



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ  
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΕΠΙΚΙΝΔΥΝΗΣ  
ΣΥΜΠΕΡΙΦΟΡΑΣ ΤΩΝ ΟΔΗΓΩΝ ΜΕΣΩ ΤΗΣ  
ΑΞΙΟΠΟΙΗΣΗΣ ΔΕΔΟΜΕΝΩΝ ΟΔΗΓΗΣΗΣ ΣΕ  
ΠΡΑΓΜΑΤΙΚΕΣ ΣΥΝΘΗΚΕΣ

---

Ελένη Μαρία Θεοδωράκη

Επιβλέπων: Γιώργος Γιαννής, Καθηγητής ΕΜΠ

Αθήνα, Μάρτιος 2024



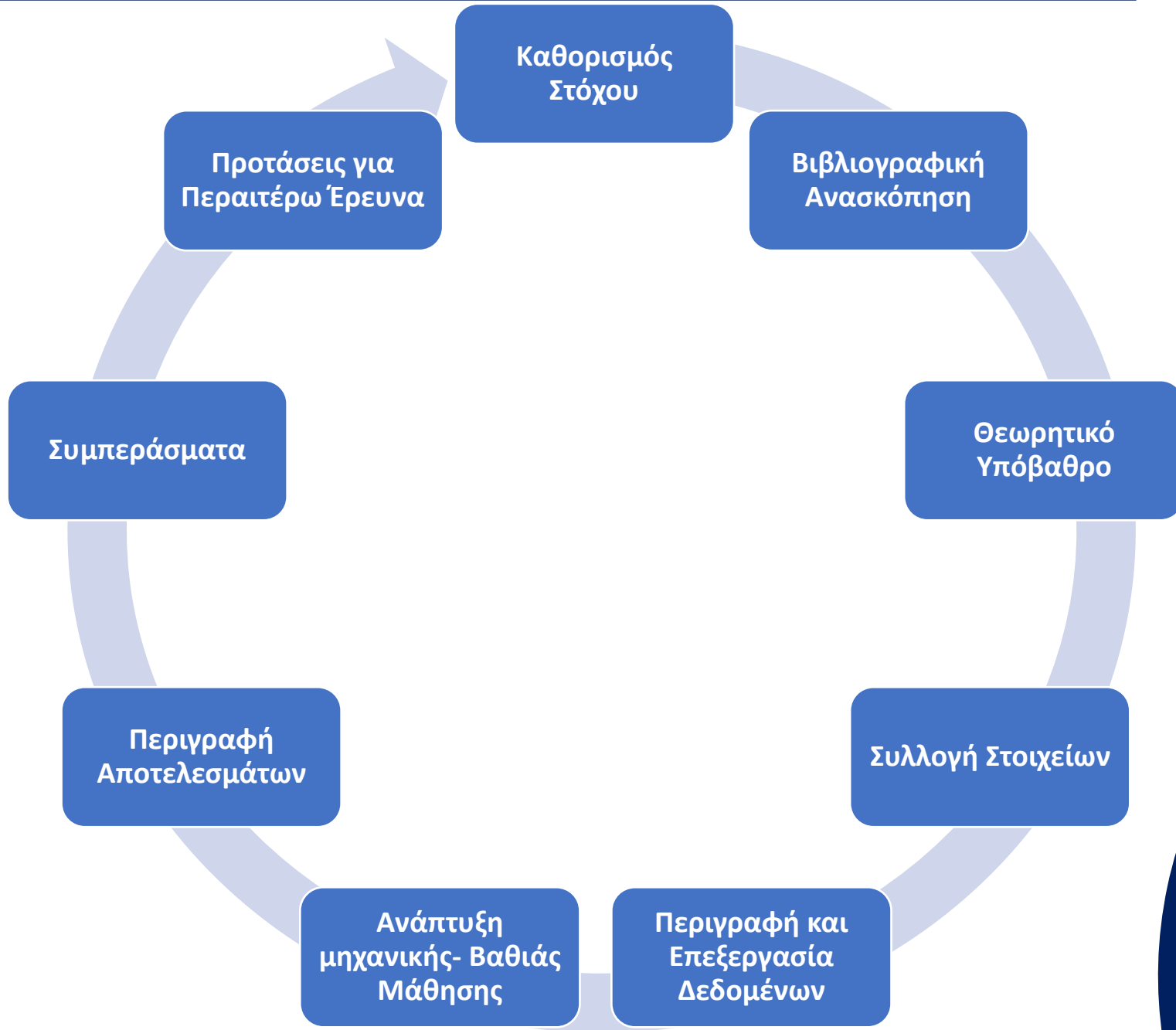
# Στόχος Διπλωματικής Εργασίας

**Εντοπισμός του επιπέδου επικίνδυνης συμπεριφοράς του οδηγού με τεχνικές μηχανικής και βαθιάς μάθησης.**

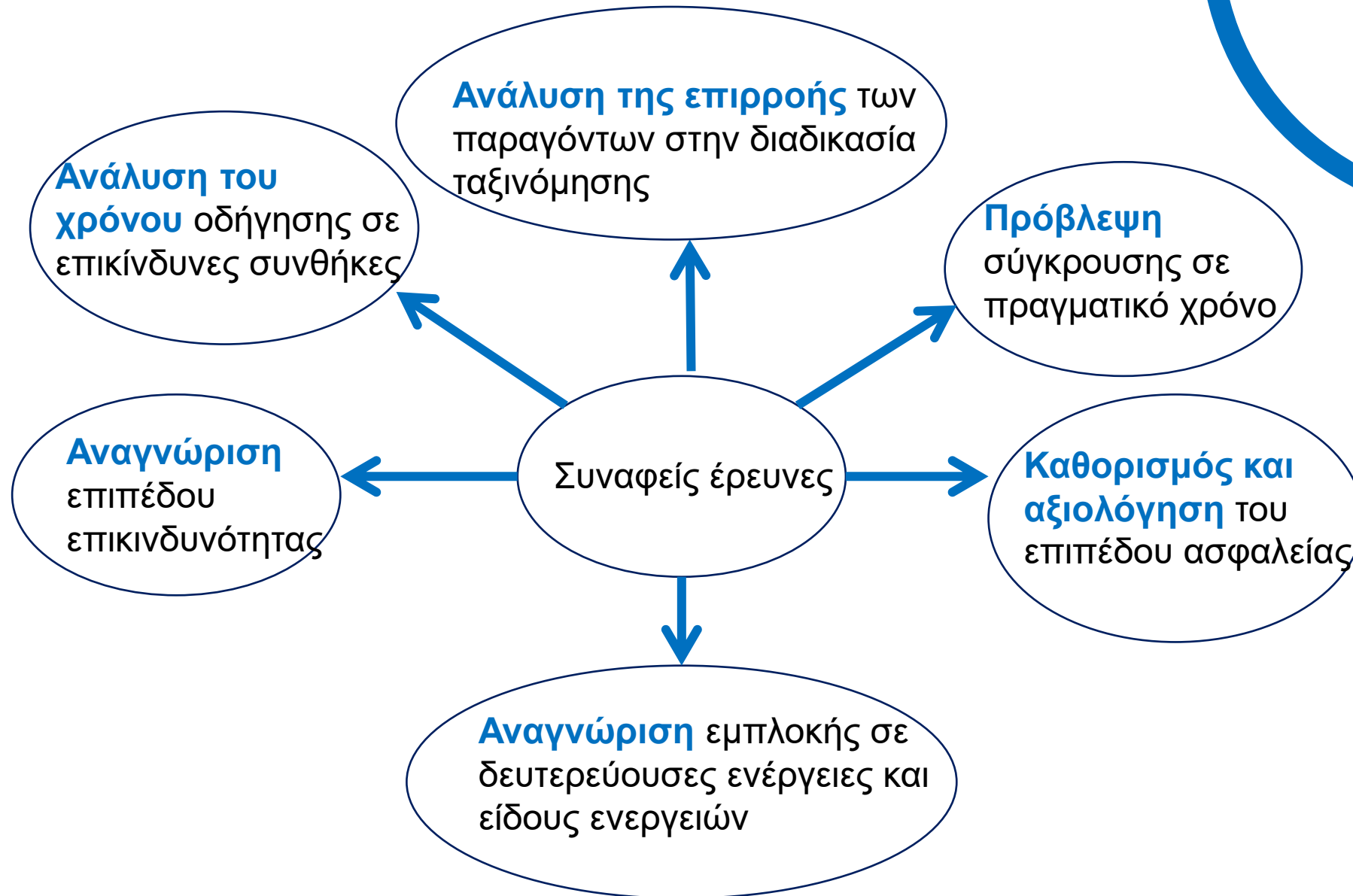
Για την επίτευξη του στόχου πραγματοποιείται:

- Ανάπτυξη συνδυασμένων αλγόριθμων για την **αναγνώριση** του επιπέδου ασφαλείας που βρίσκεται ο κάθε οδηγός
- Ανάπτυξη του αλγόριθμου Lime για την **αξιολόγηση** της επίδρασης του μεταβλητών στην επικίνδυνη οδήγηση

# Μεθοδολογία Διπλωματικής Εργασίας



# Βιβλιογραφική Ανασκόπηση (1/2)



# Βιβλιογραφική Ανασκόπηση (2/2)



**Ελλείψεις** και **περιορισμοί** ερευνών:

- **Απουσία** κατάστασης οδηγού (ψυχολογική ή κόπωση)
- **Απουσία** 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
- **Περιορισμένος** αριθμός τεχνικών αντιμετώπισης της άνισης κατανομής των δειγμάτων στα επίπεδα επικινδυνότητας



# Θεωρητικό Υπόβαθρο (1/2)

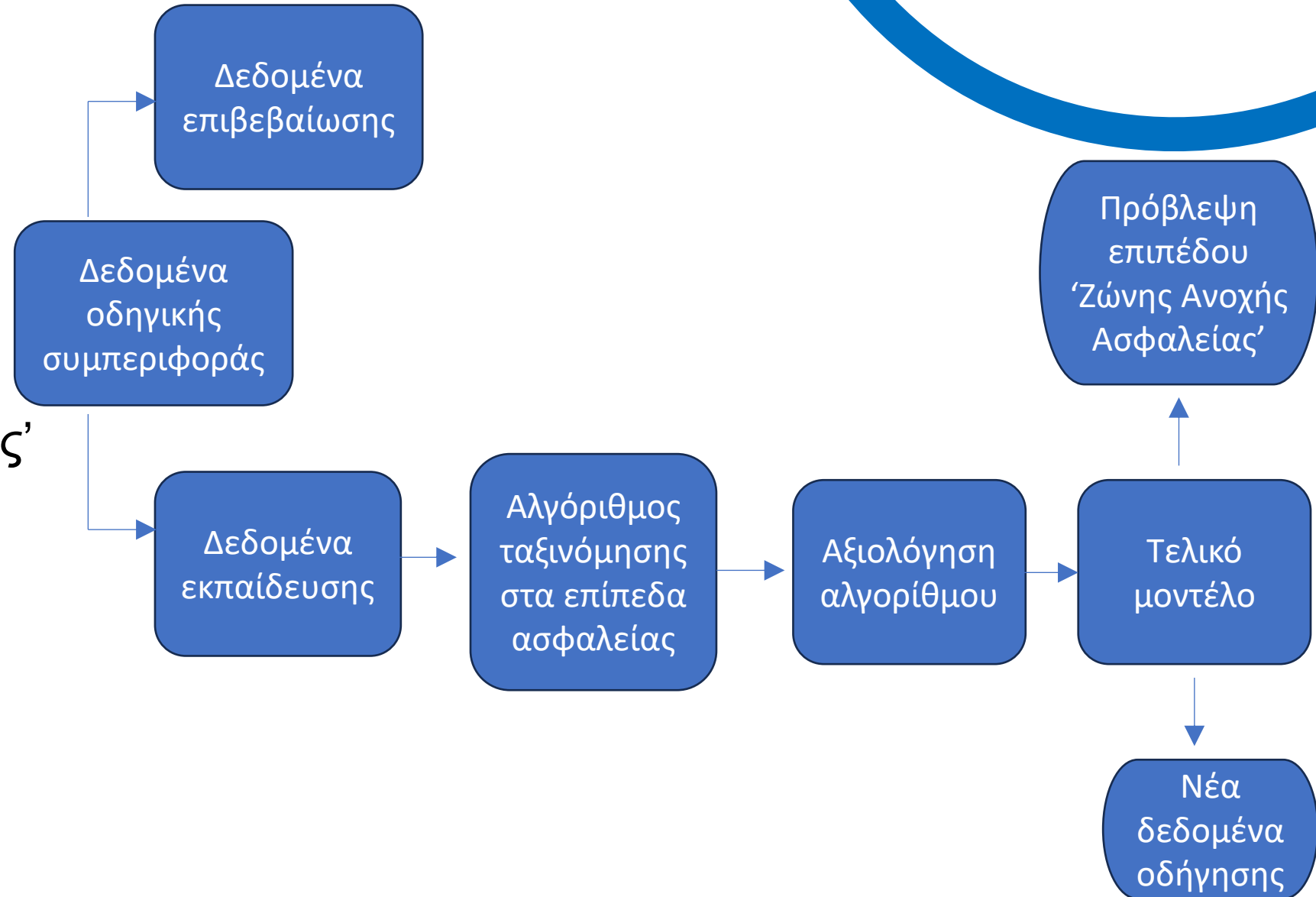


Ανάπτυξη μοντέλων  
**μηχανικής** και **βαθιάς**  
εκμάθησης:

➤ **Ταξινόμησης:** Για την  
πρόβλεψη του επιπέδου  
'Ζώνης Ανοχής Ασφαλείας'  
που βρίσκεται ο οδηγός

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

➤ Η **σημαντικότητα** στην  
διαδικασία της ταξινόμησης



# Θεωρητικό Υπόβαθρο (2/2)



Αξιολόγηση μοντέλων **μηχανικής** και **βαθιάς** μάθησης:

## Μετρικές Αξιολόγησης (Ταξινόμηση):

- Ορθότητα (Accuracy)

- Ακρίβεια (Precision)

- Ανάκλαση (Recall)

- F1 score

- False Alarm rate

- Στο πλαίσιο του ερευνητικού έργου i-Dreams τα δεδομένα συλλέχθηκαν από πείραμα σε **πραγματικές συνθήκες οδήγησης**
- **43** οδηγοί συμμετείχαν από το Βέλγιο δημιουργώντας μία βάση δεδομένων με 7.163 ταξίδια και 147.337 λεπτά
- **26** οδηγοί συμμετείχαν από την Αγγλία δημιουργώντας μία βάση δεδομένων με 8.226 ταξίδια και 118.175 λεπτά
- Το πείραμα στο δρόμο πραγματοποιήθηκε σε **4 φάσεις**



# Συλλογή Στοιχείων (2/2)



- Συλλέχθηκαν σημαντικά **οδηγικά χαρακτηριστικά**
- Χρησιμοποιήθηκαν τεχνολογίες και συστήματα αιχμής για την παρακολούθηση των δεικτών απόδοσης οδήγησης
- **Συσκευές OBD-II** εγκατεστημένες σε οχήματα για την καταγραφή εκατοντάδων παραμέτρων
- **Συστήματα Mobileye**, μέσα από το δίκτυο κινητής τηλεφωνίας συλλέγονται δεδομένα χωρίς την συμμετοχή του χρήστη



# Περιγραφή - Επεξεργασία στοιχείων (1/2)



Μεταβλητές	Επεξήγηση
ME_Car_speed	Ταχύτητα οχήματος
ME_AWS_hw_measurement	Μέτρηση της μεταβλητής Headway
ME_AWS_fcw	Προειδοποίηση μπροστινής σύγκρουσης
ME_AWS_pcw	Προειδοποίηση σύγκρουσης πεζών
ME_AWS_ldw	Προειδοποίηση αναχώρησης από τη λωρίδα
ME_AWS_pedestrian_dz	Πεζός σε επικίνδυνη ζώνη
GPS_distances	Συνολική απόσταση του ταξιδιού
GPS_spd	Ταχύτητα
ME_TSR_tsr_1_speed	Εμφάνιση κωδικού πινακίδας 1 ταχύτητας
ME_AWS_time_indicator	Υποδεικνύει συνθήκες φωτισμού
DEM_evt_ha_lvl	Δύσκολα γεγονότα επιτάχυνσης
DEM_evt_hb_lvl	Δυνατό φρενάρισμα μεσαίου επιπέδου
DEM_evt_hc_lvl	Δύσκολες στροφές

# Περιγραφή - Επεξεργασία στοιχείων (2/2)



Μεταβλητές	Επεξήγηση
ME_AWS_time_indicator	Υποδεικνύει συνθήκες φωτισμού (ημέρα, σούρουπο, νύχτα)
ME_Car_wipers	Υποδεικνύει καιρικές συνθήκες (υαλοκαθαριστήρες on/off)
ME_Car_high_beam	Μεγάλη σκάλα
IBI_value	Χρονικό διάστημα μεταξύ διαδοχικών καρδιακών παλμών
ME_LDW_Map_type_L	Προειδοποίηση αναχώρησης από την αριστερή λωρίδα
ME_LDW_Map_type_R	Προειδοποίηση αναχώρησης από την δεξιά λωρίδα

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



# Καθορισμός Επιπέδων Ασφαλείας (1/2)

Καθορισμός επιπέδου 'Ζώνης Ανοχής Ασφαλείας' βάσει των μεταβλητών `iDreams_Headway_Map_level_i` και `iDreams_Speeding_Map_level_i`.

Για κάθε 30δλ ελέγχονται οι μεταβλητές και προκύπτει το επίπεδο 'Ζώνης Ανοχής Ασφαλείας'

- **Normal** όταν η μεταβλητή που επιλέγεται είναι επιπέδου παρέμβασης  $i = 1, 0, 1$
- **Dangerous** όταν η μεταβλητή που επιλέγεται είναι επιπέδου παρέμβασης  $i = 2$
- **Avoidable Accident** όταν η μεταβλητή που επιλέγεται είναι επιπέδου παρέμβασης  $i = 3$

# Καθορισμός Επιπέδων Ασφαλείας (2/2)



## ‘Επίπεδο ‘Ζώνης Ανοχής Ασφαλείας’

	Normal	Dangerous	Avoidable Accident
Βέλγιο (Headway)	68.472	17.165	4.787
Αγγλία (Headway)	53.942	13.960	3.365
Βέλγιο (Speeding)	67.411	4.354	17.555
Αγγλία ( Speeding)	54.092	2.291	8.118

Ο καθορισμός του επιπέδου θα γίνει με την μεταβλητή **Headway**.

Οι μεταβλητές που σχετίζονται με τον παράγοντα Headway εξαιρούνται από την διαδικασία ταξινόμησης. **Κίνδυνος μεροληψίας** στα μοντέλα (bias)

# Διαδικασία Ταξινόμησης (1/7)



- Επιλογή **μεταβλητών εισόδου** στα συνδυασμένα μοντέλα ταξινόμησης με βάση τη **συσχέτιση** και τη **σημαντικότητα**
- Διαχωρισμός δεδομένων σε δεδομένα **εκπαίδευσης** και **επιβεβαίωσης** σε ποσοστό 90% και 10 % αντίστοιχα
- Αντιμετώπιση της **ανισορροπίας των δεδομένων** με την τεχνική SMOTE
- Ανάπτυξη **τριών συνδυασμένων μοντέλων**:
  1. Συνδυασμός Βαθύ Νευρωνικού Δικτύου (**DNN**)-Τυχαία Δάση (**RF**)
  2. Συνδυασμός Επαναλαμβανόμενου Νευρωνικού Δικτύου (**RNN**) – Αλγόριθμου Προσαρμοστικής Ενδυνάμωσης (**AdaBoost**)
  3. Συνδυασμός των Συνελικτικών Νευρωνικών Δικτύων (**CNN**) και Μακράς Βραχυπρόθεσμης Μνήμης (**LSTM**)



## Επιλογή μεταβλητών – Συσχέτιση

- **Υψηλή** συσχέτιση μεταξύ διαφορετικών περιγραφικών στατιστικών της ίδιας μεταβλητής
- Η μεταβλητή **ME\_Car\_speed\_mean** και **GPS\_spd\_mean** παρουσιάζουν υψηλή συσχέτιση. Αύξηση ορίου ταχύτητας συνεπάγεται αύξηση ταχύτητας οχήματος
- **Χαμηλή** συσχέτιση μεταξύ των μεταβλητών που αφορούν στην εμφάνιση απότομων ή μη περιστατικών.

# Διαδικασία Ταξινόμησης (3/7)



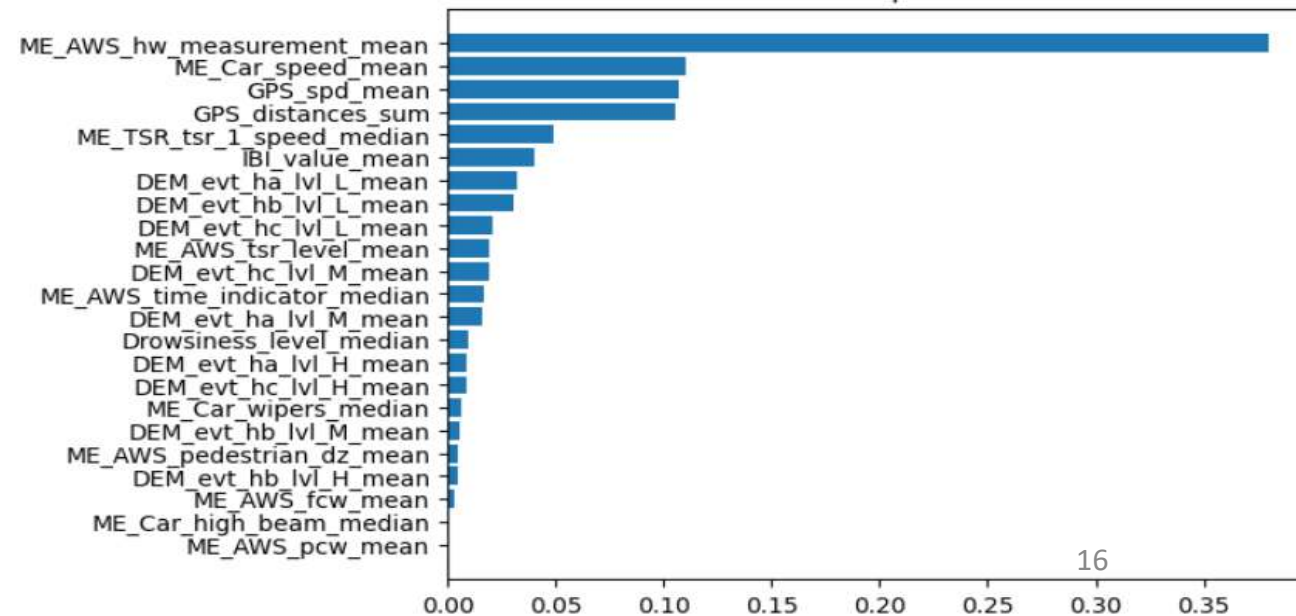
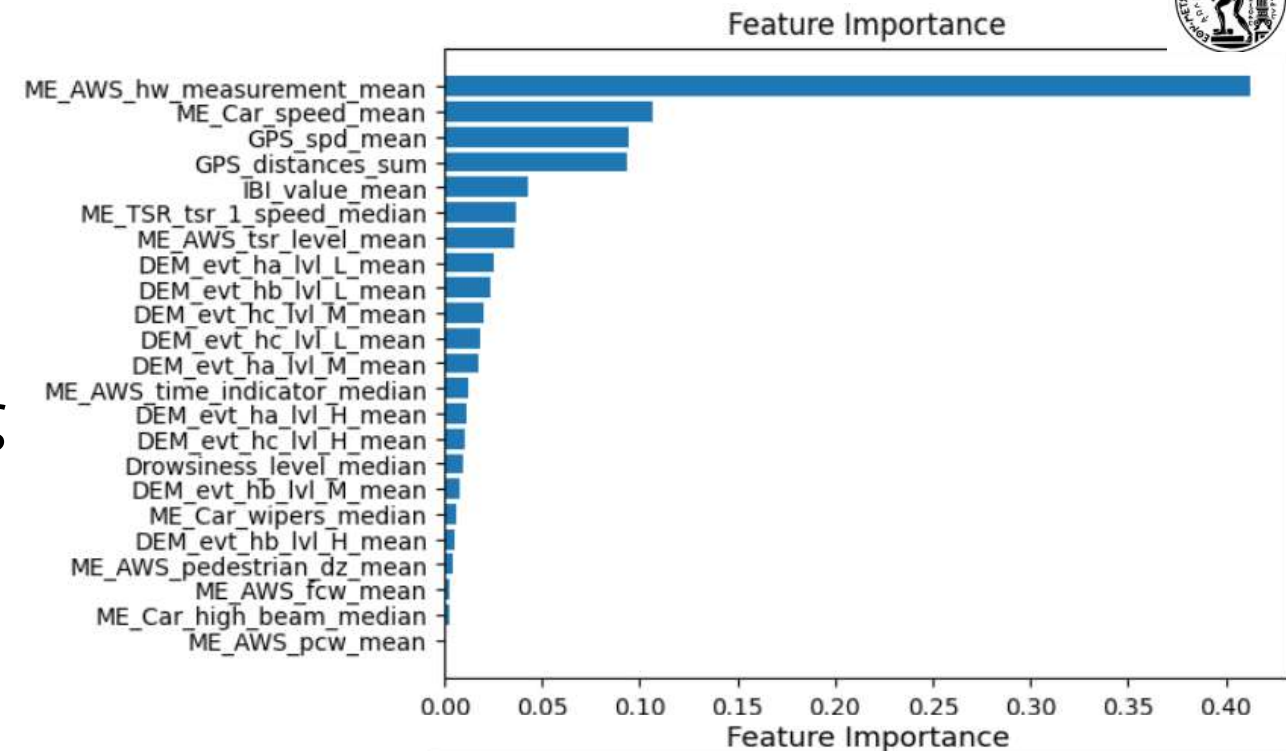
## Επιλογή μεταβλητών – Σημαντικότητα

Για το Βέλγιο και την Αγγλία:

- Μέσω του αλγορίθμου **Τυχαίων Δασών** αξιολογήθηκε η σημαντικότητα των μεταβλητών στην ταξινόμηση της 'Ζώνης Ανοχής Ασφαλείας'
- Η ME\_AWS\_hw\_measurement\_mean σχετίζεται με την μεταβλητή Headway, επομένως απορρίπτεται

## Μεταβλητές εισόδου:

1. GPS\_distances\_sum
2. GPS\_spd\_mean
3. DEM\_evt\_ha\_lvl\_L\_mean
4. DEM\_evt\_hb\_lvl\_L\_mean





# Διαδικασία Ταξινόμησης (4/7)

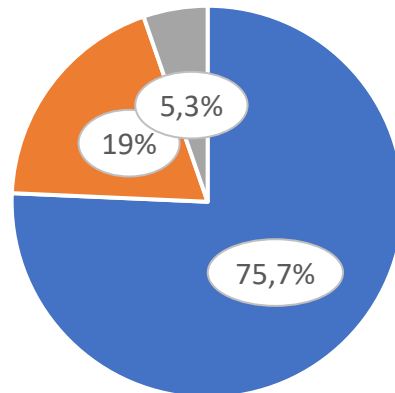


Με καθορισμό των επιπέδων προέκυψε το **πρόβλημα της ανισορροπίας** των δεδομένων στα τρία επίπεδα.

Εφαρμογή 'Τεχνικής συνθετικής υπερδειγματοληψίας μειοψηφίας' (**SMOTE**) με σκοπό:

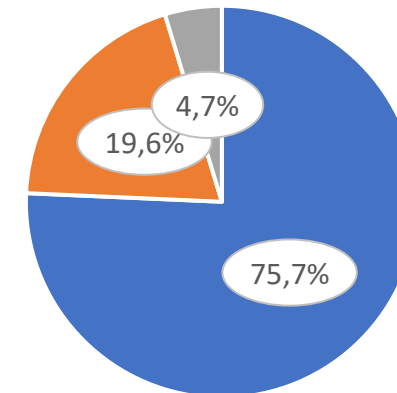
- Αντιμέτωπιση μεροληψίας μοντέλων ως προς την κυρίαρχη τάξη
- Αντιστάθμιση της άνισης κατανομής των τάξεων

Πριν την επαναδιαμόρφωση για τα δεδομένα του Βελγίου



■ Normal ■ Dangerous ■ Avoidable Accident

Πριν την επαναδιαμόρφωση για τα δεδομένα της Αγγλίας



■ Normal ■ Dangerous ■ Avoidable Accident



**Μετρικές  
Αξιολόγησης**



**Ακρίβεια**

Από τις προβλέψεις που πραγματοποιήθηκαν πόσες ήταν ορθές

**Μετρικές  
Αξιολόγησης**



## Μετρικές Αξιολόγησης

### Ακρίβεια

Από τις προβλέψεις που πραγματοποιήθηκαν πόσες ήταν ορθές

### Ανάκληση

Από τα πραγματικά θετικά στοιχεία πόσα κατάφερε να προβλέψει το μοντέλο

# Διαδικασία Ταξινόμησης (5/7)



## 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



## 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

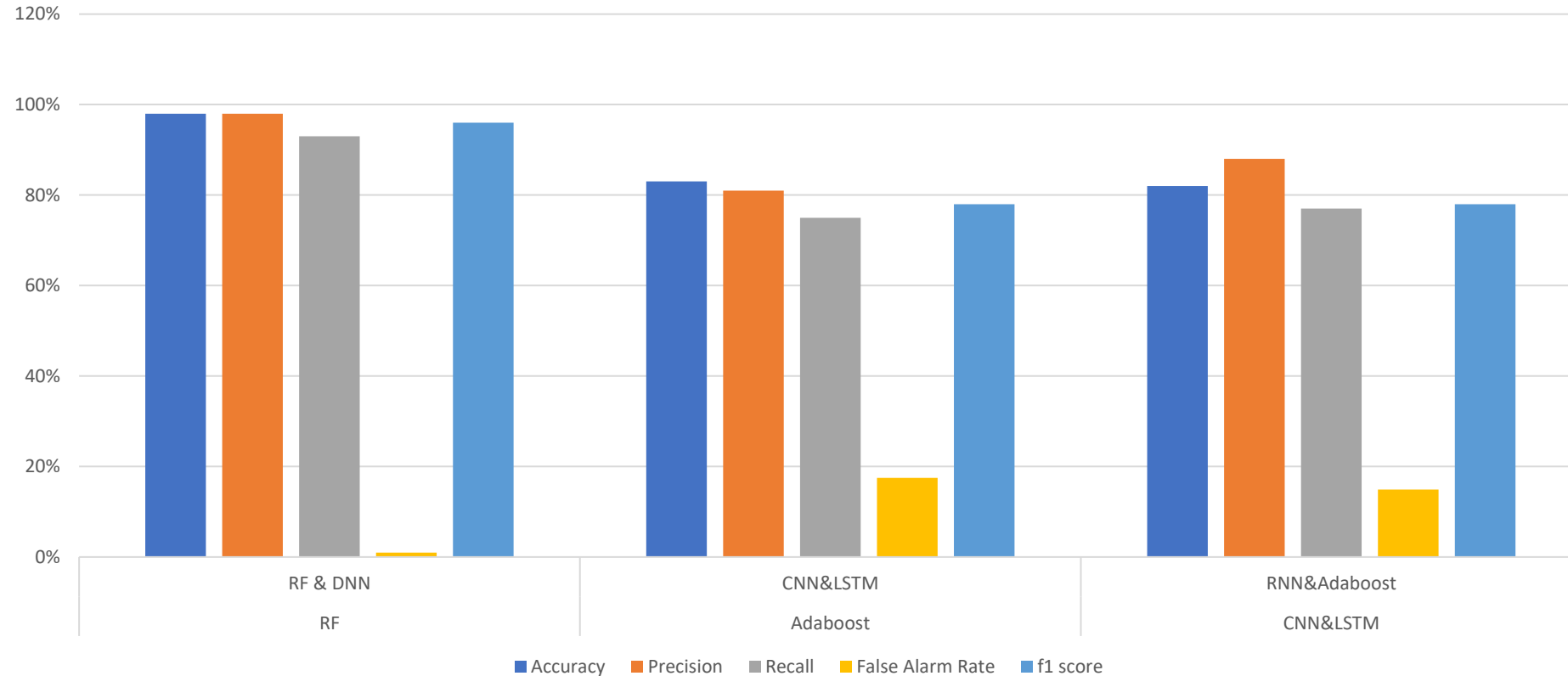
### False alarm rate

Τον αριθμό των ψευδών συναγερμών ανά συνολικό αριθμό προειδοποιήσεων

# Διαδικασία Ταξινόμησης (6/7)



Βέλγιο:

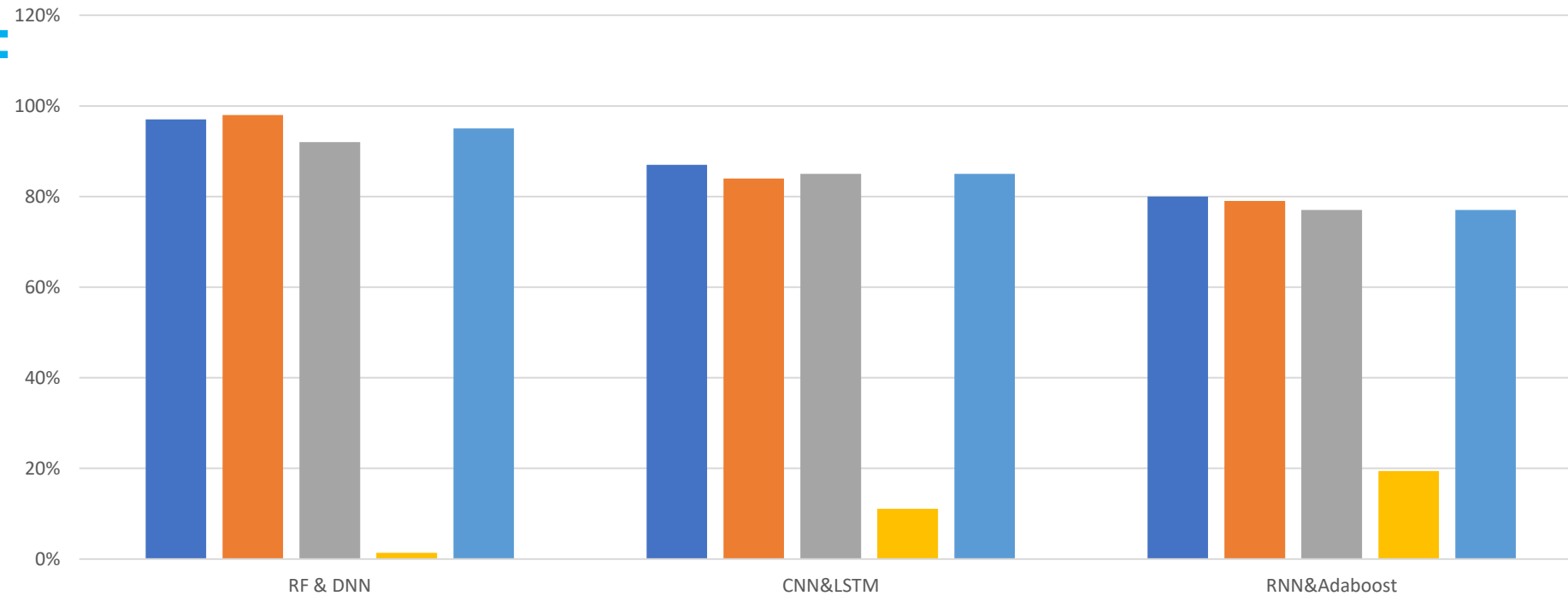


	Ορθότητα	Ακρίβεια	Ανάκληση	FPR	f1-score
RF & DNN	98%	98%	93%	0,96%	96%
RNN&Adaboost	82%	88%	77%	14,9%	78%
CNN&LSTM	83%	81%	75%	17,5%	78%

# Διαδικασία Ταξινόμησης (7/7)



Αγγλία:



■ Accuracy ■ Precision ■ Recall ■ False Alarm Rate ■ f1 score

	Ορθότητα	Ακρίβεια	Ανάκληση	FPR	f1-score
RF&DNN	97%	98%	92%	1,36%	95%
RNN&Adaboost	87%	84%	85%	11,11%	85%
CNN&LSTM	80%	79%	77%	19,4%	77%



# Αξιολόγηση επιρροής μεταβλητών



Για την ανάλυση της επιρροής των ανεξάρτητων μεταβλητών αναπτύσσεται ο **αλγόριθμος Lime**

**Σημαντικότητα** μεταβλητών για το Βέλγιο και την Αγγλία αντίστοιχα:

Μεταβλητή	Σημαντικότητα Μεταβλητής
GPS_spd_mean	0,14
GPS_distances_sum	0,31
DEM_evt_ha_lvl_L_mean	2,00
DEM_evt_hb_lvl_L_mean	0,56



**Μεγαλύτερη επιρροή** παρουσιάζουν οι μεταβλητές:  
DEM\_evt\_ha\_lvl\_L\_mean,  
DEM\_evt\_hb\_lvl\_L\_mean

Μεταβλητή	Σημαντικότητα Μεταβλητής
GPS_spd_mean	0,29
GPS_distances_sum	0,58
DEM_evt_ha_lvl_L_mean	2,00
DEM_evt_hb_lvl_L_mean	0,13



**Μεγαλύτερη επιρροή** παρουσιάζουν οι μεταβλητές:  
DEM\_evt\_ha\_lvl\_L\_mean,  
GPS\_distances\_sum

Η χρονική απόσταση του οχήματος από το προπορευόμενο όχημα είναι η σημαντικότερη μεταβλητή για την αναγνώριση της οδηγικής συμπεριφοράς.

Η συνολική διανυθείσα απόσταση, ταχύτητα του οχήματος, οι μεταβλητές εμφάνισης ή μη απότομης επιτάχυνσης και φρεναρίσματος κρίνονται σημαντικές για την αναγνώριση της οδηγικής συμπεριφοράς

Η τεχνική της 'Συνθετικής Μειονοτικής Υπερδειγματοληψίας' προσέφερε τα βέλτιστα αποτελέσματα στην διαδικασία αντιμετώπισης της άνισης κατανομής

Ο συνδυασμός 'Βαθύ Νευρωνικού Δικτύου (Deep Neural Network) - Τυχαίων Δασών (Random Forests)' σημείωσε τις υψηλότερες επιδόσεις

Οι μεταβλητές εμφάνισης ή μη απότομης επιτάχυνσης και φρεναρίσματος και η συνολική διανυθείσα απόσταση είχαν την μεγαλύτερη επιρροή στην τελική πρόβλεψη

Αξιοποίηση των συνδυαστικών **μοντέλων ταξινόμησης** για τον καθορισμό του επιπέδου ασφαλείας των οδηγών σε πραγματικές συνθήκες οδήγησης

Περαιτέρω **μελέτη των σημαντικότερων μεταβλητών** που επηρεάζουν την αναγνώριση της επικίνδυνης συμπεριφοράς

Ανάπτυξη **συστήματος αναγνώρισης** του επιπέδου 'Ζώνης Ανοχής Ασφαλείας' που βρίσκεται ο οδηγός σε πραγματικό χρόνο

Ανάπτυξη εφαρμογής σε **έξυπνα κινητά τηλέφωνα** για την αξιολόγηση της οδηγικής συμπεριφοράς

Εξοπλισμός οχημάτων με **συστήματα πρόβλεψης** ύπαρξης ατυχήματος και **προειδοποίησης** των οδηγών

Αξιοποίηση  
μεγαλύτερου όγκου  
δεδομένων

Ανάπτυξη  
εναλλακτικών  
μοντέλων τεχνικών  
εξέτασης  
σημαντικότητας  
των μεταβλητών

Διερεύνηση  
επιρροής  
πρόσθετων  
παραγόντων

Ανάπτυξη  
μοντέλων για  
ταξινόμηση  
ακολουθίας για την  
πρόβλεψη του  
επιπέδου 'Ζώνης  
Ανοχής Ασφαλείας'



**Ευχαριστώ για την προσοχή σας!** 29



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ  
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΕΠΙΚΙΝΔΥΝΗΣ  
ΣΥΜΠΕΡΙΦΟΡΑΣ ΤΩΝ ΟΔΗΓΩΝ ΜΕΣΩ ΤΗΣ  
ΑΞΙΟΠΟΙΗΣΗΣ ΔΕΔΟΜΕΝΩΝ ΟΔΗΓΗΣΗΣ ΣΕ  
ΠΡΑΓΜΑΤΙΚΕΣ ΣΥΝΘΗΚΕΣ

Ελένη Μαρία Θεοδωράκη

Επιβλέπων: Γιώργος Γιαννής, Καθηγητής ΕΜΠ

Αθήνα, Μάρτιος 2024

