



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής



Εντοπισμός επικίνδυνης συμπεριφοράς με τεχνικές μηχανικής μάθησης

Έκτορας Καμβουσιώρας



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

Αθήνα, Ιούλιος 2022



Αντικείμενο - Στόχος



Αντικείμενο - Στόχος

Διερεύνηση της **επιρροής** των διαφορετικών παραγόντων οδήγησης στην επικίνδυνη οδηγική συμπεριφορά.

Η ανάλυση πραγματοποιήθηκε με βάση δύο προσεγγίσεις:

- Μέσω της ανάπτυξης αλγορίθμων ταξινόμησης για την **αναγνώριση** του επιπέδου ασφαλείας που βρίσκεται κάθε οδηγός.
- Μέσω της εύρεσης των τιμών **Shap** για την εξήγηση λειτουργίας των μοντέλων μηχανικής μάθησης





Μεθοδολογία Διπλωματικής Εργασίας



Μεθοδολογία Διπλωματικής Εργασίας





Βιβλιογραφική Ανασκόπηση



Βιβλιογραφική ανασκόπηση

Συναφείς έρευνες (Ανάλυση οδηγικής συμπεριφοράς με τεχνικές μηχανικής εκμάθησης)

- ☐ **Πρόβλεψη** σύγκρουσης σε πραγματικό χρόνο
- ☐ **Καθορισμός** και **αξιολόγηση** επιπέδου ασφαλείας
- ☐ **Πρόβλεψη** επιπέδου επικινδυνότητας
- ☐ **Καλύτερη κατανόηση** μοντέλων μέσω τιμών Shar

Από τις τεχνικές που αναλύθηκαν:

- ☐ Τα μοντέλα 'Random Forests', 'XGBoost' σημείωσαν **υψηλά** ποσοστά ορθών προβλέψεων
- ☐ Η τεχνική SMOTE παρουσίασε **σημαντική ικανότητα** αντιμετώπισης του φαινομένου άνισης κατανομής των δεδομένων στις τάξεις
- ☐ Οι τιμές Shar βρίσκουν και εξηγούν τις πιο **σημαντικές μεταβλητές** κάθε μοντέλου





Βιβλιογραφική ανασκόπηση

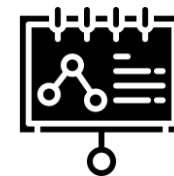
Ελλείψεις και περιορισμοί προγενέστερων ερευνών

- ☐ **Απουσία** κατάστασης οδηγού (ψυχολογική ή κόπωση)
- ☐ **Απουσία** κυκλοφοριακών και καιρικών συνθηκών
- ☐ Τα αποτελέσματα δεν προσφέρουν **σαφή απάντηση** ως προς την επιρροή των μεταβλητών και μπορεί να μην σχετίζονται και με την συμπεριφορά του οδηγού
- ☐ **Έλλειψη** στοιχείων σχετικά με τον τύπο, την γεωμετρία της οδού και των χαρακτηριστικών του οχήματος





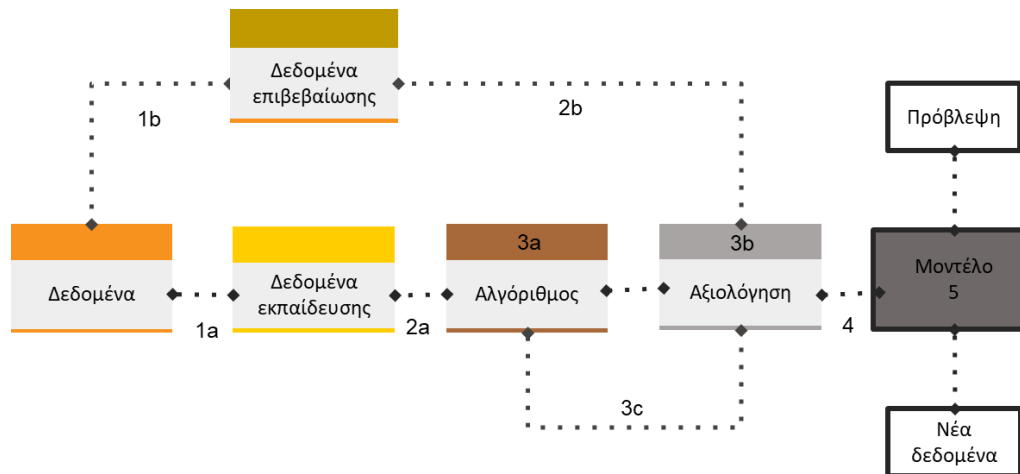
Θεωρητικό υπόβαθρο



Θεωρητικό υπόβαθρο

Ανάπτυξη μοντέλων μηχανικής
εκμάθησης

❑ Ταξινόμηση



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

❑ Σημαντικότητα στην διαδικασία της ταξινόμησης





Συλλογή Στοιχείων



Συλλογή Στοιχείων

- ❑ Τα δεδομένα προήλθαν από το ερευνητικό έργο i-DREAMS, στο οποίο 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 **οδηγικά χαρακτηριστικά** συμπεριφοράς



Περιγραφή Επεξεργασία Στοιχείων

Περιγραφή - Επεξεργασία Στοιχείων



Μεταβλητές	Επεξήγηση
TTC	Χρόνος πρόσκρουσης
Headway	Χρονική απόσταση οχημάτων
Speed	Ταχύτητα οχήματος
Distance_travelled	Απόσταση που διανύθηκε
BSAV_SpeedLimitMS	Όριο ταχύτητας
BSAV_SpeedLimitKPH	Όριο ταχύτητας
HandsOnEvent	Ένδειξη ότι τα χέρια του οδηγού βρίσκονται στο τιμόνι
FatigueEvent	KSS score
LaneDepartureWarningActive	Προειδοποίηση αλλαγής λωρίδας



LaneDepartureWarningActive_mean

Speed_mean
Speed_std
Speed_min
Speed_max

TTC_mean
TTC_std
TTC_min
TTC_max

BSAV_SpeedLimitMS_max
BSAV_SpeedLimitKPH_max

Distance
travelled_sum

HandsOnEvent_mean
HandsOnEvent_median
FatigueEvent_median

Headway_mean
Headway_std
Headway_median
Headway_min
Headway_max



Περιγραφή - Επεξεργασία Στοιχείων

Ανάλυση συσχέτισης μεταξύ των μεταβλητών



	TTC_mean	TTC_std	Headway_median	Headway_mean	Headway_std	Speed_std	HandsOnEvent_median	HandsOnEvent_mean	FatigueEvent_median	ME_ForwardCollisionWarning_mean	ME_LaneDepartureWarningActive_mean	TTC_min	TTC_max	Headway_min	Headway_max	Speed_min	Speed_max	Speed_mean_1	Distance_travelled_sum	BSAV_SpeedLimitMS_max	BSAV_SpeedLimitKPH_max
TTC_mean																					
TTC_std	0.7663																				
Headway_median	0.0176	-0.0044																			
Headway_mean	0.0413	-0.0172	0.6106																		
Headway_std	0.0448	-0.0217	0.3254	0.8881																	
Speed_std	-0.0709	-0.0371	-0.0193	0.1120	0.1897																
HandsOnEvent_median	-0.0200	0.0007	-0.0284	-0.0619	-0.0612	0.1291															
HandsOnEvent_mean	-0.0216	0.0046	-0.0178	-0.0693	-0.0771	0.1363	0.9750														
FatigueEvent_median	-0.0164	-0.0310	-0.0219	-0.0669	-0.0782	0.0822	0.2700	0.2588													
ME_ForwardCollisionWarning_mean	0.0021	0.0097	-0.0017	-0.0047	-0.0054	-0.0068	-0.0165	-0.0197	0.0205												
ME_LaneDepartureWarningActive_mean	0.0626	0.0402	-0.0089	-0.0326	-0.0403	-0.0887	-0.0315	-0.0371	-0.0114	-0.0065											
TTC_min	0.3084	-0.0028	0.0827	0.2357	0.2712	-0.1657	-0.0895	-0.0927	-0.0627	-0.0205	0.0375										
TTC_max	0.7472	0.9961	-0.0039	-0.0159	-0.0202	-0.0437	-0.0027	0.0015	-0.0325	0.0093	0.0412	0.0052									
Headway_min	0.1253	0.0227	0.3209	0.5707	0.5046	-0.0706	-0.0439	-0.0494	-0.0670	-0.0106	-0.0024	0.3876	0.0251								
Headway_max	0.0437	-0.0213	0.2491	0.7858	0.9682	0.1991	-0.0565	-0.0744	-0.0750	-0.0052	-0.0397	0.2648	-0.0198	0.4588							
Speed_min	0.1048	0.0982	-0.1431	-0.4428	-0.5275	-0.6765	-0.0485	-0.0392	-0.0445	0.0298	0.1245	-0.0795	0.0989	-0.2467	-0.5179						
Speed_max	0.1047	0.1242	-0.1771	-0.3797	-0.3936	-0.3903	0.0006	-0.0047	-0.0672	0.0328	0.1432	-0.1519	0.1234	-0.2726	-0.3710	0.8099					
Speed_mean_1	0.0967	0.1172	-0.1822	-0.4728	-0.5274	-0.5592	-0.0187	-0.0154	-0.0492	0.0325	0.1326	-0.1266	0.1174	-0.2840	-0.5098	0.9475	0.9335				
Distance_travelled_sum	-0.0819	-0.0104	-0.0827	-0.2516	-0.2955	0.0341	0.0229	0.0223	0.1779	-0.0029	0.0268	-0.2253	-0.0126	-0.2265	-0.2840	0.1608	0.0641	0.1215			
BSAV_SpeedLimitMS_max	0.1139	0.0612	-0.0300	-0.0053	0.0125	-0.4861	-0.0721	-0.0924	-0.0919	0.0413	0.1241	0.0877	0.0660	0.0056	0.0164	0.5563	0.6892	0.6245	-0.0533		
BSAV_SpeedLimitKPH_max	0.1139	0.0612	-0.0300	-0.0053	0.0125	-0.4861	-0.0721	-0.0924	-0.0919	0.0413	0.1241	0.0877	0.0660	0.0056	0.0164	0.5563	0.6892	0.6245	-0.0533	1.0000	





Περιγραφή - Επεξεργασία Στοιχείων

Ανάλυση συσχέτισης μεταξύ των μεταβλητών

- ☐ **Υψηλή συσχέτιση** μεταξύ διαφορετικών περιγραφικών στατιστικών της ίδιας μεταβλητής
- ☐ **Υψηλή συσχέτιση** μεταξύ Ταχύτητας και Όριο ταχύτητας
- ☐ **Μέτρια συσχέτιση** μεταξύ Ταχύτητας και Χρονική απόσταση οχημάτων





Εφαρμογή μεθοδολογίας



Εφαρμογή μεθοδολογίας

Καθορισμός **‘Οδηγικής Συμπεριφοράς’** μέσω δοκιμών ορίων ταχύτητας με την μέγιστη ταχύτητα και σύγκριση αποτελεσμάτων

❑ “Φυσιολογική” (class: 0):

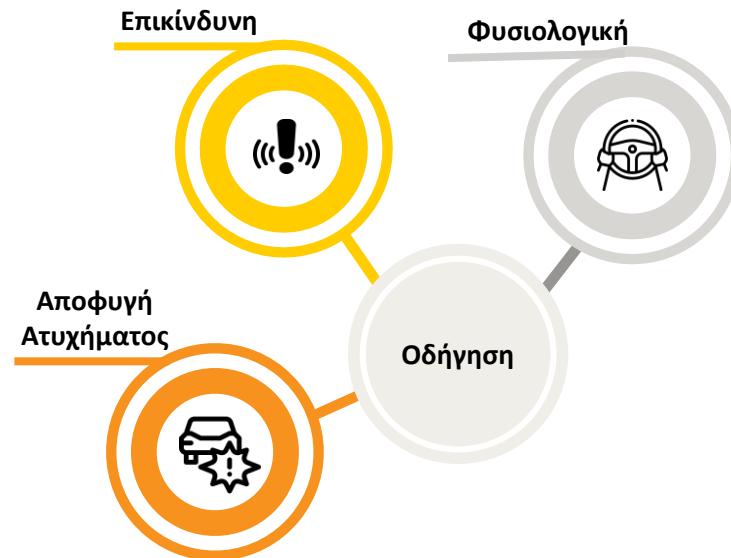
Μέγιστη Ταχύτητα $\leq 0,8 * \text{Τρέχον όριο ταχύτητας}$

❑ “Επικίνδυνη” (class: 1):

$0,8 * \text{Τρέχον όριο ταχύτητας} \leq \text{Μέγιστη Ταχύτητα} \leq \text{Τρέχον όριο ταχύτητας}$

❑ “Αποφυγή Ατυχήματος” (class: 2):

Μέγιστη Ταχύτητα $\Rightarrow \text{Τρέχον όριο ταχύτητας}$





Εφαρμογή μεθοδολογίας

Διαδικασία ταξινόμησης

- ☐ **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 70% και 30% αντίστοιχα
- ☐ **Επαναδειγματοληψία** δεδομένων εκπαίδευσης με χρήση κατάλληλου αλγορίθμου
- ☐ Ανάπτυξη **τεσσάρων αλγορίθμων ταξινόμησης**: (1) Ridge Classifier (RID), (2) Support Vector Machines (SVM), (3) Random Forest Classifier (RF) και (4) XgBoost (XG)



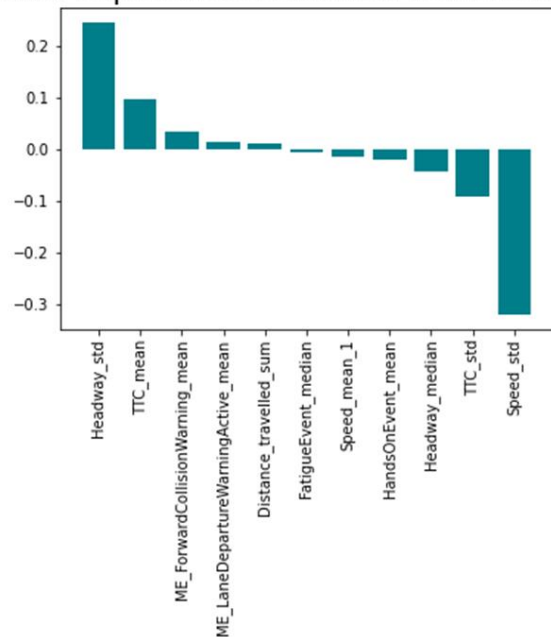


Εφαρμογή μεθοδολογίας

Επιλογή μεταβλητών - Σημαντικότητα

- ☐ Μέσω του αλγορίθμου **Ridge Classifier** αξιολογήθηκε η σημαντικότητα των μεταβλητών στην ταξινόμηση της **‘Οδηγικής Συμπεριφοράς’**
- ☐ Χρονική απόσταση οχημάτων και Ταχύτητα ήταν οι **πιο σημαντικοί παράγοντες** από όσους εξετάστηκαν
- ☐ Απόσταση που διανύθηκε , Επίπεδο κούρασης , Ένδειξη ότι τα χέρια του οδηγού βρίσκονται στο τιμόνι ήταν οι **λιγότερο σημαντικοί παράγοντες**

Feature importances obtained from coefficients





Εφαρμογή μεθοδολογίας

Επιλογή μεταβλητών - Σημαντικότητα

Ξεχωρίσαν 2 ομάδες μεταβλητών



Ομάδα (Α)

Χρόνος πρόσκρουσης, Χρονική απόσταση οχημάτων , Ταχύτητα, Προειδοποίηση αλλαγής λωρίδας



Ομάδα (Β)

Χρόνος πρόσκρουσης, Χρονική απόσταση οχημάτων, Ένδειξη ότι τα χέρια του οδηγού βρίσκονται στο τιμόνι , Επίπεδο κούρασης, Προειδοποίηση αλλαγής λωρίδας, Ταχύτητα , Απόσταση που διανύθηκε

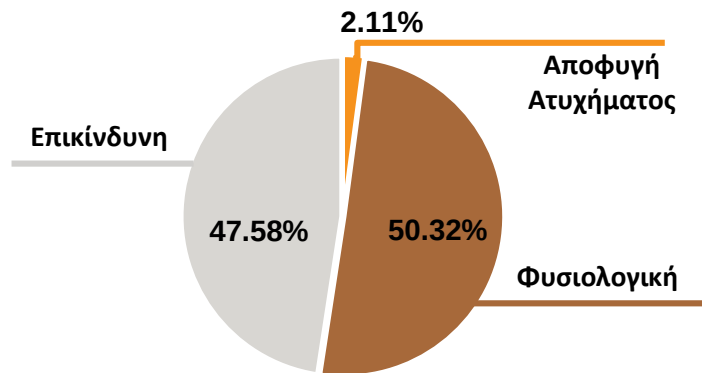




Εφαρμογή μεθοδολογίας

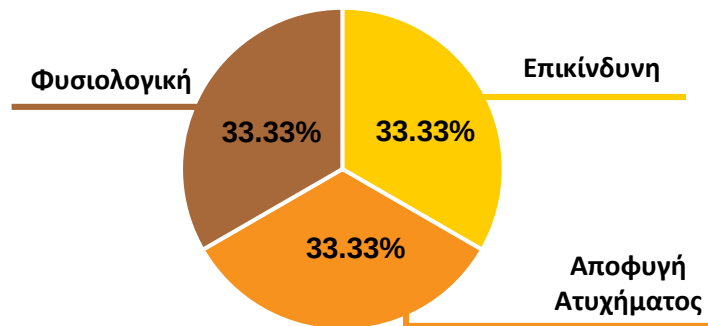
Με τον καθορισμό των επιπέδων προέκυψε ανισότητα των δειγμάτων στα τρία επίπεδα
Εφαρμόστηκε η 'Τεχνική συνθετικής υπερδειγματοληψίας μειοψηφίας' (**SMOTE**) με σκοπό την:

- ☐ Αντιμετώπιση **μεροληψίας** μοντέλων ως προς την κυρίαρχη τάξη
- ☐ **Αντιστάθμιση** της άνισης κατανομής των τάξεων



Πριν την επαναδιαμόρφωση

Μετά την επαναδιαμόρφωση



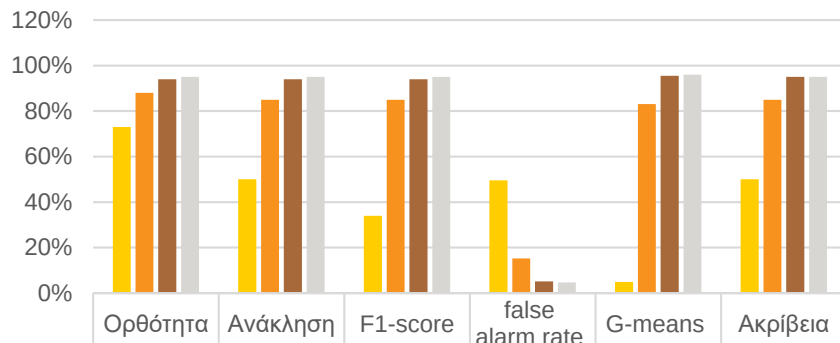
Εφαρμογή μεθοδολογίας



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

- ☐ **Ακρίβεια:** Από τις προβλέψεις που πραγματοποιήθηκαν πόσες ήταν ορθές
- ☐ **Ανάκληση:** Από τα πραγματικά θετικά στοιχεία πόσα κατάφερε να προβλέψει το μοντέλο
- ☐ **f1-score:** Αρμονικό μέτρο της ακρίβειας και της ανάκλησης
- ☐ **False alarm rate:** τον αριθμό των ψευδών συναγερμών ανά συνολικό αριθμό προειδοποιήσεων
- ☐ **G-means:** Σταθμισμένη αξιολόγηση όλων των τιμών μας

Σταθμισμένος Μέσος όρος προβλέψεων

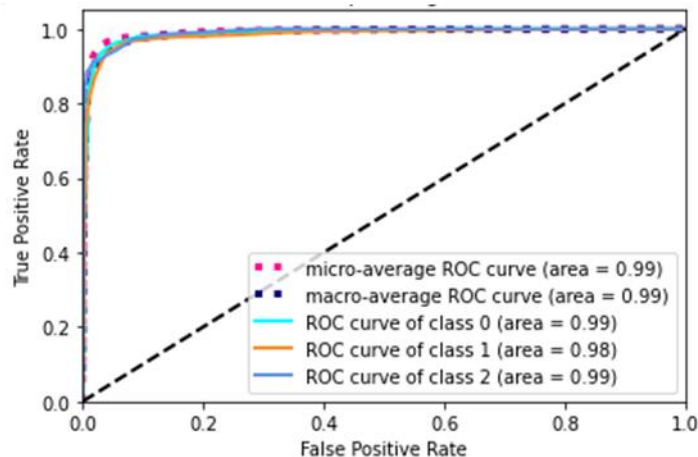


	Ορθότητα	Ανάκληση	F1-score	false alarm rate	G-means	Ακρίβεια
RidgeClassifier	73%	50%	34%	50%	5%	50%
Support Vector Machines	88%	85%	85%	15%	83%	85%
RandomForestClassifier	94%	94%	94%	5%	95%	95%
XgBoost	95%	95%	95%	5%	96%	95%

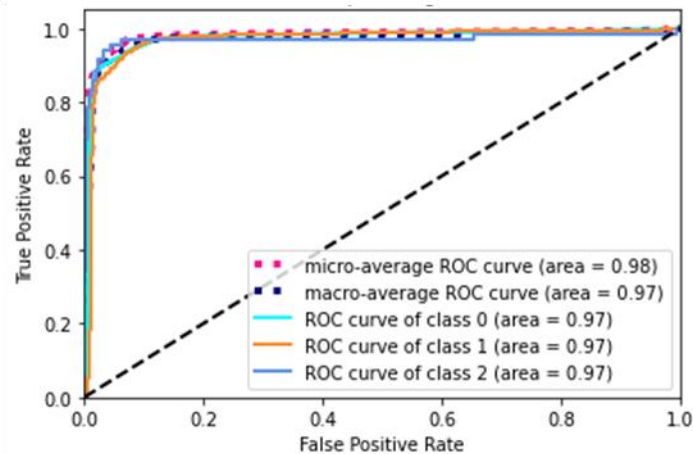




Η καμπύλη ROC-Random Forest



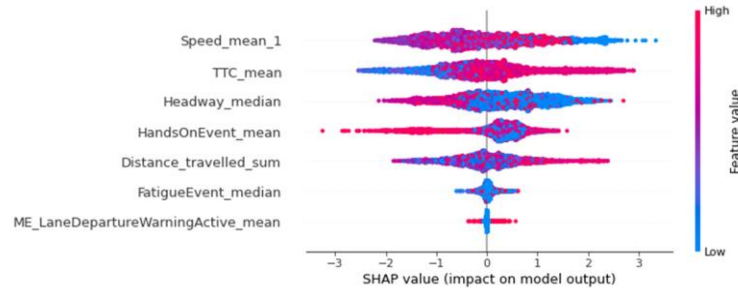
Η καμπύλη ROC-XgBoost



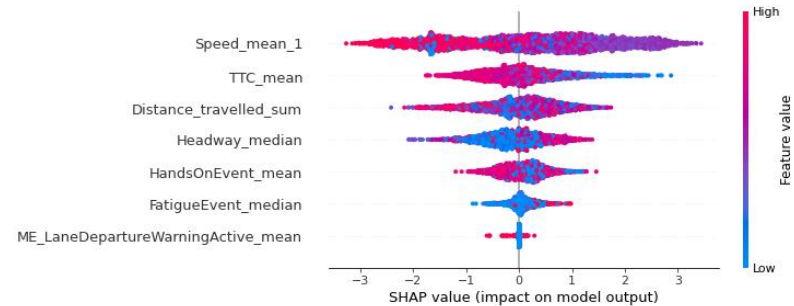


Διαδικασία Κατανόησης του μοντέλου Random Forest μέσω τιμών Shap

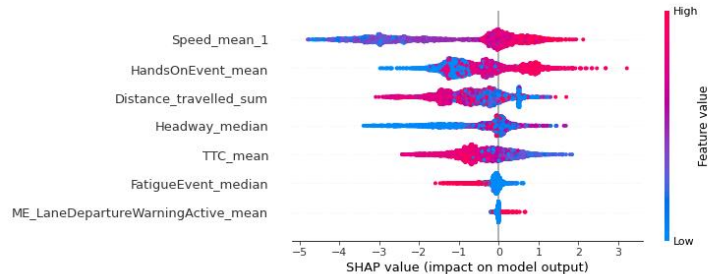
Φυσιολογική



Επικίνδυνη



Αποφυγή Ατυχήματος

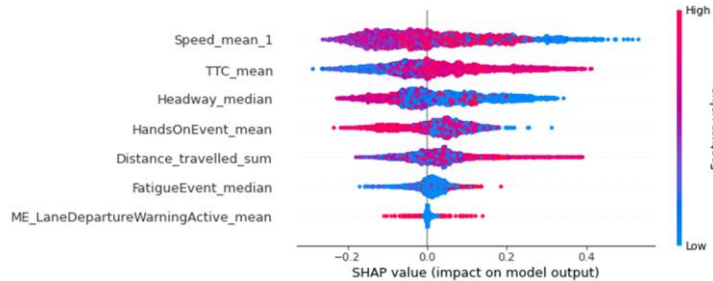


- ❑ Η ταχύτητα είναι η πιο **σημαντική μεταβλητή** σε όλες της κατηγορίες
- ❑ Ο χρόνος πρόσκρουσης **επηρεάζει** τις δύο πρώτες κατηγορίες σε μεγάλο βαθμό
- ❑ Σε κάθε κατηγορία παρουσιάζεται διαφορετική βαρύτητα για της μεταβλητές

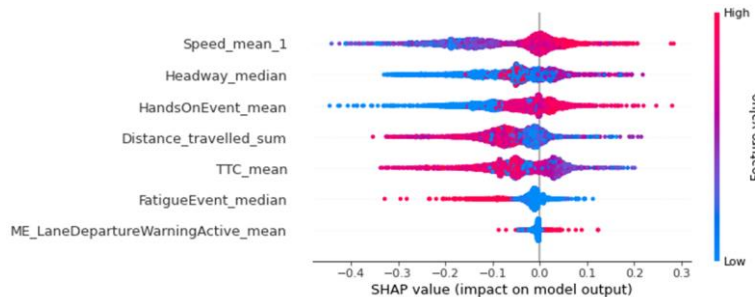




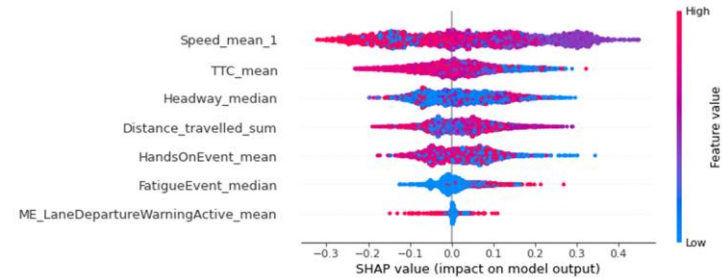
Φυσιολογική



Αποφυγή Ατυχήματος



Επικίνδυνη



- ❑ Η ταχύτητα είναι η πιο **σημαντική μεταβλητή** σε όλες της κατηγορίες
- ❑ Ο χρόνος πρόσκρουσης **επηρεάζει** τις δύο πρώτες κατηγορίες σε μεγάλο βαθμό
- ❑ Σε κάθε κατηγορία παρουσιάζεται διαφορετική βαρύτητα για της μεταβλητές





Συμπεράσματα

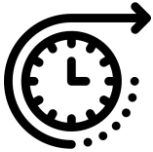


Συμπεράσματα

- ☐ Η μέθοδος Random Forest και η μέθοδος XgBoost από την ομάδα (B) σημείωσαν τις **υψηλότερες επιδόσεις** στην πλειοψηφία των μετρικών αξιολόγησης τους
- ☐ Οι πιο **σημαντικές μεταβλητές** που ωθούν τον οδηγό στην πιο επικίνδυνη οδήγηση είναι η Ταχύτητα, Χρόνος πρόσκρουσης, Χρονική απόσταση οχημάτων και Ένδειξη ότι τα χέρια του οδηγού βρίσκονται στο τιμόνι.
- ☐ Και στα τέσσερα μοντέλα η ομάδα με τον μεγαλύτερο αριθμό μεταβλητών παρουσίασε πιο **ικανοποιητικά αποτελέσματα**.
- ☐ Οι μεταβλητές Προειδοποίηση αλλαγής λωρίδας και το Επίπεδο κούρασης είχαν **μειωμένη επιρροή** στην διαδικασία της ταξινόμησης. Ωστόσο σχετίζονται άμεσα με τους υπόλοιπους παράγοντες (όπως η ταχύτητα ή διανυθείσα απόσταση)
- ☐ Οι μέθοδοι Ridge Classifier και Support Vector Machines **δεν παρουσίασαν** τόσο ικανοποιητικά αποτελέσματα όσο τα υπόλοιπα μοντέλα



Προτάσεις για περαιτέρω έρευνα



Προτάσεις για περαιτέρω έρευνα

- ☐ Αξιοποίηση **μεγαλύτερου όγκου δεδομένων**
- ☐ Ανάπτυξη **εναλλακτικών** μοντέλων μηχανικής εκμάθησης
- ☐ Ανάπτυξη **εναλλακτικών** τεχνικών εξέτασης σημαντικότητας χαρακτηριστικών
- ☐ Εξέταση μοντέλων **βαθιάς εκμάθησης**
- ☐ Εξέταση **διαφορετικών μεθόδων** για τον καθορισμό του επιπέδου «**Οδηγικής Συμπεριφοράς**»
 - Διερεύνηση **σημαντικότητας μεταβλητών** για διαφορετικά μοντέλα και σύγκριση αποτελεσμάτων
- ☐ Διερεύνηση **πρόσθετων** παραγόντων



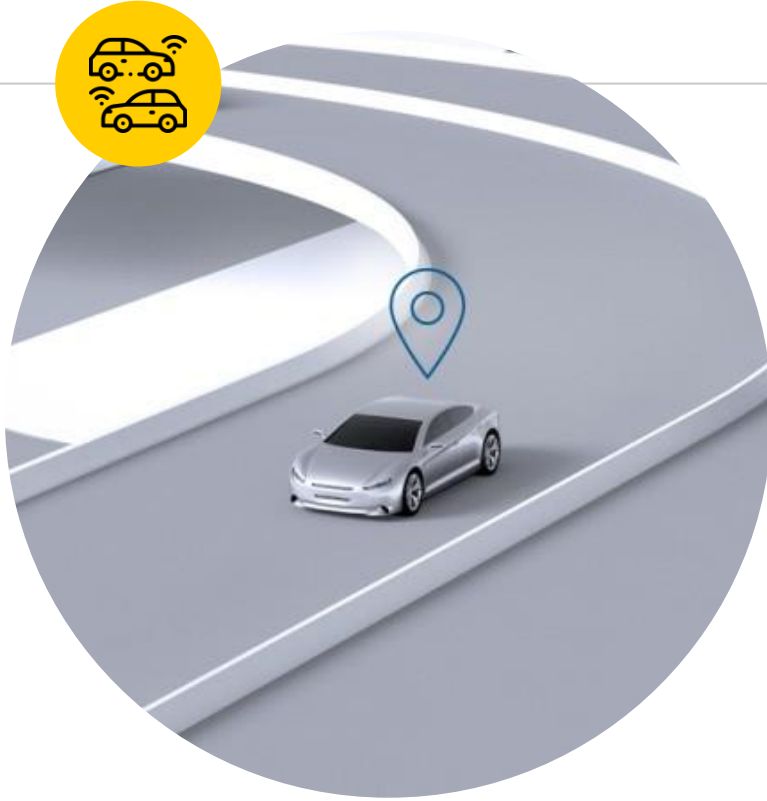


Προτάσεις για αξιοποίηση των αποτελεσμάτων



Προτάσεις αξιοποίηση των αποτελεσμάτων

- ☐ Χρήση μοντέλων κατηγοριοποίησης **προσδιορισμού επιπέδου ασφαλείας** ενός οδηγού, υπό πραγματικές συνθήκες οδήγησης
- ☐ Εξελιγμένα **συστήματα υποβοήθησης οδηγού** για τον προσδιορισμό «οδηγικής συμπεριφοράς» σε πραγματικό χρόνο (πρόβλεψη επικίνδυνων εξωτερικών παραγόντων ασφάλειας-ανάληψη δράσης-αντιμετώπισης).
- ☐ Δημιουργία μιας «**smart drive**» **εφαρμογής** για τη συλλογή εξατομικευμένων πληροφοριών για την «οδηγική ικανότητα» κάθε οδηγού (χρόνος, τρόπος και προσαρμογές οδήγησης).
- ☐ Εξοπλισμός οχημάτων με αλγοριθμικά συστήματα, που **προβλέπουν την ύπαρξη συμβάντος** και την ενημέρωση του οδηγού προς την εκτέλεση κατάλληλων ενεργειών αποφυγής του.



**Ευχαριστώ για την
προσοχή σας!**



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής



Εντοπισμός επικίνδυνης συμπεριφοράς με τεχνικές μηχανικής μάθησης

Έκτορας Καμβουσιώρας



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

Αθήνα, Ιούλιος 2022