



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής

Πρόβλεψη της οδηγικής συμπεριφοράς με
δεδομένα από διαφορετικές χώρες και με
χρήση τεχνικών μηχανικής μάθησης

Γιάννης Ρούκος

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

Αθήνα, Μάρτιος 2025



Αντικείμενο-Στόχος



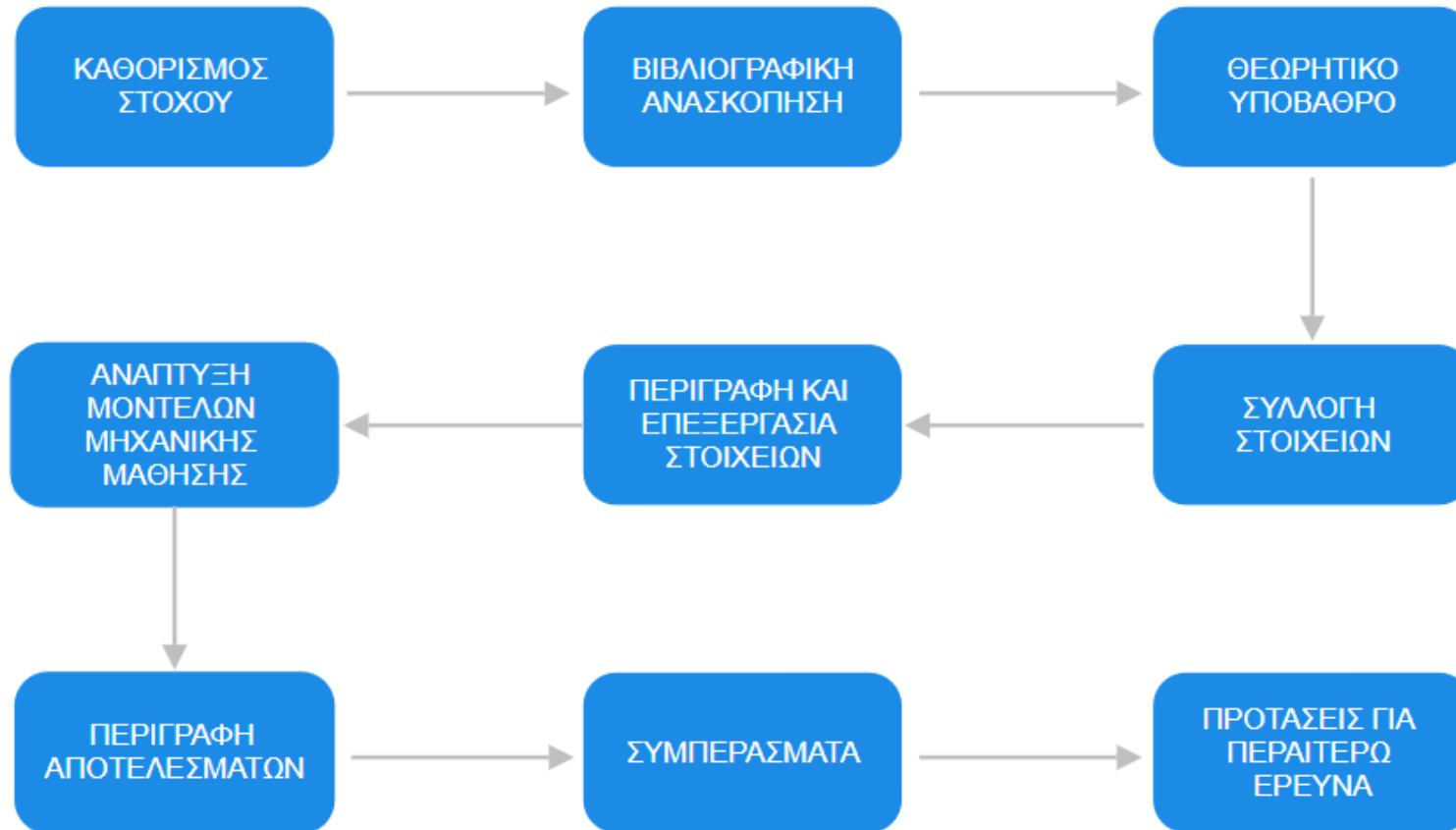
Ανάλυση **επιρροής** διαφορετικών **παραγόντων** που επηρεάζουν την επικίνδυνη οδήγηση στο **Βέλγιο** και στο **Ηνωμένο Βασίλειο** με βάση πείραμα οδήγησης σε πραγματικές συνθήκες στο πλαίσιο του i-Dreams

Η ανάλυση των συγκεκριμένων παραγόντων πραγματοποιήθηκε σε δύο κύρια στάδια:

1. Με την ανάπτυξη μοντέλων μηχανικής και βαθιάς μάθησης που ταξινομούν τους οδηγούς σε **τρία επίπεδα** "Ζώνης Ανοχής Ασφαλείας"
2. Με την αξιοποίηση της **μεθόδου SHAP** για την αξιολόγηση της συνεισφοράς κάθε χαρακτηριστικού της οδηγικής συμπεριφοράς στο επίπεδο επικινδυνότητας στην τελική πρόβλεψη



Μεθοδολογία Διπλωματικής Εργασίας



Βιβλιογραφική Ανασκόπηση (1/2)

Συναφείς έρευνες (Ανάλυση και αναγνώριση οδηγικής συμπεριφοράς)

- Αναγνώριση τύπου οδηγικής συμπεριφοράς
- Αναγνώριση της κατάστασης του οδηγού
- Εντοπισμός της κατάστασης απόσπασης της προσοχής των οδηγών
- Προσδιορισμός των δευτερευουσών ενεργειών που επηρεάζουν την απόσπαση
- Καθορισμός ομάδων επικινδυνότητας με βάση τα δημογραφικά χαρακτηριστικά
- Καθορισμός σταδίων επικινδυνότητας με βάση την πρόβλεψη της επικίνδυνης οδήγησης

Από τις τεχνικές που αναλύθηκαν:

- Τα μοντέλα "Τυχαίων Δασών", "Πολυεπίπεδου Perceptron", "Βαθμιδωτής Ενίσχυσης" κι ένα μοντέλο "Bi-LSTM" σημείωσαν τα υψηλότερα ποσοστά ορθών προβλέψεων



Βιβλιογραφική Ανασκόπηση (2/2)

Ελλείψεις και περιορισμοί προηγούμενων ερευνών

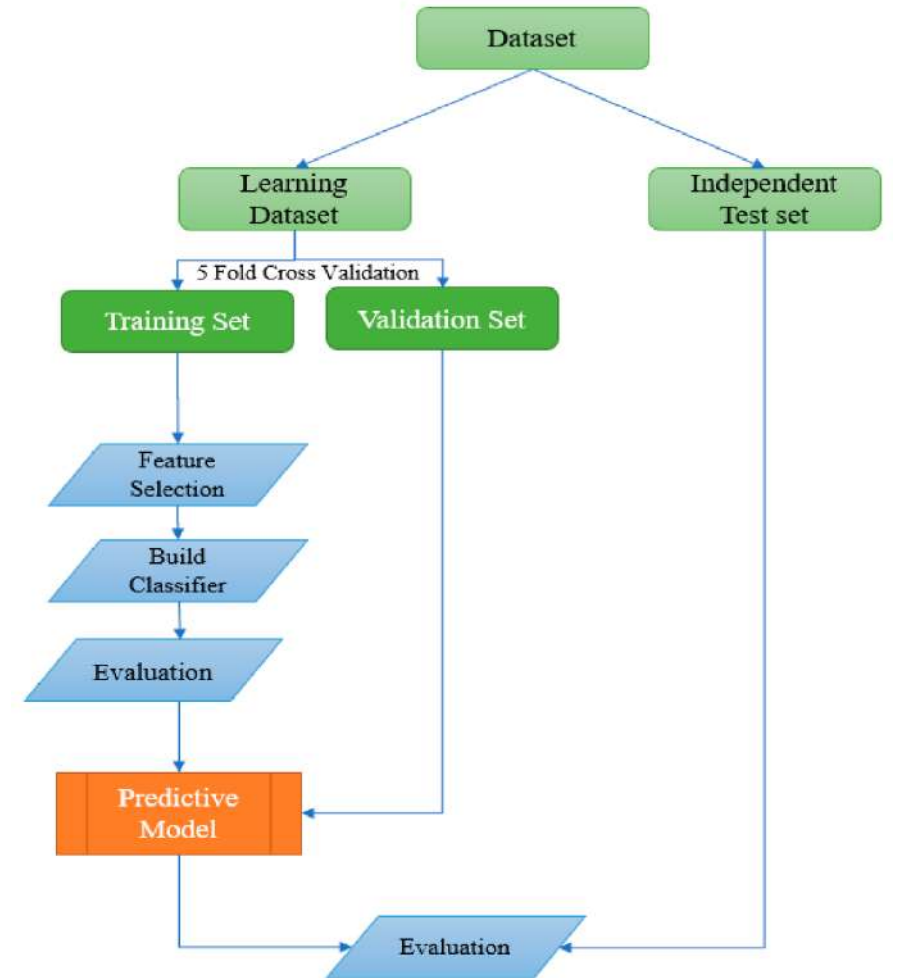
- Τα κριτήρια βαθμονόμησης που ελήφθησαν ενδέχεται να μην είναι εφαρμόσιμα σε όλους τους τύπους οδών
- Δεν ελήφθησαν υπόψη παράγοντες όπως το όριο ταχύτητας κι ο ψυχικός φόρτος εργασίας
- Δεν εξετάστηκαν τόσο η οπτική/σωματική καθώς κι η γνωστική απόσπαση
- Δεν ελήφθη υπόψη η επίδραση του τύπου οδοστρώματος, των γεωμετρικών χαρακτηριστικών και των χαρακτηριστικών του οχήματος
- Τα δεδομένα προήλθαν από υποκειμενικά ερωτηματολόγια
- Δεν καλύφθηκαν διαδικασίες αλλαγής λωρίδας ή προσπέρασης



Θεωρητικό Υπόβαθρο (1/2)

Ανάπτυξη μοντέλων μηχανικής μάθησης

- Επιλογή χαρακτηριστικών
- Αντιμετώπιση της ανισορροπίας δεδομένων μέσω της τεχνικής SMOTE
- Ανάπτυξη αλγορίθμων ταξινόμησης μετά από αρχικό διαχωρισμό των δεδομένων
- Περιγραφή μετρικών αξιολόγησης
- Κατανόηση της μεθόδου SHAP



Θεωρητικό Υπόβαθρο (2/2)

Αλγόριθμοι ταξινόμησης

- Τυχαίο Δάσος (Random Forest)
- Ελαφρύς Αλγόριθμος Βαθιάς Ενίσχυσης (LightGBM)
- Αλγόριθμος Ενίσχυσης Κατηγοριών (CatBoost)
- Πολυεπίπεδο Αντιληπτικό (MultiLayer Perceptron)

Μετρικές αξιολόγησης

- Μήτρα Σύγχυσης (Confusion Matrix)
- Ακρίβεια (Accuracy)
- Ακρίβεια Θετικής Πρόβλεψης (Precision)
- Ανάκληση (Recall)
- Ρυθμός Λανθασμένου Συναγερμού (False Alarm Rate)
- F1-score



Συλλογή κι Επεξεργασία Στοιχείων (1/4)

- Τα δεδομένα προήλθαν από τη βάση δεδομένων του ερευνητικού έργου **i-Dreams**, στα πλαίσια του οποίου πραγματοποιήθηκε ένα πείραμα οδήγησης σε **πραγματικές συνθήκες**
- Τα δεδομένα που επεξεργάζονται αφορούν δύο χώρες, το **Βέλγιο** και το **Ηνωμένο Βασίλειο**
- 96 οδηγοί συμμετείχαν στο πείραμα την τριετία 2020-2023 στις δύο χώρες που εξετάζονται
- Το πείραμα πραγματοποιήθηκε σε **τέσσερις φάσεις**
- Συλλέχθηκαν σημαντικά δεδομένα οδήγησης με διάφορα εργαλεία παρακολούθησης και παρέμβασης (Συσκευή **OBD-II**, Σύστημα **MobilEye**)



Συλλογή κι Επεξεργασία Στοιχείων (2/4)

Εργαλεία παρακολούθησης οδηγικής συμπεριφοράς

Συσκευή OBD-II



Σύστημα MobilEye



Συλλογή κι Επεξεργασία Στοιχείων (3/4)

Μεταβλητές

Απόσταση και Ταχύτητα

Μεταβλητή	Περιγραφή	Μονάδες μέτρησης	Τύπος
Headway	Χρονική απόσταση από το προπορευόμενο όχημα	Δευτερόλεπτα	Αριθμητική
ME_Car_speed_mean	Ταχύτητα οχήματος	Χιλιόμετρα ανά ώρα	Αριθμητική
ME_AWS_hw_measurement_mean	Μέτρηση της μεταβλητής headway	Δευτερόλεπτα	Αριθμητική
GPS_spd_mean	Μέση ταχύτητα ταξιδιού	Χιλιόμετρα ανά ώρα	Αριθμητική
GPS_distances_sum	Συνολική απόσταση του ταξιδιού	Χιλιόμετρα	Αριθμητική

Ασφάλεια και Προειδοποιήσεις

Μεταβλητή	Περιγραφή	Μονάδες μέτρησης	Τύπος
ME_AWS_fcw_mean	Προειδοποίηση μπροστινής σύγκρουσης	-	Διακριτή
ME_AWS_pcw_mean	Προειδοποίηση σύγκρουσης πεζών	-	Διακριτή
ME_AWS_pedestrian_dz_mean	Πεζός σε επικίνδυνη ζώνη	-	Διακριτή
ME_LDW_Map_type_L_median	Προειδοποίηση αναχώρησης από την αριστερή λωρίδα	-	Διακριτή
ME_LDW_Map_type_R_median	Προειδοποίηση αναχώρησης από τη δεξιά λωρίδα	-	Διακριτή



Συλλογή κι Επεξεργασία Στοιχείων (4/4)

Οδική Σήμανση και Περιβαλλοντικές Συνθήκες

Επικίνδυνη Οδήγηση

Μεταβλητή	Περιγραφή	Μονάδες μέτρησης	Τύπος
ME_TSR_tsr_1_sped_median	Εμφάνιση κωδικού πινακίδας 1 ταχύτητας	-	Διακριτή
ME_AWS_time_indicator_median	Υποδεικνύει συνθήκες φωτισμού (ημέρα, σούρουπο, νύχτα)	-	Διακριτή
ME_Car_wipers_median	Υποδεικνύει καιρικές συνθήκες (υαλοκαθαριστήρες on/off)	-	Διακριτή
ME_Car_high_beam_median	Μεγάλη σκάλα	-	Διακριτή

Κατηγορία	Μεταβλητές (Mean)	Μεταβλητές (Sum)
Απότομες επιταχύνσεις	DEM_evt_ha_lvl_L_mean, DEM_evt_ha_lvl_M_mean, DEM_evt_ha_lvl_H_mean	DEM_evt_ha_lvl_L_sum, DEM_evt_ha_lvl_M_sum, DEM_evt_ha_lvl_H_sum
Απότομο φρενάρισμα	DEM_evt_hb_lvl_L_mean, DEM_evt_hb_lvl_M_mean, DEM_evt_hb_lvl_H_mean	DEM_evt_hb_lvl_L_sum, DEM_evt_hb_lvl_M_sum, DEM_evt_hb_lvl_H_sum
Απότομες στροφές	DEM_evt_hc_lvl_L_mean, DEM_evt_hc_lvl_M_mean, DEM_evt_hc_lvl_H_mean	DEM_evt_hc_lvl_L_sum, DEM_evt_hc_lvl_M_sum, DEM_evt_hc_lvl_H_sum

Βιομετρικά Δεδομένα

Μεταβλητή	Περιγραφή	Μονάδες μέτρησης	Τύπος
IBI_value_mean	Χρονικό διάστημα μεταξύ διαδοχικών καρδιακών παλμών	Χιλιοστά του δευτερολέπτου	Αριθμητική



Εφαρμογή Μεθοδολογίας (1/10)

Καθορισμός Επιπέδου "Ζώνης Ανοχής Ασφαλείας" – "STZ_level"

- Η αντιστοίχιση των επιπέδων βασίστηκε στην χρονική απόσταση από το προπορευόμενο όχημα (**Headway**) και για τις δύο χώρες.

Από το πείραμα συλλέχθηκαν οι μεταβλητές `iDreams_Headway_Map_level_i`, όπου $i = -1, 0, 1, 2, 3$.

Για κάθε 30 δευτερόλεπτα προκύπτει το επίπεδο της "Ζώνης Ανοχής Ασφαλείας", το οποίο χαρακτηρίζεται ως:

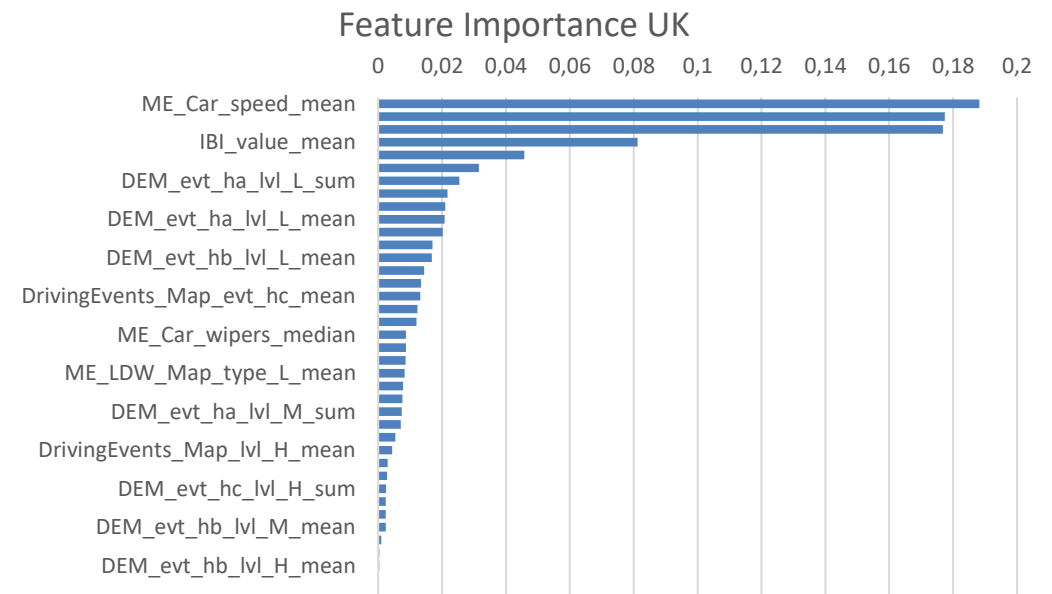
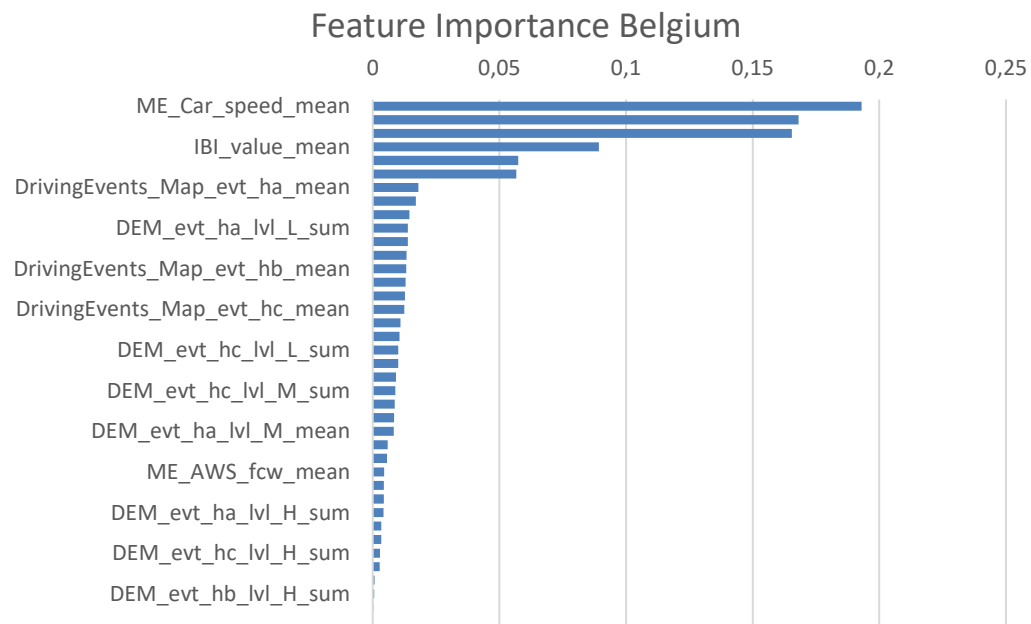
- **Κανονικό (Normal)** όταν η μεταβλητή που επιλέγεται είναι επιπέδου παρέμβασης -1,0,1
- **Επικίνδυνο (Dangerous)** όταν η μεταβλητή που επιλέγεται είναι επιπέδου παρέμβασης 2
- **Αποφυγής Ατυχήματος (Avoidable Accident)** όταν η μεταβλητή που επιλέγεται είναι επιπέδου παρέμβασης 3



Εφαρμογή Μεθοδολογίας (2/10)

Επιλογή Χαρακτηριστικών (Feature Selection)

- Με τον ταξινομητή “Τυχαίων Δασών” υπολογίστηκε η επιρροή της κάθε μεταβλητής
- Επιλέχθηκαν **εννιά μεταβλητές (δεδομένα εισόδου)** και για τα δύο datasets



Εφαρμογή Μεθοδολογίας (3/10)

Προετοιμασία δεδομένων

Δεδομένα πειράματος

- Δεδομένα εισόδου (εννιά επιλεγμένες μεταβλητές)
- Δεδομένα εξόδου (τρία επίπεδα "STZ_level")

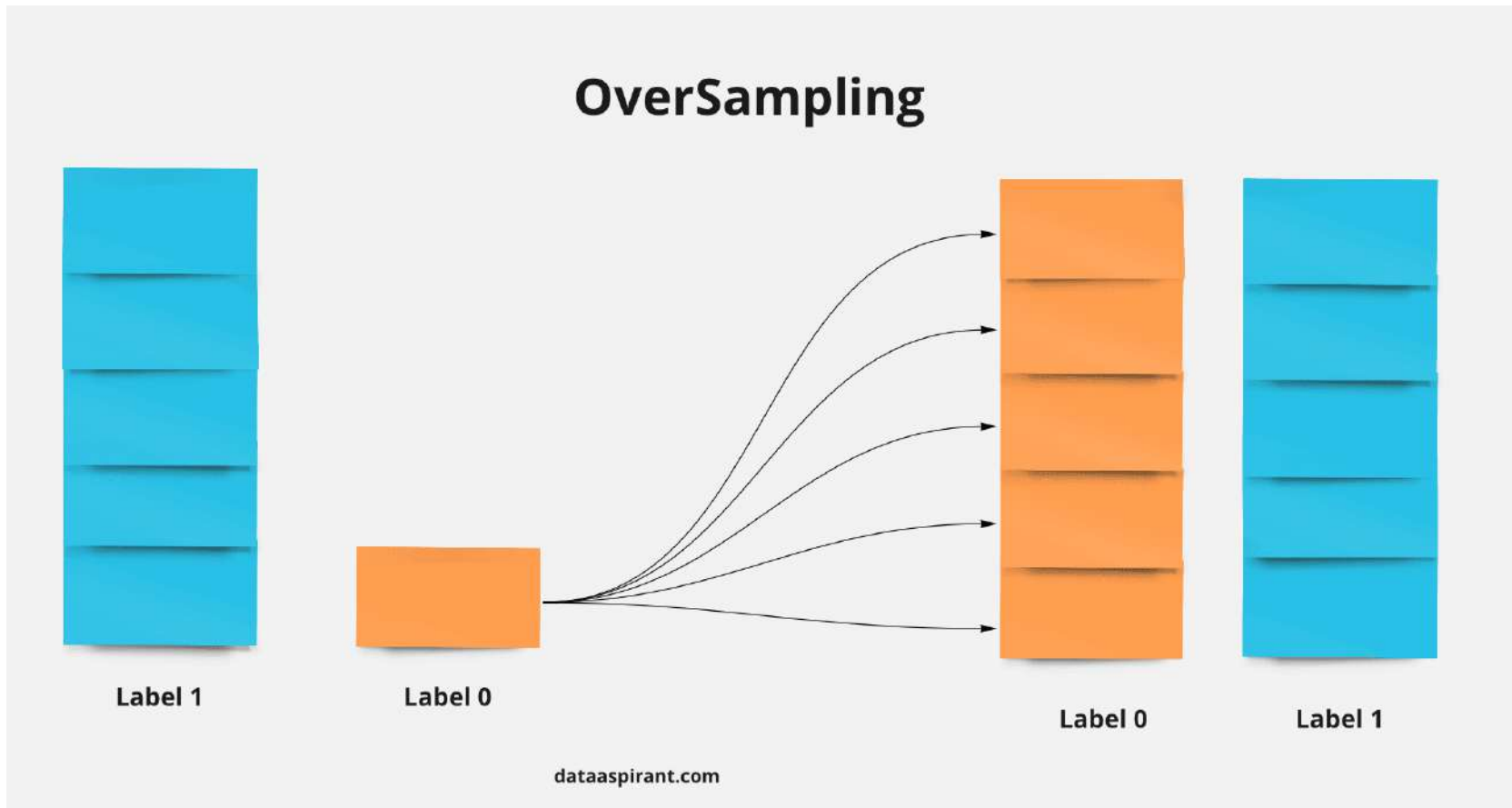
Διαχωρισμός δεδομένων

- Σετ εκπαίδευσης-training dataset (80%) - Πρόβλεψη Επιπέδου Ασφαλείας
- Σετ εξέτασης-testing dataset (20%) – Έλεγχος Ακρίβειας Μοντέλου



Εφαρμογή Μεθοδολογίας (4/10)

Αντιμετώπιση Άνισης Κατανομής Δεδομένων με τη χρήση της μεθόδου SMOTE



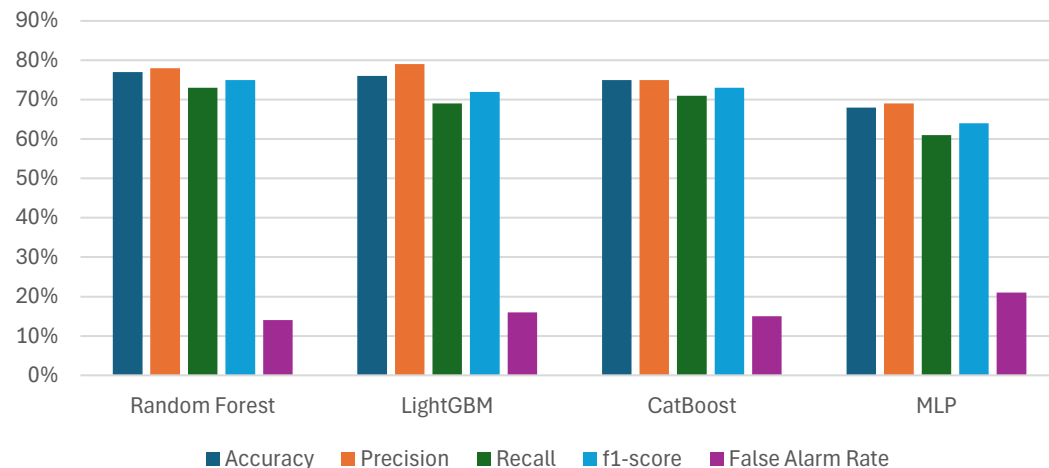
Εφαρμογή Μεθοδολογίας (5/10)

Ανάπτυξη Μοντέλων Ταξινόμησης

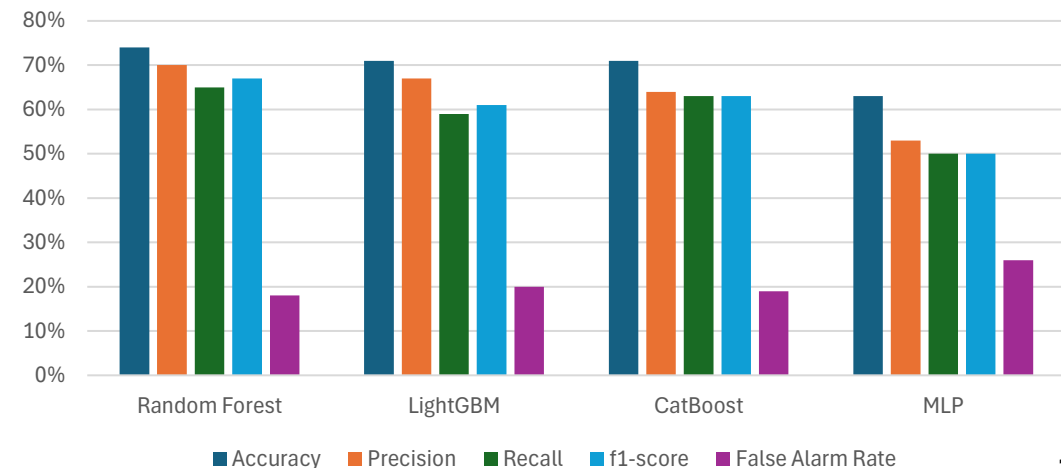
Όνομα Μοντέλου (Ελληνικά)	Όνομα Μοντέλου (Αγγλικά)	Συμβολισμός Μοντέλου
Τυχαίο Δάσος	Random Forest	RF
Ελαφρύς Αλγόριθμος Βαθιάς Ενίσχυσης	Light Gradient Boosting Machine	LightGBM
Αλγόριθμος Ενίσχυσης Κατηγοριών	Categorical Boosting	CatBoost
Πολυεπίπεδο Αντιληπτικό	Multi-Layer Perceptron	MLP

Σύγκριση μετρικών αξιολόγησης μοντέλων

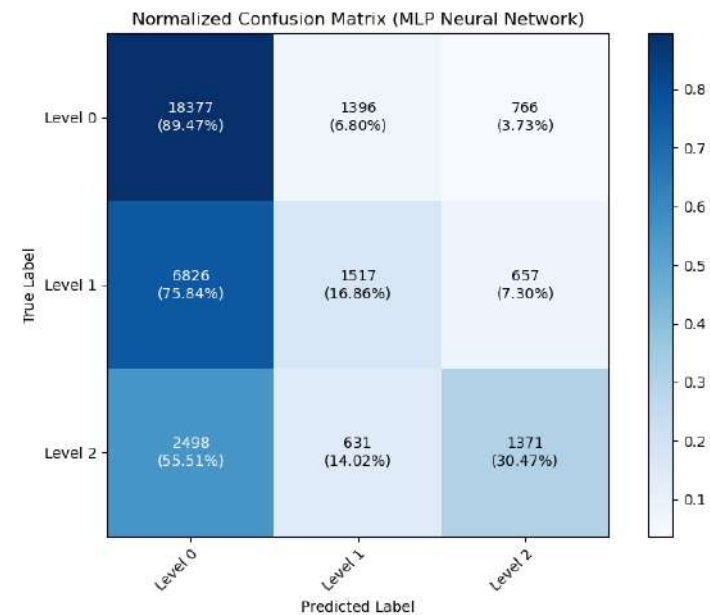
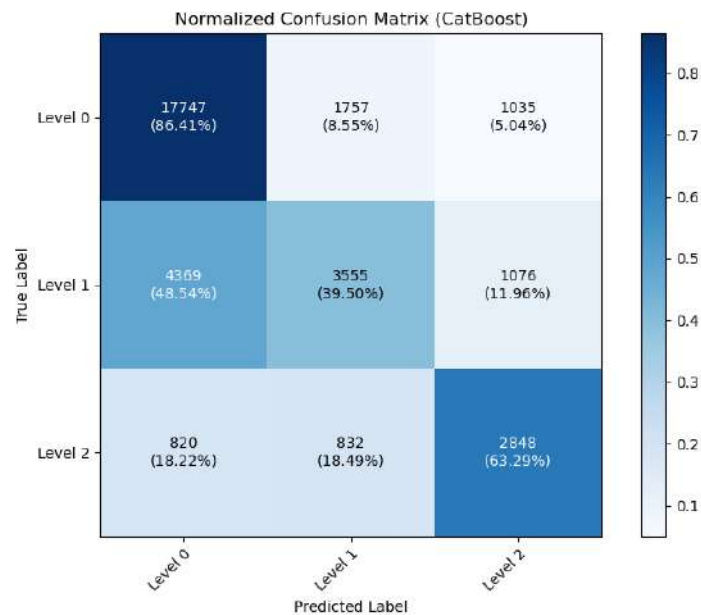
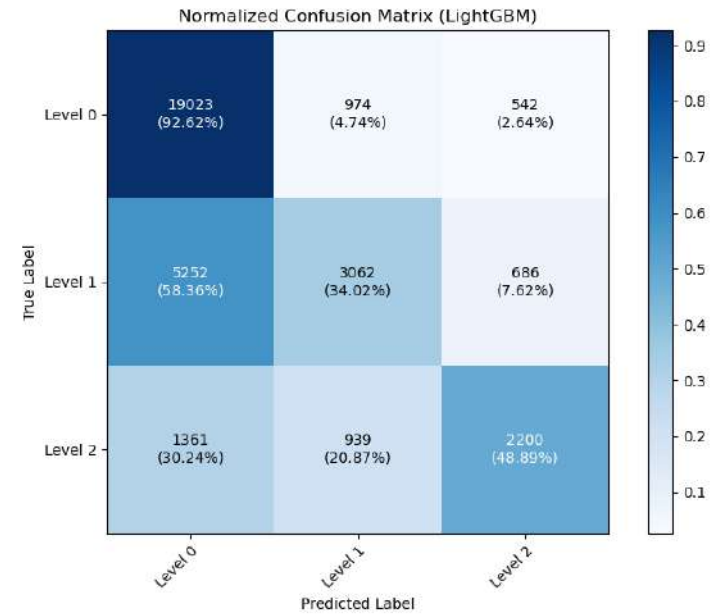
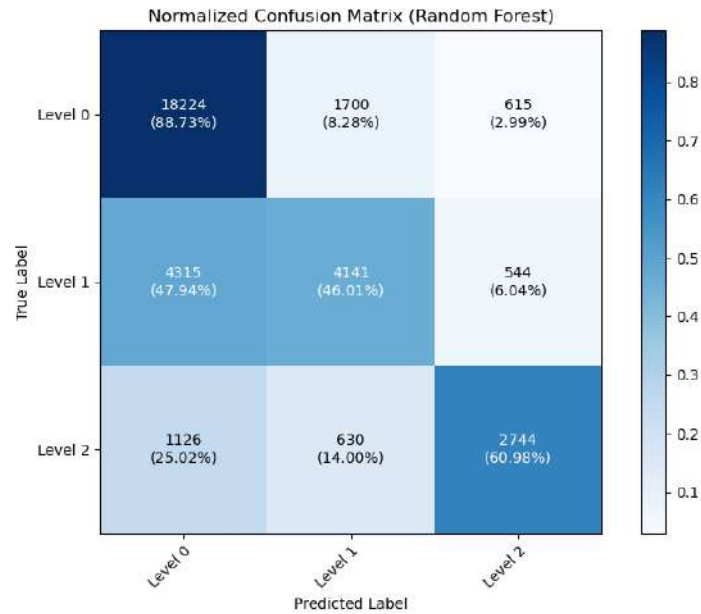
Macro Avg (Ηνωμένο Βασίλειο)



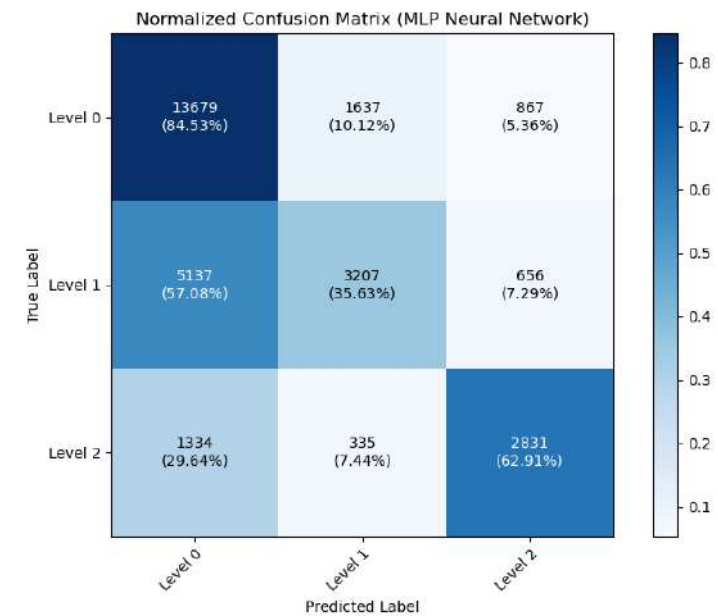
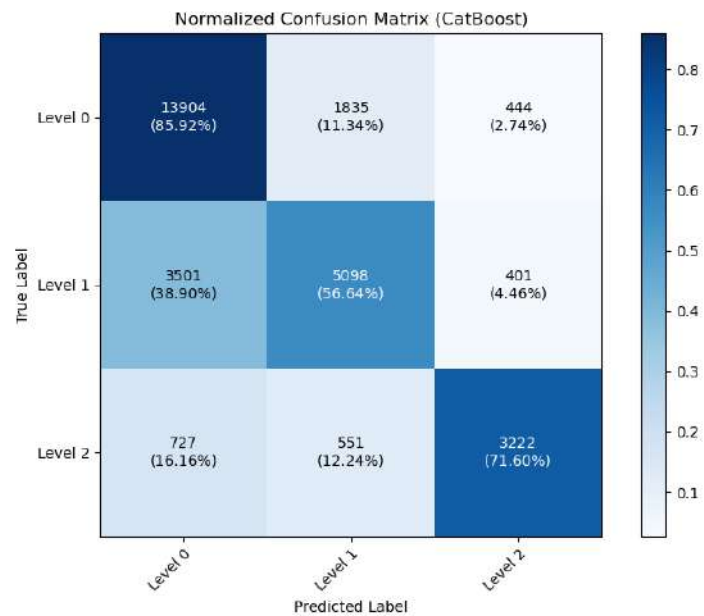
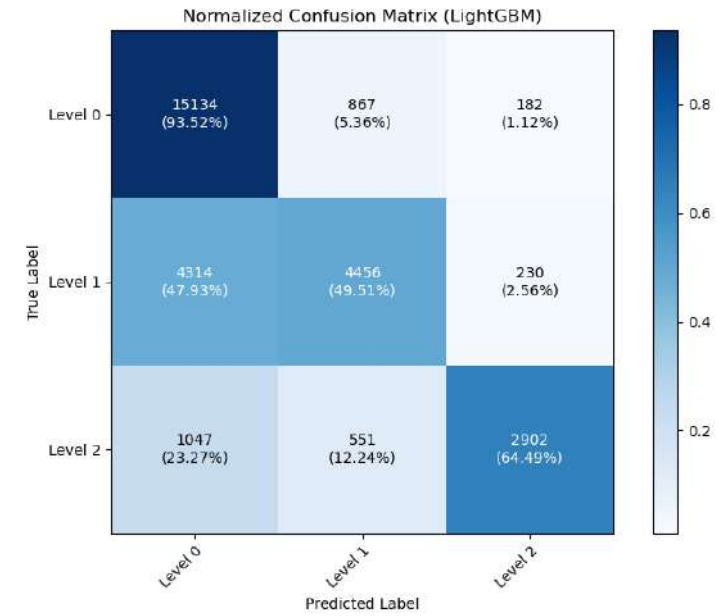
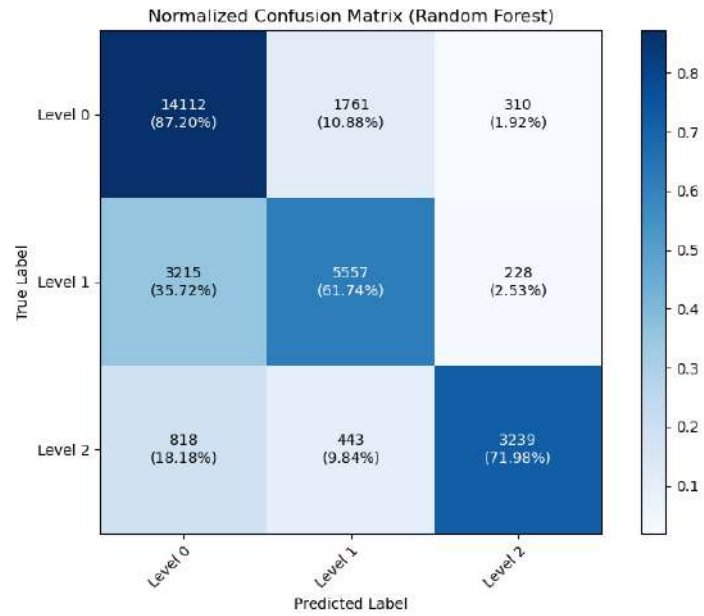
Macro Avg (Βέλγιο)



Εφαρμογή Μεθοδολογίας (6/10)-Μήτρες Σύγκυσης (Βέλγιο)



Εφαρμογή Μεθοδολογίας (7/10)-Μήτρες Σύγκρισης (Ην. Βασίλειο)



Εφαρμογή Μεθοδολογίας (8/10)

Μέθοδος SHAP

- Τα μοντέλα **CatBoost** και **LightGBM** παρουσίασαν την υψηλότερη μέση σημαντικότητα SHAP για τα δεδομένα και των δύο χωρών

Για τα δεδομένα του **Ηνωμένου Βασιλείου**:

- Για τον αλγόριθμο ταξινόμησης **CatBoost**, σημαντικότερη μεταβλητή αποδεικνύεται η μέση ταχύτητα του οχήματος (**ME_Car_speed_mean**), ιδίως για το Level 2 (Αποφυγή Ατυχήματος)
- Στην χρήση του αλγορίθμου ταξινόμησης **LightGBM**, προκύπτει ως σημαντικότερη μεταβλητή το χρονικό διάστημα μεταξύ καρδιακών παλμών (**IBI_value_mean**) και για τα τρία επίπεδα

Για τα δεδομένα του **Βελγίου**:

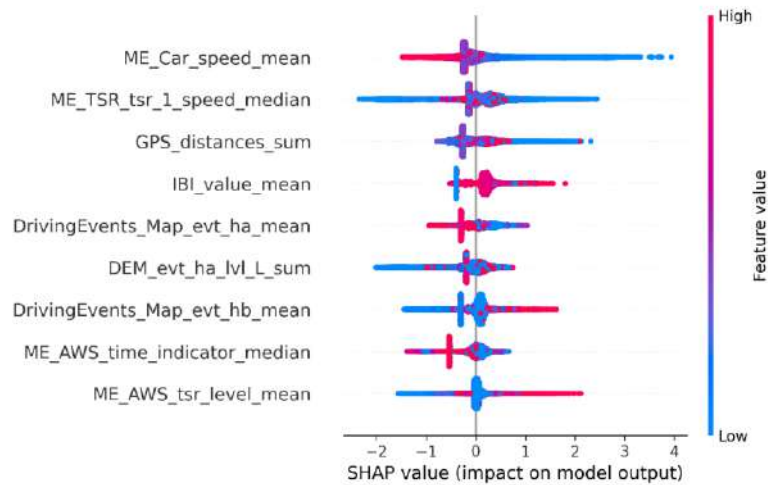
- Για τη χρήση και των δύο αλγορίθμων, σημαντικότερη μεταβλητή αναδεικνύεται το **ME_Car_speed_mean**, κυρίως για το Level 2 (Αποφυγή Ατυχήματος)



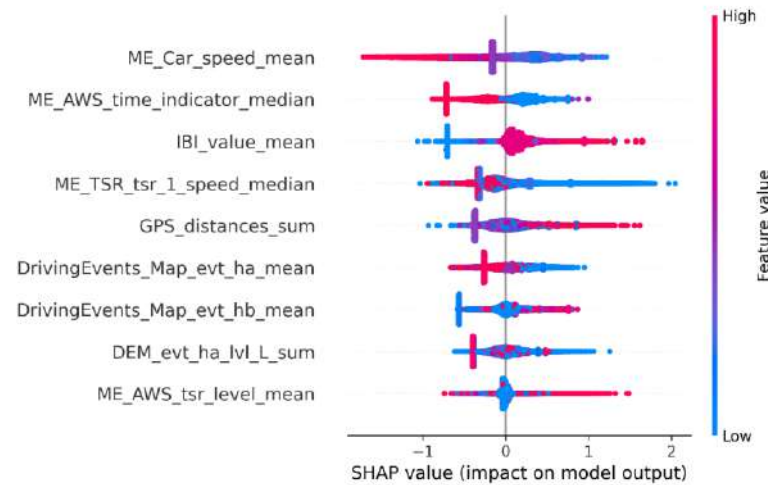
Εφαρμογή Μεθοδολογίας (9/10)- Ηνωμένο Βασίλειο

CatBoost

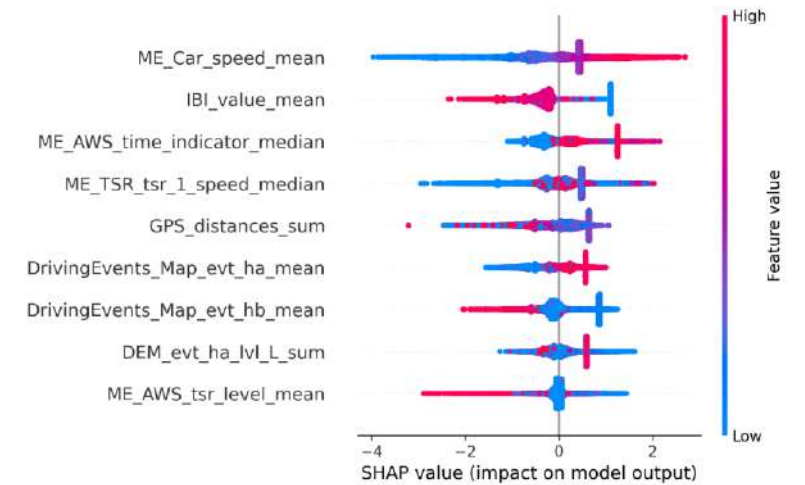
SHAP Summary Plot for STZ Level Normal



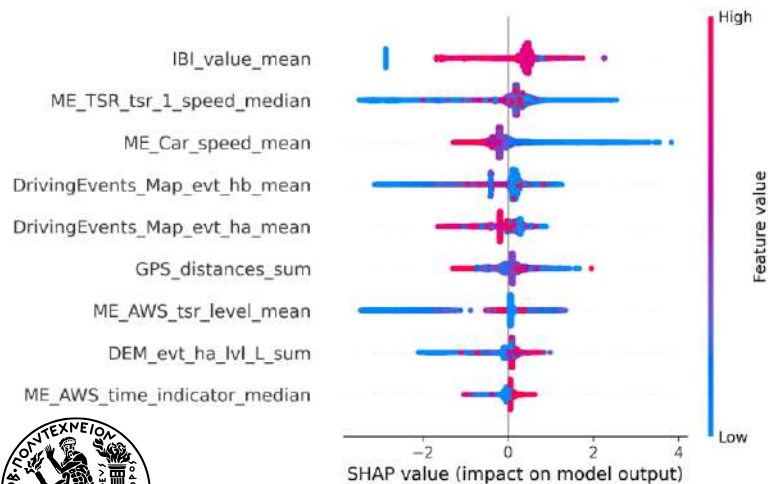
SHAP Summary Plot for STZ Level Dangerous



SHAP Summary Plot for STZ Level Avoidable Accident

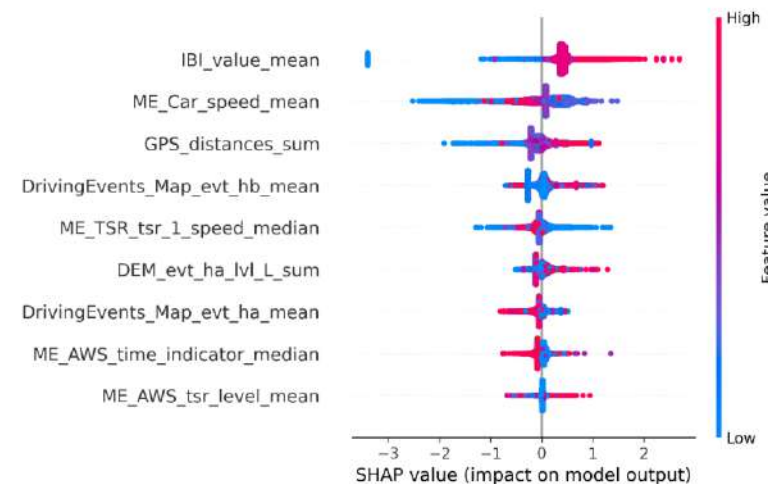


SHAP Summary Plot for STZ Level Normal

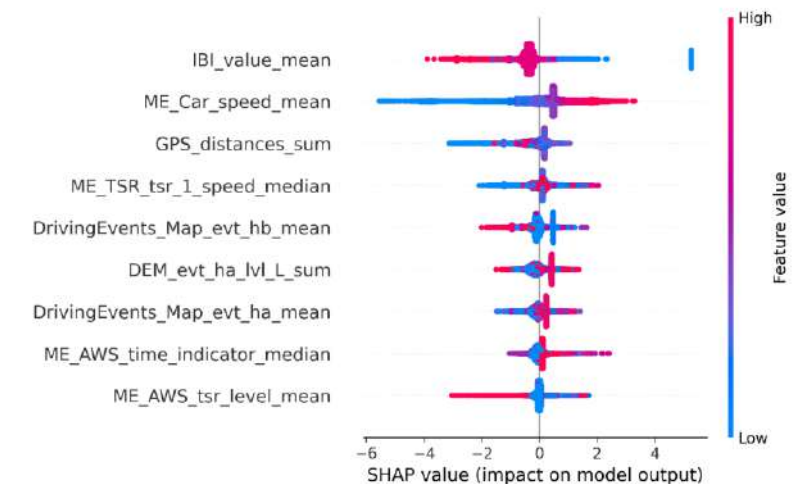


LightGBM

SHAP Summary Plot for STZ Level Dangerous



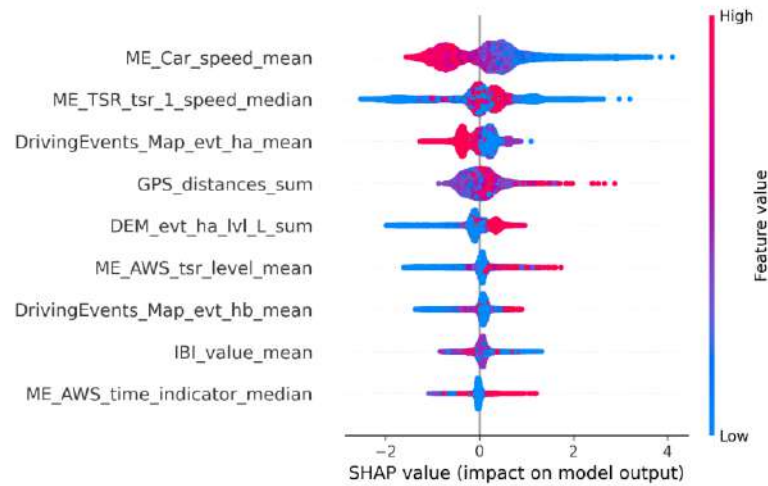
SHAP Summary Plot for STZ Level Avoidable Accident



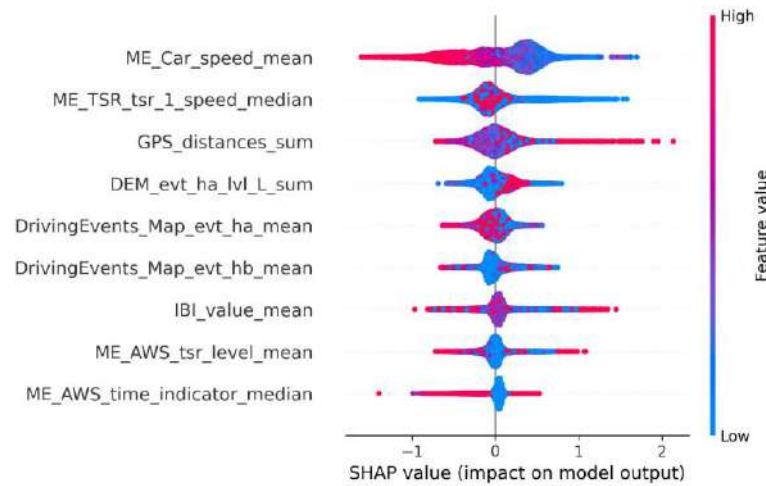
Εφαρμογή Μεθοδολογίας (10/10)- Βέλγιο

CatBoost

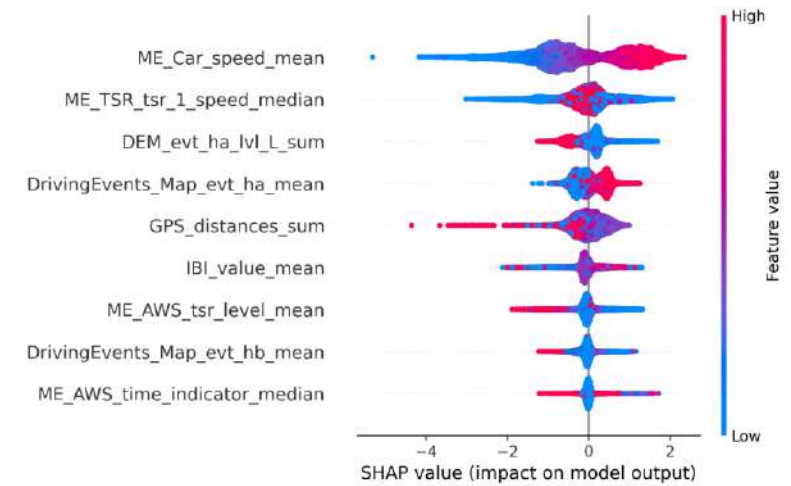
SHAP Summary Plot for STZ Level Normal



SHAP Summary Plot for STZ Level Dangerous

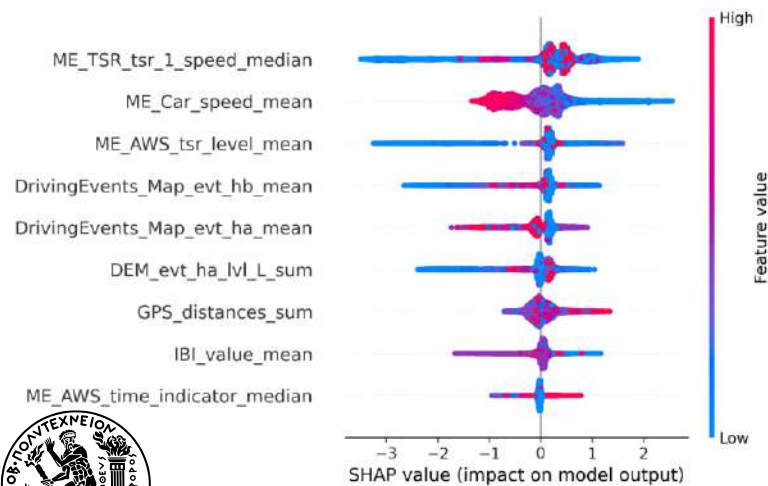


SHAP Summary Plot for STZ Level Avoidable Accident

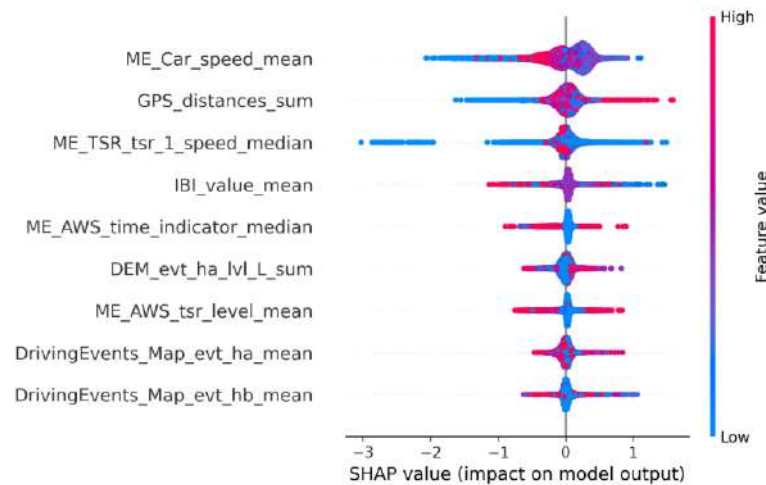


LightGBM

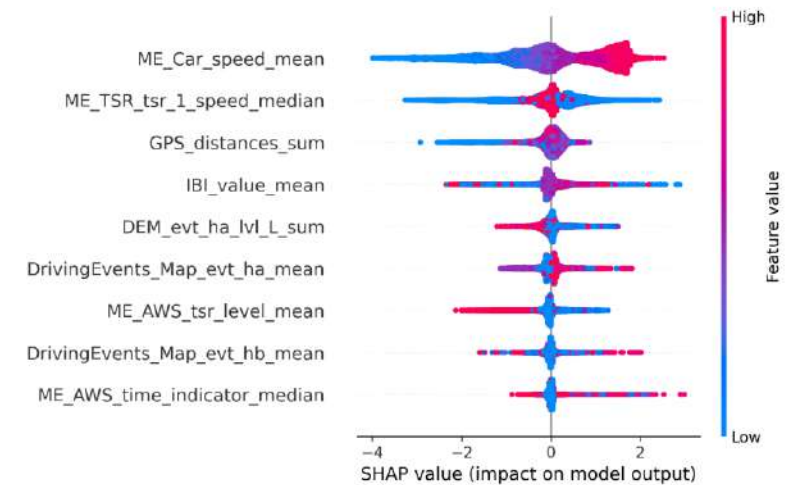
SHAP Summary Plot for STZ Level Normal



SHAP Summary Plot for STZ Level Dangerous



SHAP Summary Plot for STZ Level Avoidable Accident



Συμπεράσματα (1/2)

- Η εφαρμογή της μεθόδου **SMOTE** πράγματι αντιμετώπισε το φαινόμενο της ανισορροπίας δεδομένων μεταξύ των διαφορετικών επιπέδων
- Το μοντέλο ταξινόμησης **Multi-Layer Perceptron (MLP)** φαίνεται το πιο **αδύναμο** μοντέλο για την πρόβλεψη του επιπέδου ασφαλείας
- Τα μοντέλα **CatBoost** και **LightGBM** παρουσίασαν ικανοποιητικές μετρικές αξιολόγησης καθώς και την **υψηλότερη μέση σημαντικότητα SHAP**
- Η **μέση ταχύτητα του οχήματος** υπολογίστηκε ως η **σημαντικότερη μεταβλητή** από τον υπολογισμό σημαντικότητας μεταβλητών με τον αλγόριθμο "Τυχαίου Δάσους"
- Το ίδιο αποτέλεσμα εξήχθη κι από τη **μέθοδο SHAP** για όλες τις περιπτώσεις πλην του μοντέλου **LightGBM** για το **Ηνωμένο Βασίλειο** όπου κατατάχθηκε δεύτερη μετά το **χρονικό διάστημα μεταξύ καρδιακών παλμών**



Συμπεράσματα (2/2)

- Μεγάλη διαφορά παρατηρήθηκε στη σημαντικότητα του **χρονικού διαστήματος μεταξύ καρδιακών παλμών** μεταξύ των δύο χωρών
- Σχετικά υψηλή σημαντικότητα των **απότομων επιταχύνσεων** και των **απότομων φρεναρισμάτων**
- Η **υψηλή επιρροή της συνολικής απόστασης ταξιδιού** υπέδειξε ότι οι αποστάσεις που καλύπτονται έχουν επίσης σημαντική συνεισφορά στον τρόπο οδήγησης
- Τα αποτελέσματα των αναλύσεων έδειξαν ότι οι δύο χώρες παρουσιάζουν **παρόμοια πρότυπα οδήγησης**
- Η **ταχύτητα κατά την εμφάνιση της πρώτης πινακίδας ορίου ταχύτητας** εμφάνισε υψηλή σημαντικότητα τόσο για το Ηνωμένο Βασίλειο, όσο και για το Βέλγιο



Προτάσεις για αξιοποίηση των αποτελεσμάτων

- Χρήση **συνδυαστικών μοντέλων ταξινόμησης** για την αξιολόγηση του επιπέδου ασφάλειας των οδηγών σε πραγματικές συνθήκες
- Διερεύνηση των **κυριότερων μεταβλητών** που επηρεάζουν την αναγνώριση επικίνδυνης οδηγικής συμπεριφοράς
- Ανάπτυξη **συστήματος ανίχνευσης** της 'Ζώνης Ανοχής Ασφαλείας' του οδηγού σε **πραγματικό χρόνο**
- Δημιουργία **εφαρμογής για έξυπνα κινητά τηλέφωνα**
- Εξοπλισμός των οχημάτων με **συστήματα πρόβλεψης ατυχημάτων**



Προτάσεις για μελλοντική έρευνα

- Αξιοποίηση **μεγαλύτερου όγκου δεδομένων** για τη βελτίωση της προγνωστικής ικανότητας των μοντέλων ταξινόμησης
- Ανάπτυξη **εναλλακτικών** μεθόδων **ανάλυσης** της **σημαντικότητας** των μεταβλητών (feature importance)
- Εξέταση **πρόσθετων παραγόντων** που επηρεάζουν την οδηγική συμπεριφορά
- Ανάπτυξη εξειδικευμένων μοντέλων ταξινόμησης ακολουθίας (sequence classification) για την **πρόβλεψη του επιπέδου** 'Ζώνης Ανοχής Ασφαλείας' στο οποίο θα βρίσκεται ο οδηγός **στο επόμενο χρονικό διάστημα** των 30 δευτερολέπτων





Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Πολιτικών Μηχανικών
Τομέας Μεταφορών και Συγκοινωνιακής Υποδομής

Πρόβλεψη της οδηγικής συμπεριφοράς με
δεδομένα από διαφορετικές χώρες και με
χρήση τεχνικών μηχανικής μάθησης

Γιάννης Ρούκος

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

Αθήνα, Μάρτιος 2025

