

Σκονάκι Παρουσίασης 23/05/2024

Διαφάνεια 6 (lamda):

Σε αυτήν την διαφάνεια μπορούμε να δούμε το φιλτράρισμα του λ. Συγκεκριμένα, χρησιμοποιήσαμε το φίλτρο Chauvenet (Σοβενετ), το οποίο αφαιρεί δεδομένα με βάση κάποιο threshold (κατώτατα όρια δηλαδή για τα οποία όποια δεδομένα τα ξεπερνάνε θα τα αφαιρούμε). Το threshold το επιλέξαμε πειραματικά, συγκρίνοντας τα διαγράμματα του πριν και μετά όπως φαίνεται στην διαφάνεια. Πάνω φαίνονται τα αρχικά δεδομένα και κάτω τα δεδομένα μετά την εφαρμογή του φίλτρου. Για threshold στην συγκεκριμένη περίπτωση επιλέξαμε 10^{-4} .

Διαφάνεια 7 (Fuel Consumption):

Για μεγαλύτερη ευχέρεια επιλέξαμε ξανά το φίλτρο Chauvenet με threshold 10^{-6} αυτή την φορά. Στα υπόλοιπα δεδομένα είδαμε ότι δεν είχαν κάποιο σημαντικό αποτέλεσμα τα φίλτρα, εφαρμόσαμε όμως σε κάθε δεδομένο από ένα φίλτρο προληπτικά, ανάλογα με την κρίση μας κάθε φορά. Σημαντικό είναι επίσης να αναφερθεί ότι όταν αφαιρούσαμε με ένα φίλτρο μια τιμή από μια παράμετρο, αφαιρούσαμε από όλες τις παραμέτρους τα δεδομένα για εκείνη την χρονική στιγμή για να μην υπάρχει κάποια ασυνέχεια.

Διαφάνεια 8 (Filtering):

Εδώ βλέπουμε τα τελικά σετ δεδομένων, στα οποία έχουμε καθαρίσει τις τιμές που αναφέραμε νωρίτερα. Παρατηρούμε ότι δεν έχουν αλλάξει σημαντικά τα δεδομένα μας και ότι δεν μας προκαλούν κάποιο πρόβλημα για την συνέχεια. Οπότε μπορούμε να προχωρήσουμε στον χωρισμό των δεδομένων για να ξεκινήσει η εκπαίδευση.

Διαφάνεια 9 (Separation):

Το φιλτραρισμένο σετ δεδομένων χωρίστηκε αρχικά σε training and testing set με μια αναλογία (80%-20%). Στην συνέχεια, το training set που δημιουργήθηκε χωρίστηκε σε training and validating set ξανά με μια αναλογία (80%-20%). Η επιλογή της αναλογίας έγινε αυθαίρετα, απλά επειδή είναι γνωστό από την βιβλιογραφία και το μάθημα ότι η αναλογία 80%-20% αποτελεί μια καλή πρακτική. Τα δεδομένα χωρίστηκαν με την χρήση της cvpartition η οποία κάνει τυχαίο χωρισμό σε ένα dataset με ένα ποσοστό που θα της ορίσεις. Αφού χωρίσαμε τα δεδομένα, τα αποθηκεύσαμε, για να είμαστε σίγουροι ότι το testing set δεν θα χρησιμοποιηθεί σε καμία περίπτωση κατά το training.

Διαφάνεια 10 (Normalization):

Ακόμα, εφαρμόσαμε κανονικοποίηση (normalization) στα datasets, δηλαδή τα μεταφέραμε από την αρχική τους κλίμακα, στην κλίμακα $[0,1]$. Αυτό το κάνουμε γιατί σε αυτήν την μορφή είναι πιο εύκολο στο νευρωνικό να κάνει τους υπολογισμούς. Να τονίσουμε ότι για κάθε dataset αποθηκεύσαμε τα center and scale values, για να μπορούμε να επιστρέψουμε στο τέλος στα πραγματικά δεδομένα.

Διαφάνεια 11 (Correlation Matrix):

Το Correlation Matrix δείχνει την συσχέτιση μεταξύ των παραμέτρων. Όσο πιο κοντά είναι η τιμή στο 1, τόσο πιο γραμμική συμπεριφορά εμφανίζουν οι 2 παράμετροι, ενώ στο -1 αντίστροφη γραμμική συμπεριφορά. Αν η τιμή προσεγγίζει το 0, οι δύο παράμετροι δεν εμφανίζουν μεταξύ τους γραμμική συμπεριφορά. Χρησιμοποιείται στη συνέχεια για την επιλογή input features των νευρωνικών.

Διαφάνεια 12 (Features to Train):

Επιλέξαμε να εκπαιδεύσουμε τα NOx και το Fuel Consumption.

NOx:

- Λόγω της πολυπλοκότητας των αισθητήριων ανίχνευσης οξειδίων του αζώτου (ακριβά και δύσκολα ως προς την συντήρηση).
- Έχουν σοβαρή περιβαλλοντική επίδραση, καθώς επηρεάζει την ποιότητα του αέρα (κίνδυνος για την ανθρώπινη υγεία), αλλά και σε υψηλά επίπεδα συνεισφέρουν στην δημιουργία όξινης βροχής.
- Απαιτούν στενή παρακολούθηση καθώς πολλές περιοχές έχουν αυστηρούς κανονισμούς για τις εκπομπές ρύπων και συνεπώς απαιτείται έλεγχος για την κατάλληλη συμμόρφωση.

Fuel Consumption:

- Καθώς αποτελεί σημαντικό οικονομικό παράγοντα και επηρεάζει όσο τα λειτουργικά τόσο και τα κατασκευαστικά έξοδα.
- Επιπλέον, αποτελεί κύριο δείκτη αποδοτικότητας της μηχανής, με την βελτιστοποίηση του οποίου αυξάνεται η βιωσιμότητα και η επίδοση του κινητήρα.

Διαφάνεια 13 (Input Features):

Και για τις δύο παραμέτρους χρησιμοποιήσαμε σαν input feature τον χρόνο, καθώς η μηχανή έχει περιοδική συμπεριφορά ανάλογα με τους κύκλους λειτουργίας, οπότε καταλήξαμε στο ότι ο χρόνος μας προσφέρει μια χρήσιμη πληροφορία. Το λ παρουσίαζε μεγάλη συσχέτιση με NOx (-0.53) και Fuel Consumption (-0.66). Το Rot Speed παρουσίαζε και μεγάλη συσχέτιση με NOx (-0.44), αλλά και μικρή με το λ (0.01), οπότε μπορούμε να συμπεράνουμε ότι λειτουργούν σαν 2 ανεξάρτητες μεταβλητές.

Το Intake Pressure παρουσιάζει σχεδόν γραμμική σχέση με το Fuel Consumption (0.98) και το Exhaust Gas Temperature (0.88) επίσης. Αρχικά, δοκιμάσαμε να εκπαιδεύσουμε το νευρωνικό με βάση αυτά τα δεδομένα και παρατηρήσαμε ότι έκανε overfit. Τελικά επιλέξαμε τα EGR Command και Exhaust Gas Temperature με συσχετίσεις 0.73 και 0.63 αντίστοιχα.

Διαφάνεια 14 (Correlation Matrix):

Εδώ μπορούμε να δούμε τις συσχετίσεις που αναφέραμε νωρίτερα.

NOx:

- time
- lamda
- Rot Speed

Fuel Consumption:

- Time
- lamda
- (Intake Pressure)
- (Exhaust Gas Temperature)
- EGR Command
- Exhaust Gas Temperature