



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής



Εντοπισμός επιπέδου και διάρκειας επικίνδυνης συμπεριφοράς του οδηγού με τεχνικές μηχανικής εκμάθησης

Θοδωρής Γαρεφαλάκης

Επιβλέπων: Γιώργος Γιαννής, Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2022

Στόχος Διπλωματικής Εργασίας

Εντοπισμός του επιπέδου και της διάρκειας επικίνδυνης συμπεριφοράς του οδηγού με τεχνικές μηχανικής εκμάθησης.

Για την επίτευξη του στόχου πραγματοποιείται:

- Ανάπτυξη αλγορίθμων ταξινόμησης για την **αναγνώριση** του επιπέδου ασφαλείας που βρίσκεται κάθε οδηγός
- Ανάπτυξη αλγορίθμων παλινδρόμησης για την **πρόβλεψη** της διάρκειας οδήγησης σε κάθε επίπεδο ασφαλείας



Μεθοδολογία Διπλωματικής Εργασίας



Βιβλιογραφική ανασκόπηση (1/2)



Βιβλιογραφική ανασκόπηση (2/2)

Ελλείψεις και **περιορισμοί** προγενέστερων ερευνών

- **Απουσία** κατάστασης οδηγού (ψυχολογική ή κόπωση), κυκλοφοριακών και καιρικών συνθηκών
- **Έλλειψη** στοιχείων σχετικά με τον τύπο, την γεωμετρία της οδού και των χαρακτηριστικών του οχήματος
- **Απουσία** στάσης, αντίληψης και δημογραφικών χαρακτηριστικών των οδηγών
- Εξέταση **περιορισμένου** αριθμού τεχνικών αντιμετώπισης της άνισης κατανομής των δειγμάτων στα επίπεδα επικινδυνότητας

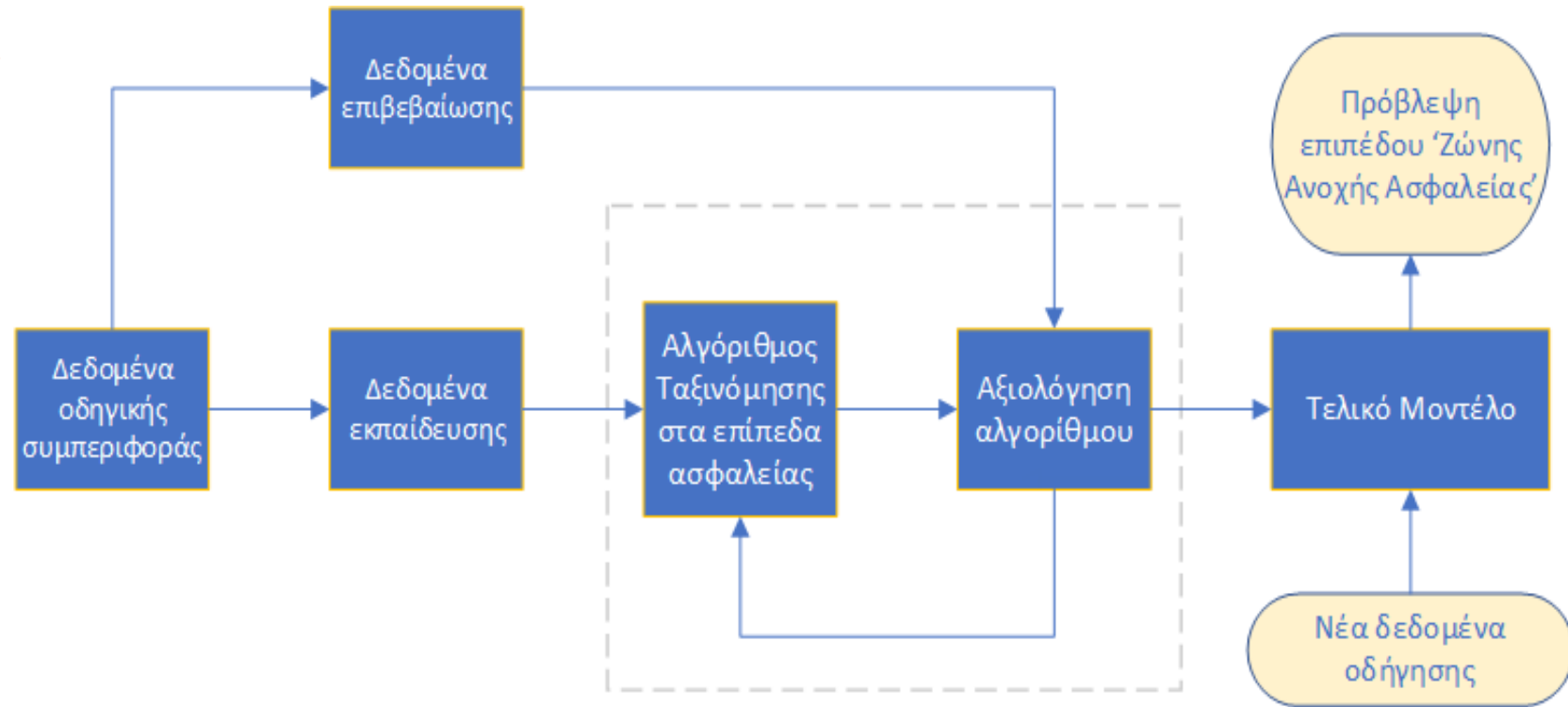


Θεωρητικό υπόβαθρο (1/2)

Ανάπτυξη μοντέλων **μηχανικής εκμάθησης**



- **Ταξινόμησης:** Για την πρόβλεψη του επιπέδου 'Ζώνης Ανοχής Ασφαλείας' που βρίσκεται ο οδηγός
- **Παλινδρόμησης:** Για την πρόβλεψη της διάρκειας σε κάθε επίπεδο ασφαλείας



Για την αξιολόγηση της **επιρροής** των μεταβλητών εξετάζονται:

- **Σημαντικότητα** στην διαδικασία της ταξινόμησης
- **Συντελεστές** στην διαδικασία της παλινδρόμησης

Θεωρητικό υπόβαθρο (2/2)

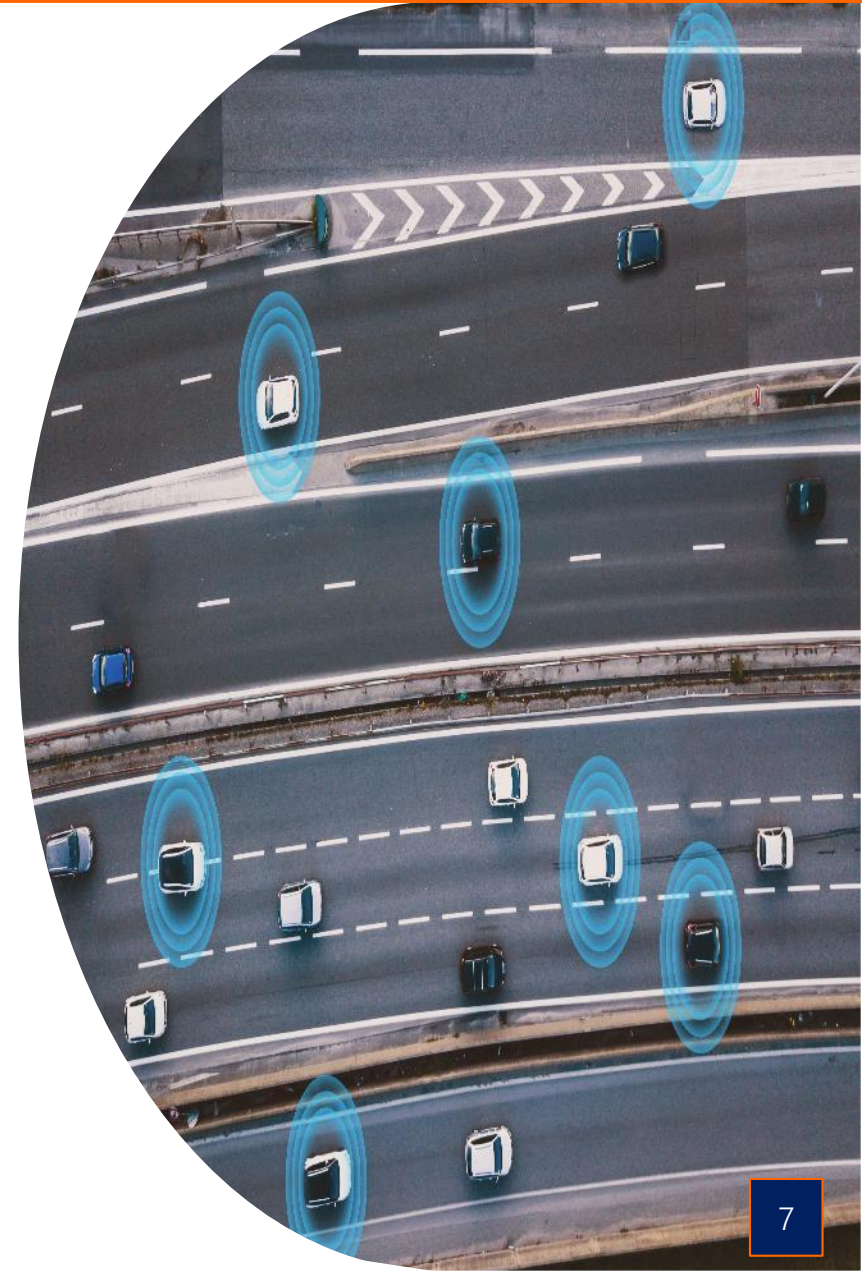
Αξιολόγηση μοντέλων μηχανικής εκμάθησης

➤ Μετρικές αξιολόγησης (ταξινόμηση)

- ⑩ Ορθότητα (Accuracy)
- ⑩ Ακρίβεια (Precision)
- ⑩ Ανάκληση (Recall)
- ⑩ f1-score
- ⑩ False Alarm Rate

➤ Κριτήρια αποδοχής μοντέλων (παλινδρόμηση)

- ⑩ R^2
- ⑩ Έλεγχος συσχέτισης
- ⑩ Έλεγχος p-value
- ⑩ Έλεγχος t-value
- ⑩ Λογική ερμηνεία συντελεστών



Συλλογή Στοιχείων

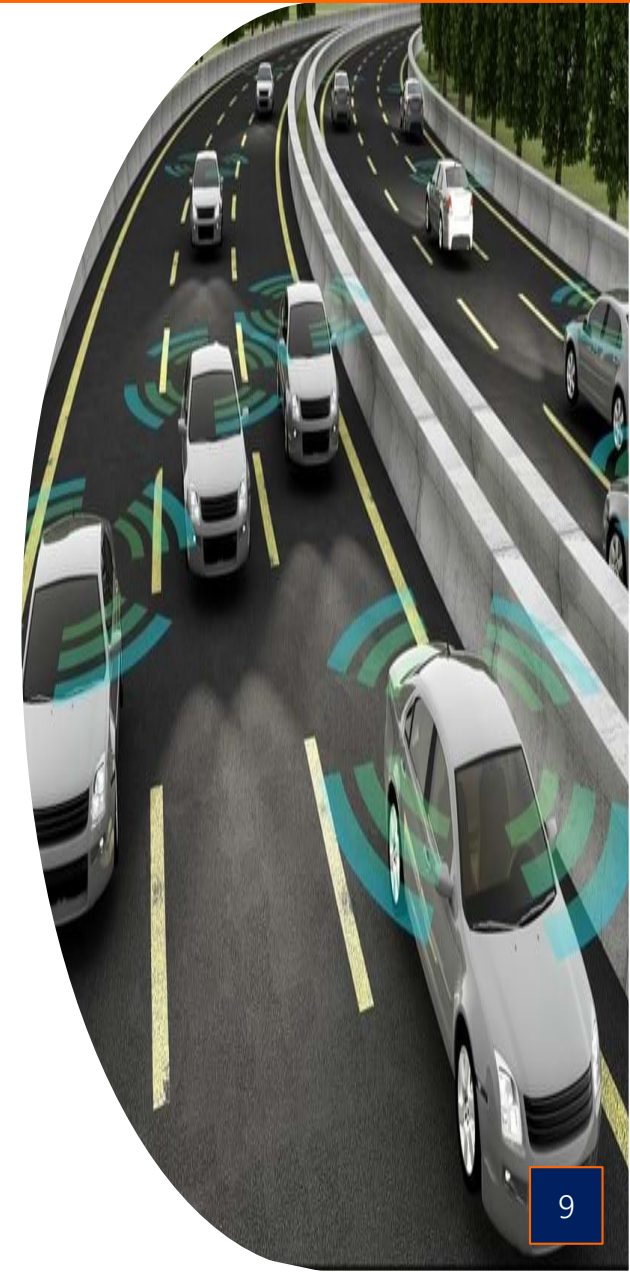
- Στο πλαίσιο του ερευνητικού έργου i-DREAMS πραγματοποιήθηκε πείραμα **προσομοιωτή οδήγησης**
- **36 οδηγοί** συμμετείχαν στο πείραμα από 7/12/2020 έως 17/01/2021
- Εφαρμόστηκαν **3 σενάρια οδήγησης** με διαφορετικά χαρακτηριστικά σε κάθε τμήμα της οδού (διαφορετικός αριθμός λωρίδων και ορίων ταχύτητας)
- Κάθε οδηγός πραγματοποίησε **3 ξεχωριστές διαδρομές** (χωρίς παρεμβάσεις, με παρεμβάσεις και παρεμβάσεις σε μεταβαλλόμενες συνθήκες)
- Συλλέχθηκαν σημαντικά **χαρακτηριστικά οδηγικής συμπεριφοράς**



Περιγραφή - Επεξεργασία Στοιχείων

Μεταβλητές	Επεξήγηση	Δείκτης για κάθε χρονικό βήμα 30 δλ.
TTC	Χρόνος πρόσκρουσης	Μέσος όρος, τυπική απόκλιση, ελάχιστη τιμή, μέγιστη τιμή
Headway	Χρονική απόσταση οχημάτων	Μέσος όρος, τυπική απόκλιση, ελάχιστη τιμή, μέγιστη τιμή, διάμεση τιμή
Speed	Ταχύτητα οχήματος	Μέσος όρος, τυπική απόκλιση, ελάχιστη τιμή, μέγιστη τιμή
Distance_travelled	Απόσταση που διανύθηκε	Άθροισμα
BSAV_SpeedLimitMS	Όριο ταχύτητας	Μέγιστη τιμή
BSAV_SpeedLimitKPH	Όριο ταχύτητας	Μέγιστη τιμή
HandsOnEvent	Ένδειξη ότι τα χέρια του οδηγού βρίσκονται στο τιμόνι	Μέσος όρος, διάμεση τιμή
FatigueEvent	Δείκτης κόυρασης	Διάμεση τιμή

Υπολογισμός περιγραφικών στατιστικών στοιχείων των μεταβλητών για κάθε χρονικό διάστημα των 30 δλ.



Καθορισμός επιπέδων ασφαλείας

Καθορισμός **επιπέδου 'Ζώνης Ανοχής Ασφαλείας'** βάσει ορίων της χρονικής απόστασης μεταξύ δύο οχημάτων

- Επίπεδο 'Normal': $\text{Headway}_{\min} > 2 \text{ δλ.}$
- Επίπεδο 'Dangerous': $1.4 \text{ δλ.} < \text{Headway}_{\min} < 2 \text{ δλ.}$
- Επίπεδο 'Avoidable Accident': $\text{Headway}_{\min} < 1.4 \text{ δλ.}$

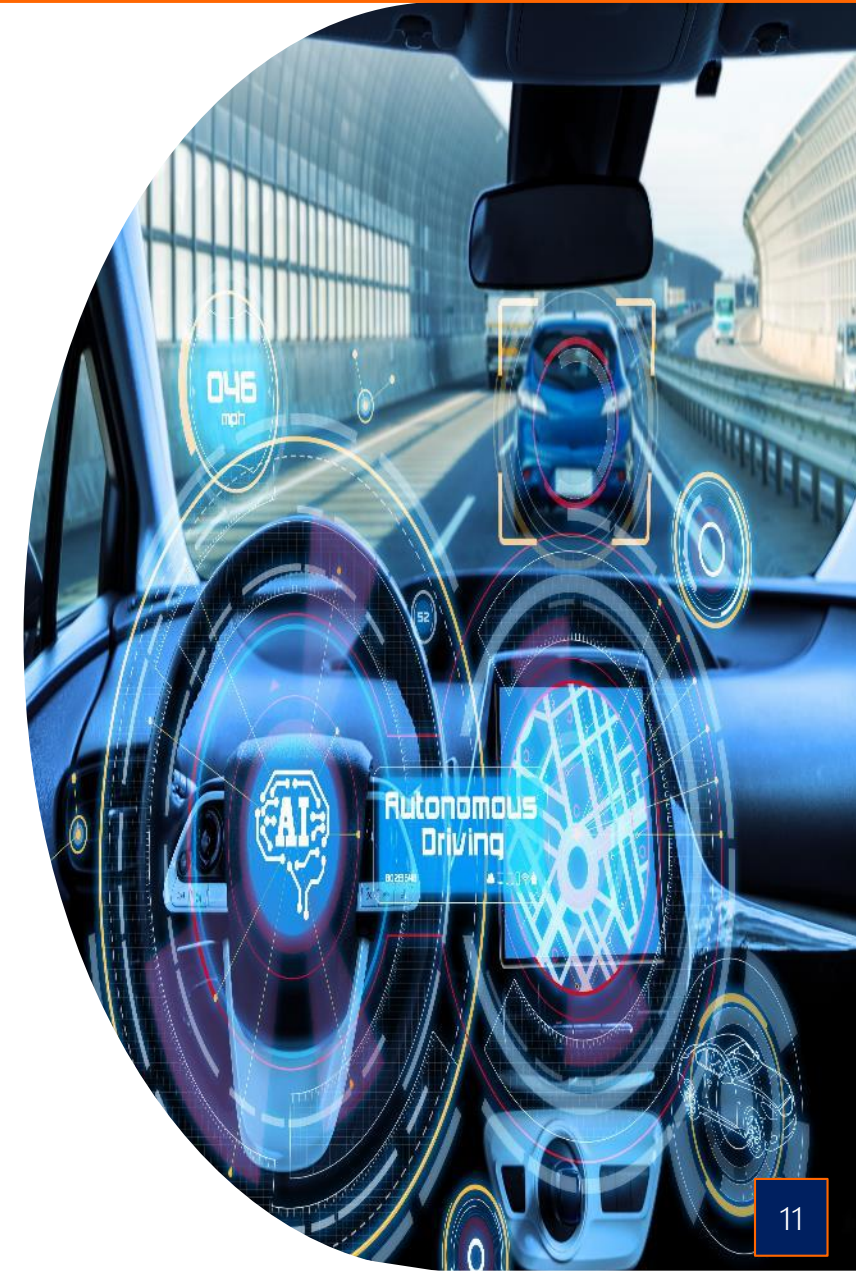


Για κάθε οδηγό υπολογίστηκε **η διάρκεια οδήγησης** σε κάθε επίπεδο της 'Ζώνης Ανοχής Ασφαλείας'

Οι μεταβλητές *Headway* και *TTC* εξαιρούνται από την διαδικασία της ταξινόμησης. Κίνδυνος **μεροληψίας** στα μοντέλα (*bias*)

Διαδικασία ταξινόμησης (1/5)

- **Επιλογή μεταβλητών** εισόδου στα μοντέλα ταξινόμησης με βάση την **συσχέτιση** και την **σημαντικότητα** των μεταβλητών
- Διαχωρισμός δεδομένων σε **δεδομένα εκπαίδευσης** και **δεδομένα επιβεβαίωσης** σε ποσοστό 90% και 10% αντίστοιχα
- **Επαναδειγματοληψία** δεδομένων εκπαίδευσης με χρήση κατάλληλου αλγορίθμου
- Ανάπτυξη **τεσσάρων αλγορίθμων ταξινόμησης**: (1) Μηχανές Διανυσμάτων Υποστήριξης (SVM), (2) Τυχαία Δάση (RF), (3) AdaBoost και (4) Πολυεπίπεδο Perceptron (MLP)



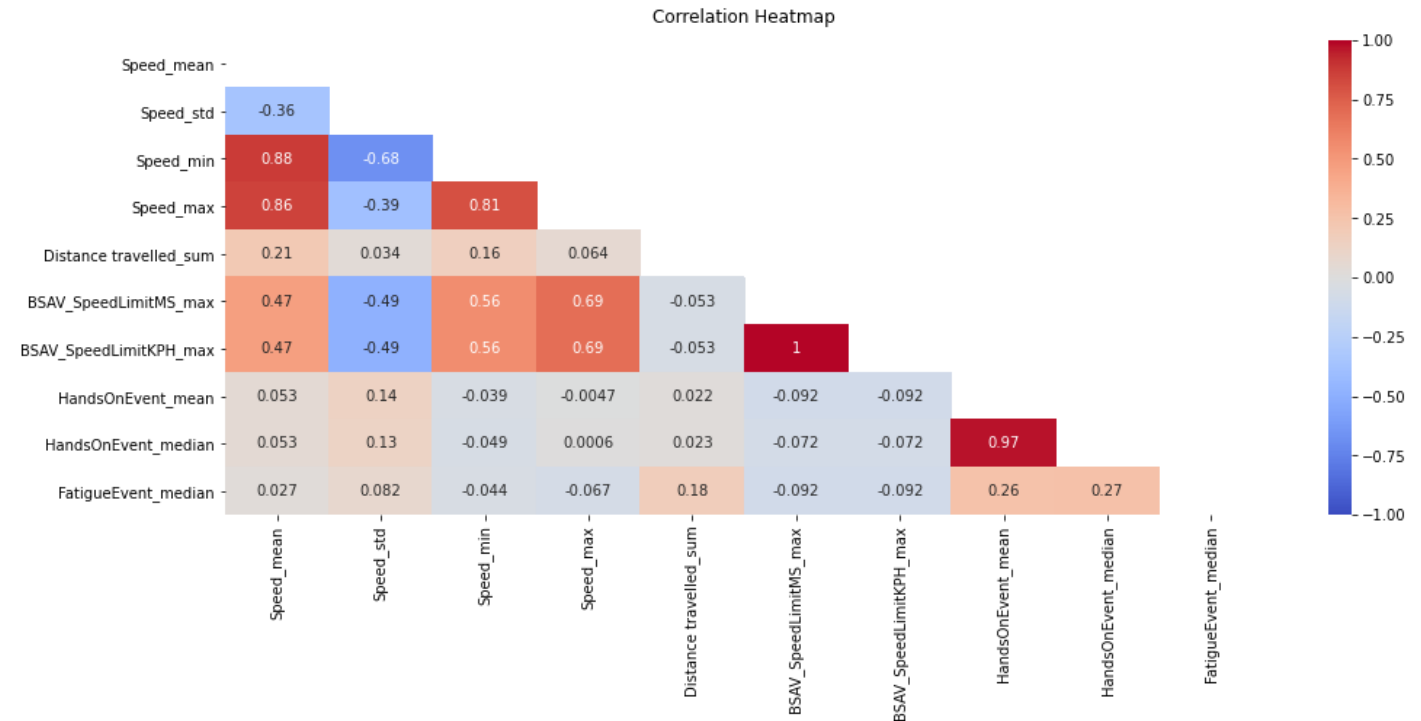
Διαδικασία ταξινόμησης (2/5)

(1) Επιλογή μεταβλητών - Συσχέτιση

➤ **Υψηλή συσχέτιση** μεταξύ διαφορετικών περιγραφικών στατιστικών της ίδιας μεταβλητής

➤ **Μέτρια συσχέτιση** μεταξύ ταχύτητας και ορίου ταχύτητας

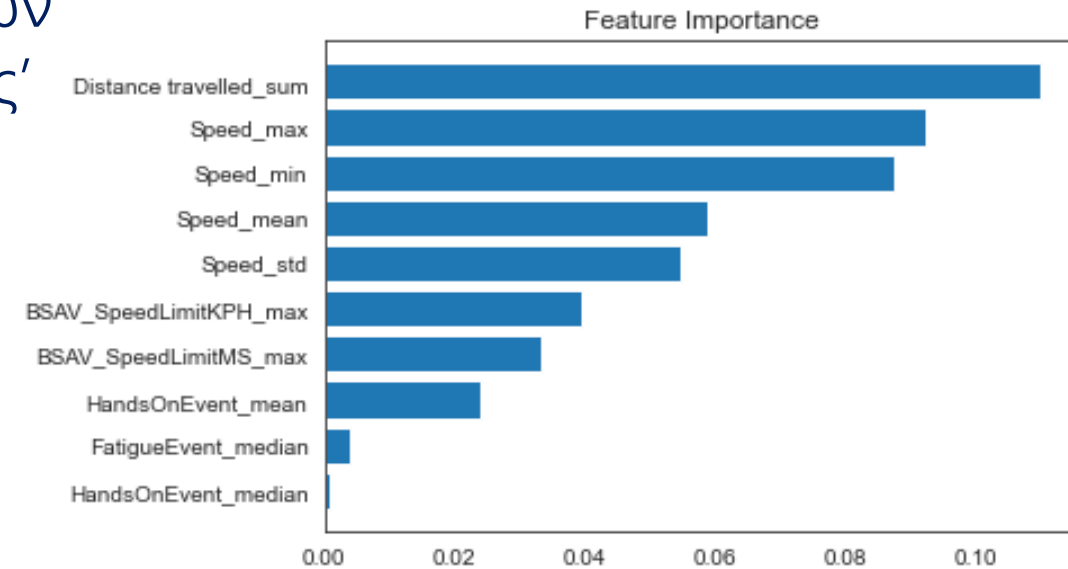
➤ **Χαμηλή συσχέτιση** μεταξύ μεταβλητών απώλειας προσοχής, κούρασης, διανυθείσας απόστασης και όλων των υπολοίπων μεταβλητών



Διαδικασία ταξινόμησης (3/5)

(2) Επιλογή μεταβλητών - Σημαντικότητα

- Μέσω του αλγορίθμου **Τυχαίων Δασών** αξιολογήθηκε η σημαντικότητα των μεταβλητών στην ταξινόμηση της 'Ζώνης Ανοχής Ασφαλείας'
- Η διανυθείσα απόσταση και η ταχύτητα ήταν **οι πιο σημαντικοί παράγοντες** από τους εξεταζόμενους
- Η μεταβλητή απόσπασης προσοχής, κούραση και το όριο ταχύτητας ήταν οι **λιγότερο σημαντικοί παράγοντες**



Μεταβλητές εισόδου:

- (1) Διανυθείσα απόσταση, (2) Μέγιστη ταχύτητα, (3) Όριο ταχύτητας

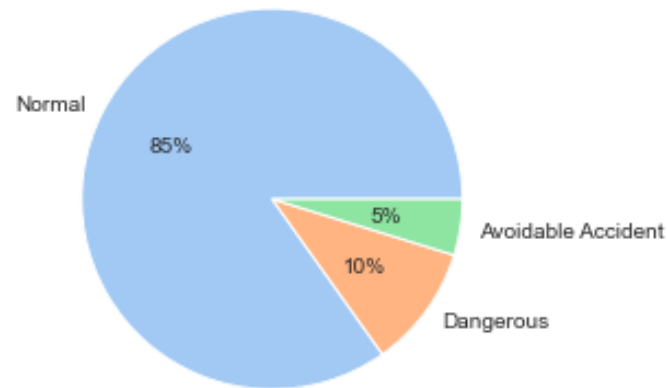
Διαδικασία ταξινόμησης (4/5)

Με τον καθορισμό των επιπέδων προέκυψε **ανισότητα** των δειγμάτων στα τρία επίπεδα

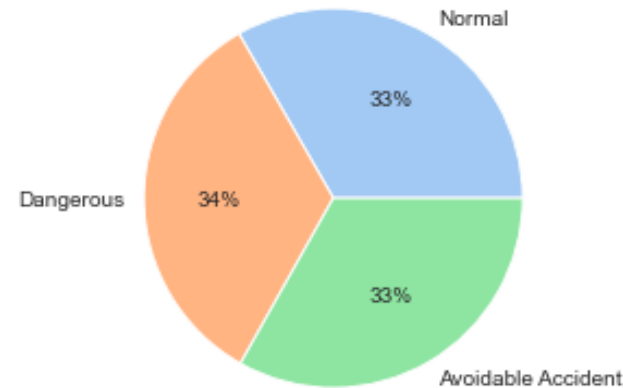
Εφαρμόστηκε η 'Προσαρμοστική Συνθετική' τεχνική (ADASYN) με σκοπό την:

➤ Αντιμετώπιση **μεροληψίας** μοντέλων ως προς την κυρίαρχη τάξη

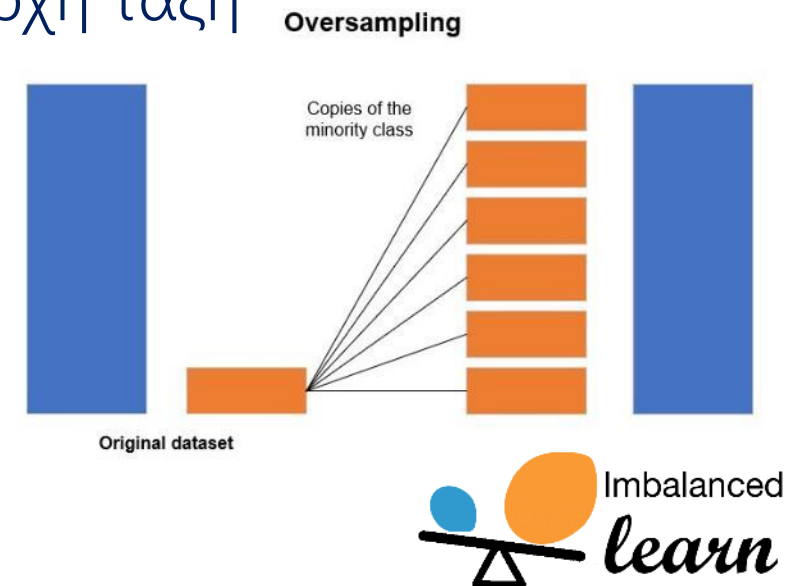
➤ **Αντιστάθμιση** της άνισης κατανομής των τάξεων



Πριν την επαναδιαμόρφωση

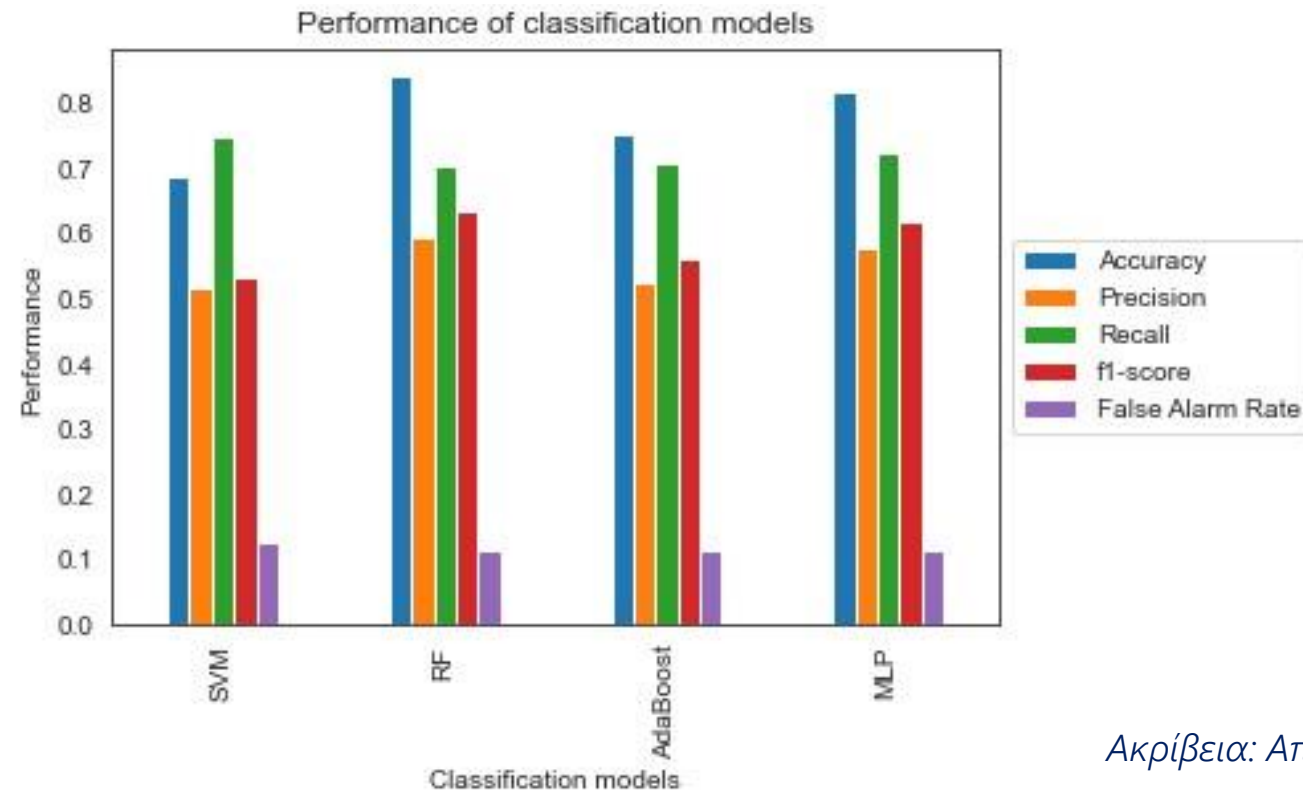


Μετά την επαναδιαμόρφωση



Διαδικασία ταξινόμησης (5/5)

- Για την σύγκριση των μοντέλων ταξινόμησης αξιοποιήθηκαν ορισμένες **μετρικές αξιολόγησης**



	Ορθότητα	Ακρίβεια	Ανάκληση	f1-score	False Alarm Rate
SVM	69%	51%	75%	53%	13%
RF	84%	59%	70%	63%	12%
AdaBoost	75%	52%	71%	56%	11%
MLP	81%	58%	72%	62%	11%

Ακρίβεια: Από τις προβλέψεις που πραγματοποιήθηκαν πόσες ήταν **ορθές**

Ανάκληση: Από τα πραγματικά στοιχεία **πόσα** κατάφερε να προβλέψει το μοντέλο

f1-score: **Αρμονικό** μέτρο της ακρίβειας και της ανάκλησης

Διαδικασία παλινδρόμησης (1/2)

- Η **διάρκεια οδήγησης** σε κάθε επίπεδο ασφαλείας αποτέλεσε την **εξαρτημένη μεταβλητή**
- Οι **ανεξάρτητες μεταβλητές** επιλέχθηκαν βάσει την μεταξύ τους **συσχέτιση** και την **στατιστική σημαντικότητα** τους
- Διαχωρισμός δεδομένων σε **δεδομένα εκπαίδευσης** και **δεδομένα επιβεβαίωσης** σε ποσοστό 85% και 15% αντίστοιχα
- Ανάπτυξη **τριών αλγορίθμων παλινδρόμησης**: (1) Παλινδρόμηση Κορυφογραμμής (RR), (2) Παλινδρόμηση Lasso (LR) και (3) Παλινδρόμηση Elastic Net (ENR)



Διαδικασία παλινδρόμησης (2/2)

- Η **διανυθείσα απόσταση** και η **μέγιστη ταχύτητα** αποτέλεσαν τις ανεξάρτητες μεταβλητές του μοντέλου

Σύνοψη μοντέλου παλινδρόμησης RR				
	Συντελεστές	Τυπική απόκλιση	t value	p value
Σταθερός όρος	9966,716	472,905	21,076	0,000
Speed_max	-112,009	2,178	-51,441	0,000
Distance travelled_sum	0,001	0,001	8,896	0,000
$R^2 = 0,8493$		Adjusted $R^2 = 0,8458$		

Σύνοψη μοντέλου παλινδρόμησης LR				
	Συντελεστές	Τυπική απόκλιση	t value	p value
Σταθερός όρος	9967,358	472,905	21,077	0,000
Speed_max	-112,017	2,177	-51,445	0,000
Distance travelled_sum	0,001	0,001	8,896	0,000
$R^2 = 0,8493$		Adjusted $R^2 = 0,8458$		

Σύνοψη μοντέλου παλινδρόμησης ENR				
	Συντελεστές	Τυπική απόκλιση	t value	p value
Σταθερός όρος	9697,044	472,981	20,459	0,000
Speed_max	-108,840	2,182	-49,873	0,000
Distance travelled_sum	0,001	0,001	8,955	0,000
$R^2 = 0,8486$		Adjusted $R^2 = 0,8451$		

Η μέγιστη ταχύτητα επηρεάζει **αρνητικά** την διάρκεια που βρίσκεται ο οδηγός σε κάθε επίπεδο ασφαλείας

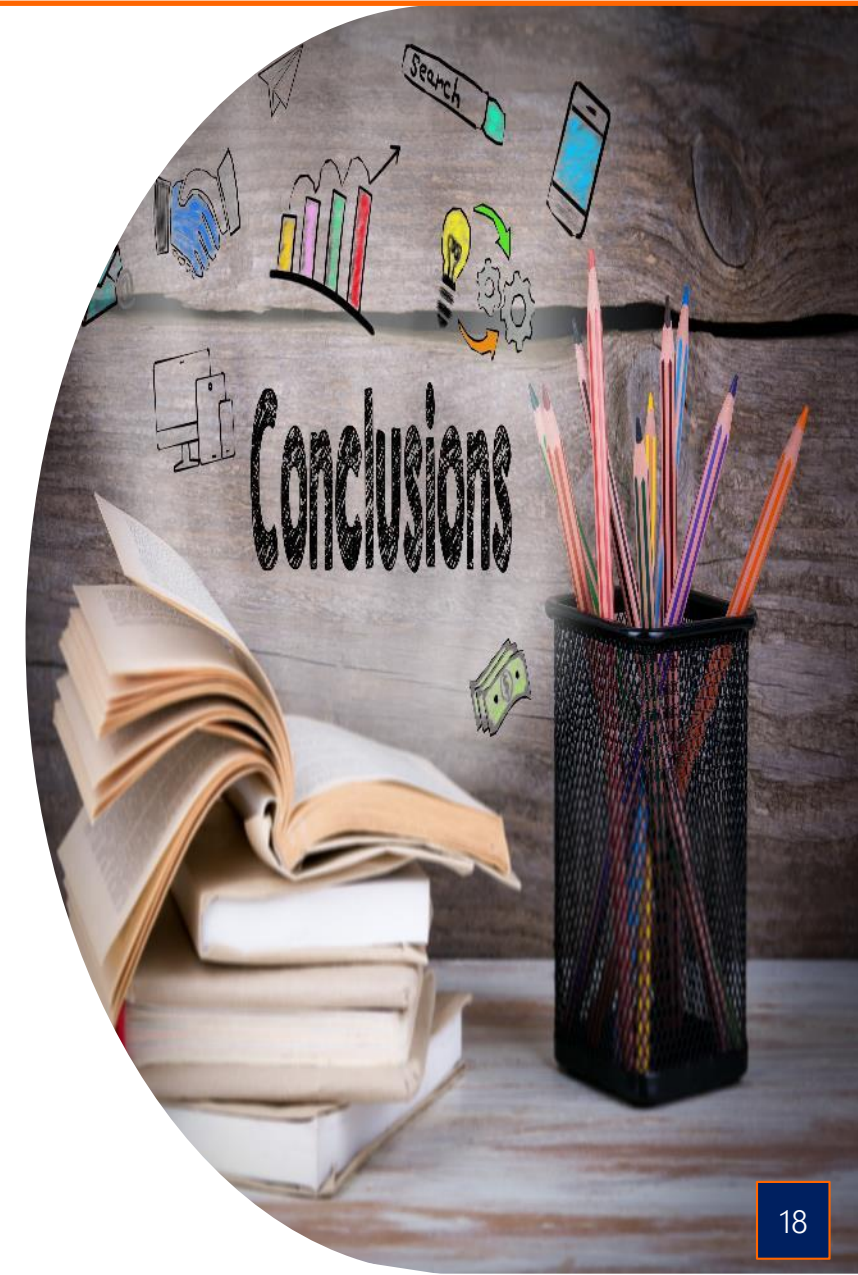
Οι ανεξάρτητες μεταβλητές είναι στατιστικά σημαντικές

$p\text{-value} < 0.05$

$|t\text{-value}| > t\text{-test}$

Συμπεράσματα

- Ο καθορισμός των τριών επιπέδων με βάση τα όρια της χρονικής απόστασης έδωσε τα **βέλτιστα αποτελέσματα**, συναφή με την διεθνή βιβλιογραφία
- Η **διανυθείσα απόσταση οδήγησης, η ταχύτητα** και **το όριο ταχύτητας** κρίνονται σημαντικές μεταβλητές για την αναγνώριση της οδηγικής συμπεριφοράς
- Τα μοντέλα **‘Τυχαίων Δασών’** και **‘Πολυεπίπεδου Perceptron’** σημείωσαν υψηλή επίδοση στην αναγνώριση της οδηγικής συμπεριφοράς.
- Η **‘Προσαρμοστική Συνθετική’ τεχνική (ADASYN)** είχε τα βέλτιστα αποτελέσματα στην διαδικασία αντιμετώπισης της άνισης κατανομής στις τάξεις
- Η **μέγιστη ταχύτητα** επηρεάζει αρνητικά την διάρκεια οδήγησης σε κάθε επίπεδο **‘Ζώνης Ανοχής Ασφαλείας’**
- Στο σύνολο τους τα τρία μοντέλα παλινδρόμησης σημείωσαν **υψηλή προγνωστική ικανότητα**



Προτάσεις για αξιοποίηση των αποτελεσμάτων

- Αξιοποίηση των **μοντέλων ταξινόμησης** για την αναγνώριση του επιπέδου ασφαλείας των οδηγών σε πραγματικές συνθήκες οδήγησης
- Ανάπτυξη κατάλληλου **συστήματος** εντός του οχήματος για την **αναγνώριση** του επιπέδου ασφαλείας που βρίσκεται ο οδηγός σε πραγματικό χρόνο
- Δημιουργία **εφαρμογής σε έξυπνα κινητά τηλέφωνα** για την αξιολόγηση της οδηγικής συμπεριφοράς
- Συντονισμός **συχνοτήτων** και **έντασης** προειδοποιήσεων κατά την διάρκεια της διαδρομής



Προτάσεις για περαιτέρω έρευνα

- Αξιοποίηση **μεγαλύτερου όγκου δεδομένων**
- Ανάπτυξη **εναλλακτικών** μοντέλων μηχανικής εκμάθησης και μοντέλων **βαθιάς εκμάθησης**
- Εξέταση **διαφορετικών μεθόδων** για τον καθορισμό του επιπέδου 'Ζώνης Ανοχής Ασφαλείας'
- Ανάπτυξη **μοντέλων ακολουθίας** για την πρόβλεψη του επιπέδου ασφαλείας στο επόμενο χρονικό βήμα
- Διερεύνηση **πρόσθετων** παραγόντων





Ευχαριστώ για την προσοχή σας!



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής



Εντοπισμός επιπέδου και διάρκειας επικίνδυνης συμπεριφοράς του οδηγού με τεχνικές μηχανικής εκμάθησης

Θοδωρής Γαρεφαλάκης

Επιβλέπων: Γιώργος Γιαννής, Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2022