

Διερεύνηση αιτιωδών σχέσεων στην ελληνική αγορά ηλεκτρικής ενέργειας με τη χρήση μεθόδων μηχανικής μάθησης

Δημήτρης Πανουρής

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΕΜΠ

29 Οκτωβρίου 2014

Περιεχόμενα

Μοντελοποίηση

Μάθηση

Ανάλυση

Συμπεράσματα

Περιεχόμενα

Μοντελοποίηση

Μάθηση

Ανάλυση

Συμπεράσματα

SMP

Τιμή ενέργειας στην χονδρεμπορική αγορά

Παράγοντες που επηρεάζουν το SMP

1. Η ζήτηση της ηλεκτρικής ενέργειας
2. Η παραγωγή από ΑΠΕ
3. Οι τιμές των καυσίμων
4. Η διαθεσιμότητα των πόρων

Σκοπός

- Δημιουργία μοντέλου για την πρόβλεψη του SMP με βάση τις παραμέτρους που το επηρεάζουν
- Κατανόηση των σχέσεων μεταξύ των μεταβλητών του μοντέλου (predictors, \mathbf{X}) και του SMP (έξοδος - response, Y)

Παραδοχές για το μοντέλο

1. Δεν υπάρχει αιτιώδης συνάφεια μεταξύ της εξόδου (SMP) και ενός ή περισσότερων μεταβλητών εισόδου.
2. Χρησιμοποιούνται όλες οι σχετικές με το μοντέλο μεταβλητές
3. Οι περιττές επεξηγηματικές μεταβλητές που προσθέτουν θόρυβο στο σύστημα αποκλείονται.
4. Αποφεύγεται η συμπερίληψη στο μοντέλο πολλαπλών συσχετισμένων μεταξύ τους μεταβλητών.

Μεταβλητές Μοντέλου

- smp** Τιμή ενέργειας στην χονδρεμπορική αγορά
- availability** Δυναμικότητα μονάδων παραγωγής λιγνίτη
- exports** Εξαγωγές ηλεκτρικής ενέργειας
- hydrogen** Παραγωγή υδροηλεκτρικών εργοστασίων
- imports** Εισαγωγές ηλεκτρικής ενέργειας
- lignite** Παραγωγή λιγνίτη
- load_forecast** Ζήτηση ηλεκτρικής ενέργειας
- ngas** Παραγωγή ενέργειας από φυσικό αέριο
- res_forecast** Πρόγνωση παραγωγής ΑΠΕ
- waters** Παραγωγή υποχρεωτικών νερών
- waip** Τιμή φυσικού αερίου

Μονάδα ενέργειας: MW

Δεδομένα: Ημερήσια 1/1/2009 - 30/11/2013

Φυσικό Αέριο

- Βιβλιογραφικές αναφορές πως επηρεάζει πολύ το SMP και περισσότερο σε υψηλή ζήτηση
- Τιμές πιο ευμετάβλητες σε σχέση με λιγνίτη

Θα αναζητήσουμε την σχέση μεταξύ φυσικού αερίου και SMP στην ανάλυσή μας.

Περιεχόμενα

Μοντελοποίηση

Μάθηση

Ανάλυση

Συμπεράσματα

Μηχανική Μάθηση

- Απόκτηση πληροφοριών από επεξεργασία δεδομένων
- Αυτοματοποιημένη ανάλυση
- Πρόβλεψη μη γραμμικότητας
- Ευρύ πλήθος εφαρμογών

Πολύ χρήσιμη όταν το πρόβλημα παρουσιάζει μοτίβο το οποίο δεν μπορούμε να περιγράψουμε με μαθηματικό τρόπο.

Αλγόριθμοι - Πλεονεκτήματα

- CART: Απεικόνιση Μοντέλου σε Δέντρα
- Random Forests: Βαθμός επιρροής μεταβλητών στην έξοδο (Σημαντικότητα Μεταβλητών)
- MARS: Δημιουργία Συνάρτησης

Η ανάλυση έγινε στη γλώσσα προγραμματισμού Python.

Διαδικασία Μάθησης

Έξοδος (Y)



Μεταβλητές (X)

Μεταβλητή Υστέρησης
(βελτίωση πρόγνωσης)Σύνολο
Εκπαίδευσης
(train test)

$\hat{Y}_{\text{predicted}}$	Y_{real}	Error
8.8	9.2	0.4
8.6	8.5	-0.1

 $\hat{f}(X_e)$

Y(t)	A(t)	B(t)	A(t-1)
?	7.5	11.1	
?	4.9	12.4	7.5

Σύνολο Επικύρωσης (X_e)
(test set)Τροφοδότηση σε Αλγόριθμο
Μηχανικής ΜάθησηςΈξοδος: Συνάρτηση \hat{f} Τροφοδότηση X_e
Πρόβλεψη Y

Δείκτες μέτρησης ακρίβειας Μοντέλων

Μέτρηση ακρίβειας μοντέλου $(-\infty, 1]$ - όσο υψηλότερο τόσο καλύτερο:

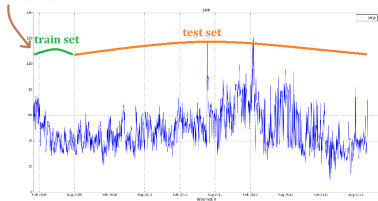
$$\text{explained_variance}(y, \hat{y}) = 1 - \frac{\text{Var}\{y - \hat{y}\}}{\text{Var}\{y\}} \quad (1)$$

Άλλοι δείκτες για εξαγωγή ολοκληρωμένων συμπερασμάτων:

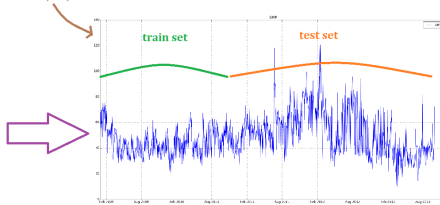
- $R^2(y, \bar{y})$: Ακρίβεια $[0, 1]$
- $MSE(y, \hat{y})$: Μέσο τετραγωνικό σφάλμα $[0, \infty)$
- $MAE(y, \hat{y})$: Μέσο απόλυτο σφάλμα $[0, \infty)$

Μέθοδος Ανάλυσης

CART, RF, MARS



CART, RF, MARS

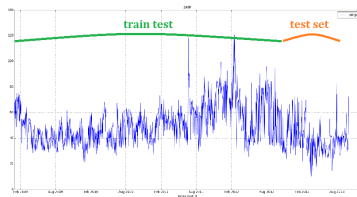


Μοντέλο 2
Variance 2
MSE 2
MAE 2
Σημαντικότητες 2



Μοντέλο 1
Variance 1
MSE 1
MAE 1
Σημαντικότητες 1

Μοντέλο N
Variance N
MSE N
MAE N
Σημαντικότητες N



CART, RF, MARS

Σταθερότητα Μοντέλων

Θέλουμε τα μοντέλα να είναι σταθερά ως προς την δομή τους:

- Να μην έχουν ακραία χαρακτηριστικά (όπως πολύ χαμηλή ακρίβεια) σε κάποια από τις εκτελέσεις
- Παρόμοιες σημαντικότητες μεταβλητών σε διαφορετικές εκτελέσεις

Περιεχόμενα

Μοντελοποίηση

Μάθηση

Ανάλυση

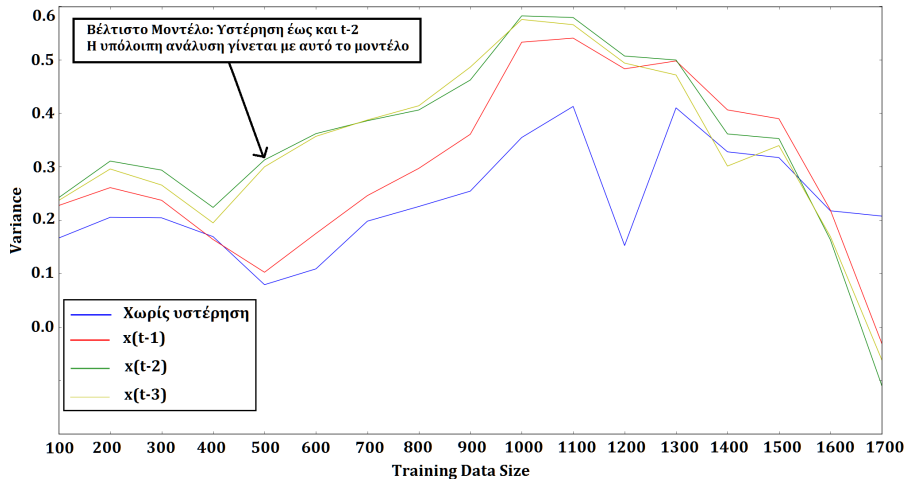
Συμπεράσματα

Ανάλυση

- Εύρεση βέλτιστης υστέρης για καλύτερο μοντέλο
- Σύγκριση αλγορίθμων και προβλεπτικής ικανότητας μοντέλων τους
- Μελέτη σημαντικότητας μεταβλητών στον καθορισμό του SMP - Partial Dependence Plots

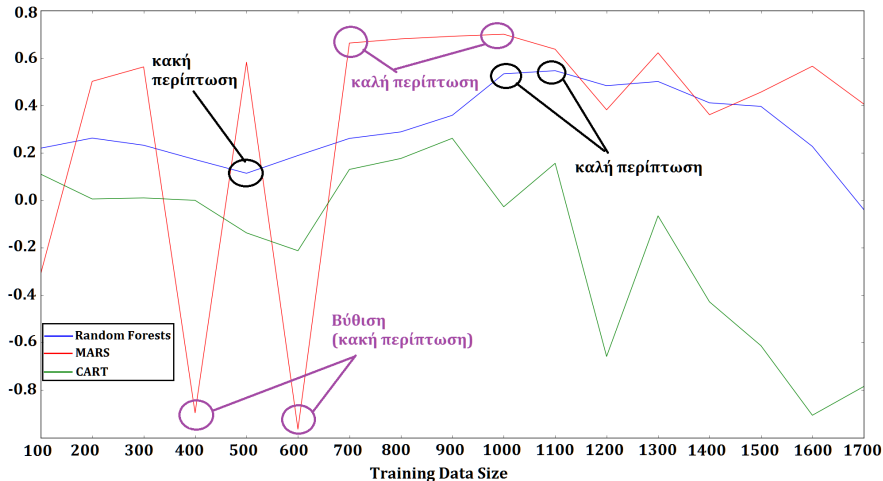
Variance - Μεταβλητές Υστέρησης

Variance for different lagged values

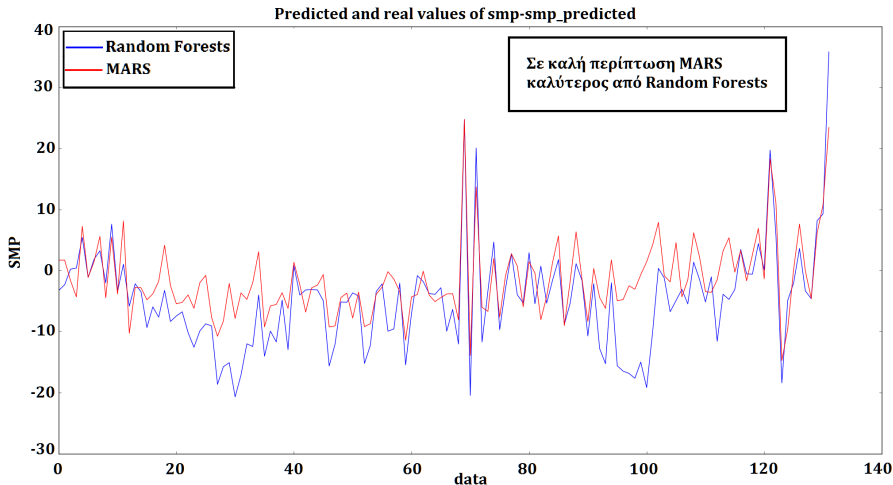


Variance - Διαφορετικοί Αλγόριθμοι

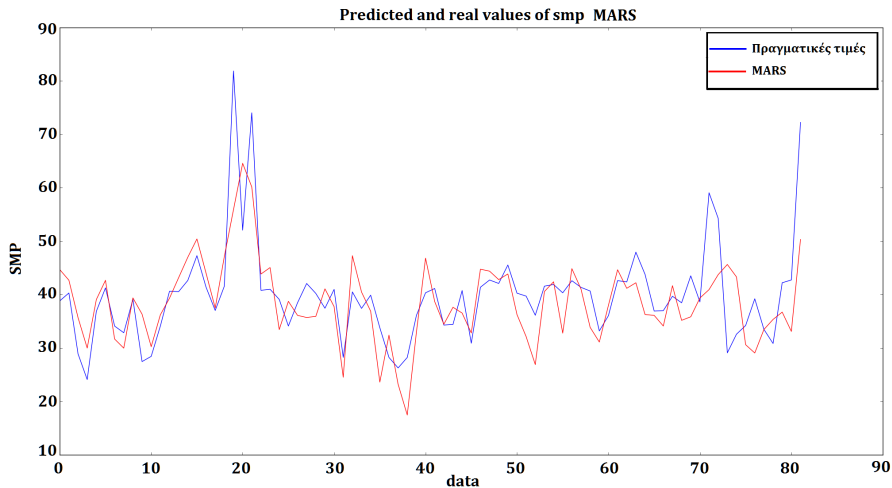
Variance - Random Forests, MARS, CART



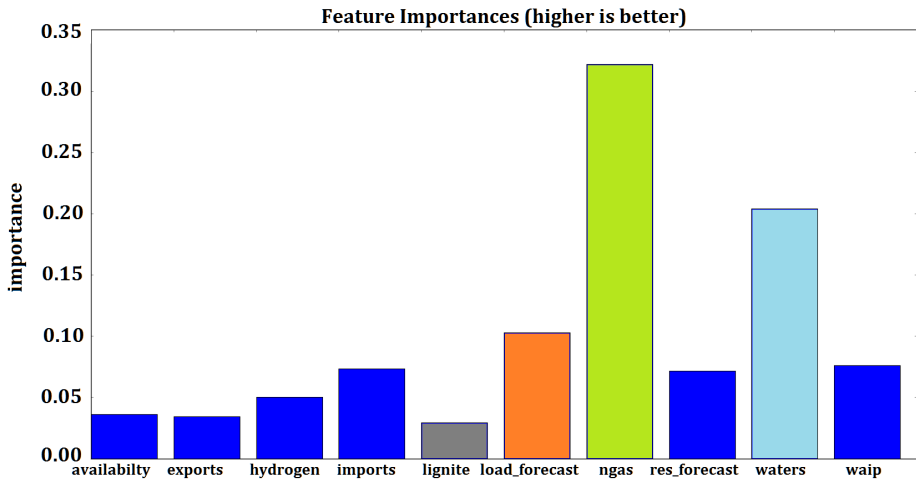
Random Forest και MARS: Διαφορά Πραγματικών-Εκτιμώμενων τιμών (καλή περίπτωση)



MARS: predicted και πραγματικές τιμές (καλή περίπτωση)

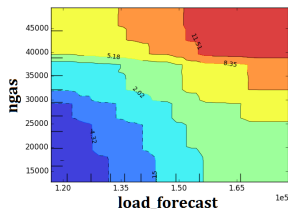
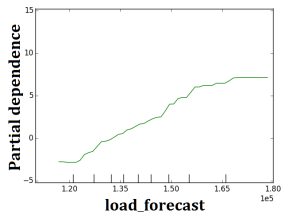
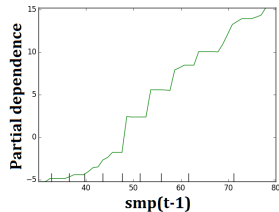
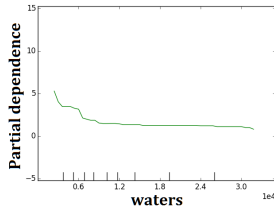
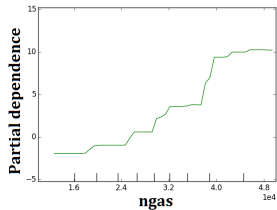


Random Forests - Σημαντικότητες Μεταβλητών



Partial Dependence - Target Features

Partial dependence ngas , $\text{smp}(t-1)$, load_forecast , waters



Περιεχόμενα

Μοντελοποίηση

Μάθηση

Ανάλυση

Συμπεράσματα

Συμπεράσματα

- MARS μπορεί να παράξει έξοδο με υψηλή ακρίβεια αλλά δεν είναι τόσο αξιόπιστος
- Random Forests κάνει μικρότερα σφάλματα αλλά δεν έχει τόσο καλή προσαρμογή
- Τα μοντέλα μπορούν να χρησιμοποιηθούν με ασφάλεια για την πρόβλεψη της τάσης της τιμής
- Σημαντικός ρόλος φυσικού αερίου και ζήτησης στον καθορισμό του SMP
- Τα παραπάνω συμπεράσματα βασίζονται στην ανάλυση περισσότερων των 100 μοντέλων που δημιουργήθηκαν μέσω μηχανικής μάθησης

Προοπτικές

- Χρησιμοποίηση επιπλέον αλγορίθμων μηχανικής μάθησης για προβλέψεις SMP
- Εξέταση αιτιωδών σχέσεων μεταξύ ανανεώσιμων και μη ανανεώσιμων πηγών ενέργειας
 - Εκτίμηση μείωσης που δημιουργούν οι ανανεώσιμες πηγές στην παραγωγή συμβατικών
 - Πρόβλεψη μη ανανεώσιμων που θα μπορούσαν να αντικατασταθούν από ΑΠΕ

Ευχαριστώ! Ερωτήσεις;