ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ευφυής έλεγχος οχημάτων για εξοικονόμηση ενέργειας

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ανδρέας Ευαγγέλου Μπουγιουκλής

Επιβλέπων : Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2017

Η σελίδα αυτή είναι σκόπιμα λευκή.

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ευφυής έλεγχος οχημάτων για εξοικονόμηση ενέργειας

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ανδρέας Ευάγγελος Μπουγιουκλής

Επιβλέπων : Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 10η Οκτωβρίου 2017.

(Υπογραφή)	*(Υπογραφή)*	*(Υπογραφή)*
.....	.....	.....
Γεώργιος Στάμου	Αντώνιος Κλαδάς	Ανδρέας-Γεώργιος Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2017

Ανδρέας Ευαγγέλου Μπουγιουκλής

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

Copyright © Ανδρέας Ευαγγέλου Μπουγιουκλής
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

Η παγκόσμια ευαισθητοποίηση για την μόλυνση του περιβάλλοντος και την κατανάλωση ενέργειας έχει ωθήσει τις αυτοκινητοβιομηχανίες στην υλοποίηση νέων οικολογικών συστημάτων. Σύμφωνα με το πρωτόκολλο του Κιότο η τάση στη παγκόσμια αυτοκινητοβιομηχανία είναι η ανάπτυξη μίας οικολογικής στρατηγικής με στόχο την ελαχιστοποίηση της καταναλισκόμενης ενέργειας των οχημάτων και την εξοικονόμηση καυσίμων. Μια τέτοια στρατηγική είναι η εύρεση της βέλτιστης ακολουθίας ταχυτήτων που θα επιτρέψουν στο αυτοκίνητο να φτάσει στο προορισμό του σε ένα επιλεγμένο χρονικό διάστημα, καταναλώνοντας την ελάχιστη δυνατή ενέργεια. Επιπλέον ο τρόπος με τον οποίο θα κινηθεί ένα όχημα πάνω στο δρόμο επηρεάζει τη μέγιστη ταχύτητα που μπορεί να μεταφέρει σε μία στροφή. Επομένως για να εκτελεστεί η βέλτιστη ακολουθία ταχυτήτων με ασφάλεια πολλές φορές είναι απαραίτητο να ακολουθηθεί η αγωνιστική γραμμή. Συνεπώς ο οδηγός πρέπει να είναι ενήμερος για το πώς θα συμπεριφερθεί στο δρόμο προκειμένου να μετακινείται με ασφάλεια αλλά και να αποφύγει την αχρείαστη κατανάλωση ενέργειας μέσω της εσκεμμένης μείωσης της ταχύτητας.

Αντικείμενο αυτής της διπλωματικής είναι ο σχεδιασμός και η υλοποίηση ενός συστήματος το οποίο θα παίρνει πληροφορίες από το οδόστρωμα, θα τις επεξεργάζεται με γνώμονα τις δυνατότητες του οχήματος και τον επιθυμητό χρόνο τερματισμού και θα εξάγει τα βέλτιστα προφίλ οδήγησης για μία συγκεκριμένη διαδρομή. Το σύστημα θα χρησιμοποιεί τεχνικές μηχανικής μάθησης προκειμένου να ενημερώνει τα προφίλ οδήγησης και θα κάνει προβλέψεις προκειμένου να εναλλάσσεται μεταξύ τους για να αντιμετωπίζει πιθανές καθυστερήσεις. Στο πλαίσιο της διπλωματικής θα υλοποιηθεί το ηλεκτρονικό κύκλωμα το οποίο θα παίρνει τα δεδομένα από το αυτοκίνητο και ο αλγόριθμος ο οποίος θα εκμεταλλεύεται τη γνώση που του δίνεται για να εξάγει τα βέλτιστα προφίλ οδήγησης. Τέλος όλα τα αποτελέσματα του συστήματος που περιγράφηκε θα εισάγονται στον υπολογιστή ταξιδίου του οχήματος και θα παρουσιάζονται στον οδηγό μέσω μιας διαδραστικής εφαρμογής που προσφέρει διορθώσεις στην οδηγική συμπεριφορά του υπό τη μορφή παιχνιδιού. Για την επίτευξη των παραπάνω θα χρησιμοποιηθεί το πρωτότυπο ερευνητικό όχημα του Εθνικού Μετσόβιου Πολυτεχνείου, Πυρφόρος.

Λέξεις Κλειδιά: <<Νευρωνικό δίκτυο, ανατροφοδοτούμενο δίκτυο, μηχανική μάθηση, επιβραβευόμενη μάθηση, αλγόριθμος εκμάθησης Q , βέλτιστη ακολουθία ταχυτήτων, αγωνιστική γραμμή, raspberry>>

Η σελίδα αυτή είναι σκόπιμα λευκή

Abstract

The Intelligent Eco Vehicle Control (IEVC) is an intelligent and complete system which interacts with the driver of an electrical vehicle by calculating the optimum speed and the racing line and teaches him/her how to use this knowledge in the track by the procedure of gamification. This report discusses five keypoints. Firstly the consumption model is implemented by a LSTM neural network which is trained by the data collected from the telemetry of the vehicle. Secondly the system learns how to optimize the velocity and predict the racing line by training an intelligent agent through the procedure of reinforcement learning. Thirdly it adapts the optimum velocity profile online and it simplifies the race by converting it into a game. And finally it offers about 24.0306% decrease of energy consumption.

Keywords : <<Neural Network, recurrent neural network, artificial intelligence, reinforcement learning, Q learning, optimum velocity profiles, racing line, raspberry>>

Ευχαριστίες

Με την ολοκλήρωση της διπλωματικής εργασίας θα ήθελα να εκφράσω τις ευχαριστίες μου σε όσους με βοήθησαν κατά την εκπόνηση της.

Ευχαριστώ τον καθηγητή μου κ. Γεώργιο Στάμου και τον υπεύθυνο υποψήφιο διδάκτορα Αντώνη Κορκοφίγκα για την εξαιρετική συνεργασία. Επιπλέον θα ήθελα να ευχαριστήσω το κ. Αντώνιο Κλαδάς που ως υπεύθυνος της ομάδας «Προμηθέας» στήριξε την υλοποίηση του παρουσιαζόμενου συστήματος.

Τέλος θα ήθελα να ευχαριστήσω θερμά όλα τα μέλη της ομάδας «Προμηθέας» τόσο για τη βοήθεια στην εργασία όσο και για το αξέχαστο κλίμα που δημιούργησαν κάνοντας την δουλειά στο εργαστήριο ευχάριστη. Ιδιαίτερα ευχαριστώ για τη βοήθεια τους στη Κυριακή Βαρμπομπίτη, το Μάρκο Διωματάρη, το Στέλιο Προκοπίου και τη Νίκη Κυλιντηρέα.

Πίνακας Περιεχομένων

Εισαγωγή 12
1.1 Σκοπός της διπλωματικής 13
1.2 Διάρθρωση εργασίας 14
Τρόποι μείωσης της κατανάλωσης των αυτοκινήτων 15
2.1 Εισαγωγή 15
2.2 Παρουσίαση οχήματος και ασκούμενων δυνάμεων 15
2.2 Εξοικονόμηση ενέργειας μέσω της πρόβλεψης των τοπογραφικών χαρακτηριστικών 17
Νευρωνικά Δίκτυα 19
3.1 Εισαγωγή 19
3.1 Το δίκτυο Perceptron 19
3.2 Το Μοντέλο Perceptron Πολλών Επιπέδων – Multilayer Perceptron (MLP) 22
3.3 Ανατροφοδοτούμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks, RNN) 23
3.3.1 Ανατροφοδοτούμενα Δίκτυα Μακράς Βραχυπρόθεσμης Μνήμης (Long Short Term Memory, LSTM) 24
3.4 Εκπαίδευση Νευρωνικών Δικτύων 26
3.4.1 Εκπαίδευση Δικτύων Προσοτροφοδότησης 26
3.4.2 Εκπαίδευση Ανατροφοδοτούμενων Δικτύων 28
3.5 Εφαρμογές Νευρωνικών Δικτύων 29
Μοντέλο Κατανάλωσης 30
4.1 Εισαγωγή 30
4.2 Προεπεξεργασία δεδομένων 31
4.2.1 Δεδομένα εισόδου-εξόδου 31
4.2.2 Μεθοδολογία προεπεξεργασίας 31
4.3 Επιλογή εισόδων-εξόδων 32
4.3.1 MLP μοντέλο κατανάλωσης 32
4.3.2 LSTM μοντέλο κατανάλωσης 34
4.3 Πειραματισμός παραμέτρων 38
4.4 Αποτελέσματα 40
Υπολογισμός βέλτιστης ακολουθίας ταχυτήτων 43
5.1 Εισαγωγή 43
5.2 Ενισχυτική μάθηση (Reinforcement Learning, RL) 44
5.3 Μεθοδολογία αλγορίθμου Q-Learning 45
5.4 Τμηματοποίηση της διαδρομής 46

5.4 Σχεδιασμός αλγορίθμου 48
5.4.1. Διαστάσεις καταστάσεων (s) και ενεργειών (a) 48
5.4.2. Συνάρτηση επιβράβευσης 48
5.4.3 Στρατηγική επιλογής ενέργειας 49
5.5 Πειραματισμός παραμέτρων 50
5.6 Παρουσίαση αποτελεσμάτων 51
5.7 Πειραματική τεκμηρίωση 53
Υπολογισμός αγωνιστικής γραμμής 56
6.1 Εισαγωγή 56
6.2 Τι είναι αγωνιστική γραμμή 57
6.3 Δεδομένα εισόδου 58
6.4 Επεξεργασία μετρήσεων 59
6.5 Ανάπτυξη αλγορίθμου 61
6.5.1 Διαστάσεις καταστάσεων (s) και ενεργειών (a) 61
6.5.2 Πίνακας Q 61
6.5.3 Συνάρτηση επιβράβευσης 61
6.5.3 Στρατηγική επιλογής ενέργειας 62
6.6 Πειραματισμός παραμέτρων 63
6.7 Αποτελέσματα 64
6.8 Συμπεράσματα 66
Περιγραφή χρησιμοποιηθέντος υλικού 67
7.1 Εισαγωγή 67
7.2 Μικροϋπολογιστικό σύστημα 67
7.3 Χαρτογράφηση πίστας 68
7.3.1 Σύστημα παγκόσμιας στιγματοθέτησης 68
7.4 Ταχύμετρο 70
7.5 Μέτρηση θέσης τιμονιού 70
7.6 Τηλεμετρία οχήματος 71
7.7 Τελική συνδεσμολογία 72
Διαδραστικό λογισμικό 74
8.1 Εισαγωγή 74
8.2 Παιχνιδοποίηση 74
8.3 Ενημέρωση λογισμικού από τους αισθητήρες 75
8.4 Επιλογή ταχύτητας 76
8.5 Επιλογή θέσης 76
8.6 Παρουσίαση συστήματος 77

Επίλογος 79
9.1 Σύνοψη και συμπεράσματα 79
9.2 Μελλοντικές επεκτάσεις 80
Βιβλιογραφία 81

Κεφάλαιο 1

Εισαγωγή

Η ανάπτυξη μίας οικολογικής στρατηγικής με σκοπό την ελαχιστοποίηση της καταναλισκόμενης ενέργειας των οχημάτων έχει βρεθεί στο επίκεντρο πολλών ερευνών από μελετητές και αυτοκινητοβιομηχανίες. Ο στόχος μίας τέτοιας στρατηγικής είναι η εύρεση της βέλτιστης ακολουθίας ταχυτήτων που θα επιτρέψουν στο αυτοκίνητο να φτάσει στο προορισμό του σε ένα επιλεγμένο χρονικό διάστημα, καταναλώνοντας την ελάχιστη δυνατή ενέργεια. Προκειμένου να επιτευχθεί αυτό εκτελείται μία πολυκριτηριακή βελτιστοποίηση που λαμβάνει υπόψη τα χαρακτηριστικά του αυτοκινήτου, τον επιθυμητό χρόνο τερματισμού και τις υψομετρικές διαφορές του οδοστρώματος. Το μεγάλο πλεονέκτημα αυτής της μεθόδου είναι πως τα αποτελέσματα μπορούν να εισαχθούν στον υπολογιστή ταξιδίου του οχήματος.

Για την καθημερινή χρήση ενός οχήματος η βέλτιστη ακολουθία ταχυτήτων, περιοριζόμενη από τις ικανότητες του αυτοκινήτου, για μία συγκεκριμένη διαδρομή είναι αρκετή προκειμένου να ελαττωθεί η καταναλισκόμενη ενέργεια. Ωστόσο στα πλαίσια του ανταγωνισμού και των αγώνων εξοικονόμησης ενέργειας τα οχήματα πρέπει να οδηγούνται στα όρια τους. Γι’ αυτό το λόγο εκτός από τη ταχύτητα του αυτοκινήτου μέγιστη σημασία έχει και η θέση του πάνω στο οδόστρωμα. Αυτό συμβαίνει διότι ο τρόπος με τον οποίο θα κινηθεί ένα όχημα πάνω στο δρόμο επηρεάζει και τη μέγιστη ταχύτητα που μπορεί να μεταφέρει σε μία στροφή. Γίνεται προφανές πως για να εκτελεστεί η βέλτιστη ακολουθία ταχυτήτων πρέπει ο οδηγός να είναι ενήμερος για το πώς θα συμπεριφερθεί στο δρόμο προκειμένου να μετακινείται με ασφάλεια αλλά και να αποφύγει την αχρείαστη κατανάλωση ενέργειας μέσω της εσκεμμένης μείωσης της ταχύτητας.

Γεννιέται λοιπόν η ανάγκη εκπαίδευσης των οδηγικού κοινού που απευθύνεται σε ανταγωνιστικούς αγώνες οικονομίας προκειμένου να είναι ικανό να χρησιμοποιήσει τις στρατηγικές εξοικονόμησης ενέργειας με ένα ασφαλές τρόπο.

1.1 Σκοπός της διπλωματικής

Σκοπός αυτής της διπλωματικής είναι η υλοποίηση ενός ευφυούς και διαδραστικού συστήματος εκμάθησης του οδηγικού κοινού για το βέλτιστο τρόπο μετακίνησης με γνώμονα την εξοικονόμησης ενέργειας.

Για να επιτευχθεί αυτό υλοποιήθηκε ένας ευφυής πράκτορας που με τεχνικές επιβραβευόμενης μάθησης, μαθαίνει να εκμεταλλεύεται ένα μοντέλο κατανάλωσης ενός οχήματος, που υλοποιήται με μεθόδους τεχνιτών νευρωνικών δικτύων, και να εξάγει τη βέλτιστη ακολουθία ταχυτήτων για μία συγκεκριμένη διαδρομή αλλά και τη θέση που πρέπει να διατηρήσει πάνω στο οδόστρωμα, τη λεγόμενη αγωνιστική γραμμή. Η γνώση αυτή μετατρέπεται σε εκπαιδευτικό εργαλείο μέσω μίας διαδραστικής εφαρμογής που προσφέρει διορθώσεις στις κινήσεις του οδηγού υπό τη μορφή παιχνιδιού.

Η εξέλιξη του παραπάνω συστήματος πραγματοποιήθηκε για το πρότυπο ηλεκτρικό όχημα του Εθνικού Μετσόβιου Πολυτεχνείου Πυρφόρος για την ερευνητική ομάδα Προμηθέας. Το όχημα είναι ένα ελαφρύ μονοθέσιο βάρους 39 κιλών, εξοπλισμένο με ένα ιδιοκατασκευασμένο σύγχρονο κινητήρα εναλλασσόμενου ρεύματος και είναι ικανό να αναπτύξει ταχύτητες έως 50 χιλιόμετρα την ώρα. Το όχημα διαγωνίζεται κάθε χρόνο στο πανευρωπαϊκό μαραθώνιο οικονομίας που διοργανώνει η Shell, Shell Eco Marathon, και πληροί όλους τους κανονισμούς του διαγωνισμού.

Εικόνα 1.1: Το όχημα Προμηθέας στη πίστα του μαραθωνίου οικονομίας της Shell 2016

1.2 Διάρθρωση εργασίας

Η εργασία οργανώνεται σε κεφάλαια με την ακόλουθη δομή:

- **Κεφάλαιο 2**: Γίνεται μία παρουσίαση του ηλεκτροκίνητου οχήματος για το οποίο υλοποιήται το παρών σύστημα και αναλύεται εν συντομία το θεωρητικό υπόβαθρο που αφορά την οικολογική στρατηγική κίνησης.

- **Κεφάλαιο 3**: Σε αυτό το κεφάλαιο γίνεται μία σύντομη παρουσίαση των θεωρητικών εννοιών των νευρωνικών δικτύων και των μεθοδολογιών που χρησιμοποιήθηκαν.

- **Κεφάλαιο 4**: Πραγματοποιείται μία εκτενής ανάλυση όλων των τεχνητών νευρωνικών δικτύων που αναπτύχθηκαν στα πλαίσια αυτής της διπλωματικής εργασίας με στόχο τη μοντελοποίηση της κατανάλωσης του πρότυπου ηλεκτρικού οχήματος Πυρφόρος.

- **Κεφάλαιο 5**: Σε αυτό το κεφάλαιο παρουσιάζεται η μεθοδολογία εύρεσης των βέλτιστων ακολουθιών ταχυτήτων που οφείλει να διατηρήσει το όχημα προκειμένου να δαπανήσει την ελάχιστη δυνατή ενέργεια όταν εκτελεί μία συγκεκριμένη διαδρομή.

- **Κεφάλαιο 6**: Χρησιμοποιώντας μεθοδολογίες τεχνητής νοημοσύνης υπολογίζεται για μία συγκεκριμένη διαδρομή ο βέλτιστος τρόπος μετακίνησης ενός οχήματος, δηλαδή η αγωνιστική γραμμή.

- **Κεφάλαιο 7**: Παρουσιάζονται όλα τα ηλεκτρονικά στοιχεία που χρησιμοποιήθηκαν προκειμένου να υλοποιηθεί η μετάδοση των πληροφοριών στον οδηγό.

- **Κεφάλαιο 8**: Αναλύεται το διαδραστικό λογισμικό που αναπτύχθηκε και η διαδικασία παιχνιδοποίησης μέσω της οποίας ενημερώνεται ο οδηγός για το βέλτιστο τρόπο μετακίνησης πάνω στη πίστα.

- **Κεφάλαιο 9**: Συνοψίζονται τα συμπεράσματα και η συνεισφορά της εργασίας και προτείνονται διορθώσεις και επεκτάσεις.

- **Κεφάλαιο 10**: Παρουσιάζεται η βιβλιογραφία που στηρίχθηκε η παρούσα εργασία.

Κεφάλαιο 2

Τρόποι μείωσης της κατανάλωσης των αυτοκινήτων

2.1 Εισαγωγή

Τα τελευταία χρόνια εξαιτίας της αύξησης της θερμοκρασίας και αλλαγής του κλίματος της Γης η ανθρωπότητα έχει ξεκινήσει να στρέφεται προς μια πιο οικολογική κατεύθυνση σε όλους τους τομείς. Αυτό γίνεται φανερό παρατηρώντας τους νέους και πιο αυστηρούς νόμους που θεσπίζουν οι χώρες και τις σοβαρές ποινές που επιρρίπτουν στους παραβάτες. Στο τομέα της αυτοκίνησης οι κατασκευαστές αλλά και οι καταναλωτές επιθυμούν να μειώσουν τη κατανάλωση του οχήματος τους για οικολογικούς σκοπούς αλλά και για οικονομικούς.

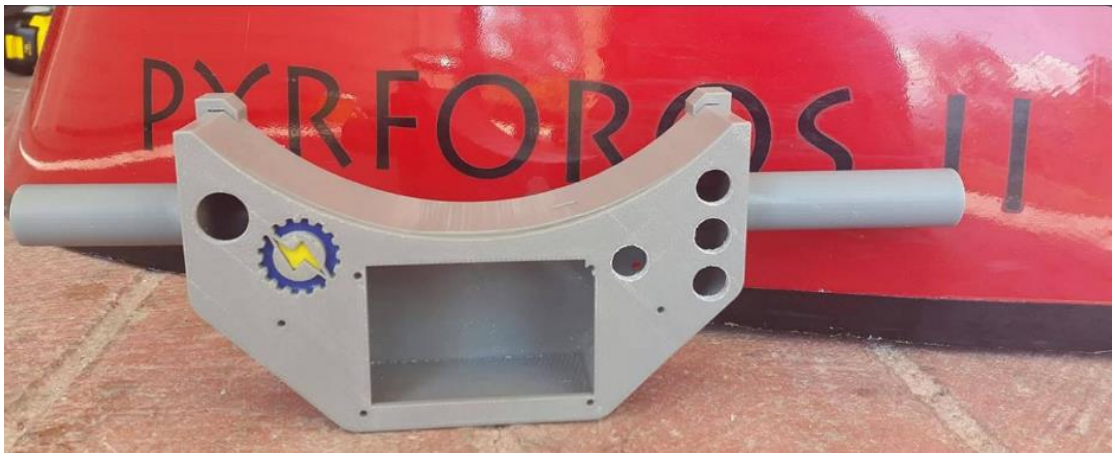
Σε αυτή την ενότητα γίνεται μία σύντομη παρουσίαση των ασκούμενων δυνάμεων πάνω στο πρότυπο ηλεκτρικό όχημα για το οποίο σχεδιάστηκε το παρουσιαζόμενο σύστημα, μία περιγραφή των χαρακτηριστικών του και τέλος γίνεται μία ανάλυση της μεθοδολογίας εξοικονόμησης ενέργειας μέσω της πρόβλεψης των τοπογραφικών συνθηκών της διαδρομής.

2.2 Παρουσίαση οχήματος και ασκούμενων δυνάμεων

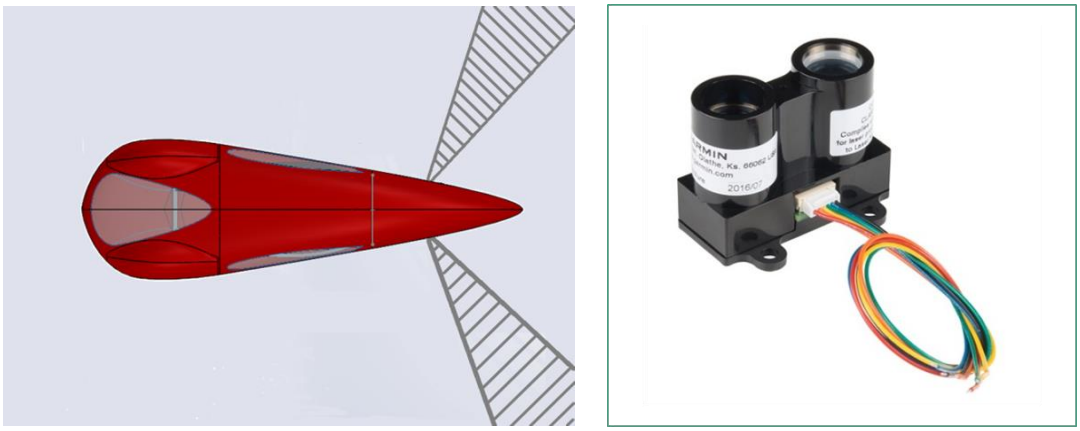
Το ηλεκτροκίνητο όχημα Πυρφόρος αποτελεί ένα πρότυπο μονοθέσιο όχημα που ζυγίζει μόλις 39 κιλά και είναι ικανό να φτάσει ταχύτητες έως 50 χιλιόμετρα την ώρα. Το αυτοκίνητο έχει κατασκευαστεί σύμφωνα με τους κανονισμούς του διαγωνισμού που συμμετάσχει, το πανευρωπαϊκό μαραθώνιο οικονομίας της Shell. Ο θεσμός αυτός έχει ως στόχο την ανάδειξη νέων τεχνολογιών με γνώμονα την μείωση της καταναλισκόμενης ενέργειας ενός αυτοκινήτου. Η ιστορία του διαγωνισμού

ξεκίνησε το 1939, όταν μία ομάδα επιστημόνων της Shell με έδρα τους ένα εργαστήριο στο Wood River του Ιλινόις των ΗΠΑ στοιχημάτισαν ποιος μπορεί να οδηγήσει το δικό του αυτοκίνητο μακρύτερα με ποσότητα καυσίμου μόλις 1 γαλόνι.

Προκειμένου να υλοποιηθεί το παρουσιαζόμενο σύστημα έγιναν μετατροπές στο όχημα για να μπορέσει να το φιλοξενήσει. Αυτές οι μετατροπές αφορούσαν τη σχεδίαση και εκτύπωση, σε τρισδιάστατο εκτυπωτή, του τιμονιού στο οποίο εγκαθίσταται το σύστημα και την τοποθέτηση αισθητήρων τεχνολογίας lidar για να εντοπίζεται με ακρίβεια η θέση του πάνω στο οδόστρωμα, περισσότερα για το υλικό παρουσιάζονται στο κεφάλαιο 8.



Εικόνα 2.1: Το τιμόνι του οχήματος που στεγάζει το σύστημα που παρουσιάζεται.

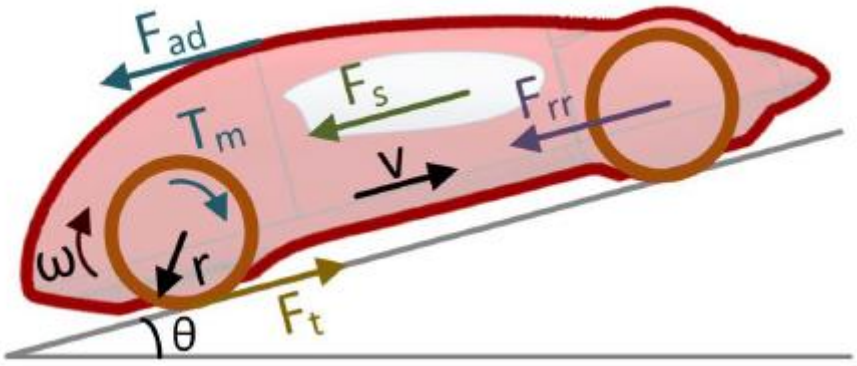


Εικόνα 2.2: Οι αισθητήρες Lidar που χρησιμοποιήθηκαν και τα σημεία τοποθέτησης τους.

	Μέγεθος	Τιμή
Κινητήρας	Μήκος (mm)	35
	Ονομαστική ροπή (Nm)	7.5
	Ονομαστική ταχύτητα περιστροφής (rpm)	300
	Αποδοτικότητα (%)	83.6
	Ολική μάζα (kg)	2.81
Όχημα	Μέγεθος ελαστικών	45/75 R 16
	Ακτίνα Τροχού (mm)	239
	Βάρος (kg)	39
	Μετάδοση	1:1
	Τελική Ταχύτητα (km/h)	50

Πίνακας 1.1: Χαρακτηριστικά οχήματος

Το όχημα έχει τρεις τροχούς, δύο μπροστά που είναι υπεύθυνοι για τη διεύθυνση του οχήματος και ένας πίσω που οδηγεί. Οι δυνάμεις που ασκούνται στο όχημα παρουσιάζονται παρακάτω και αφορούν την αντίσταση του αέρα (Fad), την τριβή των ελαστικών (Ft), η αντίσταση κύλισης (Frr) και τέλος το βάρος του οχήματος (Fs). Το μεγάλο πλεονέκτημα της χρήσης ενός νευρωνικού δικτύου για τη μοντελοποίηση του οχήματος είναι πως όλες αυτές οι δυνάμεις υπολογίζονται έμμεσα χωρίς τη χρήση Νευτώνειας μηχανικής.



Εικόνα 2.3: Ανάλυση δυνάμεων οχήματος.

2.2 Εξοικονόμηση ενέργειας μέσω της πρόβλεψης των τοπογραφικών χαρακτηριστικών

Τα σύγχρονα αυτοκίνητα είναι σύνηθες να εξοπλίζονται με υπολογιστές ταξιδίου που όταν ο οδηγός θέτει μία επιθυμητή ταχύτητα το όχημα αναλαμβάνει την επιτάχυνση και την

επιβράδυνση έτσι ώστε να την επιτύχει. Ωστόσο συστήματα όπως αυτό μπορεί πολλές φορές να επιταχύνουν απότομα το όχημα σε ανηφορικά οδοστρώματα και να επιβραδύνουν σε κατηφορικά προκειμένου να διατηρήσουν μία σταθερή ταχύτητα. Βέβαια οι συνεχείς απότομες επιταχύνσεις και επιβραδύνσεις είναι αποδεδειγμένο πως οδηγούν σε αύξηση της κατανάλωσης.

Μία διαφορετική στρατηγική μετακίνησης που εκμεταλλεύεται τα τοπογραφικά χαρακτηριστικά μπορεί να οδηγήσει σε μείωση της συνολικής κατανάλωσης χωρίς να μεταβληθεί η μέση διατηρούμενη ταχύτητα. Η στρατηγική αυτή μπορεί να υλοποιηθεί εάν τα τοπογραφικά χαρακτηριστικά της διαδρομής είναι γνωστά. Πιο συγκεκριμένα η αναμενόμενη συμπεριφορά είναι να εκμεταλλευτεί το όχημα τις κατηφορικές κλίσεις ούτως ώστε να έχει αναπτύξει μία μεγάλη ταχύτητα στην αρχή της ανηφόρας και να μη χρειαστεί να επιταχύνει για να διατηρήσει τη μέση ταχύτητα, αντιθέτως να επιβραδύνει [1].

Σύμφωνα με τα παραπάνω προκειμένου το σύστημα να υπολογίσει μία ακολουθία ταχυτήτων η οποία θα οδηγήσει σε μείωση της καταναλισκόμενης ενέργειας πρέπει να ακολουθήσει τη παραπάνω στρατηγική μετακίνησης.

Στη συνέχεια ακολουθεί μία θεωρητική περιγραφή των τεχνιτών νευρωνικών δικτύων και έπειτα γίνεται μία περιγραφή του τρόπου μοντελοποίησης της καταναλισκόμενης ενέργειας του οχήματος.

Κεφάλαιο 3

Νευρωνικά Δίκτυα

3.1 Εισαγωγή

Με τον όρο τεχνητό νευρωνικό δίκτυο (ΤΝΔ) εννοούμε κάθε αρχιτεκτονική υπολογισμού, η οποία περιλαμβάνει έναν μεγάλο αριθμό διασυνδεδεμένων νευρωνικών επεξεργαστών και απομιμείται τη λειτουργεία και τις δυνατότητες του ανθρώπινου εγκεφάλου. Συνεπώς, ένα τεχνητό νευρωνικό δίκτυο ή απλά ένα νευρωνικό δίκτυο (ΝΔ) έχει την ικανότητα να μαθαίνει από εμπειρίες, να γενικεύει την υπάρχουσα γνώση και να εκτελεί λογικές αφαιρέσεις [2].

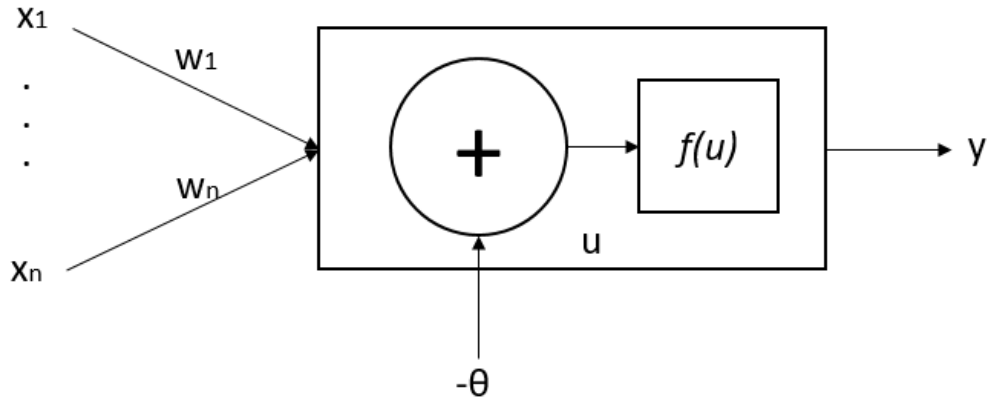
Τα ΝΔ είναι συστήματα μεγάλης κλίμακας, τα οποία περιέχουν έναν μεγάλο αριθμό μη γραμμικών επεξεργαστών ειδικού τύπου, τα νευρώνια. Κάθε ΝΔ χαρακτηρίζεται από μία κατάσταση, ένα σύνολο εισόδων με βάρη, που προέρχονται από άλλα νευρώνια, και μία εξίσωση ενεργοποίησης, η οποία περιγράφει τη δυναμική λειτουργεία του. Τα βάρη του ΝΔ ανανεώνονται μέσω μίας διαδικασίας μάθησης (εκπαίδευσης), η οποία πραγματοποιείται με την ελαχιστοποίηση κάποιας συνάρτησης κόστους. Οι βέλτιστες τιμές των βαρών αποθηκεύονται ως δυνάμεις μεταξύ των νευρώνων και χρησιμοποιούνται κατά την εκτέλεση της εργασίας για την οποία ορίστηκε το ΝΔ.

Σκοπός του κεφαλαίου αυτού είναι η συνοπτική παρουσίαση βασικών εννοιών των ΝΔ αρχίζοντας με τη περιγραφή απλών αρχιτεκτονικών και στη συνέχεια αναλύονται οι τεχνικές που χρησιμοποιήθηκαν για τη μοντελοποίηση του τρόπου κατανάλωσης ενός ηλεκτροκίνητου αυτοκινήτου.

3.1 Το δίκτυο Perceptron

Το πιο απλό νευρωνικό δίκτυο που μπορεί να σχεδιαστεί και να μελετηθεί είναι ένα δίκτυο που αποτελείται από ένα μόνο νευρώνα. Η λέξη ‹‹δίκτυο›› σε αυτή τη

περίπτωση χρησιμοποιείται καταχρηστικά αφού δεν υπάρχουν περισσότεροι του ενός νευρώνα για να συνδεθούν μεταξύ τους. Οι μόνες συνδέσεις που υπάρχουν είναι αυτές μεταξύ των εισόδων $x_1, x_2, \ldots, x_n$ και του νευρώνα, όπως φαίνεται στο σχήμα 3.1.



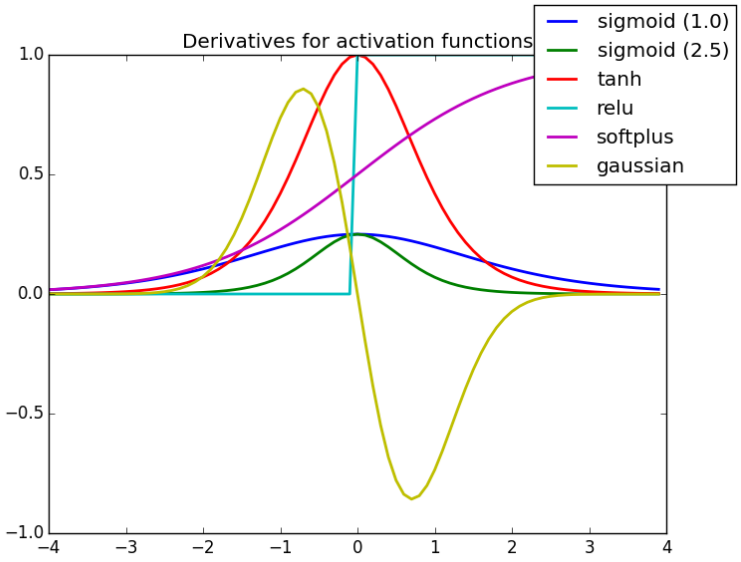
Εικόνα 3.1: Απλοποιημένο μοντέλο τεχνητού νευρώνα.

Στο μοντέλο Perceptron ο ένας και μοναδικός νευρώνας υλοποιεί την παρακάτω συνάρτηση μεταφοράς:

$$u = \sum_{i=1}^{n} w_i x_i - \theta$$
$$y = f(u)$$

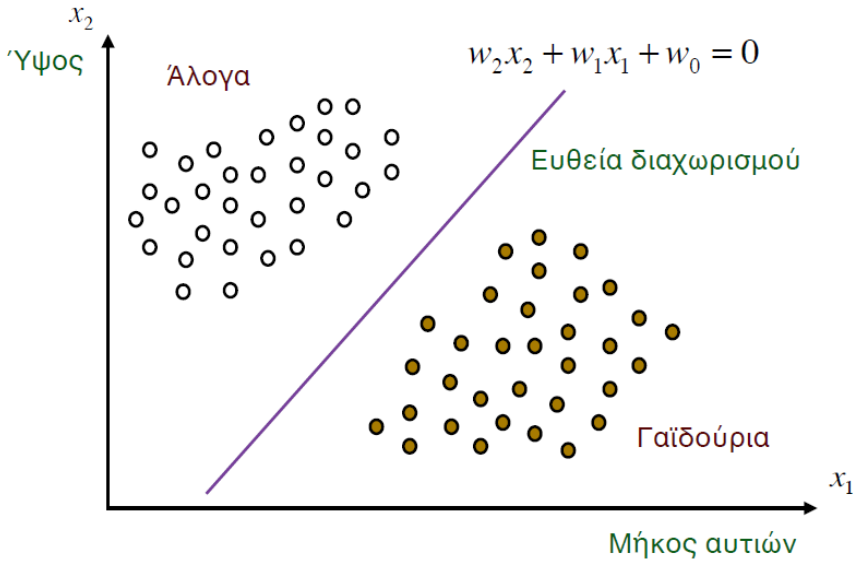
Η συνάρτηση μεταφοράς απεικονίζει το διάνυσμα εισόδου $x = [x_1, x_2, \ldots, x_n]^T$ στην έξοδο y. Οι παράμετροι $w_1, w_2, \ldots, w_n$ είναι τα συνοπτικά βάρη του νευρώνα ενώ η παράμετρος θ λέγεται κατώφλι ενεργοποίησης. Ο όρος αυτός εξηγείται από το γεγονός ότι η διέγερση u του νευρώνα η οποία προκύπτει από τις εισόδους είναι θετική αν το άθροισμα $\sum_{i=1}^{n} w_i x_i$ ξεπεράσει το όριο θ.

Η συνάρτηση ενεργοποίησης f(.) τροφοδοτείται από τη διέγερση u και δίδει την έξοδο y του νευρώνα η οποία είναι θετική όταν ξεπερνάτε το κατώφλι ενεργοποίησης. Η συνάρτηση ενεργοποίησης έχει ως έξοδο είτε στη κλασική μορφή (1/0) ή στη διπολική μορφή (1/-1). Η συγκεκριμένη συνάρτηση ονομάζεται βηματική αλλά δεν είναι η μοναδική, υπάρχει ένα μεγάλο πλήθος συναρτήσεων με διαφορετικά χαρακτηριστικά και ιδιότητες όπως παρουσιάζονται στην εικόνα 3.2. Τέλος οι παράμετροι που ρυθμίζουν ουσιαστικά τη συμπεριφορά του νευρώνα είναι το διάνυσμα των συνοπτικών βαρών $w = [w_1, w_2, \ldots, w_n]$ που μεταβάλλεται κατά τη διαδικασίας της εκπαίδευσης και το κατώφλι θ.



Εικόνα 3.2: Συνήθεις συναρτήσεις ενεργοποίησης.

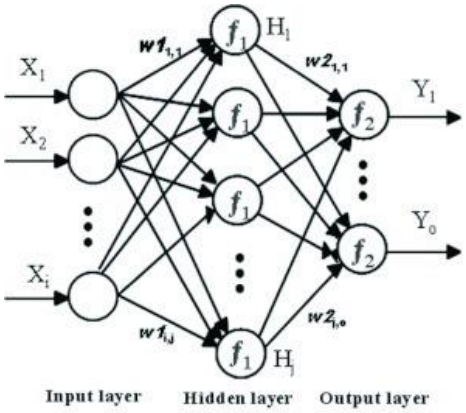
Ένα Perceptron είναι ικανό παρά την απλότητα του σχεδιασμού του να διαχωρίσει αντικείμενα σε δύο κλάσεις εάν υπάρχει μία ευθεία διαχωρισμού τέτοια ώστε να τις χωρίζει πλήρως. Τότε λέμε πως οι κλάσεις είναι γραμμικά διαχωρίσιμες. Το Perceptron ωστόσο δεν είναι δυνατό να διαχωρίσει πιο πολύπλοκες διατάξεις συνόλων [3].



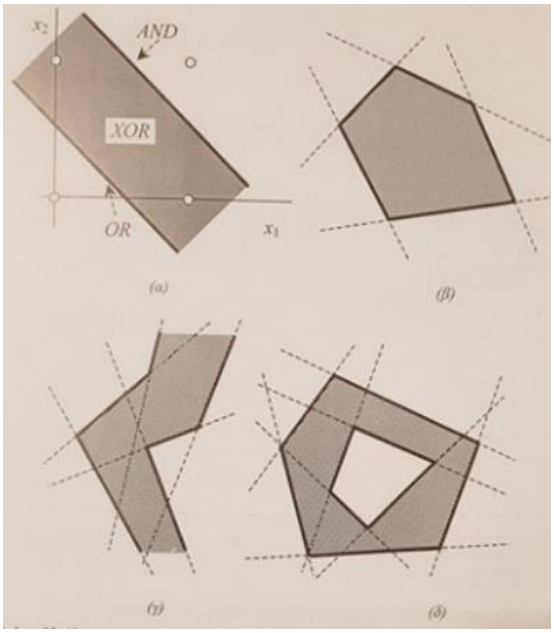
Εικόνα 3.3: Δύο γραμμικά διαχωρίσιμες κλάσεις. Για το συγκεκριμένο πρόβλημα ταξινόμησης μπορεί να υλοποιηθεί ένα δίκτυο Perceptron που με τα κατάλληλα βάρη να εκχωρεί τα στοιχεία στη κλάση που ανήκουν.

3.2 Το Μοντέλο Perceptron Πολλών Επιπέδων – Multilayer Perceptron (MLP)

Οι δυνατότητες αναπαράστασης διαχωριστικών επιφανειών είναι περιορισμένες στο δίκτυο Perceptron καθώς με ένα μόνο νευρώνα το δίκτυο μπορεί να αναπαραστήσει μόνο επίπεδες επιφάνειες. Ο περιορισμός αυτός αίρεται με τη χρήση περισσότερων νευρώνων. Το βασικό χαρακτηριστικό σε ένα δίκτυο Perceptron πολλών στρωμάτων είναι ότι κάθε στρώμα l τροφοδοτείται από τους νευρώνες του προηγούμενου στρώματος l-1 και μόνον από αυτούς. Δίκτυα που λειτουργούν με αυτό τον τρόπο καλούνται δίκτυα *προσοτροφοδότησης*. Το πρώτο στρώμα ονομάζεται στρώμα εισόδου ενώ το τελευταίο εξόδου, όλα τα ενδιάμεσα στρώματα καλούνται κρυμμένα επίπεδα. [3].



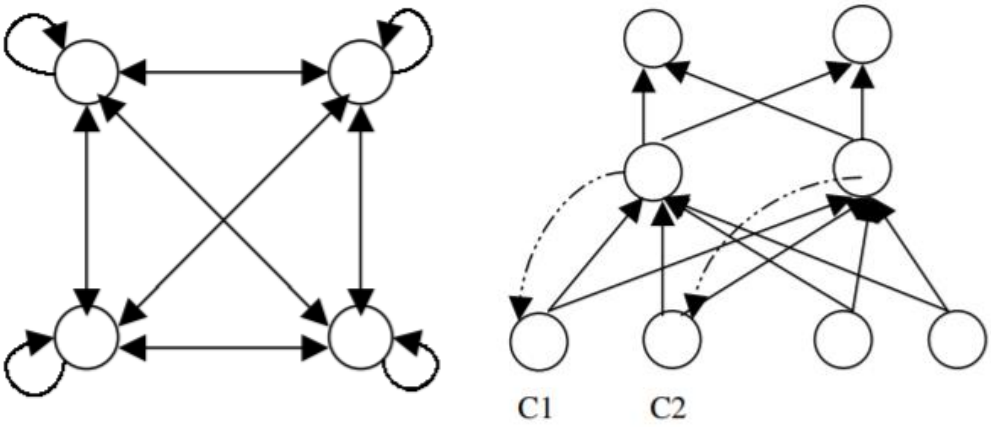
Εικόνα 3.3: Αναπαράσταση ενός πλήρους MLP δικτύου με ένα κρυφό επίπεδο. Ονομάζεται πλήρες διότι όλοι οι νευρώνες του επιπέδου i ενώνονται με του επιπέδου i+1.



Εικόνα 3.4: Χρησιμοποιώντας πολλούς νευρώνες στο κρυφό στρώμα με το κατάλληλο συνδυασμό μπορούμε να σχεδιάσουμε πολλές διαφορετικές περιοχές που διαχωρίζονται από τον υπόλοιπο χώρο.

3.3 Ανατροφοδοτούμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks, RNN)

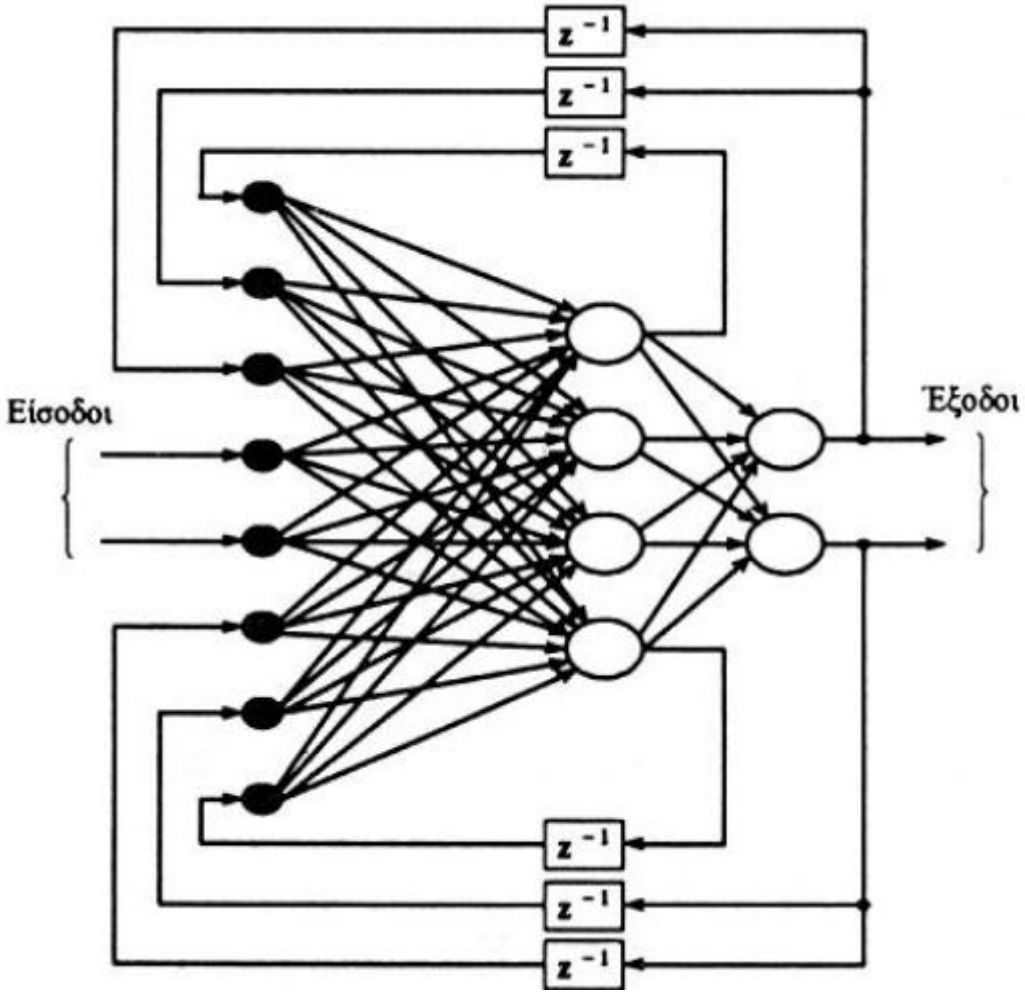
Τα ανατροφοδοτούμενα νευρωνικά δίκτυα είχαν βρεθεί στο επίκεντρο ερευνών κατά τη δεκαετία 1990-2000. Είναι σχεδιασμένα ώστε να εκπαιδεύονται σε ακολουθίες δεδομένων και να εντοπίζουν επαναλαμβανόμενα μοτίβο. Αυτή η τακτική πλησιάζει περισσότερο τον τρόπο που σκέφτονται οι άνθρωποι. Καθώς ο αναγνώστης διαβάζει αυτό το κείμενο δεν αντιλαμβάνεται κάθε λέξη ξεχωριστά αλλά αντιμετωπίζει κάθε πρόταση ως ακολουθία συμβολοσειρών και εξάγει το συνολικό νόημα σύμφωνα με ό,τι έχει διαβάσει στο παρελθόν. Τα απλά ΝΔ δεν είναι ικανά να συμπεριφερθούν με αυτό το τρόπο. Στα ανατροφοδοτούμενα ΝΔ τα κρυφά επίπεδα ανατροφοδοτούν τις εξόδους τους στις εισόδους των προηγουμένων. Με αυτό τον τρόπο αναπαρίσταται η πληροφορία του παρελθόντος και διαδίδεται χωρίς να χάνεται. Οι αρχιτεκτονικές των δικτύων ποικίλουν από πλήρως συνδεδεμένα δίκτυα σε μερικώς [8].



Εικόνα 3.5: Η αριστερή εικόνα αναπαριστά ένα πλήρως συνδεδεμένο ανατροφοδοτούμενο δίκτυο ενώ η δεξιά ένα δίκτυο μερικής συνδεσμολογίας.

Πιο συγκεκριμένα τα αποτελέσματα της ανατροφοδότησης είναι ότι, όταν ένα διάνυσμα (πρότυπο) εισόδου εισέρχεται στο αναδρομικό ΝΔ, δεν παράγει ένα πρότυπο εξόδου σε πεπερασμένο αριθμό χρονικών βημάτων, αλλά δρα με κυκλικό τρόπο, όπου τα ίδια στρώματα ενεργοποιούνται επαναληπτικά. Εάν το ΝΔ είναι αφ' εαυτού ευσταθές, πιθανώς να ταλαντωθεί για κάποιο χρονικό διάστημα προτού φτάσει σε μία σταθερή κατάσταση, στην οποία οι νευρωνικές ενεργοποιήσεις σταματήσουν να αλλάζουν με αποτέλεσμα να παραχθεί μία σταθερή έξοδος. Διαφορετικά εάν το ΝΔ δεν είναι ευσταθές οι ταλαντώσεις θα συνεχιστούν αδιάκοπα.

Ένα παράδειγμα μονοστρωματικού αναδρομικού ΝΔ απεικονίζεται παρακάτω. Οι βρόχοι ανατροφοδότησης κλείνουν μέσω μοναδιαίων καθυστερήσεων, οι οποίες συμβολίζονται με z^{-1}, όπου $z^{-1} \times y(k) = y(k-1)$ και k αναπαριστά το διακριτό χρόνο. Στη θεωρία συστημάτων το z^{-1} ονομάζεται τελεστής μοναδιαίας καθυστέρησης. [2]



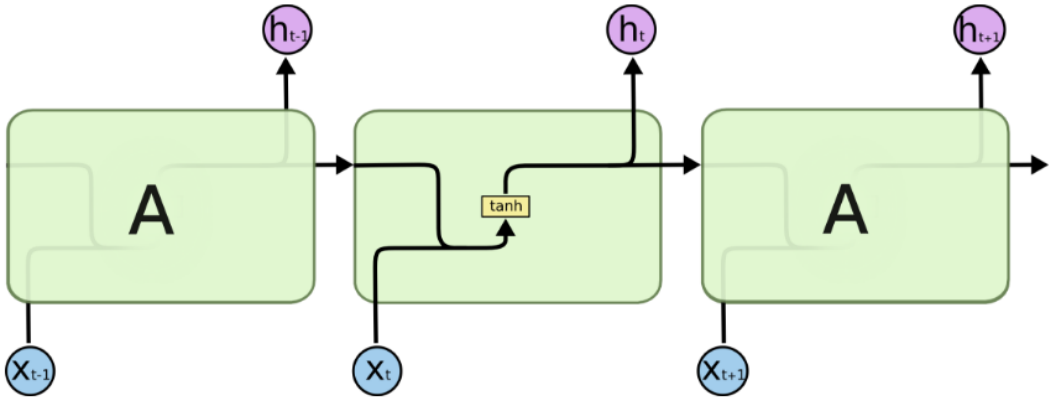
Εικόνα 3.6: Αναπαράσταση ενός πλήρους ανατροφοδοτούμενου ΝΔ όπου οι τιμές των εισόδων τροφοδοτούνται ως είσοδοι στο δίκτυο έως ότου το δικτύου συγκλίνει σε κάποια σταθερή τιμή.

Το πρόβλημα των απλών ανατροφοδοτούμενων ΝΔ είναι πως παρόλο που στη θεωρία είναι ικανά να αναγνωρίσουν μακροπρόθεσμες συσχετίσεις στη πραγματικότητα όμως δεν μπορούν όπως απέδειξαν οι Hochreiter [4] και Bengio [5].

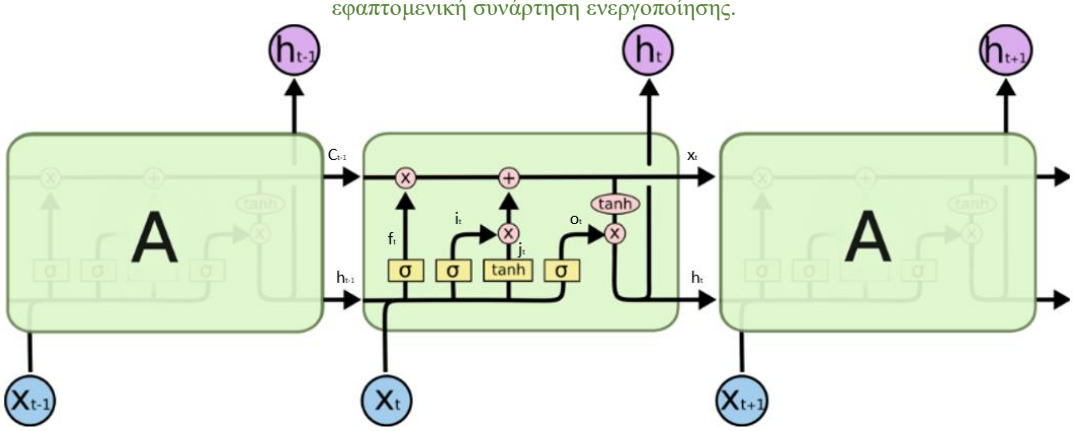
3.3.1 Ανατροφοδοτούμενα Δίκτυα Μακράς Βραχυπρόθεσμης Μνήμης (Long Short Term Memory, LSTM)

Τα δίκτυα LSTM αποτελούν μία ιδική κατηγορία των ανατροφοδοτούμενων ΝΔ που περιεγράφηκε. Πρωτοπαρουσιάστηκαν από τους Hochreiter & Schmidhuber το 1997 και μεταβλήθηκαν από πολλούς ερευνητές μέχρι σήμερα. Το πλεονέκτημα των LSTMs σε σχέση με τα απλά RNNs είναι πως μπορούν να εξάγουν εξαρτήσεις μεταξύ των δεδομένων που βρίσκονται σε μεγάλη απόσταση μέσα σε μία ακολουθία.

Τα κλασικά RNNs αναπαρίστανται ως αλυσίδα με μία πολύ απλή δομή. Τα LSTMs αποτελούν και αυτά μία αλυσίδα πυλών και πράξεων αλλά έχουν τέσσερα επίπεδα υπολογισμού.



Εικόνα 3.7: Αναπαράσταση του πυρήνα ενός απλού ανατροφοδοτούμενου νευρώνιου που αποτελείται από μία εφαπτομενική συνάρτηση ενεργοποίησης.



Εικόνα 3.8: Αναπαράσταση του πυρήνα ενός LSTM νευρώνιου.

Στην εικόνα 3.8 αναπαρίσταται ο πυρήνας ενός νευρώνα LSTM όπου κάθε γραμμή είναι ικανή να μεταφέρει ένα πολυδιάστατο διάνυσμα από την είσοδο στην έξοδο αφού το μεταβάλει. Η επάνω οριζόντια γραμμή μεταφοράς αποτελεί το κύριο μέρος του πυρήνα η οποία ασκεί μόνο γραμμικές μεταβολές στο διάνυσμα εισόδου. Το πρώτο βήμα του νευρώνα είναι η επιλογή της πληροφορίας που θα ξεχαστεί στο επόμενο. Αυτή τη απόφαση γίνεται από ένα σιγμοειδές επίπεδο που καλείται «forget gate layer». Η υλοποίηση γίνεται από τη συνάρτηση $f_t = \sigma(W_f[h_{t-1}, x_t] + b)$ μέσω της οποίας ελέγχονται οι είσοδοι h_{t-1} και x_t για κάθε στοιχείο της κατάστασης εισόδου C_{t-1}και παράγεται μία έξοδος που παίρνει τιμές εντός του εύρους [0,1] όπου η τιμή 1 σημαίνει πως τα δεδομένα κρατιόνται αναλλοίωτα και το 0 πως ξεχνιόνται, η επιλογή αυτή εξαρτάται από τη σταθερά b. Το επόμενο βήμα εισάγει τη νέα πληροφορία στο στα δεδομένα. Αυτό γίνεται σε δύο στάδια με τη βοήθεια δύο συναρτήσεων:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b)$$

&

$$j_t = tanh(W_C[h_{t-1}, x_t] + b)$$

Η πρώτη αποτελεί μία σιγμοειδής συνάρτηση μέσω της οποίας επιλέγεται η πληροφορία που θα διατηρηθεί ενώ η δεύτερη δημιουργεί ένα διάνυσμα που διευκολύνει τη πράξη της πρόσθεσης που έπεται με την έξοδο του γινομένου $C_{t-1} \times f_t$.Τέλος ακολουθεί η επιλογή της εξόδου του νευρώνα η οποία αποτελεί μία φιλτραρισμένη εκδοχή της εισόδου. Πρώτα χρησιμοποιούμε μία σιγμοειδής

συνάρτηση προκειμένου να αφαιρέσουμε ένα κομμάτι πληροφορίας και έπειτα διατηρούμε τα αποτελέσματα στο εύρος [-1,1] με τη χρήση της εφαπτομενικής συνάρτησης (tanh).

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

Η παραπάνω περιγραφή αφορά μία απλή εκδοχή ενός νευρώνα LSTM, προφανώς κατά τη πάροδο των χρόνων δημιουργήθηκαν και άλλες διατάξεις πυρήνων που εφαρμόζονται σε συγκεκριμένες εργασίας όπως είναι η ανάλυση φυσικής γλώσσας. Τα δίκτυα LSTM υπερτερούν στην εκμάθηση πολύπλοκων ακολουθιακών δεδομένων σε σχέση με τα απλά RNNs. Λόγο αυτής της ανωτερότητας επιλέχθηκαν προκειμένου να αναπαραστήσουν το μοντέλο κατανάλωσης που περιγράφεται στη συνέχεια. [7].

3.4 Εκπαίδευση Νευρωνικών Δικτύων

Τα ΝΔ πραγματοποιούν δύο βασικές λειτουργίες: την εκπαίδευση (ή μάθηση), και την ανάκληση. Μάθηση είναι η διαδικασία τροποποίησης της τιμής των βαρών του δικτύου, ώστε δοθέντος συγκεκριμένου διανύσματος εισόδου να παραχθεί συγκεκριμένο διάνυσμα εξόδου. Ανάκληση είναι η διαδικασία υπολογισμού ενός διανύσματος εξόδου για συγκεκριμένο διάνυσμα εισόδου και συγκεκριμένες τιμές [8]. Ο γενικός τρόπος με τον οποίο γίνεται η τροποποίηση των βαρών ενός ΝΔ κατά την εκπαίδευση του, επιτρέπει τη διάκριση τριών ειδών μάθησης, την επιβλεπόμενη, την ενισχυτική και τη μη επιβλεπόμενη. Η ενισχυτική μάθηση χρησιμοποιεί ένα συνολικό μέτρο επάρκειας της προκύπτουσας απόκρισης (δείκτης συμπεριφοράς) σε συνδυασμό με τα διανύσματα εισόδου, το οποίο μπορεί να οδηγήσει το δίκτυο στην επιθυμητή συμπεριφορά. Από την άλλη μεριά στη μη επιβλεπόμενη μάθηση, που καλείται και αυτό-οργανούμενη μάθηση δε χρησιμοποιείται εξωτερικός δάσκαλος, το δίκτυο οργανώνει τον πίνακα των βαρών του με τέτοιον τρόπο, ώστε να αναγνωρίζει όποια κανονικότητα υπάρχει στα διανύσματα εισόδου. Στη συνέχεια παρουσιάζεται πιο αναλυτικά η διαδικασία της επιβλεπόμενης μάθησης σε δίκτυα προσοτροφοδότησης και ανατροφοδότησης. Δεν γίνεται περεταίρω ανάπτυξη των υπόλοιπων μεθόδων εκπαίδευσης διότι δεν χρησιμοποιήθηκαν για την ανάπτυξη του συστήματος, στο κεφάλαιο 5 βέβαια αναλύονται οι έννοιες της ενισχυτικής εκπαίδευσης στο πλαίσιο αλγορίθμων εκμάθησης χωρίς τη χρήση ΝΔ.

3.4.1 Εκπαίδευση Δικτύων Προσοτροφοδότησης

Αυτό που κάνει την εκπαίδευση ενός MLP πολύ ενδιαφέρουσα είναι η ικανότητα του να προσεγγίσει οποιαδήποτε συνάρτηση δεδομένου ότι έχουμε χρησιμοποιήσει το κατάλληλο μέγεθος δικτύου. Η παραπάνω ιδιότητα δικαιολογεί τη

δημοτικότητα των αλγορίθμων εκπαίδευσης των MLP με κυριότερο εκπρόσωπο τον αλγόριθμο Back-Propagation.

Ο αλγόριθμος Back-Propagation προτάθηκε από τον Paul Werbos στη δεκαετία του 1970 στα πλαίσια ανάλυσης μοντέλων οικονομικής και πολιτικής πρόβλεψης. Βασικό χαρακτηριστικό της μεθόδου αυτής είναι η ύπαρξη στόχων (d_m) και ο σκοπός του αλγορίθμου είναι να ταυτίσει τις εξόδους (y_m) του ΝΔ με αυτούς, δοθέντος των κατάλληλων εισόδων. Ωστόσο αυτό δεν είναι πάντα εφικτό οπότε επιζητούμε τη βέλτιστη προσέγγιση της επιθυμητής κατάστασης χρησιμοποιώντας ένα κριτήριο κόστους. Το μέσο τετραγωνικό σφάλμα

$$J = \frac{1}{P}\sum_{p=1}^{P}\sum_{i=1}^{m}[d_i^{(p)} - y_i^{(p)}]^2$$

είναι ένα κλασικό κριτήριο κόστους που χρησιμοποιήται ευρέως σε πολλά προβλήματα, όπου P είναι η διάσταση του διανύσματος εισόδου και m το πλήθος των εξόδων. Έχει πλεονέκτημα ότι ελαχιστοποιεί τη τετραγωνική απόσταση των ποσοτήτων d & y και επιπλέον παραγωγίζεται οπότε μπορεί να χρησιμοποιηθεί σε μεθόδους κατάβασης δυναμικού και να σώσει συμπαγείς μαθηματικές φόρμουλες. Αυτή ακριβώς τη κατεύθυνση ακολουθεί αυτός ο αλγόριθμος. Σύμφωνα με τη μέθοδο κατάβασης δυναμικού, η μεταβολή της παραμέτρου w_{ij} ως προς το χρόνο t γίνεται με τη χρήση της παραγώγου J ως προς w_{ij}, η οποία ονομάζεται και κλίση του J ως προς w_{ij}.

$$\frac{dw_{ij}}{dt} = -\frac{dJ}{dw_{ij}} = \delta$$

- w_{ij} (l,k) το συνοπτικό βάρος που συνδέει το νευρώνα j του στρώματος l-1 με το νευρώνα i του στρώματος l κατά τη χρονική στιγμή k.

Ο αλγόριθμος ξεκινά αρχικοποιώντας τα διανύσματα των βαρών με τυχαίες τιμές και υπολογίζει μέσω αυτών τις εξόδους. Η έξοδος ενός νευρώνα υπολογίζεται από την έκφραση:

$$y_i = \sum_{j=1}^{P} w_{ij} \times x_j$$

Έπειτα υπολογίζει το σφάλμα και σύμφωνα με μία σταθερά που ονομάζεται ρυθμός μάθησης (β) μεταβάλει τα βάρη σύμφωνα με την παρακάτω εξίσωση για κάθε νευρώνα.

$$w_{ij}(l,k+1) - w_{ij}(l,k) = -\beta\,\frac{dJ}{dw_{ij}(l,k)}$$

Η διαδικασία μεταβολής των βαρών ξεκινά από τις συνδέσεις του τελευταίου κρυφού επιπέδου με τη έξοδο και συνεχίζεται προς τα πίσω, λόγο αυτής της λειτουργίας προκύπτει και η ονομασία του αλγορίθμου. Η ενέργεια αυτή επαναλαμβάνεται όσες φορές χρειαστεί έως ότου το σφάλμα σταματήσει να μειώνεται, κάθε επανάληψη που

εκτελείται ονομάζεται εποχή. Μέσω της παραπάνω διαδικασίας ένα MLP είναι ικανό να κάνει έγκυρες προβλέψεις για ένα μεγάλο σύνολο προβλημάτων [3].

3.4.2 Εκπαίδευση Ανατροφοδοτούμενων Δικτύων

Στη προηγούμενη υποενότητα αναλύθηκε ο αλγόριθμος Back-Propagation για ΝΔ προσοτροφοδότησης κατά την εκτέλεση του οποίου το σφάλμα της εξόδου μπορούσε να μεταφερθεί προς τα πίσω σε προηγούμενα επίπεδα νευρώνων. Προκειμένου να εφαρμοστεί ο ίδιος αλγόριθμος σε δίκτυα ανατροφοδότησης πρέπει το σφάλμα να μεταφερθεί και στο πεδίο του χρόνου για να εντοπιστούν μακροπρόθεσμες εξαρτήσεις.

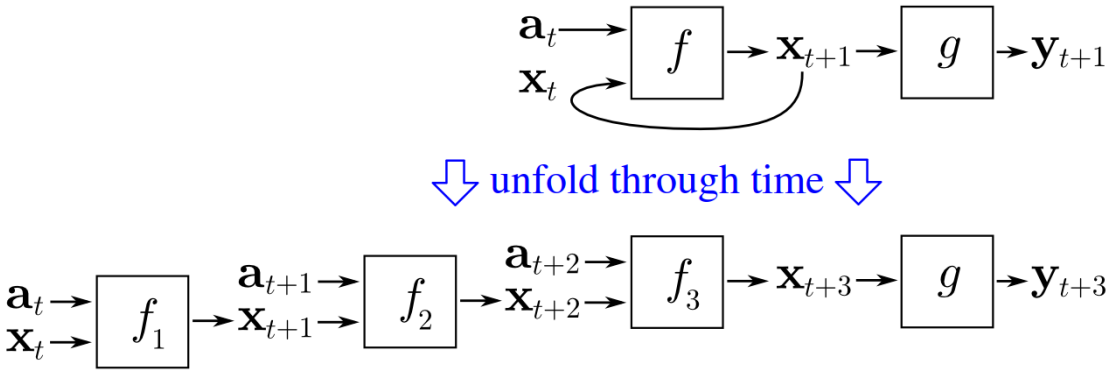
Προκειμένου να επιτευχθεί αυτό κάθε ανατροφοδοτούμενος νευρώνας εκδιπλώνεται (unfolding) στον χρόνο, εικόνα 3.9, έτσι ώστε το συνολικό δίκτυο να είναι ισοδύναμο με ένα πολυεπίπεδο δίκτυο εμπρόσθιας τροφοδότησης, ο αλγόριθμος που το επιτυγχάνει αυτό ονομάζεται Back Propagation Through Time.

Στο εκδιπλωμένο δίκτυο μπορούμε να εφαρμόσουμε τον αλγόριθμο Back Propagation για την εκπαίδευσή του, λαμβάνοντας όμως υπόψη πως τα βάρη W_a και W_x είναι τα ίδια στις ενώσεις των αντίστοιχων κόμβων για όσα επίπεδα και αν ξεδιπλώσουμε το δίκτυο, έστω τ. Έτσι υπολογίζουμε το σφάλμα για κάθε σύνδεση όπως και προηγουμένους. Έπειτα θεωρούμε ως συνολικό σφάλμα του νευρώνιου RNN ως το άθροισμα όλων των σφαλμάτων κάθε επιπέδου εκδίπλωσης του στο χρόνο.

$$\delta J = \sum_{i=1}^{\tau} J(i)$$

Έχοντας υπολογίσει το σφάλμα ανανεώνουμε τις τιμές των βαρών όπως προηγουμένους και συνεχίζουμε τη διαδικασία έως ότου σταθεροποιηθεί το σφάλμα.

Το ξεδίπλωμά, όμως στον χρόνο του δικτύου από ένα σημείο και έπειτα μπορεί να επιφέρει αποτελέσματα που επηρεάζουν αρνητικά την ικανότητα του δικτύου να μαθαίνει. Αυτό το φαινόμενο ονομάζεται φαινόμενο της εξαφανιζόμενης κλίσης (vanishing gradient) [8] και συμβαίνει μιας και όλο και περισσότερες παράμετροι συνεισφέρουν στο άθροισμά της διόρθωσης των βαρών με αποτέλεσμα να γίνεται δυσκολότερη η διαδικασία της εκπαίδευσης [9].



Εικόνα 3.9: Παρουσιάζεται ένα εκδιπλωμένο δίκτυο που φτάνει έως τρεις μονάδες χρόνου πίσω.

3.5 Εφαρμογές Νευρωνικών Δικτύων

Τα πλεονεκτήματα και τα μειονεκτήματα των ΝΔ είναι συχνά συμπληρωματικά με εκείνα των συμβατικών τεχνικών ανάλυσης δεδομένων, ενώ θα πρέπει να χρησιμοποιούνται για την επίλυση προβλημάτων, τα οποία έχουν ένα ή και περισσότερα από τα παρακάτω χαρακτηριστικά. [10]

1. Υπάρχει μεγάλο πλήθος δεδομένων για την εκπαίδευση του δικτύου.
2. Είναι δύσκολο να βρεθεί μία απλή και μοντελοποιημένη λύση.
3. Η επεξεργασία των νέων δεδομένων πρέπει να είναι ταχύτατη.
4. Η μέθοδος επεξεργασίας των δεδομένων πρέπει να είναι αποτελεσματική στην αντιμετώπιση του θορύβου στα δεδομένα εισόδου.

Τα ΝΔ έχουν χρησιμοποιηθεί σε πλήθος εφαρμογών, όπως είναι η σχεδιασμός ενεργειών (planning), ο χρονοπρογραμματισμός (scheduling), η διάγνωση λαθών σε δορυφορικές επικοινωνίες, στην ιατρική διάγνωση και ανάλυση ιατρικής εικόνας, στην ανάλυση αγοράς μετοχών αλλά και άλλες πολλές εφαρμογές.

Το πρόβλημα που αντιμετωπίζει η εργασία αυτή περιλαμβάνει και τα τέσσερα χαρακτηριστικά που περιεγράφηκαν παραπάνω και γι’ αυτό κρίθηκε σκόπιμη η χρησιμοποίηση τους για τη μοντελοποίηση της κατανάλωσης ενός ηλεκτροκίνητου οχήματος σε σχέση με μία διαδρομή και τη ταχύτητα που διατηρεί.

Κεφάλαιο 4

Μοντέλο Κατανάλωσης

4.1 Εισαγωγή

Η ανακάλυψη της ακολουθίας ταχυτήτων που θα προσφέρει την ελάχιστη κατανάλωση προϋποθέτει τη πρόβλεψη της καταναλισκόμενης ενέργειας πάνω σε μία γνωστή διαδρομή. Για να το πετύχει αυτό το σύστημα θα πρέπει να διαθέτει ένα μοντέλο κατανάλωσης του οχήματος. Το μοντέλο αυτό επιλέχθηκε να υλοποιηθεί από ένα τεχνητό νευρωνικό δίκτυο (ΤΝΔ). Αυτή η επιλογή αποσκοπεί στην απομάκρυνση του συστήματος από οποιοδήποτε θεωρητικό μοντέλο υλοποίησης που θα επέφερε σφάλματα λόγο παραμέτρων που θα ήταν αδύνατο να υπολογιστούν, όπως η τριβή των ελαστικών με το οδόστρωμα και οι απώλειες όλων των επιμέρους ηλεκτρονικών στοιχείων. Ένα ΤΝΔ μπορεί να εξετάσει το όχημα σαν σύνολο μελετώντας τη συνολική ενέργεια που καταναλώνει ανάλογα με τη θέση και τη ταχύτητα του.

Η επιλογή του δικτύου έγινε έπειτα από πολλά πειράματα όπου δοκιμάστηκαν διαφορετικές αρχιτεκτονικές και φιλοσοφίες ΤΝΔ. Προκειμένου ένα δίκτυο να χαρακτηριστεί ικανό και λειτουργικό πρέπει να είναι σε θέση να προβλέπει με ακρίβεια τη κατανάλωση του οχήματος λαμβάνοντας υπόψη όχι μόνο μία συγκεκριμένη στιγμή αλλά ένα μεγάλο κομμάτι της πορείας του οχήματος. Αυτό συμβαίνει διότι σε μία συγκεκριμένη θέση με την ίδια ταχύτητα ένα όχημα μπορεί να καταναλώνει διαφορετικά μεγέθη ενέργειας ανάλογα με τη ταχύτητα που είχε στη προηγούμενη θέση. Η ανάγκη αυτή οδήγησε την έρευνα προς τη κατεύθυνση των νευρωνικών δικτύων με ανατροφοδότηση (Recurrent Neural Networks, RNN) και ιδιαίτερα προς τα δίκτυα μακράς βραχυπρόθεσμης μνήμης (Long Short Term Memory, LSTM) λόγο της ικανότητας τους να βρίσκουν εξαρτίσεις δεδομένων που απέχουν πολύ μεταξύ τους [16].

Το επιλαχόν δίκτυο θα το χρησιμοποιήσει ο ευφυής πράκτορας για να δοκιμάσει στην ίδια διαδρομή διαφορετικές ακολουθίες ταχυτήτων ανακαλύπτοντας τη καταναλισκόμενη ενέργεια κάθε μίας. Από τα διαφορετικά σενάρια θα αποκομίζει εμπειρίες και γνώση η οποία θα του επιτρέψει να επιλέξει τη βέλτιστη ακολουθία. Στο κεφάλαιο 5 θα γίνει αναλυτική εξήγηση της μεθοδολογίας και του αλγορίθμου που χρησιμοποιήθηκε.

4.2 Προεπεξεργασία δεδομένων

4.2.1 Δεδομένα εισόδου-εξόδου

Το σύστημα υλοποιήθηκε προκειμένου να χρησιμοποιηθεί από το πρότυπο ηλεκτρικό όχημά της ομάδας Προμηθέας του Εθνικού Μετσόβιου Πολυτεχνείου. Το αυτοκίνητο διαθέτει μια ιδιοκατασκευασμένη πλακέτα τηλεμετρίας [11] η οποία είναι ικανή να μετρήσει το εναλλασσόμενο ρεύμα που καταναλώνει ο κινητήρας, τη τάση και το συνεχές ρεύμα των μπαταριών και τέλος από το Παγκόσμιο Σύστημα Στιγματοθέτησης (Global Positioning System, GPS), το υψόμετρο, τη ταχύτητα και τις συντεταγμένες. Το σύστημα τηλεμετρίας έχει συχνότητα δειγματοληψίας 50 hz. Τα δεδομένα που χρησιμοποιήθηκαν προήλθαν από τη παρουσία του οχήματος στο πανευρωπαϊκό μαραθώνιο οικονομίας Shell Eco Marathon 2016 (624.667 μετρήσεις)και 2017 (177.025 μετρήσεις) και από δοκιμές που πραγματοποιήθηκαν εντός του περιβάλλοντος του Εθνικού Μετσόβιου Πολυτεχνείου (75.756 μετρήσεις). Το σύνολο των δεδομένων απαρτίζεται από 877.448 μετρήσεις.

4.2.2 Μεθοδολογία προεπεξεργασίας

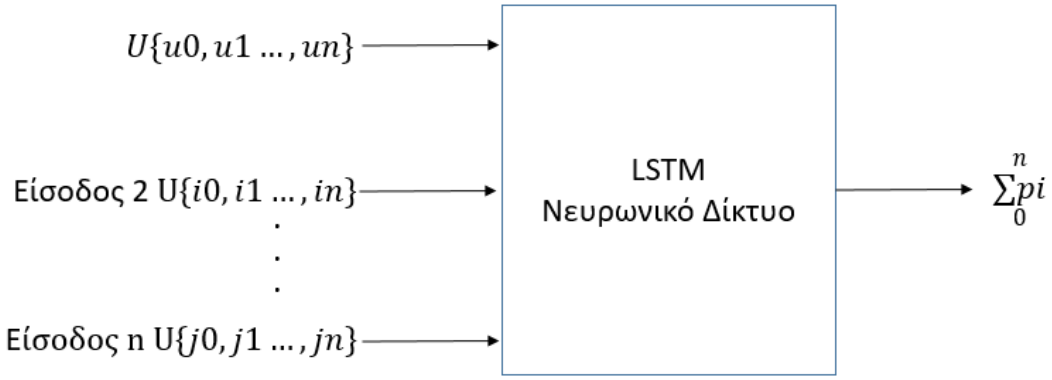
Προκειμένου να ευρεθεί το βέλτιστο ΤΝΔ που θα είναι ικανό να προβλέπει τη κατανάλωση του αυτοκινήτου σε μία άγνωστη διαδρομή πραγματοποιήθηκαν πειράματα με διαφορετικά δίκτυα και αρχιτεκτονικές. Πρωτού όμως τα δεδομένα εισαχθούν στο ΤΝΔ πρέπει να περάσουν ένα στάδιο επεξεργασίας που λέγεται προεπεξεργασία. Σε αυτό το στάδιο τα δεδομένα εισόδου και εξόδου διαμορφώνονται κατάλληλά προκειμένου το δίκτυο να μπορεί να τα διαβάσει και να διευκολυνθεί το στάδιο της εκπαίδευσης. Όλα τα δίκτυα που αναπτύχθηκαν περιγράφονται αναλυτικά στην επόμενη ενότητα.

Για τα δύο πρώτα δίκτυα πολλών επιπέδων (Multilayer Perceptron, MLP) που υλοποιήθηκαν χρησιμοποιήθηκε το περιβάλλον Matlab. Κατά τη διαδικασία προεπεξεργασίας αφαιρέθηκαν τα δεδομένα που σχετίζονταν με συντεταγμένες μηδενικών τιμών διότι σε αυτές τις περιπτώσεις το GPS ήταν αναξιόπιστο και η τιμή της ταχύτητας προέρχεται αυτό. Έπειτα σχηματίστηκαν τα διανύσματα εισόδου που αντιστοιχούσαν σε μία τιμή εξόδου. Τέλος με τη χρήση της υπάρχουσας συνάρτησης processpca με παράμετρο 0.0097 μειώθηκαν οι διαστάσεις των δεδομένων χωρίς να χαθεί σημαντική πληροφορία. Αυτό το πετυχαίνει χρησιμοποιώντας ένα ορθογώνιο μετασχηματισμό για να μετατρέψει ένα σύνολο συσχετιζόμενων τιμών σε γραμμικά ασυσχέτιστες. Αφού τελείωσε το στάδιο της προεπεξεργασίας τα δεδομένα είναι έτοιμα να εισαχθούν στο νευρωνικό δίκτυο και να ξεκινήσει η διαδικασία της εκπαίδευσης.

Όσον αφορά την υλοποίηση των δικτύων μακράς βραχυπρόθεσμης μνήμης, LSTM χρησιμοποιήθηκε η βιβλιοθήκη Keras [12] με τη χρήση του TensorFlow [13] ως σύστημα υποστήριξης και ο προγραμματισμός έγινε σε Python. Παρομοίως με πριν αφού διαβαστούν τα δεδομένα από τη τηλεμετρία του οχήματος αφαιρούνται οι μετρήσεις που αντιστοιχούν σε μηδενικές τιμές συντεταγμένων για να αποφευχθούν τα σφάλματα. Η διαφορά με ένα δίκτυο MLP είναι πως τα LSTM εκπαιδεύονται από ακολουθίες και όχι από μονοδιάστατα δεδομένα, όπως περιγράφεται στο κεφάλαιο 3, γι' αυτό ως δεύτερο βήμα της προεπεξεργασίας τα δεδομένα εισόδου χωρίζονται σε ακολουθίες μήκους *n*, στο υποκεφάλαιο 4.3 γίνεται αναλυτική περιγραφή του τρόπου επιλογής της σταθεράς *n*. Τα δεδομένα εξόδου σε ένα LSTM είναι μονοδιάστατες τιμές

γι’ αυτό προκειμένου να αναπαρασταθεί η κατανάλωση του οχήματος υπολογίζεται το άθροισμα των *n* τιμών που αντιστοιχούν στο διάνυσμα εισόδου.

Στο τέλος της προεπεξεργασίας αφαιρείται ένα ποσοστό της τάξης του 5-10 % και φυλάσσεται ξεχωριστά από τα δεδομένα εκμάθησης προκειμένου να εξεταστεί το νευρωνικό δίκτυο μετά την εκπαίδευση του και να μελετηθεί η ακρίβεια των προβλέψεων του σε ένα άγνωστο πακέτο δεδομένων.



Εικόνα 4.1: Αναπαράσταση των εισόδων και των εξόδων του LSTM δικτύου. Ως πρώτη είσοδος είναι η ταχύτητα που είναι απαραίτητο στοιχείο για να προσδιοριστή η βέλτιστη στρατηγική κίνησης. Οι υπόλοιπες είσοδοι περιγράφουν τη διαδρομή του οχήματος και διαφοροποιούνται μεταξύ των υλοποιήσεων, η τελική επιλογή παρουσιάζεται στην επόμενη υποενότητα. Ως έξοδο για την αναπαράσταση της κατανάλωσης χρησιμοποιείται το άθροισμα της καταναλισκόμενης ισχύς.

4.3 Επιλογή εισόδων-εξόδων

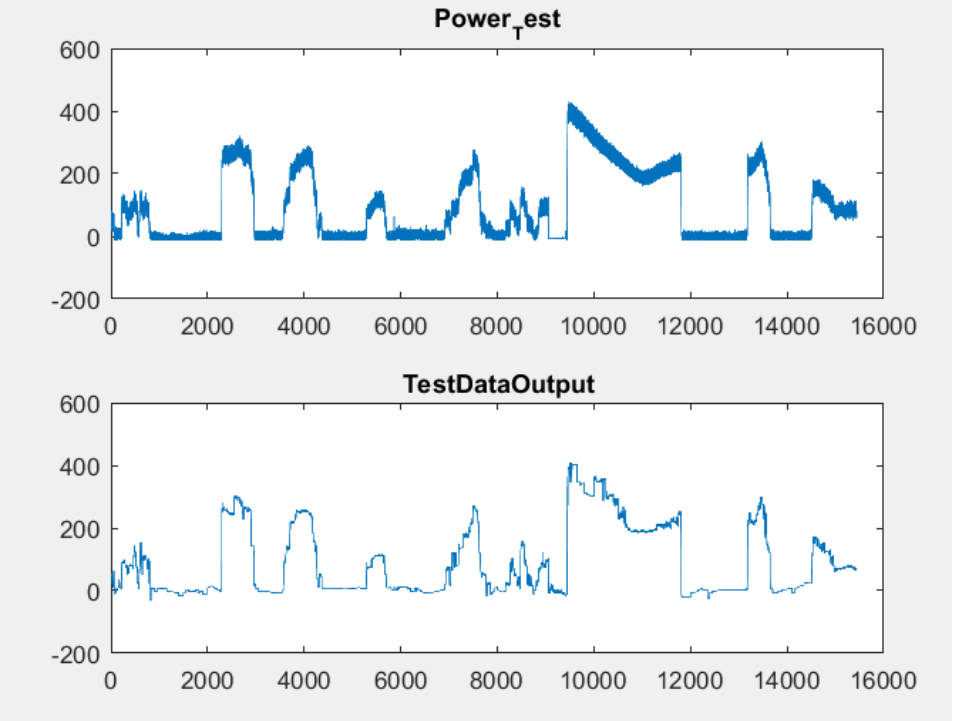
Από τα δεδομένα που είναι ικανή η τηλεμετρία να παρακολουθήσει και να αποθηκεύσει έπρεπε να γίνει η επιλογή των εισόδων και των εξόδων του ΤΝΔ.

4.3.1 MLP μοντέλο κατανάλωσης

Αρχικά παρατηρήθηκε πως το φασικό ρεύμα που καταναλώνει ο κινητήρα παρόλο που σχετίζεται άμεσα με τη καταναλισκόμενη ενέργεια δε μπορούσε να χρησιμοποιηθεί. Αυτό συνέβη διότι στο επόμενο βήμα του συστήματος ο ευφυής πράκτορας θα αλλάζει τη ταχύτητα προκειμένου να παρατηρεί τη διαφορά στη κατανάλωση, εάν είχε χρησιμοποιηθεί αυτό το ρεύμα ως είσοδος θα έπρεπε να αλλάζει και τις 2 παραμέτρους αφού η μία είναι εξαρτώμενη από την άλλη. Η εξάρτηση αυτή όμως είναι άγνωστη και δεν μπορεί εύκολα να μοντελοποιηθεί μαθηματικά με ακρίβεια, γι’ αυτό το λόγο αποφασίστηκε να μην αξιοποιηθεί.

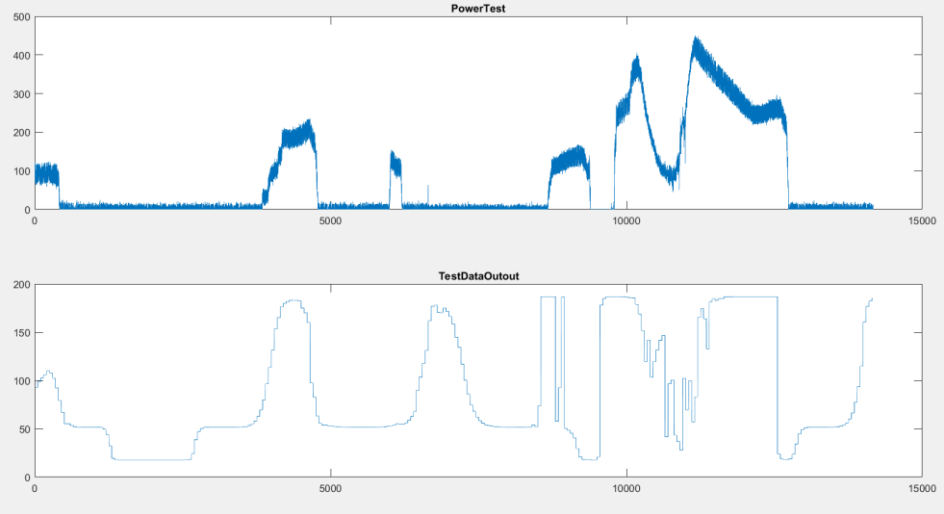
Πρωτού όμως γίνει αυτή η παρατήρηση είχε κατασκευαστεί ένα απλό νευρωνικό δίκτυο τριών επιπέδων στο περιβάλλον Matlab. Αυτό το δίκτυο έπαιρνε ως είσοδο τη ταχύτητα του οχήματος, το φασικό ρεύμα που κατανάλωνε ο κινητήρας, τη θέση του ως συντεταμένες και ως έξοδο είχε τη κατανάλωση του οχήματος κάθε χρονική στιγμή, δηλαδή το γινόμενο της τάσης της μπαταρίας με το συνεχές ρεύμα που παρήγαγαν. Για να περιγραφεί η θέση του οχήματος χρησιμοποιήθηκαν μόνο οι συντεταμένες από το GPS, αυτό έγινε υπό την υπόθεση πως η κλίση του οδοστρώματος και οι ανωμαλίες του θα εντάσσονταν έμμεσα στην εκπαίδευση. Πιο συγκεκριμένα εάν η κλίση της διαδρομής ήταν κατηφορική σε κάποιο σημείο το δίκτυο θα παρατηρούσε

μείωση της κατανάλωσης για μία ταχύτητα u ενώ για την ίδια ταχύτατα σε ένα διαφορετικό ζευγάρι που θα αναπαριστούσε ανηφόρα θα παρατηρούσε αύξηση σύμφωνα με το θεωρία που περιγράφεται στο κεφάλαιο 2. Το απλό αυτό δίκτυο ήταν ικανό να προβλέψει με μεγάλη ακρίβεια την κατανάλωση του οχήματος πάνω σε μία άγνωστη διαδρομή, Εικόνα 4.2.



Εικόνα 4.2: Στο πάνω γράφημα παρουσιάζεται η πραγματική κατανάλωση του οχήματος ενώ στο κάτω η πρόβλεψη ενός νευρωνικού δικτύου 3 επιπέδων (10,20,10) με συνάρτηση ενεργοποίησης tansig και στα τρία επίπεδα. Το δίκτυο εκπαιδεύτηκε με τη βοήθεια των συναρτήσεων trainlm και learngdm με όλες τις παραμέτρους στην προεπιλεγμένη τους τιμή. Χρησιμοποιήθηκε το 90% των δεδομένων από το Shell Eco Marathon 2016 για την εκπαίδευση και το 10% για την επιβεβαίωση. Τα γραφήματα αναπαριστούν τη κατανάλωση του οχήματος σε άγνωστο πακέτο δεδομένων. Το μέσο τετραγωνικό σφάλμα μετά την εκπαίδευση είναι 293.1503 μετά από 1086 εποχές και το σφάλμα μεταξύ της πραγματικής συνολικής καταναλισκόμενης ενέργειας και της προβλεπόμενης είναι 0.0393 %

Μετά από τη παραπάνω παρατήρηση υλοποιήθηκε ένα δίκτυο όπου σαν είσοδο είχε μόνο τη θέση του οχήματος ως συντεταγμένες και τη ταχύτητα του ενώ ως έξοδο τη κατανάλωση του οχήματος κάθε χρονική στιγμή. Σε αντίθεση με το προηγούμενο δίκτυο η αφαίρεση μίας εισόδου οδήγησε σε μη αποδεκτά αποτελέσματα με μέσο τετραγωνικό σφάλμα της τάξης του 4033.5373 όπως παρουσιάζεται στην Εικόνα 4.3.



Εικόνα 4.3: Στο πάνω γράφημα παρουσιάζεται η πραγματική κατανάλωση του οχήματος ενώ στο κάτω η πρόβλεψη του ίδιου νευρωνικού δικτύου όπως την εικόνα 4.2 με τη διαφορά ότι το εναλλασσόμενο ρεύμα δεν αποτελεί πλέον είσοδο του. Αυτή τη φορά το μέσο τετραγωνικό σφάλμα είναι 4033.5373 μετά από 239 εποχές ενώ το σφάλμα μεταξύ της πραγματικής συνολικής καταναλισκόμενης ενέργειας και της προβλεπόμενης είναι 114.2015 %

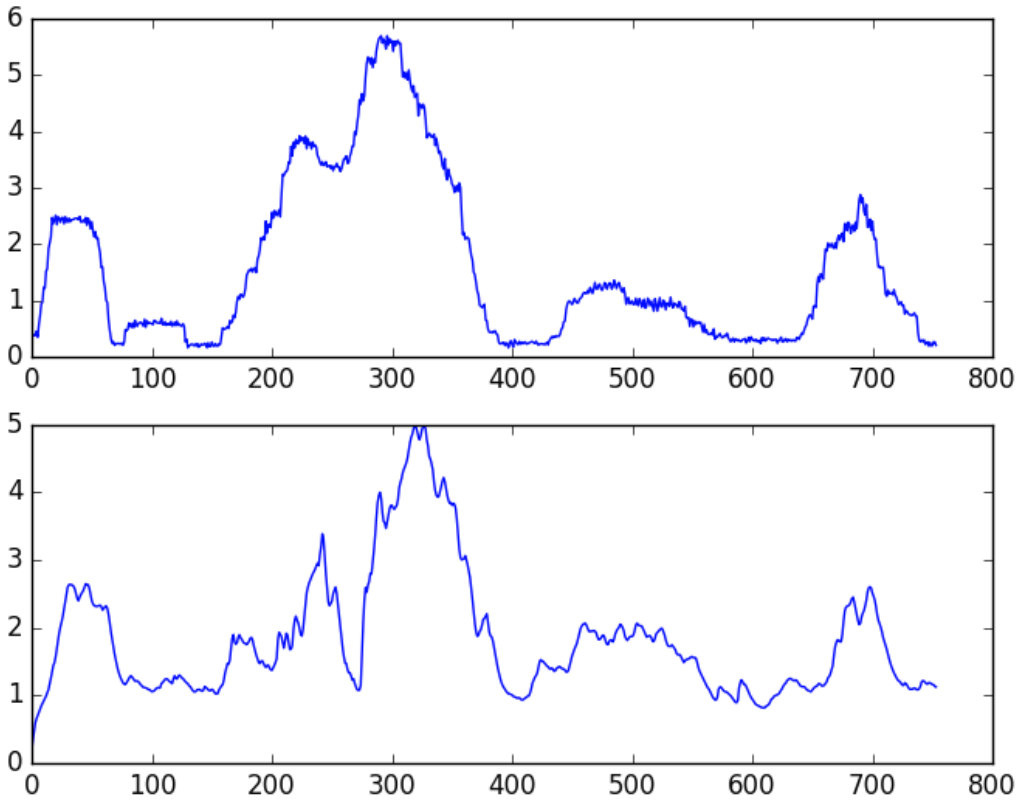
4.3.2 LSTM μοντέλο κατανάλωσης

Έπειτα η έρευνα για το βέλτιστο μοντέλο κατανάλωσης συνεχίστηκε με τη χρήση δικτύων μακράς βραχυπρόθεσμης μνήμης, LSTM, προκειμένου να εξαχθούν ορθά αποτελέσματα χωρίς το φασικό ρεύμα ως είσοδο.

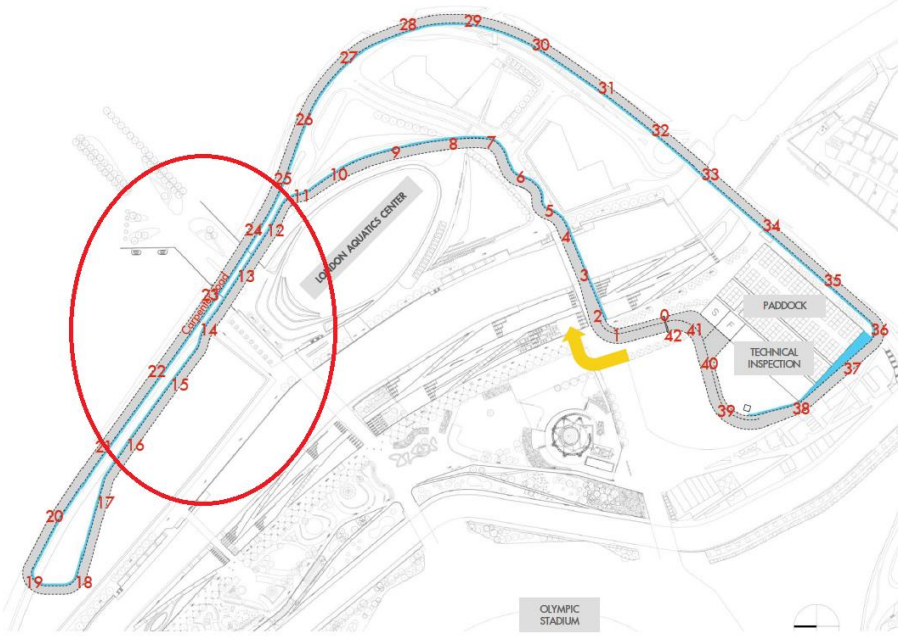
Η πρώτη δοκιμή έγινε με τις εισόδους που χρησιμοποιήθηκαν στο τελευταίο MLP προκειμένου να τονιστεί η διαφορά μεταξύ των δύο δικτύων. Η υλοποίηση πραγματοποιήθηκε μόνο με τη χρήση του TensorFlow στη γλώσσα προγραμματισμού Python. Η αρχιτεκτονική που χρησιμοποιήθηκε ήταν ενός κρυφού επιπέδου που αποτελούταν από 25 νευρώνες. Ο κάθε νευρώνας υλοποιήθηκε από ένα βασικό LSTM στοιχείο της βιβλιοθήκης TensorFlow (BasicLSTMCell) [13] Η είσοδος αποτελούταν από ακολουθίες των πενήντα στοιχείων, ως σφάλμα επιλέχθηκε το μέσο τετραγωνικό σφάλμα, η εκπαίδευση έγινε σε πακέτα των δέκα ακολουθιών, ο ρυθμός εκμάθησής ορίστηκες 0,009 και η εκπαίδευση διήρκησε 23500 εποχές. Τα αποτελέσματα παρουσιάζονται στην εικόνα 4.4. Είναι φανερό πως το συγκεκριμένο δίκτυο είναι ικανό να ξεχωρίσει πότε το όχημα καταναλώνει μεγαλύτερες ποσότητες ενέργειες σε αντίθεση με το ίδιο MLP δίκτυο. Αυτή τη φορά το μέσο τετραγωνικό σφάλμα μετά την εκπαίδευση είναι 1.815 και το σφάλμα μεταξύ της πραγματικής συνολικής καταναλισκόμενης ενέργειας και της προβλεπόμενης είναι 8.95%.

Το αποτέλεσμα είναι σαφώς ανώτερο του πρώτου δικτύου που υλοποιήθηκε, ωστόσο παρατηρήθηκε ένα άλλο σημαντικό πρόβλημα με αυτή τη διάταξη. Η πίστα του διαγωνισμού είναι κυκλική και δύο ανεξάρτητα τμήματα της με διαφορετικές κλίσης είναι πολύ κοντά μεταξύ τους και οι συντεταγμένες τους είναι σχεδόν πανομοιότητές, Εικόνα 4.5.

Προκειμένου να καταπολεμηθεί αυτό το πρόβλημα σχεδιάστηκε εκ νέου ένα LSTM δίκτυο που η διαδρομή δεν περιγραφόταν από τις συντεταγμένες αλλά από μία ακολουθία κλίσεων του οδοστρώματος.

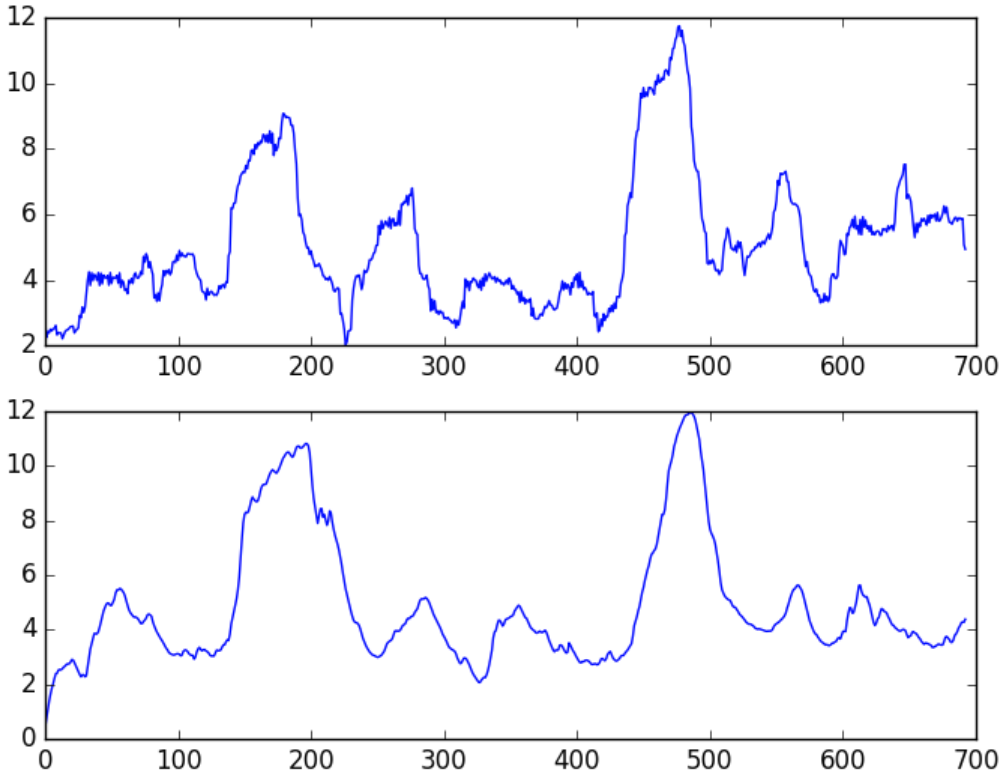


Εικόνα 4.4: Στο επάνω γράφημα παρουσιάζεται η πραγματική κατανάλωση του οχήματος ενώ στο κάτω η πρόβλεψη του LSTM νευρωνικού δικτύου που δέχεται ως είσοδο ακολουθίες των 50 στοιχείων. Οι ακολουθίες αυτές αναπαριστούν τη ταχύτητα και το γεωγραφικό πλάτος και μήκος που περιγράφουν τη θέση του οχήματος.



Εικόνα 4.5: Παρουσιάζεται ο χάρτης της πίστας του μαραθώνιου οικονομίας 2016. Εσωτερικά του κόκκινου κύκλου βρίσκονται τα τμήματα της πίστας που έχουν σχεδόν πανομοιότυπες συντεταγμένες.

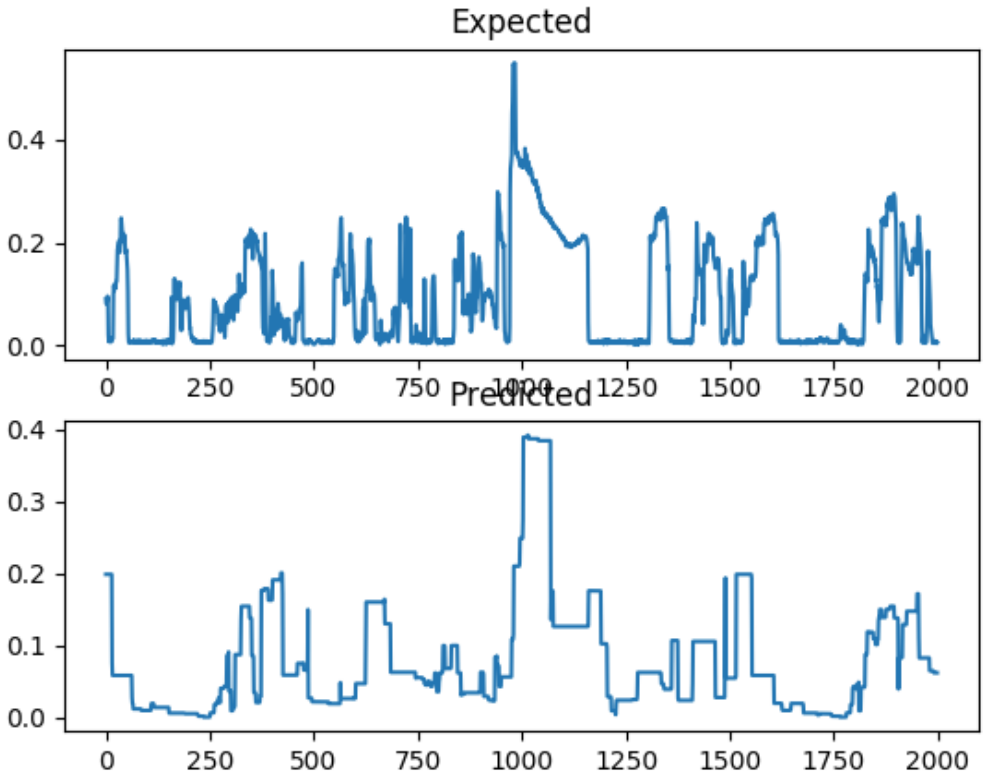
Το δεύτερο δίκτυο που υλοποιήθηκε διατηρούσε ως εισόδους τη κλίση του οδοστρώματος και τη ταχύτητα του οχήματος, η κλίση θεωρήθηκε γνωστή. Η Shell παραχωρεί ένα πλήρη χάρτη της πίστας που διαγωνίζονται τα πρότυπα οχήματα μαζί με τις κλίσεις κάθε σημείου της διαδρομής. Έτσι κατά τη διάρκεια της προεπεξεργασίας γινόταν μία ταύτιση των σημείων της πίστας, μέσω των συντεταγμένων τους, με τα στοιχεία που παραχωρεί η διοργανώτρια εταιρία. Η υλοποίηση πραγματοποιήθηκε με τη χρήση της βιβλιοθήκης Keras και το TensorFlow ως σύστημα υποστήριξης. Μετά τη βελτιστοποίηση των παραμέτρων το δίκτυο ήταν ικανό να προβλέψει τη καταναλισκόμενη ενέργεια με μέσο τετραγωνικό σφάλμα 1.212 και η απόκλιση μεταξύ της συνολικής προβλεπόμενης ενέργειας από τη πραγματική είναι 3,59% χωρίς να αλλοιωθεί η κυματομορφή, τα αποτελέσματα παρουσιάζονται στην εικόνα 4.6.



Εικόνα 4.6: Αποτελέσματα του LSTM δικτύου που δέχεται ως εισόδους 2 ακολουθίες χιλίων στοιχείων, τη ταχύτητα και τη κλίση του οδοστρώματος. Η αρχιτεκτονική του δικτύου ήταν δύο κρυφών επιπέδων το πρώτο αποτελούταν από 80 σιγμοειδής νευρώνες ενώ το δεύτερο από 40 γραμμικούς. Το σφάλμα που χρησιμοποιήθηκε ήταν το μέσο τετραγωνικό σφάλμα και ως συνάρτηση βελτιστοποίησης η RMSprop με όλες τις παραμέτρους στην προεπιλεγμένη τους τιμή.

Στη συνέχεια σε μία προσπάθεια βελτιστοποίησης του ΤΝΔ αποφασίστηκε να χρησιμοποιηθούν και οι στροφές της πίστας ως είσοδοι. Με τη χρήση των τοπογραφικών δεδομένων και την ίδια διαδικασία αντιστοίχησης των μετρήσεων που χρησιμοποιήθηκε προηγουμένως μοντελοποιήθηκαν οι στροφές της πίστας. Η μοντελοποίηση αυτή έγινε υπό τη μορφή κωδικοποίησης. Η αριστερά στροφή είχε τη τιμή -1, η ευθεία τη τιμή 0 και η δεξιά στροφή τη τιμή 1. Αυτό έγινε ως μία απόπειρα εκμετάλλευσης όλων των αδυναμιών του αυτοκινήτου, δηλαδή εάν το όχημα είχε περισσότερες τριβές σε μία αριστερή στροφή από ότι σε μία δεξιά με αυτό τον τρόπο θα μπορούσε να γίνει η παρατήρηση.

Ωστόσο τα αποτελέσματα του ΤΝΔ απέδειξαν πως η χρήση αυτής της εισόδου δημιουργεί προβλήματα και μεγαλώνει το σφάλμα των αποτελεσμάτων σε σχέση με προηγουμένως, πιο συγκεκριμένα το μέσο τετραγωνικό σφάλμα είχε τιμή 13.73 και το σφάλμα μεταξύ της συνολικής καταναλισκόμενης ενέργειας και της πρόβλεψης σε ένα άγνωστο πακέτο δεδομένων ήταν 13.21 % τα αποτελέσματα αναπαρίστανται στην εικόνα 4.7.



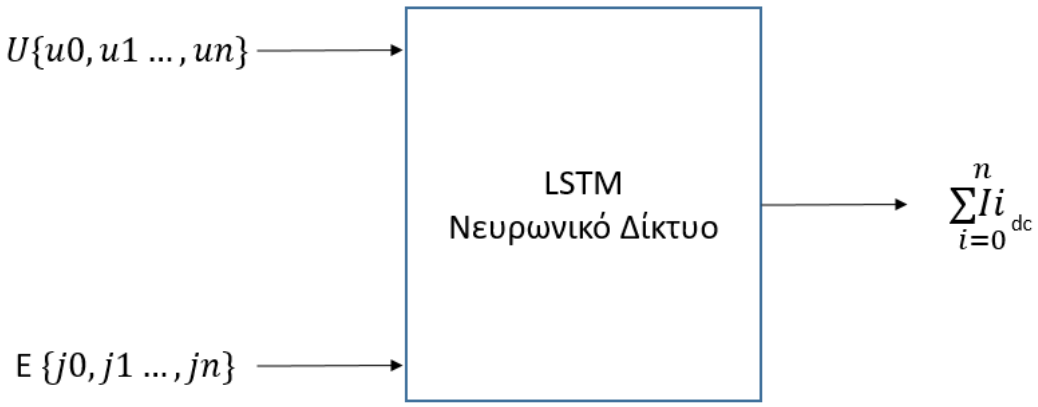
Εικόνα 4.7: Αποτελέσματα του LSTM δικτύου που δέχεται ως εισόδους 3 ακολουθίες χιλίων στοιχείων, τη ταχύτητα, τη κλίση και τη στροφή υπό τη μορφή κωδικοποίησης (-1,0,1). Η αρχιτεκτονική του δικτύου ήταν δύο κρυφών επιπέδων το πρώτο αποτελούταν από 80 σιγμοειδής νευρώνες ενώ το δεύτερο από 40 γραμμικούς. Το σφάλμα που χρησιμοποιήθηκε ήταν το μέσο τετραγωνικό σφάλμα και ως συνάρτηση βελτιστοποίησης η RMSprop με όλες τις παραμέτρους στην προεπιλεγμένη τους τιμή.

Τα παραπάνω αποτελέσματα οδήγησαν στη παρατήρηση πως κατά τη διαδικασία αντιστοίχησης των μετρήσεων με τα τοπογραφικά δεδομένα γίνονται πολλά λάθη. Αυτό συνέβη διότι τα τοπογραφικά δεδομένα δεν ήταν μεγάλης ακρίβειας και η αντιστοίχηση γινόταν μέσω του χάρτη συντεταγμένων και ενός χάρτη της πίστας που παρείχε η Shell. Ωστόσο αυτοί οι δύο χάρτες είχαν ασυνέπειες μεταξύ τους και κρίθηκε απαραίτητο να μη χρησιμοποιηθούν.

Για να καταπολεμηθεί το παραπάνω πρόβλημα και να μεγαλώσει το εύρος χρήσης του συστήματος και εκτός της πίστας του μαραθωνίου οικονομίας αποφασίστηκε να δημιουργηθεί ένα δίκτυο που θα μοντελοποιεί τη διαδρομή χρησιμοποιώντας τα δεδομένα του υψομέτρου από το GPS ενώ οι στροφές της διαδρομής κρίθηκε απαραίτητο να μη χρησιμοποιηθούν.

Εφόσον τα δεδομένα εισέρχονται υπό μορφή ακολουθίας στο νευρωνικό δίκτυο θα είναι ικανό να παρακολουθεί τη μεταβολή του υψομέτρου και να συμπεραίνει εάν το έδαφος είναι κατηφορικό ή ανηφορικό. Επιπλέον για να μειωθεί το σφάλμα από τη τηλεμετρία αποφασίστηκε ως ένδειξη της κατανάλωσης να μη μετριόνται δύο τιμές και να εξάγεται το γινόμενο ως καταναλισκόμενη ισχύ αλλά να υπολογίζεται μόνο το

καταναλισκόμενο συνεχές ρεύμα των μπαταριών. Στη συνέχεια περιγράφετε η διαδικασία βελτιστοποίησης των παραμέτρων που ακολουθήθηκε και τα αποτελέσματα του τελικού μοντέλου κατανάλωσης.



Εικόνα 4.7: Παρουσιάζονται οι τελικές είσοδοι και έξοδοι του νευρωνικού δικτύου που μοντελοποιεί τη κατανάλωση του οχήματος.

4.3 Πειραματισμός παραμέτρων

Η πρώτη παράμετρος του νευρωνικού δικτύου που έπρεπε να οριστεί ήταν το μήκος της ακολουθίας. Η τηλεμετρία του οχήματος κάνει μετρήσεις με δειγματοληψία 50Hz, δηλαδή πραγματοποιεί 50 μετρήσεις το δευτερόλεπτο. Γι' αυτό το λόγο ο αριθμός 50 χρησιμοποιήθηκε ως βήμα για να προσδιοριστεί η βέλτιστη ακολουθία στοιχείων. Επιλέχθηκε η ταχύτητα δειγματοληψίας ως βήμα διότι το όχημα κινείται με μικρές ταχύτητες και εντός ενός δευτερολέπτου δεν υπάρχει χρόνος για μεγάλες αλλαγές. Μετά από πειραματισμό στο διάστημα 50-1000 αποφασίστηκε πως η βέλτιστη ακολουθία ήταν μήκους τριακοσίων στοιχείων. Το μήκος αυτό επέτρεπε στο όχημα να κινηθεί στο χώρο και οι τιμές της ταχύτητας και του υψόμετρου να έχουν περιθώριο μεταβολής. Επιπλέον δεν ήταν πολύ μεγάλο με αποτέλεσμα να μην επιβαρύνεται η διαδικασία της εκπαίδευσης, όσο μεγαλύτερο είναι το μήκος της ακολουθίας τόσο πιο δύσκολο είναι για το νευρωνικό δίκτυο να σχηματίσει εξαρτήσεις μεταξύ των στοιχείων. Ακόμα ένα πλεονέκτημα της μικρής ακολουθίας ήταν πως το συνολικό μέγεθος των δεδομένων εκπαίδευσης δεν μίκρυνε αισθητά. Προφανώς όσο μεγαλύτερο είναι το μήκος της ακολουθίας ένα περιορισμένο σύνολο δεδομένων διαιρείται σε λιγότερα κομμάτια μεγαλύτερου μήκους.

Προκειμένου να προσδιοριστεί η βέλτιστη αρχιτεκτονική για το δίκτυο δοκιμάστηκαν όλοι οι πιθανοί συνδυασμοί νευρώνων για ένα αλλά και δύο κρυφά επίπεδα από 5 μέχρι 100 νευρώνες ανά επίπεδο, με βήμα 5. Για να γίνει αυτό υλοποιήθηκε ένας κώδικας που εκπαίδευε το ίδιο νευρωνικό δίκτυο με διαφορετικές αρχιτεκτονικές. Αυτό το πρόγραμμα εκτελέστηκε τρεις φορές και επιλέχθηκαν 5 διαφορετικές αρχιτεκτονικές που παρουσίαζαν την μικρότερη διαφορά μεταξύ της πραγματικής κατανάλωσης και της αντίστοιχης πρόβλεψης. Από αυτές τις 5 συνδεσμολογίες νευρώνων εξετάστηκαν οι γραφικές παραστάσεις των προβλέψεων τους και επιλέχθηκε μία αρχιτεκτονική δύο κρυφών επιπέδων με 75 νευρώνες στο πρώτο και 15 στο δεύτερο.

Στο τρίτο στάδιο αναγνώρισης των βέλτιστων παραμέτρων του νευρωνικού δικτύου εξετάστηκαν όλοι οι συνδυασμοί των υπάρχοντων συναρτήσεων

ενεργοποίησης στη βιβλιοθήκη Keras (softplus, softsign, relu, tanh, sigmoid, hard_sigmoid, linear, softmax) [12] Όλοι οι συνδυασμοί παρουσίαζαν αποτελέσματα παρόμοια ή χειρότερα από τη χρήση της απλής γραμμικής συνάρτηση ενεργοποίησης, linear. Γι' αυτό το λόγο και στα 2 κρυφά επίπεδα χρησιμοποιήθηκε η γραμμική συνάρτηση ενεργοποίησης.

Ως σφάλμα εκπαίδευσης χρησιμοποιήθηκε το μέσο τετραγωνικό σφάλμα διότι παρουσίασε τα καλύτερα αποτελέσματα σύγκλισης μεταξύ των υπόλοιπων διαθέσιμων σφαλμάτων (mean_squared_error, mean_absolute_error, mean_absolute_percentage_error, mean_squared_logarithmic_error, hinge, categorical_hinge, logcosh, categorical_crossentropy, squared_hinge, sparse_categorical_crossentropy, binary_crossentropy, kullback_leibler_divergence, poisson, cosine_proximity) [12] Υπεύθυνη συνάρτηση για την ελαχιστοποίηση του σφάλματος ορίστηκε η RMSprop λόγο της ανωτερότητας της μεταξύ των υπολοίπων σε ανατροφοδοτούμενα νευρωνικά δίκτυα. Αυτή η συνάρτηση ελαχιστοποίησης ξεχώρισε από τη συλλογή συναρτήσεων βελτιστοποίησης της βιβλιοθήκης Keras, οι υπόλοιπες που δοκιμάστηκαν και απορρίφθηκαν ήταν: Adagrad, Adadelta, Adam, Adamax, Nadam, TFOptimizer [12]. Ο ρυθμός εκπαίδευσης ορίστηκε 0,001 ενώ οι υπόλοιπες παράμετροι της συνάρτησης ελαχιστοποίησης σφάλματος αφέθηκαν στις προεπιλεγμένες τους τιμές.

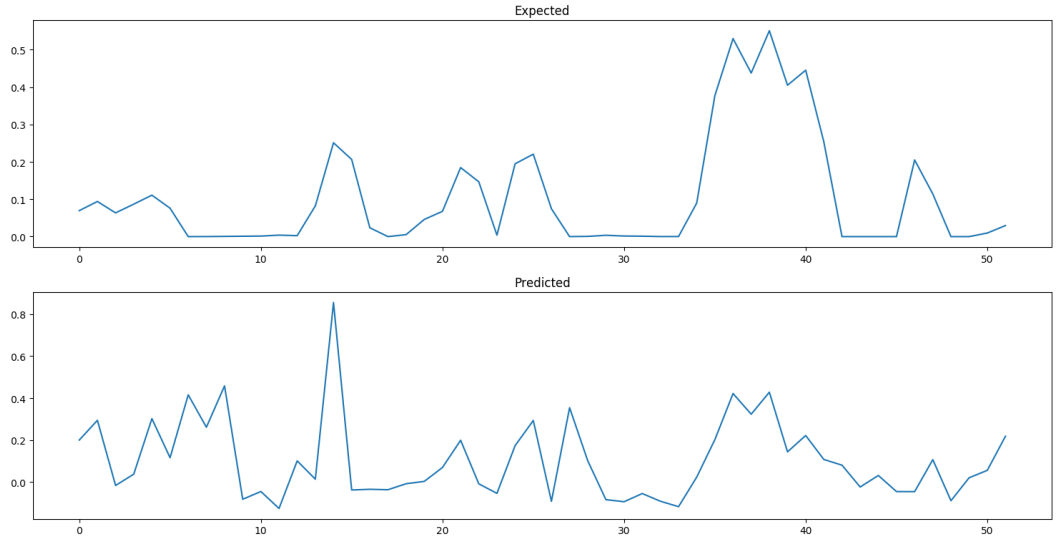
Επιπλέον η σταθερά αναμονής που εκφράζει τον αριθμό των εποχών που η διαδικασία της εκπαίδευσης συνεχίζεται παρόλο που δεν παρουσιάζεται βελτίωση του σφάλματος ορίστηκε 13.

Τέλος προκειμένου η διαδικασία της εκπαίδευσης να βελτιστοποιηθεί και να έχει θετικά αποτελέσματα χρησιμοποιήθηκε ένα πακέτο δεδομένων από το διαγωνισμό του μαραθωνίου οικονομίας 2016 που αποτελούταν από 8 γύρους της ίδιας πίστας και ένα πακέτο από δεδομένα που μετρήθηκαν κατά τη περίοδο δοκιμών. Το δεύτερο πακέτο δεδομένων αποτελούταν από ακραίες συνθήκες λειτουργείας όπως είναι η συνεχής επιτάχυνση σε ανηφορικό οδόστρωμα ή επιβράδυνση στη κατηφόρα. Αυτές οι συνθήκες προφανώς επιβαρύνουν τη κατανάλωση του οχήματος αλλά ήταν επιθυμητό το νευρωνικό δίκτυο να εκπαιδευτεί πλήρως και να γνωρίζει τι συμβαίνει στη κατανάλωση σε συνθήκες που δεν δοκιμάζονται κατά τη διαδικασία του διαγωνισμού. Τέλος χρησιμοποιήθηκε και ένα πακέτο δεδομένων που περιείχε ακολουθίες που το όχημα δεν ήταν δυνατό να εκτελέσει. Λόγο της μικρής ισχύς του κινητήρα σε πολύ ανηφορικά οδοστρώματα με κλίσεις πάνω από 15% το όχημα δεν είναι ικανό να επιταχύνει παραπάνω από 25km/h. Γι' αυτό το λόγο προστέθηκαν 50 ακολουθίες που περιείχαν δεδομένα τέτοιας μορφής επιτάχυνσης και η κατανάλωση ορίστηκε ίση με 1000, μία τιμή δηλαδή πολύ μεγαλύτερη από αυτή των μετρήσεων. Το σύνολο των δεδομένων εκπαίδευσης ήταν 812 ακολουθίες δηλαδή αποτελούταν από 243.600 μετρήσεις. Παρατηρήθηκε πως η χρήση περισσότερων δεδομένων οδηγούσε σε υπερεκπαίδευση του δικτύου. Επιπλέον για τη επιβεβαίωση της εκπαίδευσης χρησιμοποιήθηκε ένα σύνολο από 91 ακολουθίες, δηλαδή 27.300 μετρήσεις.

Οι παραπάνω βελτιστοποιήσεις και ένα πλήρες πακέτο δεδομένων είχαν σαν αποτέλεσμα να υλοποιηθεί ένα νευρωνικό δίκτυο που είναι ικανό να προβλέψει τη κατανάλωση του πρότυπου οχήματος Πυρφόρος. Στη συνέχει παρουσιάζονται τα αποτελέσματα του μοντέλου κατανάλωσης.

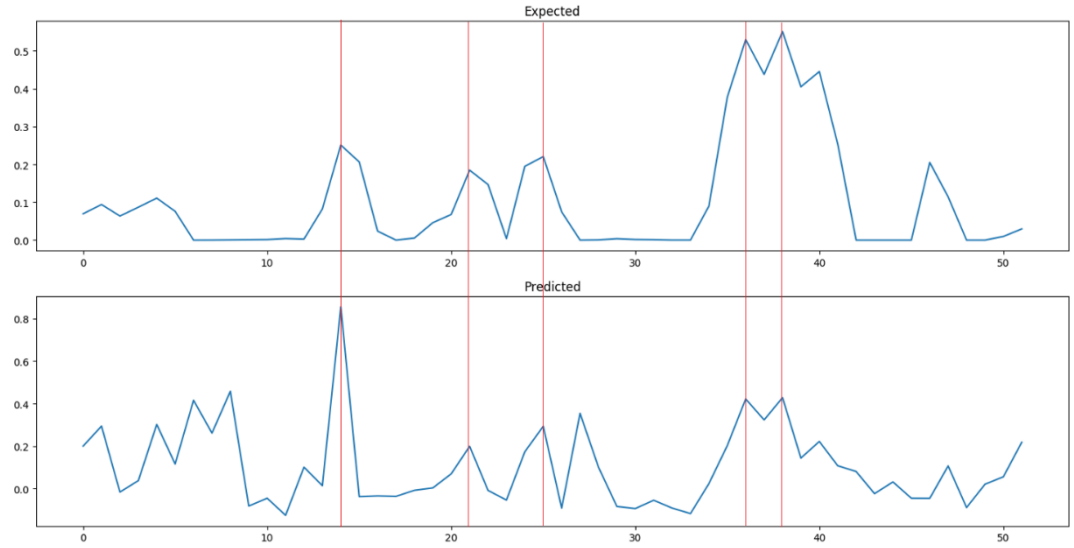
4.4 Αποτελέσματα

Προκειμένου να αξιολογηθεί το μοντέλο κατανάλωσης και να χαρακτηριστεί ως ικανό να προβλέψει τη κατανάλωση του οχήματος θα έπρεπε να μπορούσε πρώτον να προβλέψει σωστά τη κατανάλωση του οχήματος πάνω στη πίστα. Δεύτερον να καταλάβει τη αύξηση στη κατανάλωση όταν στην ίδια διαδρομή αυξανόταν η ταχύτητα του αυτοκινήτου. Και τέλος να μπορέσει να προβλέπει την αύξηση της κατανάλωσης όταν ως στρατηγική κίνησής το όχημα επιταχύνει σε ανηφορικό οδόστρωμα και επιβραδύνει σε κατηφορικό. Στη συνέχεια παρουσιάζονται διαγράμματα που αποδεικνύουν την αποτελεσματικότητα του μοντέλου κατανάλωσης που υλοποιήθηκε στις παραπάνω περιπτώσεις.



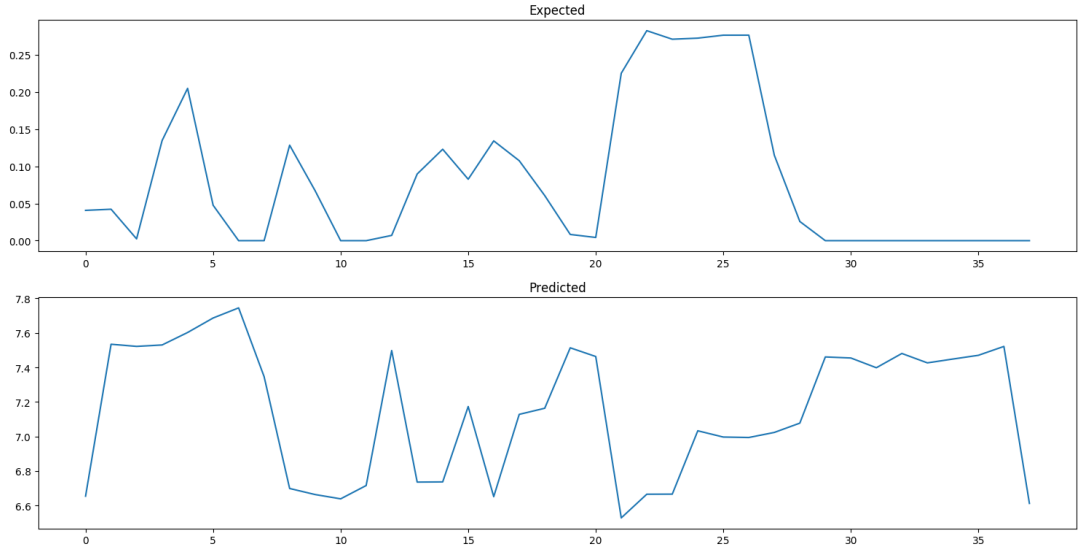
Εικόνα 4.8: Στο επάνω γράφημα απεικονίζεται η πραγματική κατανάλωση του οχήματος στη πίστα του μαραθωνίου οικονομίας ενώ στο κάτω παρουσιάζεται η πρόβλεψη του ΤΝΣ. Τα δεδομένα που χρησιμοποιήθηκαν για τη πρόβλεψη είναι διαφορετικά από αυτά που έγινε η εκπαίδευση.

Στο παραπάνω γράφημα γίνεται προφανές πως το δίκτυο είναι ικανό να προβλέψει τη κατανάλωση του οχήματος και να διατυπώσει πότε το όχημα αρχίζει να καταναλώνει μεγαλύτερες ποσότητες ενέργειας, πιο συγκεκριμένα το μέσο τετραγωνικό σφάλμα της πρόβλεψης είναι 1.212 και η ποσοστιαία διαφορά μεταξύ της συνολικής καταναλισκόμενης ενέργειας και της πρόβλεψης είναι 0.895%. Οι δύο κυματομορφές δεν είναι πανομοιότυπες αλλά το μοντέλο κατανάλωσης είναι ικανό να διαχωρίσει πότε αρχίζει μία δαπανηρή περίοδος κίνησης και πότε σταματά. Αυτό είναι αρκετό προκειμένου ο ευφυής πράκτορας στο επόμενο στάδιο να συμπεράνει πότε πρέπει να επιβραδύνει και πότε να επιταχύνει το όχημα.



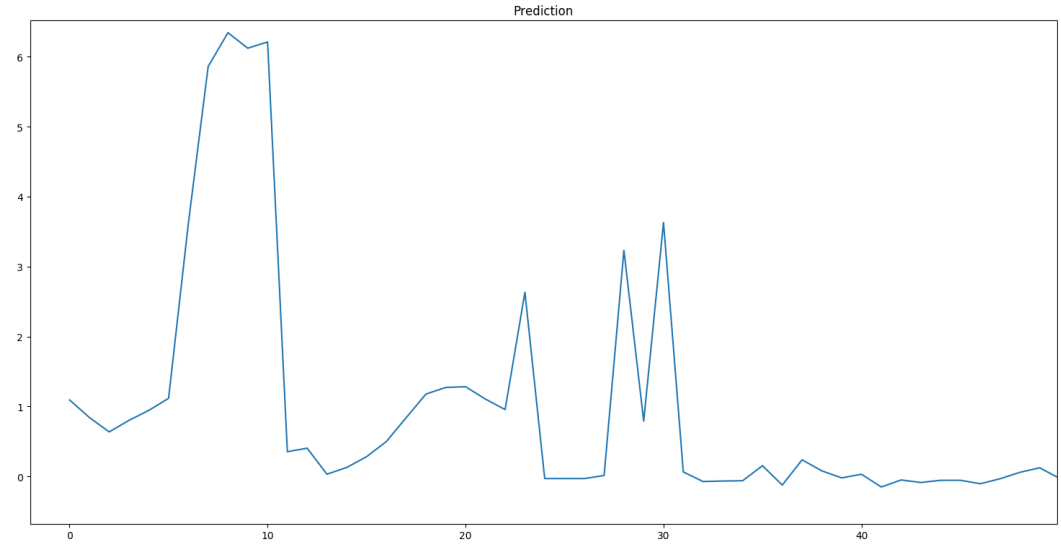
Εικόνα 4.9: Ταύτιση ακρότατων κυματομορφών.

Στο τρίτο γράφημα τονίζεται η διαφορά μεταξύ μίας ομαλής διαδρομής του οχήματος και μίας διαδρομής που διατηρεί μία υψηλή σταθερή ταχύτητα. Προφανώς το αποτέλεσμα καθιστά πολύ δαπανηρή τη χρήση υψηλών ταχυτήτων σε όλη τη διάρκεια της διαδρομής. Από το παρακάτω διάγραμμα φαίνεται πως το ΤΝΔ προβλέπει πτώση της κατανάλωσης εκεί που στη πραγματικότητα αυξάνεται. Αυτό συμβαίνει διότι κατά τη διάρκεια εκπαίδευσης του ΤΝΔ το όχημα προσεγγίζει της ανηφόρες με πολύ υψηλή ταχύτητα, προκειμένου να μη χρειαστεί να επιταχύνει κατά τη διάρκεια της, τα δεδομένα εκπαίδευσης υποδηλώνουν πως μία τέτοια συμπεριφορά μειώνει τη κατανάλωση και γι' αυτό παρατηρείται αυτό το φαινόμενο. Παρόλα αυτά η τιμή που αντιστοιχεί το ΤΝΔ στη κατανάλωση είναι πολύ μεγαλύτερη της πραγματικής κάτι το οποίο ήταν και ο πρωταρχικός στόχος.



Εικόνα 4.10: Στο επάνω γράφημα απεικονίζεται η πραγματική κατανάλωση του οχήματος στη πίστα του μαραθωνίου οικονομίας υπό κανονικές συνθήκες κίνησης ενώ στο κάτω απεικονίζεται η πρόβλεψη του μοντέλου κατανάλωσης εάν το όχημα διατηρούσε σταθερή ταχύτητα 45 km/h.

Το τελευταίο γράφημα απεικονίζει τη πρόβλεψη του μοντέλου κατανάλωσης όταν το όχημα διατηρεί μία στρατηγική που επιβαρύνει τη κατανάλωση του οχήματος. Δηλαδή επιταχύνει στα ανηφορικά και επιβραδύνει στα κατηφορικά οδοστρώματα. Ιδιαίτερα οι πρώτες 10 μετρήσεις αφορούν την μεγαλύτερη ανηφόρα που έχει η πίστα του μαραθωνίου οικονομίας στην οποία το όχημα επιχειρεί να επιταχύνει. Στη συνέχεια ακολουθεί μία απότομη κατηφόρα η οποία ακολουθείται από μικρές ανηφορικές κλήσεις.



Εικόνα 4.11: Σε αυτό το γράφημα απεικονίζεται η πρόβλεψη του ανατροφοδοτούμενου νευρωνικού δικτύου εάν το όχημα χρησιμοποιούσε στρατηγική κίνησης σύμφωνα με την οποία επιτάχυνε στην ανηφόρα και επιβράδυνε στη κατηφόρα.

Σύμφωνα με τα παραπάνω το μοντέλο κατανάλωσης που υλοποιήθηκε είναι ικανό να προβλέψει ορθά τη κατανάλωση του οχήματος σε διαφορετικά σενάρια, οποιαδήποτε στρατηγική μετακίνησης και να χρησιμοποιηθεί. Το νευρωνικό δίκτυο που περιεγράφηκε χρησιμοποιείται στη συνέχεια από τον ευφυή πράκτορα προκειμένου να εξαχθεί η βέλτιστη ακολουθία ταχυτήτων που θα επιτρέψει στο όχημα να ολοκληρώσει μία συγκεκριμένη διαδρομή σε ένα περιορισμένο χρονικό διάστημα καταναλώνοντας την ελάχιστη δυνατή ενέργεια. Το 5ο κεφάλαιο περιγράφει αναλυτικά τη διαδικασία εξαγωγής αυτής της βέλτιστης ακολουθίας ταχυτήτων.

Κεφάλαιο 5

Υπολογισμός βέλτιστης ακολουθίας ταχυτήτων

5.1 Εισαγωγή

Έχοντας υλοποιήσει το μοντέλο κατανάλωσης του οχήματος έχουμε ένα εργαλείο για να προσδιορίσουμε τη ποσότητα ενέργειας που καταναλώνεται για μία συγκεκριμένη διαδρομή και ακολουθία ταχυτήτων. Εάν θέλουμε δηλαδή μπορούμε να πειραματιστούμε με τις ταχύτητες της διαδρομής και να προβλέψουμε την τελική κατανάλωση κάθε ακολουθίας. Ο αλγόριθμος που περιγράφεται παρακάτω εκτελεί αυτήν ακριβώς τη λειτουργεία. Με άλλα λόγια εναλλάσσει τις ταχύτητες σε διαφορετικά τμήματα της διαδρομής με στόχο την ελαχιστοποίηση του κόστους, δηλαδή της κατανάλωσης.

Ερευνητές έχουν ανακαλύψει πως με τη πρόβλεψη των τοπογραφικών χαρακτηριστικών και τη προσαρμογή των ταχυτήτων μπορεί να εξοικονομηθεί έως και 30% από την καταναλισκόμενη ενέργεια, όπως περιεγράφηκε στο κεφάλαιο 2. Έχει πραγματοποιηθεί εκτεταμένη έρευνα πάνω σε αυτό το τομέα με τη χρήση γραμμικών αλγορίθμων, αλγορίθμων δυναμικού προγραμματισμού και ευριστικής βελτιστοποίησης όπως είναι οι γενετικοί αλγόριθμοι προκειμένου να ελαχιστοποιηθεί η καταναλισκόμενη ενέργεια. Σε αυτή την εφαρμογή επιλέχθηκε να ακολουθηθεί μία εναλλακτική προσέγγιση με τη χρήση μεθόδων τεχνητής νοημοσύνης και πιο συγκεκριμένα ενισχυτικής μάθησης. Η επιλογή αυτή έγινε διότι η τεχνητή νοημοσύνη είναι ικανή να εξάγει ορθά αποτελέσματα σε πολύπλοκα προβλήματα με μικρό υπολογιστικό κόστος.

Στη συνέχεια αναλύονται οι έννοιες της ενισχυτικής μάθησης, περιγράφεται αναλυτικά ο αλγόριθμος που χρησιμοποιήθηκε, η μεθοδολογία επιλογής των παραμέτρων του και τέλος παρουσιάζονται τα αποτελέσματα του υπολογισμού.

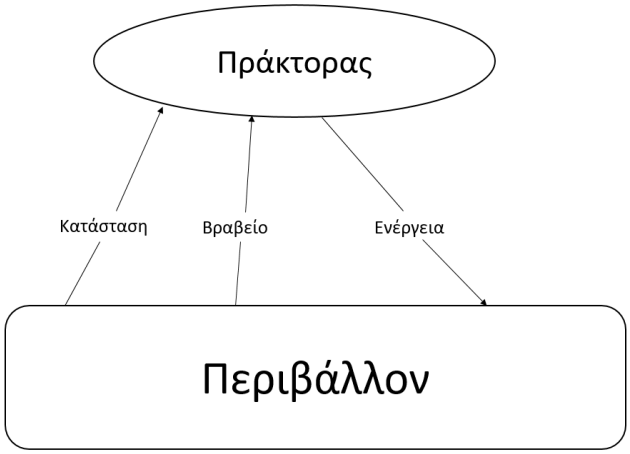
5.2 Ενισχυτική μάθηση (Reinforcement Learning, RL)

Η βασική ιδέα πίσω από την ενισχυτική μάθηση είναι βασισμένη στη μεθοδολογία εκμάθησης που χρησιμοποιήται από τους ανθρώπους. Ο χαρακτήρας που εκπαιδεύεται αποκαλείται πράκτορας (agent). Ο πράκτορας εκπαιδεύει τον εαυτό του με στόχο να ανακαλύψει τη βέλτιστη ακολουθία ενεργειών που θα του επιφέρουν το μεγαλύτερο κέρδος κατά την εκτέλεση μίας εργασίας. Προκειμένου να το πετύχει αυτό αλληλοεπιδρά με το περιβάλλον του λαμβάνοντας θετικά ή αρνητικά βραβεία. Σύμφωνα με τα βραβεία που λαμβάνει αξιολογεί κάθε ενέργεια που πραγματοποιεί και αποκτά εμπειρίες. Στόχος της εκπαίδευσης είναι να αποκτήσει αρκετές εμπειρίες ούτως ώστε να είναι ικανός να εκτελέσει μία βέλτιστη ακολουθία ενεργειών για να επιλύσει ένα πρόβλημα βασισμένος μόνο στις γνώσεις του. Ο πράκτορας δε χρειάζεται εξωτερική καθοδήγηση προκειμένου να επιλέξει τη βέλτιστη ακολουθία ενεργειών, αντίθετα αντιμετωπίζει το πρόβλημα με τη μεθοδολογία δοκιμής και απόρριψης (*trial-and-error*).

Στην ενισχυτική μάθηση γίνετε η υπόθεση πως η δυναμική του συστήματος, δηλαδή οι επιλογές του πράκτορα μπορούν να περιγραφούν σύμφωνα με τη μαθηματική σχέση του Markov (*Markov Decision Process, MDP*). Η MDP είναι μία στοχαστική διαδικασία βέλτιστης επιλογής σε διακριτό χρόνο που αναπαρίσταται ως τούπλα <s, a, p, r>. Τα στοιχεία της τούπλας είναι: η κατάσταση (state), η ενέργεια (action), οι πιθανότητες μετάβασης σε επόμενη κατάσταση και το βραβείο μετάβασης σε άλλη κατάσταση. Το βασικό χαρακτηριστικό της διαδικασίας του Markov ονομάζεται ιδιότητα του Markov και επισημάνει πως η κάθε ενέργεια είναι ανεξάρτητη από τις επιλογές του παρελθόντος. Άρα ο ευφυής πράκτορας θα μπορεί να περιγράψει τη κάθε κατάσταση χωρίς να γνωρίζει πληροφορίες για το παρελθόν.

Ο αλγόριθμος *Q-Learning* που επιλέχθηκε προκειμένου να υλοποιηθεί το παρουσιαζόμενο σύστημα είναι μία από τις πιο διαδεδομένες τεχνικές RL λόγο του μεγάλου πλήθους προβλημάτων που μπορεί να εφαρμοστεί. Παρουσιάστηκε πρώτη φορά από τον Watkins (1989) [14] και επιτρέπει στο πράκτορα να αποκτήσει εμπειρίες μέσω της αλληλεπίδρασης του με το περιβάλλον. Ο αλγόριθμος χωρίζεται σε εποχές, όπου σε κάθε εποχή ο πράκτορας μεταβαίνει σε μία κατάσταση s_n και επιλέγει μία ενέργεια a_n, $n \in [0,1 \dots, T]$ με Τ επιλογές. Η επιλογή του πράκτορα προκαλεί μία μετάβαση από μία κατάσταση σε μία άλλη και επιστρέφεται ένα κόστος που αντικατοπτρίζει την επιλογή του.

Η παρούσα εργασία πραγματεύεται την μεθοδολογία επιλογής των παραμέτρων που περιγράφουν τον αλγόριθμο *Q-Learning* και κάνει μία ανάλυση των αποτελεσμάτων του [15].



Εικόνα 5.1: Αλληλεπίδραση μεταξύ του πράκτορα και του περιβάλλοντος του.

5.3 Μεθοδολογία αλγορίθμου Q-Learning

Όπως αναφέρθηκε ο πράκτορας είναι ικανός να ανακαλύψει μέσω προσπάθειας την βέλτιστη λύση για το πρόβλημα που του παρουσιάζεται. Για να το κάνει αυτό είναι απαραίτητο να αναπαραστήσουμε τη γνώση του πράκτορα και να του προσθέσουμε την ιδιότητα της μνήμης. Αυτό επιτυγχάνεται με τη κατασκευή ενός πίνακα, το λεγόμενο πίνακα Q, με πλήθος γραμμών ίσο με τον αριθμό των καταστάσεων που μπορεί να βρεθεί και πλήθος στηλών ίσο με το πλήθος των ενεργειών που μπορεί να εκτελέσει σε κάθε κατάσταση. Ο πίνακας αυτός είναι αρχικοποιημένος με μηδενικές τιμές που συμβολίζουν την έλλειψη γνώσης για το πρόβλημα και καθώς εξελίσσεται η διαδικασία εκπαίδευσης ενημερώνεται με τον ακόλουθο τρόπο:

$$Q_t(s', a) = Q_{t\text{-}1}(s', a) + \mathrm{l}\ \{[r_t(s', a) + \gamma\ maxQ_{t+1}(s', a)] - Q_t(s', a)\}$$

Στη παραπάνω σχέση ενημερώνεται το στοιχείο του πίνακα που βρίσκεται στη γραμμή s και τη στήλη a. Η νέα τιμή ισούται με το άθροισμα της παλιάς τιμής Q(s,a) με το γινόμενο του ρυθμού εκμάθησης ($0 \leq l \leq 1$) με τη νέα εμπειρία του πράκτορα. Η νέα εμπειρία εκφράζεται ως υπολογισμός του βραβείου που δέχτηκε ο πράκτορας από το περιβάλλον, της παλιάς τιμής του Q(s,a) και τέλος το γινόμενο της σταθεράς μείωσης γ ($0 \leq \gamma \leq 1$, discount factor) με τη μέγιστη τιμή του Q στην επόμενη κατάσταση που θα βρεθεί ο πράκτορας εάν εφαρμόσει την ενέργεια a. Η σταθερά μείωσης περιγράφει την αξία που δίνει ο πράκτορας στο να δεχτεί επιβράβευση στο επόμενο βήμα, στην επόμενη κατάσταση. Εάν γ=0 δεν ενδιαφέρεται για το επόμενο βήμα ενώ εάν γ=1 θεωρεί πως το να επιβραβευτεί στο παρόν ή στο μέλλον έχει την ίδια αξία.

Στόχος του πράκτορα είναι να ανακαλύψει τη βέλτιστη ακολουθία ενεργειών που μπορεί να εκτελέσει, αυτές είναι οι μεγαλύτερες τιμές του πίνακα Q. Όσο μεγαλύτερη η τιμή μίας ενέργειας τόσο περισσότερο ο πράκτορας την εμπιστεύεται. Η ακολουθία με τις μέγιστες τιμές Q όπου κάθε μία αντιστοιχεί σε μία κατάσταση είναι η βέλτιστη ακολουθία που επιζητούσαμε εξαρχής και ονομάζεται πολιτική του πράκτορα (π(s)).

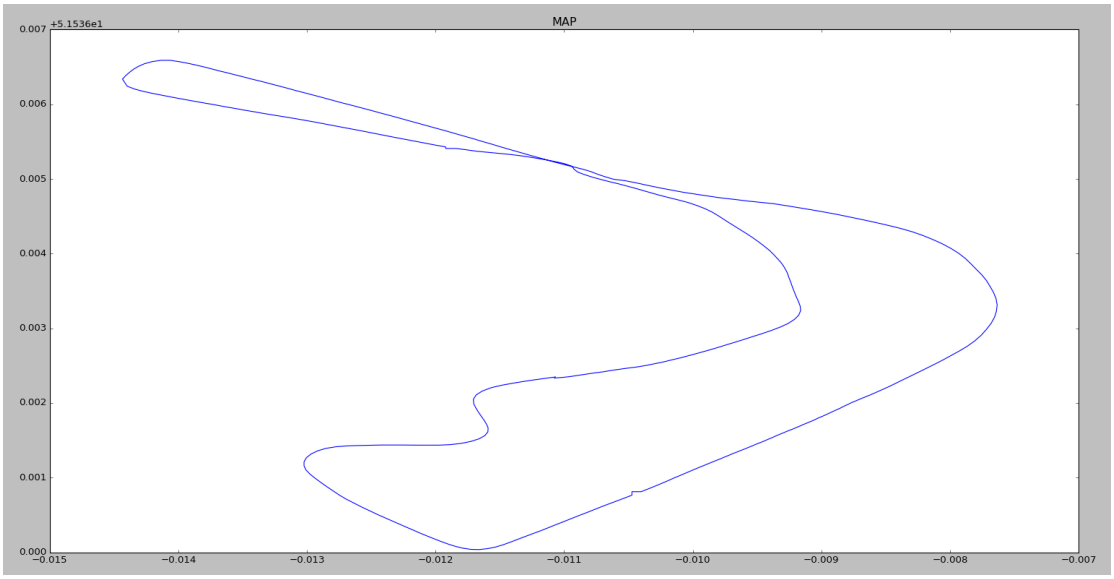
$$\pi(s)\ \in\ argmax\big(Q(s,a)\big)\ \forall\ s,a$$

Στο σύστημα που υλοποιήθηκε η διαδρομή χωρίζεται σε τμήματα, που περιγράφουν τις πιθανές καταστάσεις, σε κάθε ένα από τα οποία το όχημα κρατάει σταθερή ταχύτητα. Η ταχύτητα βέβαια που θα διατηρήσει σε κάθε τμήμα είναι η ενέργεια που επέλεξε να εκτελέσει ο πράκτορας.

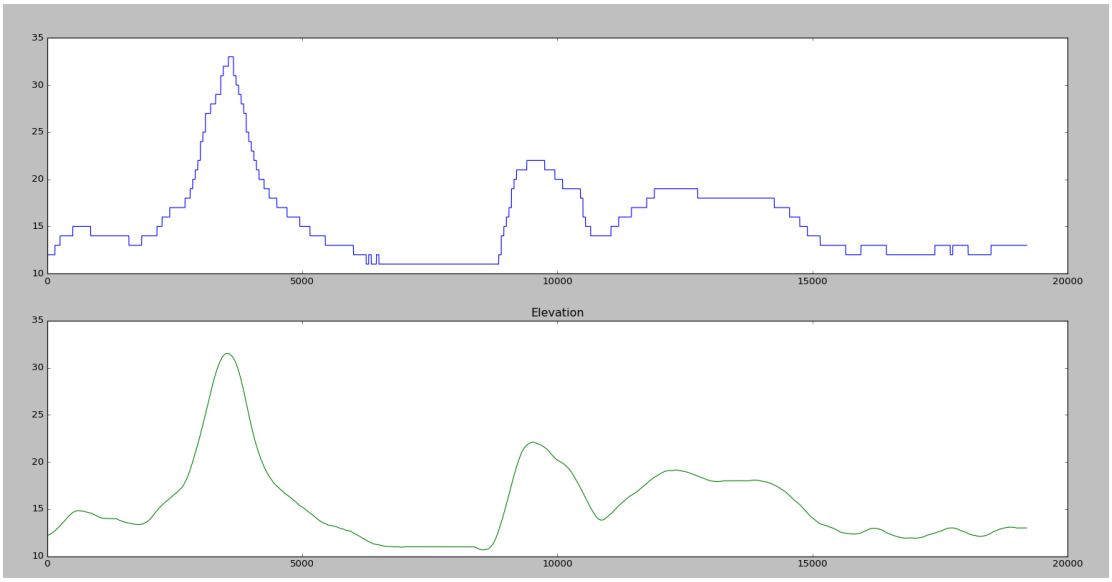
5.4 Τμηματοποίηση της διαδρομής

Προκειμένου να βρεθούν οι καταστάσεις στις οποίες θα τοποθετηθεί ο πράκτορας πρέπει η διαδρομή που θα ακολουθηθεί να χωριστεί σε επιμέρους τμήματα. Σε αυτά τα τμήματα ο πράκτορας θα διατηρεί σταθερή ταχύτητα.

Τα δεδομένα της διαδρομής προέρχονται από τη τηλεμετρία του οχήματος η οποία πραγματοποιεί μετρήσεις με συχνότητα 50Hz. Στο μοντέλο κατανάλωσης η πίστα αναπαρίσταται από τις υψομετρικές διαφορές του οδοστρώματος όπως περιεγράφηκε στο κεφάλαιο 4.3. Για λόγους ομοιομορφίας του συστήματος επιλέχθηκε η τμηματοποίηση να γίνει σύμφωνα με τις υψομετρικές διαφορές. Συνολικά η διαδρομή αποτελείται από 19.550 μετρήσεις. Ως πρώτο βήμα της διαδικασίας τα δεδομένα φιλτράρονται μέσω του φίλτρου *Savitzky-Golay, savgol*, με μέγεθος παραθύρου 1001 και πολυωνυμικής τάξης 3 [17] προκειμένου να ομαλοποιηθούν.



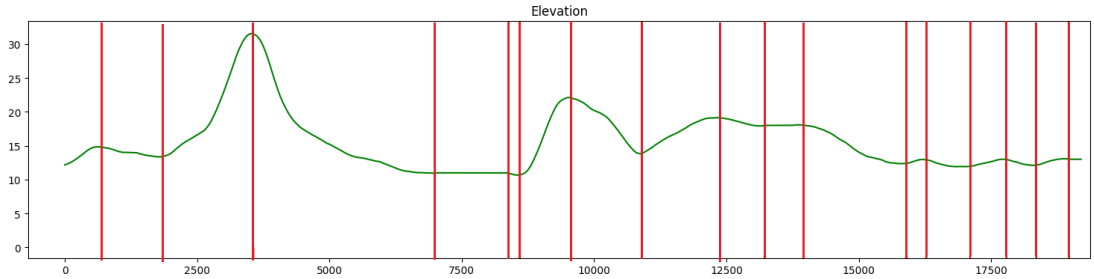
Εικόνα 5.2: Αναπαράσταση του χάρτη της διαδρομής



Εικόνα 5.3: Το πάνω γράφημα αναπαριστά τα υψομετρικά δεδομένα όπως προέρχονται από τη τηλεμετρία του οχήματος και κάτω είναι τα ίδια δεδομένα περασμένα από το φίλτρο Savitzky-Golay.

Στη συνέχεια εξετάζονται τα δεδομένα για να καθοριστούν οι καταστάσεις. Για να γίνει αυτό αρκεί να προσδιορίσουμε τον τρόπο μεταβολής της τιμής των δεδομένων. Δηλαδή εξετάζουμε εάν η τιμή του υψομέτρου κατεβαίνει, κατηφόρα, ή ανεβαίνει ανηφόρα. Θέτουμε αρχικά ένα επιτρεπτό αριθμό σφαλμάτων για να μειώσουμε το πλήθος των τμημάτων που θα έχουμε ως έξοδο. Η παράμετρος αυτή καθορίζει πόσες τιμές διαφορετικής μονοτονίας πρέπει να παρατηρηθούν πρωτού αντιστοιχίσουμε τις νέες τιμές σε διαφορετικό τμήμα της πίστας. Μετά από πειραματισμό για τη συγκεκριμένη διαδρομή ανακαλύφθηκε πως εάν θέσουμε αυτή τη παράμετρο ως 5 δεν υπάρχουν τμήματα με πάρα πολύ μικρό μήκος που θα ήταν και ανούσιο να υπολογίσουμε αφού η επίδραση τους θα ήταν σχεδόν αμελητέα.

Ως έξοδο αυτής της διαδικασίας προκύπτουν τα τμήματα που θα αντιμετωπιστούν ως οι καταστάσεις που θα βρεθεί ο πράκτορας και θα κληθεί να επιλέξει μία ενέργεια δηλαδή μία ταχύτητα. Η συγκεκριμένη διάταξη διαχώρισε τη διαδρομή σε 18 καταστάσεις με εναλλασσόμενες μονοτονίες που παρουσιάζονται παρακάτω.



Εικόνα 5.4: Οι καταστάσεις που θα βρεθεί ο πράκτορας κατά τη διαδικασία εκπαίδευσης του.

5.4 Σχεδιασμός αλγορίθμου

5.4.1. Διαστάσεις καταστάσεων (s) και ενεργειών (a)

Οι καταστάσεις του χώρου αναπαριστούν τα χαρακτηριστικά του περιβάλλοντος που είναι απαραίτητα για τη λύση του προβλήματος. Στη δική μας περίπτωση η θέση του οχήματος, ο επιθυμητός χρόνος τερματισμού και η ταχύτητα του είναι οι σημαντικοί παράγοντες για τον υπολογισμό της βέλτιστης ακολουθίας ταχυτήτων.

$$S = \{v_i, s_j, \mu\}: v_i \in V, s_i \in S \qquad \forall i, j \in \{1,2, \dots, n\}$$

Όπου v είναι η ταχύτητα του οχήματος, s είναι η κατάσταση και τέλος μ είναι η επιθυμητή μέση ταχύτητα. Το εύρος των πιθανών ταχυτήτων είναι από 10km/h έως 45km/h ενώ η ικανότητα του οχήματος να επιταχύνει και να επιβραδύνει θεωρείται στιγμιαία στα πλαίσια αυτής της εργασίας. Οι καταστάσεις προσδιορίστηκαν με τη παραπάνω μεθοδολογία. Και τέλος ο επιθυμητός χρόνος τερματισμού αναπαρίσταται ως τη μέση τιμή της ταχύτητας που πρέπει να διατηρηθεί για μία συγκεκριμένη διαδρομή ούτως ώστε να ολοκληρωθεί έγκαιρα.

5.4.2. Συνάρτηση επιβράβευσης

Η συνάρτηση επιβράβευσης είναι μία από τις πιο σημαντικές παραμέτρους που καθορίζουν την επιτυχία του συστήματος εκπαίδευσης. Αυτό συμβαίνει διότι αποτελεί το μόνο μέσω που έχει ο πράκτορας για να αξιολογήσει μία ενέργεια που εκτελεί. Ο κύριος στόχος του πράκτορα είναι να ελαχιστοποιήσει τη καταναλισκόμενη ενέργεια διατηρώντας τη μέση ταχύτητα του οχήματος σταθερή. Για αυτό το λόγο η συνάρτηση επιβράβευσης αναπαρίσταται από το άθροισμα του βραβείου κατανάλωσης και του χρονικού βραβείου.

Το βραβείο κατανάλωσης προσδιορίζεται από το πόσο απομακρύνεται η πολιτική του πράκτορα από την ακολουθία ταχυτήτων με την ελάχιστη κατανάλωση που έχει ανακαλύψει. Άρα κάθε φορά που εκτελεί μία ενέργεια και εναλλάσσει τη πολιτική του, υπολογίζεται από το ΤΝΔ του κεφαλαίου 4 η συνολική καταναλισκόμενη ενέργεια. Έπειτα υπολογίζεται η διαφορά μεταξύ της ελάχιστης κατανάλωσης και της κατανάλωσης της πολιτικής. Τέλος το αποτέλεσμα πολλαπλασιάζεται με μία σταθερά μείωσης, $0 \leq \kappa \leq 1$. Αυτό συμβαίνει διότι πολλές φορές η παραπάνω διαφορά έχει πολύ μεγάλη τιμή και θέλουμε να μειωθεί η σημασία της σε σχέση με το χρονικό βραβείο που όπως παρατηρήθηκε διατηρείται εντός ενός πεπερασμένου εύρους τιμών.

Το χρονικό βραβείο ορίστηκε σταθερό. Πιο συγκεκριμένα εάν η μέση ταχύτητα της πολιτικής είναι εντός του επιθυμητού εύρους το χρονικό βραβείο είναι 0.5 διαφορετικά είναι ίσο με -0.5. Βέβαια κάθε τμήμα της διαδρομής έχει διαφορετική έκταση και γι' αυτό προκειμένου να υπολογίσουμε τη μέση ταχύτητα της πολιτικής βρίσκουμε το σταθμισμένο μέσο όρο των ταχυτήτων.

$$\mu = \frac{\sum_1^n w_i v_i}{\sum_1^n w_i}$$

Το βάρος wi κάθε τμήματος είναι ανάλογο του μήκους του σε σχέση με τη συνολική διαδρομή.

Τελικός έχουμε μία ολοκληρωμένη συνάρτηση επιβράβευσης που λαμβάνει υπόψη της, τη καταναλισκόμενη ενέργεια και τον επιθυμητό χρόνο τερματισμού.

5.4.3 Στρατηγική επιλογής ενέργειας

Σε έναν αλγόριθμο RL, ο στόχος του πράκτορα είναι να μεγιστοποιήσει τα βραβεία που λαμβάνει κατά τη πάροδο του χρόνου. Η μεγαλύτερη δυσκολία που αντιμετωπίζει παρουσιάζεται κατά τη διαδικασία επιλογής της ενέργειας που μεγιστοποιεί το βραβείο που θα δεχτεί ο πράκτορας εκείνη τη στιγμή. Ο πράκτορας γίνεται να εγκλωβιστεί σε κάποιο τοπικό ελάχιστο της διαδικασίας και να χάσει τη βέλτιστη λύση. Για να αντιμετωπιστεί αυτό το πρόβλημα επιτρέπουμε στο πράκτορα να πειραματιστεί με τις τιμές που θα επιλέξει.

Προκειμένου να το πετύχουμε αυτό όταν ο πράκτορας καλείτε να κάνει μία επιλογή δεν τη κάνει μόνο με ένα άπληστο κριτήριο. Αντίθετα ανάλογα με το πόσο εμπιστεύεται τις γνώσεις του κάνει μία επιλογή. Εξισορροπείται δηλαδή η συμπεριφορά του πράκτορα ανάμεσα στην έρευνα και την γνώση. Αυτό το πετυχαίνουμε εισάγοντας μία πιθανοκρατική παράμετρο (p). Η παράμετρος αυτή εξαρτάται από το πλήθος των εποχών (n) που έχουν περάσει και τον ρυθμό εξερεύνησης που αναπαρίσταται από μία σταθερά $0 \leq \varepsilon \leq 1$. Όσο περισσότερες εποχές έχουν περάσει τόσο περισσότερο εμπιστεύεται τις γνώσεις του ενώ στην αρχή της εκπαίδευσης τείνει να κάνει τυχαίες επιλογές. Η μέθοδος αυτή ονομάζεται ε-άπληστη (*e-greedy*).

$$p = e^{-\varepsilon \times n}$$

Στη παραπάνω σχέση ως p έχει οριστεί η πιθανότητα να γίνει μία τυχαία επιλογή ενώ η πιθανότητα επιλογής της ταχύτητας με τη μεγαλύτερη τιμή Q είναι 1-p.

Ψευδοκώδικας αλγορίθμου

Είσοδοι: το σύνολο των καταστάσεων S, το σύνολο των ενεργειών A, τα βραβεία r_t
Είσοδοι: η σταθερά μείωσης (γ), ο ρυθμός εκμάθησης (α) και η παράμετρος άπληστης επιλογής (ε)
Είσοδοι: επιθυμητή μέση ταχύτητα (μ), εύρος ταχυτήτων (α, β)
Αρχικοποιημένος πίνακας Q(s,a) με μηδενικά σε κάθε στοιχείο του

Για όλα τα επεισόδια:
    Υπολογισμός πολιτικής
    Για όλες τις καταστάσεις s:
        Άπληστη επιλογή a από (α, β) με ρυθμό εξερεύνησης ε
        Ενημέρωση πολιτικής με τη νέα ταχύτητα
        Υπολογισμός κατανάλωσης νέας πολιτικής
        Υπολογισμός βραβείου κατανάλωσης και χρονικό βραβείο
        $Q_t(s', a) = Q_{t-1}(s', a) + 1\ \{[r_t(s', a) + \gamma\ maxQ_{t+1}(s', a)] - Q_t(s', a)\}$
        Ενημέρωση βέλτιστης πολιτικής εάν η νέα κατανάλωση είναι μικρότερη
Έως ότου σταθεροποιηθεί η πολιτική

5.5 Πειραματισμός παραμέτρων

Προκειμένου να ξεκινήσει η έρευνα των παραμέτρων έπρεπε να θεσπιστεί ένα κριτήριο τερματισμού. Ως κριτήριο τερματισμού ορίστηκε η μεταβολή της βέλτιστης κατανάλωσης που έχει ανακαλύψει ο πράκτορας. Δηλαδή εάν ο πράκτορας εντός 2000 εποχών δεν έχει ενημερώσει τη βέλτιστη πολιτική του τότε ο αλγόριθμος τερματίζεται διότι θεωρείται πως ο αλγόριθμος έχει συγκλίνει. Το μέγεθος 2000 βρέθηκε μετά από πειραματισμό εντός του εύρους (1000,10000).

Η πρώτη παράμετρος που ορίστηκε αφορούσε το παράγοντα μείωσης (γ) κατά τη διαδικασία ενημέρωσης του πίνακα Q. Η τιμή της καθορίστηκε ίση με μηδέν. Αυτό έγινε διότι στο πρόβλημα που αναλύουμε δεν επιλέγεται η επόμενη κατάσταση όπως είναι σε ένα πρόβλημα εύρεσης βέλτιστης διαδρομής, αντίθετα η επόμενη θέση είναι προκαθορισμένη και επιλέγεται μόνο ο τρόπος που θα πάμε σε αυτή, δηλαδή η ταχύτητα. Άρα ο έλεγχος του βραβείου που θα δεχτεί ο πράκτορας στην επόμενη κατάσταση είναι ανούσιος αφού θα το δεχτεί ανεξάρτητα της επιλογής του.

Η συνάρτηση επιβράβευσης χρόνου εξετάζει εάν η μέση ταχύτητα της πολιτικής του πράκτορα βρίσκεται εντός του αποδεκτού ορίου. Εάν συμβαίνει αυτό τότε επιβραβεύει το πράκτορα με 0,5 ενώ όταν αποκλίνει επιβάλει ποινή -0,5.

Έχοντας ως στόχο την εξισορρόπηση του χρονικού βραβείου με αυτό της κατανάλωσης έγινε μία ανάλυση των τιμών που λαμβάνει η συνάρτηση κατανάλωσης. Παρατηρήθηκε το εξής εύρος τιμών:

$$0{,}913 \leq Κατανάλωση \leq 71{,}57$$

Προφανώς τα παραπάνω νούμερα είναι προϊόν της υπολογιστικής προεπεξεργασίας και δεν αντιστοιχούν αριθμητικά στη πραγματική κατανάλωση αλλά τη περιγράφουν ως ένας δείκτης της. Η διαφορά της βέλτιστης κατανάλωσης με τη κατανάλωση της πολιτικής του πράκτορα βρέθηκε ότι έχει απόλυτη μέση τιμή

$$|E[Διαφορά\ Καταναλώσεων]| = 32{,}369.$$

Με τα παραπάνω στοιχεία ως δεδομένα η σταθερά μείωσης του βραβείο κατανάλωσης ορίστηκε ως

$$\kappa = 0{,}02$$

Με αυτό τον τρόπο το γινόμενο της μέσης τιμής της διαφοράς των καταναλώσεων με τη σταθερά κ είναι λίγο μεγαλύτερο του βραβείου χρόνου.

$$\kappa \times E[Διαφορά\ Καταναλώσεων] = 0{,}64738$$

Τελικός με τις παραπάνω σταθερές επιτύχαμε εξισορρόπηση των δύο βραβείων δίνοντας μεγαλύτερη έμφαση στο βραβείο κατανάλωσης.

Στη συνέχεια ορίστηκε ο ρυθμός εκπαίδευσης ίσος με 0,7. Η επιλογή έγινε από το εύρος (0,1-0,9) με βήμα 0,1. Η επιλογή αυτή οδηγήθηκε από τη στατιστική έρευνα που έγινε και τις αναμενόμενες τιμές που παρουσιάστηκαν προηγουμένως. Βρέθηκε

πως εάν l = 0,7 ο ρυθμός ανανέωσης των τιμών Q ήταν ικανοποιητικός και η σύγκλιση εντός λογικών χρονικών ορίων για την υπολογιστική δύναμη που χρησιμοποιήθηκε.

Η τελευταία παράμετρος που έπρεπε να προσδιοριστεί ήταν αυτή του ρυθμού εξερεύνησης. Προκειμένου να γίνει αυτό εξετάστηκε η ταχύτητα σύγκλισης της μεθόδου και η κατανάλωση της εκάστοτε ακολουθίας με αποδεκτό εύρος ταχυτήτων 23≤v≤27. Στο παρακάτω πίνακα περιγράφονται τα στοιχεία της έρευνας.

Τιμή ε	Ταχύτητα Σύγκλισης σε Εποχές	Κατανάλωση Ακολουθίας
0.001	4030	3.77
0.0001	5060	2.01
0.00001	2780	2.11

Πίνακας 5.1: Αποτελέσματα πειραματισμού με διαφορετικές σταθερές ε.

Από τον παραπάνω πίνακα είναι φανερό πως καλύτερη επιλογή για τη σταθερά ε είναι 0.0001.

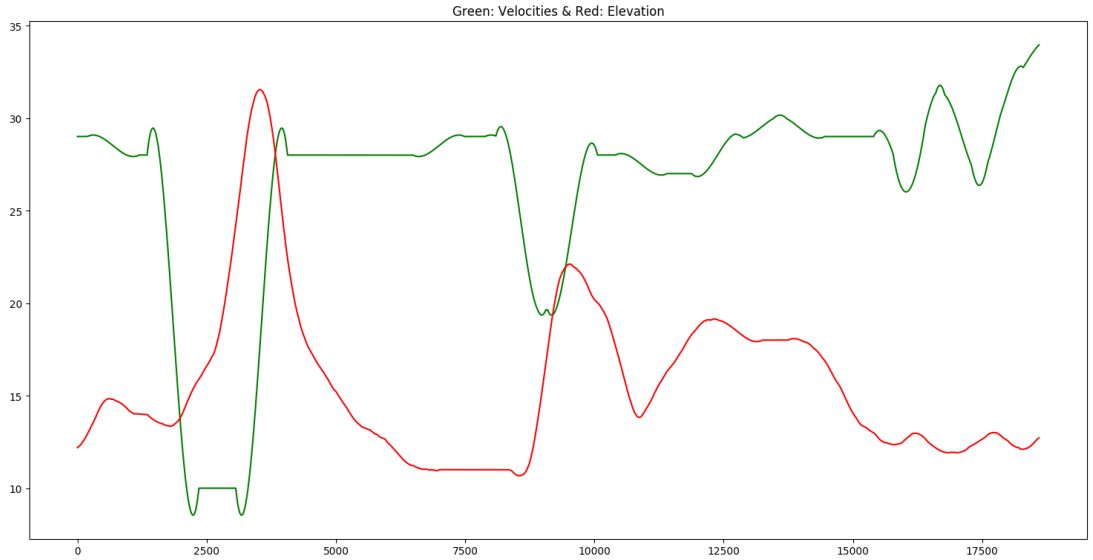
Έχει σημασία να τονιστεί η μεγάλη ταχύτητα σύγκλισης του αλγορίθμου. Σε μόλις 5060 εποχές, δηλαδή υπολογίζοντας 91.080 ακολουθίες είναι ικανός να ανακαλύψει μία ακολουθία που βρίσκεται εντός του αποδεκτού χρονικού εύρους τερματισμού με πολύ μειωμένη κατανάλωση σε σχέση με την ακολουθία σταθερής ταχύτητα 25km/h που έχει κατανάλωση 9.67 σύμφωνα με το ΤΝΔ. Επειδή η διαδρομή έχει χωριστεί σε 18 καταστάσεις και ο πράκτορας μπορεί να επιλέξει ανάμεσα σε 36 διαφορετικές ταχύτητες σε περίπτωση χρήσης ενός αλγορίθμου πλήρης αναζήτησης θα έπρεπε να γίνουν $36^{18} = 10.314.424.798.490.535.546.171.949.056 \approx 1{,}0314\ 10^{28}$ δοκιμές.

Τονίζεται πως η μεταβολή στη κατανάλωση που περιγράφει το ΤΝΔ δεν αντιστοιχεί γραμμικά σε πραγματική μείωση της κατανάλωσης. Αυτό συμβαίνει διότι το όχημα δεν είναι ικανό να αναπτύξει σε πολύ απότομες κλίσεις όλες τις ταχύτητες και γι' αυτό το νευρωνικό δίκτυο αντιστοιχεί ορισμένες ακολουθίες με πολύ μεγάλη κατανάλωση για να μην επιλεχθούν από το πράκτορα όπως αναλύθηκε στο κεφάλαιο 4.

5.6 Παρουσίαση αποτελεσμάτων

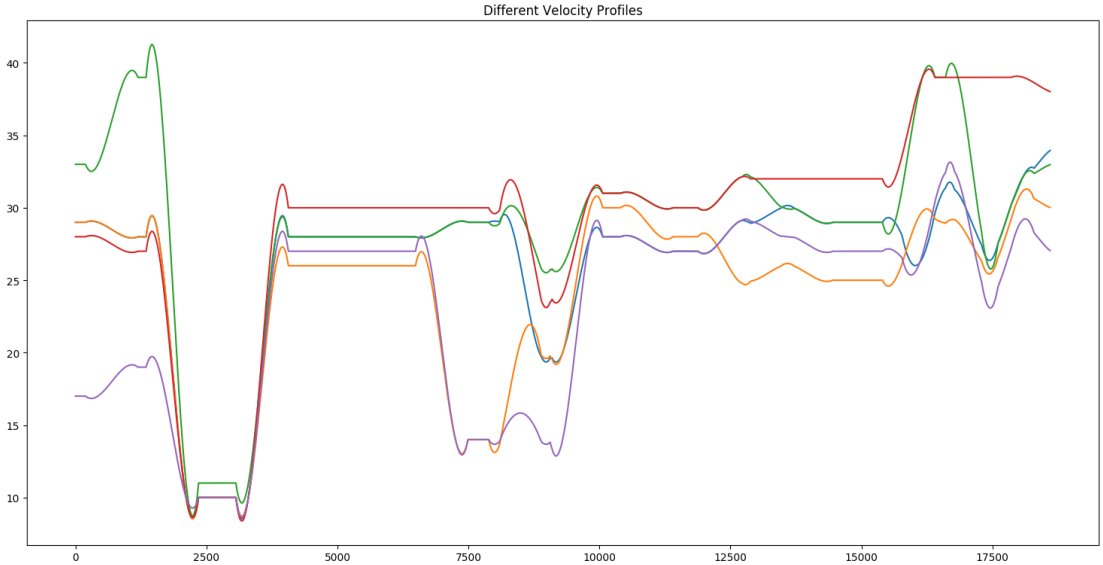
Μετά από τις παραπάνω βελτιστοποιήσεις ο αλγόριθμος σύμφωνα με τη πολυκριτηριακή του βελτιστοποίηση ήταν ικανός να παρουσιάσει ακολουθίες ταχυτήτων που βρίσκονταν σε συμφωνία με την εικόνα που παρουσιάστηκε στο κεφάλαιο 2.

Δηλαδή ο πράκτορας επιλέγει να επιταχύνει στις ανηφόρες και να επιβραδύνει στις κατηφόρες. Η συμπεριφορά αυτή είναι αποδεδειγμένη πως μειώνει τη καταναλισκόμενη ενέργεια ενός αυτοκινήτου[18],[19].



Εικόνα 5.5: Στο παραπάνω γράφημα παρουσιάζεται η επιλογή ταχυτήτων του πράκτορα με πράσινο χρώμα ενώ με κόκκινο χρώμα παρουσιάζεται το υψόμετρο της διαδρομής. Το επιθυμητό εύρος ταχυτήτων έχει οριστεί $23 \leq v \leq 27$ και η ακολουθία που παρουσιάζεται έχει μέση τιμή 26.6km/h.

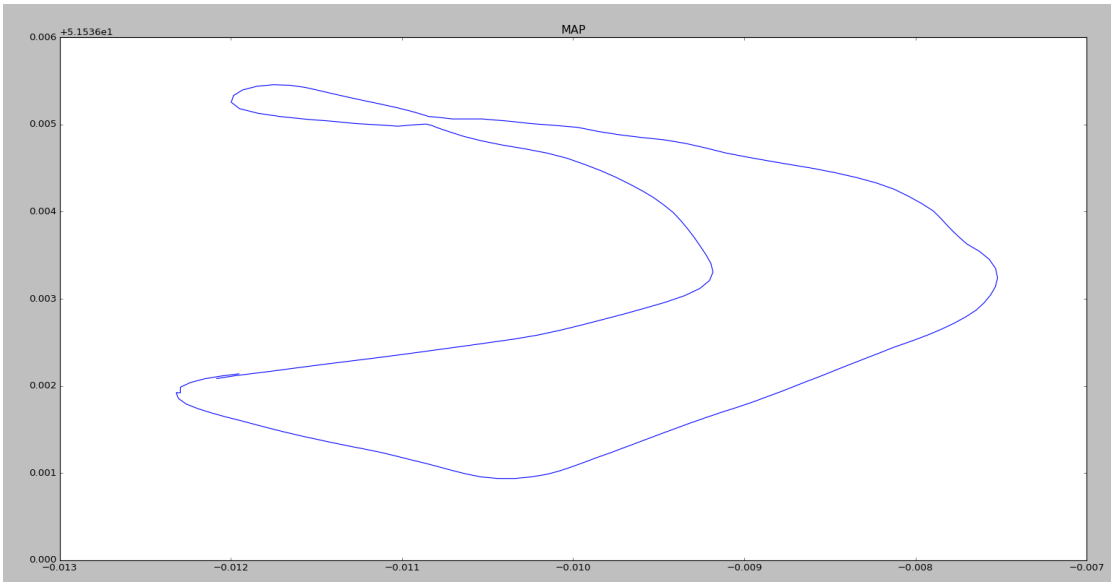
Κατά τη διάρκεια όμως μίας πραγματικής διαδρομής μπορεί να εμφανιστούν αναπάντεχα εμπόδιά τα οποία θα αναγκάσουν το όχημα να επιβραδύνει. Προκειμένου να αντιμετωπιστούν τέτοια περιστατικά ο παραπάνω αλγόριθμος εκτελέστηκε πολλές φορές με διαφορική επιθυμητή μέση τιμή ταχύτητας. Έτσι δημιουργήθηκαν ακολουθίες ταχυτήτων με διαφορετικό χρόνο τερματισμού και σε περίπτωση επιβράδυνσης εκτελείται μία πιο γρήγορη ακολουθία προκειμένου να επιτευχθεί ο επιθυμητός χρόνος τερματισμού, αλλά η επιτάχυνση γίνεται βέλτιστα. Παρακάτω παρουσιάζονται όλες οι ακολουθίες ταχυτήτων που υπολογίστηκαν και είναι φανερό πως ο πράκτορας προκειμένου να κερδίσει χρόνο αποφασίζει να αυξήσει τη ταχύτητα του στις κατηφορικές και όχι στις αναφορικές κλίσεις του οδοστρώματος.



Εικόνα 5.6: Παρουσιάζονται διαφορετικές ακολουθίες βέλτιστων ταχυτήτων με διαφορετικές μέσες τιμές και πιο συγκεκριμένα: 23.2, 24.2, 26.6, 28.8 και 29.6 χιλιόμετρα την ώρα.

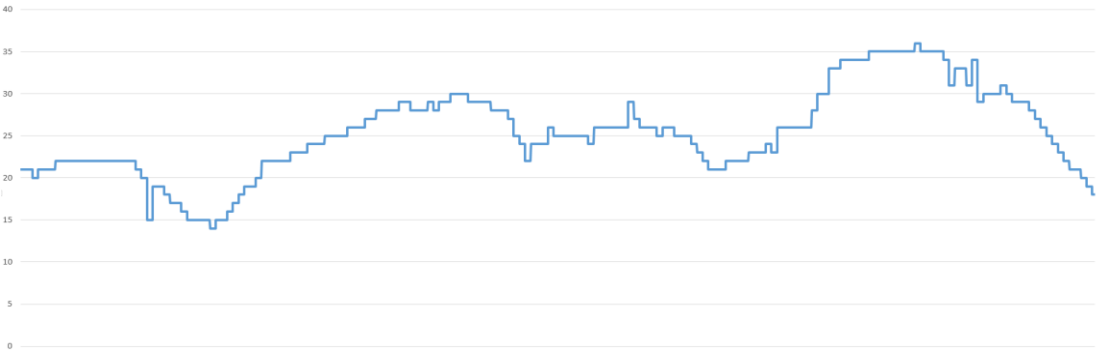
5.7 Πειραματική τεκμηρίωση

Τα δεδομένα που παρουσιάστηκαν παραπάνω αφορούσαν την μελέτη που έγινε σύμφωνα με τα δεδομένα που είχε συλλέξει η τηλεμετρία του οχήματος το έτος 2016. Το έτος 2017 ο αλγόριθμος κλήθηκε να υπολογίσει τις νέες ακολουθίες ταχυτήτων που θα επέφεραν μέγιστη εξοικονόμηση ενέργειας κατά τη διάρκεια του διαγωνισμού σε διαφορετική διαδρομή.

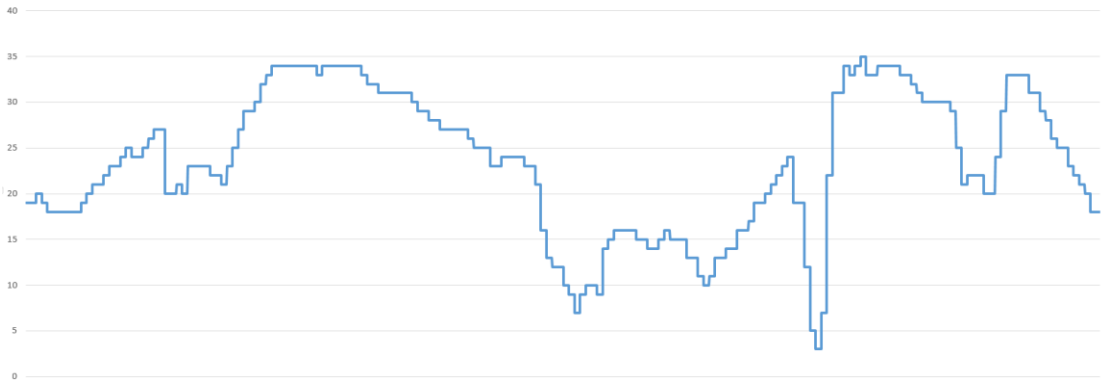


Εικόνα 5.7: Η πίστα του μαραθωνίου οικονομίας 2017 σύμφωνα με τα δεδομένα που συνέλεξε η τηλεμετρία του οχήματος.

Προκειμένου να δοκιμαστεί η βέλτιστη ακολουθία ταχυτήτων που υπολογίστηκε ζητήθηκε από τον οδηγό να διατρέξει μία φορά τη πίστα με σταθερή ταχύτητα 25 χιλιόμετρα την ώρα και στη συνέχεια να εκτελέσει τη στρατηγική που παρουσιάστηκε παραπάνω με την αντίστοιχη μέση ταχύτητα. Με αυτό τον τρόπο δοκιμάστηκε η παρουσιαζόμενη ακολουθία ταχυτήτων έναντι της στρατηγικής που χρησιμοποιήται από όλα τα εμπορικά συστήματα πλοήγησης (cruise control systems).

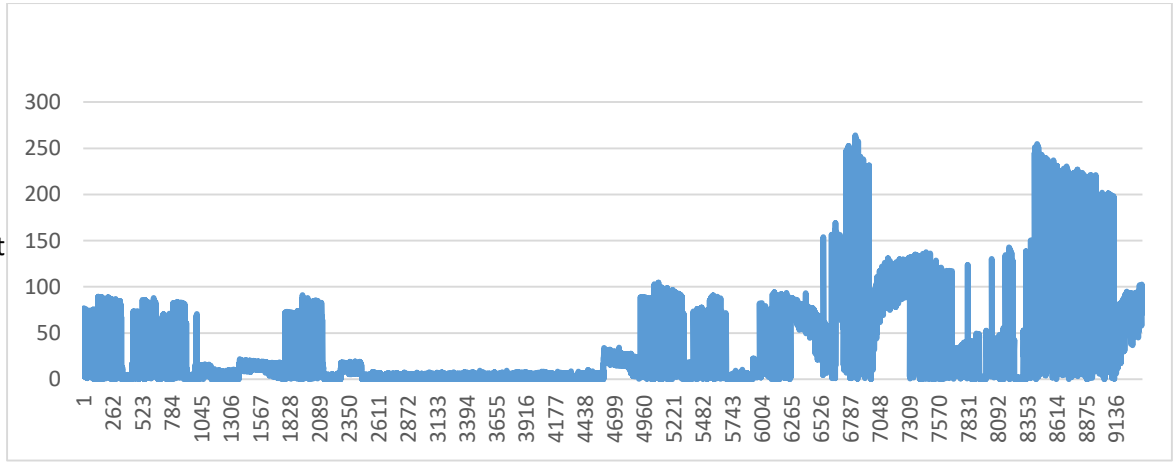


Εικόνα 5.7: Το παραπάνω γράφημα παρουσιάζει τη προσπάθεια του οδηγού να διατηρήσει μία σταθερή ταχύτατα στην απαιτητική πίστα του μαραθωνίου οικονομίας 2017. Τονίζεται πως ο οδηγός δεν επιβραδύνει το όχημα όταν το οδόστρωμα είναι κατηφορικό και γι' αυτό στο τελευταίο τμήμα της διαδρομή το όχημα επιταχύνει.

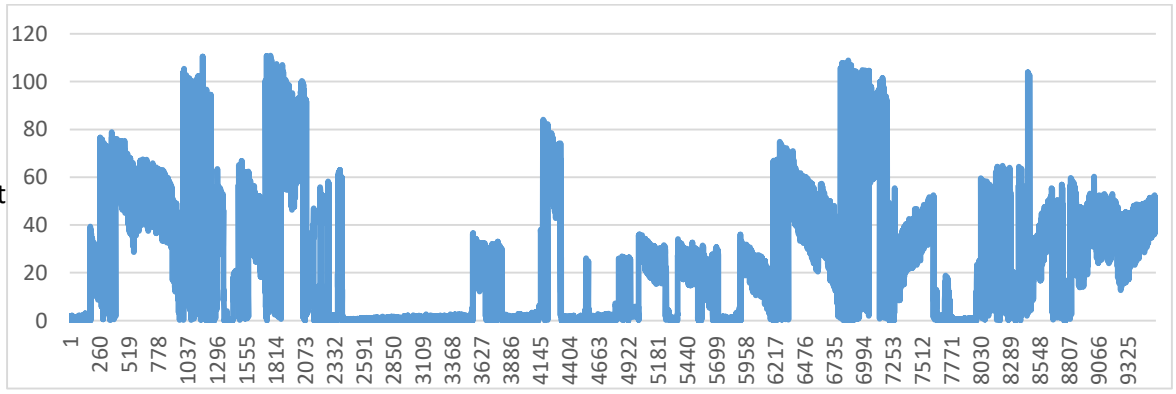


Εικόνα 5.8: Σε αυτό το γράφημα απεικονίζεται η διατηρούμενη ταχύτητα του οχήματος όταν το σύστημα είναι εν' ενεργεία και ο οδηγός προσπαθεί να τηρήσει τη βέλτιστη στρατηγική εξοικονόμησης ενέργειας.

Το πείραμα που διεξάχθηκε και περιγράφηκε παραπάνω έδειξε πως όταν το όχημα διατηρεί τις ταχύτητες που προτείνει ο ευφυής πράκτορας καταναλώνεται **24.0306%** λιγότερη ενέργεια από το κινητήριο σύστημα. Το παραπάνω αποτέλεσμα αποδεικνύει την αποδοτική λειτουργεία του παρουσιαζόμενου συστήματος και επιδεικνύει την αποτελεσματικότητα του σε διαφορετικές διαδρομές.



Εικόνα 5.9: Παρουσίαση της κατανάλωσης του οχήματος όταν διατηρείται σταθερή ταχύτατα.



Εικόνα 5.10: Παρουσίαση της κατανάλωσης του οχήματος όταν το όχημα ακολουθεί τη βέλτιστη ακολουθία ταχυτήτων.

Στη συνέχεια παρουσιάζεται ένα σύστημα παραπλήσιο με αυτό που πραγματεύεται η συγκεκριμένη ενότητα το οποίο είναι ικανό να υπολογίσει την βέλτιστη πορεία που οφείλει να διατηρήσει ένα όχημα πάνω σε μία διαδρομή προκειμένου να αξιοποιήσει τη βέλτιστη ακολουθία ταχυτήτων που υπολογίστηκε.

Κεφάλαιο 6

Υπολογισμός αγωνιστικής γραμμής

6.1 Εισαγωγή

Σε κάθε αγώνισμα μηχανοκίνητου αθλητισμού ο οδηγός προκειμένου να αναπτύξει μέγιστη ταχύτητα προσπαθεί να τοποθετήσει το όχημα του στη βέλτιστη θέση πάνω στη πίστα. Η διαδρομή που εμπεριέχει όλες αυτές τις θέσεις ονομάζεται αγωνιστική γραμμή.

Στην εφαρμογή, την οποία πραγματεύεται αυτή η εργασία, ο στόχος δεν είναι η μέγιστη ταχύτητα αλλά η βέλτιστη, με γνώμονα την εξοικονόμηση ενέργειας. Αυτό όμως προϋποθέτει την αύξηση της ταχύτητας σε κατηφορικά οδοστρώματα. Εάν, λοιπόν, η βέλτιστη ακολουθία ταχυτήτων συμβουλεύει τον οδηγό να επιταχύνει στην κατηφόρα και συμπέσει με στροφή τότε αυτός πρέπει να γνωρίζει πως θα συμπεριφερθεί προκειμένου να συνεχίσει με ασφάλεια. Γίνεται φανερό, λοιπόν, πως υπάρχουν συνθήκες που η παραμονή στην αγωνιστική γραμμή είναι απαραίτητη.

Για αυτό το λόγο αναπτύχθηκε ένας αλγόριθμος εύρεσης της αγωνιστικής γραμμής. Ο αλγόριθμος αυτός χρησιμοποιεί τις ίδιες μεθόδους τεχνητής νοημοσύνης που εφαρμόστηκαν στην εύρεση της βέλτιστης ακολουθίας ταχυτήτων.

Στη συνέχεια παρουσιάζεται ο αλγόριθμος επιβραβευόμενης μάθησης που χρησιμοποιήθηκε, η διαδικασία βελτιστοποίησης των παραμέτρων και τέλος αναλύονται τα αποτελέσματα.

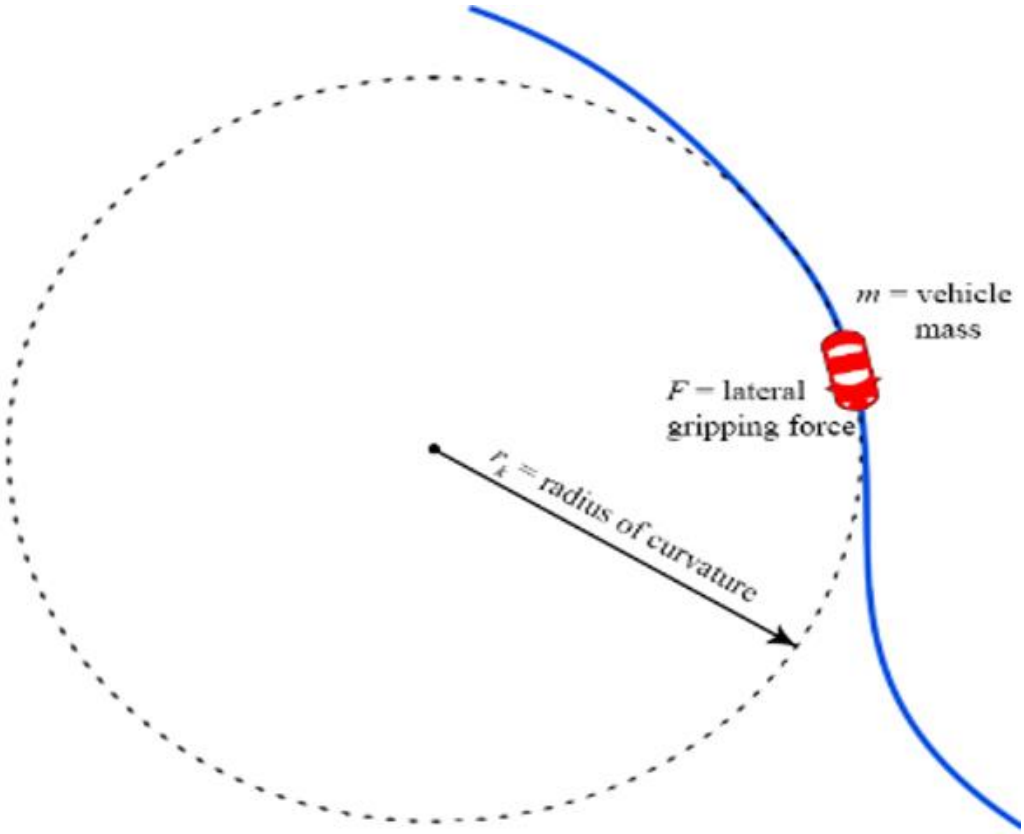
6.2 Τι είναι αγωνιστική γραμμή

Προκειμένου να μπορέσουμε να αναπτύξουμε μία μεθοδολογία που θα είναι ικανή να εξάγει την αγωνιστική γραμμή για μία διαδρομή πρέπει πρώτα να ορίσουμε την αγωνιστική γραμμή μαθηματικά.

Η τελική ταχύτητα ενός αυτοκινήτου δεν επηρεάζεται μόνο από τη δύναμη του κινητήρα αλλά και από τις εξωτερικές δυνάμεις που δέχεται. Πιο συγκεκριμένα, είναι αποδεδειγμένο πως η τελική ταχύτητα ενός οχήματος περιορίζεται από τη μάζα του (m), τη μέγιστη δύναμη συγκράτησης που μπορούν να ασκήσουν τα ελαστικά στο οδόστρωμα (F) και την ακτίνα του κύκλου που εφάπτεται στη διαδρομή (r).

$$v_{max} = \sqrt{\frac{F \times r_k}{m}}$$

Προφανώς η μάζα του οχήματος και η μέγιστη δύναμη αυτοσυγκράτησης είναι δεδομένοι αριθμοί και δεν μεταβάλλονται. Ο παράγοντας που αλλάζει είναι η ακτίνα του κύκλου που εφάπτεται στη διαδρομή, με άλλα λόγια η ίδια η διαδρομή. Όπως βλέπουμε από την παραπάνω εξίσωση, υπάρχει μία ανάλογη σχέση μεταξύ της μέγιστης ταχύτητας και της ακτίνας, δηλαδή όσο μεγαλώνει η ακτίνα τόσο μεγαλώνει και η μέγιστη ταχύτητα. Άρα ο στόχος, προκειμένου να βρεθεί η αγωνιστική γραμμή, είναι να μεγιστοποιηθεί η ακτίνα. Προφανώς εάν η διαδρομή είναι ευθεία τότε η ακτίνα του κύκλου είναι άπειρη και η μέγιστη ταχύτητα δεν περιορίζεται από τη διαδρομή [20].



Εικόνα 6.1: Προσδιορισμός δυνάμεων που επηρεάζουν τη τελική ταχύτητα ενός οχήματος.

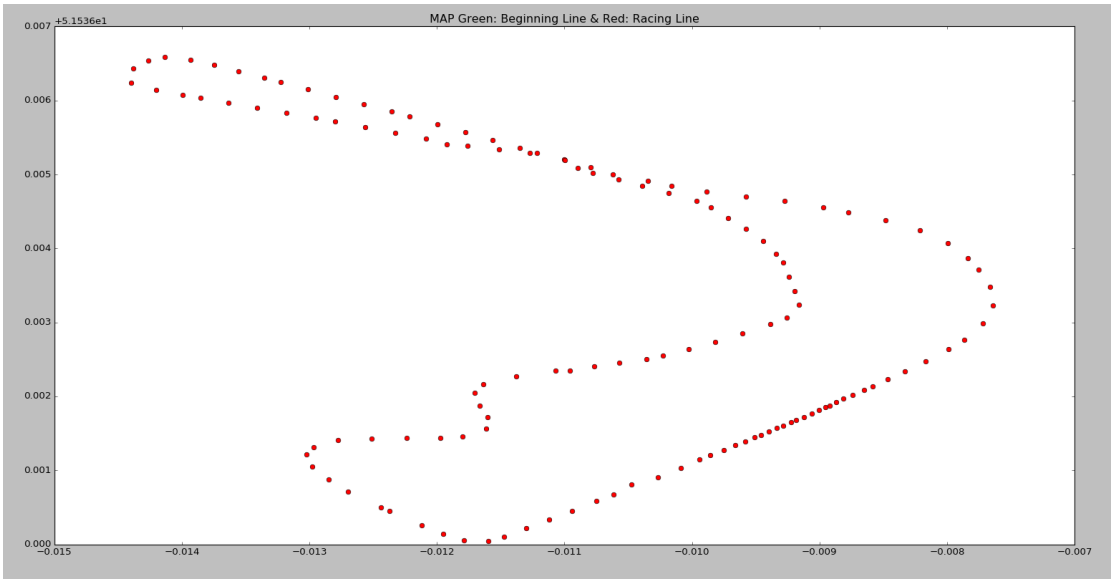
6.3 Δεδομένα εισόδου

Τα δεδομένα της διαδρομής από τα οποία θα υπολογίσουμε την αγωνιστική γραμμή προέρχονται από το σύστημα παγκόσμιας στιγματοθέτησης (GPS) του οχήματος και από τους αισθητήρες τεχνολογίας lidar.

Ως πρώτο βήμα της διαδικασίας είναι η χαρτογράφηση της πίστας. Ο οδηγός διασχίζει πρώτη φορά τη διαδρομή και μέσω του GPS παράγεται ο χάρτης της. Η διαδρομή που παρουσιάζεται παρακάτω έχει μήκος 2.240 μέτρα και απαρτίζεται από 19.950 μετρήσεις. Κατά τη διαδικασία εύρεσης της αγωνιστικής γραμμής ο αλγόριθμος θα έχει τη δυνατότητα να μετακινεί όλα αυτά τα σημεία αλλάζοντας τη πορεία του οχήματος. Ο αλγόριθμος θα τερματίζει όταν όλα τα σημεία βρίσκονται στις βέλτιστες θέσεις τους, δηλαδή στις θέσεις που μεγιστοποιούν την ακτίνα του κύκλου που εφάπτεται στη διαδρομή. Ωστόσο, ο όγκος των δεδομένων καθιστά τη διαδικασία αυτή υπολογιστικά πολύπλοκη. Κρίθηκε λοιπόν απαραίτητη η μείωση του πλήθους των μετρήσεων χωρίς όμως να αλλοιωθεί ο χάρτης της διαδρομής.

Προκειμένου να πραγματοποιηθεί μία σμίκρυνση του πλήθους των μετρήσεων αποφασίστηκε να διατηρηθεί ένα σημείο για κάθε n μετρήσεις της πίστας. Για το προσδιορισμό της σταθεράς n έγινε πειραματισμός ο οποίος βασίστηκε στην ύπαρξη σημείων με ίδιες συντεταγμένες. Δηλαδή το n μεγάλωνε έως ότου δεν υπήρχαν πλέον μετρήσεις με τις ίδιες συντεταγμένες. Ακολουθήθηκε αυτή η μεθοδολογία διότι στο επόμενο βήμα που υπολογίζεται η ακτίνα του εφαπτόμενου κύκλου πραγματοποιούνται αφαιρέσεις μεταξύ των συντεταγμένων των σημείων και η ύπαρξη μηδενικών προκαλούσε απροσδιοριστία όπως περιγράφεται παρακάτω. Τελικώς, η σταθερά ορίστηκε ίση με 140 και το τελικό πλήθος των σημείων που προέκυψε ήταν 137. Το πλήθος αυτό οδήγησε σε απλοποίηση της πολυπλοκότητας του αλγορίθμου και επιτάχυνε τη διαδικασία σύγκλισης.

Τα σημεία αυτά αναπαριστούν τη διαδρομή του οχήματος και θεωρείται πως περνάνε από τη μέση της πίστας.



Εικόνα 6.2: Αναπαράσταση των σημείων της πίστας μετά από τη μείωση του πλήθους τους.

6.4 Επεξεργασία μετρήσεων

Όπως περιγράφεται παραπάνω προκειμένου να υπολογιστεί η αγωνιστική γραμμή πρέπει τα σημεία που απαρτίζουν τη πίστα να μπορούν να μετακινούνται κατά τη διεύθυνση του πλάτους της και να είμαστε ικανοί να υπολογίσουμε την ακτίνα του κύκλου που εφάπτεται στη διαδρομή. Για όλους τους υπολογισμούς που παρουσιάζονται παρακάτω ως συντεταγμένες χρησιμοποιήθηκαν το γεωγραφικό μήκος και πλάτος των μετρήσεων.

Αρχικά, προκειμένου να βρεθεί η διεύθυνση μετακίνησης των σημείων υπολογίζεται η ευθεία που ενώνει το σημείο που θέλουμε κάθε φορά να μετακινήσουμε (x_1,y_1) με το επόμενο (x_2,y_2).

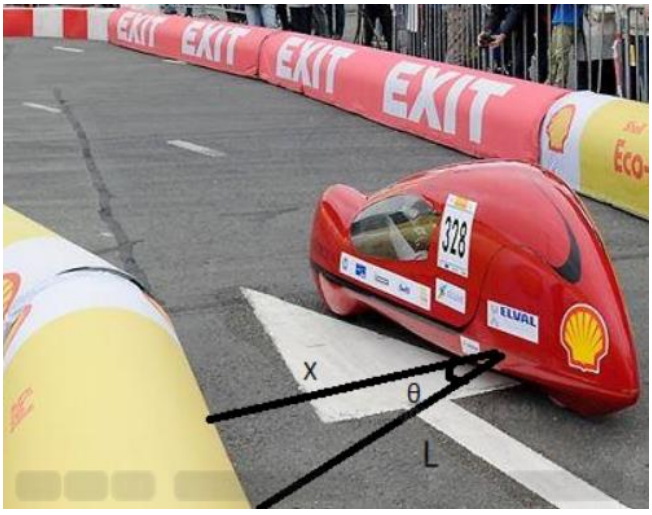
$$\lambda_1 = \frac{y_2 - y_1}{x_2 - x_1} \qquad y = \lambda_1(x - x_1) + y_1$$

Η ευθεία αυτή αντιπροσωπεύει την πορεία του οχήματος. Στη συνέχεια, υπολογίζεται η κάθετη ευθεία που περνάει από το πρώτο σημείο. Η δεύτερη ευθεία αναπαριστά τη διεύθυνση μετακίνησης του σημείου, δηλαδή τη θέση του οχήματος στο πλάτος της πίστας.

$$\lambda_2 = \frac{-1}{\lambda_1} \qquad y = \lambda_2(x - x_1) + y_1$$

Έχοντας, λοιπόν, τη διεύθυνση μετακίνησης του σημείου μπορούμε να το μετακινήσουμε και να μεταβάλουμε τη τελική πορεία του οχήματος.

Το μέγεθος αυτής της μετακίνησης εξαρτάται από το πλάτος της διαδρομής. Το πλάτος μετριέται από τους αισθητήρες lidar που είναι τοποθετημένοι στα πλαϊνά του αυτοκινήτου.



Εικόνα 6.3: Απεικόνιση λειτουργίας αισθητήρων lidar.

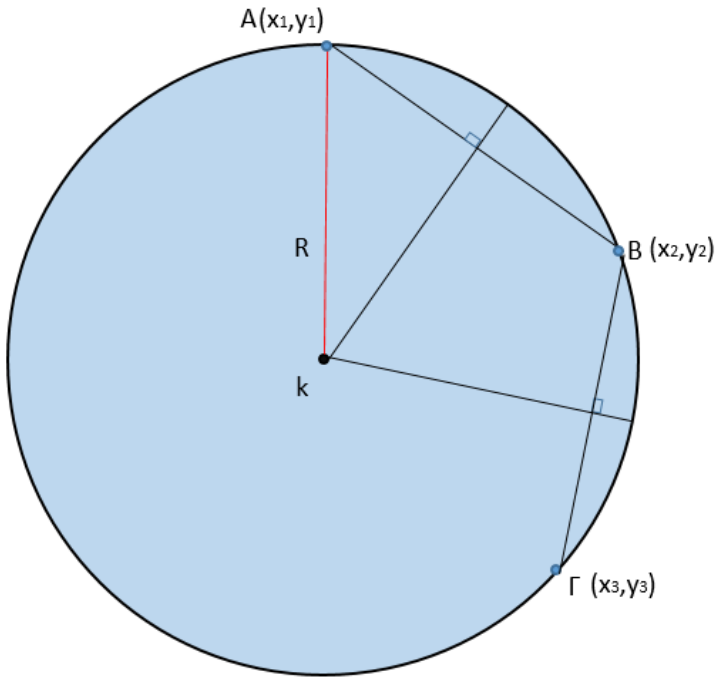
Οι αισθητήρες είναι ικανοί να μετρήσουν με ακρίβεια την απόσταση L της φωτογραφίας. Έχουν τοποθετηθεί υπό μία γνωστή γωνία θ σε σχέση με το αμάξωμα που θεωρείται πως κινείται κάθετα σε σχέση με τη διεύθυνση του πλάτους της πίστας. Συνεπώς ο υπολογισμός του πλάτους είναι θέμα επίλυσης της παρακάτω εξίσωσης:

$$x = L \times \cos\theta$$

Εφόσον υπάρχει διαθέσιμο το πλάτος της πίστας γνωρίζουμε πόσο μπορεί να μετακινηθεί το σημείο πάνω στη δεύτερη ευθεία που υπολογίστηκε.

Ο επόμενος υπολογισμός που πρέπει να γίνει είναι η εύρεση της εξίσωσης του κύκλου που εφάπτεται την διαδρομή. Επειδή, όμως, υπάρχουν πολλοί διαφορετικοί κύκλοι που εφάπτονται σε δύο σημεία αυτή τη φορά χρησιμοποιούνται τα επόμενα τρία σημεία για τους υπολογισμούς, δηλαδή αυτό που πρόκειται να μετακινηθεί και τα δύο επόμενα. Προκειμένου να υπολογίσουμε την εξίσωση του κύκλου αυτού πρώτα βρίσκουμε τις εξισώσεις των δύο ευθείων που ενώνουν το πρώτο σημείο με το δεύτερο (ΑΒ) και το δεύτερο με το τρίτο (ΒΓ). Ο κύκλος που εφάπτεται σε τρία σημεία έχει κέντρο το σημείο τομής των μεσοκαθέτων των παραπάνω ευθύγραμμων τμημάτων. Αφού υπολογίσουμε λοιπόν τις μεσοκαθέτους βρίσκουμε το σημείο τομής τους, δηλαδή τις συντεταγμένες του κέντρου του κύκλου. Το τελευταίο βήμα είναι να υπολογίσουμε την ακτίνα του κύκλου που ισούται με την απόσταση από το κέντρο k έως το Α. Για όλους τους υπολογισμούς των ευθειών χρησιμοποιήθηκαν οι εξισώσεις που περιγράφονται παραπάνω, και για την ακτίνα η επόμενη σχέση:

$$R = \sqrt{(y_k - y_A)^2 + (x_k - x_A)^2}$$



Εικόνα 6.4: Εφαπτόμενος κύκλος σε τρία σημεία. Οι μεσοκάθετες των ευθειών ΑΒ και ΒΓ τέμνονται στο κέντρο του κύκλου.

6.5 Ανάπτυξη αλγορίθμου

Όπως τονίστηκε παραπάνω, προκειμένου να υπολογιστεί η αγωνιστική γραμμή χρησιμοποιήθηκαν παρόμοιες μέθοδοι τεχνητής νοημοσύνης με τον υπολογισμό των βέλτιστων ακολουθιών ταχύτητας που περιγράφεται στο κεφάλαιο 5. Έγινε μία υλοποίηση ενός αλγορίθμου *Q Learning* ,όπου ο ευφυής πράκτορας επισκέπτεται όλα τα σημεία της πίστας και επιλέγει να τα τοποθετήσει σε μία θέση σε σχέση με το πλάτος της. Έπειτα, η ενέργεια του αξιολογείται από το περιβάλλον και του δίνεται ένα βραβείο από τη συνάρτηση επιβράβευσης. Τα χαρακτηριστικά του αλγορίθμου επιγράφονται παρακάτω.

6.5.1 Διαστάσεις καταστάσεων (s) και ενεργειών (a)

Οι καταστάσεις του χώρου αναπαριστούν τα χαρακτηριστικά του περιβάλλοντος που είναι απαραίτητα για τη λύση του προβλήματος. Στη δική μας περίπτωση οι θέσεις του σημείου που επεξεργαζόμαστε και των δύο επόμενων είναι οι σημαντικοί παράγοντες για τον υπολογισμό της αγωνιστικής γραμμής.

$$S = \{s_i, s_j, s_k\}:\ s_i\ \in S \qquad \forall i, j, k \in \{1,2, \dots, n\}$$

Όπου s είναι η θέση των σημείων i,j,k. Όπως τονίστηκε προηγουμένως, αρχικά θεωρείται πως τα σημεία βρίσκονται στο κέντρο της πίστας και ο αλγόριθμος μπορεί να τα μετακινήσει έως τα όρια του πλάτους της. Η πίστα έχει μέσο πλάτος έξι μέτρα και γι' αυτό ως βήμα μετακίνησης θεωρήθηκε το μέτρο, δηλαδή ο αλγόριθμος μετακινεί το κάθε σημείο ένα μέτρο τη φορά. Σε γεωγραφικές συντεταγμένες το ένα μέτρο μεταφράζεται σε μετακίνηση κατά περίπου 10^{-9} για τη τοποθεσία που έγινε το πείραμα.

6.5.2 Πίνακας Q

Η μεγάλη διαφορά με τον προηγούμενο αλγόριθμο που παρουσιάστηκε είναι πως ο πράκτορας έχει διαφορετικό πλήθος ενεργειών να επιλέξει σε κάθε σημείο. Αυτό συμβαίνει διότι κάθε τμήμα της πίστας έχει διαφορετικό πλάτος και εφόσον το βήμα μετακίνησης είναι σταθερό σε κάποιες περιπτώσεις τα σημεία μπορούν να απομακρυνθούν περισσότερο από την αρχική τους θέση σε σχέση με άλλα. Ωστόσο, ο πίνακας Q έχει σταθερές διαστάσεις, γι’ αυτό σχεδιάστηκε μία συνάρτηση μέσα στον αλγόριθμο που ελέγχει το τμήμα που είναι κάθε φορά ο πράκτορας και τον περιορίζει στις ενέργειες που μπορεί να επιλέξει σύμφωνα με τον περιορισμό του πλάτους. Ο πίνακας Q έχει 137 γραμμές, όσες είναι και οι μετρήσεις (καταστάσεις) και 8 στήλες, όσο είναι το μεγαλύτερο πλάτος που παρατηρήθηκε, δηλαδή 8 μέτρα.

6.5.3 Συνάρτηση επιβράβευσης

Η συνάρτηση επιβράβευσης είναι το μόνο μέσο που έχει ο πράκτορας για να αξιολογήσει κάθε φορά την ενέργεια που αποφασίζει να εκτελέσει. Το επιθυμητό είναι να μεγιστοποιήσει την ακτίνα του κύκλου που εφάπτεται στη διαδρομή. Όπως

περιγράφηκε παραπάνω όταν υπολογίζονται οι ακτίνες των εφαπτόμενων κύκλων χρησιμοποιούνται τρία σημεία.

Προκειμένου να αξιολογηθεί η ενέργεια του πράκτορα υπολογίζονται οι ακτίνες 3 εφαπτόμενων κύκλων. Ο πρώτος γίνεται με τη βοήθεια των σημείων που έπονται και του σημείου ενδιαφέροντος, ο δεύτερος είναι μεταξύ των επόμενων δύο σημείων και του σημείου ενδιαφέροντος και ο τελευταίως τοποθετεί το σημείο ενδιαφέροντος στη μέση και χρησιμοποιεί τα διπλανά του. Αυτό έγινε προκειμένου ο πράκτορας να δέχεται μία πιο ολοκληρωμένη εικόνα της συνέπειας της ενέργειας του.

Τελικώς, το βραβείο υπολογίζεται ως το άθροισμα των διαφορών των ακτίνων που προκλήθηκαν από τη μετακίνηση, και των ακτίνων της πολιτικής του πράκτορα.

$$r = (r_{μπροστά} - r_{πολιτική μπορστά}) \times \alpha + (r_{πίσω} - r_{πολιτική πίσω}) \times \beta + (r_{μέση} - r_{πολιτική μέση}) \times \gamma$$

Όπου $r_{μπροστά}$, $r_{πίσω}$ και $r_{μέση}$ είναι οι ακτίνες των κύκλων που εφάπτονται στις διαδρομές που σχηματίζουν οι αντίστοιχες τριάδες. Ενώ τα $r_{πολιτική μπορστά}$, $r_{πολιτική πίσω}$ και $r_{πολιτική μέση}$ αφορούν τις ακτίνες των αντίστοιχων κύκλων που εφάπτονται στην πολιτική του πράκτορα. Τέλος τα α, β, γ είναι συντελεστές εντός του εύρους τιμών [0,1] προκειμένου να επηρεάσουν τη σημασία κάθε διαφοράς.

6.5.3 Στρατηγική επιλογής ενέργειας

Όπως και στον αλγόριθμο του κεφαλαίου 5 ο πράκτορας, όταν πρέπει να πάρει μία απόφαση για τη μετακίνηση των σημείων δεν επιλέγει άπληστα το σημείο με τη μεγαλύτερη τιμή Q. Αντιθέτως, και σε αυτή τη περίπτωση χρησιμοποιείται η ε-άπληστη στρατηγική που περιγράφηκε.

Στη συνέχεια ακολουθεί η διαδικασία επιλογής των παραμέτρων που βελτιστοποιούν τον αλγόριθμο.

Ψευδοκώδικας αλγορίθμου

Είσοδοι: το σύνολο των καταστάσεων S, το σύνολο των ενεργειών A, το βραβείο r_t
Είσοδοι: η σταθερά μείωσης (γ), ο ρυθμός εκμάθησης (α) και η παράμετρος άπληστης επιλογής (ε)
Είσοδοι: τα μέγιστα πλάτη όλων των καταστάσεων
Αρχικοποιημένος πίνακας Q(s,a) με μηδενικά σε κάθε στοιχείο του

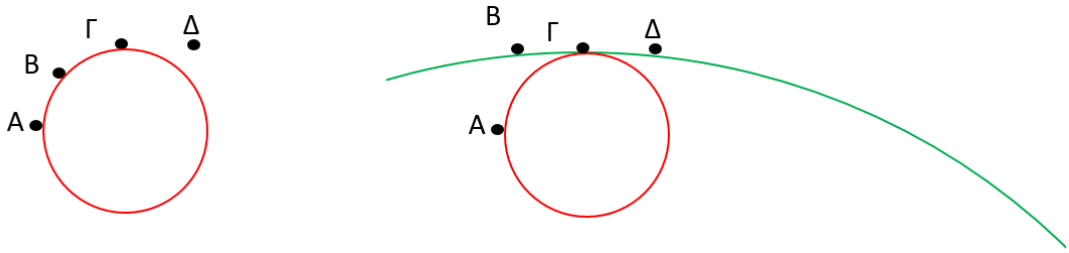
Για όλα τα επεισόδια:
    Υπολογισμός πολιτικής
    Για όλες τις καταστάσεις s:
        Άπληστη επιλογή a με ρυθμό εξερεύνησης ε
        Ενημέρωση πολιτικής με τη νέα θέση σημείου
        Υπολογισμός νέας ακτίνας εφαπτόμενου κύκλου
        Υπολογισμός βραβείου
        $Q_t(s', a) = Q_{t-1}(s', a) + 1\ \{[r_t(s', a) + \gamma\ maxQ_{t+1}(s', a)] - Q_t(s', a)\}$
        Ενημέρωση βέλτιστης πολιτικής s εάν η νέα ακτίνα είναι μεγαλύτερη
Έως ότου σταθεροποιηθεί η πολιτική

6.6 Πειραματισμός παραμέτρων

Προκειμένου να ξεκινήσει η έρευνα των παραμέτρων έπρεπε να θεσπιστεί ένα κριτήριο τερματισμού. Ως κριτήριο τερματισμού ορίστηκε η μεταβολή της βέλτιστης πολιτικής του πράκτορα. Δηλαδή εάν ο πράκτορας εντός 1000 εποχών δεν έχει ανακαλύψει καλύτερη πολιτική τότε ο αλγόριθμος τερματίζεται διότι θεωρείται πως έχει συγκλίνει στη βέλτιστη τοποθέτηση των σημείων.

Όπως και στον αλγόριθμο του κεφαλαίου 5, η επόμενη κατάσταση είναι δεδομένη και δεν επηρεάζεται από την ενέργεια του πράκτορα. Γι' αυτό το λόγο και σε αυτή τη περίπτωση η σταθερά μείωσης γ πήρε μηδενική τιμή.

Παρόλο που η επιλογή του πράκτορα δεν επηρεάζει ποια κατάσταση θα επισκεφτεί στο επόμενο βήμα, τροποποιεί σημαντικά ένα μεγάλο τμήμα της διαδρομής. Αυτό συμβαίνει διότι η μεταβολή ενός σημείου αλλοιώνει το μέγεθος της ακτίνας του εφαπτόμενου κύκλου στην υπόλοιπη διαδρομή. Αυτό παρουσιάζεται στην εικόνα 6.5.



Εικόνα 6.5: Το Α τοποθετείται στη θέση που φαίνεται στο αριστερό σχήμα και σχηματίζεται ο κόκκινος εφαπτόμενος κύκλος. Στο επόμενο βήμα μετακινείται το σημείο Β. Αυτό έχει ως αποτέλεσμα στα σημεία Β,Γ,Δ να εφάπτεται ο πράσινος κύκλος, αλλά η ακτίνα που υπολογίστηκε στο προηγούμενο βήμα, του κόκκινου κύκλου, να μην ισχύει πλέον αφού έχει αλλοιωθεί η διαδρομή και έχει μεταβληθεί ο εφαπτόμενος κύκλος.

Για να καταπολεμηθεί το παραπάνω πρόβλημα αποφασίστηκε να χρησιμοποιηθεί ένα πολύ μικρό βήμα εκπαίδευσης ως μία προσπάθεια εξισορρόπησης των βέλτιστων θέσεων των σημείων σε διαφορετικές τριάδες. Αυτό συμβαίνει διότι οι τιμές του Q ενημερώνονται με πολύ αργό ρυθμό με συνέπεια το τελικό αποτέλεσμα να αποτελεί μία μέση λύση των βέλτιστων θέσεων των μετρήσεων για κάθε τριάδα. Πιο συγκεκριμένα το βήμα εκπαίδευσης τέθηκε ίσο με 0.001.

Στη συνέχεια εξετάστηκε η συνάρτηση επιβράβευσης και βρέθηκε πως τα καλύτερα αποτελέσματα έπονται όταν τα α, β, γ παίρνουν τις τιμές 1, 1 και 0.2 αντίστοιχα. Ο προσδιορισμός αυτών των τιμών έγινε πειραματικά από το διάστημα [0,1] με βήμα 0.1. Οι τιμές αυτές επηρεάστηκαν από το γεγονός πως η πίστα είναι δεξιόστροφη και δεν περιλαμβάνει αριστερές στροφές. Σε κάποια άλλη διαδρομή οι τιμές αυτές μπορεί να χρειάζονται τροποποίηση.

Η τελευταία παράμετρος που έπρεπε να προσδιοριστεί ήταν αυτή του ρυθμού εξερεύνησης. Προκειμένου να γίνει αυτό, εξετάστηκε η ταχύτητα σύγκλισης της μεθόδου, το ποσοστό των κελιών του πίνακα Q που επισκέφτηκε ο πράκτορας και η κυματομορφή της εκάστοτε ακολουθίας. Στον παρακάτω πίνακα περιγράφονται τα στοιχεία της έρευνας.

Τιμή ε	Ταχύτητα Σύγκλισης σε Εποχές	% Κελιών του πίνακα Q που εξερευνήθηκαν
0.01	1.800	84.877
0.001	9.600	85.34
0.0001	68.000	85.031
0.00001	999.996	94.256

Πίνακας 6.1: Αποτελέσματα πειραματισμού με διαφορετικές σταθερές ε.

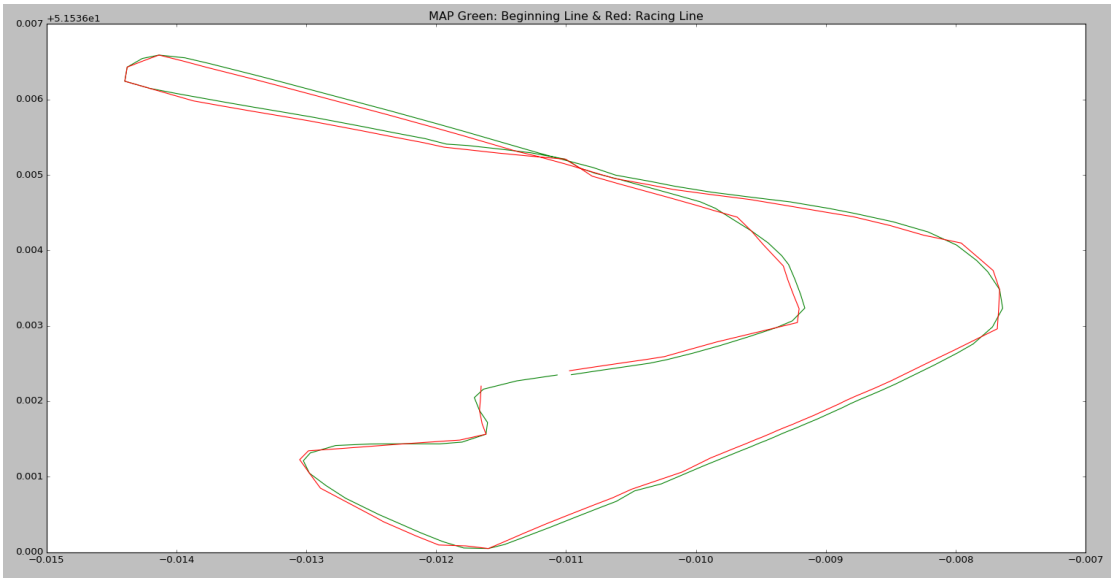
Από τον παραπάνω πίνακα επιλέχθηκε η τιμή 0.001 διότι συγκλίνει σε ένα λογικό αριθμό εποχών και είχε τα καλύτερα αποτελέσματα όσων αφορά τις θέσεις των σημείων, όπως παρουσιάζονται παρακάτω.

Έχει σημασία να τονιστεί η μεγάλη ταχύτητα σύγκλισης του αλγορίθμου. Σε μόλις 9600 εποχές, δηλαδή υπολογίζοντας 1.315.200 ακολουθίες είναι ικανός να ανακαλύψει την αγωνιστική γραμμή. Η μεγάλη ταχύτητα σύγκλισης γίνεται αισθητή όταν υπολογίζουμε το συνολικό πλήθος των πιθανών συνδυασμών με τη χρήση 137 μετρήσεων και 8 διαφορετικές θέσεις για το καθένα που είναι:

$$8^{137} \approx 5{,}288 \times 10^{123}$$

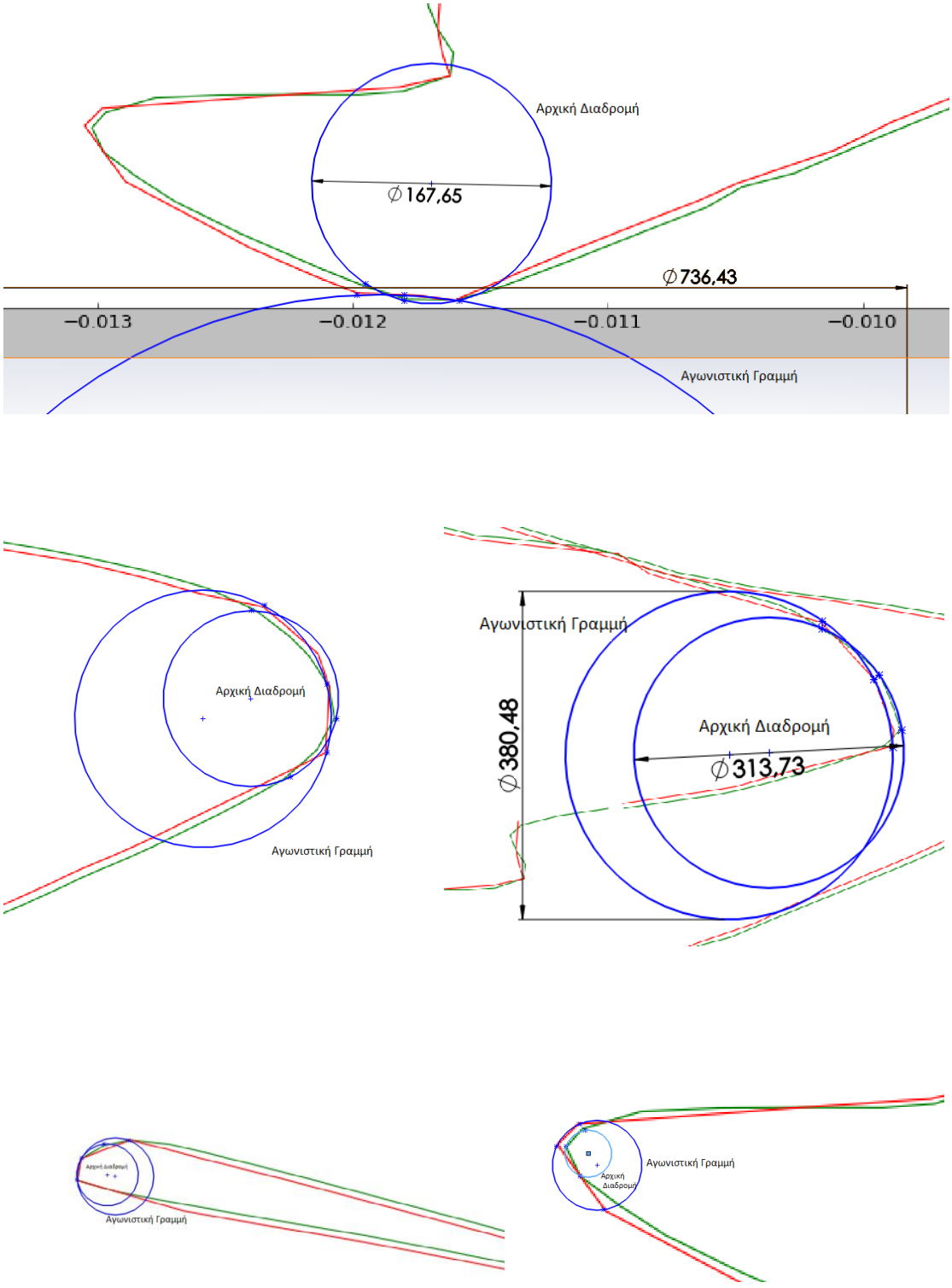
6.7 Αποτελέσματα

Η βελτιστοποίηση των παραμέτρων και η διαδικασία που περιγράφηκε οδήγησε στο παρακάτω αποτέλεσμα.

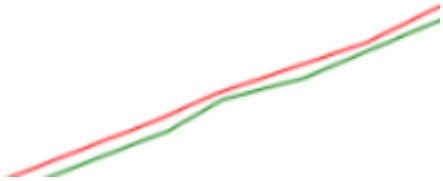


Εικόνα 6.6: Παρουσίαση αγωνιστικής γραμμής με κόκκινο χρώμα ενώ με πράσινο αναπαρίσταται η αρχική διαδρομή.

Προκειμένου να αξιολογηθούν τα αποτελέσματα του αλγορίθμου και να διαπιστωθεί εάν όντως η παραπάνω διαδρομή είναι η αγωνιστική γραμμή της πίστας αυτής ελέγχθηκαν τα σημεία ενδιαφέροντος, δηλαδή οι στροφές, και συγκρίθηκαν οι ακτίνες των εφαπτόμενων κύκλων της αρχικής διαδρομής με αυτές της αγωνιστικής γραμμής. Τα αποτελέσματα φαίνονται παρακάτω για κάθε στροφή της πίστας.



Εικόνα 6.7: Παρουσιάζονται διαφορετικές στροφές της πίστας και οι εφαπτόμενοι κύκλοι στην αγωνιστική γραμμή με κόκκινο χρώμα και την αρχική διαδρομή με πράσινο χρώμα. Οι κύκλοι αυτοί σχηματίστηκαν με τη χρήση των σημείων που καθορίζουν την έναρξη, τη μέση και το τέλος της στροφής. Γίνεται φανερό πως οι εφαπτόμενοι κύκλοι στην αγωνιστική γραμμή έχουν μεγαλύτερη ακτίνα σε σχέση με αυτούς της αρχικής διαδρομής.



Εικόνα 6.8: Παρουσιάζεται ένα τμήμα της διαδρομής στο οποίο ο αλγόριθμος παρατηρεί τη λάθος κίνηση του οδηγού να εκτρέψει το όχημα (με πράσινο χρώμα) και τη διορθώνει σε μία πιο ευθεία διαδρομή (με κόκκινο χρώμα).

Όπως προκύπτει από τα παραπάνω διαγράμματα ο αλγόριθμος είναι ικανός να εξάγει την αγωνιστική γραμμή για μία συγκεκριμένη διαδρομή. Η πολύ μεγάλη ταχύτητα σύγκλισης τον καθιστά ιδιαίτερα αποδοτικό και ικανό να χρησιμοποιηθεί από οποιοδήποτε σύστημα πλοήγησης.

6.8 Συμπεράσματα

Όπως μελετήθηκε, μέσω της βελτιστοποίησης των παραμέτρων του αλγορίθμου ο πράκτορας είναι ικανός να ανακαλύψει ποια είναι η βέλτιστη θέση του οχήματος για κάθε σημείο στη πίστα. Συνεπώς, ο αλγόριθμος που περιγράφηκε παραπάνω μπορεί να υπολογίσει με μεγάλη ταχύτητα και αποδοτικότητα την αγωνιστική γραμμή οποιασδήποτε διαδρομής.

Συνδυάζοντας τα αποτελέσματα αυτής της ενότητας με αυτά της προηγούμενης έχουμε συγκεντρώσει όλα τα απαραίτητα στοιχεία προκειμένου το όχημα να κινηθεί βέλτιστα μέσα στη πίστα και να δαπανήσει την ελάχιστη δυνατή ενέργεια ολοκληρώνοντας τη διαδρομή σε ένα προκαθορισμένο χρονικό διάστημα. Στη συνέχεια περιγράφονται τα ηλεκτρονικά μέσα που χρησιμοποιήθηκαν προκειμένου να παρουσιαστούν όλα τα δεδομένα στον οδηγό και παρουσιάζεται το διαδραστικό λογισμικό που εξηγεί τις κινήσεις που πρέπει να εκτελέσει ο οδηγός υπό τη μορφή παιχνιδιού.

Κεφάλαιο 7

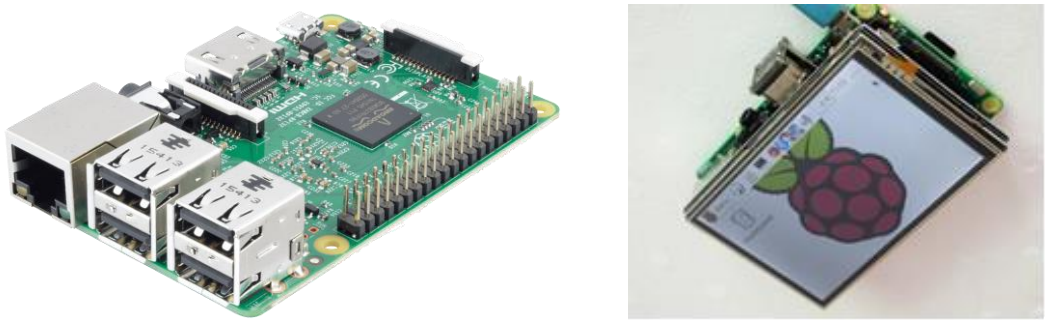
Περιγραφή χρησιμοποιηθέντος υλικού

7.1 Εισαγωγή

Στα προηγούμενα κεφάλαια περιγράφηκε αναλυτικά η υπολογιστική διαδικασία μέσω της οποίας ανακαλύπτεται ο βέλτιστος τρόπος μετακίνησης για μία συγκεκριμένη διαδρομή. Σε αυτήν την ενότητα αναλύονται τα ηλεκτρονικά στοιχεία που χρησιμοποιήθηκαν προκειμένου να λαμβάνονται σε πραγματικό χρόνο τα δεδομένα που περιγράφουν τη κίνηση του οχήματος και τις ενέργειες του οδηγού. Πιο συγκεκριμένα παρουσιάζονται τα στοιχεία που υπολογίζουν με ακρίβεια τη θέση του αυτοκινήτου πάνω στη πίστα, τη ταχύτητα του και η θέση του τιμονιού κάθε στιγμή.

7.2 Μικροϋπολογιστικό σύστημα

Όλο το κύκλωμα που παρουσιάζεται σχεδιάστηκε προκειμένου να λειτουργεί με το μικροϋπολογιστικό σύστημα raspberry και πιο συγκεκριμένα με τον μικροϋπολογιστή raspberry pi 3.



Εικόνα 7.1: Raspberry Pi 3 μαζί με την οθόνη αφής 3.5 ιντσών που χρησιμοποιήθηκε.

Επιλέχθηκε αυτός ο μικροϋπολογιστής λόγο της μεγάλης επεξεργαστικής δύναμης του και του μεγάλου πλήθους ψηφιακών εισόδων-εξόδων που διαθέτει. Επιπλέον υπάρχει ένα μεγάλο πλήθος διαφορετικών αισθητήρων που είναι συμβατοί μαζί του. Στη συνέχεια ακολουθεί ο πίνακας με τα χαρακτηριστικά του μικροϋπολογιστή.

SoC:	Broadcom BCM2837
Επεξεργαστής (CPU):	4× ARM Cortex-A53, 1.2GHz
Μονάδα επεξεργασίας γραφικών(GPU):	Broadcom VideoCore IV
Μνήμη τυχαίας προσπέλασης (RAM):	1GB LPDDR2 (900 MHz)
Υποστηριζόμενα Δίκτυα:	10/100 Ethernet,2.4GHz 802.11n wireless
Μνήμη:	microSD
Θήρες εισόδου εξόδου:	40 ψηφιακές θήρες
Λοιπές θήρες:	DMI, 3.5mm analogue audio-video jack, 4× USB 2.0, Ethernet, Camera Serial Interface (CSI), Display Serial Interface (DSI)

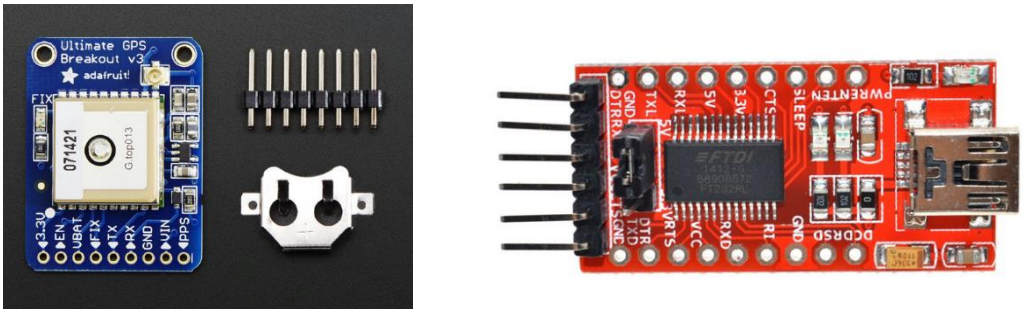
Πίνακας 7.1: Χαρακτηριστικά του Raspberry Pi 3

7.3 Χαρτογράφηση πίστας

Προκειμένου να χαρτογραφηθεί η πίστα αποφασίστηκε πως είναι απαραίτητη η χρήση δύο συστημάτων, το πρώτο θα είναι υπεύθυνο για το προσδιορισμό της θέσης του αυτοκινήτου στο μήκος της διαδρομής και το δεύτερο στο πλάτος. Με αυτό τον τρόπο γίνεται με ακρίβεια ο προσδιορισμός της θέσης του οχήματος και πραγματοποιούνται οι μετρήσεις για τις υψομετρικές διαφορές, που χρησιμοποιήθηκαν για την εκπαίδευση του νευρωνικού δικτύου, και το πλάτος της πίστας σε κάθε σημείο, που χρησιμοποιήται για τον υπολογισμό της αγωνιστικής γραμμής.

7.3.1 Σύστημα παγκόσμιας στιγματοθέτησης

Η χαρτογράφηση της πίστας και ο προσδιορισμός της θέσης του οχήματος στο μήκος της δεν χρειάζεται να είναι μεγάλης ακρίβειας. Γι’ αυτό το λόγο χρησιμοποιήθηκε ένα σύστημα παγκόσμιας στιγματοθέτησης (GPS) και πιο συγκεκριμένα το Adafruit Ultimate GPS. Αυτή η συσκευή επικοινωνεί με το raspberry μέσω ενός ftdi serial adapter και οι μετρήσεις αποθηκεύονται σε αρχεία για μελλοντική χρήση.



Εικόνα 7.2: Αριστερά παρουσιάζεται το adafruit ultimate GPS και αριστερά το ftdi serial adapter μέσω του οποίου επιτυγχάνεται η επικοινωνία με το raspberry pi.

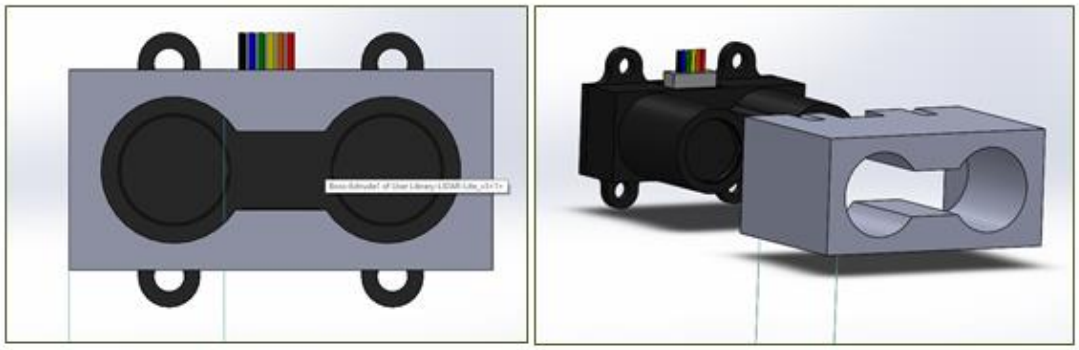
7.3.2 Αισθητήρες μέτρησης απόστασης από τα άκρα της διαδρομής

Προκειμένου να προσδιοριστεί η θέση του οχήματος στο πλάτος της πίστας εξετάστηκαν πολλά διαφορετικά αισθητήρια συστήματα θέσης. Το σύστημα πρέπει να είναι ικανό να μετράει την απόσταση του οχήματος από τα άκρα της πίστας που στο μαραθώνιο οικονομίας οριοθετούνται από μεγάλα προστατευτικά τοιχώματα.

Η έρευνα που έγινε κατέληξε στους αισθητήρες Lidar Lite v3 για τρεις σημαντικούς λόγους. Πρώτον πρόκειται για μία πολυχρησιμοποιημένη συσκευή που φημίζεται για την αξιοπιστία της και την ακρίβεια των μετρήσεων της. Δεύτερων είναι συμβατή με το πρωτόκολλο επικοινωνίας I2C κάτι που ήταν πολύ σημαντικό εφόσον έπρεπε να συνδεθούν δύο αισθητήρες στο μικροϋπολογιστικό σύστημα (ένας για να μετράει την απόσταση από το αριστερό άκρο και ο άλλος από το δεξί) και να ταυτοποιείται κάθε στιγμή πια μέτρηση αντιστοιχεί σε ποιον αισθητήρα. Και τρίτων ο ρυθμός μέτρησης και η ευαισθησία του αισθητήρα μεταβάλλονται κάτι που είναι πολύ σημαντικό προκειμένου να προσαρμοστεί η λειτουργεία του αισθητήρα συγκεκριμένα για το σύστημα που κατασκευάστηκε.

Για τους παραπάνω λόγους επιλέχθηκε ο συγκεκριμένος αισθητήρας ενώ για τη τοποθέτηση του σχεδιάστηκε και εκτυπώθηκε σε τρισδιάστατο εκτυπωτή μία προστατευτική θήκη που εφαρμόζει τον αισθητήρα στο σώμα του οχήματος.

Εικόνα 7.3: Αισθητήρας Lidar Lite v3



Εικόνα 7.4: Τα σχέδια των βάσεων των αισθητήρων

7.4 Ταχύμετρο

Προκειμένου να υλοποιηθεί ένα ταχύμετρο που θα ενημερώνει το raspberry pi για τη ταχύτητα του οχήματος χρησιμοποιήθηκε ένας μαγνητικός διακόπτης και ένας μόνιμος μαγνήτης. Ο μαγνητικός διακόπτης είναι τοποθετημένος σε ένα σταθερό σημείο κοντά στη ρόδα και ο μόνιμος μαγνήτης τοποθετήθηκε πάνω στις ακτίνες του τροχού. Ο μαγνητικός διακόπτης έκλινε το κύκλωμα κάθε φορά που ο μαγνήτης περνούσε από κοντά του. Στη συνέχεια γνωρίζοντας την ακτίνα του τροχού και το χρονικό διάστημα ανάμεσα σε δύο διαδοχικές μεταβολές του διακόπτη ο μικροϋπολογιστής υπολογίζει τη ταχύτητα του οχήματος.

Εικόνα 7.5: Το ταχύμετρο

7.5 Μέτρηση θέσης τιμονιού

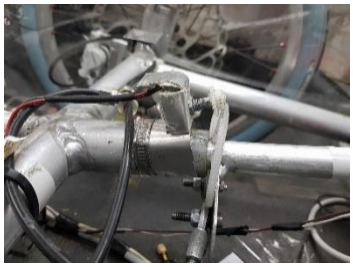
Ο στόχος του συστήματος είναι να παρέχει συμβουλές που θα καθοδηγούν τον οδηγό με το βέλτιστο τρόπο μέσα στη πίστα. Όσων αφορά την αγωνιστική γραμμή το σύστημα προκειμένου να προσφέρει διορθώσεις πρέπει να γνωρίζει τη θέση του τιμονιού κάθε στιγμή.

Ο προσδιορισμός της θέσης του τιμονιού γίνεται με τη χρήση ενός γραμμικού ποτενσιόμετρου που είναι στηριγμένο στη βάση της κολώνας του τιμονιού. Με αυτό τον τρόπο προσδιορίζεται η θέση του τιμονιού αφού η μεταβολή του επηρεάζει την αντίσταση του ποτενσιόμετρου και η αντιστοίχηση γίνεται γραμμικά.

Πιο συγκεκριμένα χρησιμοποιήθηκε ένα γραμμικό ποτενσιόμετρο μέτρησης από 0 έως 5 Kohm. Η βάση στήριξης κατασκευάστηκε από αλουμίνιο και παρουσιάζεται παρακάτω.

Τα δεδομένα όμως του ποτενσιόμετρου είναι αναλογικά και το raspberry pi έχει μόνο ψηφιακές εισόδους. Προκειμένου λοιπόν να επιτευχθεί η επικοινωνία των συσκευών χρησιμοποιήθηκε ένας μετασχηματιστής σήματος που μετατρέπει τα αναλογικά σήματα σε ψηφιακά.

Μέσο αυτής της λύσης το σύστημα πλέον μπορεί να γνωρίζει τις κινήσεις του οδηγού και να προβεί σε διορθώσεις.



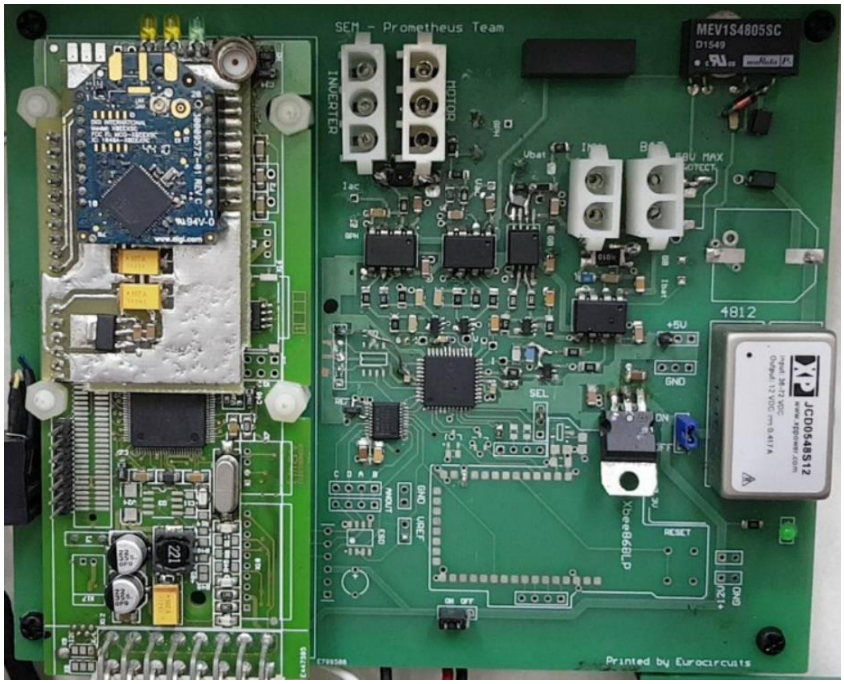
Εικόνα 7.6: Αριστερά φαίνεται το ποτενσιόμετρο που χρησιμοποιήθηκε και οι άλλες δύο εικόνες παρουσιάζουν τη βάση στήριξης του.

7.6 Τηλεμετρία οχήματος

Τα ηλεκτρονικά στοιχεία που παρουσιάστηκαν παραπάνω αφορούν την ενημέρωση του raspberry pi με όλες τις πληροφορίες που χρειάζεται προκειμένου να αξιολογήσει τις κινήσεις του οδηγού και να προβεί σε βελτιώσεις. Όσων αφορά όμως το πακέτο δεδομένων μέσω του οποίου εκπαιδεύτηκε το νευρωνικό δίκτυο του κεφαλαίου 4 οι μετρήσεις έγιναν από το σύστημα τηλεμετρίας του οχήματος [11].

Το σύστημα αυτό έχει την εποπτεία της κατανάλωσης ενέργειας και των κυριότερων μεγεθών των συνιστωσών του κινητηρίου συστήματος του πρότυπου ηλεκτρικού οχήματος. Συγκεκριμένα με συχνότητα δειγματοληψίας 50hz μετράει το εναλλασσόμενο ρεύμα που καταναλώνει ο κινητήρας, το συνεχές ρεύμα των μπαταριών, τη τάση των μπαταριών και μέσω ενός συστήματος παγκόσμιας στιγματοθέτησης υπολογίζει το γεωγραφικό μήκος και πλάτος και τέλος τη ταχύτητα του οχήματος.

Τα παραπάνω δεδομένα στέλνονται ασύρματα σε έναν ηλεκτρονικό υπολογιστή ο οποίος τα αποθηκεύει σε μορφή αρχείου για μελλοντική χρήση.



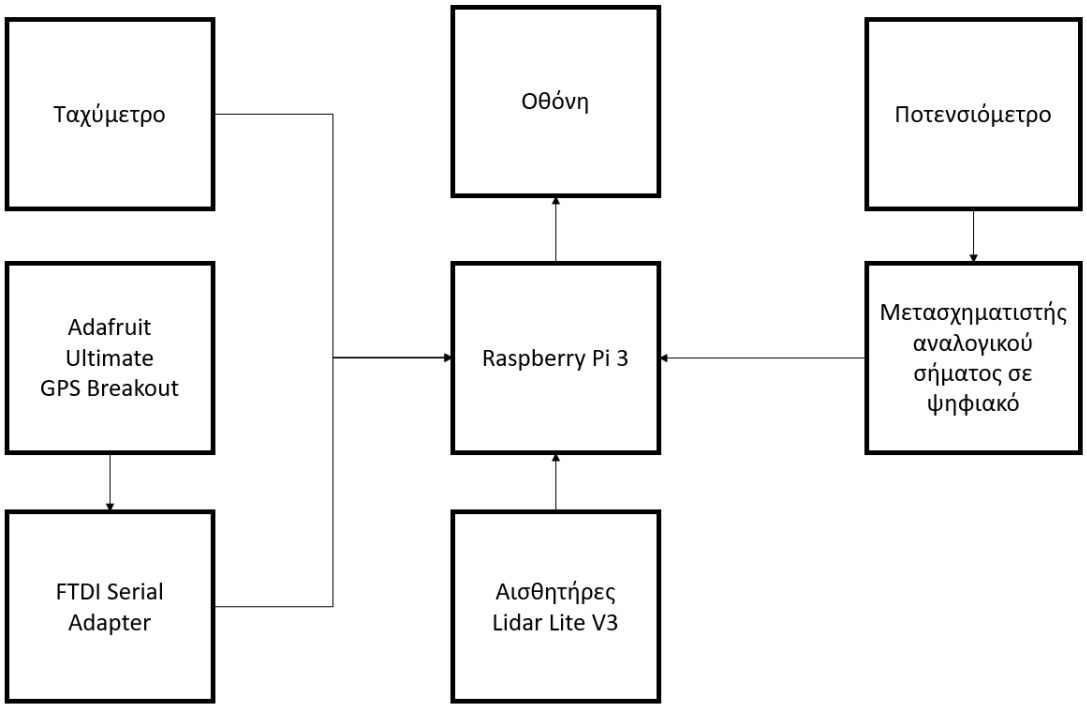
Εικόνα 7.7: Το σύστημα τηλεμετρίας του οχήματος.

7.7 Τελική συνδεσμολογία

Όλα τα παραπάνω υποκυκλώματα καταλήγουν στο τιμόνι του οχήματος που στεγάζει το μικροϋπολογιστή. Το τιμόνι σχεδιάστηκε και εκτυπώθηκε σε τρισδιάστατο υπολογιστή συγκεκριμένα για να φιλοξενήσει το παρουσιαζόμενο σύστημα.

Εικόνα 7.7: Αριστερή εικόνα παρουσιάζει τη πρόσοψη του τιμονιού και η δεξιά το πίσω μέρος και όλα τα καλώδια των υποκυκλωμάτων που εισέρχονται σε αυτό μέσω βυσμάτων.

Με τη χρήση όλων των ηλεκτρονικών στοιχείων που παρουσιάστηκαν παραπάνω το κύκλωμα είναι ικανό να συνεργαστεί με το σύστημα που παρουσιάζεται και να πραγματοποιηθούν όλες οι απαραίτητες μετρήσεις.



Εικόνα 7.7: Παρουσίαση όλης της συνδεσμολογίας του κυκλώματος.

Κεφάλαιο 8

Διαδραστικό λογισμικό

8.1 Εισαγωγή

Στα προηγούμενα κεφάλαια παρουσιάστηκε αναλυτικά η διαδικασία που ακολουθήθηκε προκειμένου να μελετηθεί ο τρόπος κίνησης του πρότυπου ηλεκτρικού οχήματος πάνω στη πίστα του μαραθωνίου οικονομίας. Όλες αυτές οι πληροφορίες πρέπει να μεταδοθούν στον οδηγό με ένα κατανοητό και απλό τρόπο προκειμένου να μην τον αποσπούν από την οδήγηση.

Προκειμένου να επιτευχθεί αυτό υλοποιήθηκε ένα διαδραστικό λογισμικό που εγκαταστάθηκε στον μικροϋπολογιστή του τιμονιού και ενημερώνει τον οδηγό για τις σωστές κινήσεις που πρέπει να κάνει με γνώμονα την εξοικονόμηση ενέργειας.

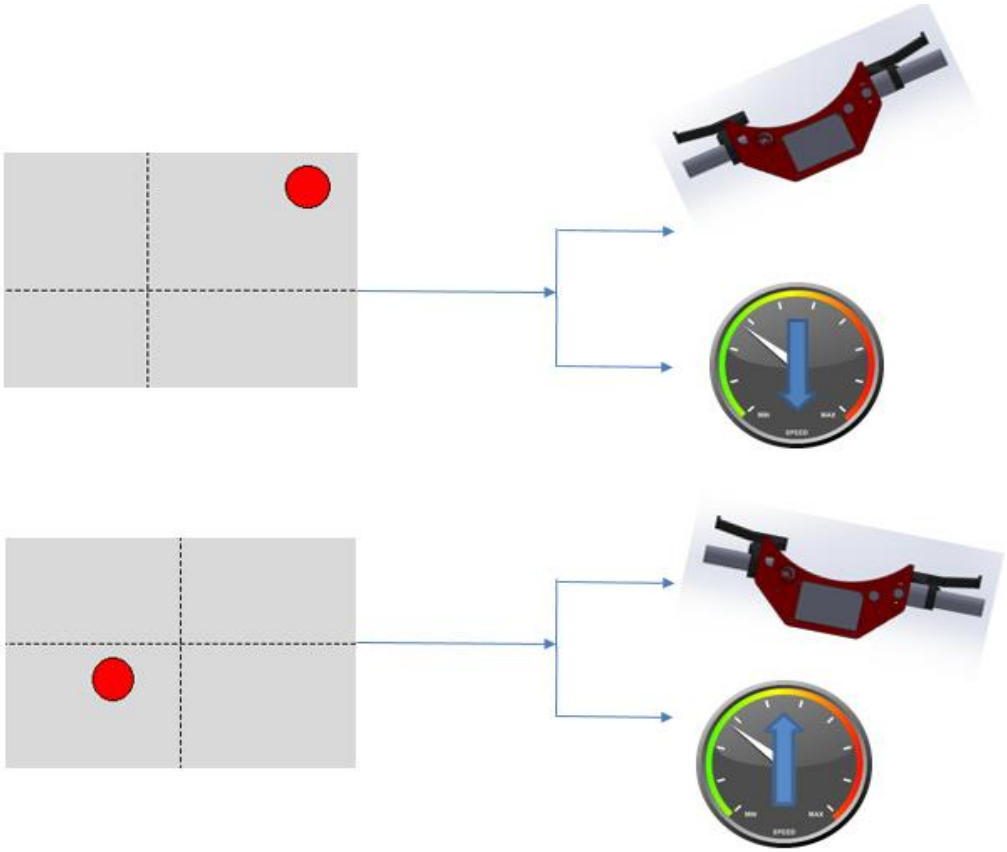
8.2 Παιχνιδοποίηση

Η κεντρική ιδέα του διαδραστικού λογισμικού που αλληλοεπιδρά με τον οδηγό βασίστηκε σε ένα απλό παιχνίδι. Στην οθόνη του μικροϋπολογιστή παρουσιάζεται μία κόκκινη σφαίρα και ο στόχος του παιχνιδιού είναι να τη διατηρήσει ο οδηγός στο κέντρο. Η σφαίρα έχει τη τάση να μετακινείτε προς τα επάνω όταν ο οδηγός διατηρεί μεγαλύτερη ταχύτητα από τη προβλεπόμενη ενώ πηγαίνει προς τα κάτω όταν διατηρεί μικρότερη ταχύτητα. Αντίστοιχα μετακινείται προς τα αριστερά όταν ο οδηγός πρέπει να στρίψει το τιμόνι δεξιά και δεξιά όταν πρέπει να στρίψει αριστερά.

Μέσο της παραπάνω διαδικασίας ο οδηγός εναλλάσσει τις ταχύτητες που διατηρεί και τις κλήσεις που εφαρμόζει στο τιμόνι προκειμένου να κερδίσει το παιχνίδι,

δηλαδή να διατηρήσει τη σφαίρα στο κέντρο και με αυτό τον τρόπο εφαρμόζει τη βέλτιστη στρατηγική μετακίνησης.

Το λογισμικό που παρουσιάζεται γράφτηκε σε Python και για την υλοποίηση των γραφικών αξιοποιήθηκε η βιβλιοθήκη Tkinter.



Εικόνα 8.1: Τρόπος λειτουργίας του διαδραστικού λογισμικού

8.3 Ενημέρωση λογισμικού από τους αισθητήρες

Στην ενότητα επτά παρουσιάστηκαν όλοι οι αισθητήρες που εγκαταστάθηκαν στο όχημα προκειμένου να επιτευχθεί η λειτουργεία του συστήματος. Όλες οι μετρήσεις καταλήγουν στο raspberry pi αλλά η κάθε μία έχει τη δική της συχνότητα εμφάνισης. Για παράδειγμα το ταχύμετρο μεταδίδει μία μέτρηση κάθε φορά που ο μόνιμος μαγνήτης, που βρίσκεται στο τροχό, περνάει κοντά από το μαγνητικό διακόπτη ενώ τα lidar έχουν συχνότητα δειγματοληψίας 50 Hz. Κρίθηκε σκόπιμο να μην αλλοιωθεί ο ρυθμός άφιξης των μετρήσεων από κανένα αισθητήρα. Γι' αυτό το λόγο σε κάθε αισθητήρα προσδόθηκε ένας μοναδικός κωδικός μεγέθους 4 bit. Με αυτό τον τρόπο όποτε το raspberry λαμβάνει μία μέτρηση ελέγχει το τετραψήφιο κωδικό και ανανεώνει την απαραίτητη μέτρηση. Ο κώδικας που είναι υπεύθυνος για το διάβασμα των αισθητήρων και την εισαγωγή του τετραψήφιου κωδικού γράφτηκε σε Bash.

8.4 Επιλογή ταχύτητας

Στην ενότητα 5 παρουσιάστηκε η μεθοδολογία εύρεσης των βέλτιστων ακολουθιών ταχυτήτων που αποδεδειγμένα μειώνουν την τελική κατανάλωση του οχήματος. Επιπλέον αναφέρθηκε πως σε περίπτωση καθυστέρησης το όχημα οφείλει να ακολουθήσει διαφορετική στρατηγική ταχυτήτων με διαφορετικό χρόνο τερματισμού προκειμένου να τερματίσει εντός των χρονικών ορίων. Στη μνήμη του μικροϋπολογιστή έχουν αποθηκευτεί πέντε ακολουθίες ταχυτήτων με διαφορετική μέση ταχύτητά η κάθε μία.

Το λογισμικό που αναπτύχθηκε μετράει το χρόνο και την απόσταση που απομένει, προκειμένου να επιλέξει την κατάλληλη ακολουθία. Όπως έχει αναφερθεί και παραπάνω, η διαδρομή χωρίζεται σε τμήματα και κάθε τμήμα έχει τη δική του βέλτιστη ταχύτητα και ένα συγκεκριμένο μήκος. Από τη μέτρηση των τμημάτων που απομένουν υπολογίζεται η απόσταση που πρέπει να καλυφθεί μέχρι το τερματισμό. Μέσω του παραπάνω υπολογισμού γίνεται γνωστή η εναπομείνασα απόσταση και μέσω του ρολογιού του μικροϋπολογιστή μετρείται ο χρόνος που απομένει. Συνεπώς μπορεί να υπολογιστεί η απαραίτητη μέση ταχύτητα που πρέπει να διατηρηθεί και να επιλεγεί η αντίστοιχη στρατηγική.

Όντας γνωστή η απαραίτητη στρατηγική ταχυτήτων απομένει να επιλεγεί η σωστή ταχύτητα που πρέπει να εφαρμοστεί σε κάθε τμήμα της διαδρομής. Αυτό γίνεται με τη βοήθεια του συστήματος παγκόσμιας στιγματοθέτησης. Δηλαδή ελέγχεται μέσω GPS η θέση του οχήματος και γίνεται η αντιστοίχιση με τη σωστή ταχύτητα της στρατηγικής.

Τέλος μέσω του ταχύμετρου ελέγχεται η ταχύτητα του οχήματος και μετακινείται η σφαίρα στο κάθετο άξονα σύμφωνα με τη ποσοστιαία διαφορά που έχει από τη βέλτιστη ταχύτητα.

Ακολουθώντας τη παραπάνω διαδικασία επιτυγχάνεται η μετακίνηση της σφαίρας στον κάθετο άξονα και η ενημέρωση του οδηγού για τη βέλτιστη ταχύτητα που οφείλει να διατηρήσει.

8.5 Επιλογή θέσης

Όπως προαναφέρθηκε η μετακίνηση της σφαίρας στον οριζόντιο άξονα αφορά τη θέση του οχήματος πάνω στη πίστα και την απόκλιση της από την αγωνιστική γραμμή. Στο κεφάλαιο 6 που αναλύθηκε η διαδικασία εύρεσης της αγωνιστικής γραμμής παρουσιάστηκε η μεθοδολογία τμηματοποίησης της πίστας σε πολλά σημεία. Το κάθε σημείο απέχει μία συγκεκριμένη απόσταση από τα άκρα της πίστας. Με κέντρο αυτά τα σημεία χωρίστηκε η πίστα σε πολλά διαφορετικά τμήματα και σε κάθε τμήμα το όχημα οφήλει να διατηρήσει μία συγκεκριμένη απόσταση από τα άκρα.

Παρομοίως με τη προηγούμενη διαδικασία ο προσδιορισμός του τμήματος που βρίσκεται κάθε φορά το όχημα γίνεται σύμφωνα με το GPS. Έπειτα πραγματοποιείται μία μέτρηση από τα δύο Lidar του οχήματος και γίνεται γνωστή η ακριβής θέση του αυτοκινήτου. Τέλος μέσω του ποτενσιόμετρου που έχει εγκατασταθεί στο τιμόνι γίνεται γνωστή η κλίση που ασκεί ο οδηγός σε αυτό.

Από τις μετρήσεις των Lidar και τη βέλτιστη θέση του οχήματος για το συγκεκριμένο τμήμα υπολογίζεται η ποσοστιαία απόκλιση από τη σωστή θέση του οχήματος. Έπειτα ελέγχεται εάν ο οδηγός προσπαθεί να διορθώσει τη θέση του οχήματος από το ποτενσιόμετρο.

Τέλος η σφαίρα μετακινείται στον οριζόντιο άξονα σύμφωνα με τη ποσοστιαία απόκλιση του αυτοκινήτου από τη βέλτιστη θέση εάν ο οδηγός δεν προσπαθεί να διορθώσει τη θέση του οχήματος. Εάν προσπαθεί να τη διορθώσει η σφαίρα πάλι κινήται προς τα δεξιά ή τα αριστερά αλλά κατά μικρότερο ποσοστό προσφέροντας μία οπτική επιβράβευση στον οδηγό για τη κίνηση που ασκεί.

Μέσο της παραπάνω διαδικασίας επιτυγχάνεται η παιχνιδοποίηση της διαδικασίας διόρθωσης των ενεργειών του οδηγού που έχουν σχέση με τη τοποθέτηση του οχήματος πάνω στη πίστα.

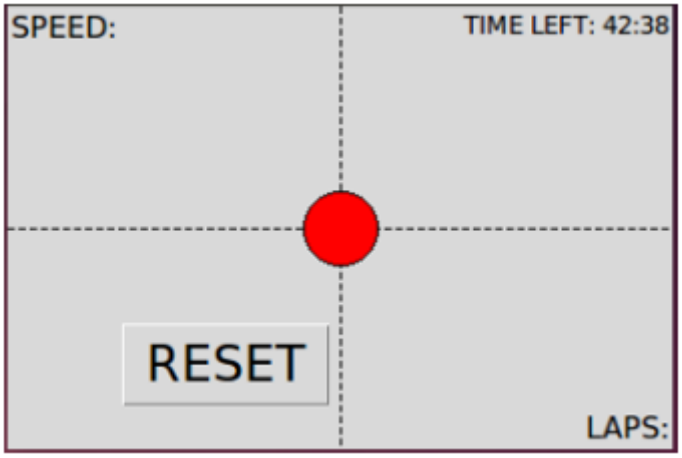
8.6 Παρουσίαση συστήματος

Ψευδοκώδικας αλγορίθμου

Είσοδοι: οι ακολουθίες ταχυτήτων (u), επιθυμητός χρόνος τερματισμού (t), μήκος κάθε τμήματος (m),
Είσοδοι: αριθμός γύρων που θα εκτελεστούν (g)
Είσοδοι: ακολουθία βέλτιστων θέσεων αγωνιστικής γραμμής (s)

Έως ο αριθμός γύρων που εκτελέστηκαν = g
    Μέτρηση απόστασης μέχρι το τερματισμό
    Μέτρηση εναπομένοντα χρόνου
    Υπολογισμός απαραίτητης μέσης ταχύτητας
    Επιλογή βέλτιστης ακολουθίας ταχυτήτων
    Προσδιορισμός θέσης από το GPS
    Προσδιορισμός βέλτιστης ταχύτητας
    Μετακίνηση σφαίρας στον κάθετο άξονα

    Προσδιορισμός βέλτιστης θέσης σύμφωνα με τη μέτρηση από το GPS και την ακολουθία m
    Διάβασμα μέτρησης ποτενσιόμετρου
    Διάβασμα μέτρησης Lidar
    Μετακίνηση σφαίρας στον οριζόντιο άξονα
    Προσδιορισμός γύρων που έχουν εκτελεστεί



Εικόνα 8.2: Το διαδραστικού λογισμικού όπως παρουσιάζεται στον οδηγό.

Το σύστημα που παρουσιάστηκε παραπάνω έχει τη δυνατότητα να προσφέρει σε αληθινό χρόνο διορθώσεις στις κινήσεις οποιουδήποτε οδηγού να και να λάβει τη θέση ενός δασκάλου αγωνιστικής και οικονομικής οδήγησης μέσω ενός διαδραστικού και ευχάριστου παιχνιδιού.

Κεφάλαιο 9

Επίλογος

Στο παρών κεφάλαιο παρουσιάζεται μία σύνοψη της εργασίας που πραγματοποιήθηκε και σχολιάζονται τα πειραματικά αποτελέσματα από τη διεξαγωγή του μαραθωνίου οικονομίας 2017 στο Λονδίνου. Εν κατακλείδι, δίνονται προτάσεις οι οποίες μπορούν να αποτελέσουν τη βάση για τη βελτίωση του παρουσιαζόμενου συστήματος.

9.1 Σύνοψη και συμπεράσματα

Το σύστημα που αναπτύχθηκε αντιμετωπίζει το ζητούμενο της οικονομικής οδήγησης από δύο πλευρές. Πρώτα σχεδιάστηκε και υλοποιήθηκε ένα ανατροφοδοτούμενο τεχνητό νευρωνικό δίκτυο που προσομοιώνει τον τρόπο κατανάλωσης του οχήματος και μέσω αυτού με μεθοδολογίες τεχνητής μάθησης υπολογίζονται οι βέλτιστες ταχύτητες που πρέπει να διατηρηθούν. Δεύτερων υπολογίζεται η αγωνιστική γραμμή μέσω της οποίος διασφαλίζετε η ασφαλής εκτέλεση των βέλτιστων ταχυτήτων.

Το ΤΝΔ είναι σε θέση να προβλέψει την αύξηση και μείωση της κατανάλωσης του οχήματος όταν αυτό επιταχύνει και επιβραδύνει σε οδοστρώματα διαφορετικών κλίσεων. Επιπλέον αποδείχτηκε πειραματικά πως όταν εφαρμόζονται οι βέλτιστες ταχύτητες του συστήματος το όχημα καταναλώνει 24.0306 % λιγότερη ενέργεια σε σχέση με τη διατήρηση σταθερής ταχύτητας. Συνεπώς η χρήση συστημάτων παρόμοιων με αυτό που πραγματοποιήθηκε μπορούν να μειώσουν το οικολογικό αποτύπωμα των οχημάτων και να αντικαταστήσουν τα διαδεδομένα συστήματα πλοήγησης που υπάρχουν στην αγορά (cruise control systems) . Η παραπάνω πρόταση ενισχύεται από τη πολύ μεγάλη ταχύτητα σύγκλισης του προγράμματος όπως παρουσιάστηκε στο κεφάλαιο 5.

Επιπλέον η μοντελοποίηση της διαδρομής οδήγησε σε επιτυχημένη σχεδίαση της αγωνιστικής γραμμής για τη συγκεκριμένη πίστα που διαθέτει αυξημένες απαιτήσεις λόγο των πολλών και έντονων στροφών. Μέσω αυτής της διαδικασίας μπορεί να γίνει ο προσδιορισμός λοιπόν της αγωνιστικής γραμμής και συστήματα όπως αυτό γίνεται να χρησιμοποιηθούν στο προσδιορισμό της διαδρομής που θα ακολουθήσει ένα αυτόνομο όχημα ή ρομπότ.

Τέλος η παρουσίαση των αποτελεσμάτων γίνεται μέσω μίας απλής διαδραστικής εφαρμογής που είναι ικανή να εκπαιδεύσει οποιοδήποτε οδηγό να οδηγεί είτε οικολογικά είτε αγωνιστικά με έναν ευχάριστο τρόπο.

Το σύστημα λοιπόν έχει την ικανότητα να εκπαιδευτεί από μόνο του, να υπολογίσει τις βέλτιστες κινήσεις μετακίνησης και να εκτελέσει το έργο του εκπαιδευτή, δηλαδή στα πλαίσια αυτής της διπλωματικής αναπτύχθηκε ένα υπολογιστικό πρόγραμμα που πρώτα εκπαιδεύει τον εαυτό του και ύστερα έναν άνθρωπο.

9.2 Μελλοντικές επεκτάσεις

Το παρών σύστημα υλοποιήθηκε και αξιολογήθηκε με τη χρήση δεδομένων που προήλθαν από την ιδιοκατασκευασμένη πλακέτα τηλεμετρίας και από τα αισθητήρια συστήματα του οχήματος Πυρφόρος. Προκειμένου λοιπόν να προβεί σε αποτελέσματα ακριβείας κρίνεται απαραίτητη η εξέταση του με τη χρήση μετρητικών πολύ μικρών σφαλμάτων.

Όσων αφορά το ΤΝΔ δίκτυο που υλοποιήθηκε προκειμένου να μπορέσει να εφαρμοστεί σε οποιαδήποτε διαδρομή κρίνεται απαραίτητη η περαιτέρω εκπαίδευση του με δεδομένα από οδοστρώματα ποικίλων κλίσεων. Αυτό συμβαίνει διότι όπως προαναφέρθηκε η εκπαίδευση έγινε από τα δεδομένα που διέθετε η ομάδα από τη συμμετοχή της στο διαγωνισμό οικονομίας 2016 και από δοκιμές που πραγματοποίησε εντός του πολυτεχνείου της Αθήνας.

Τέλος κατά τη διαδικασία ταυτοποίησης της θέσης του οχήματος πάνω στη διαδρομή χρησιμοποιήθηκαν τα δεδομένα που μετρήθηκαν από το σύστημα παγκόσμιας στιγματοθέτησης του οχήματος. Ωστόσο παρατηρήθηκε πως το σύστημα αυτό ορισμένες στιγμές αδρανούσε και έχανε το σήμα του με αποτέλεσμα να μεταδίδει λανθασμένες μετρήσεις. Συνεπώς κρίνεται σκόπιμη η αναβάθμιση του.

Κεφάλαιο 10

Βιβλιογραφία

[1] Yu-Chen Lin and Ha Ly Thi Nguyen, Development of an Eco-Cruise Control System Based on Digital Topographical Data, Feng Chia University, 12 September 2016.

[2] Τζαφέστας, Γ.Σ., Υπολογιστική νοημοσύνη, Τόμος Α: Μεθοδολογίες, Αθήνα, 1992.

[3] Κωνσταντίνος Διαμαντάρας, Τεχνητά νευρωνικά δίκτυα, Εκδώσεις Κλειδάριθμος, Αθήνα, 2007

[4] Josef Hochreiter, Untersuchungen zu dynamischen neuronalen Netzen, Technische Universit¨at M¨unchen, 1991

[5] Yoshua Bengio, Patrice Simard and Paolo Frasconi, Learning Long-Term dependencies with gradient descent is difficult, 1994

[6] Christopher Olah, Understanding LSTM Networks, 2015, http://colah.github.io/posts/2015-08-Understanding-LSTMs/

[7] Βλαχάβας Ι., Κεφάλας Π., Βασιλειάδης Ν., Ρεφάνίδης Ν., Κόκκορας Φ., Σακελαρίου Η., Τεχνητή Νοημοσύνη, Εκδόσεις Γαργατάνη, Θεσσαλονίκη, 2002

[8] Y Bengio, P Simard, and P Frasconi, ‘‘Learning long-term dependencies with gradient descent is difficult’’, IEEE Transactions on Neural Networks, , No. 5, 157–166, 1994

[9] Jiang Guo, BackPropagation Through Time, 2013

[10] Bishop, C.M., Neural networks and their applications, American Institute of Physics, Rev. Sci. Instrum., vol. 65, no 6, pp. 183-1932, 1994

[11] Σταύρος Δάβαρης, διπλωματική εργασία, Σχεδιασμός και κατασκευή διάταξης τηλεμετρίας ηλεκτρικού κινητήριου συστήματος, Αθήνα, 2016

[12] Keras: The Python Deep Learning library, https://keras.io/

[13] TensotFlow, An open-source software library for Machine Intelligence, https://www.tensorflow.org/

[14] Watkins, C. J. C. H. (1989). Learning from delayed rewards. University of Cambridge England.

[15] Hasitha Dilshani Gamage, Dr. Jinwoo (Brian) Lee Machine Learning Approach for Self-Learning Eco-Speed Control, Queensland University of Technology, Australia, 2016

[16] M. A. Abou-Nasr, Dynamic Adaptation of a Vehicle's Cruising Speed with Recurrent Neural Networks for Enhancing Fuel Economy, Intelligent Control Research Ford Motor Company Dearborn, MI. USA, 2013

[17] Savitzky A., and Golay, M.J.E. 1964, Analytical Chemistry, vol. 36, pp. 1627–1639.

[18] B. Saerens, H.A. Rakha, M. Diehl, E. Van den Bulck, A methodology for assessing eco-cruise control for passenger vehicles, Katholieke Universiteit Leuven, 2013

[19] Christos T. Krasopoulos, Minos E. Beniakar, Antonios G. Kladas, Velocity and Torque Profile Optimization of Electric Vehicle Including Limited Overload, Athens, 2016

[20] Gordon Köfner, Using Genetic Algorithms for Racing Line Optimization: Finding and Optimizing a Racing Line through a Tubular Track Using Genetic Algorithms, 2014

[21] B. Saerens, H.A. Rakha, M. Diehl, E. Van den Bulck, A methodology for assessing eco-cruise control for passenger vehicles, Virginia Tech Transportation Institute, 2013

[22] Changwoo Park, Namju Jeon, A Study of Adaptive Cruise Control System to Improve Fuel Efficiency, Republic of Korea, 2015

[23] Pıter Gαspαr, Balαzs Nımeth, Design of adaptive cruise control for road vehicles using topographic and traffic information, Hungary, 2014

[24] Lei Kang, Bozhao Qi, Dan Janecek and Suman Banerjee, EcoDrive: A Mobile Sensing and Control System for Fuel Efficient Driving, Paris,, 2015

[25] M.A.S. Kamal, Masakazu Mukai, Taketoshi Kawabe, Ecological Vehicle Control on Roads With Up-Down Slopes, IEEE Transactions on Intelligent Transportation Systems,2011

[26] Philipp Themann, Julian Bock, Lutz Eckstein, Energy efficient adaptive cruise control utilizing V2X information, Germany, 2013

[27] Tenjin-cho, Kiryu-shi Gunma, Evaluation for Energy Saving Effect and Simulation Research on Vehicle Platoon Control for City Traffic, Japan, 2016

[28] Matthew Barth, Guoyuan Wu, and Kanok Boriboonsomsin, Intelligent Transportation Systems for Improving Traffic Energy Efficiency and Reducing GHG Emissions from Roadways, University of California, Riverside, 2015

[29] Hesham A. Rakha, Kyoungho Ahn, and Sangjun Park, Predictive Eco-Cruise Control (ECC) System: Model Development, Modeling, and Potential Benefits, Virginia Tech Transportation Institute

[30] Danil V. Prokhorov, Toyota Prius HEV neurocontrol and diagnostics, Toyota Technical Center, A division of Toyota Motor Engineering and Manufacturing North America, 2008

[31] Ahmad Mozaffari, Mahyar Vajedi, Nasser L. Azad, A robust safety-oriented autonomous cruise control scheme for electric vehicles based on model predictive control and online sequential extreme learning machine with a hyper-level fault tolerance-based supervisor, University of Waterloo, 2015

[32] Ronald J. Williams, David ZIpser, A learning for continyally running fully recurrent neural networks, Northeastern University Boston, Massachusetts, 1989

[33] Xue-Bin Liang, Jun Wang, A Recurrent Neural Network for Nonlinear Optimization with a Continuously Differentiable Objective Function and Bound Constraints, IEEE TRANSACTIONS ON NEURAL NETWORKS, 2000

[34] Guang-Bin Huang, P. Saratchandran, An Efficient Sequential Learning Algorithm for Growing and Pruning RBF (GAP-RBF) Networks, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, 2004

[35] Jyh-Shing Roger Jang, Input Selction for ANSIS Learning, National Tsing Hua University Hsinchu, Taiwan

[36] Włodzisław Duch and Jerzy Korczak, Optimization and global minimization methods suitable for neural networks, Torun, 1998

[37] Herbert Jaeger, The “echo state” approach to analysing and training recurrent neural networks – with an Erratum note1, Fraunhofer Institute for Autonomous Intelligent Systems, 2010

[38] James Martens and Ilya Sutskever, Training Deep and Recurrent Networks with Hessian-Free Optimization, Department of Computer Science, University of Toronto

[39] J. Kacprzyk, Application of soft computing, Springer, 2009

[40] Omer F. Demirel, Selim Zaim, Forecasting Electricity Consumption with Neural Networks and Support Vector Regression, 8th International Strategic Management Conference, 2012

[41] James Martens Ilya Sutskever, Learning Recurrent Neural Networks with Hessian-Free Optimization, University of Toronto, Canada, 2011

[42] Dongsung Huh, and Emanuel Todorov, Real-Time Motor Control using Recurrent Neural Networks, In IEEE, 2009