάλματος.



Selected features: (0, 5)  
MSE: 10.120028860319591

Μπορούμε εύκολα να παρατηρήσουμε πως το το μεγαλύτερο αρνητικό μέσο τετραγωνικό σφάλμα επιτυγχάνετε για αριθμό χαρακτηριστικών ίσο με 2. Το αντίστοιχο καλύτερο μέσο τετραγωνικό σφάλμα είναι ίσο με 10.12 και επιτυγχάνετε χρησιμοποιώντας την πρώτη και έκτη στήλη του πίνακα δεδομένων.

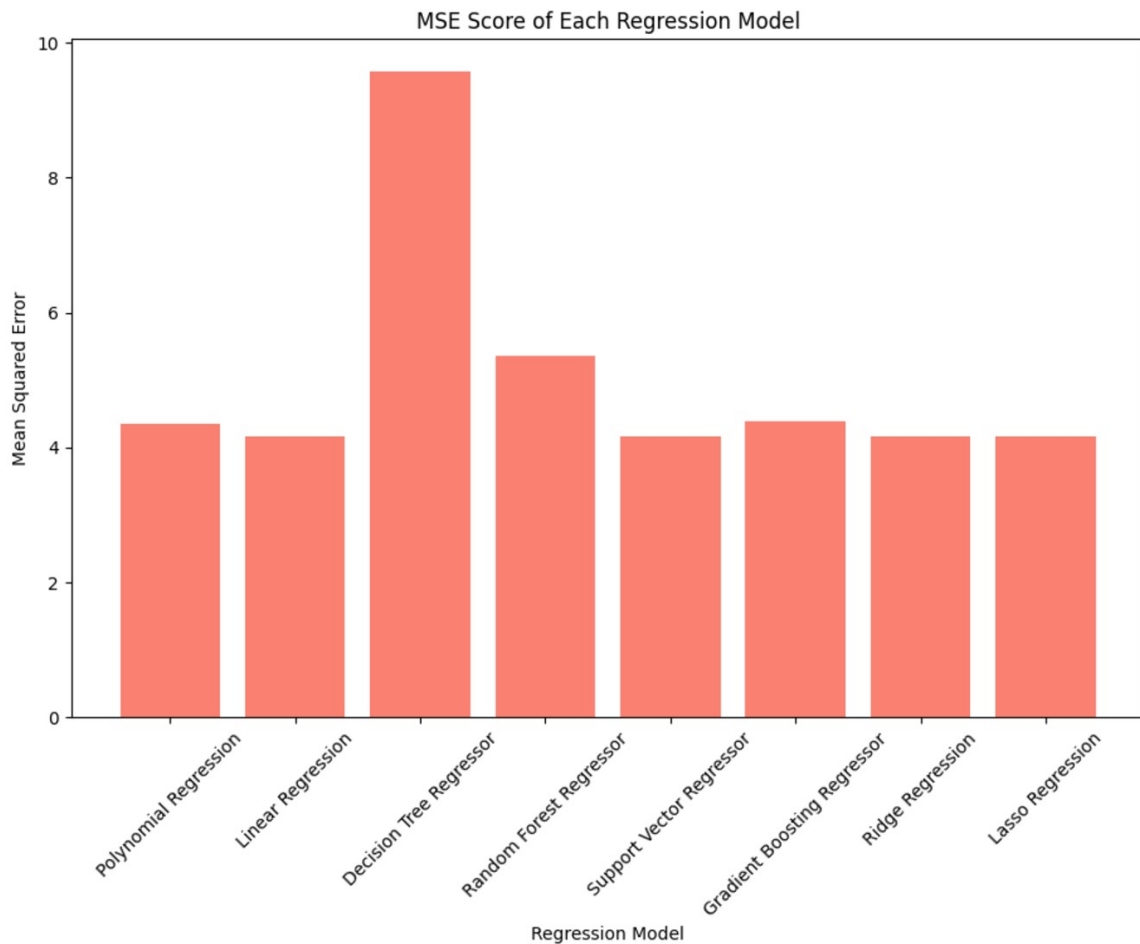
## 2.4 Επιλογή υπερπαραμέτρων

Για κάθε ένα από τα μοντέλα παλινδρόμησης που ορίσαμε παραπάνω, πραγματοποιήσαμε βελτιστοποίηση υπερπαραμέτρων (hyperparameter tuning) χρησιμοποιώντας τη μέθοδο GridSearchCV. Η βελτιστοποίηση έγινε για κάθε μοντέλο παλινδρόμησης χρησιμοποιώντας τα αντίστοιχα χαρακτηριστικά που επιλέχθηκαν ως βέλτιστα μέσω της μεθόδου Sequential Forward Selection (SFS). Στη διαδικασία, θέσαμε cross-validation = 10 και επιλέξαμε ως μετρική απόδοσης το μέσο τετραγωνικό σφάλμα(MSE). Η διαδικασία βελτιστοποίησης πραγματοποιήθηκε σε δύο στάδια. Αρχικά, εξετάσαμε ένα εύρος παραμέτρων γύρω από τις προεπιλεγμένες τιμές (default parameters), ώστε να αποκτήσουμε μια αρχική εικόνα για τη συμπεριφορά του μοντέλου. Στη συνέχεια, εστιάσαμε στις παραμέτρους που ανέδειξε η πρώτη διαδικασία ως πιο σημαντικές, προσαρμόζοντας το εύρος των τιμών τους για πιο λεπτομερή αναζήτηση. Με αυτόν τον τρόπο καταφέραμε να βελτιώσουμε περαιτέρω την απόδοση των μοντέλων μας, εξασφαλίζοντας καλύτερα αποτελέσματα.

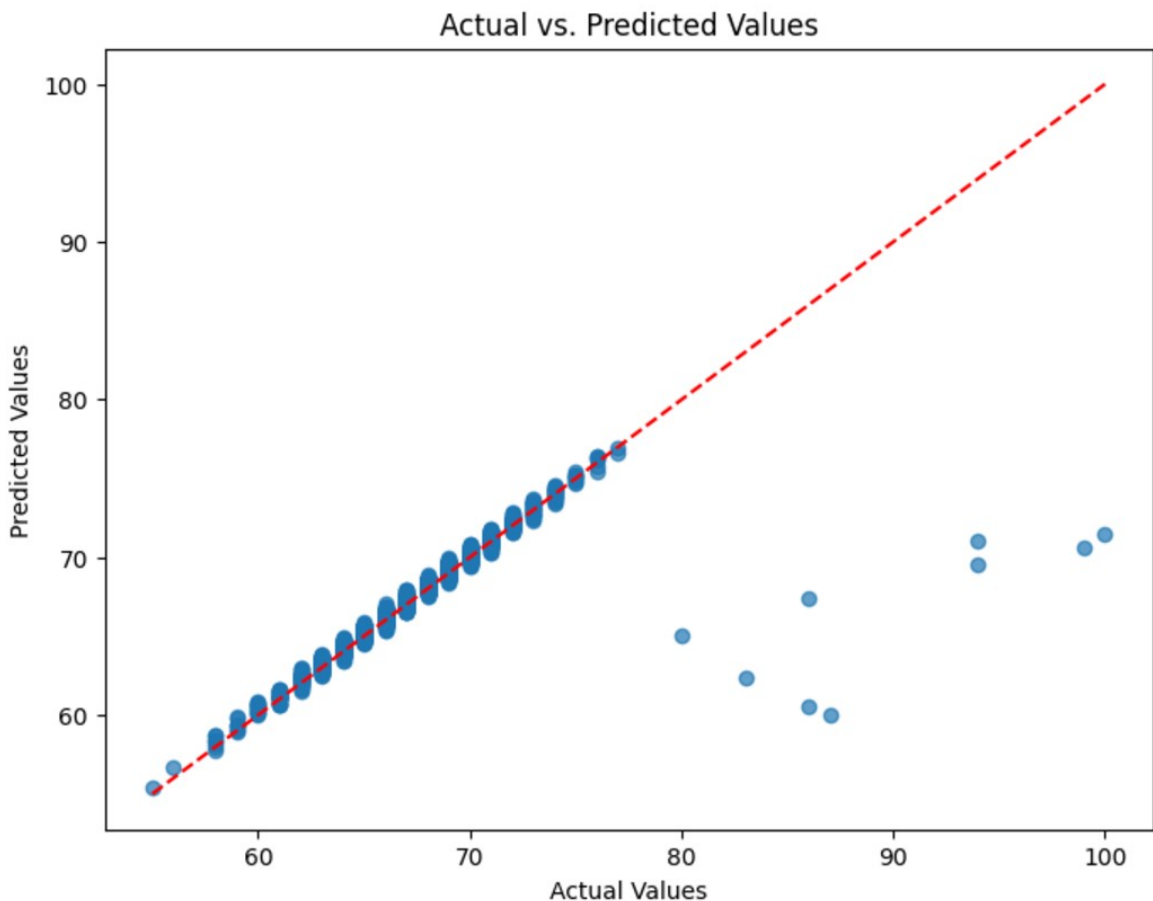
Η μέθοδος GridSearchCV λειτουργεί ως ακολούθως. Το σύνολο εκπαίδευσης χωρίζεται σε  $k$ =cross-validation ίσα μέρη(folds). Στη συνέχεια, για κάθε συνδυασμό υπερπαραμέτρων που ορίζει ο χρήστης το μοντέλο εκπαιδεύεται  $k$  φορές, χρησιμοποιώντας κάθε φορά ένα διαφορετικό fold ως σύνολο δοκιμής (test set), ενώ τα υπόλοιπα  $k-1$  folds χρησιμοποιούνται ως σύνολο εκπαίδευσης(training set). Η απόδοση κάθε συνδυασμού υπερπαραμέτρων αξιολογείται με βάση τον μέσο όρο της απόδοσης για κάθε επιλογή fold ως test set. Τέλος, επιλέγεται ο συνδυασμός υπερπαραμέτρων που βελτιστοποιεί τη μετρική απόδοσης.

## 2.5 Επιλογή μοντέλου

Για να βρούμε το τελικό βέλτιστο μοντέλο παλινδρόμησης από αυτά που εξετάσαμε κατασκευάσαμε το ραυδόγραμμα του μέσου τετραγωνικού σφάλματος για κάθε ένα από τα μοντέλα παλινδρόμησης που βελτιστοποιήσαμε. Το ραυδόγραμμα αυτό παρουσιάζεται παρακάτω:



Το μοντέλο το οποίο περιγράφει καλύτερα τα δεδομένα μας και προβλέπει σε καλύτερο βαθμό τον βαθμό του μαθητή με βάση την μετρική απόδοσης MSE είναι το μοντέλο παλινδρόμησης Lasso με μέσο τετραγωνικό σφάλμα ίσο με 4.155, επομένως έχουμε  $RMSE = \sqrt{MSE} = 2.038$  και αυτό σημαίνει πως κατά μέσο όρο η πρόβλεψη μας για κάθε βαθμό εξέτασης διαφέρει κατά 2.038 μονάδες από την πραγματική τιμή. Με βάση την ανάλυση που κάναμε για τις υπερπαραμέτρους θέσαμε ρυθμό κανονικοποίησης(α) ίσο με 0.0001, μέγιστο αριθμό επαναλήψεων(max\_iter) ίσο με 500 και ανοχή σύγκλισης(tol) ίση με 0.1. Για να εξετάσουμε περαιτέρω την απόδοση του μοντέλου μας υπολογίσαμε τον συντελεστή προσδιορισμού  $R^2$ , ο οποίος ισούται με 0.733 επομένως το μοντέλο μας εξηγά το 73.3% της μεταβλητότητας του βαθμού του μαθητή(Exam Score). Τέλος, κατασκευάσαμε το γράφημα που φαίνεται παρακάτω όπου παρουσιάζει τις προβλεπόμενες τιμές του συνόλου δοκιμής σε σχέση με τις αντίστοιχες πραγματικές τιμές.



Μπορούμε να παρατηρήσουμε πως το μοντέλο μας με βάση την παραπάνω γραφική παράσταση υπολογίζει σε αρκετά ικανοποιητικό βαθμό τον βαθμό του μαθητή για τιμές μικρότερες από 80. Για τιμές μεγαλύτερες από 80 δεν μπορούμε να πούμε το ίδιο. Αυτό οφείλετε στο ότι δεν υπάρχουν αρκετοί μαθητές οι οποίοι πήραν βαθμό μεγαλύτερο από 80 στο training set (35 σε αριθμό), επομένως το μοντέλο μηχανικής μάθησης που δημιουργήσαμε δεν έχει αρκετά δεδομένα για να προβλέψει ικανοποιητικά τον τελικό βαθμό του μαθητή. Επίσης λόγω της μεγάλης διαφοράς στο πλήθος των δεδομένων στο training set για βαθμούς μικρότερους από 80 σε σχέση με μεγαλύτερους από αυτό, μπορεί να οδήγησε το μοντέλο μας σε υπερβολική εστίαση στα δεδομένα μικρότερα από 80 (biased ως προς αυτά τα δεδομένα) αφού τελικά θα οδηγούσε σε μικρότερο μέσο τετραγωνικό σφάλμα σε αντίθεση με το να το έκανε γενίκευση (generalize) και για τιμές μεγαλύτερες από 80.



### **3 Το μοντέλο παλινδρόμησης Lasso**

#### **3.1 Έγκριση ειδικού πλαισίου χωροταξικού σχεδιασμού**

Αναλύεται η διαδικασία έγκρισης του πλαισίου.

#### **3.2 Καθορισμός περιοχών αιολικής προτεραιότητας**

Αναλύονται οι παράγοντες επιλογής περιοχών.

#### **3.3 Παράμετροι που καθορίζουν τις αποστάσεις**

Παρουσιάζονται τα κριτήρια για τις αποστάσεις υποδομών.

4    **Σύγκριση με έρευνες**

hb

## 5 Βιβλιογραφία