



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΑΛΛΗΛΕΠΙΔΡΑΣΗΣ
ΑΥΤΟΝΟΜΩΝ ΟΧΗΜΑΤΩΝ ΚΑΙ ΠΕΖΩΝ ΜΕ ΧΡΗΣΗ
ΠΡΟΧΩΡΗΜΕΝΩΝ ΜΕΘΟΔΩΝ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ



Δημήτριος Γ. Τριάντης

Διπλωματική Εργασία

Επιβλέπουσα: Ελένη Βλαχογιάννη,
Αναπληρώτρια Καθηγήτρια Σχολής Πολιτικών Μηχανικών ΕΜΠ

Οκτώβριος 2022



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΜΕΤΑΦΟΡΩΝ ΚΑΙ ΣΥΓΚΟΙΝΩΝΙΑΚΗΣ ΥΠΟΔΟΜΗΣ

ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΑΛΛΗΛΕΠΙΔΡΑΣΗΣ
ΑΥΤΟΝΟΜΩΝ ΟΧΗΜΑΤΩΝ ΚΑΙ ΠΕΖΩΝ ΜΕ ΧΡΗΣΗ
ΠΡΟΧΩΡΗΜΕΝΩΝ ΜΕΘΟΔΩΝ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ

Δημήτριος Γ. Τριάντης

Διπλωματική Εργασία

Επιβλέπουσα: Ελένη Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια ΕΜΠ

Οκτώβριος 2022



NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF CIVIL ENGINEERING
DEPT. OF TRANSPORTATION PLANNING AND ENGINEERING

**MODELLING OF INTERACTIONS
BETWEEN AUTONOMOUS VEHICLES AND PEDESTRIANS
USING ADVANCED DEEP LEARNING METHODS**

Dimitrios G. Triantis

Diploma Thesis

Supervisor: Eleni Vlahogianni, Assistant Professor N.T.U.A.

October 2022

Ευχαριστίες

Με την παρούσα Διπλωματική Εργασία ολοκληρώνεται ο κύκλος των προπτυχιακών μου σπουδών στη Σχολή Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου.

Αρχικά θα ήθελα να ευχαριστήσω θερμά την κα Ελένη Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια ΕΜΠ για την ανάθεση του θέματος της παρούσας διπλωματικής εργασίας, την καθοδήγηση, το ενδιαφέρον και την υποστήριξη της σε όλη τη διάρκεια της εκπόνησης της.

Θα ήθελα επίσης να ευχαριστήσω τον Εμμανουήλ Καμπιτάκη, υποψήφιο διδάκτορα ΕΜΠ για την άριστη συνεργασία και βοήθεια που μου προσέφερε τους τελευταίους αυτούς μήνες. Η σημαντική συμβολή του στην ολοκλήρωση της διπλωματικής μου εργασίας δεν μπορεί παρά να αναγνωρισθεί.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου για την στήριξη που μου προσέφερε καθ' όλη τη διάρκεια των προπτυχιακών σπουδών μου.

Αθήνα, Οκτώβριος 2022
Δημήτριος Γ. Τριάντης

ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΑΛΛΗΛΕΠΙΔΡΑΣΗΣ ΑΥΤΟΝΟΜΩΝ ΟΧΗΜΑΤΩΝ ΚΑΙ ΠΕΖΩΝ ΜΕ ΧΡΗΣΗ ΠΡΟΧΩΡΗΜΕΝΩΝ ΜΕΘΟΔΩΝ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ

Δημήτριος Γ. Τριάντης

Επιβλέπουσα: Ελένη Βλαχογιάννη, Αναπληρώτρια Καθηγήτρια ΕΜΠ

ΣΥΝΟΨΗ

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη προτύπων βαθιάς μάθησης για την πρόβλεψη των τροχιών αυτόνομων οχημάτων και πεζών κατά την αλληλεπίδραση τους σε αστικά οδικά δίκτυα. Τα δεδομένα που χρησιμοποιήθηκαν συγκεντρώθηκαν σε κεντρικές αστικές περιοχές των Βρυξελλών και του Λέβεν και αφορούν σε αλληλεπιδράσεις δύο πρακτόρων (ενός πεζού και ενός αυτόνομου οχήματος). Τα εν λόγω δεδομένα περιείχαν στην ανεπεξέργαστη μορφή τους τις συντεταγμένες των πρακτόρων ανά 0.1 δευτερόλεπτα. Για την καλύτερη απόδοση των χρησιμοποιούμενων προτύπων, από τις συντεταγμένες εξάχθηκαν τα εξής χαρακτηριστικά: συνιστώσες ταχυτήτων, διανυσματικές ταχύτητες, διαφορά ταχυτήτων μεταξύ πρακτόρων, διαφορά γωνιών κίνησης μεταξύ πρακτόρων, ευκλείδεια απόσταση, κατακόρυφες και οριζόντιες αποστάσεις από τις οπτικές καθενός από τους δύο πράκτορες. Τα πρότυπα βαθιάς μάθησης που εφαρμόστηκαν είναι επτά σε αριθμό και όλα έχουν αρχιτεκτονική Encoder-Decoder και εμπεριέχουν κελιά LSTM. Και τα επτά πρότυπα εκπαιδεύτηκαν για τα ίδια επεισόδια (70% των δεδομένων) με και χωρίς τα χαρακτηριστικά που εξάχθηκαν, με σκοπό να διερευνηθεί ο βαθμός επιρροής αυτών στην ακρίβεια κάθε αρχιτεκτονικής. Χρησιμοποιώντας ως μετρικές αξιολόγησης των προτύπων το Μέσο Απόλυτο Σφάλμα (MAE) και το Μέσο Τετραγωνικό Σφάλμα (MSE) προέκυψε ότι ακριβέστερο εξ όλων των προτύπων είναι το Stacking Ensemble με τα χαρακτηριστικά, το οποίο συνδυάζει τα προηγούμενα έξι. Από τα έξι επιμέρους πρότυπα ακριβέστερο στις προβλέψεις προέκυψε το BiLSTM encoder-decoder με ένα επίπεδο CNN. Τέλος, όσον αφορά την περαιτέρω διερεύνηση του εν λόγω προβλήματος, έγινε μια σειρά από σχετικές προτάσεις.

Λέξεις-κλειδιά: Αυτόνομα οχήματα, πεζοί, αστικό περιβάλλον, πρόβλεψη τροχιών, βαθιά μάθηση, κωδικοποιητής-αποκωδικοποιητής, LSTM, αμφίδρομο LSTM, ανατροφοδοτούμενα νευρωνικά δίκτυα, συνελκτικά νευρωνικά δίκτυα, μηχανισμός έμφασης

MODELLING OF INTERACTIONS BETWEEN AUTONOMOUS VEHICLES AND PEDESTRIANS USING ADVANCED DEEP LEARNING METHODS

Dimitrios G. Triantis

Supervisor: Eleni Vlahogianni, Assistant Professor N.T.U.A.

ABSTRACT

The aim of this thesis is to develop deep learning models that predict the trajectories of autonomous vehicles and pedestrians during their interactions in urban environments. The data that were used for this analysis were gathered at dense urban areas in the cities of Brussels and Leuven, Belgium and pertain to interactions between two agents (one pedestrian and one autonomous vehicle per interaction). In their raw form, these data are comprised of the coordinates of each of the two agents per 0.1 seconds. In order to improve the performance of the used models, the following features were extracted from the said coordinates: both agents' speeds and velocities, velocity and angle difference between the two agents, euclidean distance, lateral and longitudinal distances from each of the two agents' perspectives. The deep learning models that were used are seven in number and all have an Encoder-Decoder architecture while at the same time including LSTM cells. All seven models were trained for the same interaction episodes (70% of the data) with and without the extracted features, in order to investigate the level of influence that they have in each of the architectures. Using the Mean Absolute Error (MAE) and Mean Square Error (MSE) as performance metrics it was concluded that the most precise model is the Stacking Ensemble one, which uses the other six models to achieve the optimal prediction accuracy. From these six models, the one with the greatest accuracy is the BiLSTM encoder-decoder one with an additional CNN layer. Lastly, when it comes to further investigation of such interactions, a few recommendations were made.

Keywords: autonomous vehicles, pedestrians, urban environment, trajectory prediction, deep learning, encoder-decoder, LSTM, bidirectional LSTM, recurrent neural networks, RNN, convolutional neural networks, CNN, attention mechanism

MODELLING OF INTERACTIONS BETWEEN AUTONOMOUS VEHICLES AND PEDESTRIANS USING ADVANCED DEEP LEARNING METHODS

Dimitrios G. Triantis

Supervisor: Eleni Vlahogianni, Assistant Professor N.T.U.A.

EXTENDED SUMMARY

Motivation

In the last fifty years or so, transportation with the use of motor vehicles has been an incremental of most people's everyday lives. Traffic safety is maybe the biggest challenge that is faced up to this point when it comes to such modes of transportation. With the three main factors leading to traffic accidents being road user behavior, road infrastructure and vehicle control, it is evident that technological advancements concerning those factors are in dire need. Indeed, among academics and professionals all throughout the world, continuous advancements are being made in order to better the situation. One of the most recent and promising developments in road safety are autonomous vehicles (AVs). Of course, road safety is not the only area where AVs are in a position to have major impacts, with decreased gas emissions, increased accessibility and economic growth being some of the most notable potential benefits of AV use.

Now when it comes to a road accident, among the previously mentioned factors, driver behavior is established by specialists as the most critical one. Through the adoption of AV use, driver behavior errors will almost certainly be greatly decreased, leading to a greatly increased level of road safety worldwide.

In any road environment, among the potential road users, pedestrians are established to be the most vulnerable. Therefore, during interactions between them and AVs, it is incredibly critical for the AVs to be in a position to adjust their movement accordingly. In this framework, it was decided that this thesis' aim is to develop different deep learning models in order to achieve the required level of trajectory prediction accuracy in AV-pedestrian interactions. Due to the dynamic relationship between the two agents in such interactions, it is decided that the used deep learning models will be designed to make predictions for both of their trajectories.

Dataset

In order to create models that will be able to precisely predict the trajectories of AVs and pedestrian during their interactions, the Euro-PVI dataset (Bhattacharyya et al., 2021) is used. This dataset contains data for 6163 different interactions, taking place in central urban areas of Brussels and Leuven, Belgium with high pedestrian density. Each of these interactions contains the coordinates of the two participating agents (an AV and a pedestrian) with a timestep of 0.1 seconds. When it comes to what constitutes an interaction in the setting of the Euro-PVI dataset, any event in which the movement of

an agent led to a change of speed for the other agent is deemed to be an interaction. In addition, in the said dataset the Wizard-of-Oz technique is used. The use of this technique means that the vehicle that is used as an AV, is actually driven by a person but is designed to be perceived as autonomous by pedestrians. This same vehicle was responsible for the gathering of the data, as it was equipped with a LIDAR system along with a positioning system and cameras.

Most interactions within the dataset consist of around 50 timesteps, rendering the creation of long-term predictions models (5-8 seconds) unfeasible. Therefore, it was decided to create short term models that require data from 30 timesteps (3 seconds) in the past in order to make predictions for 20 timesteps (2 seconds) in the future. This meant that all interactions with less than 50 timesteps had to be discarded. The remaining data consists of 1443 interactions, with an average number of $\mu = 69.93$ timesteps per interaction and standard deviation of $\sigma = 16.96$. In the following figure, the distribution of the number of timesteps per interaction is visualized through a bar chart.

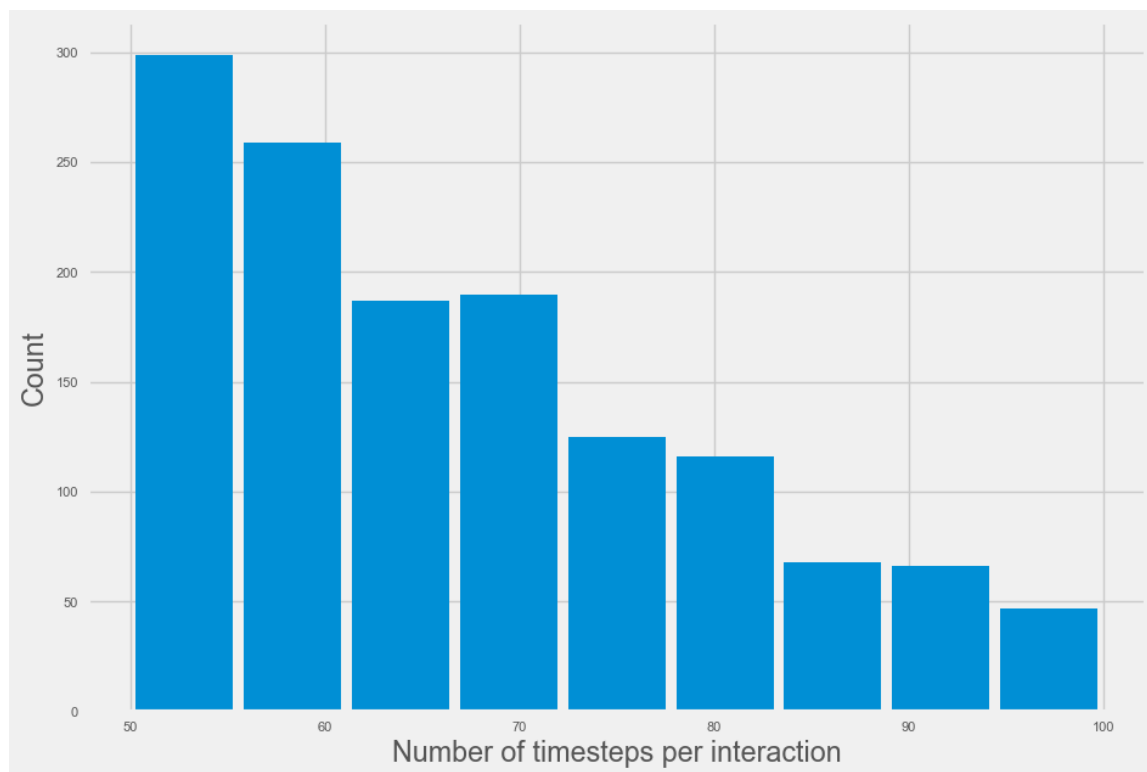


Figure 1: Distribution bar chart of the number of timesteps per interaction

In order to increase the accuracy of the models, it is decided that the extraction of a number of figures is a necessity, as the coordinates of the agents by themselves are not deemed to be adequate for the training process. Specifically, the features that are extracted from the data are:

- The AV's and pedestrian's speeds per timestep (in m/s) for the x and y axes. This feature is calculated using equation 1, where j is the axis (x or y) and x_i is the displacement:

$$v_i^j = \frac{\Delta x_i^j}{\Delta t} = \frac{(x_i - x_{i-1})^j}{0.1}, \text{ όπου } j: x \text{ ή } y \quad (1)$$

- Each agent's velocity per timestep (in m/s). This feature is calculated using equation 2:

$$u_i = \sqrt{v_x^2 + v_y^2} \quad (2)$$

- The angle difference between the two agents, per timestep (in rad). In order to calculate the angle difference, the angle of movement needs to be calculated using equation 3:

$$\theta_i = \arctan\left(\frac{v_i^y}{v_i^x}\right) \quad (3)$$

Then the angle difference is calculated as expected, using equation 4:

$$\Delta\theta_i = |\theta_i^{AV} - \theta_i^{ped}| \quad (4)$$

- Each agent's acceleration per timestep (in m/s^2), using the equation 5:

$$a_i = \frac{\Delta u_i}{\Delta t} = \frac{u_i - u_{i-1}}{0.1} \quad (5)$$

- Velocity difference between the two agents, per timestep (in m/s), using equation 6:

$$\Delta u_i = |u_i^{AV} - u_i^{ped}| \quad (6)$$

- Euclidean distance between the two agents, which are modelled as circles. The diameter of the AV model is equal to $3.5m$ while the diameter of the pedestrian model is equal to $0.7m$. The euclidean distance is calculated using the equations 7 and 8 (d_i is the distance between the circles of the two circles):

$$d_i = \sqrt{(x_i^{AV} - x_i^{ped})^2 + (y_i^{AV} - y_i^{ped})^2} \quad (7)$$

$$d_i^{EUC} = d_i - 1.75 - 0.35 \quad (8)$$

- The longitudinal and lateral distances between the two agents from each of their perspectives. In order to make such calculations, the absolute difference z_i of the slope φ_i of the line that cross both circles' centers ($-\pi \leq \varphi_i \leq \pi$) minus the angle of movement of each agent needs to be calculated and reduced to the first quadrilateral. Equations 9-13 are used to calculate these features. The calculated distances are visible in figure 2.

$$\varphi_i = \arctan \left(\frac{\Delta y_i}{\Delta x_i} \right) \quad (9)$$

$$z_i = |\varphi_i - \theta_i| \quad (10)$$

$$z'_i = \begin{cases} z_i, \gamma \text{ if } 0 \leq z_i < \pi/2 \\ \pi - z_i, \gamma \text{ if } \pi/2 \leq z_i < \pi \\ z_i - \pi, \gamma \text{ if } \pi \leq z_i < 3\pi/2 \\ 2\pi - z_i, \gamma \text{ if } 3\pi/2 \leq z_i < 2\pi \end{cases} \quad (11)$$

$$d_{lat,j}^i = \sin(z'_{i,j}) \cdot d_i - 1.75 - 0.35 \quad (12)$$

$$d_{long,j}^i = \cos(z'_{i,j}) \cdot d_i - 1.75 - 0.35 \quad (13)$$

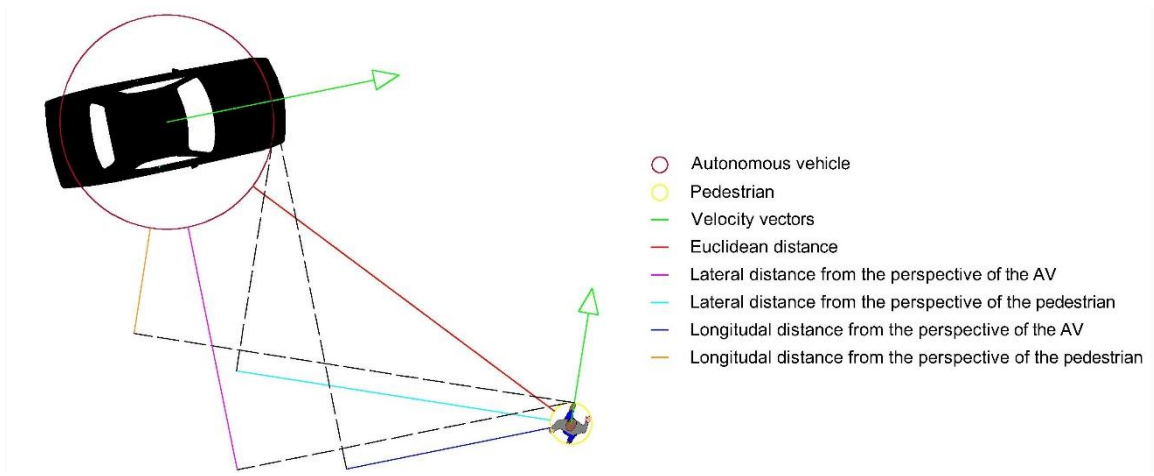


Figure 2: Extracted distances

Lastly after extracting the features, the noise of the data needed to be removed. Specifically, there were three major types of noise that needed to be tended to:

- The first category refers to a few instances where timesteps were shuffled. This type of noise was removed just by fixing the timestep sequences.
- The second category refers to cases where one agent (usually a pedestrian) does not move between two consecutive timesteps. Since the data were extracted using mostly visual methods, even if the said agent is not moving, their coordinates slightly change from one timestep to the next. So, in order to remove this type of noise three solutions were implemented:
 - In cases where the velocity speed between two timesteps was less than 0.3 m/s, the coordinates between the two timesteps were equated, and the speed and acceleration were nullified.
 - In the cases where the distance between two consecutive timesteps was lesser than 0.8m for the AV and 0.3m for the pedestrian, the coordinates between the two timesteps were equated.

- In cases where in more than 85% of the timesteps, the velocity of an agent was 0, then their velocity was nullified for the whole interaction.
- Lastly, in the very few cases where the velocity of one of the two agents was too high (over 72km/h for AVs and 20km/h for pedestrians) the whole interaction was discarded.

In general, the final form of the data seemed to be pretty logical based on the conditions that the data were collected. It is worth noting that for a decent number of timesteps within the data, one of the agents was not moving. That is most likely the result of them waiting for the other agent to pass them so they can get to their destination. In addition, the distances between the agents have a small but non-negligible frequency in values over 25m. The most possible explanation for such big distances is that they exist on very early or very late timesteps, where the interaction has just started or is about to end.

Basic theoretical background

In this thesis, seven different deep learning models are used. All of these models have an Encoder-Decoder architecture. A model with such an architecture consists of three separate parts: the encoder, the encoder vector and the decoder. The encoder is a neural network which consists of one or more memory cells. Each of these cells accepts a single element of the input sequence, collects information for that element and propagates it forward. The aim of the encoder is to transform the incoming series of data into the encoder vector. This vector then goes on to enter the decoder. The decoder is also a neural network which consists of a number of memory cells which use the encoder vector to produce a new series of output data, thus completing the prediction. In the following figure a typical encoder-decoder architecture is visualized.

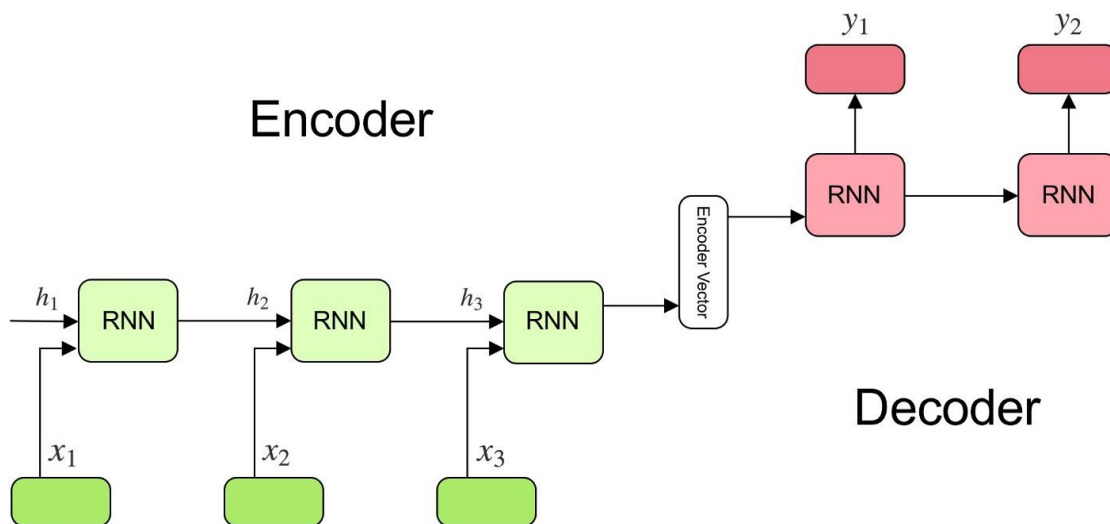


Figure 3: Visualization of the encoder-decoder architecture

[Source: towardsdatascience.com]

In addition, all the used models contain LSTM cells (Long Short-Term Memory [Hochreiter & Schmidhuber, 1997]). LSTM cells can be used much like normal memory cells-layers

with the difference that they yield much better results. The main idea around LSTM cells is that the neural network can learn what to store in the long-term state, what to throw away, and what to read from it. Bidirectional LSTM (BiLSTM) is based on the LSTM cell with the main difference that the analysis is done in both directions separately, i.e. from the first to the last timestep and vice versa. In that way, the model is able to learn the relationship between the values at each timestep better and therefore yield results of better quality and precision.

In two of the used models a technique called attention mechanism is used. This technique basically enables a neural network to reinforce certain parts of the input and silence some others in a way that attends to salient features of the output.

Lastly, two of the models include an additional convolutional layer [LeCun et al., 1998]. A convolutional layer uses the mathematical concept of convolution. Through the application of convolution with multiple filters to the input of the layer, the number of activations of the neurons on points where the filters match the input is increased, thus locating possible output characteristics.

Model methodology

As mentioned above, all seven models that are used have an encoder-decoder architecture, consisting of an LSTM or BiLSTM layer among others. When it comes to the extracted features, all seven of the models are tested with and without them, in order to assess their level of influence on the accuracy of the predictions.

In all of these models the Huber loss is used as the loss function during their training. This particular loss function is very popular in deep learning mainly due to its ability to combine much of the sensitivity of the mean-unbiased, minimum-variance estimator of the mean (using the quadratic loss function) and the robustness of the median-unbiased estimator (using the absolute value function).

When it comes to performance metrics, the mean absolute error (MAE) and the mean squared error (MSE) are used. Both metrics are indicators of how close the predictions of each model are to the ground truth. MAE is a more direct indicator as it basically shows the mean difference between the predictions and ground truths. MSE is less direct but has the great benefit of taking the weight of each prediction into consideration, through the squaring of each difference.

The training set, which is used for the training of each model, accounts to 70% of the data. The validation set, which is used for the validation of each model during its training, amounts to 15% of the data, as does the testing set.

All models are trained for 100 epochs with early stopping being in order when the loss function ceases to decrease. When it comes to the learning rate, which is an indicator of the rate of change of each model's parameters, a learning rate scheduler of exponential decay is used. Lastly, the Adam optimizer is used, as it considered to be one of the best if not the best optimizer available.

Then seven models that are trained and tested in this thesis are the following:

- LSTM encoder-decoder (LSTM_ed): This model consists of an encoder and a decoder. An LSTM cell exists within each of the two. This model is the simplest out of all seven, and basically acts as the foundation for the rest. Its architecture is visible in figure 4.

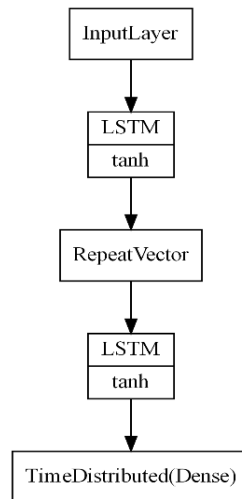


Figure 4: LSTM_ed architecture

- LSTM encoder-decoder with attention mechanism (LSTM_ed_attention): This model builds on the architecture of LSTM_ed by adding the attention mechanism. Its architecture is visible in figure 5.

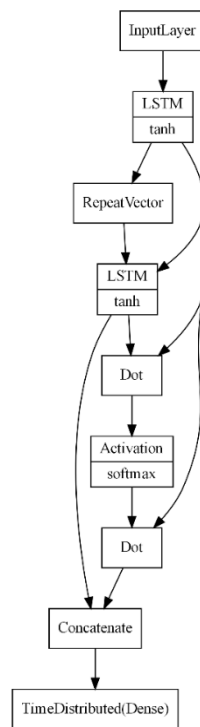


Figure 5: LSTM_ed_attention architecture

- LSTM encoder-decoder with an additional convolutional layer and attention mechanism (CNN LSTM ed attention): This model uses the same architecture as LSTM_ed_attention, while adding an one dimensional convolutional layer right before the encoder LSTM cell. Its architecture is visible in figure 6.

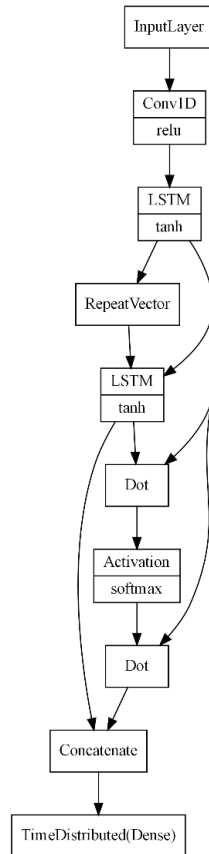


Figure 6: CNN_LSTM_ed_attention architecture

- BiLSTM encoder-decoder (BiLSTM ed): This model is the first one that uses the bidirectional RNN technique on LSTM. The two following models are based on its architecture, which is visible on figure 7.

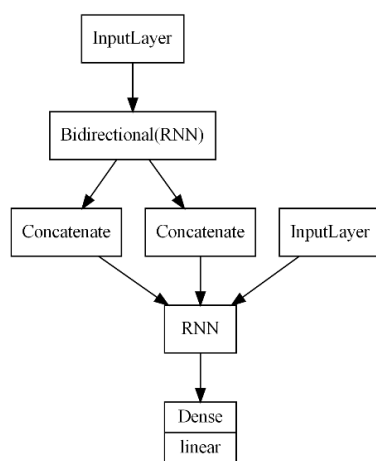


Figure 7: BiLSTM_ed architecture

- BiLSTM encoder-decoder with attention mechanism (BiLSTM ed attention): Much like LSTM_ed_attention with LSTM_ed, this model uses BiLSTM_ed's architecture and adds the attention mechanism to it. Its architecture is visible in figure 8.

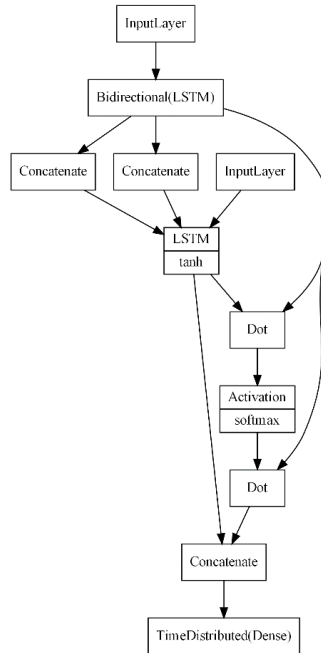


Figure 8: BiLSTM_ed_attention architecture

- BiLSTM encoder decoder with an additional one-dimensional convolutional layer (CNN BiLSTM ed): This particular model builds on BiLSTM_ed's architecture, adding an one dimensional convolutional layer to it. Its architecture is visible in figure 9.

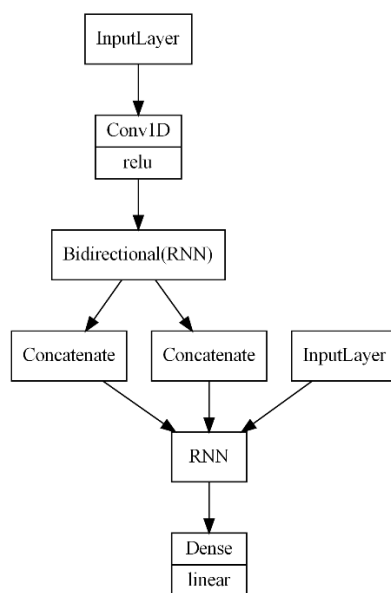


Figure 9: CNN_BiLSTM_ed architecture

- Stacking Ensemble:** This layer uses the Stacking Ensemble methodology in order to calibrate the six previous models in a way that keeps the good characteristics of each one of them, thus achieving the highest possible precision level. Its architecture is visible in figure 10.

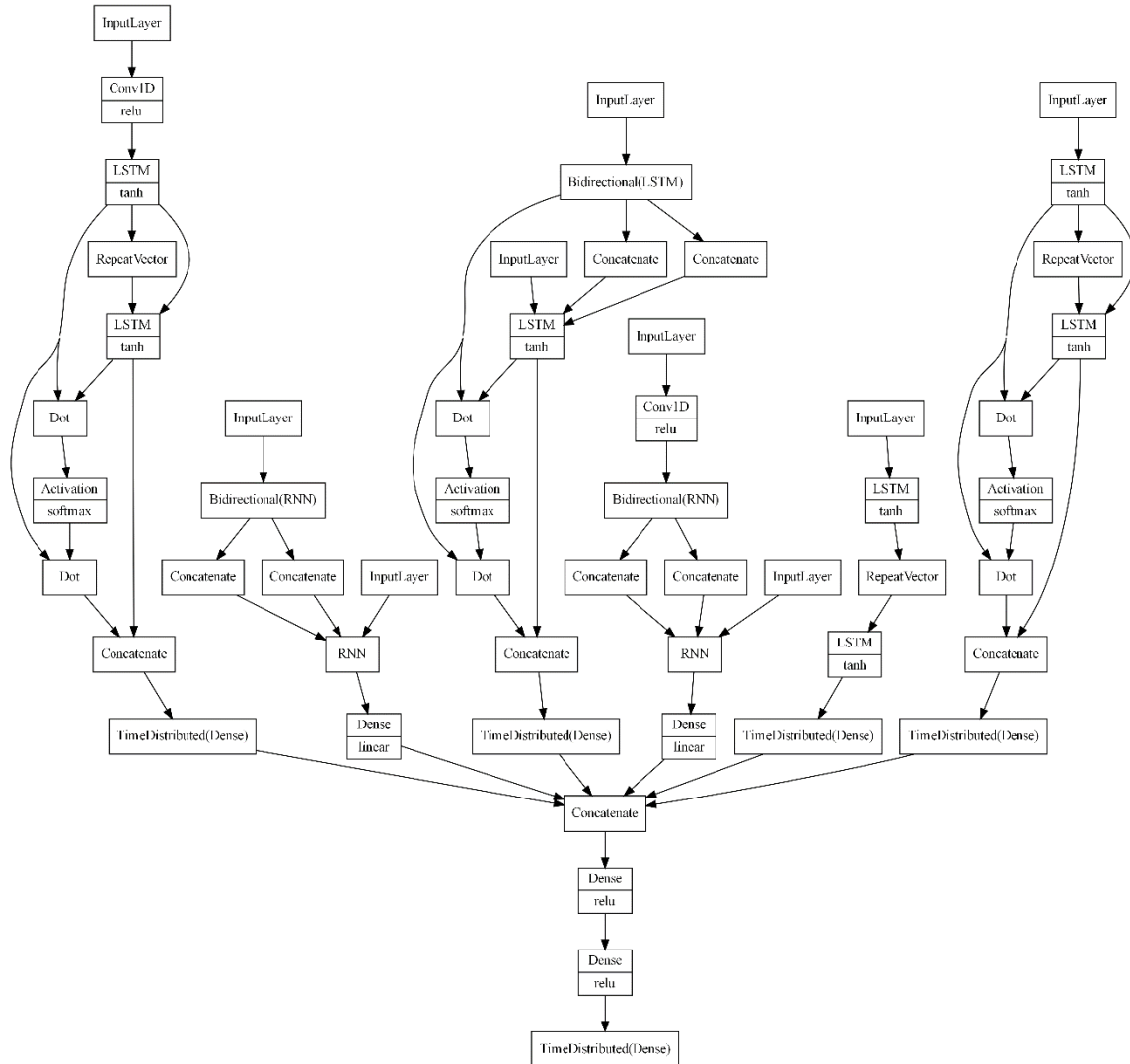


Figure 10: Stacking Ensemble architecture

Results and conclusions

Due to the fact that designing the models into making predictions for the coordinates of each agent seemed to make it much more complex than it needs to be, the predictions were made on the component speeds of each agent on the x and y axes, and were then reduced to trajectory predictions. The resulting mean metrics of each of the models on the test set is visible on the following table.

Table 1: Mean trajectory metrics for each model and agent

Models	+/- features	AV		Pedestrian	
		traj. MAE	traj. MSE	traj. MAE	traj. MSE
LSTM_ed	+	0.068	0.020	0.074	0.016
LSTM_ed_attention	+	0.064	0.019	0.070	0.015
CNN_LSTM_ed_attention	+	0.065	0.019	0.070	0.015
BiLSTM_ed	-	0.058	0.018	0.069	0.016
BiLSTM_ed_attention	+	0.058	0.016	0.060	0.011
CNN_BiLSTM_ed	+	0.057	0.017	0.060	0.010
Stacking Ensemble	+	0.051	0.015	0.056	0.009

As seen in table 1, by far the best out of the seven models is the Stacking Ensemble one (with features), as expected. Following it closely are BiLSTM_ed_attention and CNN_BiLSTM_ed, both producing almost identical metrics, thus showing that the effect of convolutional layers and the attention mechanism provide a similar level of improvement for the BiLSTM encoder-decoder architecture. After those three, the BiLSTM_ed without the features seems to be the fourth best, based on its performance metrics. The fact that for this particular architecture the features seemed to have negligible effects is an indication that the bidirectional analysis that takes place in a bidirectional RNN does not require additional features to produce precise results. LSTM_ed_attention and CNN_LSTM_ed_attention produce the same metrics and are equally ranked as the fifth best models. This equality in performance metrics shows that the additional convolutional layer does nothing to improve an LSTM encoder-decoder where the attention mechanism is already applied. The worst out of the seven models is the LSTM_ed, something that is expected due to its simple architecture as well as the fact that all the other models are structured based on it. At this point, it needs to be stressed that all models produce very close predictions with each other. That means that even the worse model, i.e. the LSTM_ed, is in a position to make quite precise trajectory predictions.

When it comes to the development of the trajectory metrics in the horizon of the 20 timesteps, it is quite similar for all seven models mostly since they all have similar architectures and philosophies. As seen in diagram 1 for the Stacking Ensemble model, the AV MAE's value seems to be much higher and changes much more rapidly for every timestep compared to the pedestrian MAE. That is probably explained from the fact that AVs are able to move much faster and with much higher accelerations than pedestrians. As seen in diagram 2 again for the Stacking Ensemble model, trajectory MSE's development in the horizon of two seconds is quite similar and somewhat linear for both agents. That is an indication of the fact that the models gradually lower their prediction accuracy level in an approximately linear rate. Again, the AVs seem to have a higher variance and greater values per timesteps, for the same reason as above.

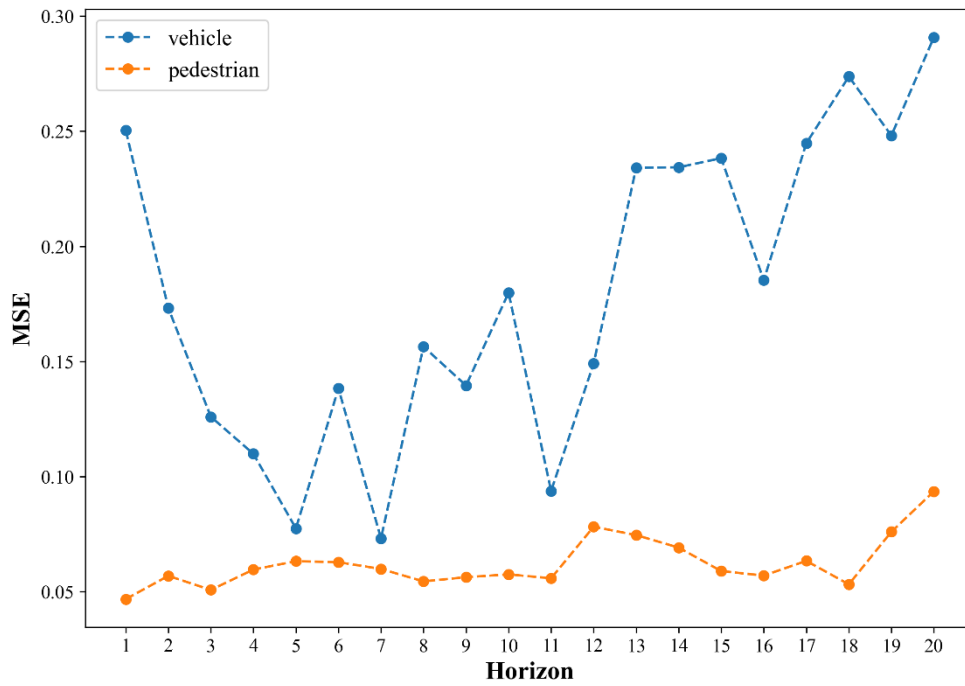


Diagram 1: MAE values in the horizon of 20 timesteps for the Stacking Ensemble model

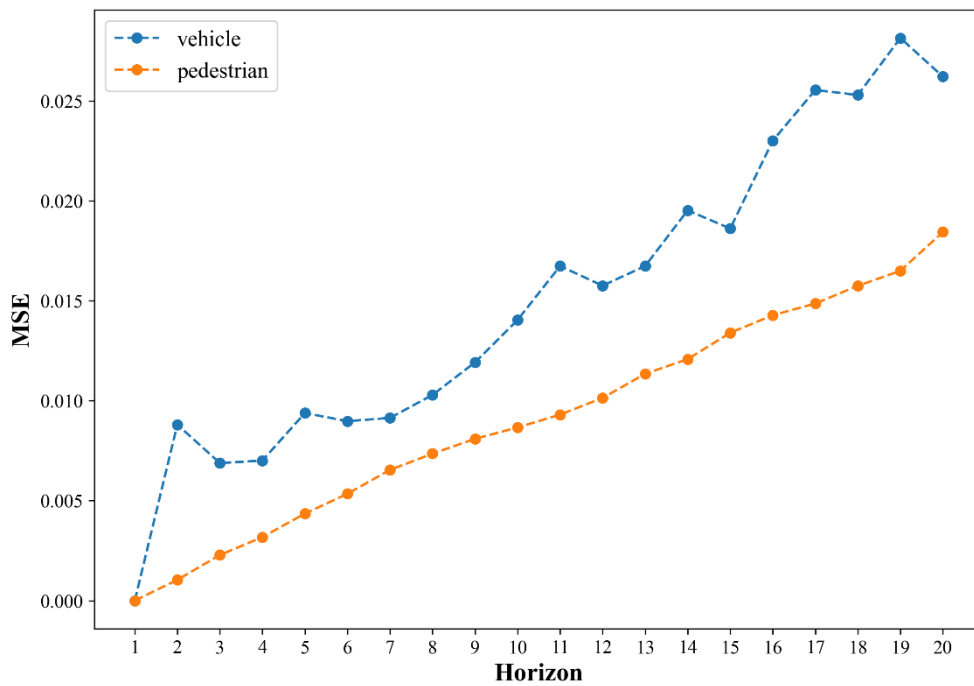


Diagram 2: MSE values in the horizon of 20 timesteps for the Stacking Ensemble model

In figure 11, the predicted and ground truth trajectories of both agents for four random interactions can be seen for the Stacking Ensemble model. The great accuracy that is achieved is evident.

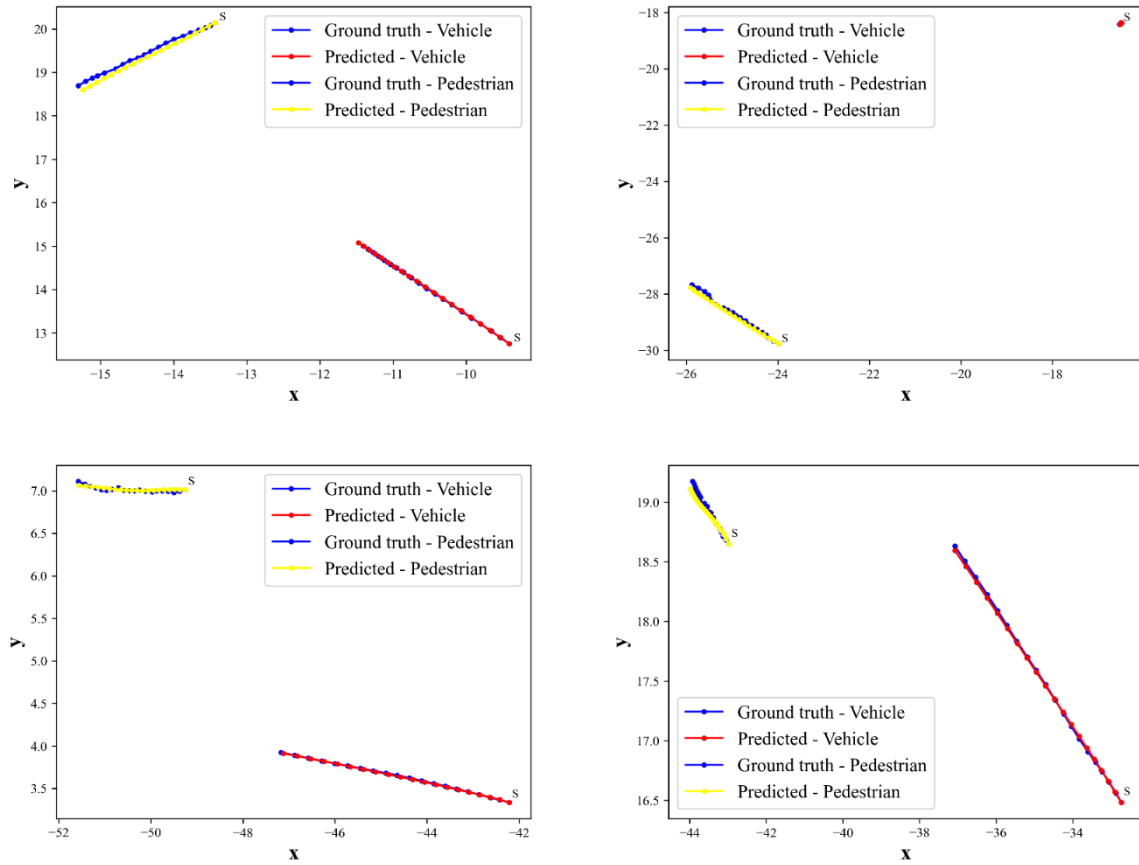


Figure 11: Ground truth VS predicted trajectories for four random interactions (Stacking Ensemble model)

ΠΕΡΙΕΧΟΜΕΝΑ

ΛΙΣΤΑ ΔΙΑΓΡΑΜΜΑΤΩΝ	xxvii
ΛΙΣΤΑ ΕΙΚΟΝΩΝ	xxx
ΛΙΣΤΑ ΠΙΝΑΚΩΝ.....	xxxι
Κεφάλαιο 1: Εισαγωγή.....	1
1.1 Γενικά	1
1.2 Αυτόνομα οχήματα και σκοπιμότητα αυτών.....	2
1.3 Σκοπός της Εργασίας.....	5
1.4 Διάρθρωση της Εργασίας.....	6
Κεφάλαιο 2: Βιβλιογραφική ανασκόπηση.....	9
2.1 Βαθιά Μάθηση στις Μεταφορές και την Κυκλοφοριακή Τεχνική.....	9
2.2 Πρότυπα Πρόβλεψης Κίνησης Αυτόνομων Οχημάτων	11
2.2.1 Κατηγορίες Δεδομένων.....	11
2.2.2 Μέθοδοι Ανάλυσης και Πρόβλεψης.....	12
2.3 Πρότυπα πρόβλεψης κίνησης πεζών και οχημάτων στηριζόμενα στη μάθηση	13
2.4 Συμπεράσματα βιβλιογραφικής ανασκόπησης.....	15
Κεφάλαιο 3: Θεωρητικό υπόβαθρο.....	17
3.1 Βαθιά Μάθηση.....	17
3.1.1 Νευρωνικά δίκτυα και βαθιά νευρωνικά δίκτυα	17
3.1.2 Μαθηματική έκφραση βαθιού νευρωνικού δικτύου	19
3.2 Ανατροφοδοτούμενα νευρωνικά δίκτυα - RNNs.....	20
3.2.1 Αρχιτεκτονική Encoder-Decoder.....	21
3.2.2 Το Κελί LSTM	23
3.2.3 Αμφίδρομο LSTM (BiLSTM).....	25
3.2.4 Μηχανισμός Έμφασης (Attention Mechanism).....	26
3.3 Συνελκτικά νευρωνικά δίκτυα – CNNs.....	27
3.3.1 CNN-RNN βαθύ νευρωνικό δίκτυο	28
3.4 Συνάρτηση απώλειας.....	29
3.5 Μετρικές αξιολόγησης.....	29
Κεφάλαιο 4: Μεθοδολογική προσέγγιση	31
4.1 Περιγραφή δεδομένων	31
4.2 Επεξεργασία δεδομένων	32
4.2.1 Εξαγωγή χαρακτηριστικών (feature extraction).....	33
4.2.2 Αφαίρεση σφαλμάτων στα δεδομένα	36

4.2.3	Περιγραφή δεδομένων	36
4.3	Μεθοδολογία ανάπτυξης και εκπαίδευσης προτύπων.....	40
4.3.1	Πρότυπο LSTM encoder-decoder (LSTM_ed)	41
4.3.3	Πρότυπο CNN-LSTM encoder-decoder με έμφαση (CNN_LSTM_ed_attention)	43
4.3.4	Πρότυπο BiLSTM encoder-decoder (BiLSTM_ed)	43
4.3.5	Πρότυπο BiLSTM encoder-decoder με έμφαση (BiLSTM_ed_attention).....	44
4.3.6	Πρότυπο CNN-BiLSTM encoder-decoder (CNN_BiLSTM_ed)	45
4.3.7	Stacking Ensemble	45
	Κεφάλαιο 5: Αποτελέσματα.....	47
5.1	Αποτελέσματα ανά πρότυπο	47
5.1.1	Αποτελέσματα LSTM_ed.....	47
5.1.2	Αποτελέσματα LSTM_ed_attention.....	52
5.1.3	Αποτελέσματα CNN_LSTM_ed_attention	56
5.1.4	Αποτελέσματα BiLSTM_ed.....	58
5.1.5	Αποτελέσματα BiLSTM_ed_attention.....	63
5.1.6	Αποτελέσματα CNN_BiLSTM_ed	67
5.1.7	Αποτελέσματα Stacking Ensemble.....	71
5.2	Συγκριτικά αποτελέσματα	75
	Κεφάλαιο 6: Συμπεράσματα.....	79
6.1	Βασικά συμπεράσματα.....	79
6.2	Προτάσεις για περαιτέρω έρευνα	80
	Αναφορές	83
	ΠΑΡΑΡΤΗΜΑ	87
	Μέρος 1: Πίνακες.....	87
	Μέρος 2: Διαγράμματα.....	105

ΛΙΣΤΑ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 2. 1: Ενδεικτικά ποσοστά κατανομής συγκοινωνιακών ερευνών με χρήση Βαθιάς Μάθησης ανά αντικείμενο	10
Διάγραμμα 2. 2: Ενδεικτική ποσοστά κατανομής συγκοινωνιακών ερευνών με βάση τις δημοφιλέστερες μεθόδους Βαθιάς Μάθησης.....	10
Διάγραμμα 4. 1: Ραβδόγραμμα αριθμού χρονικών βημάτων ανά αλληλεπίδραση.....	33
Διάγραμμα 4. 2: Ραβδόγραμμα κατανομής διανυσματικής ταχύτητας (A) αυτόνομου οχήματος και (B) πεζών.....	37
Διάγραμμα 4. 3: Ραβδόγραμμα κατανομής επιτάχυνσης (A) αυτόνομου οχήματος και (B) πεζών	38
Διάγραμμα 4. 4: Ραβδόγραμμα κατανομής της διαφοράς (A) γωνιών κίνησης και (B) ταχυτήτων μεταξύ αυτόνομου οχήματος και πεζού	39
Διάγραμμα 4. 5: Ραβδόγραμμα κατανομής ευκλείδειας απόστασης.....	40
Διάγραμμα 5. 1: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_wf.....	48
Διάγραμμα 5. 2: MAE θέσης στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_wf.....	49
Διάγραμμα 5. 3: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_wf.....	49
Διάγραμμα 5. 4: MSE θέσης στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_wf.....	50
Διάγραμμα 5. 5: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (LSTM_ed_wf)	51
Διάγραμμα 5. 6: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_attention_wf.....	52
Διάγραμμα 5. 7: MAE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_attention_wf.....	53
Διάγραμμα 5. 8: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_attention_wf.....	53
Διάγραμμα 5. 9: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_attention_wf.....	54
Διάγραμμα 5. 10: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (LSTM_ed_attention_wf)	55
Διάγραμμα 5. 11: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (CNN_LSTM_ed_attention_wf).....	57
Διάγραμμα 5. 12: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_nf.....	59
Διάγραμμα 5. 13: MAE θέσης στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_nf	60
Διάγραμμα 5. 14: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_nf.....	60
Διάγραμμα 5. 15: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_nf.....	61

Διάγραμμα 5. 16: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (BiLSTM_ed_nf).....	62
Διάγραμμα 5. 17: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_attention_wf.....	64
Διάγραμμα 5. 18: MAE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_attention_wf.....	64
Διάγραμμα 5. 19: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_attention_wf.....	65
Διάγραμμα 5. 20: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_attention_wf.....	65
Διάγραμμα 5. 21: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (BiLSTM_ed_attention_wf).....	66
Διάγραμμα 5. 22: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το CNN_BiLSTM_ed_wf	68
Διάγραμμα 5. 23: MAE θέσης στον ορίζοντα 20 χρονικών βημάτων για το CNN_BiLSTM_ed_wf	68
Διάγραμμα 5. 24: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το CNN_BiLSTM_ed_wf	69
Διάγραμμα 5. 25: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το CNN_BiLSTM_ed_wf	69
Διάγραμμα 5. 26: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (CNN_BiLSTM_ed_wf).....	70
Διάγραμμα 5. 27: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το Stacking Ensemble	72
Διάγραμμα 5. 28: MAE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το Stacking Ensemble	72
Διάγραμμα 5. 29: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το Stacking Ensemble	73
Διάγραμμα 5. 30: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το Stacking Ensemble	73
Διάγραμμα 5. 31: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (Stacking Ensemble).....	74
Διάγραμμα Π. 1: Εξέλιξη απώλειας Huber ανά εποχή για το LSTM_ed_wf.....	105
Διάγραμμα Π. 2: Εξέλιξη MAE ταχύτητας ανά εποχή για το LSTM_ed_wf	105
Διάγραμμα Π. 3: Εξέλιξη MSE ανά εποχή για το LSTM_ed_wf.....	106
Διάγραμμα Π. 4: Εξέλιξη απώλειας Huber ανά εποχή για το LSTM_ed_attention_wf. 106	
Διάγραμμα Π. 5: Εξέλιξη MSE ανά εποχή για το LSTM_ed_attention_wf	107
Διάγραμμα Π. 6: Εξέλιξη MSE ανά εποχή για το LSTM_ed_attention_wf	107
Διάγραμμα Π. 7: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το CNN_LSTM_ed_attention_wf	108
Διάγραμμα Π. 8: MAE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το CNN_LSTM_ed_attention_wf	108
Διάγραμμα Π. 9: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το CNN_LSTM_ed_attention_wf	109

Διάγραμμα Π. 10: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το CNN_LSTM_ed_attention_wf	109
Διάγραμμα Π. 11: Εξέλιξη απώλειας Huber ανά εποχή για το CNN_LSTM_ed_wf	110
Διάγραμμα Π. 12: Εξέλιξη MSE ανά εποχή για το CNN_LSTM_ed_wf	110
Διάγραμμα Π. 13: Εξέλιξη MSE ανά εποχή για το LSTM_ed_attention_wf	111
Διάγραμμα Π. 14: Εξέλιξη απώλειας Huber ανά εποχή για το BiLSTM_ed_nf	111
Διάγραμμα Π. 15: Εξέλιξη MSE ανά εποχή για το BiLSTM_ed_nf	112
Διάγραμμα Π. 16: Εξέλιξη MSE ανά εποχή για το BiLSTM_ed_nf	112
Διάγραμμα Π. 17: Εξέλιξη απώλειας Huber ανά εποχή για το BiLSTM_ed_attention_wf	113
Διάγραμμα Π. 18: Εξέλιξη MSE ανά εποχή για το BiLSTM_ed_attention_wf	113
Διάγραμμα Π. 19: Εξέλιξη MSE ανά εποχή για το BiLSTM_ed_attention_wf	114
Διάγραμμα Π. 20: Εξέλιξη απώλειας Huber ανά εποχή για το CNN_BiLSTM_ed_wf ..	114
Διάγραμμα Π. 21: Εξέλιξη MSE ανά εποχή για το CNN_BiLSTM_ed_wf	115
Διάγραμμα Π. 22: Εξέλιξη MSE ανά εποχή για το CNN_BiLSTM_ed_wf	115
Διάγραμμα Π. 23: Εξέλιξη απώλειας Huber ανά εποχή για το Stacking Ensemble.....	116
Διάγραμμα Π. 24: Εξέλιξη MSE ανά εποχή για το Stacking Ensemble	116
Διάγραμμα Π. 25: Εξέλιξη MSE ανά εποχή για το Stacking Ensemble	117

ΛΙΣΤΑ ΕΙΚΟΝΩΝ

Εικόνα 1. 1: Ποσοστά συμβολής των τριών παραγόντων στα οδικά ατυχήματα	2
Εικόνα 1. 2: Τα 6 επίπεδα αυτοματισμού	3
Εικόνα 1. 3: Οφέλη της αυτοματοποιημένης κινητικότητας	5
Εικόνα 1. 4: Διάρθρωση διπλωματικής εργασίας.....	7
Εικόνα 3. 1: Οι δύο επικρατούσες απόψεις για τη σχέση της Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης και της Βαθιάς Μάθησης	17
Εικόνα 3. 2: Αρχιτεκτονική νευρωνικού δικτύου με ένα κρυφό επίπεδο	18
Εικόνα 3. 3: Αρχιτεκτονική βαθιού νευρωνικού δικτύου	19
Εικόνα 3. 4: Ανατροφοδοτούμενος νευρώνας, ανηγμένος στον χρόνο	20
Εικόνα 3. 5: Επίπεδο ανατροφοδοτούμενων νευρώνων, ανηγμένο στον χρόνο.....	21
Εικόνα 3. 6: Αρχιτεκτονική ενός πρότυπου Encoder-Decoder	23
Εικόνα 3. 7: Αρχιτεκτονική κελιού LSTM	25
Εικόνα 3. 8: Αρχιτεκτονική πρότυπου ανατροφοδοτούμενου RNN	26
Εικόνα 3. 9: Οπτικοποίηση του μηχανισμού έμφασης-attention	27
Εικόνα 3. 10: Εφαρμογή απλής συνέλιξης σε μονοδιάστατη είσοδο	28
Εικόνα 3. 11: Οπτικοποίηση CNN-RNN νευρωνικού δικτύου	28
Εικόνα 4. 1: Παραδείγματα αλληλεπιδράσεων στο Euro-PVI dataset	32
Εικόνα 4. 2: Εξαγόμενες Αποστάσεις	35
Εικόνα 4. 8: Αρχιτεκτονική πρότυπου LSTM_ed.....	41
Εικόνα 4. 9: Αρχιτεκτονική πρότυπου LSTM_ed_attention.....	42
Εικόνα 4. 10: Αρχιτεκτονική πρότυπου CNN_LSTM_ed_attention	43
Εικόνα 4. 11: Αρχιτεκτονική πρότυπου BiLSTM_ed.....	44
Εικόνα 4. 12: Αρχιτεκτονική πρότυπου BiLSTM_ed_attention.....	44
Εικόνα 4. 13: Αρχιτεκτονική πρότυπου CNN_BiLSTM_ed	45
Εικόνα 4. 14: Αρχιτεκτονική πρότυπου Stacking Ensemble.....	46

ΛΙΣΤΑ ΠΙΝΑΚΩΝ

Πίνακας 2. 1: Συγκεντρωτική βιβλιογραφία στην προτυποποίηση πρόβλεψης της τροχιάς οχημάτων και πεζών	16
Πίνακας 5. 1: Μέσα MAE και MSE για το LSTM_ed.....	47
Πίνακας 5. 2: Μέσα MAE και MSE για το LSTM_ed_attention.....	52
Πίνακας 5. 3: Μέσα MAE και MSE για το CNN_LSTM_ed_attention	56
Πίνακας 5. 4: Μέσα MAE και MSE για το BiLSTM_ed.....	58
Πίνακας 5. 5: Μέσα MAE και MSE για το BiLSTM_ed_attention	63
Πίνακας 5. 6: Μέσα MAE και MSE για το CNN_BiLSTM_ed	67
Πίνακας 5. 7: Μέσα MAE και MSE για το CNN_BiLSTM_ed	71
Πίνακας 5. 8: Συγκριτικές μετρικές τροχιών για τους δύο πράκτορες	75
Πίνακας 5. 9: Βαθμολόγηση προτύπων	77
Πίνακας Π. 1: Σύνοψη πρότυπου LSTM_ed_wf.....	87
Πίνακας Π. 2: Σύνοψη πρότυπου LSTM_ed_attention_wf.....	87
Πίνακας Π. 3: Σύνοψη πρότυπου CNN_LSTM_ed_attention_wf	88
Πίνακας Π. 4: Σύνοψη πρότυπου BiLSTM_ed_wf.....	88
Πίνακας Π. 5: Σύνοψη πρότυπου BiLSTM_ed_attention_wf.....	89
Πίνακας Π. 6: Σύνοψη πρότυπου CNN_BiLSTM_ed_wf.....	89
Πίνακας Π. 7. 1: Σύνοψη πρότυπου Stacking Ensemble - Μέρος 1.....	90
Πίνακας Π. 8: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το LSTM_ed.....	95
Πίνακας Π. 9: Εξέλιξη μετρικών ανά πρότυπο και εποχή.....	96
Πίνακας Π. 10: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το LSTM_ed_attention.....	98
Πίνακας Π. 11: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το CNN_LSTM_ed_attention	99
Πίνακας Π. 12: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το BiLSTM_ed.....	100
Πίνακας Π. 13: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το BiLSTM_ed_attention.....	101
Πίνακας Π. 14: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το CNN_BiLSTM_ed.....	102
Πίνακας Π. 15: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το Stacking Ensemble.....	103
Πίνακας Π. 16: MAE και MSE ανά πράκτορα, κατηγορία και πρότυπο	104

ΚΕΦΑΛΑΙΟ 1

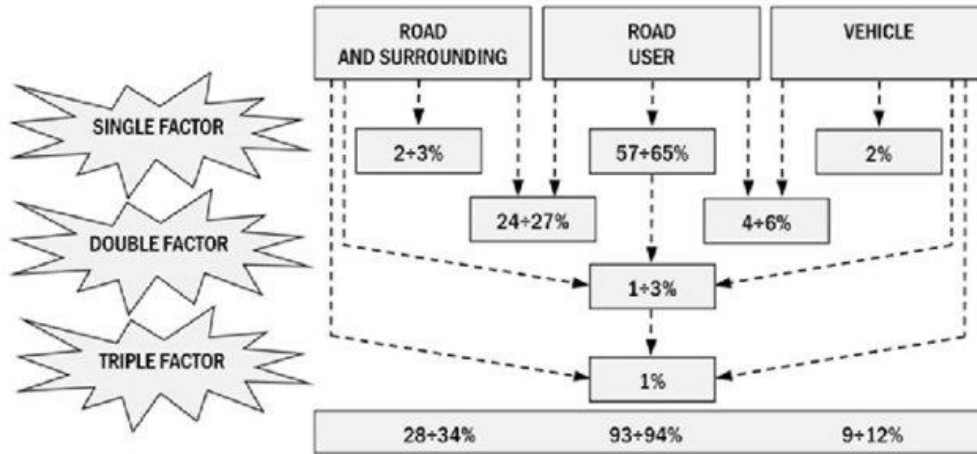
ΕΙΣΑΓΩΓΗ

1.1 Γενικά

Ένα αδιαμφισβήτητα σημαντικό κομμάτι της καθημερινότητας των ανθρώπων σχεδόν σε όλες της χώρες του κόσμου, εδώ και περίπου μισό αιώνα είναι η μετακίνηση τους με τη βοήθεια μηχανοκίνητων οχημάτων. Καθίσταται σαφές ότι για τη βέλτιστη και ασφαλή λειτουργία των μετακινήσεων των ανθρώπων με τη χρήση των οχημάτων οποιασδήποτε μορφής (ΜΜΜ, Ι.Χ., δίκυκλα) θα πρέπει όλοι οι παράγοντες που επηρεάζουν τις μετακινήσεις να αντιμετωπίζονται με την δέουσα προσοχή. Το γεγονός όμως ότι οι προαναφερθείσες μετακινήσεις, σε συνδυασμό με την εκάστοτε οδική υποδομή, αλλά και τους υπόλοιπους χρήστες αυτής (πεζοί, ποδηλάτες κλπ.), δομούν ένα περιβάλλον σημαντικών αλληλεπιδράσεων μεταξύ των χρηστών, οδηγεί υπό τις υπάρχουσες τεχνολογίες σε σημαντικό αριθμό οδικών ατυχημάτων αλλά και λοιπών προβλημάτων τα οποία έχουν σημαντικό αντίκτυπο στην υπάρχουσα κατάσταση.

Τα οδικά ατυχήματα που προκαλούνται κατά την αλληλεπίδραση των χρηστών μιας οδού οδηγούν σε μεγάλο αριθμό τραυματισμών και θανάτων παγκοσμίως. Αξίζει να σημειωθεί ότι κάθε έτος, περίπου 1,35 εκατομμύρια θάνατοι από οδικά ατυχήματα σημειώνονται παγκοσμίως, ένας αριθμός που κατ' εκτίμηση θα αντιστοιχούσε περίπου με 3.500 αεροπορικά ατυχήματα ανά έτος. Το μέγεθος των θυμάτων αλλά και τραυματιών υπό την υπάρχουσα κατάσταση είναι σημαντικό και είναι απαραίτητο να συνεχιστούν και ίσως και εντατικοποιηθούν όλες οι προσπάθειες που γίνονται επί του παρόντος για την βελτίωση του εν λόγω θέματος.

Εν γένει, είναι αποδεκτό από την επιστημονική κοινότητα ότι οι σημαντικότερες αιτίες των οδικών ατυχημάτων είναι η αυξημένη ταχύτητα, η οδήγηση υπό την επήρεια αλκοόλ, η μη χρήση ζώνης/κράνους και η απόσπαση της προσοχής του οδηγού. Οι αιτίες αυτές αφορούν τους οδηγούς οχημάτων, οι οποίοι αποτελούν τον ένα από τους τρεις παράγοντες που επηρεάζουν σημαντικά την συχνότητα και την σοβαρότητα των ατυχημάτων, με τους άλλους δύο να είναι η οδική υποδομή και τα οχήματα. Μια προσεγγιστική κατανομή των ατυχημάτων ανά παράγοντα είναι ορατή στην Εικόνα 1.1.



Εικόνα 1. 1: Ποσοστά συμβολής των τριών παραγόντων στα οδικά ατυχήματα [Πηγή: Rumar Kare, 1998]

Με βάση την παραπάνω Εικόνα καθίσταται σαφές ότι η οδηγική συμπεριφορά των χρηστών είναι ο σημαντικότερος παράγοντας, χωρίς όμως αυτό να σημαίνει ότι οι υπόλοιποι δύο θα πρέπει να ληφθούν υπόψιν σε μικρότερο βαθμό. Υπάρχουν πολυάριθμοι τρόποι που εφαρμόζονται επί του παρόντος για την βελτίωση της οδικής ασφάλειας όσον αφορά καθέναν από τους τρεις παράγοντες. Φερεπειύν, στην περίπτωση της οδηγικής συμπεριφοράς μπορεί να εφαρμοστεί αυστηρή αστυνόμευση ή αύξηση της δυσκολίας της λήψης διπλώματος αυτοκινήτου, ενώ στην περίπτωση του οδικού περιβάλλοντος μπορεί να γίνει εντοπισμός των επικίνδυνων θέσεων και εφαρμογή κατάλληλων παρεμβάσεων στην υποδομή. Μια από τις πιο σύγχρονες λύσεις που βρίσκεται υπό εξέλιξη είναι τα αυτόνομα οχήματα, τα οποία έχουν πολυάριθμα πλεονεκτήματα και μπορούν να βελτιώσουν σημαντικά όχι μόνο τον παράγοντα του οχήματος αλλά και την συμπεριφορά των οδηγών.

1.2 Αυτόνομα οχήματα και σκοπιμότητα αυτών

Τις τελευταίες δεκαετίες, διάφορες ιδιωτικές εταιρίες διεθνούς κύρους με συμμετοχή στον τομέα της αυτοκινητοβιομηχανίας, εκμεταλλεόμενες τις τεχνολογικές εξελίξεις σε διάφορους τομείς με κύριο από αυτούς την επιστήμη υπολογιστών, έχουν κάνει σημαντικές επενδύσεις με σκοπό τον σχεδιασμό αυτόνομων οχημάτων. Σκοπός των οχημάτων αυτών είναι η αντιμετώπιση και ελαχιστοποίηση κάποιων από τα σημαντικότερα προβλήματα και θέματα τα οποία προκύπτουν από την οδήγηση.

Αυτόνομα ονομάζονται τα οχήματα τα οποία έχουν εγκατεστημένο αυτοματοποιημένο σύστημα οδήγησης. Με τον όρο αυτό χαρακτηρίζονται τα συστήματα οδήγησης εντός του οχήματος που αναλαμβάνουν όλες τις λειτουργίες που είναι απαραίτητες για την οδήγηση του οχήματος σε πραγματικό χρόνο, χωρίς την ανάγκη παρέμβασης ανθρώπινου χειριστή. Με αυτόν τον τρόπο τα οχήματα αυτά είναι ικανά να ανιχνεύουν το οδικό περιβάλλον και να πλοηγούνται χωρίς ανθρώπινη βοήθεια.

Αξίζει να σημειωθεί ότι στην παρούσα φάση, καθώςον το εγχείρημα των αυτόνομων οχημάτων δεν βρίσκεται υπό καμία έννοια στα τελικά στάδια, υπάρχουν 5+1 επίπεδα συστημάτων οδήγησης στα οποία υπόκειται ο χαρακτηρισμός του αυτοματισμού:

- **Επίπεδο 0:** Ο έλεγχος του οχήματος υπόκειται πλήρως στους χειρισμούς του οδηγού. Δεν υπάρχει κανένας αυτοματισμός.
- **Επίπεδο 1:** Το σύστημα αυτοματισμού είναι σε θέση να επιταχύνει ή επιβραδύνει, λαμβάνοντας υπόψιν τις συνθήκες που επικρατούν στο περιβάλλον του. Κατά τα άλλα ο οδηγός είναι αυτός που ελέγχει την κίνηση του οχήματος.
- **Επίπεδο 2:** Σε προσθήκη του προηγούμενου επιπέδου, το όχημα έχει τη δυνατότητα να μεταβάλλει την ταχύτητα αυτόματα, μέσω της χρήσης περισσότερων του ενός συστημάτων υποβοήθησης. Και σε αυτή τη περίπτωση ο οδηγός διατηρεί τον έλεγχο του οχήματος.
- **Επίπεδο 3:** Το σύστημα αυτοματισμού είναι σε θέση να λειτουργεί πλήρως αυτοματοποιημένα, με την προσδοκία όμως ότι ο οδηγός θα μπορεί να επέμβει στην περίπτωση που χρειαστεί.
- **Επίπεδο 4:** Το όχημα σχεδιάζεται έτσι ώστε να μπορεί να λειτουργήσει πλήρως αυτόματα, χωρίς την βοήθεια οδηγού. Όμως, σε ειδικές περιπτώσεις (δυσμενείς καιρικές συνθήκες ή συνθήκες κυκλοφορίας) ο οδηγός θα πρέπει να ανακτήσει τον έλεγχο.
- **Επίπεδο 5:** Το όχημα λειτουργεί πλήρως αυτοματοποιημένα. Η μόνη ενέργεια που μπορεί να επιτελέσει ο χρήστης είναι ο προγραμματισμός της διαδρομής.



Εικόνα 1. 2: Τα 6 επίπεδα αυτοματισμού [Πηγή: geospatialworld.net]

Ίσως η σημαντικότερη προσφορά των αυτόνομων οχημάτων είναι η σημαντική δυνητική μείωση των ατυχημάτων αλλά και της σοβαρότητας αυτών. Αν αντιληφθεί κανείς ότι το συντριπτικό ποσοστό των ατυχημάτων οφείλεται στην οδική συμπεριφορά των χρηστών (σύμφωνα με τον οργανισμό NHTSA το ποσοστό ανέρχεται στο 95%), τότε μπορεί

εύκολα να γίνει κατανοητό το πόσο ευεργετική μπορεί να είναι η χρήση αυτόνομων οχημάτων όσον αφορά την οδική ασφάλεια.

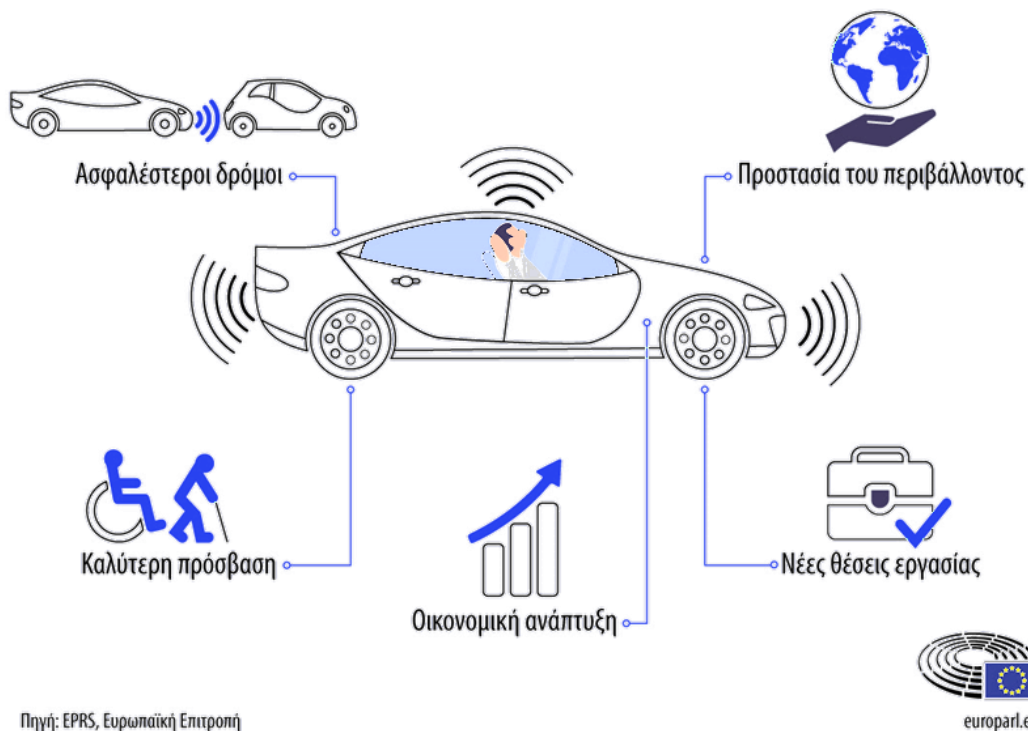
Στην σύγχρονη εποχή, οι τεχνολογίες σε όλους τους τομείς έχουν κάνει σημαντικά βήματα προς τα εμπρός, με σκοπό την βελτίωση της ζωής των ανθρώπων αλλά και την εξυπηρέτηση απαραίτητων λειτουργιών με όλο και πιο βελτιωμένες διαδικασίες. Στο ρεύμα αυτό εντάσσονται και τα αυτόνομα οχήματα, τα οποία συμβάλλουν στην άνετη κίνηση των χρηστών. Συγκεκριμένα, με την εφαρμογή της αυτόνομης οδήγησης δίνεται η δυνατότητα στους χρήστες, αναλόγως με το επίπεδο αυτοματισμού, να εκμεταλλευτούν τον χρόνο κίνησης τους πραγματοποιώντας άλλες δραστηριότητες. Επιπλέον, ο αυτοματισμός της οδήγησης μπορεί να έχει ιδιαιτέρως ευεργετικά αποτελέσματα και στην μετακίνηση ατόμων που προηγουμένως δεν ήταν σε θέση να κινηθούν μόνα τους, όπως για παράδειγμα κάποιες κατηγορίες ατόμων με ειδικές ανάγκες, ανήλικα ή υπέργηρα άτομα.

Τα αυτόνομα οχήματα έχουν επίσης την δυνατότητα μείωσης του κόστους μετακίνησης. Αυτό δυνητικά μπορεί να επιτευχθεί μέσω της μείωσης του αριθμού των ατυχημάτων, καθώς τότε το κόστος επισκευής των οχημάτων και το κόστος της ασφάλισης θα μειωθούν σε πολύ σημαντικό βαθμό. Επιπλέον, με την εδραίωση των αυτόνομων οχημάτων θα ευνοηθεί η πρακτική του ride-sharing, μειώνοντας έτσι σημαντικά τα έξοδα κίνησης των χρηστών.

Όσον αφορά τον οικονομικό τομέα, η εξέλιξη και εδραίωση των αυτόνομων οχημάτων μπορεί να συμβάλλει θετικά, μέσω της δημιουργίας μεγάλου αριθμού νέων θέσεων εργασίας. Αξίζει να σημειωθεί ότι το 2019 εκτιμήθηκε από την Ευρωπαϊκή Ένωση ότι μέχρι το 2025 οι αυτοκινητοβιομηχανίες που δραστηριοποιούνται στην Ευρώπη θα έχουν προσκομίσει κέρδη αξίας 620 δις ευρώ μέσω των αυτόνομων οχημάτων καθώς και 180 δις ευρώ για τον κλάδο της ηλεκτρονικής.

Η βελτιστοποιημένη κίνηση των αυτόνομων οχημάτων, ειδικότερα στην περίπτωση όπου αυτά είναι διασυνδεδεμένα μεταξύ τους, μπορεί να έχει σημαντικά θετικό αντίκτυπο σε περιβαλλοντικό και κυκλοφοριακό επίπεδο. Οι πρακτικές κίνησης που εφαρμόζονται στα αυτόνομα οχήματα, όπως για παράδειγμα η διατήρηση επαρκούς απόστασης, σε συνδυασμό με την μείωση του αριθμού των ατυχημάτων, μπορούν να μειώσουν σημαντικά την κυκλοφοριακή συμφόρηση. Επιπλέον, οι πρακτικές αυτές είναι σχεδιασμένες έτσι ώστε να έχουν σημαντικά μειωμένο αντίκτυπο όσον αφορά τις εκπομπές των οχημάτων και την κατανάλωση καυσίμων. Αν μάλιστα ληφθεί υπόψιν ότι στην παρούσα φάση ο μεγαλύτερος αριθμός αυτόνομων οχημάτων συνδυάζεται με ηλεκτρική ή υβριδική κίνηση, το οικολογικό αποτύπωμα αυτών είναι εντόνως βελτιωμένο σε σχέση με αυτό των συμβατικών οχημάτων.

Συνολικά, όπως καθίσταται σαφές από τα παραπάνω, η εδραίωση και εξέλιξη των αυτόνομων οχημάτων μπορεί να έχει ευεργετικά αποτελέσματα σε πολλούς διαφορετικούς τομείς, με σημαντικότερους από αυτούς την ασφάλεια, την άνεση των χρηστών, την οικονομική ανάπτυξη και την προστασία του περιβάλλοντος. Τα παραπάνω συνοψίζονται στην Εικόνα 1.3.



Εικόνα 1. 3: Οφέλη της αυτοματοποιημένης κινητικότητας
[Πηγή: Ευρωπαϊκό Κοινοβούλιο]

1.3 Σκοπός της Εργασίας

Όπως αναφέρθηκε παραπάνω, ένας από τους σημαντικότερους παράγοντες στους οποίους δρα θετικά η υιοθέτηση της αυτόνομης οδήγησης είναι η οδική ασφάλεια. Γενικότερα, ένα οδικό περιβάλλον συντίθεται από τους χρήστες της οδού, δηλαδή τα οχήματα, τους τυχόντες πεζούς και ποδηλάτες, και από την οδική υποδομή.

Είναι ευρέως γνωστό ότι τα ατυχήματα με τη μεγαλύτερη σοβαρότητα είναι αυτά στα οποία υπάρχει συμμετοχή ευάλωτων χρηστών. Οι κατ' εξοχήν ευάλωτοι χρήστες είναι οι πεζοί. Επομένως, είναι πολύ σημαντικό να διερευνηθεί η αλληλεπίδραση μεταξύ πεζών και αυτόνομων οχημάτων. Επιπλέον, η διερεύνηση της δυναμικής σχέσης μεταξύ πεζών και αυτόνομων οχημάτων είναι πολύ σημαντική προς τον σχεδιασμό των πολιτικών κίνησης των αυτόνομων οχημάτων, λαμβάνοντας υπόψιν την συμπεριφορά των πεζών όταν έρχονται σε επαφή με αυτά. Έτσι η ακριβής πρόβλεψη της κίνησης των αυτόνομων οχημάτων και των πεζών είναι απαραίτητη για τον ορθό σχεδιασμό των εν λόγω οχημάτων.

Σκοπός της παρούσας διπλωματικής είναι η ανάπτυξη προτύπων βαθιάς μάθησης για την πρόβλεψη των τροχιών αυτόνομων οχημάτων και πεζών κατά την αλληλεπίδραση τους σε πυκνές αστικές περιοχές. Στο πλαίσιο της παρούσας διπλωματικής εργασίας εφαρμόζονται επτά διαφορετικές μεθοδολογίες βαθιάς μάθησης, μέσω των οποίων καθίσταται δυνατή η πρόβλεψη της κίνησης των αυτόνομων οχημάτων και των πεζών. Τα προκύπτοντα πρότυπα χρησιμοποιούν και στις δύο περιπτώσεις ως δεδομένα τις χρονοσειρές των κινήσεων των δύο πρακτόρων (agents) για έναν πεπερασμένο αριθμό επεισοδίων. Έτσι γνωρίζοντας τις κινήσεις και κάποια άλλα χαρακτηριστικά των

πρακτόρων τα τελευταία 3 δευτερόλεπτα, έχουν την δυνατότητα να προβλέπουν την κίνηση αυτών για τα επόμενα 2 δευτερόλεπτα.

Και τα επτά πρότυπα της παρούσας εργασίας χρησιμοποιούν την μεθοδολογία Long Short-Term Memory – LSTM (Hochreiter & Schmidhuber, 1997). Η εν λόγω μεθοδολογία αποτελεί ένα τεχνητό νευρωνικό δίκτυο το οποίο έχει ευρέα χρήση στους κλάδους της τεχνητής νοημοσύνης και της βαθιάς μάθησης. Χρησιμοποιείται για πολυάριθμες εφαρμογές οι οποίες καλύπτουν ένα ευρύ διεπιστημονικό φάσμα. Ενδεικτικά αναφέρεται ότι το LSTM εκτός από τις εφαρμογές που μπορεί να έχει σε αναμενόμενους κλάδους όπως η ρομποτική και οι μεταφορές, χρησιμοποιείται έως και στον σχεδιασμό φαρμάκων και την σύνθεση μουσικής. Μια από τις πιο κύριες εφαρμογές της μεθόδου είναι η πρόβλεψη χρονοσειρών (time series forecasting), με βάση την οποία χρησιμοποιείται και στην παρούσα εργασία.

Όλα τα πρότυπα που εφαρμόζονται στην παρούσα εργασία είναι σχεδιασμένα με αρχιτεκτονική Encoder-Decoder (Κωδικοποιητής – Αποκωδικοποιητής) (Cho et al., 2014). Ο εν λόγω τύπος πρότυπου έχει χρησιμοποιηθεί σε πολυάριθμες εφαρμογές, με κάποιες από αυτές να αφορούν την επιστήμη της γλωσσολογίας, την επεξεργασία εικόνων. Μεταξύ άλλων, τα LSTM Encoder-Decoder εν γένει προσδίδουν πολύ ενθαρρυντικά αποτελέσματα στη πρόβλεψη χρονοσειρών. Για αυτόν τον λόγο επιλέγεται να χρησιμοποιηθεί στην παρούσα εργασία. Τα χρησιμοποιούμενα πρότυπα γενικότερα διαφοροποιούνται μεταξύ τους στο ότι χρησιμοποιούν διάφορα επιμέρους νευρωνικά δίκτυα και αρχιτεκτονικές με σκοπό να διερευνηθεί ποιο από αυτά προσδίδει τα πιο ακριβή αποτελέσματα.

1.4 Διάρθρωση της Εργασίας

Η παρούσα διπλωματική εργασία χωρίζεται σε πέντε κεφάλαια. Πρώτο από αυτό είναι το τρέχον, το οποίο αφορά το γενικότερο πλαίσιο της παρούσας διπλωματικής εργασίας, τον σκοπό αυτής και μια σύντομη αναφορά των εργαλείων που θα χρησιμοποιηθούν.

Το δεύτερο κεφάλαιο αφορά την βιβλιογραφική ανασκόπηση. Στο κεφάλαιο αυτό συνοψίζονται οι μεθοδολογίες οι οποίες έχουν χρησιμοποιηθεί για πρόβλεψη των τροχιών πεζών και οχημάτων.

Ακολουθεί το κεφάλαιο του θεωρητικού υποβάθρου. Στο εν λόγω κεφάλαιο γίνεται μια σύντομη περιγραφή των αλγορίθμων και των μαθηματικών εργαλείων που χρησιμοποιούνται στην παρούσα εργασία.

Το τέταρτο κεφάλαιο αφιερώνεται στην περιγραφή και επεξήγηση των δεδομένων που χρησιμοποιούνται στη παρούσα εργασία, καθώς και στην μεθοδολογική προσέγγιση που ακολουθείται για τον σχεδιασμό των προτύπων.

Επόμενο κεφάλαιο είναι αυτό της παρουσίασης και σχολιασμού των αποτελεσμάτων της ανάλυσης. Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα των προτύπων και εξάγονται τα βασικότερα συμπεράσματα όσον αφορά την συγκριτική ακρίβεια τους.

Στο έκτο και τελευταίο κεφάλαιο της παρούσας διπλωματικής εργασίας παρουσιάζονται τα συμπεράσματα που προέκυψαν από αυτή. Στο εν λόγω κεφάλαιο συμπεριλαμβάνονται και οι προτάσεις για περαιτέρω έρευνα. Τα κεφάλαια της παρούσας διπλωματικής εργασίας συνοψίζονται στην εικόνα 1.4.



Εικόνα 1. 4: Διάρθρωση διπλωματικής εργασίας

ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1 Βαθιά Μάθηση στις Μεταφορές και την Κυκλοφοριακή Τεχνική

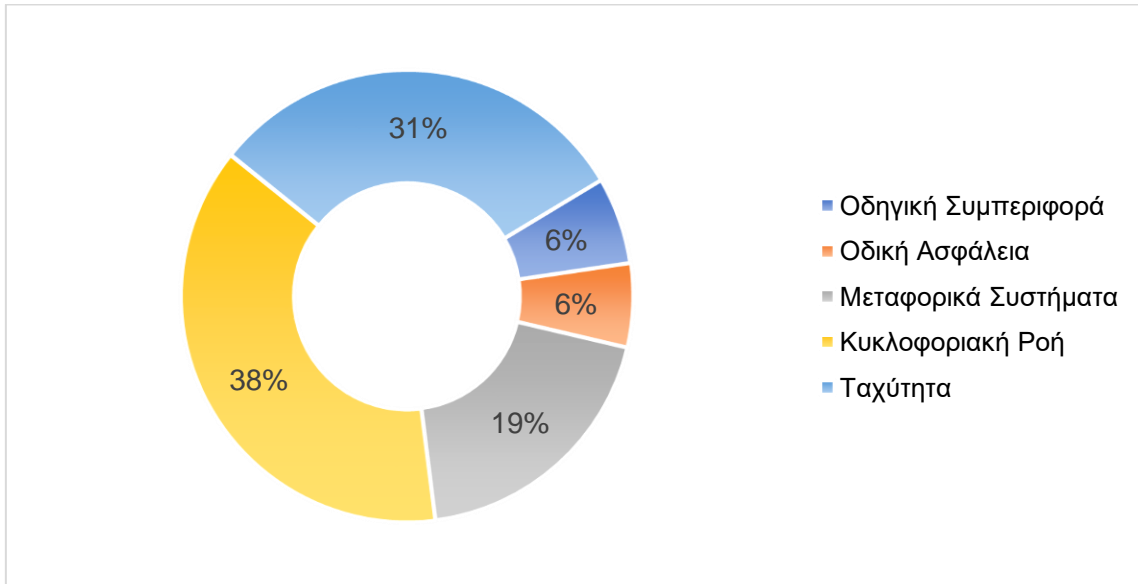
Η ραγδαία εξέλιξη της επιστήμης υπολογιστών τις τελευταίες δεκαετίες αποτελεί γεγονός. Ένα ιδιαιτέρως ευρύ φάσμα επιστημών και θεματικών πεδίων χρησιμοποιούν τις μεθοδολογίες που έχουν προκύψει από την εξέλιξη αυτή σε όλο και αυξανόμενο βαθμό. Κάτι τέτοιο συμβαίνει καθώς κλάδοι της επιστήμης υπολογιστών όπως η βαθιά μάθηση δίνουν την δυνατότητα σε επαγγελματίες και ερευνητές παγκοσμίως να παράγουν ποιοτικότερα και πιο ακριβή αποτελέσματα όποιος και αν είναι ο σκοπός τους. Όπως είναι φυσικό, ο συγκοινωνιακός σχεδιασμός ακολουθεί αυτό το ρεύμα, με την βαθιά μάθηση να χρησιμοποιείται ολοένα και περισσότερο.

Ήδη από την αρχή της προαναφερθείσας «επανάστασης», άρχισαν να εφαρμόζονται μεθοδολογίες τεχνητής νοημοσύνης σε συγκοινωνιακά θέματα. Ένα παράδειγμα είναι η χρήση κυκλοφοριακών δεδομένων από παγκόσμια συστήματα στιγματοθέτησης (GPS) με σκοπό την παροχή οδηγιών πλοήγησης στους χρήστες. Επί του παρόντος οι τεχνικές βαθιάς μάθησης εφαρμόζονται σε πολυάριθμα συγκοινωνιακά θέματα. Η πρόβλεψη των φόρτων, ταχυτήτων, διαδρομών και αφίξεων αποτελούν μόνο κάποιες από αυτές. Λόγω του γεγονότος ότι συστήματα όπως τα GPS και τα κλειστά κυκλώματα τηλεόρασης (CCTV) έχουν εδραιωθεί την τελευταία δεκαετία, έχει δημιουργηθεί μια σημαντική και συνεχώς αυξανόμενη ροή δεδομένων. Αυτή η ροή σε συνδυασμό με την συνεχή ανάπτυξη των τεχνικών βαθιάς μάθησης και την δυνατότητα τους να έχουν υπό συνθήκες καθολική εφαρμογή, έχουν οδηγήσει στην προτίμηση τους από τους συγκοινωνιολόγους.

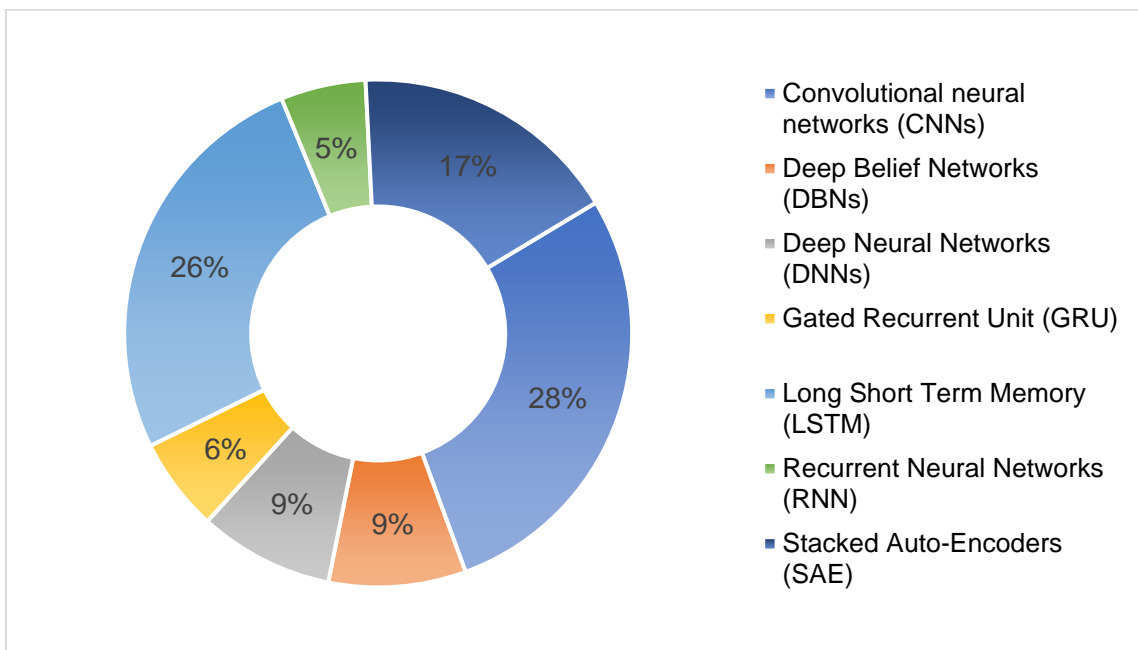
Όπως αναφέρεται και από τους Varghese et al. (2020), στην παρούσα φάση υπάρχουν τρεις σημαντικές προκλήσεις όσον αφορά την εφαρμογή της βαθιάς μάθησης σε συγκοινωνιακές έρευνες. Πρώτον, το γεγονός ότι σε αρκετές περιπτώσεις τα προκύπτοντα πρότυπα δεν συμβαδίζουν με τις εδραιωμένες στο επιστημονικό κοινό θεωρίες. Επομένως η βελτιστοποίηση της ακρίβειας των προτύπων δεν θα πρέπει να είναι ο μοναδικός σκοπός των μελετητών. Δεύτερον, για να είναι ποιοτικά τα παραγόμενα πρότυπα θα πρέπει τα δεδομένα από τα οποία προκύπτουν να είναι κατάλληλα και με αρκετά μεγάλο μέγεθος δείγματος. Μπορεί κάτι τέτοιο εν γένει να ακούγεται απλό, αλλά η συγκέντρωση μεγάλων και ποιοτικών δειγμάτων είναι σημαντικά κοστοβόρα. Τέλος, το γεγονός ότι ποτέ δεν είναι σίγουρο από πριν ότι μια τεχνική βαθιάς μάθησης θα παράγει τα επιθυμητά αποτελέσματα σημαίνει ότι οι μελετητές θα πρέπει να είναι προετοιμασμένοι να αντιμετωπίσουν μια τέτοια κατάσταση και να προσαρμόσουν τις χρησιμοποιούμενες τεχνικές με βάση αυτήν.

Η βαθιά μάθηση αποτελείται από πολλές διαφορετικές κατηγορίες μεθόδων, με κάποιες από τις σημαντικότερες να είναι τα Συνελικτικά Νευρωνικά Δίκτυα - Convolutional

Neural Networks (CNNs), τα Ανατροφοδοτούμενα Νευρωνικά Δίκτυα - Recurrent Neural Networks (RNNs) και μια υποκατηγορία αυτών, τα LSTMs. Στο Διαγράμματα 2.1 και 2.2 φαίνεται η κατανομή των μεθόδων βαθιάς μάθησης ανά συγκοινωνιακό αντικείμενο αλλά και η κατανομή ανά μέθοδο βαθιάς μάθησης. Αξίζει να σημειωθεί ότι οι έρευνες από τις οποίες προκύπτουν τα παρακάτω σχήματα πληρούν συγκεκριμένα κριτήρια που καθορίζονται από τους (Varghese et al., 2020).



Διάγραμμα 2. 1: Ενδεικτικά ποσοστά κατανομής συγκοινωνιακών ερευνών με χρήση Βαθιάς Μάθησης ανά αντικείμενο



Διάγραμμα 2. 2: Ενδεικτική ποσοστά κατανομής συγκοινωνιακών ερευνών με βάση τις δημοφιλέστερες μεθόδους Βαθιάς Μάθησης

2.2 Πρότυπα Πρόβλεψης Κίνησης Αυτόνομων Οχημάτων

Η δημιουργία έμπιστων προτύπων πρόβλεψης των κινήσεων των πεζών είναι απαραίτητη για τον ορθό σχεδιασμό των αυτόνομων οχημάτων. Γενικότερα οι αλληλεπίδραση των αυτόνομων οχημάτων με τους πεζούς αλλά και γενικότερα με το περιβάλλον τους αποτελεί μια πιθανοτική κατάσταση. Αυτό οφείλεται κυρίως στο γεγονός ότι η εν λόγω αλληλεπίδραση είναι έντονα απρόβλεπτη, καθώς εκτός από το ότι τα χαρακτηριστικά της αλληλεπίδρασης δεν εκτιμώνται πάντα ορθά και πλήρως, συχνά εκτελούνται αυθόρμητες ενέργειες από τους συμμετέχοντες και η οποιαδήποτε επιλογή καθενός από αυτούς επηρεάζει σημαντικά και τους υπόλοιπους [Karle et al., 2022].

2.2.1 Κατηγορίες Δεδομένων

Όσον αφορά στα δεδομένα με βάση τα οποία μπορεί να δημιουργηθεί ένα πρότυπο πρόβλεψης κίνησης για αυτόνομα οχήματα υπάρχουν εν γένει τρεις κύριες κατηγορίες πηγών:

Προσομοιώσεις:

Με τον όρο προσομοίωση νοούνται δύο διαφορετικές επιμέρους κατηγορίες. Συγκεκριμένα, υπάρχουν πρότυπα εικονικής πραγματικότητας στα οποία συμμετέχουν άνθρωποι – είτε ως πεζοί είτε ως οδηγοί – με βάση τις ενέργειες των οποίων μπορούν να συγκεντρωθούν δεδομένα. Οι Burns et al. (2019) δημιούργησαν 18 διαφορετικά σενάρια διέλευσης αυτόνομων οχημάτων τα οποία διαφοροποιήθηκαν με βάση παραμέτρους όπως η ταχύτητα των οχημάτων ή η θέση των πεζών. Στα σενάρια αυτά συμμετείχαν 13 διαφορετικοί άνθρωποι και με βάση τις ενέργειες αυτών δημιουργήθηκε μια βάση δεδομένων. Αντιστοίχως οι Mahadevan et al. (2019) δημιούργησαν ένα περιβάλλον εικονικής πραγματικότητας για πεζούς και αυτόνομα οχήματα στο οποίο παραμετροποιήθηκαν χαρακτηριστικά όπως το επίπεδο αυτοματισμού και η ρύθμιση της κυκλοφορίας.

Επιπλέον στην παρούσα κατηγορία υπάγονται και πρότυπα προσομοίωσης στα οποία γίνεται δόμηση του περιβάλλοντος που διερευνάται και πιθανοτική πρόβλεψη της κίνησης των συμμετεχόντων. Οι Gindele et al. (2010) δημιούργησαν ένα περιβάλλον (υπό τη μορφή Δυναμικού Μπεϋζιανού Δικτύου) που αντιλαμβάνεται την επικρατούσα κατάσταση και είναι σε θέση να καθορίσει τις μελλοντικές τροχιές των συμμετεχόντων. Οι Gill et al. (2019) δόμησαν μια προσομοίωση με βάση την οποία, γνωρίζοντας τις παρούσες καταστάσεις όλων των συμμετεχόντων πρακτόρων και αναλύοντας τους κινδύνους που προκύπτουν από τις προβλεπόμενες τροχιές, γίνεται πρόβλεψη της κίνησης των πρακτόρων.

Τεχνική Wizard-of-Oz:

Στην περίπτωση αυτή χρησιμοποιούνται πραγματικά δεδομένα αλληλεπίδρασης, με την προϋπόθεση ότι το όχημα που δρα ως αυτόνομο στην πραγματικότητα δεν είναι. Συγκεκριμένα το εν λόγω όχημα σχεδιάζεται με τέτοιο τρόπο ώστε να γίνεται αντιληπτό στους υπόλοιπους χρήστες ως αυτόνομο, ενώ στην πραγματικότητα είναι οδηγούμενο από κάποιον άνθρωπο. Ο οδηγός μπορεί να βρίσκεται εντός του οχήματος ή και να το οδηγεί από απόσταση. Χαρακτηριστικά παραδείγματα χρήσης αυτής της τεχνικής για

τον πρόβλεψη κίνησης αυτόνομων οχημάτων και πεζών είναι οι Lagstrom, & Lundgren (2015) και Palmeiro et al. (2018).

Τα δεδομένα που χρησιμοποιούνται και στην παρούσα εργασία είναι αυτής της μορφής [Bhattacharyya et al., 2021] και θα αναλυθούν σε επόμενο κεφάλαιο. Η εν λόγω κατηγορία δεδομένων επιλέχθηκε καθώς έχει ορισμένα πλεονεκτήματα, με το σημαντικότερο από αυτά να είναι η δυνατότητα ανάλυσης της κίνησης αυτόνομων οχημάτων εν την απουσία αυτών, είτε γιατί δεν ήταν δυνατό να χρησιμοποιηθούν είτε γιατί δεν υπάρχουν αυτόνομα οχήματα αποδεκτού επιπέδου ακόμα στην αγορά.

Πραγματικά Δεδομένα:

Στην συγκεκριμένη μέθοδο ένα αυτόνομο όχημα κινείται σε μια πραγματική θέση και με χρήση ειδικά τοποθετημένων και ρυθμισμένων εικονοληπτικών μηχανών καταγράφονται οι τροχιές όλων των συμμετεχόντων στο επεισόδιο. Η εν λόγω κατηγορία εν γένει προσδίδει τα ποιοτικότερα δεδομένα αλλά στην παρούσα φάση έχει κάποιους συγκεκριμένους περιορισμούς, με κάποιους από αυτούς να είναι:

- Μη ύπαρξη αυτόνομου οχήματος αποδεκτού επιπέδου αυτοματισμού
- Μειωμένη δυνατότητα ελέγχου του οδικού περιβάλλοντος και των σχετικών παραμέτρων
- Κίνδυνος μη ακριβών δεδομένων και ανάγκη για τακτική ρύθμιση των εικονοληπτικών μηχανών

2.2.2 Μέθοδοι Ανάλυσης και Πρόβλεψης

Σύμφωνα με τους Rudenko et al. (2020), υπάρχουν τρεις διαφορετικές κατηγορίες μεθοδολογιών οι οποίες μπορούν να εφαρμοστούν για την προτυποποίηση προβλέψεων:

Πρόβλεψη στηριζόμενη στην φυσική: Βασίζεται στην πιθανοτική πρόβλεψη της κίνησης των πεζών με χρήση κινηματικών προτύπων και Μπεϋζιανού φίλτρου. Στην παρούσα κατηγορία υπάγονται και μέθοδοι που κατασκευάζουν «χάρτες» κατάληψης του χώρου με βάση τους χωρικούς περιορισμούς που προκύπτουν από το σύνολο των πιθανών καταστάσεων. Είναι κατάλληλη προς χρήση στην περίπτωση βραχυπρόθεσμων προβλέψεων (έως δύο δευτερόλεπτα) και για τον έλεγχο αποφυγής ατυχημάτων.

Πρόβλεψη στηριζόμενη σε πρότυπα: Η συγκεκριμένη κατηγορία διαφοροποιείται από την προηγούμενη κυρίως στο ότι προστίθεται το βήμα της μηχανικής μάθησης. Χωρίζεται σε τρεις βασικές υποκατηγορίες:

- Ομαδοποίηση (Clustering): Σχηματισμός ομάδων δεδομένων τροχιών με σκοπό τον προσδιορισμό πρωτότυπων τροχιών.
- Κατηγοριοποίηση (Classification): Κατηγοριοποίηση τροχιών σε προδιακεκριμένες κλάσεις για τον προσδιορισμό μελλοντικών ελιγμών (Support Vector Machine, Hidden Markov Models, RNNs).

- Encoder – Decoder: Εξαγωγή πληροφοριών από τα δεδομένα εισαγωγής και παραγωγή μελλοντικών τροχιών με τη χρήση νευρωνικών δικτύων. Χρησιμοποιείται κυρίως για μακροπρόθεσμες προβλέψεις (5 – 8 δευτερόλεπτα).

Πρόβλεψη στηριζόμενη στον σχεδιασμό: Η συγκεκριμένη κατηγορία χρησιμοποιεί την θεωρία των μαρκοβιανών παιγνίων (Markov games) και χωρίζεται σε δύο υποκατηγορίες:

- Μάθηση της βέλτιστης συμπεριφοράς με χρήση Ανάστροφης Ενισχυτικής Μάθησης (Inverse Reinforcement Learning) ή άμεσης εκτίμησης της πολιτικής (Imitation Learning). Χρησιμοποιείται κυρίως για μακροπρόθεσμες προβλέψεις (5 – 8 δευτερόλεπτα).
- Σχεδιασμός υπό αβεβαιότητα: Ένταξη αβεβαιοτήτων μέσω της χρήσης Μαρκοβιανών διαδικασιών αποφάσεων (MDPs) ή τεχνικών θεωρίας παιγνίων.

2.3 Πρότυπα πρόβλεψης κίνησης πεζών και οχημάτων στηριζόμενα στη μάθηση

Και οι επτά μεθοδολογίες βαθιάς μάθησης που χρησιμοποιούνται στην παρούσα εργασία υπάγονται στην κατηγορία «Πρόβλεψη στηριζόμενη σε πρότυπα», όπως αυτή καθορίστηκε στην προηγούμενη παράγραφο. Επομένως ακολούθως αναφέρονται δημοσιεύσεις και έρευνες οι οποίες αφορούν το αντικείμενο της πρόβλεψης κίνησης πεζών και αυτόνομων οχημάτων χρησιμοποιώντας παρόμοιες μεθοδολογίες.

Αρχικά, έχουν αναπτυχθεί πρότυπα τα οποία προβλέπουν μόνον τις πορείες των πεζών, είτε αυτές αφορούν σε επεισόδια πεζών μεταξύ τους είτε μεταξύ οχημάτων και πεζών. Οι Nikhil & Morris (2018) κατασκεύασαν ένα μονοτροπικό πρότυπο πρόβλεψης της τροχιάς πεζού σε επεισόδια αυτού με άλλους πεζούς με χρήση CNN, χρησιμοποιώντας πραγματικά δεδομένα όπως αυτά του ETH dataset [Pellegrini et al., 2010]. Οι Styles, Ross & Sanchez (2019) κατασκεύασαν ένα μονοτροπικό πρότυπο βαθιάς μάθησης για την πρόβλεψη της τροχιάς πεζού σε επεισόδια με οχήματα, στο οποίο χρησιμοποιήθηκαν δεδομένα από κάμερες εγκατεστημένες σε οχήματα, τα οποία στην συνέχεια επεξεργάστηκαν μηχανικά ούτως ώστε να βελτιωθεί η ποιότητα τους. Σε αντίθεση με τα δύο προαναφερθέντα μονοτροπικά πρότυπα, οι Gupta et al. (2018), συνδυάζοντας τις μεθόδους LSTM και Generative Adversarial Networks (GAN), δημιούργησαν ένα πολυτροπικό πρότυπο πρόβλεψης της τροχιάς πολλών διαφορετικών πεζών, το οποίο εφαρμόστηκε και πάλι για επεισόδια της μορφής του ETH dataset. Οι Poibrenski et al. (2020) κατασκεύασαν με χρήση μιας δομής RNN encoder-decoder, ένα πολυτροπικό πρότυπο που εξάγει τις τρεις πιο πιθανές μελλοντικές τροχιές καθενός από τους πεζούς που συμμετέχουν σε ένα επεισόδιο το οποίο μπορεί να συμπεριλαμβάνει και οχήματα.

Αντίστοιχα, στο εγγύτερο παρελθόν έχουν κατασκευαστεί πολυάριθμα πρότυπα πρόβλεψης της κίνησης οχημάτων με κύριο σκοπό σχεδόν σε όλες τις περιπτώσεις τον σχεδιασμό των συστημάτων πλοήγησης των αυτόματων οχημάτων. Οι Hosseini et al.

(2020) χρησιμοποιούν την μέθοδο της ομαδοποίησης (clustering), για να κατασκευάσουν ένα μονοτροπικό πρότυπο πρόβλεψης της τροχιάς οχήματος κατά την αλληλεπίδραση του με άλλα οχήματα. Στην εν λόγω δημοσίευση επίσης αναπτύχθηκε μεθοδολογία βελτίωσης παραγωγής τεχνητών δεδομένων με χρήση GAN. Οι Dai, Li & Li (2019) δημιούργησαν ένα μονοτροπικό πρότυπο επεξεργασμένων LSTM για την πρόβλεψη της τροχιάς οχήματος. Οι Zhao et al. (2019) ανέπτυξαν με τη βοήθεια των Deep Convolutional Networks (DCN) ένα πολυτροπικό πρότυπο πρόβλεψης της κίνησης των οχημάτων κατά την αλληλεπίδραση τους με άλλους πράκτορες οποιασδήποτε μορφής. Ακόμη, οι Zhao et al. (2019) συνδύασαν τις μεθόδους LSTM, CNN και GAN για να κατασκευάσουν ένα πολυτροπικό πρότυπο πρόβλεψης της κίνησης πρακτόρων, είτε αυτοί είναι οχήματα είτε είναι πεζοί.

Η αλληλεπίδραση ενός οχήματος με πεζούς ή ποδηλάτες αποτελεί βέβαιο γεγονός σχεδόν σε κάθε διαδρομή του. Έτσι, με σκοπό τον ορθό σχεδιασμό των αυτόνομων οχημάτων είναι απαραίτητο να υπάρχουν αξιόπιστα πρότυπα τα οποία είναι σε θέση να προβλέπουν τις τροχιές όλων των πρακτόρων που συμμετέχουν σε ένα επεισόδιο. Πρόσφατα, με την εξέλιξη των μεθόδων βαθιάς μάθησης έχουν αναπτυχθεί πολλά πρότυπα που είναι σε θέση να κάνουν ακριβώς αυτό. Ένα παράδειγμα τέτοιου προτύπου είναι αυτό που ανέπτυξαν οι Lee et al. (2017). Το πρότυπο αυτό αποτελείται από επίπεδα RNN και έχει αρχιτεκτονική encoder-decoder και γίνεται χρήση πραγματικών δεδομένων αλληλεπίδρασης πεζών με αυτόνομα οχήματα. Οι Ivanovic & Pavone (2018) κατασκεύασαν ένα αντίστοιχο πρότυπο συνδυάζοντας τρεις εδραιωμένες μεθοδολογίες (Convolutional Variational Autoencoder (CVAE), LSTM, Dynamic Spatiotemporal Structure). Οι Salzmann et al. (2021) δημοσίευσαν μια βελτιωμένη έκδοση του παραπάνω προτύπου, εντάσσοντας τα Recurrent Graph NN.

Η μεθοδολογία Transformer Vaswani et al. (2017) έχει αναπτυχθεί αρκετά πρόσφατα αλλά παρόλα αυτά, λόγω των πολύ ποιοτικών αποτελεσμάτων που προσδίδει, ήδη εφαρμόζεται σε πολλούς διαφορετικούς τομείς με έναν από αυτούς να είναι οι προβλέψεις κίνησης πεζών και οχημάτων. Ένα παράδειγμα σχεδιασμού προτύπου με χρήση transformer είναι αυτό των Ngiam et al. (2021). Το εν λόγω πρότυπο αποτελείται από τρία στάδια: αρχικά γίνεται ενσωμάτωση των συμμετεχόντων πρακτόρων και του περιβάλλοντα χώρου σε έναν πολυδιάστατο χώρο, στην συνέχεια εφαρμόζεται ένα δίκτυο βασιζόμενο στον μηχανισμό έμφασης (attention-based) για την κωδικοποίηση (encoding) των αλληλεπιδράσεων και τέλος εφαρμόζεται ένα δίκτυο βασιζόμενο στην προσοχή (attention-based) για την αποκωδικοποίηση (decoding) αυτών. Ένα άλλο παράδειγμα προτύπου που χρησιμοποιεί την τεχνική Transformer είναι αυτό των Yuan et al. (2021). Στο πρότυπο αυτό χρησιμοποιείται ένα transformer που αναπτύσσεται από τους συγγραφείς και χρησιμοποιεί ένα νέο μηχανισμό προσοχής που είναι σε θέση να ξεχωρίζει τους πράκτορες μεταξύ τους και να τους αντιμετωπίζει με βάση την ταυτότητά τους. Οι Chen et al. (2021) δομούν ένα πρότυπο όπου χρησιμοποιεί ένα χωροχρονικό transformer για την περίπτωση χωροχρονικών αλληλεπιδράσεων και ένα χρονικό transformer για την περίπτωση χρονικών αλληλεπιδράσεων. Στον Πίνακα 2.1 συνοψίζονται όλες οι προαναφερθείσες δημοσιεύσεις.

2.4 Συμπεράσματα βιβλιογραφικής ανασκόπησης

Η βιβλιογραφική ανασκόπηση δείχνει ότι τα τελευταία χρόνια η εξέλιξη των μεθοδολογιών της επιστήμης υπολογιστών έχει ως αποτέλεσμα την ανάπτυξη προτύπων με χρήση πολλών διαφορετικών μεθοδολογιών. Είναι ξεκάθαρο ότι τα νευρωνικά δίκτυα έχουν κεντρικό ρόλο στα αναπτυσσόμενα πρότυπα, με μεθοδολογίες όπως τα ανατροφοδοτούμενα νευρωνικά δίκτυα (RNN) και το LSTM να εφαρμόζονται κατά κόρον.

Τα πρότυπα πρόβλεψης κίνησης πεζών και οχημάτων στην πλειοψηφία τους αναπτύσσονται με απώτερο σκοπό τον ορθό σχεδιασμό των συστημάτων πλοήγησης αυτόνομων οχημάτων. Για αυτό το σκοπό έχουν αναπτυχθεί πρότυπα για την πρόβλεψη κίνησης καθενός από τους πιθανούς συμμετέχοντες σε μια αλληλεπίδραση αυτόνομου οχήματος (πεζοί, οχήματα, ποδηλάτες). Η πλειοψηφία των προτύπων που έχουν αναπτυχθεί αφορά σε πρόβλεψη για μόνο μια κατηγορία πρακτόρων, χωρίς αυτό να σημαίνει ότι δεν έχουν αναπτυχθεί και πρότυπα πολλαπλών πρακτόρων.

Το Transformer αποτελεί μια πρόσφατα ανεπτυγμένη μεθοδολογία (2017) με πολύ ενθαρρυντικά αποτελέσματα όσον αφορά την πρόβλεψη τροχιών. Η εφαρμογή της εν λόγω μεθοδολογίας στο αντικείμενο των αυτόνομων οχημάτων έχει ξεκινήσει πολύ πρόσφατα -όπως φαίνεται και από τα έτη έκδοσης στον Πίνακα 2.1 – και έχει ακόμα πολλά περιθώρια ανάπτυξης.

Πίνακας 2. 1: Συγκεντρωτική βιβλιογραφία στην προτυποποίηση πρόβλεψης της τροχιάς οχημάτων και πεζών

Τίτλος	Συγγραφείς	Έτος δημοσίευσης	Εφαρμοζόμενες Μέθοδοι	Πράκτορες	Δεδομένα	Αυτόνομα Οχήματα*
Convolutional Neural Network for Trajectory Prediction	Nikhil & Morris	2018	CNN	Πεζός (Unimodal)	Πραγματικά δεδομένα	ΝΑΙ
Forecasting Pedestrian Trajectory with Machine-Annotated Training Data	Styles, Ross & Sanchez	2019	CNN	Πεζός (Unimodal)	Μηχανικά επεξεργασμένα, πραγματικά δεδομένα	ΌΧΙ
Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks	Gupta et al.	2018	GAN + LSTM	Πεζοί (Multimodal)	Πραγματικά δεδομένα	ΟΧΙ
M2P3: Multimodal Multi-Pedestrian Path Prediction by Self-Driving Cars With Egocentric Vision	Poibrenski et al.	2020	RNN encoder-decoder	Πεζοί (Multimodal)	Πραγματικά δεδομένα	ΝΑΙ
A Generic Framework for Clustering Vehicle Motion Trajectories	Hoseini et al.	2020	Clustering, GAN	Όχημα (Unimodal)	Τεχνητά με χρήση GAN δεδομένα	ΝΑΙ
Modeling Vehicle Interactions via Modified LSTM Models for Trajectory Prediction	Dai, Li & Li	2019	LSTM	Όχημα (Unimodal)	Πραγματικά δεδομένα	ΌΧΙ
Multi-Agent Tensor Fusion for Contextual Trajectory Prediction	Zhao et al.	2019	CNN + LSTM + GAN	Πεζοί ή Οχήματα (Multimodal)	Πραγματικά δεδομένα	ΌΧΙ
Multimodal Trajectory Predictions for Autonomous Driving using Deep Convolutional Networks	Cui et al.	2019	DCN	Οχήματα (Multimodal)	Πραγματικά δεδομένα	ΝΑΙ
DESIRE: Distant Future Prediction in Dynamic Scenes with Interacting Agents	Lee et al.	2017	Encoder-Decoder, RNN	Πεζοί και Οχήματα (Multi-Agent)	Πραγματικά δεδομένα	ΝΑΙ
The Trajectron: Probabilistic Multi-Agent Trajectory Modeling With Dynamic Spatiotemporal Graphs	Boris Ivanovic, Marco Pavone	2018	CVAE, LSTM, Dynamic Spatiotemporal Graphical Structure	Πεζοί και Οχήματα (Multi-Agent)	Πραγματικά δεδομένα	ΟΧΙ
Trajectron++: Dynamically-Feasible Trajectory Forecasting With Heterogeneous Data	Salzmann et al.	2021	Recurrent Graph NN	Πεζοί και Οχήματα (Multi-Agent)	Πραγματικά δεδομένα	ΟΧΙ
Scene Transformer: A unified architecture for predicting multiple agent trajectories	Ngiam et al.	2021	Transformer	Πεζοί και Οχήματα (Multi-Agent)	Πραγματικά δεδομένα	ΟΧΙ
AgentFormer: Agent-Aware Transformers for Socio-Temporal Multi-Agent Forecasting	Yuan et al.	2021	Transformer	Πεζοί και Οχήματα (Multi-Agent)	Πραγματικά δεδομένα	ΝΑΙ
S2TNet: Spatio-Temporal Transformer Networks for Trajectory Prediction in Autonomous Driving	Chen et al.	2021	Transformer	Πεζοί, ποδηλάτες και Οχήματα (Multi-Agent)	Πραγματικά δεδομένα	ΝΑΙ

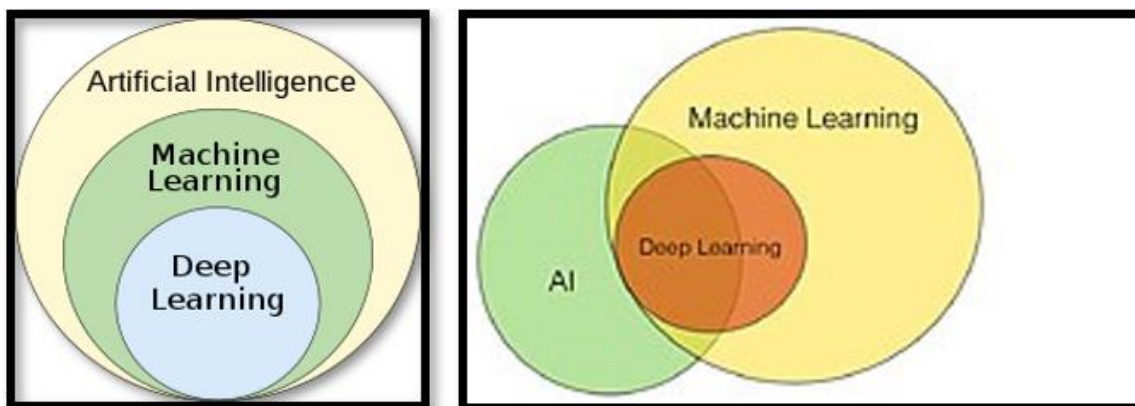
*: Όλες οι μεθοδολογίες είναι εφαρμόσιμες για αυτόνομα οχήματα. Η στήλη αφορά τα δεδομένα που χρησιμοποιήθηκαν για επαλήθευση των μοντέλων

ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

Η παρούσα εργασία έχει ως κύριο άξονα την χρήση μεθόδων βαθιάς μάθησης με σκοπό την πρόβλεψη των τροχιών αυτόνομων οχημάτων και πεζών κατά την αλληλεπίδραση τους. Στο παρόν κεφάλαιο, αρχικά θα γίνει μια σύντομη επεξήγηση των βασικών εννοιών και χαρακτηριστικών της βαθιάς μάθησης. Στην συνέχεια θα γίνει σύντομη ανάλυση του θεωρητικού υποβάθρου που χρησιμοποιείται από καθεμία από τις μεθοδολογίες της παρούσας εργασίας.

3.1 Βαθιά Μάθηση

Όλες οι τεχνικές μάθησης που εκτελούνται με τη βοήθεια ηλεκτρονικού υπολογιστή, χρησιμοποιώντας ανθρώπινα μοτίβα ως πρότυπα χαρακτηρίζονται ως μεθοδολογίες τεχνητής νοημοσύνης. Αν και τα τελευταία χρόνια ορισμένοι ερευνητές το αμφισβητούν [Langley, 2011] [James et al., 2014], η μηχανική μάθηση αποτελεί μια σημαντική υποκατηγορία της τεχνητής νοημοσύνης. Το χαρακτηριστικό που διαφοροποιεί την μηχανική μάθηση από τις υπόλοιπες υποκατηγορίες της τεχνητής νοημοσύνης είναι το ότι βασίζεται σε συγκεκριμένα και απaráλλακτα δεδομένα για να κάνει προβλέψεις, ενώ η τεχνητή νοημοσύνη βασίζεται στις επιλογές ενός ή περισσότερων πρακτόρων που βρίσκονται εντός ενός συγκεκριμένου περιβάλλοντος και μαθαίνουν ποιες είναι οι ενέργειες που πρέπει να επιτελέσουν για να επιτύχουν τον στόχο τους. Οι δύο διαφορετικές αντιλήψεις των ειδικών σε σχέση με τη κατηγοριοποίηση των μεθόδων μάθησης φαίνεται στην Εικόνα 3.1.

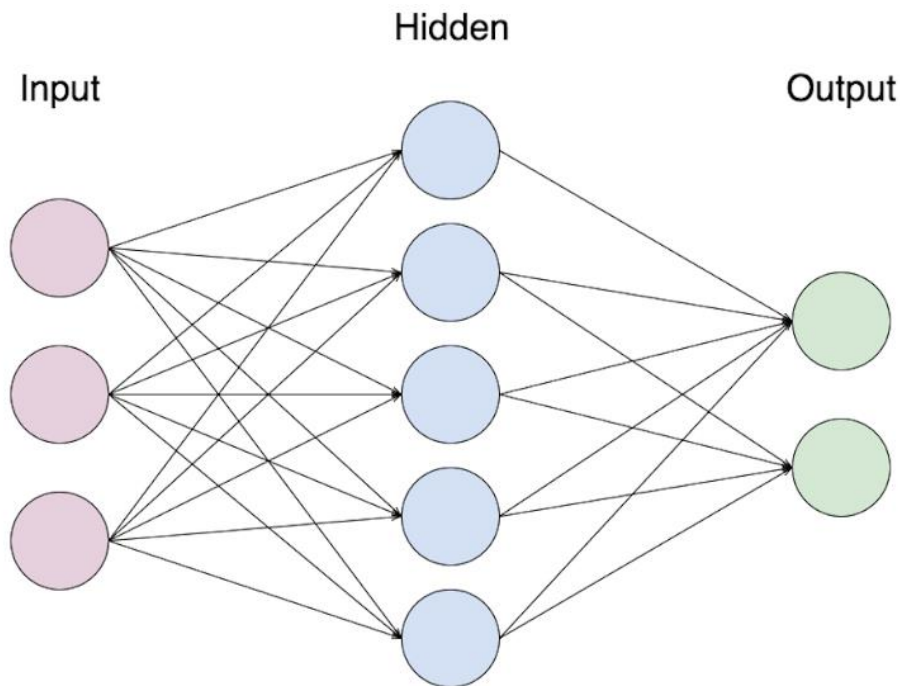


Εικόνα 3. 1: Οι δύο επικρατούσες απόψεις για τη σχέση της Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης και της Βαθιάς Μάθησης [Πηγή: Machine Learning, Wikipedia]

3.1.1 Νευρωνικά δίκτυα και βαθιά νευρωνικά δίκτυα

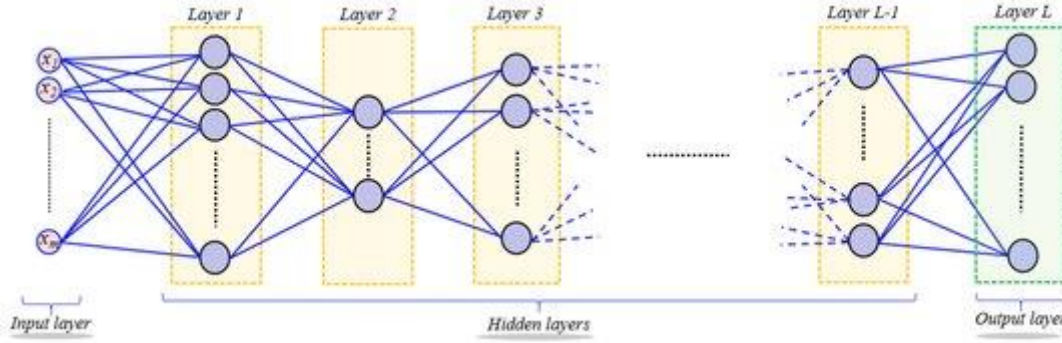
Ένα νευρωνικό δίκτυο εν γένει έχει σχετικά απλή δομή. Σε πρώτη φάση αποτελείται από ένα επίπεδο εισαγωγής δεδομένων. Το επίπεδο αυτό αποτελείται από N νευρώνες (όσες είναι και οι μεταβλητές εισόδου) που επεξεργάζονται τα εισαγόμενα δεδομένα με τέτοιο τρόπο ώστε να τα μορφοποιήσουν στην κατάλληλη μορφή για την ακόλουθη χρήση

τους. Ύστερα ακολουθεί το πρώτο (μπορεί να είναι παράλληλα και το τελευταίο) κρυφό επίπεδο, το οποίο επίσης αποτελείται από N νευρώνες στους οποίους εισάγεται μια σειρά από τιμές (x_1, x_2, \dots, x_n) προερχόμενες από τους νευρώνες του επιπέδου εισόδου με κάθε τιμή εντός της σειράς να προέρχεται από έναν από αυτούς. Στην συνέχεια οι τιμές πολλαπλασιάζονται με το διάνυσμα βάρους του κρυφού επιπέδου (w_1, w_2, \dots, w_n) και ο νευρώνας αθροίζει τις τιμές του παραγόμενου διανύσματος $(x_1 \cdot w_1 + x_2 \cdot w_2 + \dots + x_n \cdot w_n)$. Ακολούθως, αφού εφαρμόσει μια συνάρτηση ενεργοποίησης, ο εκάστοτε νευρώνας μεταφέρει την τιμή που προέκυψε στο επόμενο επίπεδο. Τελευταίο επίπεδο σε ένα νευρωνικό δίκτυο είναι το επίπεδο εξόδου, στο οποίο γίνεται ο τελικός υπολογισμός των προβλέψεων με βάση τα πραγματικά δεδομένα. Η αρχιτεκτονική ενός δικτύου με ένα κρυφό επίπεδο φαίνεται στην Εικόνα 3.2.



Εικόνα 3. 2: Αρχιτεκτονική νευρωνικού δικτύου με ένα κρυφό επίπεδο
[Πηγή: towardsdatascience.com]

Όλες οι μεθοδολογίες που χρησιμοποιούνται σε αυτή την εργασία αποτελούν βαθιά νευρωνικά δίκτυα. Η κύρια διαφοροποίηση των εν λόγω δικτύων είναι ότι δομούνται από πολλαπλά κρυφά επίπεδα. Με αυτόν τον τρόπο – μαζί με άλλες διαφοροποιήσεις μικρότερης σημασίας – τα βαθιά δίκτυα είναι σε θέση να παράγουν πολύ ποιοτικότερα αποτελέσματα σε σχέση με τα απλά νευρωνικά δίκτυα του ενός κρυφού επιπέδου. Στην Εικόνα 3.3 φαίνεται η αρχιτεκτονική ενός βαθιού νευρωνικού δικτύου.



Εικόνα 3. 3: Αρχιτεκτονική βαθιού νευρωνικού δικτύου [Πηγή: Altaf et al., 2019]

3.1.2 Μαθηματική έκφραση βαθιού νευρωνικού δικτύου

Ορίζοντας με αυστηρά μαθηματικό τρόπο ένα βαθύ δίκτυο, αυτό αποτελεί ένα ιεραρχικό πρότυπο στο οποίο κάθε επίπεδο εφαρμόζει μια γραμμική μετατροπή ακολουθούμενη από μια μη γραμμική στο προηγούμενο επίπεδο. Συγκεκριμένα:

- Έστω $X \in \mathbf{R}^{N \times D}$ η είσοδος, όπου κάθε γραμμή του X αποτελείται από στοιχεία των δεδομένων D διαστάσεων (για παράδειγμα μια χρονοσειρά των τιμών της ταχύτητας ενός οχήματος για D συνεχόμενες χρονικές στιγμές) και N είναι ο αριθμός των παραδειγμάτων εκπαίδευσης του πρότυπου (για παράδειγμα ο αριθμός των διαφορετικών συμβάντων αλληλεπίδρασης μεταξύ αυτόνομου οχήματος και πεζού).
- Έστω $W_k \in \mathbf{R}^{d_{k-1} \times d_k}$ ο πίνακας των βαρών (weights) που πολλαπλασιάζεται γραμμικά με το διάνυσμα εξόδου του επιπέδου $k-1$, $X_{k-1} \in \mathbf{R}^{N \times d_{k-1}}$. Υπολογίζει την d_k -διάστατη εκτίμηση του $X_{k-1}W^k \in \mathbf{R}^{N \times d_k}$ στο επίπεδο k .
- Έστω $\psi_k: \mathbf{R} \rightarrow \mathbf{R}$ μια μη γραμμική συνάρτηση ενεργοποίησης η οποία μπορεί να πάρει διάφορες τιμές με κάποιες από τις πιο χρησιμοποιούμενες να είναι:
 - Η υπερβολική εφαιπτομένη $\psi_k(x) = \tanh(x)$
 - Το σιγμοειδές $\psi_k(x) = (1 + e^{-x})^{-1}$
 - Η συνάρτηση διορθωμένης γραμμικής μονάδας (Rectified Linear Unit – ReLU) $\psi_k(x) = \max\{0, x\}$

Η μη γραμμική συνάρτηση ενεργοποίησης εφαρμόζεται σε κάθε είσοδο $X_{k-1}W^k$ στο επίπεδο και έτσι παράγεται το k επίπεδο του νευρωνικού δικτύου ως $X_k = \psi_k(X_{k-1}W^k)$. Η έξοδος X_k του δικτύου εκφράζεται μαθηματικά ως εξής:

$$\Phi(X, W^1, \dots, W^K) = \psi_K(\psi_{K-1}(\dots \psi_2(\psi_1(XW^1)W^2) \dots W^{K-1})W^K) \quad (3.1)$$

όπου Φ είναι ένας Πίνακας $N \times C$, όπου $C = d_K$ είναι η διάσταση της εξόδου του δικτύου. Παρατηρώντας την παραπάνω σχέση μπορεί να θεωρηθεί ότι ο Φ αποτελεί συνάρτηση των βαρών του δικτύου $W = \{W^k\}_{k=1}^K$ με είσοδο X .

Στον παραπάνω ορισμό των νευρωνικών δικτύων ορίστηκε το πως ακριβώς καθορίζονται η είσοδος σε κάθε επίπεδο καθώς και οι βασικότερες μορφές που μπορεί

να λάβει η συνάρτηση ενεργοποίησης. Όσον αφορά τα βάρη $W = \{W^k\}_{k=1}^K$, ένα βαθύ δίκτυο τα «μαθαίνει» μέσω N παραδειγμάτων εκπαίδευσης της μορφής (X, Y) . Σε προβλήματα προβλέψεων-παλινδρόμησης όπως αυτό της παρούσας εργασίας οι γραμμές του $Y \in \mathbb{R}^{N \times C}$ εκφράζουν τις εξαρτημένες μεταβλητές για τις γραμμές του X . Η μάθηση των προαναφερθέντων βαρών ενός βαθιού νευρωνικού δικτύου εκφράζονται από το ακόλουθο πρόβλημα βελτιστοποίησης:

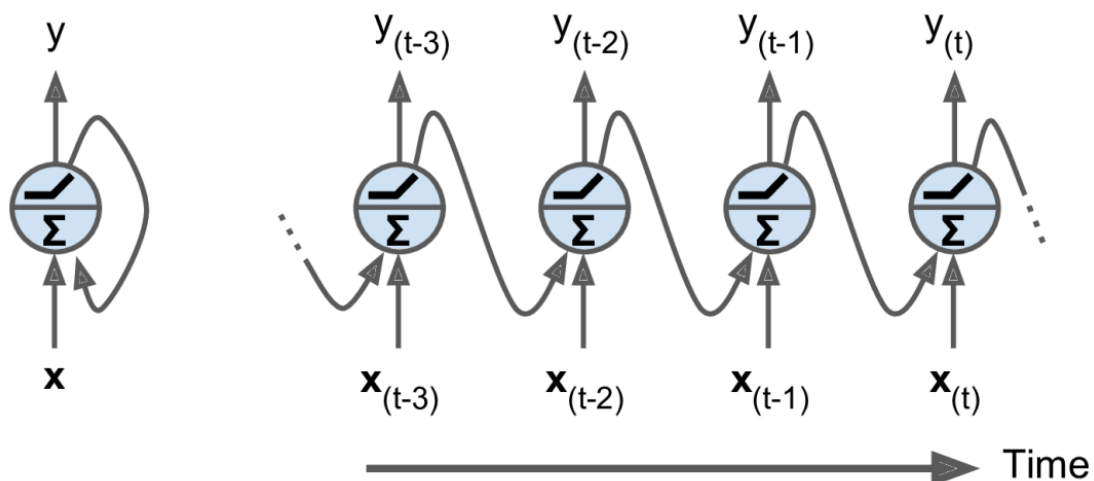
$$\min_{W=\{W^k\}_{k=1}^K} \ell(Y, \Phi(X, W^1, \dots, W^K)) + \lambda \theta(X, W^1, \dots, W^K) \quad (3.2)$$

όπου $\ell(Y, \Phi)$ είναι μια συνάρτηση απώλειας (loss function) που υπολογίζει την διαφορά μεταξύ της πραγματικής εξόδου Y και της πρόβλεψης της εξόδου $\Phi(X, W)$, θ είναι μια συνάρτηση κανονικοποίησης (regularization function) σχεδιασμένη ώστε να αποτρέπει την υπερεκπαίδευση (overfitting).

3.2 Ανατροφοδοτούμενα νευρωνικά δίκτυα - RNNs

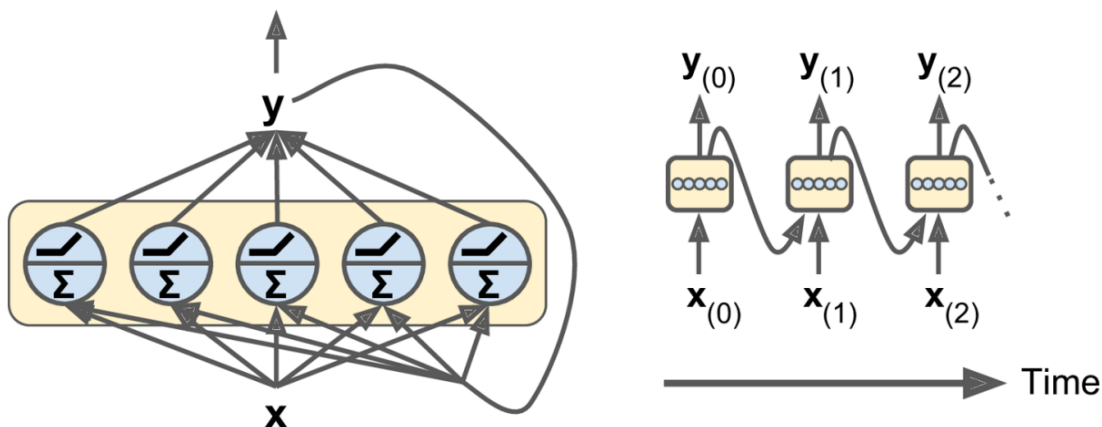
Τα RNNs είναι μια κατηγορία νευρωνικών δικτύων που είναι σχεδιασμένα ώστε να μπορούν να προβλέψουν το μέλλον με σημαντική ακρίβεια. Λόγω της δυνατότητας τους να προβλέπουν το μέλλον με χρήση δεδομένων χρονοσειρών σε υψηλό επίπεδο, έχουν ευρεία χρήση σε πολλά επιστημονικά πεδία με κάποια από αυτά να είναι τα χρηματιστηριακά, η γλωσσολογία αλλά και η αυτόνομη οδήγηση.

Ένα RNN έχει δομή ίδια με αυτή ενός τυπικού νευρωνικού δικτύου εμπρόσθιας μετάδοσης (Feed-Forward Neural Network) με την διαφορά ότι εμπεριέχει και συνδέσεις με ανεστραμμένη φορά. Για να γίνει κατανοητό αυτό, έστω ένα RNN με μόνον ένα νευρώνα, ο οποίος δέχεται είσοδο και παράγει έξοδο. Σε κάθε χρονικό βήμα t , ο εν λόγω ανατροφοδοτούμενος νευρώνας δέχεται την είσοδο $x_{(t)}$ καθώς και την έξοδο που προέκυψε από το προηγούμενο χρονικό βήμα $y_{(t-1)}$. Η παραπάνω ιδέα φαίνεται στην Εικόνα 3.4, όπου παρουσιάζεται ο νευρώνας ανηγμένος στον χρόνο.



Εικόνα 3. 4: Ανατροφοδοτούμενος νευρώνας, ανηγμένος στον χρόνο
[Πηγή: Geron, 2019]

Λόγω του ότι στο παραπάνω υπεραπλοσθευτικό RNN υπάρχει μόνο ένας ανατροφοδοτούμενος νευρώνας οι εισόδοι και έξοδοι αποτελούν μονόμετρα μεγέθη. Εντάσσοντας πολλαπλούς τέτοιους νευρώνες, δημιουργείται ένα επίπεδο ανατροφοδοτούμενων νευρώνων, το οποίο έχει διανύσματα ως εισόδους και εξόδους. Στην παρακάτω Εικόνα 3.5 απεικονίζεται ένα τέτοιο επίπεδο.



Εικόνα 3. 5: Επίπεδο ανατροφοδοτούμενων νευρώνων, ανηγμένο στον χρόνο
[Πηγή: Geron, 2019]

Όσον αφορά τα βάρη κάθε νευρώνα, υπάρχει ένα για τις εισόδους $x(t)$ που συμβολίζεται ως w_x και ένα για τις εξόδους του προηγούμενου χρονικού βήματος $y(t-1)$ που συμβολίζεται ως w_y . Τα βάρη αυτά συνθέτουν για κάθε επίπεδο τους αντίστοιχους πίνακες βαρών W_x και W_y . Το διάνυσμα εξόδου ενός επιπέδου ανατροφοδοτούμενων νευρώνων υπολογίζεται με τη χρήση της παρακάτω σχέσης:

$$y(t) = \psi(W_x^T x(t) + W_y^T y(t-1) + b) \quad (3.3)$$

όπου b είναι το διάνυσμα κατωφλιού και ψ είναι η συνάρτηση ενεργοποίησης, όπως αυτή ορίστηκε στη προηγούμενη παράγραφο.

Η έξοδος ενός ανατροφοδοτούμενου νευρώνα στο χρονικό βήμα t είναι συνάρτηση των εισόδων όλων των προηγούμενων χρονικών βημάτων. Με βάση αυτό το γεγονός, γίνεται κατανοητό ότι ένας τέτοιος νευρώνας έχει κατά κάποια έννοια «μνήμη». Γενικότερα οποιοδήποτε τμήμα ενός νευρωνικού δικτύου είναι σε θέση να διατηρεί μια κατάσταση για ένα σύνολο χρονικών βημάτων καλείται κελί μνήμης (memory cell) ή απλώς κελί. Ένας ανατροφοδοτούμενος νευρώνας καθώς και ένα επίπεδο ανατροφοδοτούμενων νευρώνων μπορεί να χαρακτηριστεί ως κελί μνήμης. Σε επόμενη παράγραφο αναλύεται και το κελί LSTM όπου αποτελεί ένα πιο περίπλοκο είδος κελιού που χρησιμοποιείται εκτενώς στην παρούσα εργασία.

3.2.1 Αρχιτεκτονική Encoder-Decoder

Όλα τα πρότυπα που εφαρμόζονται σε αυτήν την εργασία αποτελούν νευρωνικά δίκτυα Encoder-Decoder. Η εν λόγω αρχιτεκτονική αναπτύχθηκε με σκοπό την δημιουργία ενός

πρότυπου το οποίο είναι σε θέση να λαμβάνει ακολουθίες δεδομένων και να τις χρησιμοποιεί για να παραγάγει νέες ακολουθίες (sequence to sequence – Seq2Seq).

Ένα πρότυπο encoder-decoder αποτελείται από τρία επιμέρους τμήματα: το encoder (κωδικοποιητή), το διάνυσμα του κωδικοποιητή (encoder vector) και το decoder (αποκωδικοποιητή). Το encoder είναι ένα δίκτυο το οποίο ουσιαστικά αποτελεί ένα σύνολο συνεχόμενων κελιών μνήμης. Καθένα από αυτά αφορά ένα συγκεκριμένο στοιχείο της ακολουθίας εισόδου. Σκοπός του δικτύου είναι να μετατρέψει τις εισερχόμενη ακολουθία δεδομένων σε διάνυσμα, το οποίο αποτελεί το προαναφερθέν διάνυσμα του αποκωδικοποιητή. Το διάνυσμα αυτό στην συνέχεια εισέρχεται στο decoder. Το decoder επίσης αποτελεί ένα δίκτυο το οποίο συντελείται από μια σειρά κελιών μνήμης, τα οποία χρησιμοποιούν το εισερχόμενο διάνυσμα για να παραγάγουν μια νέα ακολουθία εξόδου, εκπληρώνοντας έτσι τον σκοπό του πρότυπου.

Η εν λόγω αρχιτεκτονική προτάθηκε από τους Cho et al. (2014) και Sutskever et al. (2014) όσον αφορά τα RNNs. Στο πλαίσιο ενός πρότυπου τέτοιας αρχιτεκτονικής το encoder διαβάζει την είσοδο η οποία αποτελεί μια ακολουθία διανυσμάτων $x = (x_1, \dots, x_i)$ και την μετατρέπει στο διάνυσμα του κωδικοποιητή (encoder vector) c^2 . Το βαθμωτό αυτό μέγεθος προκύπτει από την εφαρμογή ενός ή περισσότερων κρυφών επιπέδων (για παράδειγμα κελιών LSTM). Μαθηματικά αυτή η μετατροπή εκφράζεται από τις σχέσεις:

$$h_t = f(x_t, h_{t-1}) \quad (3.4)$$

$$c = q(\{h_1, \dots, h_{T_x}\}) \quad (3.5)$$

Όπου h_t είναι το κρυφό επίπεδο στην χρονική στιγμή t , c είναι το διάνυσμα του encoder και f, q είναι μη γραμμικές συναρτήσεις καθοριζόμενες από το κρυφό επίπεδο.

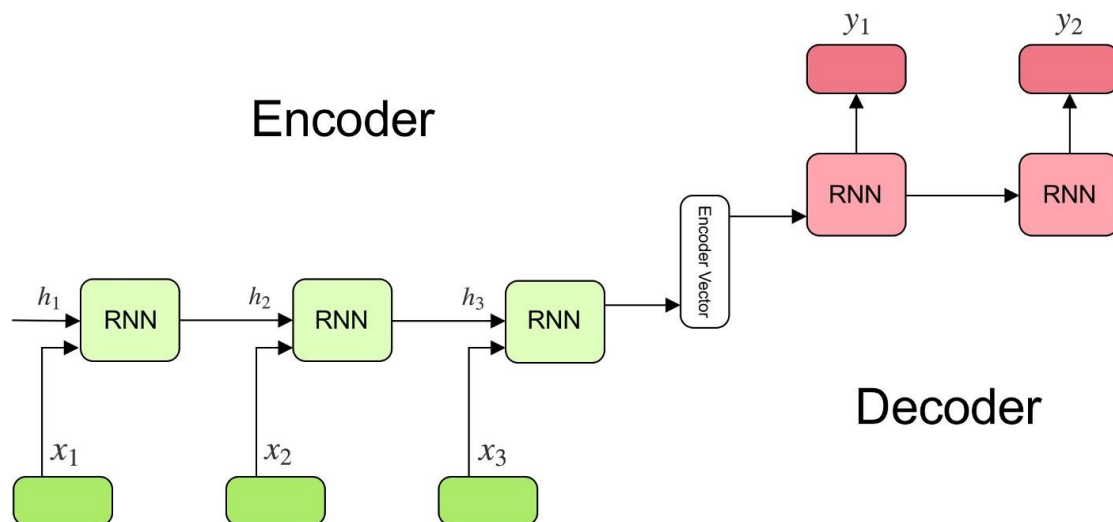
Όσον αφορά το decoder, με εισόδους το προαναφερθέν διάνυσμα του encoder και τις προβλέψεις στα προηγούμενα χρονικά βήματα $\{y_1, \dots, y_{t-1}\}$ στοχεύει στην πρόβλεψη κατά το παρόν χρονικό βήμα t , y_t . Με άλλα λόγια, το decoder υπολογίζει την πιθανότερη ορθή ακολουθία προβλέψεων $y = \{y_1, \dots, y_{T_y}\}$ ως εξής:

$$p(y) = \prod_{t=1}^T p(y_t | \{y_1, \dots, y_{t-1}\}, c) \quad (3.6)$$

Κάθε επιμέρους δεσμευμένη πιθανότητα προτυποποιείται με τη βοήθεια μιας μη γραμμικής, ενδεχομένως πολυεπίπεδης συνάρτησης g και του κρυφού επιπέδου s_t ως εξής:

$$p(y_t | \{y_1, \dots, y_{t-1}\}, c) = g(y_{t-1}, s_t, c) \quad (3.7)$$

Η απλοποιημένη αρχιτεκτονική ενός τέτοιου νευρωνικού δικτύου φαίνεται στην Εικόνα 3.6.



Εικόνα 3. 6: Αρχιτεκτονική ενός πρότυπου Encoder-Decoder
[Πηγή: towardsdatascience.com]

3.2.2 Το Κελί LSTM

Το κελί Long Short-Term Memory (LSTM) [Hochreiter & Schmidhuber, 1997] εδώ και δεκαετίες έχει ευρεία χρήση στον χώρο των προβλέψεων, με συνεχείς βελτιώσεις στην πορεία του χρόνου. Όλες οι μεθοδολογίες που αναπτύσσονται στην παρούσα εργασία στηρίζονται άμεσα ή έμμεσα σε αυτό το κελί. Το κελί LSTM μπορεί να χρησιμοποιηθεί ομοίως με ένα βασικό κελί όπως αυτό περιγράφηκε παραπάνω, με τη διαφορά ότι είναι πολύ πιο αποδοτικό.

Η κατάσταση (state) ενός κελιού LSTM χωρίζεται σε δύο διαφορετικά διανύσματα: το $h_{(t)}$ που αφορά τη βραχυπρόθεσμη κατάσταση και το $c_{(t)}$ που αφορά τη μακροπρόθεσμη κατάσταση. Η κύρια ιδέα γύρω από την εν λόγω αρχιτεκτονική είναι το δίκτυο είναι σε θέση να μάθει τι πρέπει να αποθηκεύσει στην μακροπρόθεσμη κατάσταση, τι να διαγράψει αλλά και τι να διαβάσει από αυτή. Όσο το διάνυσμα $c_{(t-1)}$ «κινείται» από αριστερά προς δεξιά εντός του κελιού, πρώτα διαπερνά μια πύλη απώλειας μνήμης (forget gate) στην οποία διαγράφει κάποια δεδομένα μνήμης, και στην συνέχεια αποκτά κάποια νέα δεδομένα μνήμης μέσω ενός τελεστή άθροισης από την πύλη εισόδου (input gate). Το προκύπτον διάνυσμα $c_{(t)}$ εξάγεται απευθείας από το κελί, χωρίς καμία περαιτέρω επεξεργασία. Επιπλέον, μετά την προαναφερθείσα εφαρμογή του τελεστή άθροισης, η μακροπρόθεσμη κατάσταση αντιγράφεται και μετασχηματίζεται μέσω της συνάρτησης υπερβολικής εφαπτομένης (tanh), με το αποτέλεσμα να φιλτράρεται μέσω της πύλης εξόδου (output gate). Έτσι παράγεται η βραχυπρόθεσμη κατάσταση $h_{(t)}$, η οποία είναι ίση με την έξοδο του κελιού $y_{(t)}$.

Όσον αφορά τον τρόπο με τον οποίο παράγει την έξοδο $y_{(t)}$ το κελί, σε πρώτη φάση εισάγονται σε αυτό το παρόν διάνυσμα εισόδου $x_{(t)}$ και η προηγούμενη βραχυπρόθεσμη κατάσταση $h_{(t-1)}$.

Συγκεκριμένα τα παραπάνω εισάγονται σε τέσσερα διαφορετικά πλήρως διασυνδεδεμένα επίπεδα, όπου το καθένα από αυτά εξυπηρετεί έναν διαφορετικό σκοπό:

- Το κύριο επίπεδο είναι εκείνο που έχει ως έξοδο το διάνυσμα $g_{(t)}$. Το επίπεδο αυτό αναλύει τις παρούσες εισόδους $x_{(t)}$ και την προηγούμενη βραχυπρόθεσμη κατάσταση $h_{(t-1)}$. Σε αντίθεση με ένα τυπικό κελί μνήμης όπου δεν υπάρχει τίποτα παραπάνω από αυτό το επίπεδο, η έξοδος αυτού του επιπέδου σε ένα LSTM κελί δεν οδηγείται απευθείας εκτός αυτού αλλά αποθηκεύεται μερικώς στην μακροπρόθεσμη κατάσταση.
- Τα υπόλοιπα τρία επίπεδα αποτελούν διαχειριστές της κάθε πύλης (gate controllers). Λόγω του ότι χρησιμοποιούν την λογιστική συνάρτηση ενεργοποίησης, οι έξοδοι τους έχουν εύρος από 0 έως 1. Όπως φαίνεται και στην Εικόνα 3.7, οι έξοδοι αυτών των τριών επιπέδων εισάγονται σε τελεστές πολλαπλασιασμού ανά στοιχείο, και επομένως άμα έχουν έξοδο μηδενικά «κλείνουν» την πύλη, ενώ άμα έχουν έξοδο άσσους την «ανοίγουν» πλήρως. Συγκεκριμένα, οι πύλες είναι οι παρακάτω:
 - Η πύλη απώλειας μνήμης (ελεγχόμενη από το $f_{(t)}$) ελέγχει ποια τμήματα της μακροπρόθεσμης κατάστασης πρέπει να διαγραφούν.
 - Η πύλη εισόδου (ελεγχόμενη από το $i_{(t)}$) ελέγχει ποια τμήματα του $g_{(t)}$ θα πρέπει να προστεθούν στην μακροπρόθεσμη κατάσταση.
 - Η πύλη εξόδου (ελεγχόμενη από το $o_{(t)}$) ελέγχει ποια τμήματα της μακροπρόθεσμης κατάστασης πρέπει να διαβαστούν και να εξαχθούν στο παρόν χρονικό βήμα στο $y_{(t)}$ (και στο $h_{(t)}$).

Εν ολίγοις, ένα κελί LSTM μαθαίνει να αναγνωρίζει μια σημαντική είσοδο, να την αποθηκεύει στην μακροπρόθεσμη κατάσταση, να την διατηρεί όσο χρειάζεται και να την εξάγει όπου χρειάζεται. Οι σχέσεις εκφράζουν τον τρόπο υπολογισμού της μακροπρόθεσμης και βραχυπρόθεσμης κατάστασης αλλά και της εξόδου σε κάθε χρονικό βήμα.

$$i_{(t)} = \sigma(W_{xi}^T \cdot x_{(t)} + W_{hi}^T \cdot h_{(t-1)} + b_i) \quad (3.8)$$

$$f_{(t)} = \sigma(W_{xf}^T \cdot x_{(t)} + W_{hf}^T \cdot h_{(t-1)} + b_f) \quad (3.9)$$

$$o_{(t)} = \sigma(W_{xo}^T \cdot x_{(t)} + W_{ho}^T \cdot h_{(t-1)} + b_o) \quad (3.10)$$

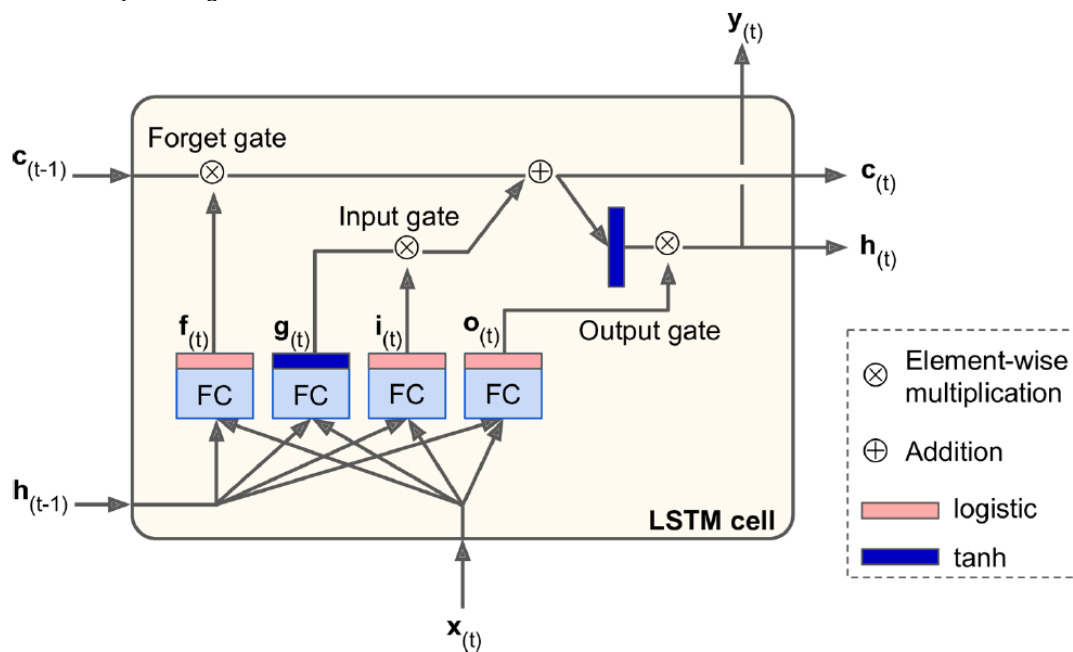
$$g_{(t)} = \sigma(W_{xg}^T \cdot x_{(t)} + W_{hg}^T \cdot h_{(t-1)} + b_g) \quad (3.11)$$

$$c_{(t)} = f_{(t)} \otimes c_{(t-1)} + i_{(t)} \otimes g_{(t)} \quad (3.12)$$

$$y_{(t)} = h_{(t)} = o_{(t)} \otimes \tanh(c_{(t)}) \quad (3.13)$$

Όπου:

- $W_{xi}, W_{xf}, W_{xo}, W_{xg}$ είναι οι πίνακες των βαρών για κάθε ένα από τα 4 επίπεδα όσον αφορά τις συνδέσεις τους με το διάνυσμα εισόδου $x(t)$.
- $W_{hi}, W_{hf}, W_{ho}, W_{hg}$ είναι οι πίνακες των βαρών για κάθε ένα από τα 4 επίπεδα όσον αφορά τις συνδέσεις τους με το διάνυσμα εισόδου $h(t-1)$.
- b_i, b_f, b_o, b_g είναι οι όροι κατωφλίου για κάθε ένα από τα 4 επίπεδα.

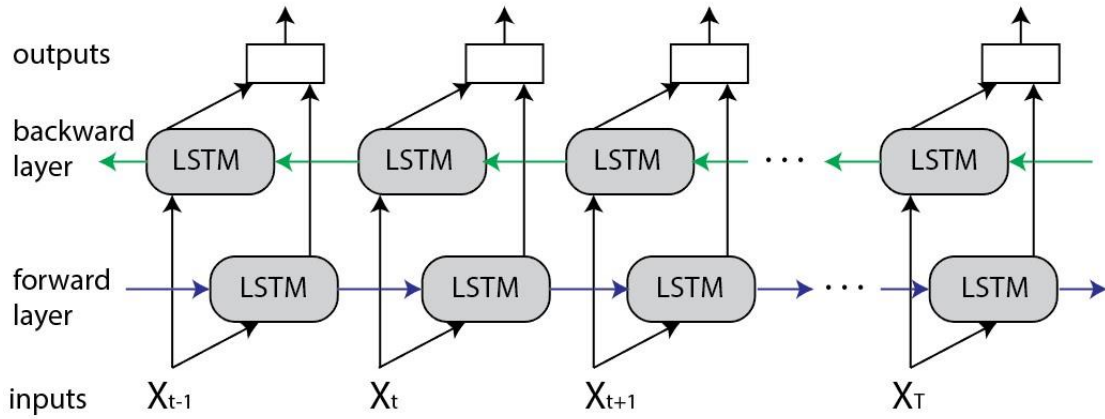


Εικόνα 3. 7: Αρχιτεκτονική κελιού LSTM [Πηγή: Geron, 2019]

3.2.3 Αμφίδρομο LSTM (BiLSTM)

Τα αμφίδρομα ανατροφοδοτούμενα νευρωνικά δίκτυα (bidirectional RNNs) αναπτύχθηκαν από τους Schuster & Paliwal (1997) κυρίως με σκοπό να αντιμετωπίσουν τον περιορισμό των παραδοσιακών RNNs όσον αφορά στο ποσοστό των δεδομένων που χρησιμοποιούνται για την πρόβλεψη. Τα αμφίδρομα RNNs διαφοροποιούνται με τα υπόλοιπα RNNs στο ότι πέρα από την θετική πορεία των χρονικών βημάτων, εφαρμόζεται παράλληλα και αρνητική χρονική πορεία. Δηλαδή κάθε τέτοιο πρότυπο εξελίσσεται από το πρώτο στο τελευταίο χρονικό βήμα και παράλληλα από το τελευταίο στο πρώτο. Οι έξοδοι των δύο αυτών κατευθύνσεων δεν συνδέονται μεταξύ τους.

Το αμφίδρομο LSTM (bidirectional LSTM), όπου εφεξής θα αναφέρεται για συντομία ως BiLSTM, λειτουργεί με βάση αυτήν την λογική. Η ονομασία του αναφέρεται στο γεγονός ότι αποτελεί ένα αμφίδρομο RNN όπου χρησιμοποιεί τα κελία μνήμης LSTM, ακριβώς όπως περιγράφηκαν προηγουμένως. Στην Εικόνα 3.8 φαίνεται ένα BiLSTM πρότυπο ανηγμένο στον χρόνο, στο οποίο είναι εμφανώς ορατό το επίπεδο θετικής χρονικής εξέλιξης αλλά και αυτό της αρνητικής.



Εικόνα 3. 8: Αρχιτεκτονική πρότυπου ανατροφοδοτούμενου RNN
[Πηγή: baeldung.com]

3.2.4 Μηχανισμός Έμφασης (Attention Mechanism)

Στις περιπτώσεις όπου κάποιες εισαγόμενες ακολουθίες πληροφοριών που μεταφέρονται από και προς κάθε επίπεδο ενός RNN έχουν μεγάλο μήκος, τα πρότυπα που χρησιμοποιούνται έχουν σχετικά μειωμένη απόδοση. Προς αντιμετώπιση αυτού του προβλήματος αναπτύχθηκε ο μηχανισμός έμφασης (attention mechanism) (Bahdanau et al., 2015), ο οποίος ουσιαστικά επιτρέπει σε ένα RNN να ενισχύσει κάποια κομμάτια μιας εισόδου και να σιγάσει κάποια άλλα με τέτοιο τρόπο ώστε να δοθεί έμφαση στα εξέχοντα χαρακτηριστικά (salient features) της εισόδου.

Ο εν λόγω μηχανισμός αφορά πρότυπα που χρησιμοποιούν την αρχιτεκτονική encoder-decoder. Η διαφοροποίηση του με την παραδοσιακή αρχιτεκτονική είναι ότι παράγεται διακριτό διάνυσμα encoder c_i για κάθε χρονικό βήμα. Κάθε τέτοιο διάνυσμα εξαρτάται από την ακολουθία χαρακτηριστικών (annotations) (h_1, \dots, h_{T_x}) . Καθένα από αυτά τα επίπεδα περιέχει πληροφορίες σχετικές με τη συνολική ακολουθία εισόδου με έμφαση στα κομμάτια που βρίσκονται κοντά στο παρόν χρονικό βήμα i . Συγκεκριμένα υπολογίζεται ως εξής:

$$c_i = \sum_{j=1}^{T_x} a_{ij} h_j \quad (3.14)$$

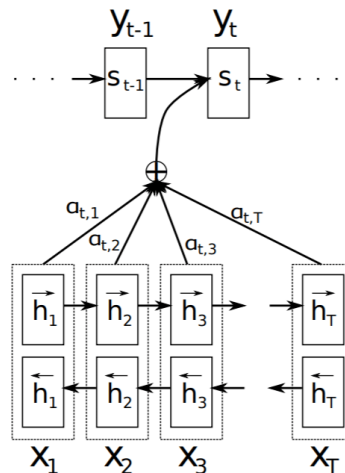
Ως a_{ij} συμβολίζεται το βάρος κάθε h_j και υπολογίζεται από την σχέση 3.15:

$$a_i = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})} \quad (3.15)$$

$$e_{ij} = a(s_{i-1}, h_j) \quad (3.16)$$

Όπου e_{ij} είναι ένα πρότυπο ευθυγράμμισης (alignment model) το οποίο μετρά πόσο «ταιριάζουν» οι εισοδοί στην θέση j και οι έξοδοι στην θέση i . Αυτή η μέτρηση βασίζεται στο κρυφό επίπεδο s_{i-1} και στο χαρακτηριστικό h_j .

Η παραπάνω σύντομη επεξήγηση του τρόπου λειτουργίας του μηχανισμού Attention αφορά στην μεθοδολογία η οποία εισάχθηκε από τους Bahdanau et al. (2015). Από την δημοσίευση του μηχανισμού το 2015 έχουν αναπτυχθεί πολυάριθμες άλλες μεθοδολογίες οι οποίες βασίζονται στην εν λόγω ιδέα. Στην Εικόνα 3.9 φαίνεται η κύρια ιδέα του μηχανισμού.

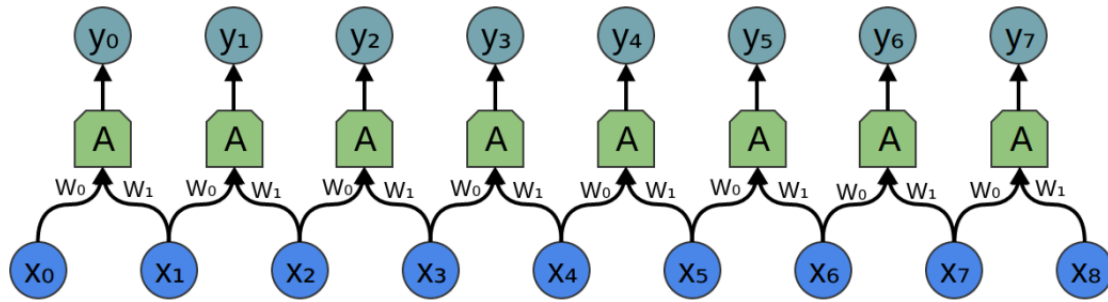


Εικόνα 3. 9: Οπτικοποίηση του μηχανισμού έμφασης-attention [Πηγή: Bahdanau et al., 2015]

3.3 Συνελικτικά νευρωνικά δίκτυα – CNNs

Τα συνελικτικά νευρωνικά δίκτυα (convolutional neural networks – CNNs) εισήχθησαν από τους LeCun et al. (1998). Τα τελευταία χρόνια, λόγω της συνεχούς αύξησης της υπολογιστικής δύναμης των υπαρχόντων τεχνολογιών, αποφέρουν ιδιαίτερως ποιοτικά αποτελέσματα σε πολυάριθμα επιστημονικά πεδία που εφαρμόζονται. Αν και είναι σχεδιασμένα και ευρέως χρησιμοποιούμενα στον τομέα της οπτικής ανάλυσης, έχουν εφαρμογή και σε πολλούς άλλους τομείς με κάποιους αυτούς να είναι η επεξεργασία φυσικής γλώσσας (natural language processing) και ο σχεδιασμός των αυτόνομων οχημάτων. Μπορούν να χρησιμοποιηθούν και στις προβλέψεις χρονοσειρών.

Ως βαθιά νευρωνικά δίκτυα, τα CNNs έχουν την μορφή που περιγράφηκε παραπάνω και φαίνεται στην Εικόνα 3.3. Η ειδοποιός διαφορά τους από τις υπόλοιπες τεχνικές βαθιάς μάθησης είναι ότι για κρυφά επίπεδα χρησιμοποιούν τα συνελικτικά επίπεδα (convolutional layers). Η λειτουργία ενός τέτοιου επιπέδου βασίζεται στην συνέλιξη. Μέσω της εφαρμογής συνέλιξης με πολλαπλά φίλτρα στην είσοδο του κρυφού επιπέδου αυξάνονται σημαντικά οι ενεργοποιήσεις των επιμέρους νευρώνων στα σημεία που τα φίλτρα «ταιριάζουν» με την είσοδο, εντοπίζοντας έτσι πιθανά χαρακτηριστικά της εξόδου. Συνδυάζοντας αυτά τα χαρακτηριστικά το επίπεδο είναι σε θέση να μάθει χαρακτηριστικά υψηλότερου επιπέδου, όπως φαίνεται για την περίπτωση απλής συνέλιξης σε μονοδιάστατη είσοδο (μπορεί να εφαρμοστεί και για είσοδο δύο διαστάσεων) στην Εικόνα 3.10.



Εικόνα 3. 10: Εφαρμογή απλής συνέλιξης σε μονοδιάστατη είσοδο [Πηγή: Olah, 2014]

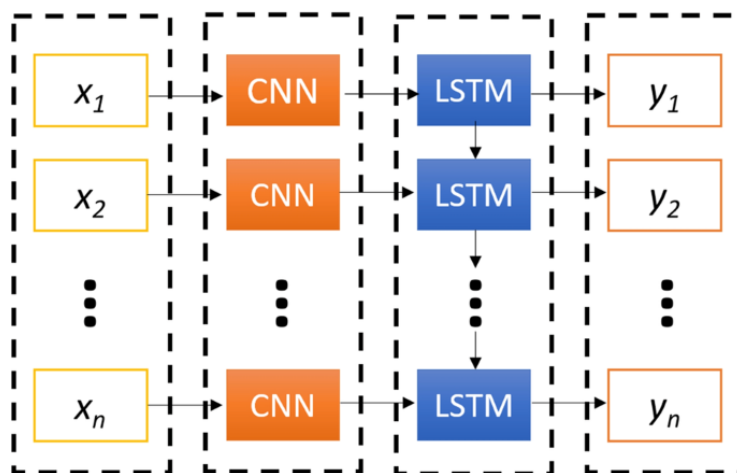
Στην Εικόνα 3.10 εφαρμόζεται σε κάθε ζεύγος x_i, x_{i+1} συνάρτηση με μορφή αντίστοιχη με αυτή των πυλών απώλειας μνήμης, εισόδου και εξόδου του κελιού LSTM, με τα W_i που φαίνονται στην Εικόνα να αποτελούν τα βάρη για κάθε είσοδο i . Όσον αφορά την εφαρμογή της συνέλιξης, οι διακριτές εξισώσεις που εφαρμόζονται σε μονοδιάστατα (3.17) και δυοδιάστατα (3.18) δεδομένα αντίστοιχα είναι οι παρακάτω:

$$(f * g)[n] = \sum_{m=-M}^M f(m)g(n - m) \quad (3.17)$$

$$(f * g)[i, j] = \sum_{m=-M}^M \sum_{n=-N}^N f(m, n)g(i - m, j - n) \quad (3.18)$$

3.3.1 CNN-RNN βαθύ νευρωνικό δίκτυο

Ο λόγος που παραπάνω επεξηγούνται τα CNNs είναι ότι κάποια από τα εφαρμοζόμενα πρότυπα αποτελεί συνδυασμό CNN και RNN επιπέδων. Η εν λόγω αρχιτεκτονική εισήχθη γενικευμένα για συνδυασμό CNN και RNN προτύπων (Donahue et al., 2014) και δεν αποτελεί τίποτα παραπάνω από ένα βαθύ νευρωνικό δίκτυο στο οποίο ανατροφοδοτούμενα επίπεδα όπως το κελί LSTM ακολουθούν συνελικτικά επίπεδα. Στην Εικόνα 3.11 είναι ορατή η οπτικοποίηση ενός τέτοιου πρότυπου.



Εικόνα 3. 11: Οπτικοποίηση CNN-RNN νευρωνικού δικτύου [Πηγή: Tasdelen & Sen, 2021]

3.4 Συνάρτηση απώλειας

Κατά την εκπαίδευση ενός πρότυπου, η συνάρτηση απώλειας (loss function) αποτελεί μια αριθμητική έκφραση των παραμέτρων του πρότυπου. Σε οποιοδήποτε πρόβλημα βελτιστοποίησης σκοπός είναι η ελαχιστοποίηση της εν λόγω συνάρτησης.

Υπάρχουν πολλές διαφορετικές τέτοιες συναρτήσεις. Στην παρούσα εργασία αποφασίζεται η χρήση της απώλειας Huber, η οποία περιγράφει την ποινή που προκαλείται από μια διαδικασία εκτίμησης f (penalty of an estimation procedure). Έτσι η εξίσωση υπολογισμού της απώλειας Huber είναι η εξής:

$$L_{\delta}(a) = \begin{cases} \frac{1}{2}a^2, & |a| \leq \delta \\ \delta \cdot (|a| - \frac{1}{2}\delta), & \text{αλλού} \end{cases} \quad (3.19)$$

Η εξίσωση 3.19 είναι δευτεροβάθμια για μικρές τιμές του a και γραμμική για μεγάλες τιμές, με ίσες τιμές και κλίση για $|a| = \delta$. Ορίζοντας ως a τα κατάλοιπα, ισχύει $a = y - f(x)$ και έτσι η σχέση 3.19 μπορεί να γενικευθεί ως ακολούθως:

$$L_{\delta}(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2, & |y - f(x)| \leq \delta \\ \delta \cdot (|y - f(x)| - \frac{1}{2}\delta), & \text{αλλού} \end{cases} \quad (3.20)$$

Η απώλεια Huber είναι ιδιαίτερα δημοφιλής λόγω του γεγονότος ότι είναι σε θέση να συνδυάσει την ευαισθησία της μέσης αμερόληπτης εκτιμήτριας ελάχιστης διασποράς (χρησιμοποιώντας την δευτεροβάθμια εξίσωση) και την ευρωστία της μέσης αμερόληπτης εκτιμήτριας (χρησιμοποιώντας την συνάρτηση απόλυτης τιμής).

3.5 Μετρικές αξιολόγησης

Στην στατιστική υπάρχει μια σειρά από μεγέθη τα οποία χρησιμοποιούνται στην επιστήμη δεδομένων για την αξιολόγηση των χρησιμοποιούμενων προτύπων. Στην συγκεκριμένη εργασία ως μετρικές αξιολόγησης για όλα τα πρότυπα χρησιμοποιήθηκαν το MAE και το MSE.

Το Μέσο Απόλυτο Σφάλμα (Mean Absolute Error) ή MAE αποτελεί μια μετρική εν γένει αποτελεί μέτρο σύγκρισης της διαφοράς μεταξύ ενός ζεύγους παρατηρήσεων. Στην περίπτωση της πρόβλεψης χρονοσειρών η σύγκριση αυτή γίνεται μεταξύ των πραγματικών και των προβλεπόμενων τιμών. Η εξίσωση υπολογισμού του MAE είναι η εξής:

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (3.21)$$

Το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error) ή MSE έχει τον ίδιο ακριβώς ρόλο με τη διαφορά ότι αναφέρεται στην τετραγωνισμένη διαφορά μεταξύ του ζεύγους παρατηρήσεων. Η εξίσωση υπολογισμού του MSE είναι η ακόλουθη:

$$MSE = \frac{\sum_{i=1}^n (y_i - x_i)^2}{n} \quad (3.22)$$

ΜΕΘΟΔΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

Ύστερα από την παρουσίαση του θεωρητικού υποβάθρου ακολουθεί το κεφάλαιο της μεθοδολογικής προσέγγισης. Στο κεφάλαιο αυτό αναλύεται η μεθοδολογία που ακολουθήθηκε όσον αφορά τα δεδομένα που χρησιμοποιήθηκαν καθώς και τα πρότυπα που εφαρμόστηκαν σε αυτά.

4.1 Περιγραφή δεδομένων

Για την πρόβλεψη της κίνησης αυτόνομων οχημάτων και πεζών χρειάζονται δεδομένα τροχιών τέτοιων αλληλεπιδράσεων. Όπως αναφέρεται και στην παράγραφο 2.2.1 υπάρχουν τρεις βασικές κατηγορίες συγκέντρωσης δεδομένων για προβλήματα αυτού του είδους: οι προσομοιώσεις, η τεχνική Wizard-of-Oz και τα πραγματικά δεδομένα. Παρόλο που εν γένει τα πραγματικά δεδομένα είναι τα ποιοτικότερα, υπάρχουν ορισμένοι περιορισμοί που κάνουν την συγκέντρωσή τους σχετικά δύσκολη. Αντιθέτως η τεχνική Wizard-of-Oz αναιρεί αυτούς τους περιορισμούς καθώς χρησιμοποιεί συμβατικό όχημα με οδηγό, «μεταμφιεσμένο» σε αυτόνομο. Για τον λόγο αυτό τα δεδομένα που χρησιμοποιούνται στην παρούσα εργασία είναι αυτής της μορφής.

Συγκεκριμένα, χρησιμοποιείται το Euro-PVI Dataset [Bhattacharyya et al., 2021]. Το όχημα που χρησιμοποιήθηκε για την συγκέντρωση των δεδομένων ήταν εξοπλισμένο με σύστημα LIDAR (Light Detection And Ranging) 360°, σύστημα τοποθέτησης (positioning system) και κάμερες με κατεύθυνση προς τα εμπρός. Στο όχημα αυτό υπήρχε οδηγός και συνοδηγός κατά την συγκέντρωση των δεδομένων. Τα δεδομένα συγκεντρώθηκαν στο Βέλγιο, και συγκεκριμένα σε περιοχές με μεγάλη κίνηση πεζών εντός του Λέβεν και των Βρυξελλών. Τα δεδομένα συγκεντρώθηκαν σε περίοδο δύο εβδομάδων, κατά την οποία ο οδηγός και ο συνοδηγός πλοηγούνταν οκτώ ώρες τη μέρα σε προκαθορισμένες περιοχές. Συγκεντρώθηκαν οι τροχιές του οχήματος καθώς και των ευάλωτων χρηστών (πεζών και ποδηλατών) οι οποίοι συμμετείχαν σε αλληλεπιδράσεις. Στο πλαίσιο της συγκέντρωσης των δεδομένων, ως αλληλεπίδραση ορίστηκε κάθε συμβάν κατά το οποίο η παρουσία είτε του οχήματος είτε του ευάλωτου χρήστη οδηγεί σε μεταβολή της ταχύτητας του άλλου. Παραδείγματα αλληλεπίδρασης είναι η συνάντηση του οχήματος με ένα πεζό σε σχετική διάβαση (Εικόνα 4.1 αριστερά) αλλά και η επικοινωνία μεταξύ των εμπλεκόμενων που οδηγεί το όχημα στο να μειώσει ταχύτητα (Εικόνα 4,2 δεξιά).



Εικόνα 4. 1: Παραδείγματα αλληλεπιδράσεων στο Euro-PVI dataset
[Πηγή: Bhattacharyya et al., 2021]

4.2 Επεξεργασία δεδομένων

Το Euro-PVI dataset είναι διαθέσιμο σε μορφή json. Συγκεκριμένα διατίθενται δύο φάκελοι:

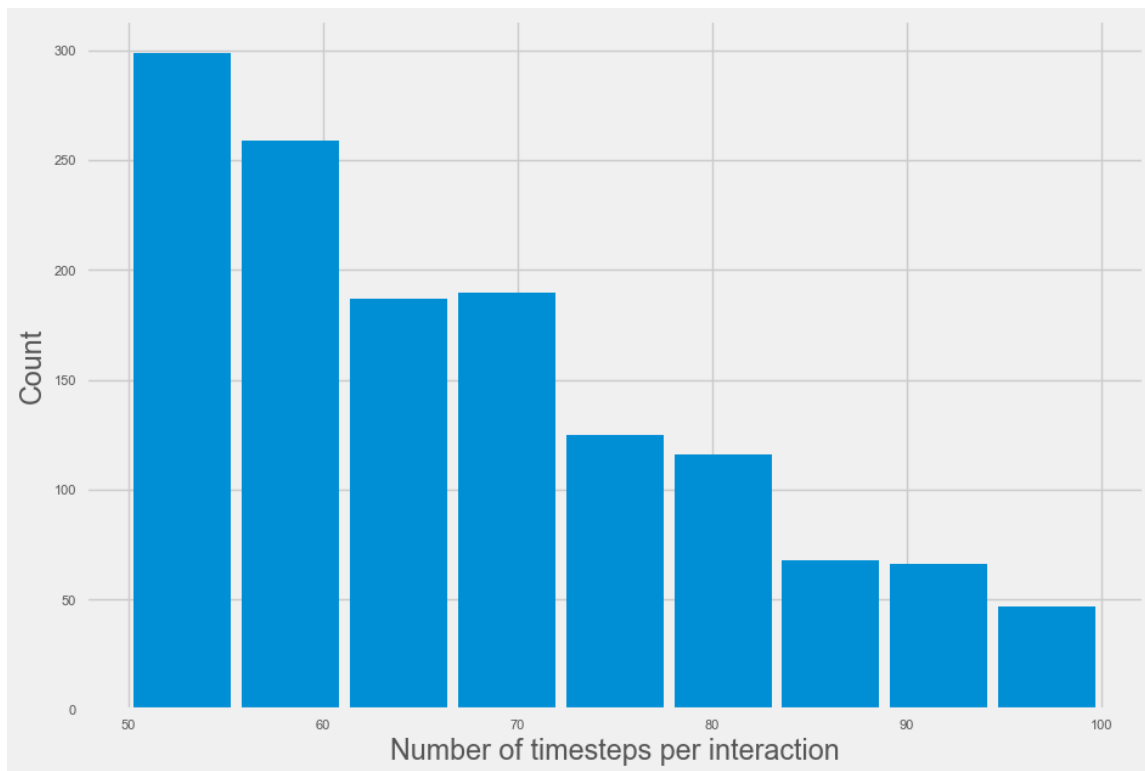
- Ο φάκελος *raw_data*, ο οποίος εμπεριέχει μια σειρά από αρχεία στα οποία είναι καταγεγραμμένα τα ανεπεξέργαστα δεδομένα που συγκεντρώθηκαν από κάθε τεχνολογία μέτρησης που ήταν εγκατεστημένη στο όχημα.
- Ο φάκελος *traj_data*, ο οποίος εμπεριέχει δύο αρχεία μορφής json, το *valtest.json* και το *train.json*.

Τα αρχεία εντός του φακέλου *traj_data* επεξεργάστηκαν με τον κατάλληλο τρόπο ώστε να μετατραπούν σε μια σειρά από αρχεία μορφής csv. Συγκεκριμένα, κάθε σκηνή αλληλεπιδράσεων αποτέλεσε έναν φάκελο, μέσα στον οποίο υπάρχει ένα αρχείο csv για κάθε αλληλεπίδραση. Εντός καθενός από αυτά τα αρχεία είναι καταγεγραμμένες οι συντεταγμένες του αυτόνομου οχήματος και του πεζού (οι αλληλεπιδράσεις με ποδηλάτες διαγράφηκαν καθώς δεν σχετίζονται με το θέμα της εργασίας), κάθε 0.1 δευτερόλεπτα από την αρχή της αλληλεπίδρασης. Συνολικά οι σκηνές είναι 960, με 6163 επιμέρους αλληλεπιδράσεις να είναι διαθέσιμες.

Όπως αναφέρεται και στην παράγραφο 2.2.2, υπάρχουν οι κατηγορίες βραχυπρόθεσμων (έως δύο δευτερόλεπτα) και μακροπρόθεσμων προβλέψεων (5-8 δευτερόλεπτα). Αν και οι εφαρμοζόμενες τεχνικές είναι σε θέση να παράγουν πρότυπα μακροπρόθεσμων προβλέψεων, ο περιορισμένος όγκος διαθέσιμων δεδομένων οδήγησε τη παρούσα εργασία στον υπολογισμό βραχυπρόθεσμων προβλέψεων. Συγκεκριμένα, αποφασίστηκε πρόβλεψη έως δύο δευτερόλεπτα στο μέλλον από κάθε χρονικό βήμα, για την οποία απαιτούνται τρία δευτερόλεπτα μέχρι το βήμα αυτό. Εν ολίγοις απαιτούνται 30 χρονικά βήματα πίσω, για να γίνει πρόβλεψη της θέσης στα επόμενα 20 χρονικά βήματα.

Είναι εμφανές από τα παραπάνω ότι για να είναι εφαρμόσιμη μια αλληλεπίδραση στην εκπαίδευση και επαλήθευση καθενός από τα πρότυπα, θα πρέπει να αποτελείται από τουλάχιστον 50 χρονικά βήματα (5 δευτερόλεπτα). Επομένως όσες αλληλεπιδράσεις δεν πληρούν αυτήν την προϋπόθεση διαγράφηκαν από τα δεδομένα. Έτσι, ο τελικός αριθμός διαθέσιμων αλληλεπιδράσεων είναι $N = 1443$ με μέσο αριθμό χρονικών

βημάτων $\mu = 69.93$ και τυπική απόκλιση $\sigma = 16.96$. Στο διάγραμμα 4.1 φαίνεται το αντίστοιχο ραβδόγραμμα.



Διάγραμμα 4. 1: Ραβδόγραμμα αριθμού χρονικών βημάτων ανά αλληλεπίδραση

4.2.1 Εξαγωγή χαρακτηριστικών (feature extraction)

Στην παρούσα εργασία εφαρμόζονται επτά βαθιά νευρωνικά δίκτυα. Γενικότερα, αυτού του είδους τα πρότυπα κάνουν ακριβέστερες προβλέψεις όταν εισάγονται σε αυτά και διάφορα επιμέρους χαρακτηριστικά. Στην συγκεκριμένη περίπτωση, οι συντεταγμένες από μόνες τους δεν είναι επαρκείς για να είναι ορθή η πρόβλεψη. Όποτε θα πρέπει από αυτές να εξαχθούν τα ακόλουθα χαρακτηριστικά:

- Ταχύτητα αυτόνομου οχήματος και πεζού (σε m/s). Αποτελεί την κύρια έκφραση των συντεταγμένων και προκύπτει από αυτές.
- Επιτάχυνση (σε m/s^2).
- Διαφορά γωνιών (σε ακτίνια).
- Διαφορά ταχυτήτων (σε m/s).
- Ευκλείδεια απόσταση (σε m).
- Κάθετη και οριζόντια απόσταση από την οπτική του αυτόνομου οχήματος (σε m).
- Κάθετη και οριζόντια απόσταση από την οπτική του πεζού (σε m).

Αρχικά όσον αφορά **την ταχύτητα**, αυτή υπολογίζεται για το αυτόνομο όχημα και τον πεζό από τις συντεταγμένες. Συγκεκριμένα η ταχύτητα στο χρονικό βήμα i κατά x και y αντίστοιχα v_i^j υπολογίζεται ως εξής:

$$v_i^j = \frac{\Delta x_i^j}{\Delta t} = \frac{(x_i - x_{i-1})^j}{0.1}, \text{ όπου } j: x \text{ ή } y \quad (4.1)$$

Έχοντας υπολογίσει τις ταχύτητες των δύο πρακτόρων κατά x και y , υπολογίζεται η **διανυσματική ταχύτητα** u_i και η **γωνία κίνησης** θ_i στο χρονικό βήμα i ως εξής:

$$u_i = \sqrt{v_x^2 + v_y^2} \quad (4.2)$$

$$\theta_i = \arctan\left(\frac{v_i^y}{v_i^x}\right) \quad (4.3)$$

Παραγωγίζοντας την διανυσματική ταχύτητα προκύπτει η **διανυσματική επιτάχυνση** a_i :

$$a_i = \frac{\Delta u_i}{\Delta t} = \frac{u_i - u_{i-1}}{0.1} \quad (4.4)$$

Γνωρίζοντας την γωνία κίνησης του αυτόνομου οχήματος θ_i^{AV} και του πεζού θ_i^{ped} κατά το χρονικό βήμα i , υπολογίζεται η **διαφορά γωνιών κίνησης** $\Delta\theta_i$:

$$\Delta\theta_i = |\theta_i^{AV} - \theta_i^{ped}| \quad (4.5)$$

Αντίστοιχα υπολογίζεται και η **διαφορά διανυσματικών ταχυτήτων** Δu_i :

$$\Delta u_i = |u_i^{AV} - u_i^{ped}| \quad (4.6)$$

Όσον αφορά τις αποστάσεις, για να ληφθεί υπόψιν η χωρική διάσταση των πρακτόρων αποφασίστηκε να θεωρηθούν κύκλοι. Συγκεκριμένα, ο κάθε πεζός θεωρήθηκε ότι καλύπτει έναν κύκλο με διάμετρο 0.7 m ενώ το αυτόνομο όχημα έναν κύκλο με διάμετρο 3.5 m . Η καθεμία από τις υπολογιζόμενες αποστάσεις φαίνονται στην Εικόνα 4.3.

Η **ευκλείδεια απόσταση** d_i^{EUC} υπολογίζεται αφαιρώντας τις ακτίνες των δύο κύκλων από την απόσταση των συντεταγμένων των δύο πρακτόρων, ως εξής:

$$d_i = \sqrt{(x_i^{AV} - x_i^{ped})^2 + (y_i^{AV} - y_i^{ped})^2} \quad (4.7)$$

$$d_i^{EUC} = d_i - 1.75 - 0.35 \quad (4.8)$$

Για τον υπολογισμό των κατακόρυφων και οριζόντιων αποστάσεων πρέπει πρώτα να υπολογιστεί η **κλίση** φ_i της ευθείας που ενώνει τα κέντρα των δύο κύκλων (σχέση 4.9), η οποία λαμβάνει τιμές από $-\pi$ ως π . Αφού η αρνητικές τιμές αναχθούν σε θετικές, υπολογίζεται η **γωνία** z_i , η οποία ισούται με την διαφορά μεταξύ της γωνίας κίνησης του εκάστοτε πράκτορα και της γωνίας φ_i (σχέση 4.10), αναγόμενη στο 1^ο τεταρτημόριο (σχέση 4.11). Για τον τελικό υπολογισμό της οριζόντιας απόστασης από την οπτική του πράκτορα j , υπολογίζεται η αντίστοιχη συνιστώσα όπως φαίνεται στην Εικόνα 4.3 (σχέση 4.12). Αναλόγως υπολογίζεται και η κατακόρυφη απόσταση (σχέση 4.13).

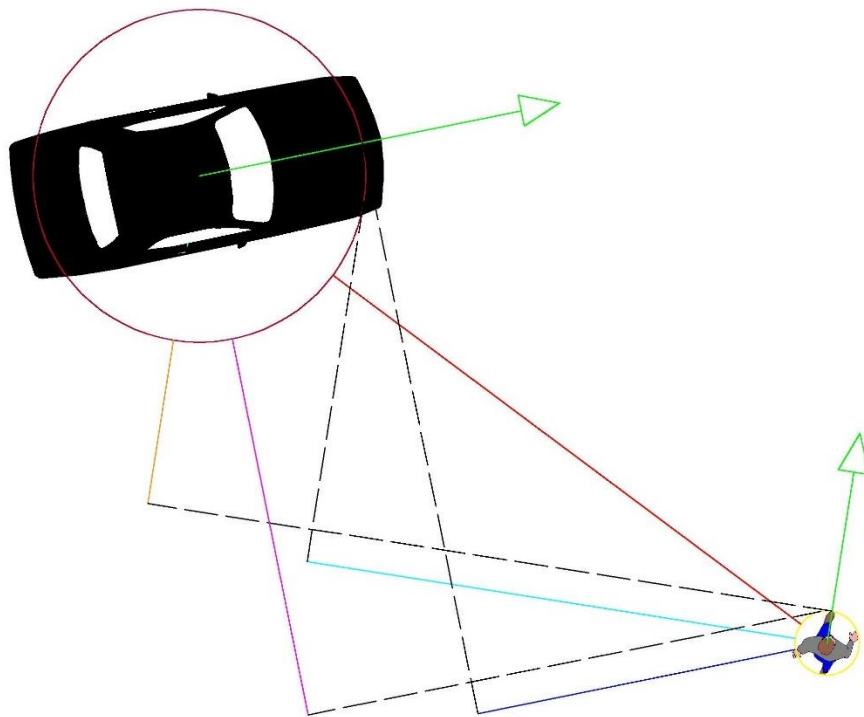
$$\varphi_i = \arctan \left(\frac{\Delta y_i}{\Delta x_i} \right) \quad (4.9)$$

$$z_i = |\varphi_i - \theta_i| \quad (4.10)$$

$$z'_i = \begin{cases} z_i, & \text{για } 0 \leq z_i < \pi/2 \\ \pi - z_i, & \text{για } \pi/2 \leq z_i < \pi \\ z_i - \pi, & \text{για } \pi \leq z_i < 3\pi/2 \\ 2\pi - z_i, & \text{για } 3\pi/2 \leq z_i < 2\pi \end{cases} \quad (4.11)$$

$$d_{lat,j}^i = \sin(z'_{i,j}) \cdot d_i - 1.75 - 0.35 \quad (4.12)$$

$$d_{long,j}^i = \cos(z'_{i,j}) \cdot d_i - 1.75 - 0.35 \quad (4.13)$$



- Αυτόνομο όχημα
- Πεζός
- Διανύσματα ταχυτήτων
- Ευκλείδεια απόσταση
- Οριζόντια απόσταση από την οπτική του αυτόνομου οχήματος
- Οριζόντια απόσταση από την οπτική του πεζού
- Κατακόρυφη απόσταση από την οπτική του αυτόνομου οχήματος
- Κατακόρυφη απόσταση από την οπτική του πεζού

Εικόνα 4. 2: Εξαγόμενες Αποστάσεις

4.2.2 Αφαίρεση σφαλμάτων στα δεδομένα

Στην πλειοψηφία των δεδομένων που συγκεντρώνονται στηριζόμενα σε οπτικές μεθόδους, υπάρχουν ορισμένα σφάλματα στα δεδομένα τα οποία οφείλονται σε ατέλειες των συστημάτων μέτρησης. Στο Euro-PVI dataset εντοπίστηκαν τρεις κατηγορίες σφαλμάτων, τα οποία στην συνέχεια απαλείφθηκαν.

Η πρώτη κατηγορία σφαλμάτων αφορά ορισμένες περιπτώσεις που τα χρονικά βήματα δεν ήταν τοποθετημένα με ορθή σειρά. Στην περίπτωση αυτή τα χρονικά βήματα τοποθετήθηκαν στη σωστή σειρά με βάση τον δείκτη κάθε γραμμής στο σχετικό αρχείο.

Η δεύτερη κατηγορία σφαλμάτων αφορά κυρίως τους πεζούς. Συγκεκριμένα, όταν ο πεζός -και πιο σπάνια το αυτόνομο όχημα - σε μια αλληλεπίδραση είναι ακίνητος για δύο συνεχή χρονικά βήματα $i, i + 1$, θα έπρεπε να καταγράφονται οι ίδιες συντεταγμένες στις δύο αυτές χρονικές στιγμές. Όμως λόγω των μικρών ατελειών του αλγορίθμου καταγραφής των συντεταγμένων, κάτι τέτοιο δεν συμβαίνει. Έτσι, σε κάθε περίπτωση όπου η διαφορά στη διανυσματική ταχύτητα είναι μικρότερη από 0.3 (σχέση 4.14), οι συντεταγμένες ταυτίστηκαν και η ταχύτητα (και επομένως και η επιτάχυνση) μηδενίστηκαν.

$$\Delta u_i = |u_i - u_{i-1}| > 0.3 \text{ m/s} \quad (4.14)$$

Ένα επιπλέον μέτρο διασφάλισης ότι η εν λόγω κατηγορία σφαλμάτων θα απαλειφθεί ήταν η τοποθέτηση των ίδιων συντεταγμένων στο επόμενο χρονικό βήμα, στην περίπτωση όπου η διανυσματική απόσταση μεταξύ των δύο χρονικών βημάτων είναι μικρότερη από 0.86m για το αυτόνομο όχημα και 0.3m για τον πεζό. Ύστερα, στις περιπτώσεις όπου για έναν πράκτορα εντοπίστηκε μηδενική ταχύτητα τουλάχιστον στο 85% των χρονικών βημάτων, αυτή μηδενίστηκε για το σύνολο της αλληλεπίδρασης, καθώς θεωρήθηκε ότι ο πράκτορας παραμένει ακίνητος καθ' όλη τη διάρκεια αυτής.

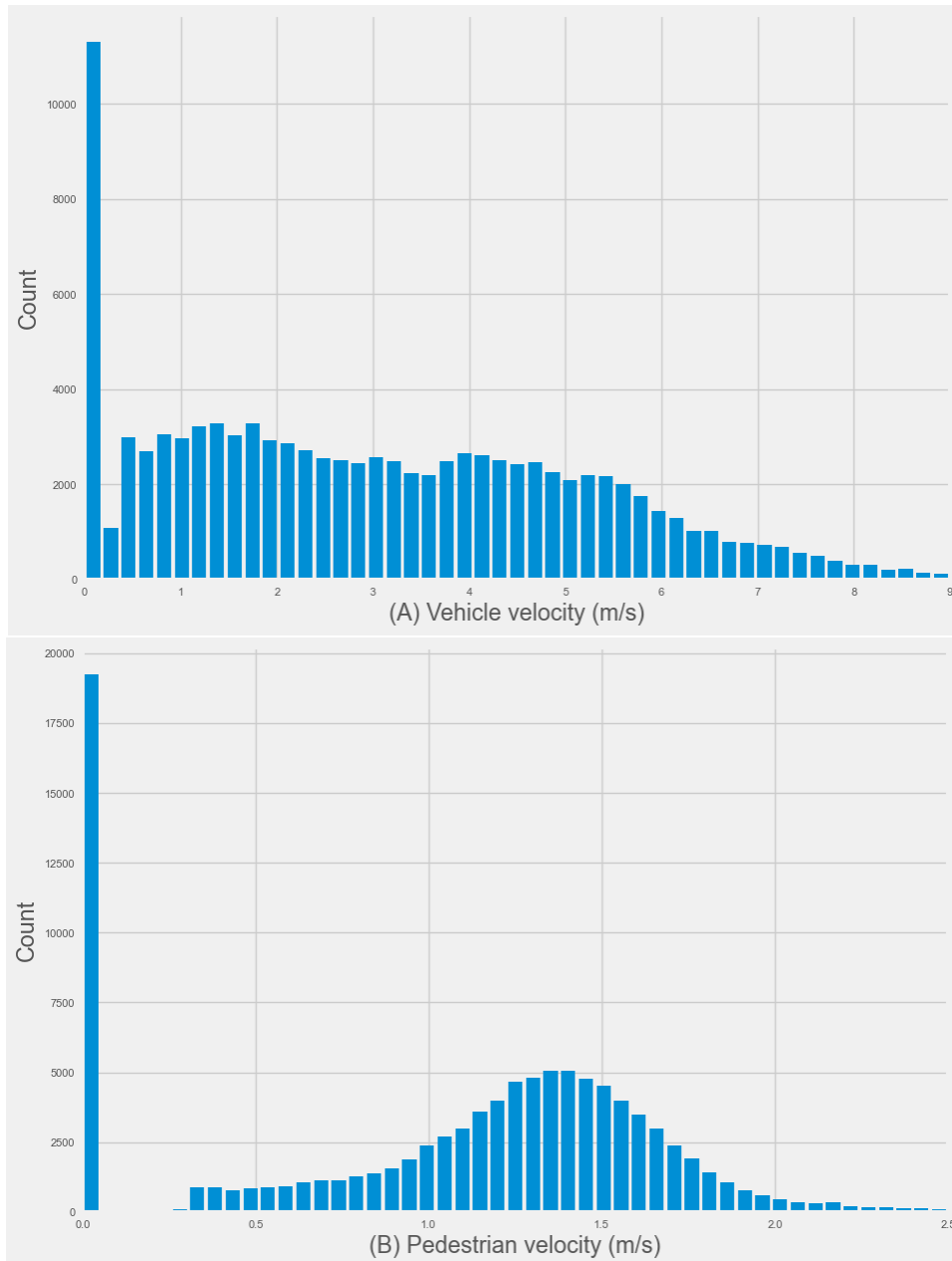
Η τρίτη και τελευταία κατηγορία σφαλμάτων αφορά την μέγιστη ταχύτητα κίνησης. Σε ορισμένες αλλά πολύ περιορισμένες περιπτώσεις παρατηρήθηκαν ιδιαιτέρως μεγάλες ταχύτητες. Στις περιπτώσεις όπου η ταχύτητα του οχήματος ξεπέρασε τα 72 km/h ή η ταχύτητα του πεζού ξεπέρασε τα 28 km/h (8 m/s), οι αλληλεπιδράσεις διαγράφηκαν.

4.2.3 Περιγραφή δεδομένων

Στα Διαγράμματα 4.2-4.5 είναι ορατά τα ραβδογράμματα για 9 από τις κατηγορίες χαρακτηριστικών που εξάχθηκαν, με κύριο σκοπό την ποσοτική περιγραφή του τύπου των αλληλεπιδράσεων του Euro-PVI dataset, ύστερα από την αφαίρεση των σφαλμάτων.

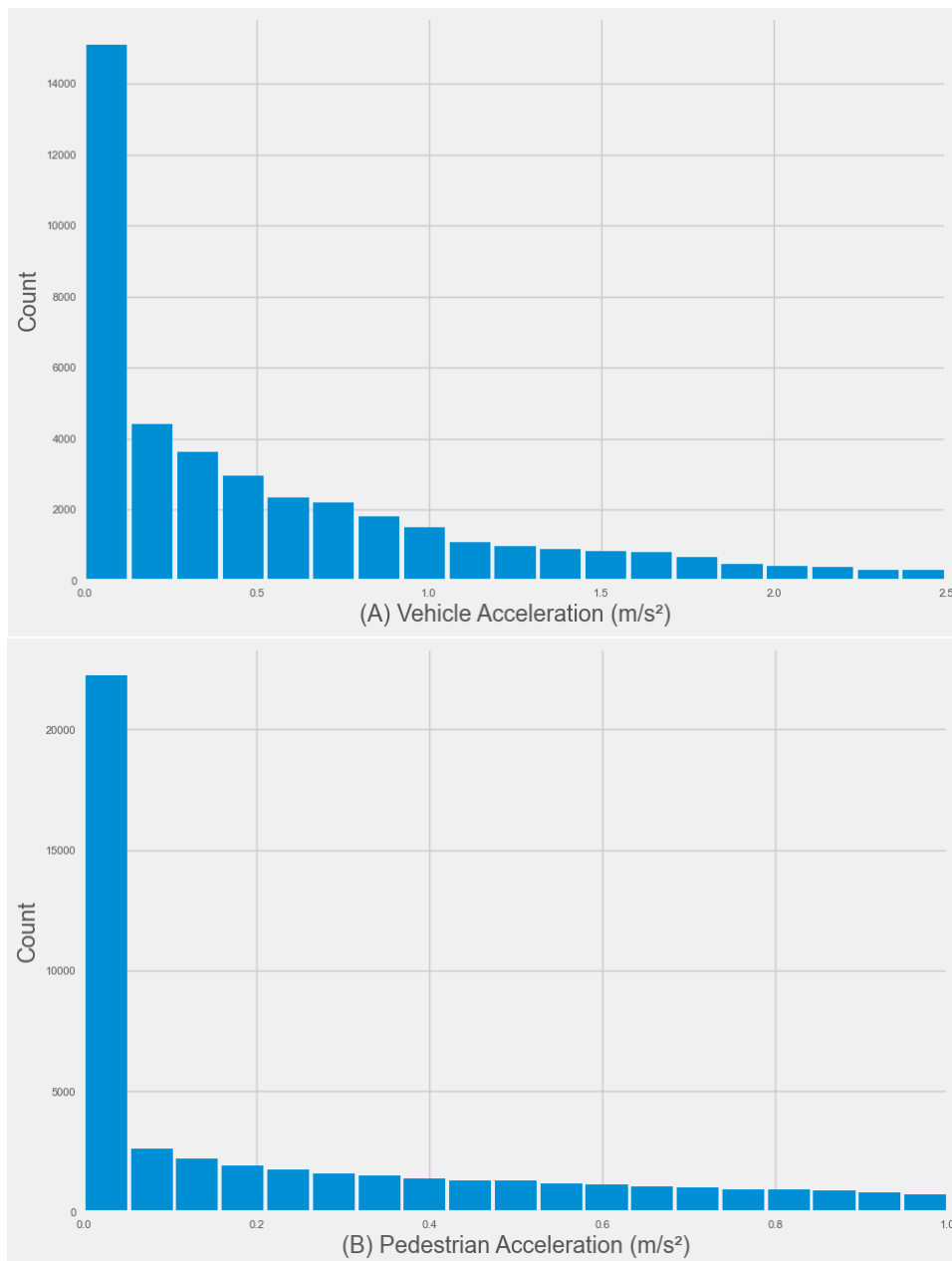
Στο διάγραμμα 4.2 φαίνονται τα ραβδογράμματα της διανυσματικής ταχύτητας (A) των αυτόνομων οχημάτων και (B) των πεζών. Και στις δύο περιπτώσεις φαίνεται ότι σε ένα σημαντικό αριθμό αλληλεπιδράσεων οι πράκτορες παραμένουν ακίνητοι. Εφόσον οι αλληλεπιδράσεις καταγράφονται σε πυκνό αστικό ιστό αυτό είναι απολύτως φυσιολογικό, καθώς ο ένας από τους δύο πράκτορες σχεδόν πάντα χρειάζεται να ακινητοποιηθεί να κινείται για να μπορεί να περάσει ο άλλος. Στην περίπτωση των οχημάτων, με εξαίρεση τη μηδενική ταχύτητα, οι περισσότερες τιμές κυμαίνονται

μεταξύ 0.5 και 5 m/s, κάτι απολύτως λογικό δεδομένων των συνθηκών μέτρησης. Οι πεζοί από την άλλη, με εξαίρεση την περίπτωση της μηδενικής ταχύτητας όπου υπάρχει εξαιρετικά μεγάλη συχνότητα, παρουσιάζουν μια κατανομή που προσεγγίζει εντυπωσιακά πολύ την «καμπάνα» της κανονικής κατανομής, με την μέση τιμή να είναι περίπου ίση με 1.2 m/s, μια τυπική ταχύτητα κίνησης πεζού.



Διάγραμμα 4. 2: Ραβδόγραμμα κατανομής διανυσματικής ταχύτητας (A) αυτόνομου οχήματος και (B) πεζών

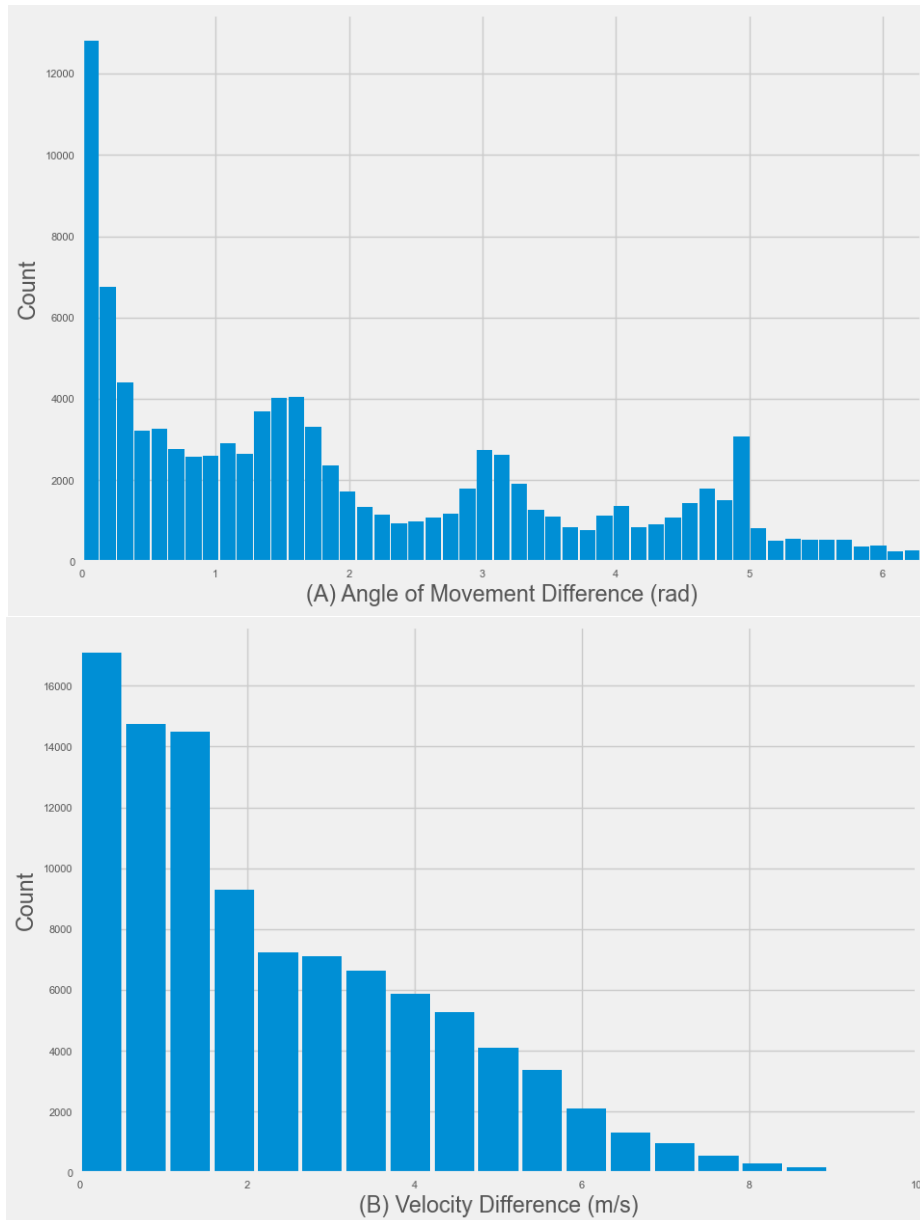
Στο διάγραμμα 4.3 φαίνεται η κατανομή της επιτάχυνσης των οχημάτων και των πεζών, η οποία είναι απολύτως συμβατή με τα παραπάνω.



Διάγραμμα 4. 3: Ραβδόγραμμα κατανομής επιτάχυνσης (A) αυτόνομου οχήματος και (B) πεζών

Όσον αφορά τις διαφορές σε ταχύτητα και γωνία κίνησης, η κατανομή τους φαίνεται στο διάγραμμα 4.4. Η διαφορά σε γωνία κίνησης μεταξύ των δύο πρακτόρων λαμβάνει τιμές από το 0 έως το 2π. Παρατηρείται μια μειούμενη περιοδικότητα, με μέγιστη τιμή με διαφορά το 0. Κάτι τέτοιο δεν είναι ιδιαίτερα αναμενόμενο δεδομένου του ορισμού των αλληλεπιδράσεων και ενδεχομένως να οφείλεται στο ότι ορισμένες φορές οι δύο πράκτορες κινούνται παράλληλα στην προσπάθεια του πεζού να φτάσει σε διάβαση πεζών καθώς και στο ότι ένας ακίνητος πράκτορας έχει εξ ορισμού γωνία ίση με 0. Από

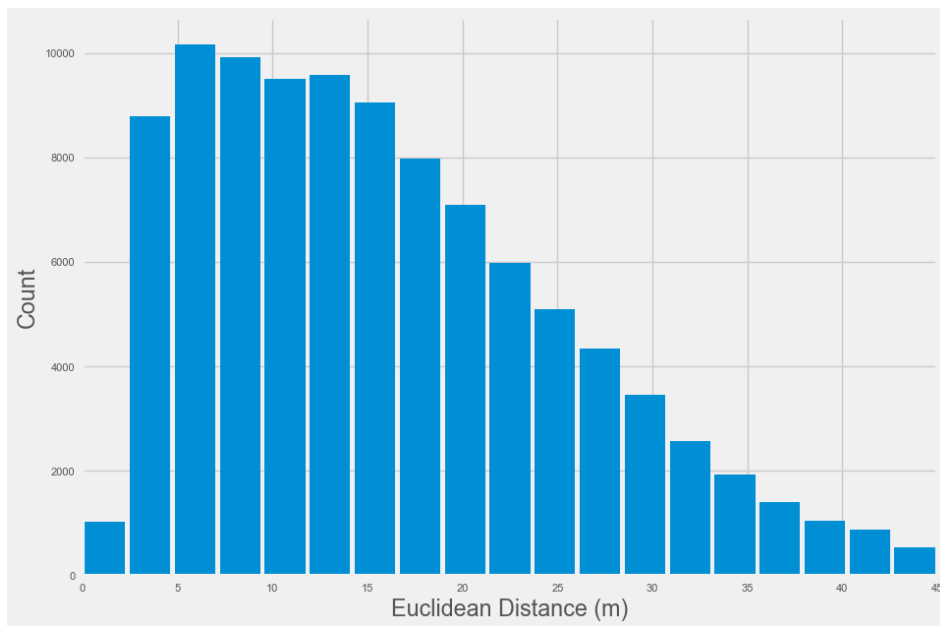
την άλλη, η διαφορά ταχύτητας είναι μειούμενη σχετικά ομαλά, με μεγαλύτερη συχνότητα στις τιμές από 0 έως 2. Αυτό κατά πάσα πιθανότητα οφείλεται στο ότι οι μετρήσεις έγιναν σε αστικές περιοχές με μεγάλη κίνηση πεζών και επομένως χαμηλές ταχύτητες κίνησης για τα αυτόνομα οχήματα.



Διάγραμμα 4. 4: Ραβδόγραμμα κατανομής της διαφοράς (Α) γωνιών κίνησης και (Β) ταχυτήτων μεταξύ αυτόνομου οχήματος και πεζού

Τέλος ως ένδειξη των αποστάσεων μεταξύ των πρακτόρων κατασκευάστηκε το ραβδόγραμμα κατανομής της ευκλείδειας απόστασης μεταξύ των πρακτόρων (Εικόνα 4.7). Οι αποστάσεις είναι αυξανόμενες από τα 0m έως τα 7.5m. Στην συνέχεια από τα 7.5m έως τα 17.5m οι συχνότητες μεγιστοποιούνται. Ύστερα μειώνονται σταθερά μέχρι να μηδενισθούν περίπου στα 40m. Μια τέτοια κατανομή είναι φυσιολογική αν και το

εύρος είναι λίγο μεγαλύτερο από το αναμενόμενο. Η μη αμελητέα συχνότητα σε μεγάλες αποστάσεις (20m) ίσως να αναφέρεται σε αρχικά ή τελικά στάδια αλληλεπιδράσεων, όπου οι πράκτορες δεν είχαν πλησιάσει επαρκώς ακόμη.



Διάγραμμα 4. 5: Ραβδόγραμμα κατανομής ευκλείδειας απόστασης

4.3 Μεθοδολογία ανάπτυξης και εκπαίδευσης προτύπων

Στην παρούσα εργασία εφαρμόζονται επτά πρότυπα πρόβλεψης. Όλα τα πρότυπα έχουν αρχιτεκτονική encoder-decoder και περιλαμβάνουν κελιά LSTM. Όσον αφορά τα εξαγόμενα χαρακτηριστικά, καθένα από τα πρότυπα εφαρμόστηκε με και χωρίς αυτά, για να διαπιστωθεί ο βαθμός επιρροής τους. Σε όλα τα πρότυπα ως συνάρτηση απώλειας χρησιμοποιήθηκε η απώλεια Huber (παράγραφος 3.4). Επιπλέον ως μετρικές αξιολόγησης χρησιμοποιήθηκαν το μέσο απόλυτο σφάλμα (MAE) και το μέσο τετραγωνικό σφάλμα (MSE) (παράγραφος 3.5). Τα δεδομένα, όπως είθισται στην βαθιά μάθηση χωρίστηκαν σε τρία μέρη:

- Στο training set, το οποίο αποτέλεσε το 70% των διαθέσιμων χρονοσειρών 50 χρονικών βημάτων και χρησιμοποιήθηκε για την εκπαίδευση των προτύπων.
- Το validation set, το οποίο αποτέλεσε το 15% των διαθέσιμων χρονοσειρών και χρησιμοποιήθηκε για επαλήθευση κατά την εκπαίδευση των προτύπων.
- Το testing set, το οποίο επίσης αντιστοιχεί στο υπόλοιπο 15% των χρονοσειρών και χρησιμοποιήθηκε για την δοκιμή των προτύπων.

Όλα τα πρότυπα εκπαιδεύτηκαν για 100 εποχές (epochs), υπό την εφαρμογή πρόωρης διακοπής (early stopping), όταν η συνάρτηση απώλειας και οι μετρικές σταματήσουν να βελτιώνονται. Όσον αφορά τον ρυθμό εκπαίδευσης (learning rate), αποτελεί ένα μέτρο του ρυθμού μεταβολής των παραμέτρων ενός προτύπου κατά την διάρκεια της εκπαίδευσης. Είναι πολύ σημαντικό να λαμβάνει κατάλληλη τιμή καθώς η επιλογή προβληματικής τιμής μπορεί να οδηγήσει σε λανθασμένα πρότυπα (είτε

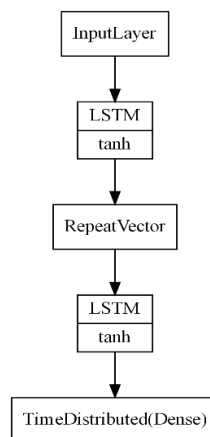
υπερκπαιδευμένα είτε υπόεκπαιδευμένα). Στην συγκεκριμένη περίπτωση, για όλα τα πρότυπα επιλέγεται να εφαρμοστεί ρυθμιστής ρυθμού εκπαίδευσης (learning rate scheduler). Συγκεκριμένα επιλέγεται ρυθμιστής που εκφράζεται από την εξίσωση 4.15, όπου ως i ορίζεται η i -οστή εποχή. Η εξίσωση αυτή χαρακτηρίζεται ως εκθετικά αποσυντιθέμενης μορφής (exponential decay) και είναι ιδιαίτερως αποτελεσματική για αυτού του είδους προβλήματα.

$$lr_i = 10^{-3} \cdot 0.9^{lr_{i-1}} \quad (4.15)$$

Σε επίπεδο βελτιστοποίησης των παραμέτρων των προτύπων επιλέχθηκε ο βελτιστοποιητής (optimizer) Adam. Ο εν λόγω βελτιστοποιητής είναι ευρέως αποδεκτό ότι είναι ένας από τους καλύτερους. Αυτό διότι διατηρεί διαφορετικούς ρυθμούς εκπαίδευσης ανά παράμετρο του πρότυπου, οι οποίοι προσαρμόζονται με βάση τον μέσο των πρόσφατων μεγεθών (magnitudes) των κλίσεων σε σχέση με τα βάρη. Επιπλέον, μεταξύ διαφόρων άλλων βελτιστοποιητών που δοκιμάστηκαν, ο Adam προσέδιδε τα πιο ποιητικά πρότυπα.

4.3.1 Πρότυπο LSTM encoder-decoder (LSTM_ed)

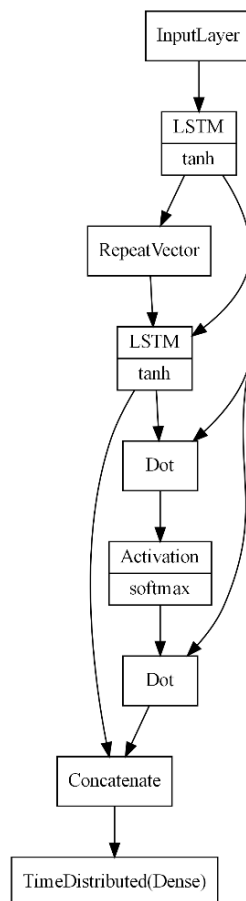
Το πρώτο και ίσως απλούστερο πρότυπο είναι έχει την μορφή encoder-decoder με δύο επίπεδα LSTM. Εν συντομία θα αναφέρεται ως LSTM_ed (_wf με τα εξαγόμενα χαρακτηριστικά και _nf χωρίς αυτά). Η αρχιτεκτονική του φαίνεται στην Εικόνα 4.16. Εντός των δύο κελιών LSTM χρησιμοποιείται ως συνάρτηση ενεργοποίησης η υπερβολική εφαπτομένη. Το ενδιάμεσο επίπεδο RepeatVector δεν κάνει τίποτα παραπάνω από επανάληψη της εισόδου του και μετατροπή της από διάσταση (N, 150) σε (N, 20, 150), ουσιαστικά λαμβάνοντας τον ρόλο του διανύσματος encoder. Το τελικό επίπεδο TimeDistributed(Dense) μετατρέπει την είσοδο του καταλλήλως έτσι ώστε να λάβει την μορφή (N, 20, 4), όπου N είναι ο αριθμός των προβλέψεων, 20 είναι ο αριθμός των προβλεπόμενων χρονικών βημάτων ανά πρόβλεψη και 4 είναι ο αριθμός των χαρακτηριστικών που προβλέπονται (οριζόντιες και κατακόρυφες ταχύτητες αυτόνομου οχήματος και πεζών). Αποτελείται από 279.604 παραμέτρους (274.204 χωρίς τα χαρακτηριστικά), τις λιγότερες από όλα τα πρότυπα. Η δομή του συνοψίζεται στον Πίνακα Π.1 του παραρτήματος.



Εικόνα 4. 3: Αρχιτεκτονική πρότυπου LSTM_ed

4.3.2 Πρότυπο LSTM encoder-decoder με έμφαση (LSTM_ed_attention)

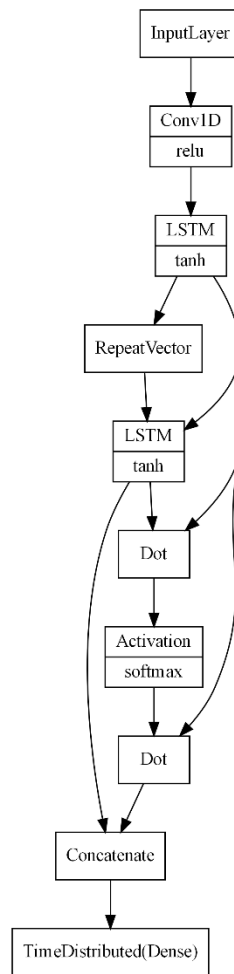
Το πρότυπο LSTM_ed_attention διαφοροποιείται με το LSTM_ed στο ότι εφαρμόζεται η τεχνική της έμφασης (attention). Αποτελείται από δύο επίπεδα LSTM (ένα στο encoder και ένα στο decoder) με χρήση της συνάρτησης της υπερβολικής εφαιπτομένης όπως και παραπάνω. Το επίπεδο Activation μαζί με τα δύο επίπεδα Dot πίσω από αυτό εφαρμόζουν τον μηχανισμό έμφασης. Στο πρώτο επίπεδο Dot υπολογίζεται το εσωτερικό γινόμενο μεταξύ της στοίβας encoder και της στοίβας decoder (encoder and decoder stacks), με αυτό τον τρόπο ανάγοντας την έξοδο του δεύτερου κελιού LSTM, που είναι της μορφής (N, 20, 150) σε (N, 20, 30), όπου 30 είναι τα χρησιμοποιούμενα χρονικά βήματα. Στην συνέχεια, το Activation επίπεδο εφαρμόζει την συνάρτηση softmax, η οποία μετατρέπει την είσοδο σε κατανομή πιθανοτήτων των ίδιων διαστάσεων. Το δεύτερο επίπεδο Dot υπολογίζει το εσωτερικό γινόμενο της στοίβας encoder με την έξοδο του προηγούμενου επιπέδου. Το επίπεδο concatenate, όπως γίνεται αντιληπτό και από το όνομα του, συνδέει σειριακά όλες του τις εισόδους. Το εν λόγω πρότυπο αποτελείται από 280.204 παραμέτρους (274.804 χωρίς τα χαρακτηριστικά), δηλαδή ελάχιστα περισσότερες από το LSTM_ed. Η δομή του συνοψίζεται στον Πίνακα Π.2 του παραρτήματος.



Εικόνα 4. 4: Αρχιτεκτονική πρότυπου LSTM_ed_attention

4.3.3 Πρότυπο CNN-LSTM encoder-decoder με έμφαση (CNN_LSTM_ed_attention)

Το τρίτο εφαρμοζόμενο πρότυπο συνδυάζει ένα επίπεδο CNN με δύο επίπεδα LSTM και εφαρμογή μηχανισμού έμφασης. Το επίπεδο CNN εφαρμόζεται πριν τα υπόλοιπα επίπεδα και είναι μονοδιάστατο, συνδυάζοντας τα χρονικά βήματα κάθε αλληλεπίδρασης ανά δύο. Κατά τα άλλα η αρχιτεκτονική είναι ακριβώς η ίδια με αυτή του LSTM_ed_attention. Εμπεριέχει 311.700 παραμέτρους (311.124 χωρίς τα χαρακτηριστικά), ξεπερνώντας μόνον τα LSTM_ed και LSTM_ed_attention ως προς τον αριθμό τους. Η δομή του συνοψίζεται στον Πίνακα Π.3 του παραρτήματος.

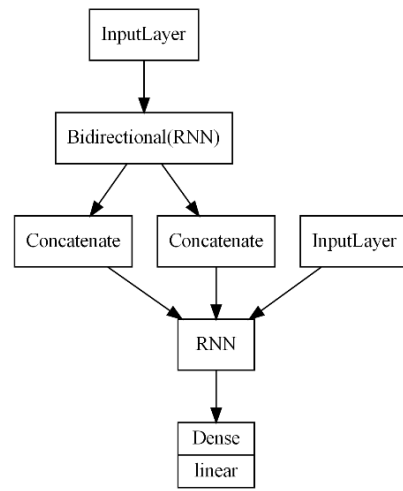


Εικόνα 4. 5: Αρχιτεκτονική πρότυπου CNN_LSTM_ed_attention

4.3.4 Πρότυπο BiLSTM encoder-decoder (BiLSTM_ed)

Το πρότυπο BiLSTM_ed έχει μια σχετικά απλή αρχιτεκτονική. Βασίζεται στο δίκτυο BiLSTM, το οποίο εξελίσσεται και προς τις δύο κατευθύνσεις, όπως περιγράφηκε σε προηγούμενο κεφάλαιο. Με αυτόν τον τρόπο οι σχέσεις μεταξύ των χαρακτηριστικών κάθε χρονικού βήματος «μαθαίνονται» καλύτερα από ότι στην περίπτωση ενός απλού LSTM δικτύου. Το πρώτο BiLSTM επίπεδο ανήκει στον encoder και οι παραγόμενες τους καταστάσεις συνενώνονται σειριακά στα δύο επίπεδα Concatenate. Όσον αφορά το δεύτερο επίπεδο εισόδου, αυτό ουσιαστικά ψευδοεπίπεδο (dummy layer), καθώς αποτελείται από μηδενικά και έχει διαστάσεις ίδιες με αυτές των στοχευμένων

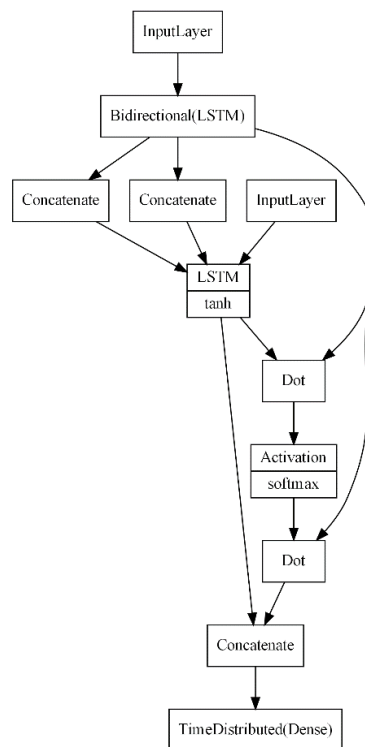
δεδομένων (target data). Εμπεριέχει 413.700 παραμέτρους (404.484 χωρίς τα χαρακτηριστικά) και έτσι κατατάσσεται 4^ο ως προς τον αριθμό τους. Η δομή του συνοψίζεται στον Πίνακα Π.4 του παραρτήματος.



Εικόνα 4. 6: Αρχιτεκτονική πρότυπου BiLSTM_ed

4.3.5 Πρότυπο BiLSTM encoder-decoder με έμφαση (BiLSTM_ed_attention)

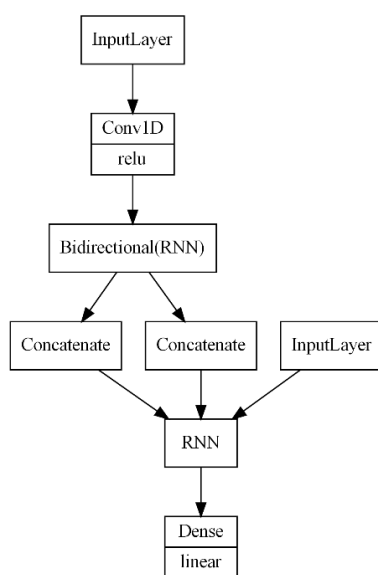
Το πρότυπο BiLSTM_ed_attention συνδυάζει την τεχνική BiLSTM και τον μηχανισμό έμφασης, όπως φαίνεται και στην Εικόνα 4.11. Έτσι εν τέλει αποκτά μια σχετικά πολύπλοκη αρχιτεκτονική, κάτι που δικαιολογεί το γεγονός ότι εμπεριέχει 565.204 παραμέτρους (554.404 χωρίς τα χαρακτηριστικά). Η δομή του συνοψίζεται στον Πίνακα Π.5 του παραρτήματος.



Εικόνα 4. 7: Αρχιτεκτονική πρότυπου BiLSTM_ed_attention

4.3.6 Πρότυπο CNN-BiLSTM encoder-decoder (CNN_BiLSTM_ed)

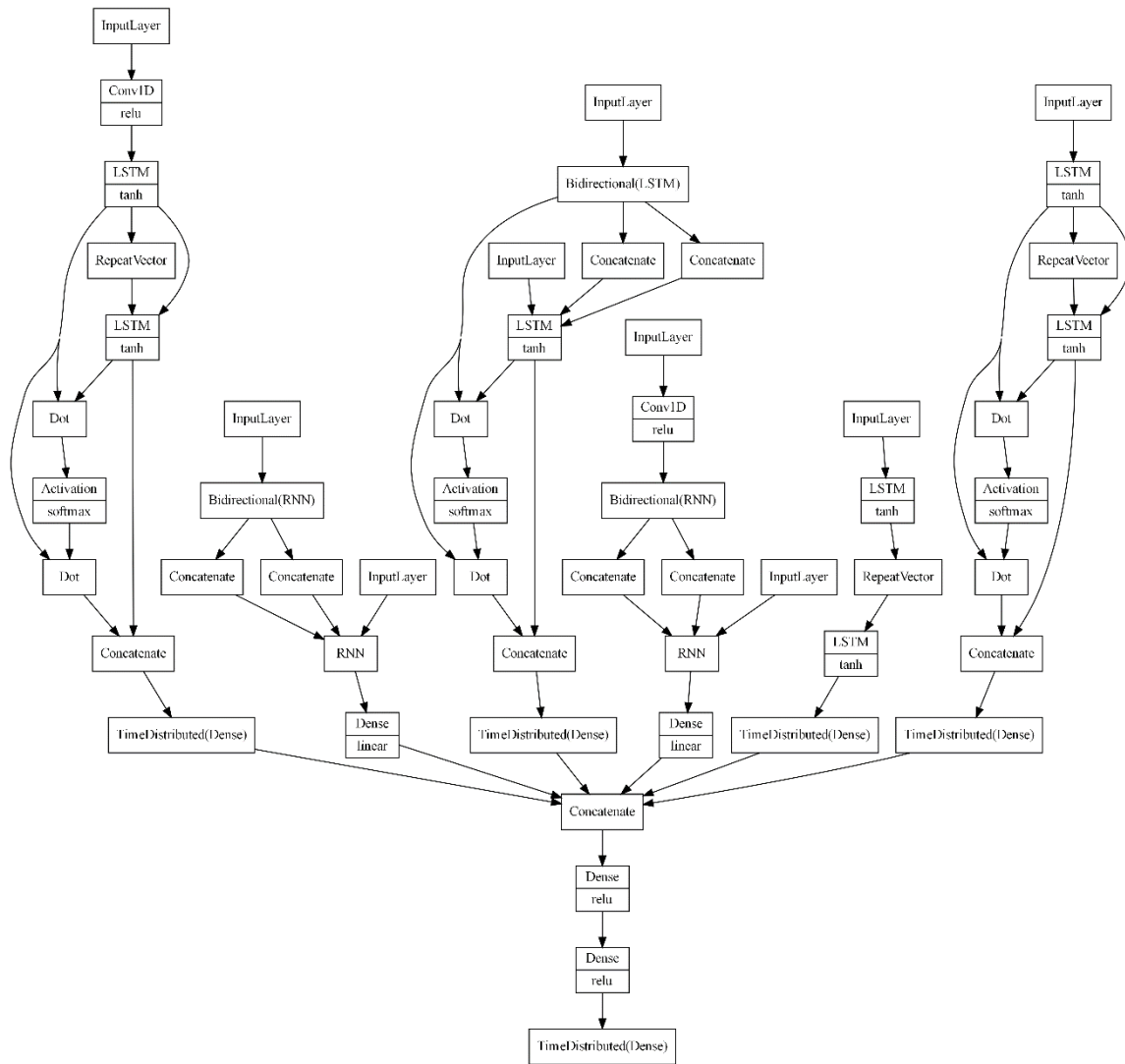
Το πρότυπο CNN_BiLSTM_ed χρησιμοποιεί δύο BiLSTM επίπεδα (ένα στο encoder και ένα στο decoder), σε συνδυασμό με ένα μονοδιάστατο CNN επίπεδο. Το CNN επίπεδο τοποθετείται ακριβώς μετά το επίπεδο εισόδου, όπως φαίνεται στην Εικόνα 4.13. Εμπεριέχει 466.820 παραμέτρους (466.244 χωρίς τα χαρακτηριστικά) και έτσι κατατάσσεται 3^ο ως προς τον αριθμό τους. Η δομή του συνοψίζεται στον Πίνακα Π.6 του παραρτήματος.



Εικόνα 4. 8: Αρχιτεκτονική πρότυπου CNN_BiLSTM_ed

4.3.7 Stacking Ensemble

Το έβδομο και τελευταίο πρότυπο αποτελεί συνδυασμό των προηγούμενων έξι όπως φαίνεται στην Εικόνα 4.14. Η τεχνική stacking ensemble ουσιαστικά σταθμίζει τις προβλέψεις διαφορετικών προτύπων, με σκοπό να δημιουργήσει ένα βέλτιστο πρότυπο πρόβλεψης. Οι έξοδοι των έξι προτύπων συνδέονται σειριακά στο αντίστοιχο επίπεδο. Στην συνέχεια, τα δύο επίπεδα Dense (το ένα για 128 και το άλλο για 64 νευρώνες εισόδου) εφαρμόζουν την συνάρτηση ενεργοποίησης ReLU μεταξύ της εισόδου στο επίπεδο και ενός διανύσματος βαρών το οποίο δημιουργείται από αυτό. Προφανώς, το εν λόγω πρότυπο είναι το πιο περίπλοκο από τα επτά και αυτό εκφράζεται από το γεγονός ότι έχει 2.328.948 παραμέτρους (2.296.980 χωρίς τα χαρακτηριστικά). Η δομή του συνοψίζεται στον Πίνακα Π.7 του παραρτήματος.



Εικόνα 4. 9: Αρχιτεκτονική πρότυπου Stacking Ensemble

ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 Αποτελέσματα ανά πρότυπο

Στο προηγούμενο κεφάλαιο εξηγήθηκε η μεθοδολογία που ακολουθήθηκε όσον αφορά τα επτά εφαρμοζόμενα πρότυπα. Στο παρόν κεφάλαιο θα παρουσιαστούν και σχολιαστούν τα αποτελέσματα των προτύπων. Κύριος σκοπός της εφαρμογής τόσων διαφορετικών προτύπων είναι ανάγκη για μεγιστοποιημένη ακρίβεια στις προβλέψεις. Επιπλέον, όπως έχει ήδη αναφερθεί, τα πρότυπα εκπαιδεύτηκαν με και χωρίς τα χαρακτηριστικά (features) που εξάχθηκαν, για να διαπιστωθεί ο βαθμός επιρροής τους στο τελικό αποτέλεσμα. Για το καθένα από τα επτά πρότυπα, μεταξύ των αποτελεσμάτων της εκπαίδευσης με ή χωρίς τα χαρακτηριστικά, θα παρουσιάζονται μόνον αυτά τα οποία παρουσίασαν καλύτερες (εν προκειμένω μικρότερες) μετρικές στο testing set.

5.1.1 Αποτελέσματα LSTM_ed

Το πρότυπο LSTM_ed χρειάστηκε 50 εποχές για να ελαχιστοποιηθεί η απώλεια Huber (52 χωρίς τα χαρακτηριστικά). Λόγω του γεγονότος ότι σε αλληλεπιδράσεις μεταξύ αυτόνομων οχημάτων και πεζών οι εμπλοκές αποτελούν θέμα διαφοράς εκατοστών, είναι απαραίτητο οι μετρικές MAE και MSE να είναι όσο το δυνατόν ελάχιστες. Όσον αφορά τα χαρακτηριστικά, όπως φαίνεται στον Πίνακα 5.1, βελτιώνουν σε αρκετά σημαντικό βαθμό το πρότυπο, δεδομένου του προβλήματός. Οι αναλυτικές τιμές των μετρικών για όλα τα χρονικά βήματα στον Πίνακα Π.8 του παραρτήματος.

Πίνακας 5. 1: Μέσα MAE και MSE για το LSTM_ed

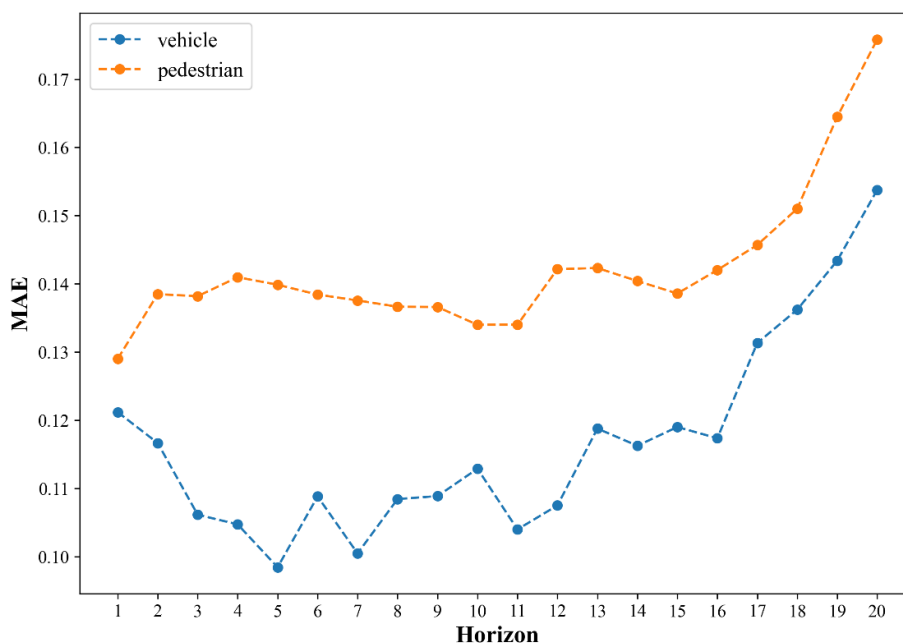
	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
Με χαρακτ.	0.117	0.068	0.190	0.020	0.142	0.074	0.075	0.016
Χωρίς χαρακτ.	0.129	0.078	0.197	0.025	0.160	0.092	0.087	0.028

Στο παρόν πρότυπο όπως φαίνεται και στο Διαγράμματα 5.1, από το 1^ο έως το 15^ο χρονικό βήμα παρουσιάζεται μια σχετική σταθερή εξέλιξη γύρω από την τιμή των 0.135 m/s όσον αφορά το MAE ταχυτήτων για τους πεζούς, με την τιμή του να αυξάνεται σημαντικά από το 16^ο χρονικό βήμα και ύστερα. Το γεγονός αυτό υποδηλώνει ότι από το 16^ο χρονικό βήμα και ύστερα, η ποιότητα του πρότυπου όσον αφορά την ταχύτητα των πεζών αρχίζει να μειώνεται με ταχύ ρυθμό. Το MAE ταχυτήτων για το αυτόνομο όχημα από την άλλη έχει γενικότερα αρκετά μικρότερη τιμή. Αυτό συμβαίνει κυρίως λόγω του γεγονότος ότι ο θόρυβος (noise) των δεδομένων του οχήματος είναι πολύ χαμηλότερος σε σχέση με τους πεζούς, καθώς ο τρόπος κίνησης τους είναι πολύ πιο

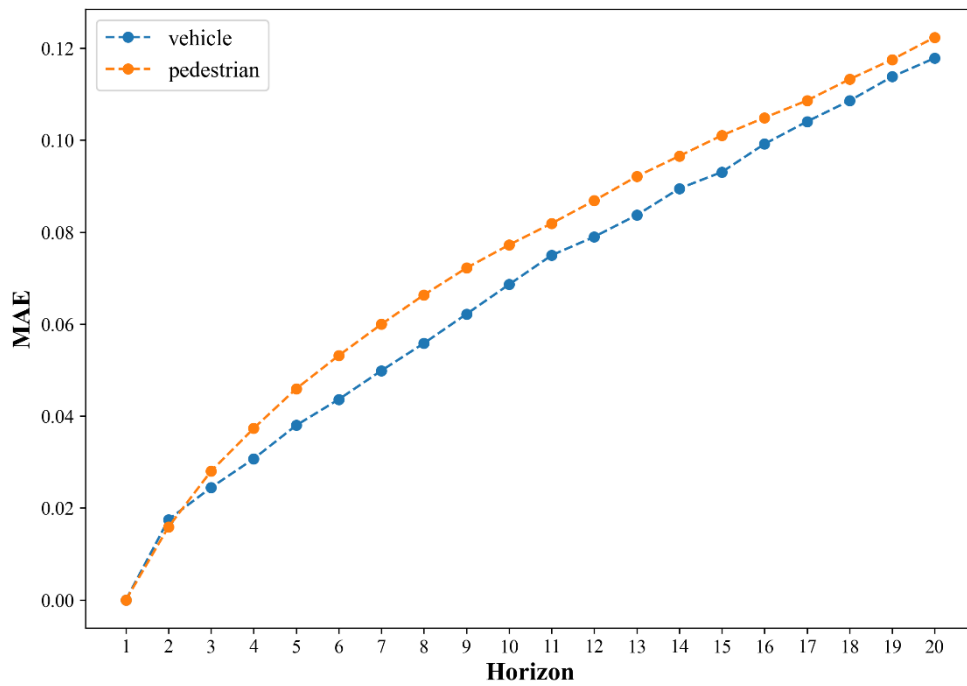
«ρευστός» σε σχέση με ένα όχημα. Παρόλα αυτά, και το MAE ταχυτήτων οχημάτων αποκτά έντονη κλίση από το 17^ο έως το 20^ο χρονικό βήμα.

Στο διάγραμμα 5.2 παρουσιάζεται η εξέλιξη του MAE της θέσης των πρακτόρων στο βάθος 20 χρονικών βημάτων. Σε αντίθεση με τα MAE ταχυτήτων, ο ρυθμός μεταβολής για τα MAE θέσεων είναι σχεδόν σταθερός. Αυτό αποκαλύπτει ότι η σχέση ταχύτητας και θέσης δεν είναι γραμμική. Επιπλέον από το διάγραμμα αυτό μπορεί να διαπιστωθεί ότι ο ουσιαστικός στόχος του πρότυπου, δηλαδή η πρόβλεψη των τροχιών των δύο πρακτόρων, μειώνεται σε ακρίβεια σχεδόν γραμμικά όσο αυξάνεται η χρονική απόσταση.

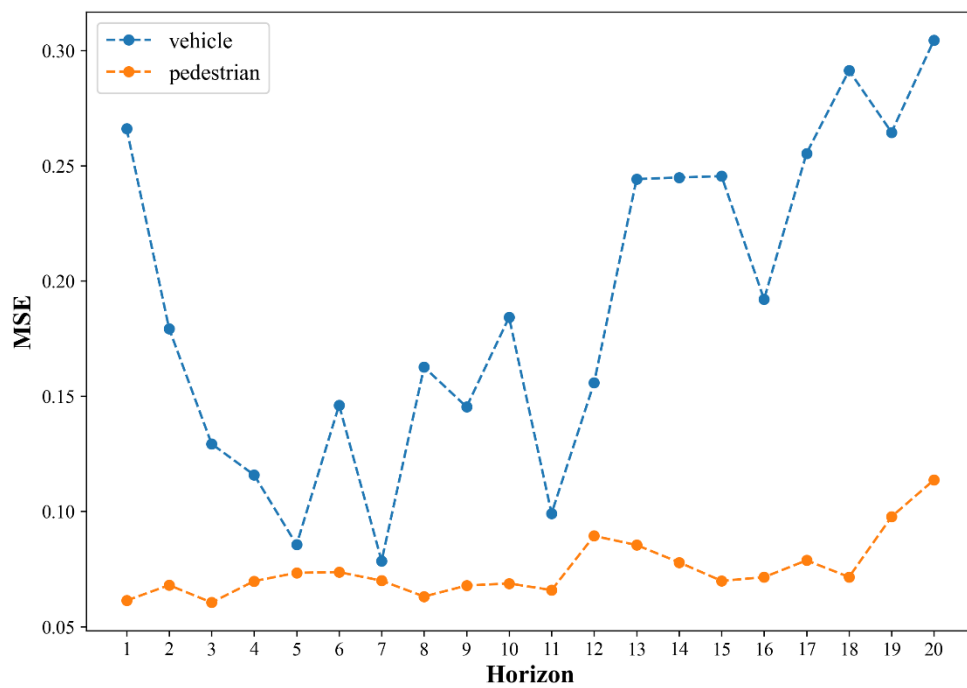
Όσον αφορά το MSE, η διαφορά του με το MAE έγκειται στο ότι λόγω του τετραγωνισμού κάθε διαφοράς μεταξύ της πρόβλεψης και της πραγματικής τιμής, είναι σε θέση να σταθμίσει ακριβέστερα κάθε τέτοια διαφορά. Στο διάγραμμα 5.3 παρουσιάζεται η εξέλιξη του MSE της ταχύτητας στον ορίζοντα των 20 χρονικών βημάτων. Είναι εντυπωσιακό ότι σε αντίθεση με το MAE, το MSE της ταχύτητας των πεζών είναι σχετικά σταθερό κατά και τα 20 χρονικά βήματα. Το γεγονός αυτό δείχνει ότι για το παρόν πρότυπο η διαφορά μεταξύ των προβλέψεων και των πραγματικών τιμών των ταχυτήτων είναι ομοιόμορφη για τους πεζούς. Κάτι τέτοιο δεν συμβαίνει για τα οχήματα, με τις τιμές του MSE να παρουσιάζουν μεγάλες διακυμάνσεις μεταξύ των χρονικών βημάτων. Ο κύριος λόγος που μπορεί να συμβαίνει κάτι τέτοιο είναι ότι ένα αυτόνομο όχημα έχει την ικανότητα να κινείται με πολύ υψηλότερη ταχύτητα και επιτάχυνση από έναν πεζό, και επομένως με πολύ χαμηλότερο επίπεδο προεβλεψιμότητας. Από την άλλη, το MSE θέσεων για τους δύο πράκτορες (διάγραμμα 5.4) παρουσιάζει παρόμοια συμπεριφορά με το MAE θέσεων (διάγραμμα 5.2).



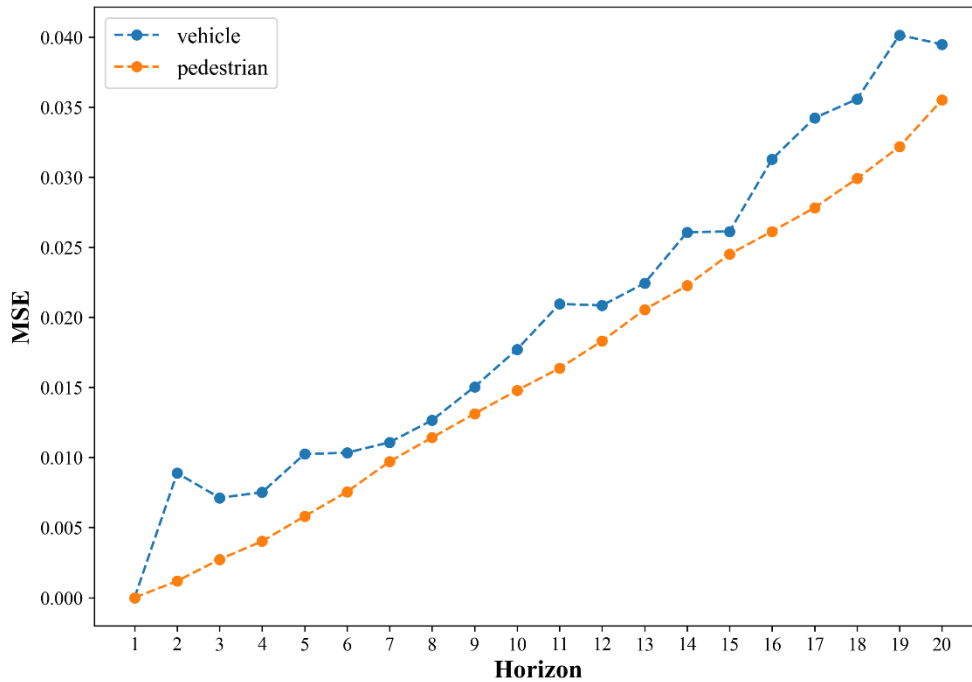
Διάγραμμα 5. 1: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_wf



Διάγραμμα 5. 2: MAE θέσης στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_wf



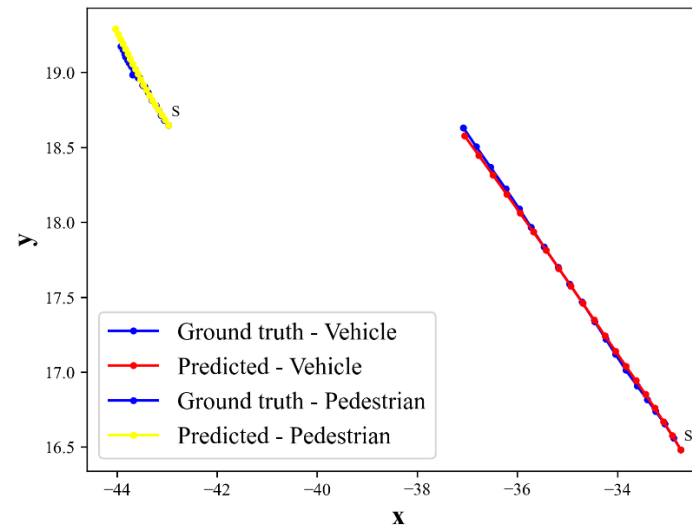
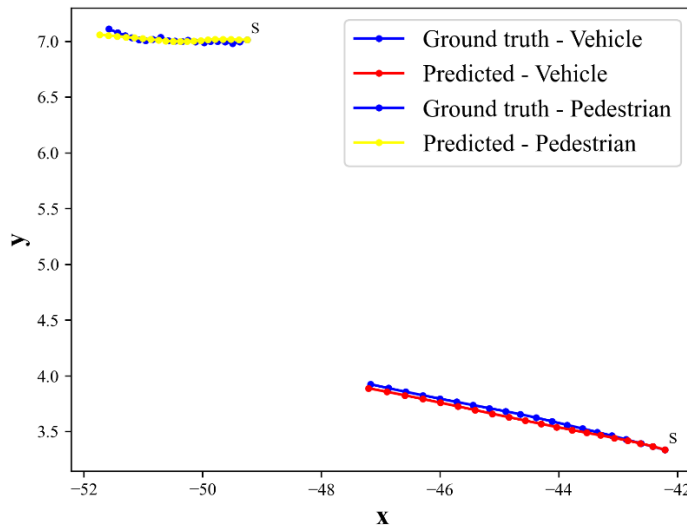
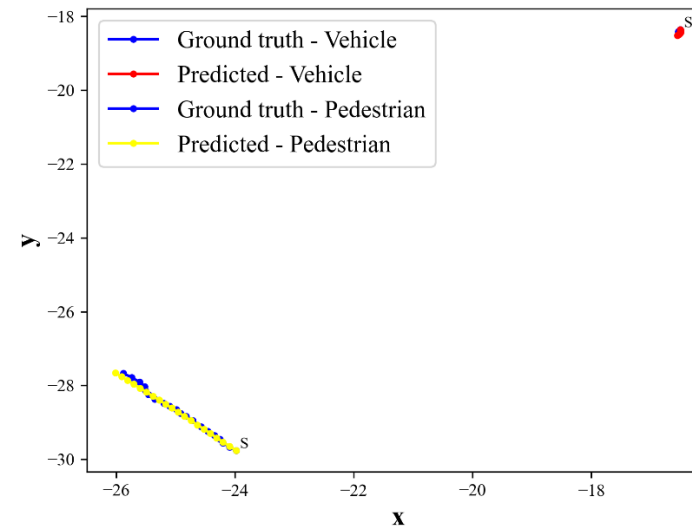
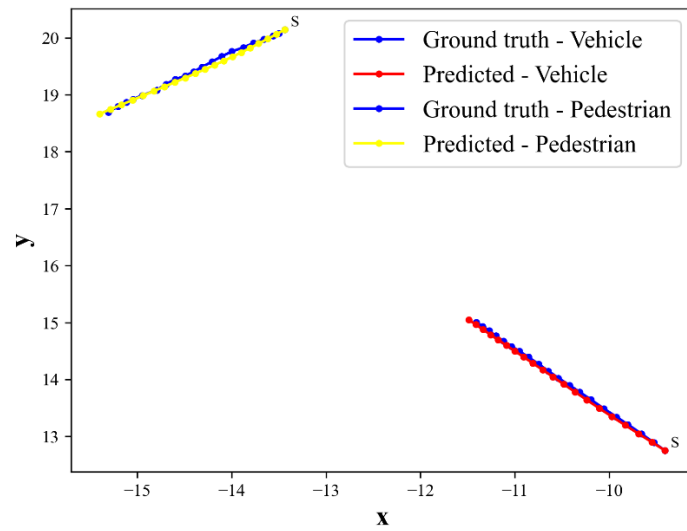
Διάγραμμα 5. 3: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_wf



Διάγραμμα 5. 4: MSE θέσης στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_wf

Η σύγκλιση της απώλειας Huber και των MAE και MSE ανά εποχή για το LSTM_ed απεικονίζεται στα Διαγράμματα Π.1 με Π.3 του παραρτήματος. Η μορφή τους μοιάζει με αυτή της λογαριθμικής συνάρτησης, με τις τιμές τους να σταθεροποιούνται στην τελευταία εποχή. Η αριθμητική εξέλιξη των μέσων MAE και MSE ανά εποχή παρουσιάζεται στον Πίνακα Π.9 του παραρτήματος.

Επιπλέον, στην Εικόνα 5.1 παρουσιάζονται ενδεικτικά οι προβλέψεις τεσσάρων τυχαίων αλληλεπιδράσεων σε σχέση με τις πραγματικές τιμές. Παρόλο που το εν λόγω πρότυπο είναι το απλούστερο που εφαρμόζεται, είναι εντυπωσιακή η ακρίβεια που επιτυγχάνει.



Διάγραμμα 5. 5: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (LSTM_ed_wf)

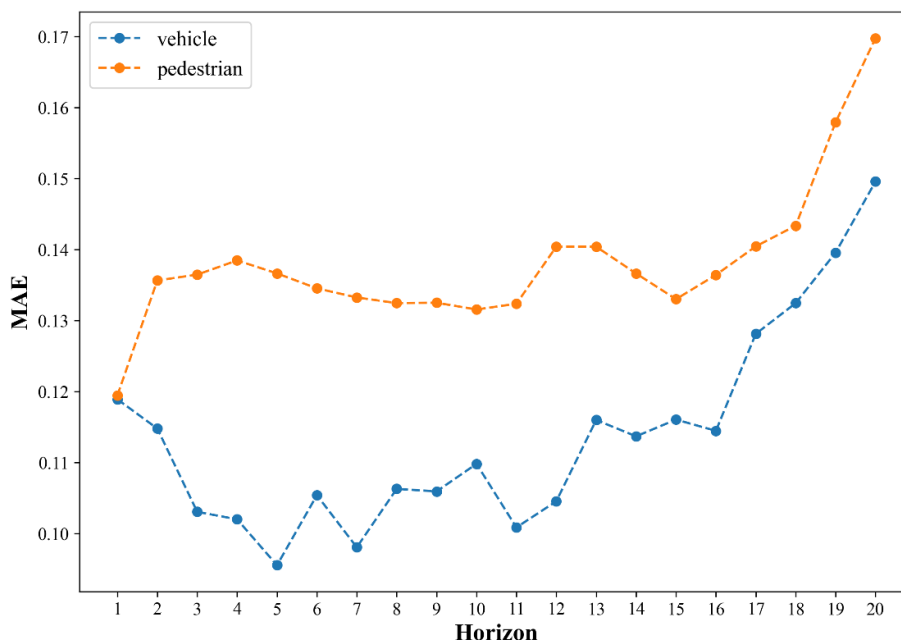
5.1.2 Αποτελέσματα LSTM_ed_attention

Το πρότυπο LSTM_ed_attention διαφοροποιείται με το LSTM_ed στο ότι συμπεριλαμβάνει μηχανισμό έμφασης. Το πρότυπο απαίτησε 50 εποχές για να βελτιστοποιηθεί (και 47 χωρίς τα χαρακτηριστικά, οι λιγότερες από όλα τα πρότυπα), δηλαδή ακριβώς ίσες με το LSTM_ed. Το πρότυπο που συμπεριλαμβάνει τα χαρακτηριστικά (features) έχει ελαφρώς μικρότερα μέσα MAE και MSE, όπως φαίνεται στον Πίνακα 5.2, αποδεικνύοντας έτσι τη χρησιμότητα τους για την επίτευξη ενός βέλτιστου πρότυπου.

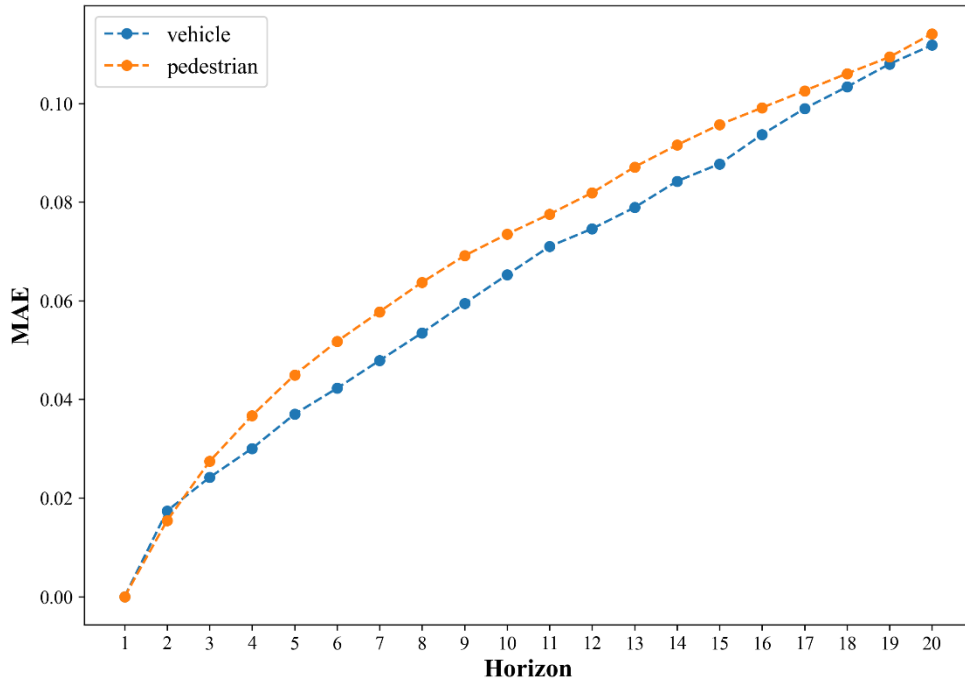
Πίνακας 5. 2: Μέσα MAE και MSE για το LSTM_ed_attention

	Αυτόνομο Οχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
Με χαρακτ.	0.114	0.064	0.188	0.019	0.138	0.07	0.072	0.015
Χωρίς χαρακτ.	0.127	0.077	0.195	0.025	0.159	0.091	0.087	0.028

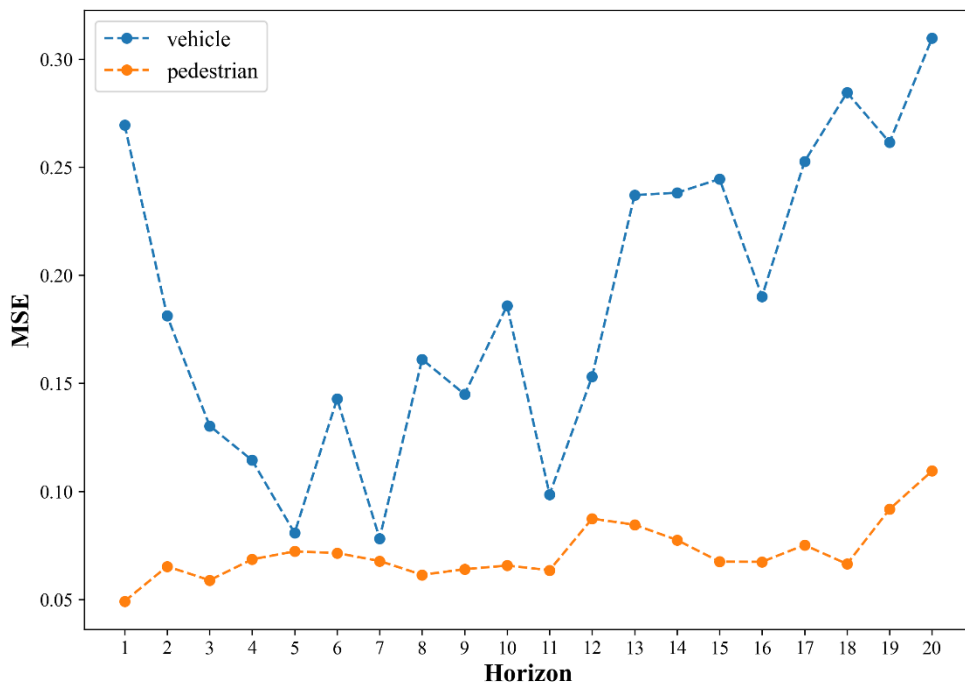
Λόγω του ότι το εν λόγω πρότυπο έχει πολύ παρόμοια αρχιτεκτονική με το LSTM_ed, η εξέλιξη των μετρικών ταχυτήτων και θέσεων και τους δύο πεζούς παρουσιάζουν σχεδόν ίδια συμπεριφορά, όπως φαίνεται στα Διαγράμματα 5.6-5.9. Όμως λόγω της παρουσίας του μηχανισμού έμφασης, παρουσιάζεται ορατή μείωση των μετρικών ανά χρονικών βήμα. Επιπλέον η αριθμητική εξέλιξη των MAE και MSE ταχυτήτων και θέσεων παρουσιάζεται στον Πίνακα Π.10 του παραρτήματος.



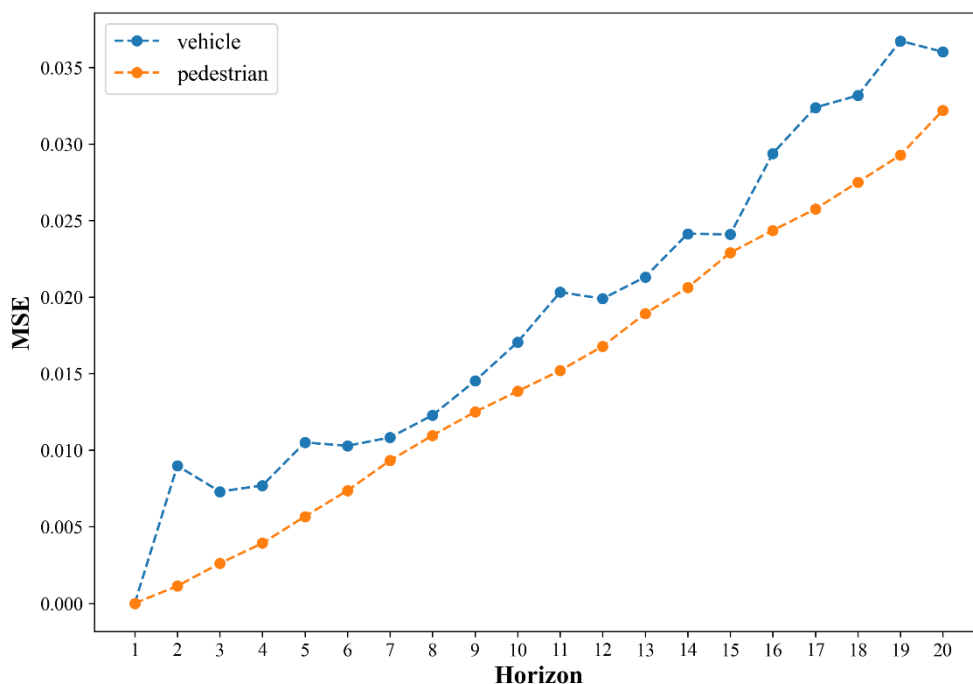
Διάγραμμα 5. 6: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_attention_wf



Διάγραμμα 5. 7: MAE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_attention_wf



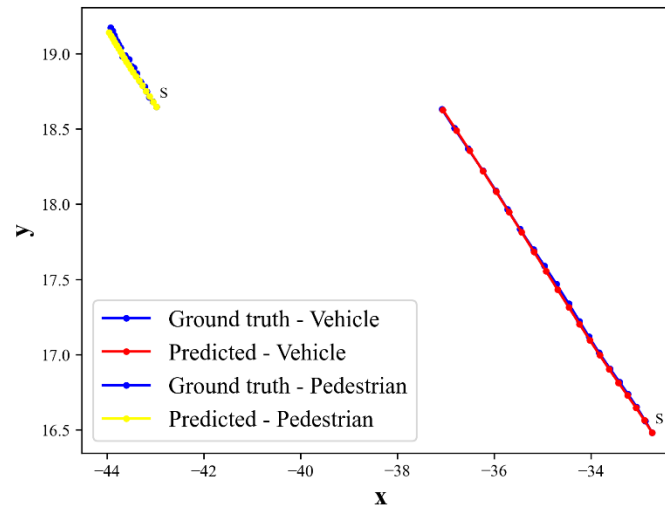
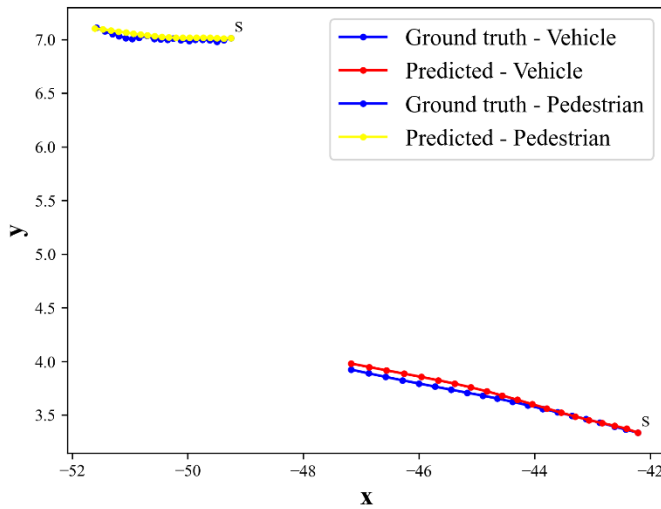
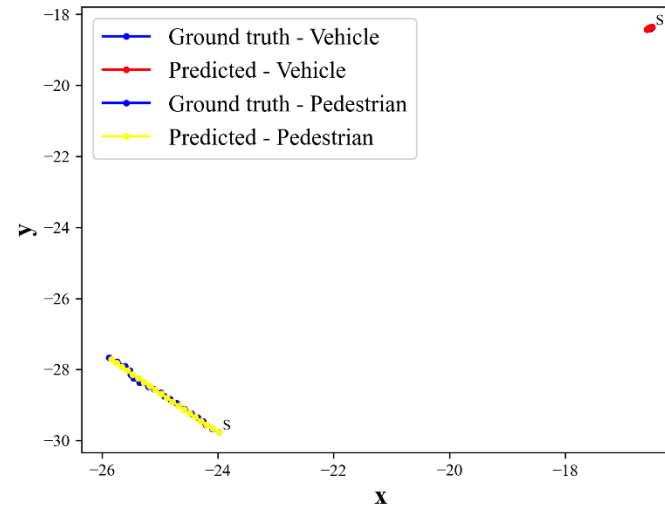
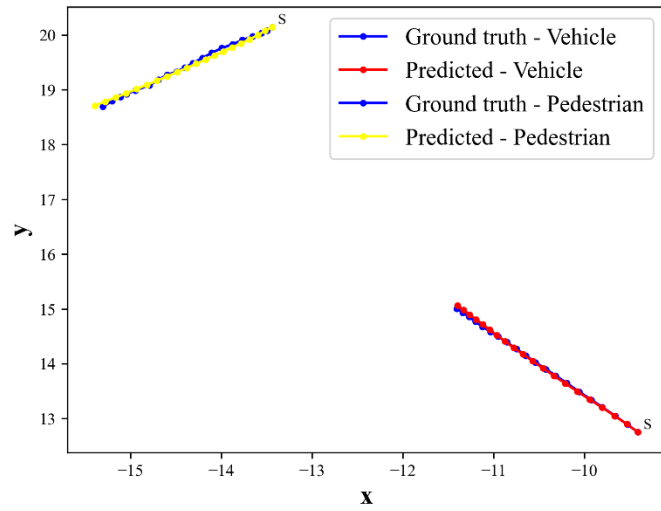
Διάγραμμα 5. 8: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_attention_wf



Διάγραμμα 5. 9: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το LSTM_ed_attention_wf

Η σύγκλιση της απώλειας Huber και των MAE και MSE ανά εποχή για το LSTM_ed απεικονίζεται στα Διαγράμματα Π.4 με Π.6. Η αριθμητική εξέλιξη των μέσων MAE και MSE ανά εποχή παρουσιάζεται στον Πίνακα Π.9 του παραρτήματος.

Στο διάγραμμα 5.10 παρουσιάζονται οι προβλεπόμενες και οι πραγματικές τροχιές τεσσάρων διαφορετικών επεισοδίων (τα ίδια με του Διαγράμματος 5.5), όπως προέκυψαν από το πρότυπο LSTM_ed_attention. Παρόλο που το LSTM_ed οπτικά έχει πολύ ακριβή αποτελέσματα, η βελτίωση με τη προσθήκη του μηχανισμού έμφασης είναι εμφανής.



Διάγραμμα 5. 10: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (LSTM_ed_attention_wf)

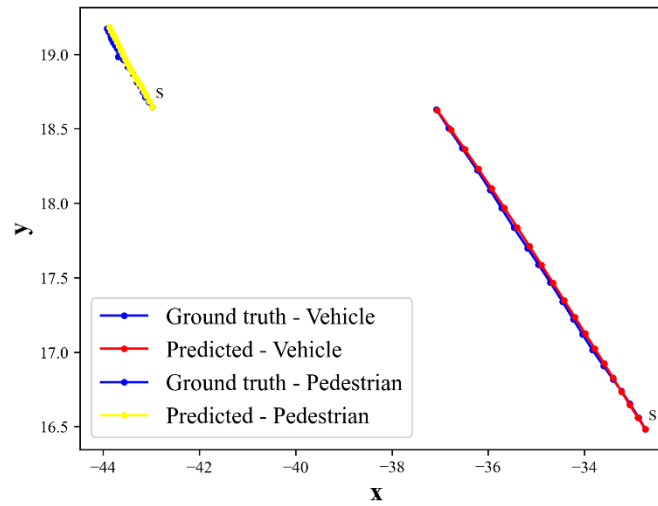
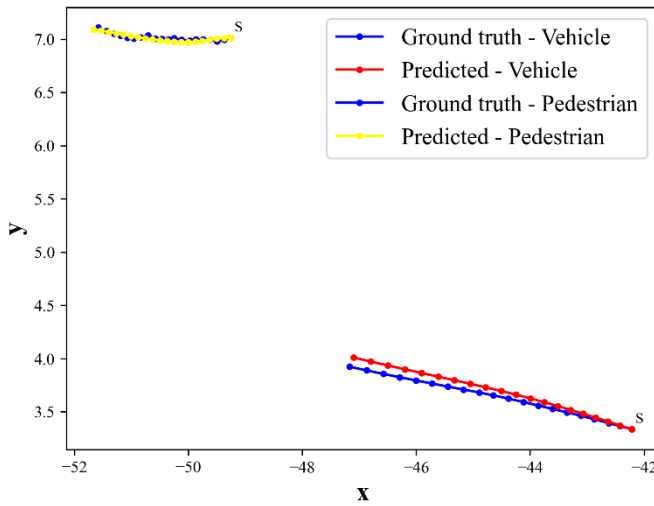
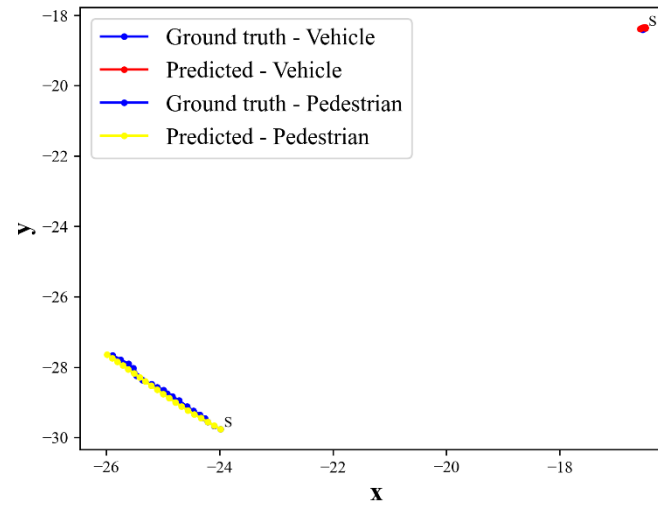
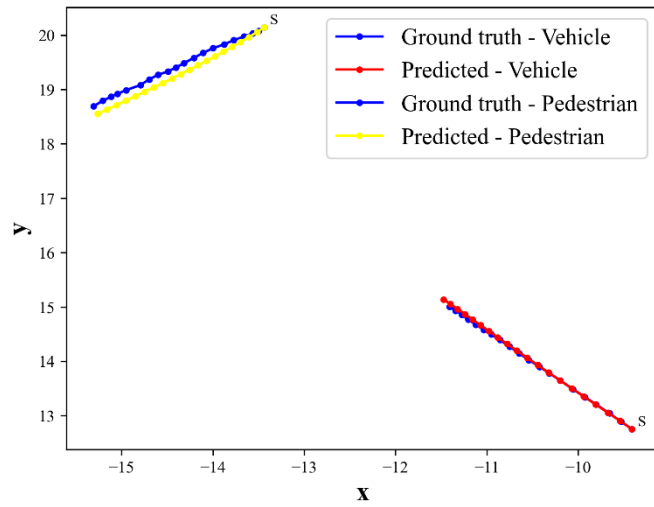
5.1.3 Αποτελέσματα CNN_LSTM_ed_attention

Το πρότυπο CNN_LSTM_ed_attention αποτελεί βελτίωση του LSTM_ed_attention, λόγω της προσθήκης του επιπέδου CNN, στο οποίο εφαρμόζεται μονοδιάστατη συνέλιξη στην είσοδο. Για να ελαχιστοποιηθεί η απώλεια Huber το παρόν πρότυπο χρειάστηκε 51 εποχές (και 53 χωρίς τα χαρακτηριστικά). Όπως και στα προηγούμενα δύο πρότυπα, η ακρίβεια του προτύπου βελτιώνεται με την χρήση των χαρακτηριστικών, όπως φαίνεται στον Πίνακα 5.3. Συγκρίνοντας τους πίνακες 5.2 και 5.3 φαίνεται ότι οι επιμέρους μετρικές προκύπτουν ίσες μέχρι το τρίτο δεκαδικό, υποδεικνύοντας έτσι ότι η προσθήκη του επιπέδου CNN δεν οδηγεί σε βελτίωση της ακρίβειας για την συγκεκριμένη αρχιτεκτονική. Όπως είναι σαφές, τα MAE, MSE ταχυτήτων και θέσεων ανά χρονικό βήμα είναι σχεδόν τα ίδια με αυτά του LSTM_ed_attention (βλ. Πίνακα Π.11 στο παράρτημα). Για τον λόγο αυτό παραλείπεται η παρουσίαση των σχετικών διαγραμμάτων. Για λόγους πληρότητας συμπεριλαμβάνονται στο παράρτημα (Διαγράμματα Π.7 με Π.10). Επιπλέον, στο παράρτημα είναι ορατά τα Διαγράμματα σύγκλισης των μετρικών και της απώλειας Huber ανά εποχή για το παρόν πρότυπο (Π.11 με Π.14). Η αριθμητική εξέλιξη των μέσων MAE και MSE ανά εποχή παρουσιάζεται στον Πίνακα Π.9 του παραρτήματος.

Πίνακας 5. 3: Μέσα MAE και MSE για το CNN_LSTM_ed_attention

	Αυτόνομο Οχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
Με χαρακτ.	0.114	0.064	0.188	0.019	0.138	0.07	0.072	0.015
Χωρίς χαρακτ.	0.122	0.071	0.194	0.022	0.152	0.083	0.081	0.022

Στο διάγραμμα 5.11 παρουσιάζονται οι προβλεπόμενες και πραγματικές τροχιές των πρακτόρων που προκύπτουν από το πρότυπο CNN_LSTM_ed_attention για τέσσερα επεισόδια (τα ίδια με τα υπόλοιπα σχετικά Διαγράμματα). Παρά το γεγονός ότι για το παρόν πρότυπο έχει προστεθεί ένα επιπλέον επίπεδο (CNN) στην αρχιτεκτονική του LSTM_ed_attention, οι προβλέψεις μοιάζουν να αποκλίνουν ελαφρώς περισσότερο από τις πραγματικές τιμές, επιβεβαιώνοντας ότι η προσθήκη ενός επιπέδου CNN στην αρχιτεκτονική του LSTM_ed_attention δεν έχει κανένα νόημα.



Διάγραμμα 5. 11: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (CNN_LSTM_ed_attention_wf)

5.1.4 Αποτελέσματα BiLSTM_ed

Το πρότυπο BiLSTM_ed είναι το πρώτο και απλούστερο από τα τρία πρότυπα που χρησιμοποιούν την τεχνική BiLSTM, η οποία εφαρμόζει ταυτόχρονη αμφίδρομη ανάλυση της εισόδου. Για την ελαχιστοποίηση της απώλειας Huber και άρα την πλήρη εκπαίδευση του πρότυπου απαιτήθηκαν 51 εποχές (και 53 χωρίς τα χαρακτηριστικά). Το παρόν πρότυπο είναι το μόνο στο οποίο τα εξαγόμενα χαρακτηριστικά μοιάζουν να μην βελτιώνουν πολύ την ακρίβεια των προβλέψεων. Όπως φαίνεται στον Πίνακα 5.3, και για τις μετρικές ταχυτήτων και για τις μετρικές θέσεων, οι τιμές με και χωρίς χαρακτηριστικά είναι περίπου οι ίδιες. Μάλιστα το μέσο MAE θέσης και το μέσο MSE ταχύτητας για το αυτόνομο όχημα έχουν χαμηλότερη τιμή χωρίς τα χαρακτηριστικά παρά με αυτά. Για τον λόγω αυτό, μεταξύ των δύο αποφασίζεται να αναλυθούν τα αποτελέσματα που προέκυψαν χωρίς να ληφθούν υπόψιν τα χαρακτηριστικά.

Πίνακας 5. 4: Μέσα MAE και MSE για το BiLSTM_ed

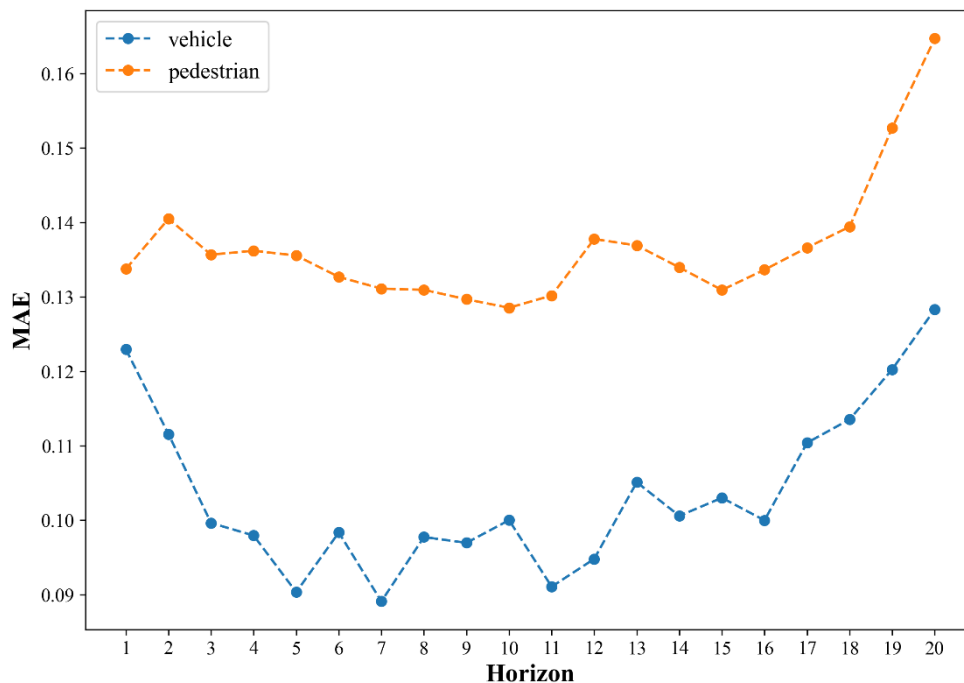
	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
Με χαρακτ.	0.103	0.059	0.185	0.017	0.130	0.063	0.067	0.011
Χωρίς χαρακτ.	0.104	0.058	0.183	0.018	0.137	0.069	0.072	0.016

Η αριθμητική παρουσίαση της εξέλιξης των μετρικών στο χρονικό ορίζοντα και ανά εποχή παρουσιάζονται στους πίνακες Π.12 και Π.9 του παραρτήματος. Όπως φαίνεται στο διάγραμμα 5.12, η εξέλιξη του MAE ταχυτήτων στον ορίζοντα των 2 δευτερολέπτων είναι παρόμοια με αυτή των προηγούμενων 3 προτύπων. Αυτό συμβαίνει λόγω του ότι όλα τα πρότυπα που χρησιμοποιούνται στην παρούσα εργασία έχουν παρόμοια αρχιτεκτονική και φιλοσοφία. Παρόλα αυτά, παρατηρώντας το σχετικό διάγραμμα, πέρα από το γεγονός ότι τα MAE ταχυτήτων ανά χρονικό βήμα είναι μικρότερο σε σχέση με τα αντίστοιχα των προτύπων χωρίς BiLSTM γίνεται αντιληπτό και ότι η εξέλιξη της εν λόγω μετρικής παρουσιάζει μια ελαφρά ηπιότητα σε σχέση με τα προηγούμενα πρότυπα, κυρίως όσον αφορά το αυτόνομο όχημα.

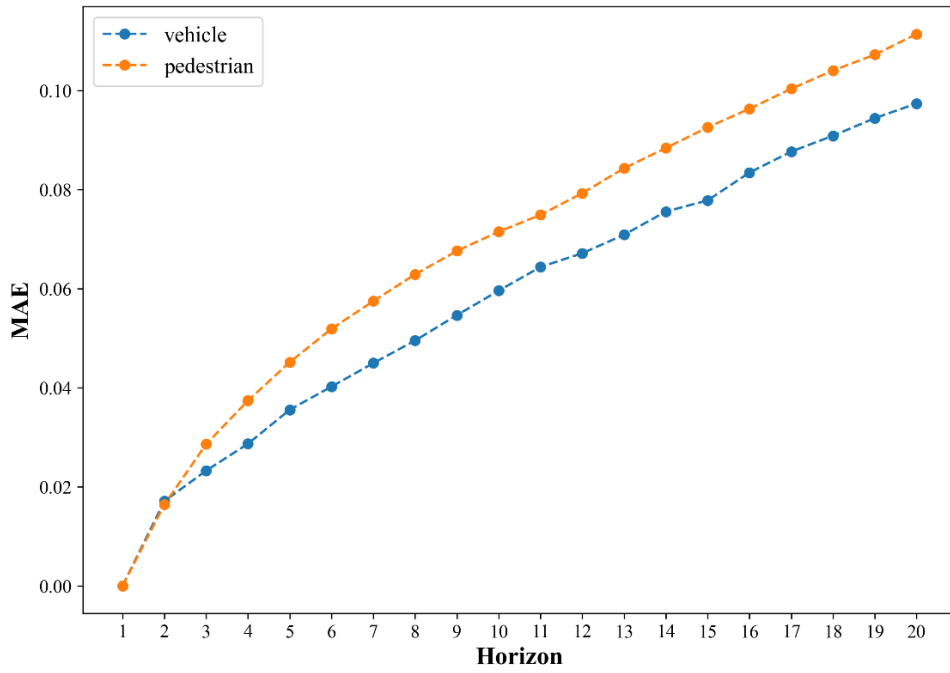
Η εξέλιξη των MAE θέσεων στον χρονικό ορίζοντα των 2 δευτερολέπτων για το παρόν πρότυπο παρουσιάζεται στο διάγραμμα 5.13. Σε αντίθεση με τα προηγούμενα τρία πρότυπα στα οποία η εν λόγω μετρική ήταν αρκετά κοντά για τους δύο πράκτορες, είναι ευδιάκριτο ότι οι προβλέψεις των τροχιών του αυτόνομου οχήματος είναι σημαντικά πιο ακριβείς σε σχέση με αυτές των πεζών. Επιπλέον, από την περιοχή γύρω από το 10^ο χρονικό βήμα και ύστερα, ο ρυθμός αύξησης του MAE ανά χρονικό βήμα δείχνει να σταθεροποιείται. Με αυτόν τον τρόπο γίνεται αντιληπτό ότι υπό τον υπάρχοντα όγκο δεδομένων, η πρόβλεψη πέρα των 20 χρονικών βημάτων δεν θα ήταν ιδιαίτερα ακριβής.

Όσον αφορά την εξέλιξη του MSE στον χρονικό ορίζοντα, στην περίπτωση των ταχυτήτων (διάγραμμα 5.14) φαίνεται να έχει σχεδόν την ίδια συμπεριφορά με τα προηγούμενα τρία πρότυπα. Από την άλλη, το MSE τροχιών (διάγραμμα 5.15)

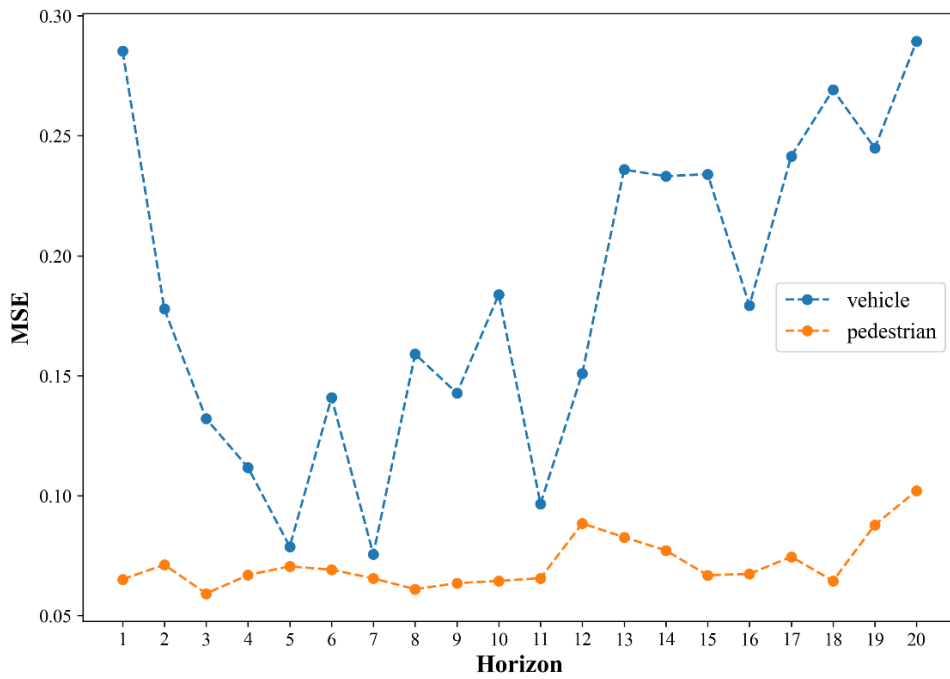
παρουσιάζει μια αυξητική τάση από το πρώτο κιόλας χρονικό βήμα, η οποία διαφοροποιείται ελαφρώς μεταξύ των δύο πρακτόρων. Το MSE θέσεων του αυτόνομου οχήματος παρουσιάζει μια αυξητική τάση με εν γένει σταθερό ρυθμό και αρκετές μικρές διακυμάνσεις μεταξύ των χρονικών βημάτων. Από την άλλη, η προκύπτουσα γραφική παράσταση εξέλιξης του MAE τροχιών των πεζών δεν παρουσιάζει σχεδόν καθόλου διακυμάνσεις μεταξύ των χρονικών βημάτων. Αντιθέτως οπτικά προσεγγίζει κοίλη καμπύλη με χαμηλή καμπυλότητα. Το γεγονός αυτό υποδηλώνει ότι η πρόβλεψη της κίνησης των πεζών είναι εξίσου ακριβής με του αυτόνομου οχήματος, αλλά παρουσιάζει μια ελαφρώς μεγαλύτερη σταθερότητα, ενδεχομένως λόγω του ότι σε σημαντικό αριθμό αλληλεπιδράσεων οι πεζοί παρέμειναν ακίνητοι. Παράλληλα η κλίση αύξησης του MAE θέσεων πεζών είναι ιδιαίτερα αυξημένη κατά τα τελευταία χρονικά βήματα, υποδηλώνοντας την δραστικά μειούμενη ακρίβεια πρόβλεψης όσο αυξάνονται τα χρονικά βήματα.



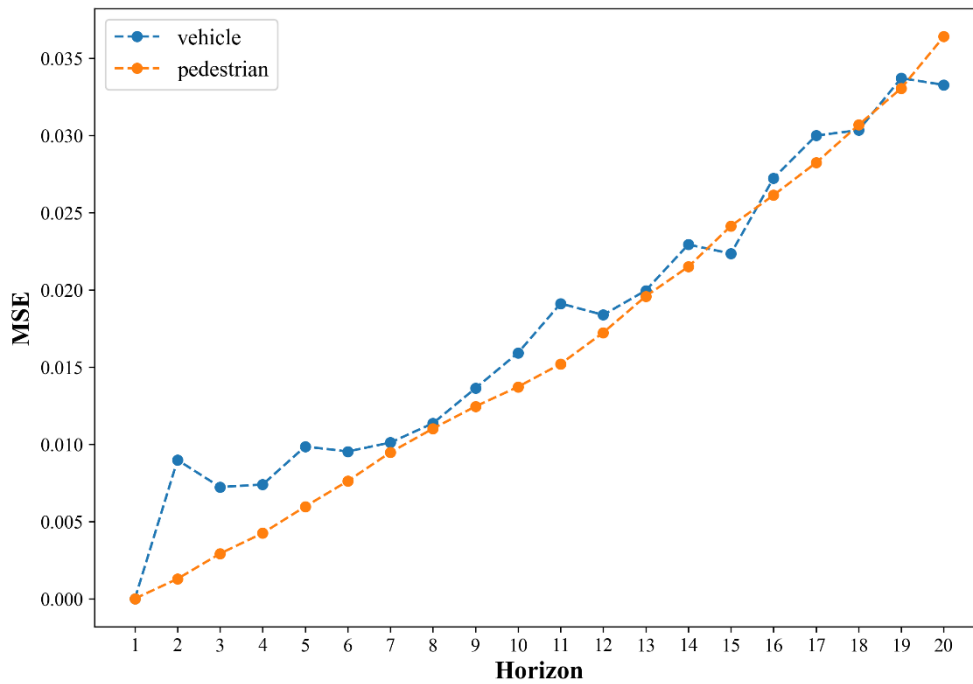
Διάγραμμα 5. 12: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_nf



Διάγραμμα 5. 13: MAE θέσης στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_nf



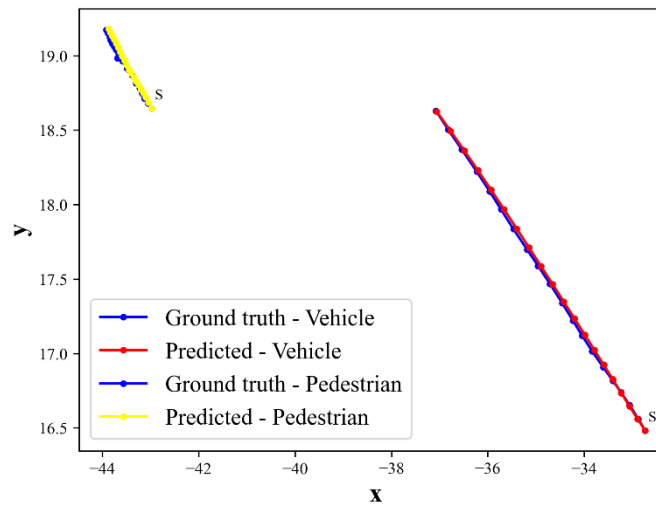
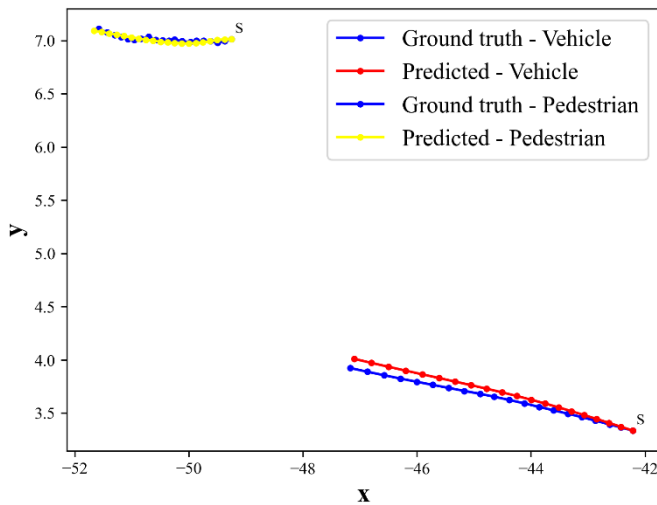
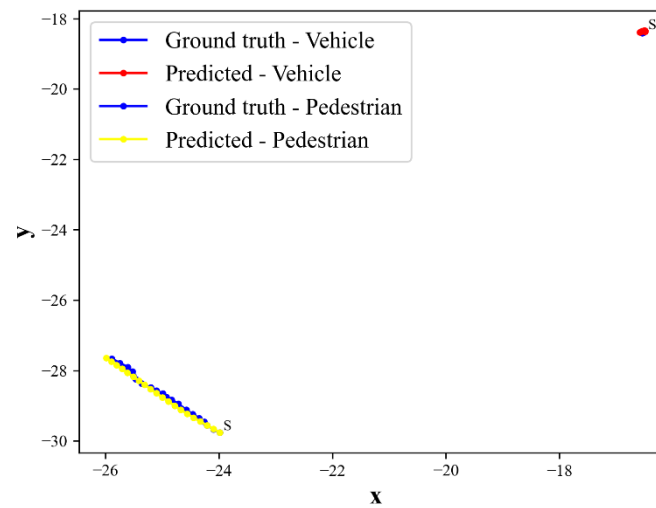
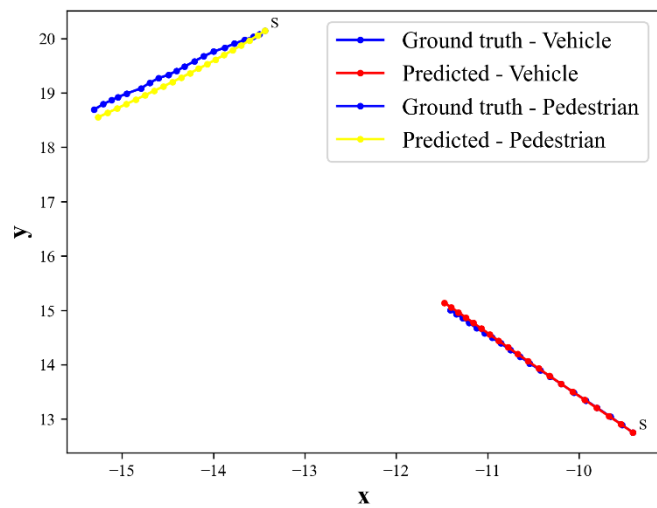
Διάγραμμα 5. 14: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_nf



Διάγραμμα 5. 15: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_nf

Η σύγκλιση της απώλειας Huber και των MAE και MSE ανά εποχή για το BiLSTM_ed_nf απεικονίζεται στα Διαγράμματα Π.14 με Π.16 του παραρτήματος.

Στο διάγραμμα 5.16 παρουσιάζονται οι προβλεπόμενες και οι πραγματικές τροχιές τεσσάρων διαφορετικών επεισοδίων (τα ίδια με του Διάγραμματος 5.5), όπως προέκυψαν από το πρότυπο BiLSTM_ed_nf. Παρατηρώντας το διάγραμμα, είναι εμφανές ότι οι προβλέψεις όσον αφορά τις τροχιές των οχημάτων είναι ιδιαίτερα ακριβείς. Από την άλλη, οι τροχιές των πεζών μοιάζουν να είναι ελαφρώς πιο ανακριβείς. Ενδεχομένως αυτό να είναι μεμονωμένο τυχαίο γεγονός. Σε κάθε περίπτωση το εν λόγω πρότυπο, όπως και τα υπόλοιπα επιτυγχάνει υψηλό επίπεδο ακρίβειας, με εντυπωσιακά αποτελέσματα όσον αφορά τα οχήματα.



Διάγραμμα 5. 16: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (BiLSTM_ed_nf)

5.1.5 Αποτελέσματα BiLSTM_ed_attention

Το πρότυπο BiLSTM_ed_attention έχει την ίδια αρχιτεκτονική με το BiLSTM_ed με την διαφορά ότι συμπεριλαμβάνει μηχανισμό έμφασης. Για την ελαχιστοποίηση της απώλειας Huber το παρόν πρότυπο χρειάστηκε 49 εποχές με τα χαρακτηριστικά και 54 χωρίς αυτά. Σε αντίθεση με το προηγούμενο πρότυπο, το BiLSTM_ed_attention έχει αισθητά καλύτερη απόδοση όταν λαμβάνονται υπόψιν τα χαρακτηριστικά, όπως φαίνεται στον Πίνακα 5.5.

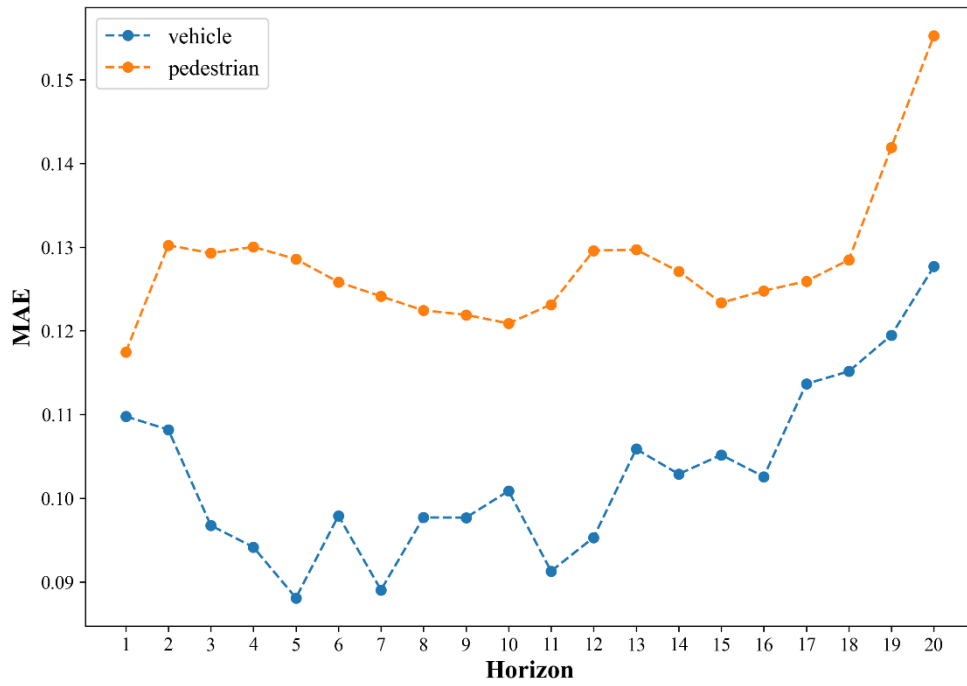
Πίνακας 5. 5: Μέσα MAE και MSE για το BiLSTM_ed_attention

	Αυτόνομο Οχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
Με χαρακτ.	0.103	0.059	0.185	0.017	0.130	0.063	0.067	0.011
Χωρίς χαρακτ.	0.115	0.068	0.190	0.021	0.146	0.080	0.078	0.021

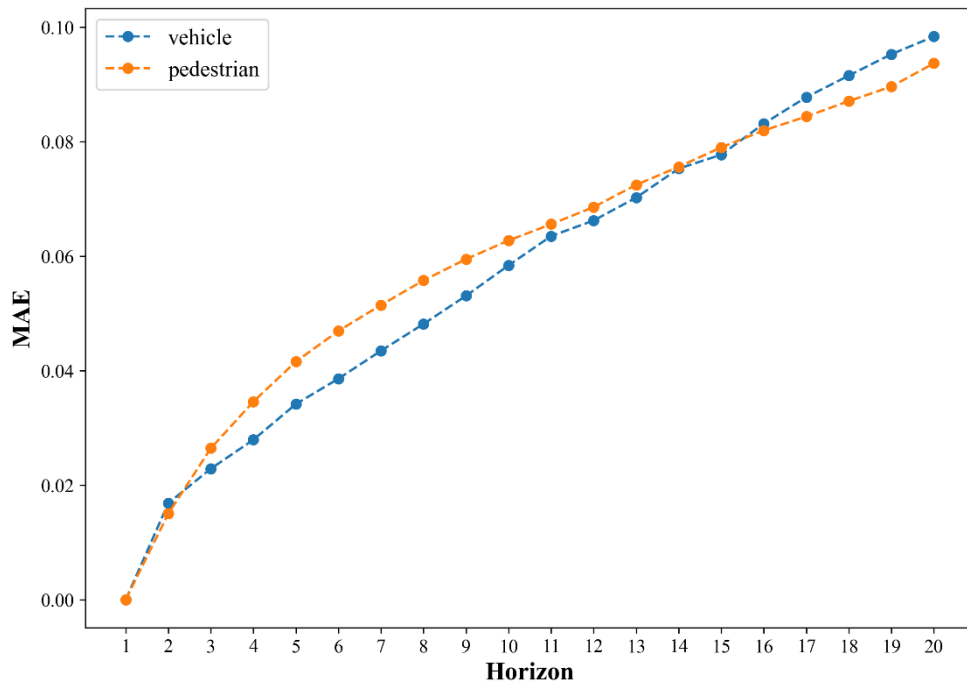
Η προσθήκη του μηχανισμού έμφασης στην αρχιτεκτονική του πρότυπου BiLSTM_ed αποδίδει τις ίδιες τιμές μετρικών έως το τρίτο δεκαδικό στην περίπτωση συμπερίληψης των χαρακτηριστικών. Στην περίπτωση που τα χαρακτηριστικά δεν λαμβάνονται υπόψιν, η προσθήκη του μηχανισμού έμφασης προσδίδει πολύ υψηλότερες τιμές MAE και MSE. Επομένως ο μηχανισμός έμφασης δεν βελτιώνει την ακρίβεια για την παρούσα αρχιτεκτονική, αλλά αντιθέτως μπορεί μέχρι και να την μειώσει. Παρόλα αυτά συγκρίνοντας το BiLSTM_ed_attention_wf με το BiLSTM_nf, το πρώτο έχει αρκετά χαμηλότερες τιμές μέσω μετρικών.

Η εξέλιξη του MAE και του MSE ταχυτήτων στον ορίζοντα των 2 δευτερολέπτων δείχνει να είναι σχεδόν ίδια με των υπολοίπων BiLSTM προτύπων, όπως φαίνεται στα Διαγράμματα 5.17 και 5.19. Από την άλλη, οι μετρικές θέσεων παρουσιάζουν μια ελαφρά μεταβολή ως προς την εξέλιξη τους στον προαναφερθέντα ορίζοντα. Παρατηρώντας το διάγραμμα 5.19 μπορεί να γίνει αντιληπτό ότι η χρονική εξέλιξη του MAE τροχιών είναι περίπου ίση μεταξύ των δύο πρακτόρων, κάτι που δεν συμβαίνει στο BiLSTM_ed. Αντιθέτως, συγκρίνοντας τα Διαγράμματα 5.14 και 5.20 γίνεται ξεκάθαρο ότι η παρουσία του μηχανισμού έμφασης είναι ιδιαιτέρως ευεργετική για το MSE τροχιών των πεζών, με την εξέλιξη του MSE τροχιών αυτόνομων οχημάτων να μην μεταβάλλεται ιδιαιτέρως σε σχέση με αυτές των προηγούμενων προτύπων. Η συνολική αριθμητική αποτύπωση της εξέλιξης των μετρικών αξιολόγησης ανά κατηγορία και πράκτορα παρουσιάζεται στον Πίνακα Π.13 του παραρτήματος.

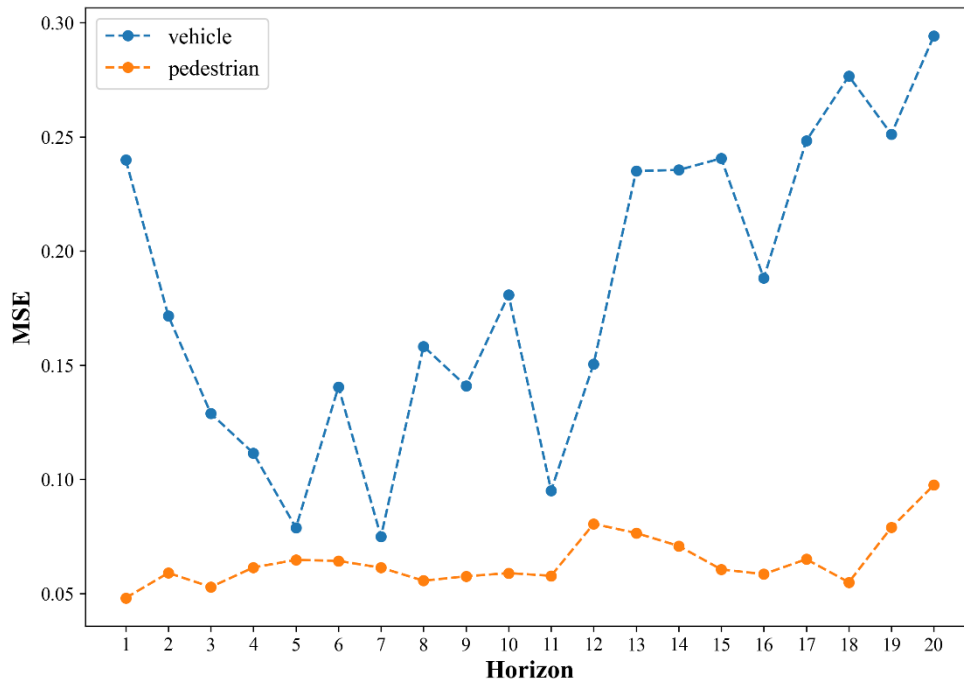
Η οπτικοποίηση της σύγκλισης της απώλειας Huber και των μετρικών αξιολόγησης ανά εποχή είναι ορατή στα Διαγράμματα Π.17 με Π.19 του παραρτήματος. Επιπλέον στον Πίνακα Π.9 του παραρτήματος παρουσιάζεται η εξέλιξη των μετρικών ανά εποχή.



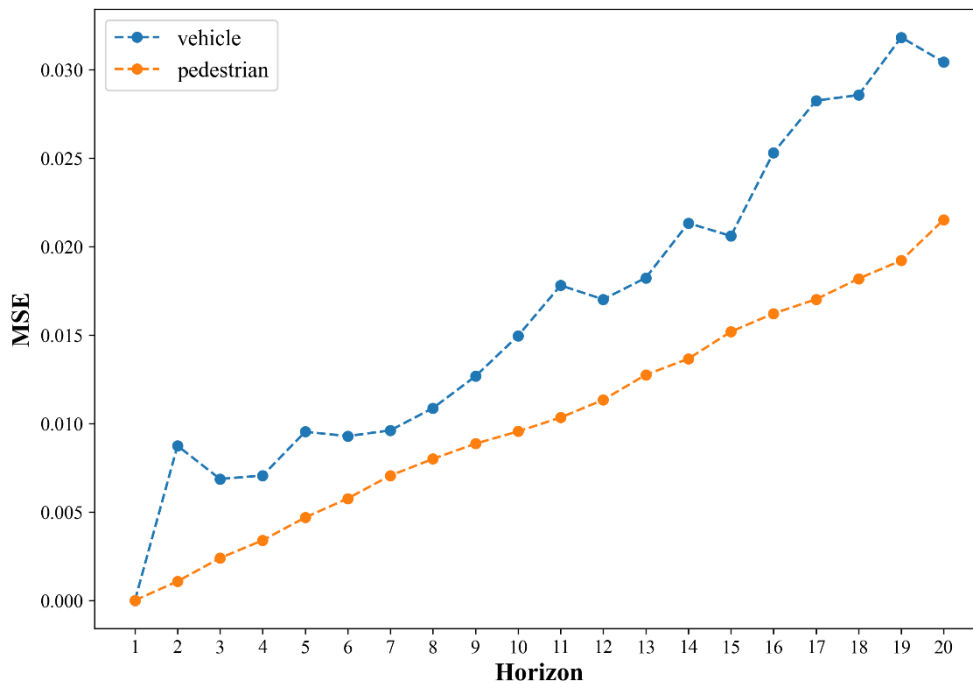
Διάγραμμα 5. 17: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_attention_wf



Διάγραμμα 5. 18: MAE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_attention_wf

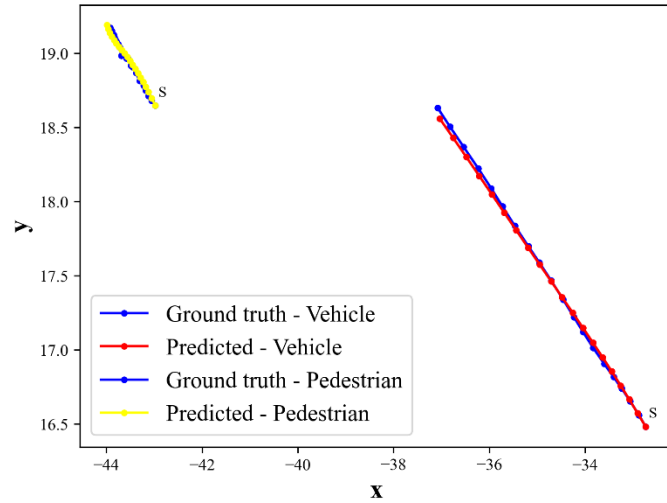
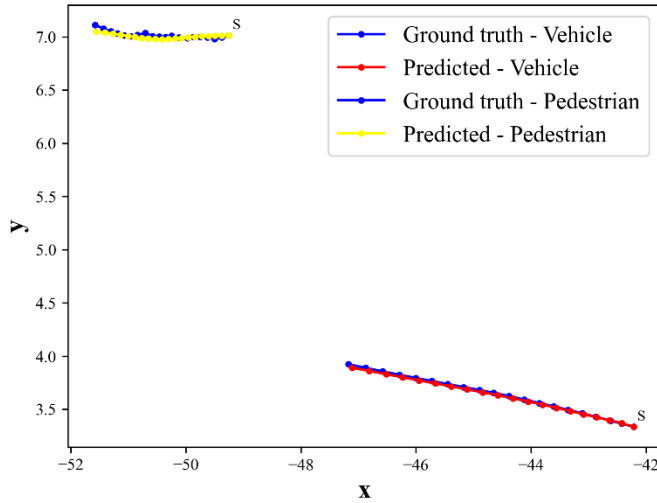
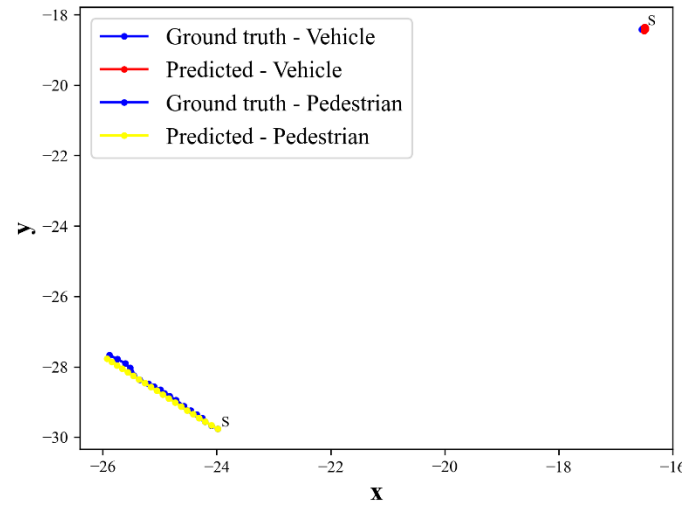
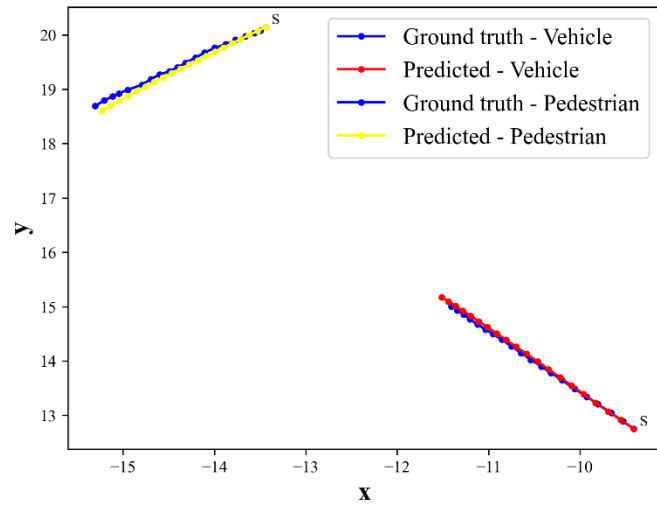


Διάγραμμα 5. 19: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_attention_wf



Διάγραμμα 5. 20: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το BiLSTM_ed_attention_wf

Στο διάγραμμα 5.17 απεικονίζονται οι πραγματικές και προβλεπόμενες τροχιές των ίδιων τεσσάρων επεισοδίων αλληλεπίδρασης, όπως προέκυψαν από το πρότυπο BiLSTM_ed_attention_wf. Οι προβλέψεις φαίνονται να είναι πολύ κοντά στις πραγματικές τιμές, κάτι το οποίο επιβεβαιώνεται από τις αρκετά χαμηλές τιμές των μετρικών αξιολόγησης.



Διάγραμμα 5. 21: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (BiLSTM_ed_attention_wf)

5.1.6 Αποτελέσματα CNN_BiLSTM_ed

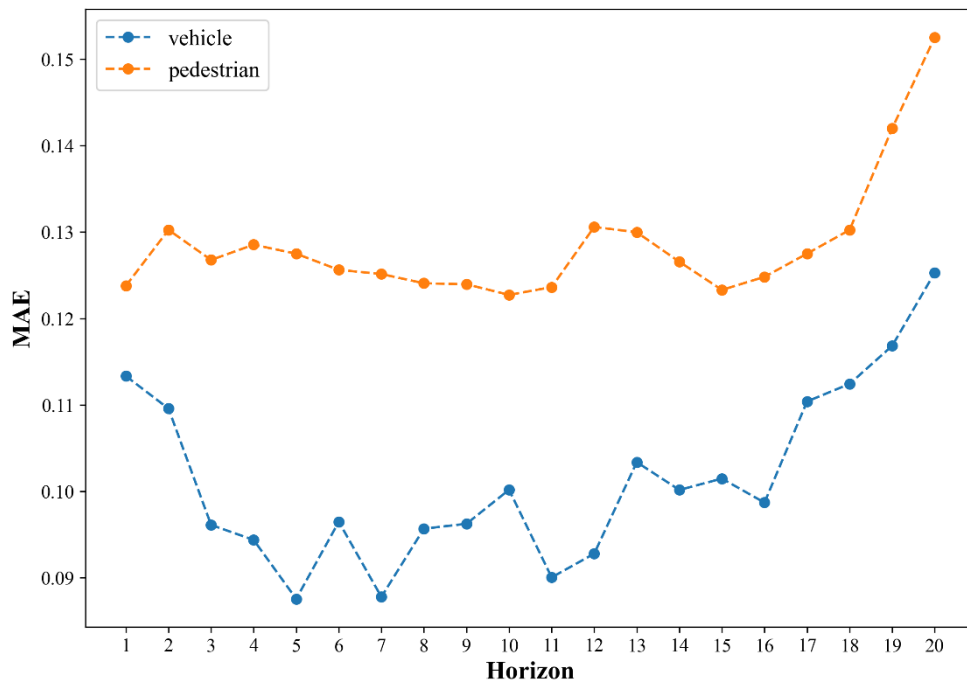
Το πρότυπο CNN_BiLSTM_ed στηρίζεται στην αρχιτεκτονική του BiLSTM_ed, αλλά διαφοροποιείται στο ότι συμπεριλαμβάνεται και ένα επίπεδο CNN. Για την πλήρη εκπαίδευση το παρόν πρότυπο χρειάστηκε 48 εποχές χωρίς τα χαρακτηριστικά και 54 με αυτά. Στα πρότυπα όπου αποτελούνταν από απλά κελιά LSTM η παρουσία του CNN δεν έδειξε να προσφέρει ιδιαίτερη βελτίωση. Στην περίπτωση του BiLSTM με τα χαρακτηριστικά να λαμβάνονται υπόψιν, η παρουσία ενός επιπέδου CNN είναι πιο ευεργετική, όπως φαίνεται συγκρίνοντας τους πίνακες 5.4 και 5.6. Βεβαίως αυτό δεν σημαίνει ότι η βελτίωση της ακρίβειας είναι σημαντική. Στην περίπτωση εκπαίδευσης χωρίς τα χαρακτηριστικά, το πρότυπο που συμπεριλαμβάνει το επίπεδο CNN οδηγεί σε μειωμένη ακρίβεια, αποδεικνύοντας έτσι την σημαντικότητα τους για την συγκεκριμένη αρχιτεκτονική.

Πίνακας 5. 6: Μέσα MAE και MSE για το CNN_BiLSTM_ed

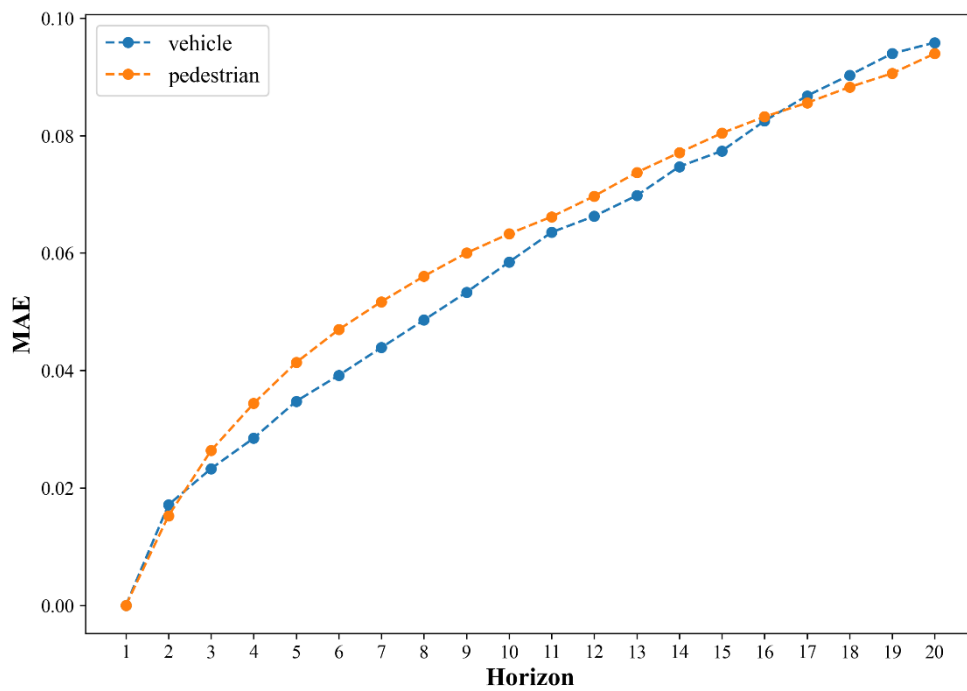
	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
Με χαρακτ.	0.101	0.057	0.185	0.017	0.128	0.060	0.066	0.010
Χωρίς χαρακτ.	0.112	0.066	0.189	0.020	0.144	0.077	0.077	0.019

Η εξέλιξη του MAE και του MSE ταχυτήτων και τροχιών στον ορίζοντα των 20 χρονικών βημάτων δείχνει να είναι σχεδόν ίδια με του BiLSTM_ed_attention_wf, όπως φαίνεται στα Διαγράμματα 5.22 με 5.25. Το γεγονός αυτό σημαίνει ότι η προσθήκη μηχανισμού έμφασης και η προσθήκη επιπέδου CNN επηρεάζουν με παρόμοιο τρόπο την αρχιτεκτονική BiLSTM_ed, χωρίς αυτό να σημαίνει ότι τα παραγόμενα πρότυπα έχουν το ίδιο επίπεδο ακρίβειας στις προβλέψεις τους. Συγκρίνοντας τους πίνακες 5.5 και 5.6 προκύπτει ότι τα δύο πρότυπα έχουν παρόμοιες τιμές μετρικών, χωρίς να μπορεί να γίνει ξεκάθαρο ποιο από τα δύο πρότυπα παράγει ακριβέστερες προβλέψεις και για τους δύο πράκτορες, αλλά θα ειπωθούν περισσότερα περί αυτού στην επόμενη παράγραφο. Η συνολική αριθμητική αποτύπωση της εξέλιξης των μετρικών αξιολόγησης ανά κατηγορία και πράκτορα παρουσιάζεται στον Πίνακα Π.14 του παραρτήματος.

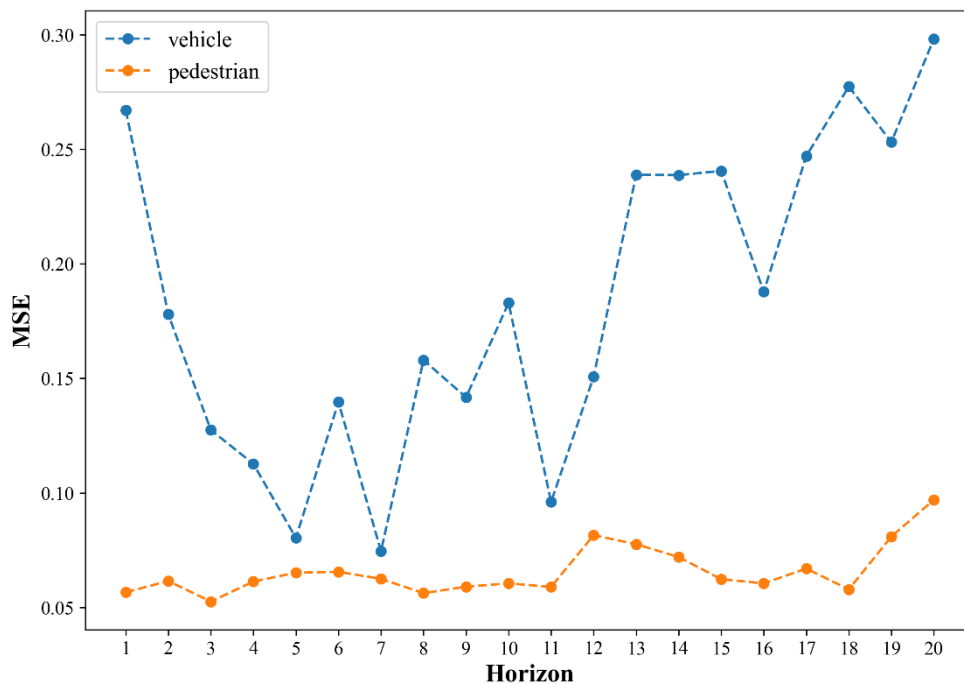
Η σύγκλιση της απώλειας Huber και των MAE και MSE ανά εποχή για το BiLSTM_ed_nf απεικονίζεται στα Διαγράμματα Π.20 με Π.22 του παραρτήματος. Επιπλέον, η εξέλιξη των δύο μετρικών ανά πράκτορα και εποχή παρουσιάζεται στον Πίνακα Π.9 του παραρτήματος.



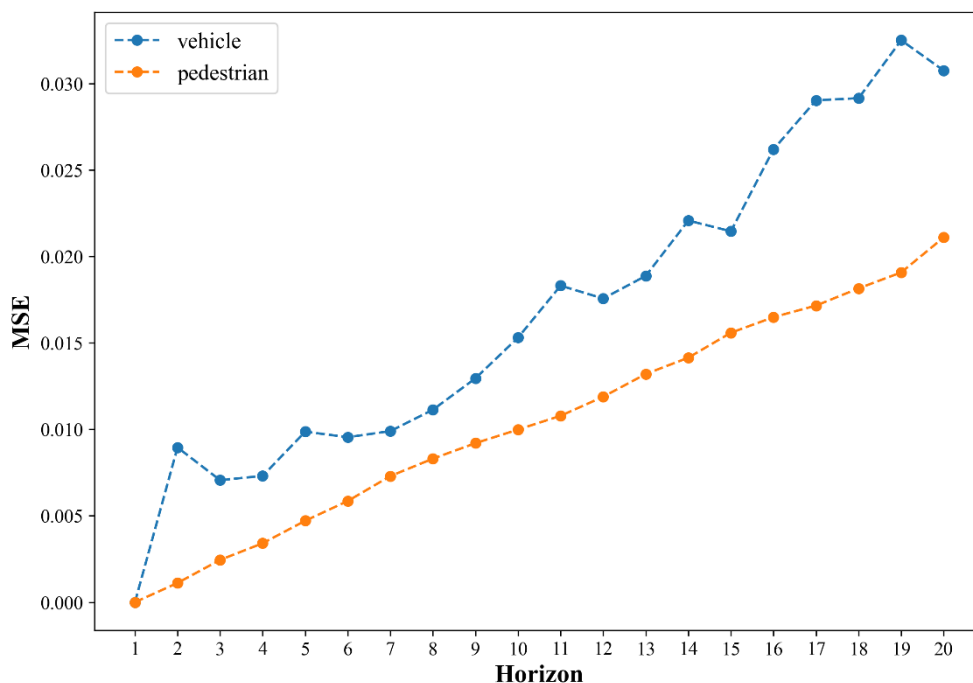
Διάγραμμα 5. 22: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το CNN_BiLSTM_ed_wf



Διάγραμμα 5. 23: MAE θέσης στον ορίζοντα 20 χρονικών βημάτων για το CNN_BiLSTM_ed_wf

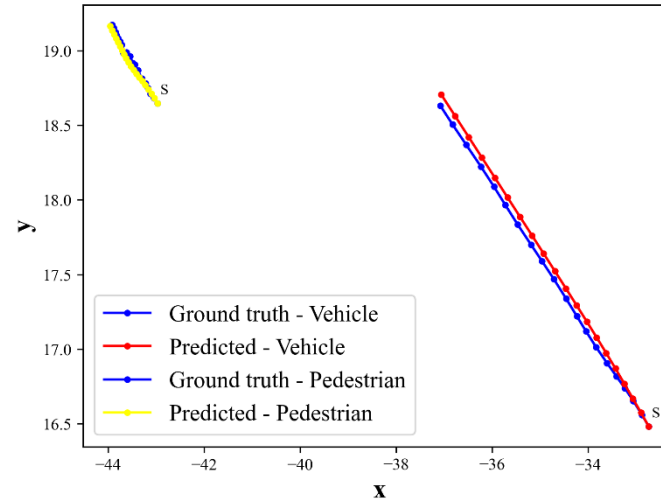
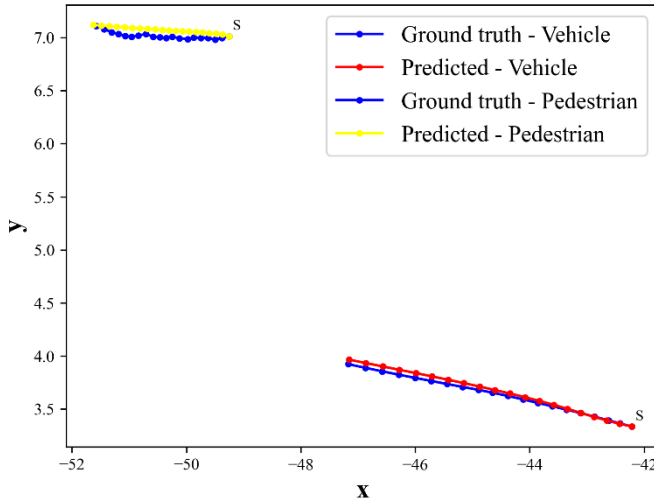
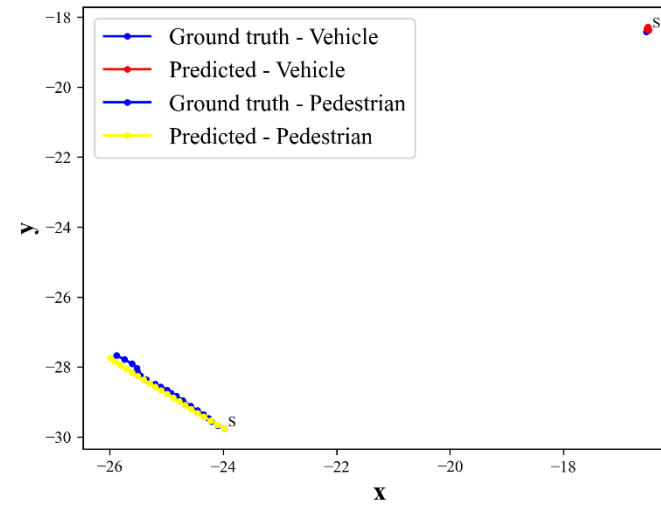
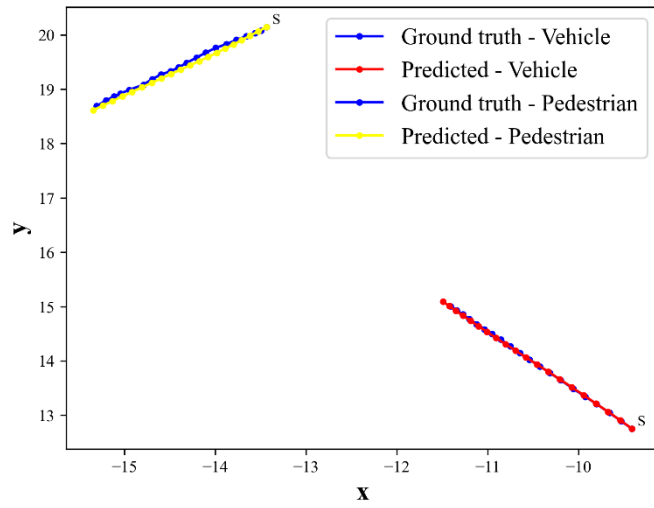


Διάγραμμα 5. 24: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το CNN_BiLSTM_ed_wf



Διάγραμμα 5. 25: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το CNN_BiLSTM_ed_wf

Στο διάγραμμα 5.22 παρουσιάζονται οι προβλεπόμενες και πραγματικές τροχιές για τέσσερις αλληλεπιδράσεις (τις ίδιες με τα άλλα πρότυπα), όπως προέκυψαν από το πρότυπο CNN_BiLSTM_ed_wf.



Διάγραμμα 5. 26: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (CNN_BiLSTM_ed_wf)

5.1.7 Αποτελέσματα Stacking Ensemble

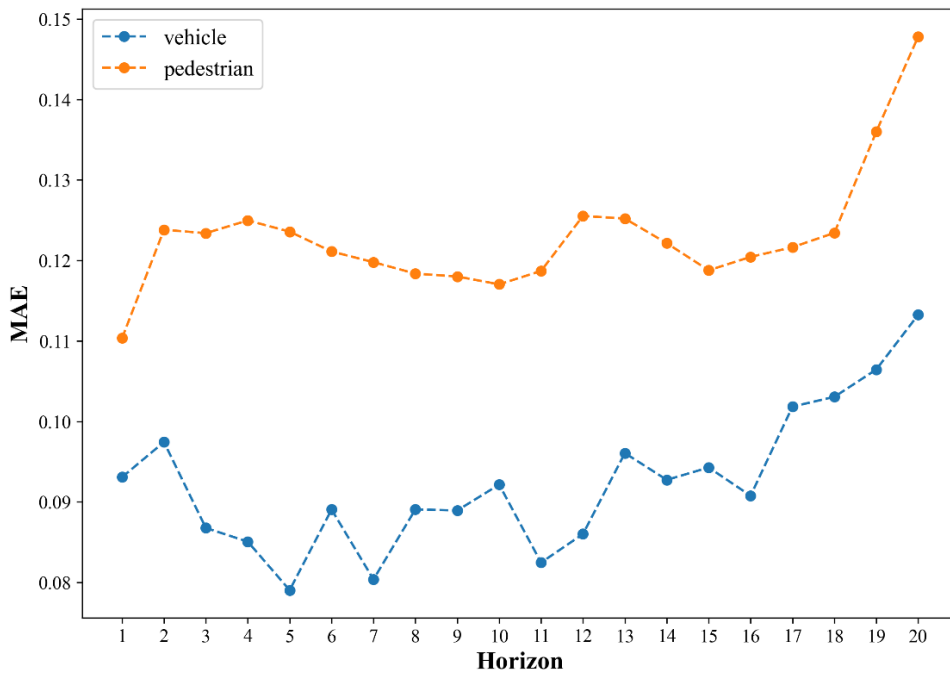
Το πρότυπο Stacking Ensemble συνδυάζει τις αρχιτεκτονικές των υπολοίπων έξι προτύπων, σταθμίζοντας τες με τέτοιο τρόπο ώστε να επιτευχθεί η μέγιστη δυνατή ακρίβεια προβλέψεων. Για τον λόγο αυτό για την εκπαίδευση του απαιτήσε μόλις 16 εποχές με τα χαρακτηριστικά και 9 χωρίς αυτά. Όπως είναι αναμενόμενο, η προσθήκη των χαρακτηριστικών έδωσε προβλέψεις με αισθητά μεγαλύτερη ακρίβεια, όπως φαίνεται στον Πίνακα 5.7. Εντυπωσιακό είναι το γεγονός ότι κυρίως όσον αφορά στα οχήματα, οι προβλέψεις του Stacking Ensemble χωρίς τα χαρακτηριστικά είναι περίπου οι ίδιες με τις καλύτερες που επιτευχθήκαν ξεχωριστά από τα προηγούμενα έξι πρότυπα.

Πίνακας 5. 7: Μέσα MAE και MSE για το CNN_BiLSTM_ed

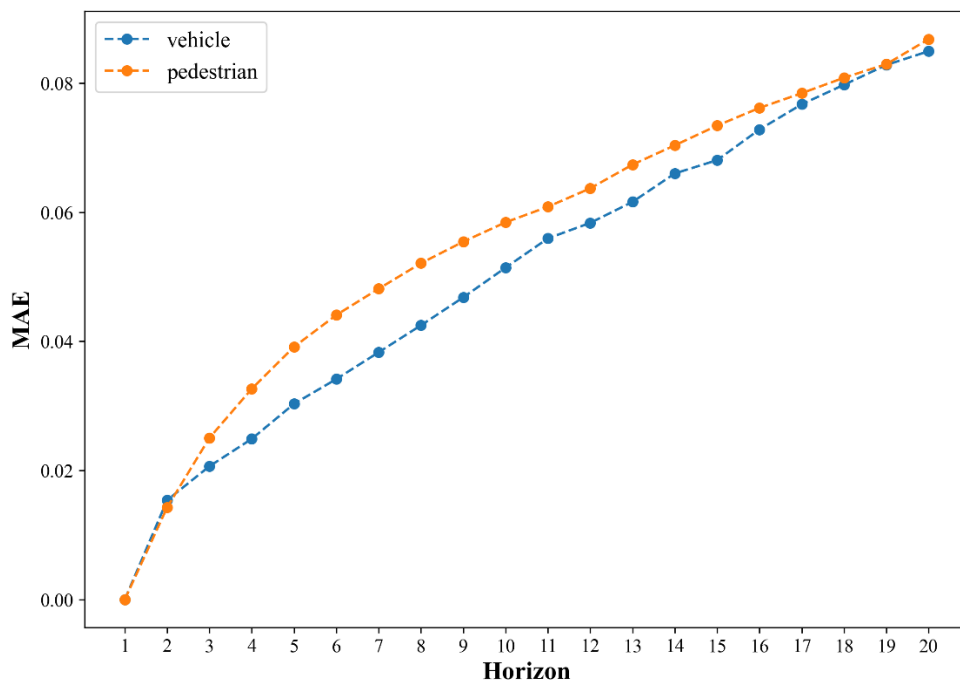
	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
Με χαρακτ.	0.092	0.051	0.181	0.015	0.123	0.056	0.062	0.009
Χωρίς χαρακτ.	0.101	0.058	0.184	0.018	0.136	0.070	0.072	0.016

Όπως φαίνεται στο διάγραμμα 5.23, ο ρυθμός μεταβολής του MAE ταχυτήτων στον ορίζοντα των 20 χρονικών βημάτων είναι αντίστοιχος με των υπολοίπων προτύπων. Αυτό οφείλεται στο γεγονός ότι όλες οι χρησιμοποιούμενες αρχιτεκτονικές έχουν αρκετά παρόμοια αρχιτεκτονική, καθώς αποτελούν πρότυπα encoder-decoder που περιέχουν κελιά LSTM. Για τους ίδιους λόγους η εξέλιξη όλων των μετρικών ανά κατηγορία και πράκτορα έχουν ρυθμό μεταβολής ανά χρονικό βήμα παρόμοιο με των υπολοίπων προτύπων, όπως φαίνεται στα Διαγράμματα 5.24 με 5.26. Στον Πίνακα Π.15 του παραρτήματος παρουσιάζονται οι εν λόγω εξελίξεις των μετρικών.

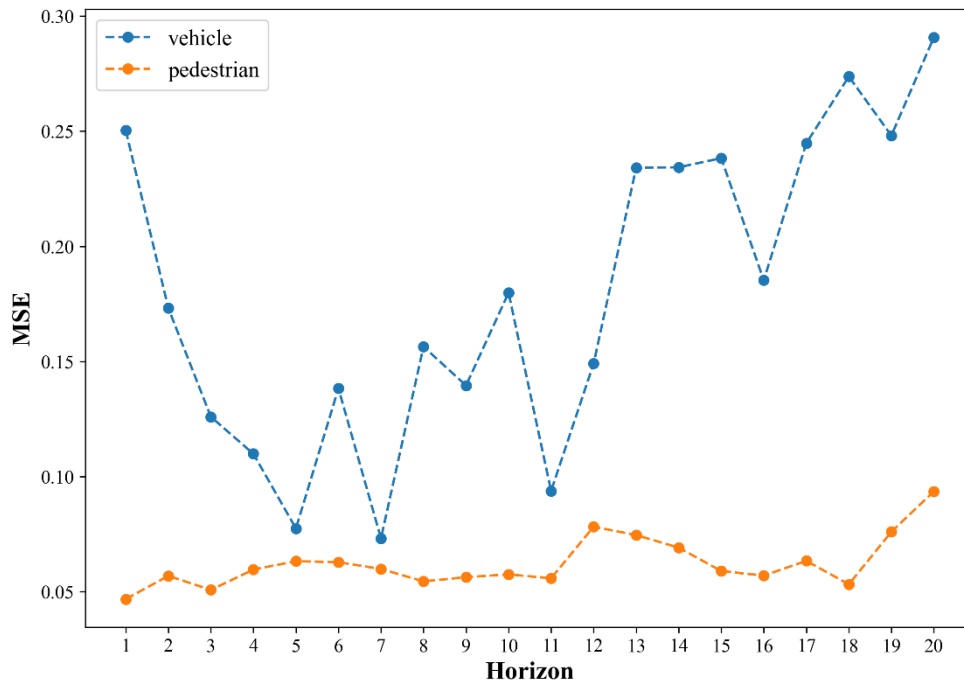
Όσον αφορά τις εξελίξεις των μετρικών (Πίνακας Π.9 του παραρτήματος) και της απώλειας Huber ανά εποχή, όπως φαίνεται στα Διαγράμματα Π.23 με Π.25 του παραρτήματος, η σύγκλιση είναι πολύ πιο απότομη και γρήγορη σε σχέση με τα άλλα πρότυπα. Αυτό οφείλεται και πάλι στο γεγονός ότι το πρότυπο Stacking Ensemble χτίζεται επάνω στα υπόλοιπα πρότυπα όποτε δεν χρειάζεται εκτενή εκπαίδευση (μη συμπεριλαμβάνοντας τον χρόνο που απαιτήθηκε για την εκπαίδευση των υπόλοιπων έξι προτύπων).



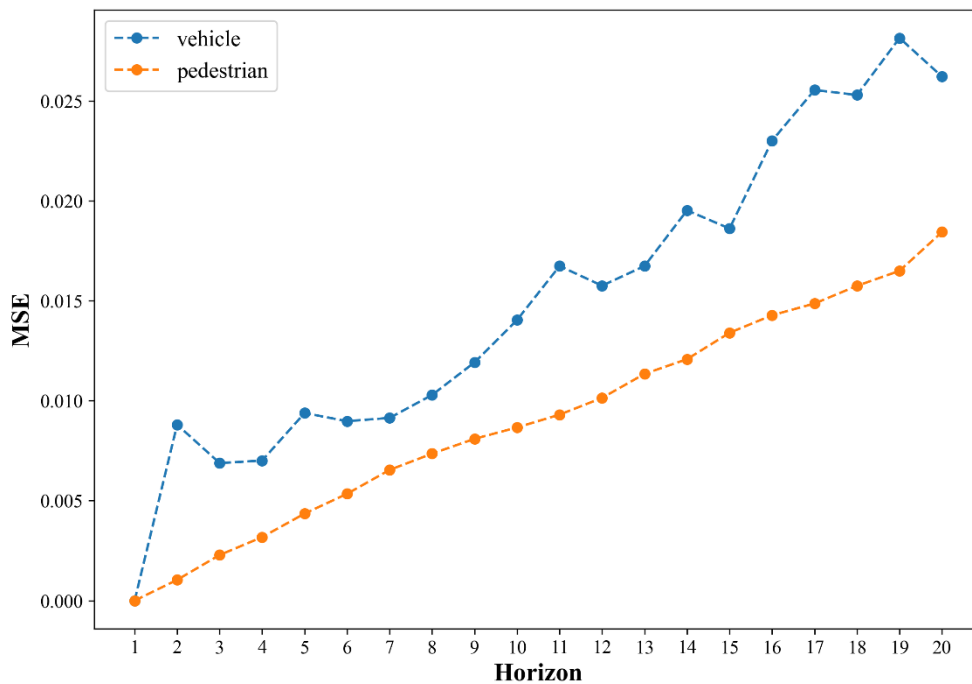
Διάγραμμα 5. 27: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το Stacking Ensemble



Διάγραμμα 5. 28: MAE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το Stacking Ensemble

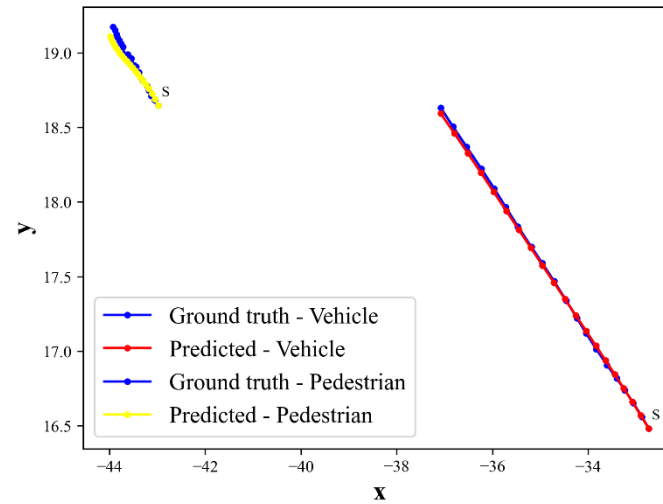
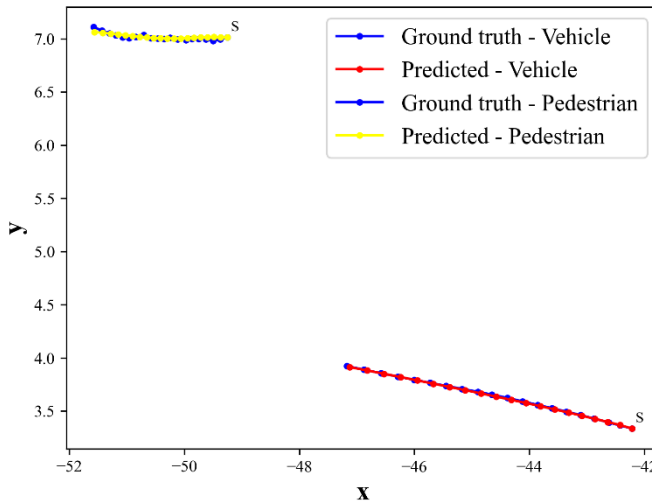
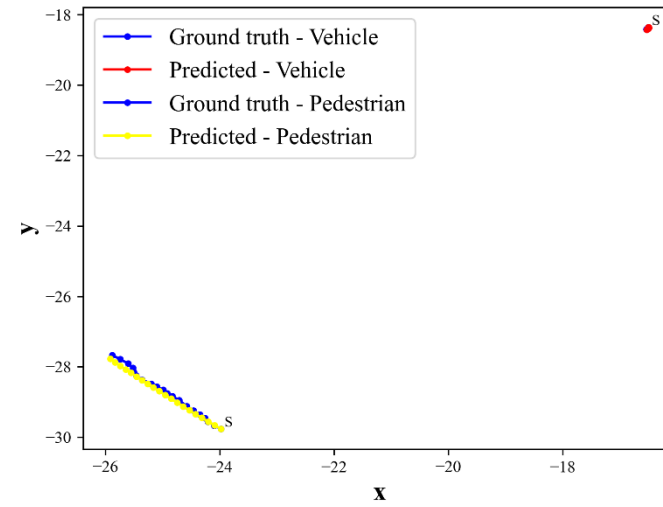
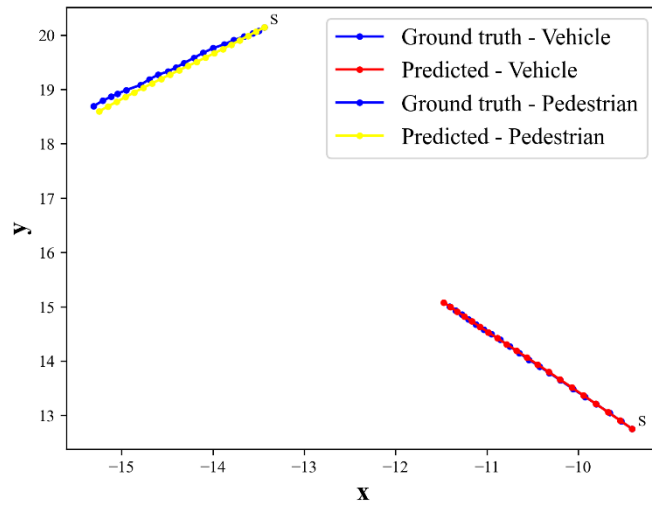


Διάγραμμα 5. 29: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το Stacking Ensemble



Διάγραμμα 5. 30: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το Stacking Ensemble

Στο διάγραμμα 5.27 παρουσιάζονται οι προβλεπόμενες και πραγματικές τροχιές για τέσσερις αλληλεπιδράσεις (τις ίδιες με τα άλλα πρότυπα), όπως προέκυψαν από το πρότυπο CNN_BiLSTM_ed_wf. Τα αποτελέσματα του Stacking Ensemble είναι εντυπωσιακά, με τις προβλέψεις των τροχιών (κυρίως των οχημάτων) σχεδόν να ταυτίζονται με τις πραγματικές τιμές αυτών.



Διάγραμμα 5. 31: Προβλέψεις και πραγματικές τροχιές για τέσσερα τυχαία επιλεγμένα συμβάντα (Stacking Ensemble)

5.2 Συγκριτικά αποτελέσματα

Στην παρούσα παράγραφο παρουσιάζεται η συγκριτική ακρίβεια προβλέψεων των επτά προτύπων. Από την προηγούμενη ήδη παράγραφο κατέστη ξεκάθαρο ότι το καλύτερο πρότυπο είναι το Stacking Ensemble με τα χαρακτηριστικά. Από την άλλη, το πρότυπο με τη χαμηλότερη ακρίβεια προβλέψεων με βάση τις τιμές των μετρικών αξιολόγησης των προτύπων είναι αυτό με την απλούστερη αρχιτεκτονική από όλα, δηλαδή το LSTM_ed. Βέβαια το εν λόγω πρότυπο, παρόλο που είναι το χειρότερο μεταξύ των επτά, παράγει ακριβείς προβλέψεις (<8cm) ακόμη και με τον περιορισμένο αριθμό δεδομένων που είναι διαθέσιμος.

Για την πιο εποπτική σύγκριση μεταξύ των τιμών των μετρικών αξιολόγησης μεταξύ των προτύπων, αποφασίστηκε να γίνει στάθμιση αυτών. Συγκεκριμένα η μέγιστη τιμή ανά κατηγορία μετρικής εξισώθηκε με την αυθαίρετη τιμή 100 ενώ η ελάχιστη (άρα και καλύτερη) με 0, όπως φαίνεται στον Πίνακα 5.8. Με αυτόν τον τρόπο είναι δυνατή η πιο εύκολη και άμεση σύγκριση των προτύπων. Επιπλέον στον σχετικό Πίνακα αναφέρονται οι μέγιστες (100) και οι ελάχιστες (0) τιμές ανά κατηγορία μετρικών καθώς και το αν συμπεριλαμβάνονται ή όχι τα χαρακτηριστικά σε κάθε πρότυπο. Ακόμη, λόγω του ότι σκοπός της εργασίας είναι η πρόβλεψη των τροχιών αποφασίστηκε στον εν λόγω Πίνακα να μην παρουσιαστούν οι μετρικές ταχυτήτων, αλλά μόνον οι μετρικές τροχιών. Ο αντίστοιχος Πίνακας με τις πραγματικές τιμές των μετρικών συμπεριλαμβάνεται στο παράρτημα, με αρίθμηση Π.16.

Πίνακας 5. 8: Συγκριτικές μετρικές τροχιών για τους δύο πράκτορες

Πρότυπα	Με/χωρίς χαρακτ.	Αυτόνομο Όχημα		Πεζός	
		Σταθμισμένο MAE	Σταθμισμένο MSE	Σταθμισμένο MAE	Σταθμισμένο MSE
LSTM_ed	Με	100	100	100	100
LSTM_ed_attention	Με	76	80	78	86
CNN_LSTM_ed_attention	Με	82	80	78	86
BiLSTM_ed	Χωρίς	41	60	72	100
BiLSTM_ed_attention	Με	41	20	22	29
CNN_BiLSTM_ed	Με	35	40	22	14
Stacking Ensemble	Με	0	0	0	0
	Διάστημα	0.051(0) με 0.068(100)	0.015(0) με 0.020(100)	0.056(0) με 0.074(100)	0.009(0) με 0.016(100)

Μοναδικό πρότυπο στο οποίο τα χαρακτηριστικά έδειξαν να μην έχουν ιδιαίτερη αποτελεσματικότητα είναι το BiLSTM_ed. Το πρώτο συμπέρασμα αυτής της διαπίστωσης είναι ότι η παρουσία των χαρακτηριστικών εν γένει είναι ιδιαιτέρως πρόσφορη για πρότυπα αρχιτεκτονικής encoder-decoder με χρήση BiLSTM. Επίσης μέσα από αυτή τη διαπίστωση συμπεραίνεται ότι η αρχιτεκτονική του BiLSTM_ed, δηλαδή ένα πρότυπο encoder-decoder με παρουσία ενός κελίου BiLSTM στο encoder και στο

decoder αντίστοιχα, δεν επωφελείται από την παρουσία των συγκεκριμένων χαρακτηριστικών που εξάχθηκαν. Με βάση αυτό, η μοναδική επεξήγηση για το γεγονός ότι στην περίπτωση της ίδιας αρχιτεκτονικής με προσθήκη μηχανισμού έμφασης ή ενός επιπέδου CNN η παρουσία των χαρακτηριστικών βελτιώνει αισθητά την ακρίβεια των προβλέψεων, είναι ότι η απόδοση των δύο αυτών επιπρόσθετων μεθοδολογιών βελτιώνεται σημαντικά μέσω των χαρακτηριστικών.

Η παρουσία του μηχανισμού έμφασης στην αρχιτεκτονική LSTM encoder-decoder δείχνει να έχει πολύ σημαντική επίδραση μειώνοντας αισθητά όλες τις μετρικές τροχιών και για τους δύο πράκτορες. Από την άλλη, η προσθήκη ενός επιπέδου CNN σε ένα πρότυπο LSTM encoder-decoder με μηχανισμό έμφασης, δείχνει να μην έχει σχεδόν καμία επίδραση στις μετρικές. Επομένως, καθίσταται σαφές ότι η παρουσία ενός επιπέδου CNN σε συνδυασμό με μηχανισμό έμφασης σε πρότυπο αρχιτεκτονικής encoder-decoder δεν έχει ιδιαίτερο νόημα.

Συγκρίνοντας τις σταθμισμένες μετρικές των προτύπων LSTM με τα πρότυπα BiLSTM, είναι ξεκάθαρο ότι η αμφίδρομη ανάλυση που επιτελείται μέσω της τεχνικής BiLSTM έχει ιδιαίτερως ευεργετικά αποτελέσματα. Συγκρίνοντας τώρα το πρότυπο BiLSTM_ed_nf με το BiLSTM_ed_attention_wf, παρατηρείται ότι όσον αφορά τα αυτόνομα οχήματα οι προβλέψεις έχουν το ίδιο MAE αλλά σημαντικά μικρότερο MSE. Αυτό υποδηλώνει ότι γενικότερα οι μέσες προβλέψεις είναι αντίστοιχου επιπέδου, αλλά ο μηχανισμός έμφασης μειώνει σημαντικά το βάρος των προβλέψεων, ουσιαστικά μαζεύοντας τις απομακρυσμένες προβλέψεις του BiLSTM_ed κοντύτερα στη συμπεριφορά των προβλέψεων του πρότυπου BiLSTM_ed_attention. Αντιθέτως, όσον αφορά τους πεζούς η σημαντική μείωση και των δύο μετρικών τροχιών υποδηλώνει ότι το πρότυπο BiLSTM_ed_attention έχει πολύ καλύτερη ποιότητα προβλέψεων. Επομένως είναι σαφές ότι η παρουσία μηχανισμού έμφασης σε ένα πρότυπο encoder-decoder με χρήση BiLSTM μπορεί να προσφέρει αισθητή βελτίωση.

Το πρότυπο CNN_BiLSTM_ed_wf έχει σχεδόν τις ίδιες τιμές MAE με το BiLSTM_ed_attention και για τους δύο πράκτορες, όμως τα MSE έχουν αισθητή διαφορά μεταξύ τους. Συγκεκριμένα το σταθμισμένο MSE του πρότυπου με το CNN επίπεδο για τα αυτόνομα οχήματα έχει τη διπλάσια τιμή ενώ για τους πεζούς έχει την μίση. Ουσιαστικά τα δύο πρότυπα παράγουν ισοδύναμες προβλέψεις, υποδεικνύοντας έτσι ότι ένα CNN επίπεδο έχει περίπου τα ίδια αποτελέσματα με τον μηχανισμό έμφασης όσον αφορά το BiLSTM encoder-decoder.

Έχοντας αναλύσει τις σχέσεις μεταξύ των προτύπων, το μόνο που απομένει είναι η κατάταξη τους. Όπως φαίνεται στην δεύτερη στήλη του Πίνακα 5.9, υπολογίζοντας τον μέσο όρο των σταθμισμένων μετρικών για καθένα από τις κατηγορίες προκύπτει μια σχετική κατάταξη. Όμως λαμβάνοντας υπόψιν το σχετικό βάρος κάθε μετρικής γίνεται μια επιπλέον σταθμισμένη βαθμολόγηση των προτύπων. Συγκεκριμένα, λόγω του ότι η εν λόγω κατάταξη αποτελεί βαθμολογία, οι μέσες μετρικές ανάγονται ώστε βαθμολογία 100 να έχει το καλύτερο πρότυπο και 0 το χειρότερο (σε αντίθεση με τις μέσες μετρικές). Τα MAE ουσιαστικά υποδηλώνουν την ακριβή απόκλιση των προβλέψεων με τις πραγματικές τιμές και έτσι αποφασίζεται να έχουν μεγαλύτερο βάρος. Επιπλέον, λόγω

του ότι η κίνηση των πεζών δεν ελέγχεται από το ίδιο το αυτόνομο όχημα, η ακριβής πρόβλεψη της τροχιάς αυτών αποτελεί προτεραιότητα. Για αυτόν τον λόγο επιλέγεται αυξημένο βάρος στις μετρικές των πεζών. Η σταθμισμένη βαθμολογία για κάθε πρότυπο προκύπτει από την σχέση:

$$\Sigma. B. = 0.3 \cdot MAE_{AV} + 0.1 \cdot MSE_{AV} + 0.4 \cdot MAE_{ped} + 0.2 \cdot MSE_{ped} \quad (5.1)$$

Με βάση τη μέση σταθμισμένη τιμή των μετρικών κάθε πρότυπου (στήλη 2 Πίνακα 5.9) αλλά και της σταθμισμένης βαθμολογίας του (στήλη 3 Πίνακα 5.9) προκύπτει η προσεγγιστική συγκριτική βαθμολογία κάθε πρότυπου, όπως αυτή φαίνεται στην τελευταία στήλη του Πίνακα 5.9 (Α το καλύτερο, Ε το χειρότερο). Σε αυτό το σημείο τονίζεται ότι οι βαθμολογίες αυτές είναι συγκριτικές και ότι τα διαστήματα των μετρικών είναι πολύ μικρά (τελευταία γραμμή Πίνακα 5.8) και επομένως όλα τα πρότυπα βρίσκονται περίπου στο ίδιο επίπεδο ακρίβειας προβλέψεων.

Πίνακας 5. 9: Βαθμολόγηση προτύπων

Πρότυπα	Μέση τιμή σταθμισμένων μετρικών	Σταθμισμένη βαθμολογία	Συγκριτική Βαθμολογία
LSTM_ed	100	0	E
LSTM_ed_attention	80	21	D
CNN_LSTM_ed_attention	81	19	D
BiLSTM_ed	68	33	C
BiLSTM_ed_attention	28	71	B
CNN_BiLSTM_ed	28	74	B
Stacking Ensemble	0	100	A

ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 Βασικά συμπεράσματα

Η παρούσα διπλωματική εργασία έχει ως αντικείμενο την εφαρμογή διαφόρων προτύπων βαθιάς μάθησης με σκοπό την πρόβλεψη των τροχιών αυτόνομων οχημάτων και πεζών κατά την αλληλεπίδραση τους. Τα πρότυπα που εφαρμόστηκαν ήταν επτά σε αριθμό. Τα έξι από αυτά έχουν αρχιτεκτονική encoder-decoder. Τα τρία από αυτά τα έξι αποτελούνται από ένα κελί LSTM στο encoder και ένα στο decoder. Τα υπόλοιπα τρία λειτουργούν αντίστοιχα αλλά με εφαρμογή αμφίδρομου LSTM. Το έβδομο και τελευταίο πρότυπο συνδυάζει τα υπόλοιπα έξι, σταθμίζοντας τα με τέτοιο τρόπο ώστε να μεγιστοποιήσει την ακρίβεια των προβλέψεων. Όσον αφορά τα εξαγόμενα χαρακτηριστικά, τα πρότυπα δοκιμάστηκαν με και χωρίς αυτά, ώστε να προσδιοριστεί ο βαθμός επιρροής τους.

Αρχικά όσον αφορά τις τεχνικές που δοκιμάστηκαν, η αρχιτεκτονική encoder-decoder δείχνει να έχει μεγάλη εφαρμοσιμότητα για τις προβλέψεις χρονοσειρών, καθώς όλα τα πρότυπα είχαν εντυπωσιακά αποτελέσματα. Επιπλέον, παρόλο που τα πρότυπα που χρησιμοποίησαν την τεχνική LSTM είχαν αρκετά μεγάλη ακρίβεια στις προβλέψεις τους, η αμφίδρομη ανάλυση που γίνεται στα BiLSTM πρότυπα έδειξε να αυξάνει αισθητά την ποιότητα προβλέψεων των προτύπων.

Σχετικά με τα δεδομένα, ήταν περιορισμένα σε όγκο, κάτι το οποίο μείωσε την ακρίβεια των προτύπων. Επιπλέον η εκτενής ανάγκη για απαλοιφή ορισμένων σφαλμάτων (noise) σε αυτά, αποτέλεσε ένα μικρό πρόβλημα, το οποίο βέβαια αντιμετωπίστηκε επαρκώς. Τέλος, το γεγονός ότι οι περισσότερες σκηνές ήταν περιορισμένης διάρκειας (<5 δευτερόλεπτα) δημιούργησε ορισμένους περιορισμούς ως προς την μεθοδολογία, με κυριότερο από αυτούς να είναι η αδυναμία ανάπτυξης προτύπων μακροπρόθεσμων προβλέψεων (5-8 δευτερόλεπτα). Παρόλα αυτά, τα δεδομένα ήταν επαρκή για βραχυπρόθεσμη πρόβλεψη τροχιών, και με ορισμένη επεξεργασία ήταν σε θέση να παράγουν εντυπωσιακά ακριβή πρότυπα.

Όσον αφορά τα χαρακτηριστικά (features), επιλέχθηκε να γίνει εξαγωγή των βασικότερων εξ αυτών. Με τις ταχύτητες και τις αποστάσεις να είναι δεδομένες, οι διαφορές αυτών και των γωνιών κίνησης θεωρήθηκε ότι επηρεάζουν σημαντικά τις επιλογές των πρακτόρων. Η δοκιμή των προτύπων με και χωρίς αυτά, επιβεβαίωσε την σημαντικότητα τους καθώς όλα τα πρότυπα εκτός του BiLSTM_ed είχαν βελτιωμένη απόδοση εν τη παρουσία τους. Το BiLSTM_ed είχε παρόμοιες αποδόσεις με και χωρίς τα χαρακτηριστικά και αυτό οφείλεται κυρίως στο γεγονός ότι η αμφίδρομη ανάλυση που γίνεται εντός αυτών είναι επαρκής από μόνη της και δεν δείχνει να απαιτεί επιπλέον δεδομένα στην μορφή χαρακτηριστικών ανά χρονικό βήμα. Βεβαίως, τα υπόλοιπα πρότυπα BiLSTM βελτιώσαν την απόδοσή τους με την συμπερίληψη των

χαρακτηριστικών, αλλά κάτι τέτοιο κατά κύριο λόγο έχει να κάνει με τα επιπλέον επίπεδα τους (CNN και μηχανισμός έμφασης-attention αντίστοιχα).

Τέλος, μεταξύ των επτά προτύπων, όπως είναι προφανές, το ακριβέστερο είναι το Stacking Ensemble, λόγω του ότι συνδυάζει τα υπόλοιπα έξι. Από την άλλη χειρότερο επίπεδο προβλέψεων παρήγαγε το πρώτο πρότυπο που δοκιμάστηκε, δηλαδή το LSTM_ed. Κάτι τέτοιο ήταν αναμενόμενο καθώς το πρότυπο αυτό αποτελεί την βάση των υπολοίπων και έχει την απλούστερη αρχιτεκτονική. Η επίδραση του μηχανισμού έμφασης έδειξε να είναι σημαντική στην περίπτωση των LSTM αλλά και των BiLSTM προτύπων, αποδεικνύοντας έτσι την ισχύ στις αρχιτεκτονικές encoder-decoder. Από την άλλη, ενώ στο πρότυπο BiLSTM η προσθήκη του επιπέδου CNN έδειξε να βελτιώνει την ακρίβεια των προβλέψεων, κάτι τέτοιο δεν συνέβη στην περίπτωση του LSTM encoder-decoder με έμφαση (attention). Ενδεχομένως αυτό το γεγονός να οφείλεται στο ότι η εφαρμογή επιπέδου CNN παράλληλα με τον μηχανισμό έμφασης ίσως είναι παραπανίσια. Ακόμη μπορεί να οφείλεται στο ότι η επίδραση ενός επιπέδου CNN σε πρότυπα LSTM encoder-decoder ενδεχομένως είναι μηδαμινή. Σε κάθε περίπτωση, καλύτερο πρότυπο προέκυψε το Stacking Ensemble και μεταξύ των υπολοίπων έξι το CNN_BiLSTM_ed_wf, όντας πολύ κοντά με το BiLSTM_ed_attention_wf.

6.2 Προτάσεις για περαιτέρω έρευνα

Με βάση τα παραπάνω συμπεράσματα προκύπτει μια σειρά από ευκαιρίες για περαιτέρω διερεύνηση της θεματικής περιοχής της παρούσας διπλωματικής εργασίας. Οι ευκαιρίες αυτές αφορούν κυρίως τις τεχνικές πρόβλεψης και τα δεδομένα που χρησιμοποιούνται. Με βάση αυτές τις ευκαιρίες προκύπτουν οι ακόλουθες προτάσεις για περαιτέρω έρευνα.

Ως προς τα διαθέσιμα δεδομένα, προτείνεται η ανάλυση με μεγαλύτερο αριθμό αλληλεπιδράσεων με σκοπό την βελτίωση της ακρίβειας των προβλέψεων. Επιπλέον, για να καταστεί δυνατός ο σχεδιασμός προτύπων μακροπρόθεσμης πρόβλεψης προτείνεται η χρήση δεδομένων με μεγαλύτερη διάρκεια ανά αλληλεπίδραση (τουλάχιστον 10 δευτερόλεπτα). Τέλος, προτείνεται η χρήση δεδομένων με μικρότερο αριθμό στάσιμων πρακτόρων καθώς οι αλληλεπιδράσεις με τέτοια χαρακτηριστικά μπορούν δυνητικά να οδηγήσουν σε παραπλανητικά επίπεδα ακρίβειας.

Σε σχέση με τις αρχιτεκτονικές βαθιάς μάθησης, με σκοπό την περαιτέρω βελτίωση της ακρίβειας τους θα μπορούσαν να δοκιμαστούν αρχιτεκτονικές που περιλαμβάνουν περισσότερα επίπεδα LSTM, BiLSTM αλλά και CNN. Ακόμη, στο πλαίσιο των προτύπων με αρχιτεκτονική encoder-decoder υπάρχουν πολυάριθμες τεχνικές, όπως για παράδειγμα τα προσφάτως ανεπτυγμένα Transformers, τα οποία υπό συνθήκες μπορούν να παράγουν πολύ ακριβή πρότυπα πρόβλεψης χρονοσειρών. Πέρα όμως από τα πρότυπα με την εν λόγω αρχιτεκτονική υπάρχουν πολλαπλές άλλες αρχιτεκτονικές βαθιάς μάθησης με εφαρμογή στην πρόβλεψη χρονοσειρών που μπορούν να εφαρμοστούν.

Σε σχέση με τη μηχανική χαρακτηριστικών, προτείνεται η εξαγωγή και δοκιμή περαιτέρω χαρακτηριστικών (features) με σκοπό την βελτίωση της ακρίβειας

προβλέψεων. Κάποια τέτοια χαρακτηριστικά είναι ο χρόνος για σύγκρουση (time to collision) και η απόσταση από τον προορισμό.

ΑΝΑΦΟΡΕΣ

Ευρωπαϊκό Κοινοβούλιο. *Αυτόνομα αυτοκίνητα στην ΕΕ: από επιστημονική φαντασία...σε απτή πραγματικότητα* (2019).

www.europarl.europa.eu/news/el/headlines/edconomy/20190110STO23102/autonoma-autokinita-stin-ee-apo-epistimoniki-fantasia-se-apti-pragmatikotita.

Altaf, Fouzia & Islam, Syed & Akhtar, Naveed & Janjua, Naeem. (2019). Going Deep in Medical Image Analysis: Concepts, Methods, Challenges and Future Directions.

Amini R.E., Katrakazas C., Riener A. & Antoniou C. (2021). Interaction of automated driving systems with pedestrians: challenges, current solutions, and recommendations for eHMI, *Transport Reviews*, 41:6, pp. 788-813.

Bahdanau, D., Cho, K., and Bengio, Y. (2014). Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv.org*.

Benefits of Self-Driving Vehicles. Coalition for future mobility. (2018). coalitionforfuturemobility.com/benefits-of-self-driving-vehicles.

Bhattacharyya, A., Olmeda Reino, D., Fritz, M., and Schiele, B. (2021). Euro-PVI: Pedestrian Vehicle Interactions in Dense Urban Centers. *arXiv.org*.

Breiman L. (2001). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, Volume 16(3), 2001, pp. 199-231.

Burns, C. G., Oliveira, L., Hung, V., Thomas, P., & Birrell, S. (2019). Pedestrian Attitudes to Shared-Space Interactions with Autonomous Vehicles – A Virtual Reality Study. In N. Stanton (Ed.), *Advances in Intelligent Systems and Computing*, pp. 307-316.

Chen, W., Wang, F. & Sun, H. (2021). S2TNet: Spatio-Temporal Transformer Networks for Trajectory Prediction in Autonomous Driving. *Proceedings of The 13th Asian Conference on Machine Learning*, in *Proceedings of Machine Learning Research*, pp. 454-469.

Cho, K., van Merriënboer, B., Gulcehre, C., Bougares, F., Schwenk, H., and Bengio, Y. (2014a). Learning phrase representations using RNN encoder-decoder for statistical machine translation. In *Proceedings of the Empirical Methods in Natural Language Processing (EMNLP 2014)*.

Christopher Olah (2014). *Conv Nets: A Modular Perspective*. *github.io*. <https://colah.github.io/posts/2014-07-Conv-Nets-Modular/>.

Cui H., Radosavljevic V., Chou F. C., Lin T. H., Nguyen T., Huang T. K., Schneider J. & Djuric N. (2018). Multimodal Trajectory Predictions for Autonomous Driving using Deep Convolutional Networks. *arXiv.org*.

Dai, S., Li, L., & Li, Z. (2019). Modeling Vehicle Interactions via Modified LSTM Models for Trajectory Prediction. *IEEE Access, Volume 7*, pp. 38287-38296.

Deep Learning - Wikipedia. En.wikipedia.org. (2022). Retrieved 12 September 2022, from https://en.wikipedia.org/wiki/Deep_learning#Definition.

Donahue J., Hendricks L. A., Rohrbach M., Venugopalan S., Guadarrama S., Saenko K., & Darrell T. (2014). Long-term Recurrent Convolutional Networks for Visual Recognition and Description. arXiv.org.

Geron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow : Concepts, Tools, and Techniques to Build Intelligent Systems (2nd ed.). O'Reilly Media, Incorporated.

Gindele T, Bredhtel S, Dillmann R (2010). A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments In: Proc. IEEE intelligent transportation systems conference, pp. 1625–1631.

Goodfellow I., Bengio Y., & Courville A. (2018). Deep learning. *Genet Program Evolvable Mach 19*, pp. 305–307.

Gupta A., Johnson J., Fei-Fei L., Savarese S., & Alahi A. (2018). Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. arXiv.org.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780.

Hoseini, F. S., Rahrovani, S., & Haghiri Chehreghani, M. (2020). A Generic Framework for Clustering Vehicle Motion Trajectories. arXiv.org.

Ivanovic B. & Pavone M. (2018) “The Trajectron: Probabilistic Multi-Agent Trajectory Modeling With Dynamic Spatiotemporal Graphs. arXiv.org.

James G., Witten D., Hastie T. & Tibshirani R. (2013). An introduction to statistical learning: with applications in R. New York :Springer.

Karle P., Geisslinger M., Betz J. & Lienkamp M., Scenario Understanding and Motion Prediction for Autonomous Vehicles—Review and Comparison. (2022). *IEEE Journals & Magazine | IEEE Xplore*.

Kostadinov, S. (2019). Understanding Encoder-Decoder Sequence to Sequence Model. <https://towardsdatascience.com/understanding-encoder-decoder-sequence-to-sequence-model-679e04af4346>.

Kyunghyun Cho K., van Merriënboer B., Gulcehre C., Bahdanau D., Bougares F., Schwenk H., Bengio Y. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv.org.

Langley, P. (2011). The changing science of machine learning. *Mach Learn* 82, pp. 275–279.

LeCun Y., Bottou L., Bengio Y. & P. Haffner P. (1998). Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, Volume 86, No. 11, pp. 2278-2324.

Lee N., Choi W., Vernaza P., Choy C. B., Torr H. S., & Chandraker M. (2017) DESIRE: Distant Future Prediction in Dynamic Scenes with Interacting Agents. [arXiv.org](https://arxiv.org/abs/1703.07492).

Machine Learning - Wikipedia. [En.wikipedia.org](https://en.wikipedia.org/wiki/Machine_learning#Artificial_intelligence). (2022). Retrieved 10 September 2022, from https://en.wikipedia.org/wiki/Machine_learning#Artificial_intelligence.

Mahadevan K., Sanoubari E., Somanath S., Young J.E. & Sharlin E., (2019). AV-Pedestrian Interaction Design Using a Pedestrian Mixed Traffic Simulator. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*.

Ngiam J. (2021). Scene Transformer: A unified architecture for predicting multiple agent trajectories. [arXiv.org](https://arxiv.org/abs/2103.04732).

Nikhil, N. & Morris B.T. (2018). Convolutional Neural Network for Trajectory Prediction. [arXiv.org](https://arxiv.org/abs/1808.08112).

Poibrenski A., Klusch M., Vozniak I., & Müller C. (2020). M2P3: multimodal multi-pedestrian path prediction by self-driving cars with egocentric vision. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing (SAC '20)*. Association for Computing Machinery, New York, NY, USA, pp. 190–197.

Rodríguez Palmeiro A., van der Kint S., Vissers L., Farah H., C.F. de Winter J., Hagenzieker M. (2018). Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment, *Transportation Research Part F: Traffic Psychology and Behaviour*, Volume 58, 2018, pp. 1005-1020.

Rudenko A., Palmieri L., Herman M., Kitani K.M., Gavrila D.M., Arras K.O. (2020). Human motion trajectory prediction: a survey. *The International Journal of Robotics Research*. 39(8):895-935.

Rumelhart, D., Hinton, G. and Williams, R. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), pp. 533-536.

Salzmann T., Ivanovic B., Chakravarty P., & Pavone, M. (2020) Trajectron++: Dynamically-Feasible Trajectory Forecasting With Heterogeneous Data. [arXiv.org](https://arxiv.org/abs/2006.04544).

Singh Gill, J., Pisu, P., Krovi, V. N., and Schmid, M. J. (2019). Behavior Identification and Prediction for a Probabilistic Risk Framework. [arXiv.org](https://arxiv.org/abs/1903.08761).

Styles O., Ross A. & Sanchez V. (2019). Forecasting Pedestrian Trajectory with Machine-Annotated Training Data. [arXiv.org](https://arxiv.org/abs/1903.08761).

Sutskever, I., Vinyals, O., and Le, Q. (2014). Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems (NIPS 2014)*.

Varghese, V., Chikaraishi, M. & Urata, (2021). Deep Learning in Transport Studies: A Meta-analysis on the Prediction Accuracy. SpringerLink.

Vidal, R., Bruna, J., Giryes, R., and Soatto, S. (2017). Mathematics of Deep Learning. arXiv.org.

Wiederer, T. (2016). Neural Networks in Javascript. webkid.io. <https://webkid.io/blog/neural-networks-in-javascript/>.

Yuan Y., Weng X., Ou Y., and Kitani K. (2021). AgentFormer: Agent-Aware Transformers for Socio-Temporal Multi-Agent Forecasting. arXiv.org.

Zhao T., Xu Y., Monfort M., Choi W., Baker C., Zhao Y., Wang Y. & Wu Y. N. (2019). Multi-Agent Tensor Fusion for Contextual Trajectory Prediction. arXiv.org.

Zvornicanin, E. (2022). Differences Between Bidirectional and Unidirectional LSTM. baeldung.com. <https://www.baeldung.com/cs/bidirectional-vs-unidirectional-lstm>.

ΠΑΡΑΡΤΗΜΑ

Μέρος 1: Πίνακες

Πίνακας Π. 1: Σύνοψη πρότυπου LSTM_ed_wf

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 150)	98400
repeat_vector (RepeatVector)	(None, 20, 150)	0
lstm_1 (LSTM)	(None, 20, 150)	180600
time_distributed (TimeDistributed)	(None, 20, 4)	604

=====
Total params: 279,604
Trainable params: 279,604
Non-trainable params: 0

Πίνακας Π. 2: Σύνοψη πρότυπου LSTM_ed_attention_wf

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 30, 13)]	0	[]
lstm (LSTM)	[(None, 30, 150), (None, 150), (None, 150)]	98400	['input_1[0][0]']
repeat_vector (RepeatVector)	(None, 20, 150)	0	['lstm[0][1]']
lstm_1 (LSTM)	(None, 20, 150)	180600	['repeat_vector[0][0]', 'lstm[0][1]', 'lstm[0][2]']
dot (Dot)	(None, 20, 30)	0	['lstm_1[0][0]', 'lstm[0][0]']
attention_vec (Activation)	(None, 20, 30)	0	['dot[0][0]']
dot_1 (Dot)	(None, 20, 150)	0	['attention_vec[0][0]', 'lstm[0][0]']
concatenate (Concatenate)	(None, 20, 300)	0	['dot_1[0][0]', 'lstm_1[0][0]']
time_distributed (TimeDistributed)	(None, 20, 4)	1204	['concatenate[0][0]']

=====
Total params: 280,204
Trainable params: 280,204
Non-trainable params: 0

Πίνακας Π. 3: Σύνοψη πρότυπου CNN_LSTM_ed_attention_wf

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 30, 13)]	0	[]
conv1d (Conv1D)	(None, 30, 64)	896	['input_1[0][0]']
lstm (LSTM)	[(None, 30, 150), (None, 150), (None, 150)]	129000	['conv1d[0][0]']
repeat_vector (RepeatVector)	(None, 20, 150)	0	['lstm[0][1]']
lstm_1 (LSTM)	(None, 20, 150)	180600	['repeat_vector[0][0]', 'lstm[0][1]', 'lstm[0][2]']
dot (Dot)	(None, 20, 30)	0	['lstm_1[0][0]', 'lstm[0][0]']
attention_vec (Activation)	(None, 20, 30)	0	['dot[0][0]']
dot_1 (Dot)	(None, 20, 150)	0	['attention_vec[0][0]', 'lstm[0][0]']
concatenate (Concatenate)	(None, 20, 300)	0	['dot_1[0][0]', 'lstm_1[0][0]']
time_distributed (TimeDistributed)	(None, 20, 4)	1204	['concatenate[0][0]']
=====			
Total params: 311,700			
Trainable params: 311,700			
Non-trainable params: 0			

Πίνακας Π. 4: Σύνοψη πρότυπου BiLSTM_ed_wf

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 30, 13)]	0	[]
bidirectional (Bidirectional)	[(None, 256), [(None, 128), (None, 128)], [(None, 128), (None, 128)]]	145408	['input_1[0][0]']
input_2 (InputLayer)	[(None, None, 4)]	0	[]
concatenate (Concatenate)	(None, 256)	0	['bidirectional[0][1]', 'bidirectional[0][3]']
concatenate_1 (Concatenate)	(None, 256)	0	['bidirectional[0][2]', 'bidirectional[0][4]']
rnn_1 (RNN)	[(None, None, 256), [(None, 256), (None, 256)]]	267264	['input_2[0][0]', 'concatenate[0][0]', 'concatenate_1[0][0]']
dense (Dense)	(None, None, 4)	1028	['rnn_1[0][0]']
=====			
Total params: 413,700			
Trainable params: 413,700			
Non-trainable params: 0			

Πίνακας Π. 5: Σύνοψη πρότυπου BiLSTM_ed_attention_wf

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 30, 13)]	0	[]
bidirectional (Bidirectional)	[(None, 30, 300), (None, 150), (None, 150), (None, 150), (None, 150)]	196800	['input_1[0][0]']
input_2 (InputLayer)	[(None, None, 4)]	0	[]
concatenate (Concatenate)	(None, 300)	0	['bidirectional[0][1]', 'bidirectional[0][3]']
concatenate_1 (Concatenate)	(None, 300)	0	['bidirectional[0][2]', 'bidirectional[0][4]']
lstm_1 (LSTM)	[(None, None, 300), (None, 300), (None, 300)]	366000	['input_2[0][0]', 'concatenate[0][0]', 'concatenate_1[0][0]']
dot (Dot)	(None, None, 30)	0	['lstm_1[0][0]', 'bidirectional[0][0]']
activation (Activation)	(None, None, 30)	0	['dot[0][0]']
dot_1 (Dot)	(None, None, 300)	0	['activation[0][0]', 'bidirectional[0][0]']
concatenate_2 (Concatenate)	(None, None, 600)	0	['dot_1[0][0]', 'lstm_1[0][0]']
time_distributed (TimeDistributed)	(None, None, 4)	2404	['concatenate_2[0][0]']
=====			
Total params: 565,204			
Trainable params: 565,204			
Non-trainable params: 0			

Πίνακας Π. 6: Σύνοψη πρότυπου CNN_BiLSTM_ed_wf

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, None, 13)]	0	[]
conv1d (Conv1D)	(None, None, 64)	896	['input_1[0][0]']
bidirectional (Bidirectional)	[(None, 256), [(None, 128), (None, 128)], [(None, 128), (None, 128)]]	197632	['conv1d[0][0]']
input_2 (InputLayer)	[(None, None, 4)]	0	[]
concatenate (Concatenate)	(None, 256)	0	['bidirectional[0][1]', 'bidirectional[0][3]']
concatenate_1 (Concatenate)	(None, 256)	0	['bidirectional[0][2]', 'bidirectional[0][4]']
rnn_1 (RNN)	[(None, None, 256), [(None, 256), (None, 256)]]	267264	['input_2[0][0]', 'concatenate[0][0]', 'concatenate_1[0][0]']
dense (Dense)	(None, None, 4)	1028	['rnn_1[0][0]']
=====			
Total params: 466,820			
Trainable params: 466,820			
Non-trainable params: 0			

Πίνακας Π. 7. 1: Σύνοψη πρότυπου Stacking Ensemble - Μέρος 1

Layer (type) Connected to	Output Shape	Param #	
ensemble_5_input_1 (InputLayer)	[(None, 30, 13)]	0	[]
ensemble_2_input_1 (InputLayer)	[(None, 30, 13)]	0	[]
ensemble_5_conv1d (Conv1D) ['ensemble_5_input_1[0][0]']	(None, 30, 64)	896	
ensemble_6_input_1 (InputLayer)	[(None, 30, 13)]	0	[]
ensemble_2_bidirectional (Bidirectional) ['ensemble_2_input_1[0][0]']	[(None, 30, 300), (None, 150), (None, 150), (None, 150), (None, 150)]	196800	
ensemble_5_lstm (LSTM) ['ensemble_5_conv1d[0][0]']	[(None, 30, 150), (None, 150), (None, 150)]	129000	
ensemble_6_lstm (LSTM) ['ensemble_6_input_1[0][0]']	[(None, 30, 150), (None, 150), (None, 150)]	98400	
ensemble_2_input_2 (InputLayer)	[(None, None, 4)]	0	[]
ensemble_2_concatenate (Concatenate) ['ensemble_2_bidirectional[0][1]', 'ensemble_2_bidirectional[0][3]']	(None, 300)	0	,
ensemble_2_concatenate_1 (Concatenate) ['ensemble_2_bidirectional[0][2]', 'ensemble_2_bidirectional[0][4]']	(None, 300)	0	,
ensemble_5_repeat_vector (RepeatVector) ['ensemble_5_lstm[0][1]']	(None, 20, 150)	0	

Πίνακας Π. 7.2: Σύνοψη πρότυπου Stacking Ensemble - Μέρος 2

ensemble_6_repeat_vector (RepeatVector)	(None, 20, 150)	0	
ensemble_2_lstm_1 (LSTM)	[(None, None, 300), (None, 300), (None, 300)]	366000	
ensemble_4_input_6 (InputLayer)	[(None, None, 13)]	0	[]
ensemble_5_lstm_1 (LSTM)	(None, 20, 150)	180600	
ensemble_6_lstm_1 (LSTM)	(None, 20, 150)	180600	
ensemble_2_dot (Dot)	(None, None, 30)	0	
ensemble_3_input_1 (InputLayer)	[(None, 30, 13)]	0	[]
ensemble_4_conv1d_2 (Conv1D)	(None, None, 64)	896	
ensemble_5_dot (Dot)	(None, 20, 30)	0	
ensemble_6_dot (Dot)	(None, 20, 30)	0	
lstm_input (InputLayer)	[(None, 30, 13)]	0	[]

Πίνακας Π. 7.3: Σύνοψη πρότυπου Stacking Ensemble - Μέρος 3

ensemble_2_activation (Activation) ['ensemble_2_dot[0][0]']	(None, None, 30)	0	
ensemble_3_bidirectional (Bidirectional) ['ensemble_3_input_1[0][0]']	[(None, 256), (None, 128), (None, 128), (None, 128), (None, 128)]	145408	
ensemble_4_bidirectional_2 (Bidirectional) ['ensemble_4_conv1d_2[0][0]']	[(None, 256), (None, 128), (None, 128), (None, 128), (None, 128)]	197632	
ensemble_5_attention_vec (Activation) ['ensemble_5_dot[0][0]']	(None, 20, 30)	0	
ensemble_6_attention_vec (Activation) ['ensemble_6_dot[0][0]']	(None, 20, 30)	0	
ensemble_7_lstm (LSTM) ['lstm_input[0][0]']	(None, 150)	98400	
ensemble_2_dot_1 (Dot) ['ensemble_2_activation[0][0]', 'ensemble_2_bidirectional[0][0]']	(None, None, 300)	0]
ensemble_3_input_2 (InputLayer)	[(None, None, 4)]	0	[]
ensemble_3_concatenate (Concatenate) ['ensemble_3_bidirectional[0][1]', 'ensemble_3_bidirectional[0][3]']	(None, 256)	0	, ,]
ensemble_3_concatenate_1 (Concatenate) ['ensemble_3_bidirectional[0][2]', 'ensemble_3_bidirectional[0][4]']	(None, 256)	0	, ,]
ensemble_4_input_7 (InputLayer)	[(None, None, 4)]	0	[]

Πίνακας Π. 7.4: Σύνοψη πρότυπου Stacking Ensemble - Μέρος 4

ensemble_4_concatenate_4 (Concatenate)	(None, 256)	0	
['ensemble_4_bidirectional_2[0][1]]',
ensemble_4_concatenate_4]',
'ensemble_4_bidirectional_2[0][3			']']
ensemble_4_concatenate_5 (Concatenate)	(None, 256)	0	
['ensemble_4_bidirectional_2[0][2]]',
ensemble_4_concatenate_5]',
'ensemble_4_bidirectional_2[0][4			']']
ensemble_5_dot_1 (Dot)	(None, 20, 150)	0	
['ensemble_5_attention_vec[0][0]'			,
'ensemble_5_lstm[0][0]'			,
ensemble_6_dot_1 (Dot)	(None, 20, 150)	0	
['ensemble_6_attention_vec[0][0]'			,
'ensemble_6_lstm[0][0]'			,
ensemble_7_repeat_vector (RepeatVector)	(None, 20, 150)	0	
['ensemble_7_lstm[0][0]'			,
ensemble_2_concatenate_2 (Concatenate)	(None, None, 600)	0	
['ensemble_2_dot_1[0][0]'			,
ensemble_2_concatenate_2]',
'ensemble_2_lstm_1[0][0]']',
ensemble_3_rnn_1 (RNN)	[(None, None, 256), 267264		
['ensemble_3_input_2[0][0]'			,
ensemble_3_rnn_1	[(None, 256),		,
'ensemble_3_concatenate[0][0]'			,
ensemble_3_rnn_1	(None, 256)]]		,
'ensemble_3_concatenate_1[0][0]']
ensemble_4_rnn_5 (RNN)	[(None, None, 256), 267264		
['ensemble_4_input_7[0][0]'			,
ensemble_4_rnn_5	[(None, 256),		,
'ensemble_4_concatenate_4[0][0]'			,
ensemble_4_rnn_5	(None, 256)]]		,
'ensemble_4_concatenate_5[0][0]'			']
ensemble_5_concatenate (Concatenate)	(None, 20, 300)	0	
['ensemble_5_dot_1[0][0]'			,
ensemble_5_concatenate]',
'ensemble_5_lstm_1[0][0]']',

Πίνακας Π. 7.5: Σύνοψη πρότυπου Stacking Ensemble - Μέρος 5

ensemble_6_concatenate (Concatenate)	(None, 20, 300)	0	
['ensemble_6_dot_1[0][0]', enate) 'ensemble_6_lstm_1[0][0]']			
ensemble_7_lstm_1 (LSTM)	(None, 20, 150)	180600	
['ensemble_7_repeat_vector[0][0]']			
]			
ensemble_2_time_distributed (TimeDistributed)	(None, None, 4)	2404	
['ensemble_2_concatenate_2[0][0]']			
imeDistributed)			
]			
ensemble_3_dense (Dense)	(None, None, 4)	1028	
['ensemble_3_rnn_1[0][0]']			
ensemble_4_dense_2 (Dense)	(None, None, 4)	1028	
['ensemble_4_rnn_5[0][0]']			
ensemble_5_time_distributed (TimeDistributed)	(None, 20, 4)	1204	
['ensemble_5_concatenate[0][0]']			
imeDistributed)			
ensemble_6_time_distributed (TimeDistributed)	(None, 20, 4)	1204	
['ensemble_6_concatenate[0][0]']			
imeDistributed)			
ensemble_7_time_distributed (TimeDistributed)	(None, 20, 4)	604	
['ensemble_7_lstm_1[0][0]']			
imeDistributed)			
concatenate (Concatenate)	(None, 20, 24)	0	
['ensemble_2_time_distributed[0][0]']			
0]'],			
'ensemble_3_dense[0][0]'],			
'ensemble_4_dense_2[0][0]'],			
'ensemble_5_time_distributed[0][0]']			
0]'],			
'ensemble_6_time_distributed[0][0]']			
0]'],			
'ensemble_7_time_distributed[0][0]']			
0]']			
dense (Dense)	(None, 20, 128)	3200	
['concatenate[0][0]']			
dense_1 (Dense)	(None, 20, 64)	8256	
['dense[0][0]']			

Πίνακας Π.7.6: Σύνοψη πρότυπου Stacking Ensemble - Μέρος 6

```

time_distributed (TimeDistribu (None, 20, 4)      260
['dense_1[0][0]']
ted)

=====
Total params: 2,328,948
Trainable params: 11,716
Non-trainable params: 2,317,232
    
```

Πίνακας Π. 8: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το LSTM_ed

Χρονικά Βήματα	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
1	0.121	0.000	0.266	0.000	0.129	0.000	0.061	0.000
2	0.117	0.017	0.179	0.009	0.138	0.016	0.068	0.001
3	0.106	0.024	0.129	0.007	0.138	0.028	0.060	0.003
4	0.105	0.031	0.116	0.008	0.141	0.037	0.070	0.004
5	0.098	0.038	0.086	0.010	0.140	0.046	0.073	0.006
6	0.109	0.044	0.146	0.010	0.138	0.053	0.074	0.008
7	0.100	0.050	0.079	0.011	0.138	0.060	0.070	0.010
8	0.108	0.056	0.163	0.013	0.137	0.066	0.063	0.011
9	0.109	0.062	0.145	0.015	0.137	0.072	0.068	0.013
10	0.113	0.069	0.184	0.018	0.134	0.077	0.069	0.015
11	0.104	0.075	0.099	0.021	0.134	0.082	0.066	0.016
12	0.108	0.079	0.156	0.021	0.142	0.087	0.089	0.018
13	0.119	0.084	0.244	0.022	0.142	0.092	0.085	0.021
14	0.116	0.089	0.245	0.026	0.140	0.097	0.078	0.022
15	0.119	0.093	0.246	0.026	0.139	0.101	0.070	0.025
16	0.117	0.099	0.192	0.031	0.142	0.105	0.071	0.026
17	0.131	0.104	0.255	0.034	0.146	0.109	0.079	0.028
18	0.136	0.109	0.291	0.036	0.151	0.113	0.072	0.030
19	0.143	0.114	0.265	0.040	0.164	0.118	0.098	0.032
20	0.154	0.118	0.305	0.039	0.176	0.122	0.114	0.036
M.O.	0.117	0.068	0.190	0.020	0.142	0.074	0.075	0.016
M.O. χωρίς χαρακτηριστικά	0.129	0.078	0.197	0.025	0.160	0.092	0.087	0.028

Πίνακας Π. 9: Εξέλιξη μετρικών ανά πρότυπο και εποχή

Epoch	LSTM_ed		LSTM_ed_attention		CNN_LSTM_ed_attention		BiLSTM_ed		BiLSTM_ed_attention		CNN_BiLSTM_ed		Stacking Ensemble	
	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
1	0.3504	0.3659	0.3449	0.3550	0.3664	0.3818	0.2289	0.2011	0.3502	0.3543	0.3844	0.4037	0.1294	0.1522
2	0.2715	0.2425	0.2650	0.2344	0.2755	0.2428	0.2144	0.1889	0.2610	0.2297	0.2751	0.2416	0.1022	0.1193
3	0.2510	0.2203	0.2430	0.2113	0.2539	0.2203	0.1999	0.1774	0.2410	0.2092	0.2508	0.2171	0.1012	0.1185
4	0.2347	0.2038	0.2264	0.1952	0.2352	0.2031	0.1879	0.1684	0.2257	0.1942	0.2331	0.2001	0.1005	0.1179
5	0.2213	0.1914	0.2103	0.1818	0.2214	0.1903	0.1794	0.1624	0.2091	0.1804	0.2145	0.1845	0.1002	0.1177
6	0.2084	0.1806	0.1969	0.1712	0.2076	0.1792	0.1710	0.1566	0.1960	0.1699	0.1999	0.1733	0.0997	0.1174
7	0.1973	0.1720	0.1860	0.1634	0.1967	0.1707	0.1640	0.1523	0.1844	0.1617	0.1873	0.1642	0.0996	0.1173
8	0.1882	0.1652	0.1760	0.1567	0.1865	0.1633	0.1571	0.1483	0.1731	0.1540	0.1758	0.1564	0.0990	0.1171
9	0.1791	0.1591	0.1684	0.1516	0.1771	0.1572	0.1532	0.1459	0.1643	0.1484	0.1668	0.1507	0.0987	0.1170
10	0.1714	0.1540	0.1616	0.1475	0.1702	0.1529	0.1468	0.1424	0.1569	0.1442	0.1577	0.1455	0.0985	0.1169
11	0.1646	0.1500	0.1560	0.1442	0.1626	0.1482	0.1426	0.1399	0.1487	0.1398	0.1515	0.1420	0.0983	0.1167
12	0.1595	0.1470	0.1498	0.1410	0.1564	0.1447	0.1394	0.1380	0.1435	0.1370	0.1456	0.1390	0.0981	0.1167
13	0.1541	0.1441	0.1455	0.1386	0.1518	0.1421	0.1351	0.1359	0.1381	0.1343	0.1396	0.1361	0.0978	0.1166
14	0.1499	0.1418	0.1412	0.1365	0.1472	0.1397	0.1334	0.1350	0.1341	0.1323	0.1357	0.1342	0.0977	0.1165
15	0.1456	0.1396	0.1382	0.1350	0.1435	0.1379	0.1301	0.1333	0.1304	0.1306	0.1315	0.1323	0.0975	0.1165
16	0.1427	0.1381	0.1352	0.1336	0.1399	0.1361	0.1274	0.1319	0.1268	0.1289	0.1279	0.1308	0.0974	0.1164
17	0.1394	0.1366	0.1325	0.1323	0.1367	0.1345	0.1252	0.1306	0.1239	0.1277	0.1251	0.1296	-	-
18	0.1370	0.1354	0.1303	0.1314	0.1341	0.1334	0.1235	0.1297	0.1217	0.1266	0.1228	0.1286	-	-
19	0.1344	0.1343	0.1283	0.1304	0.1315	0.1322	0.1224	0.1291	0.1193	0.1255	0.1206	0.1277	-	-
20	0.1327	0.1335	0.1264	0.1296	0.1298	0.1313	0.1202	0.1279	0.1174	0.1246	0.1185	0.1268	-	-
21	0.1308	0.1326	0.1248	0.1288	0.1280	0.1305	0.1192	0.1275	0.1156	0.1238	0.1170	0.1261	-	-
22	0.1294	0.1319	0.1236	0.1283	0.1262	0.1297	0.1176	0.1266	0.1143	0.1231	0.1153	0.1254	-	-
23	0.1279	0.1313	0.1222	0.1277	0.1246	0.1290	0.1168	0.1261	0.1129	0.1225	0.1139	0.1248	-	-
24	0.1266	0.1307	0.1211	0.1272	0.1233	0.1285	0.1157	0.1255	0.1115	0.1218	0.1127	0.1244	-	-

25	0.1254	0.1302	0.1201	0.1267	0.1223	0.1280	0.1149	0.1250	0.1106	0.1213	0.1117	0.1239	-	-
26	0.1245	0.1298	0.1192	0.1264	0.1213	0.1276	0.1142	0.1247	0.1095	0.1208	0.1108	0.1235	-	-
27	0.1236	0.1294	0.1184	0.1260	0.1203	0.1272	0.1137	0.1243	0.1088	0.1204	0.1099	0.1231	-	-
28	0.1228	0.1291	0.1177	0.1257	0.1196	0.1268	0.1128	0.1238	0.1080	0.1200	0.1091	0.1228	-	-
29	0.1221	0.1287	0.1170	0.1254	0.1188	0.1265	0.1124	0.1234	0.1074	0.1197	0.1085	0.1225	-	-
30	0.1214	0.1285	0.1165	0.1252	0.1182	0.1262	0.1119	0.1231	0.1068	0.1193	0.1078	0.1223	-	-
31	0.1208	0.1282	0.1160	0.1250	0.1175	0.1260	0.1115	0.1228	0.1063	0.1190	0.1073	0.1220	-	-
32	0.1203	0.1280	0.1155	0.1248	0.1171	0.1257	0.1111	0.1225	0.1058	0.1187	0.1069	0.1218	-	-
33	0.1199	0.1278	0.1151	0.1246	0.1166	0.1255	0.1108	0.1222	0.1054	0.1185	0.1064	0.1216	-	-
34	0.1194	0.1276	0.1148	0.1244	0.1161	0.1253	0.1104	0.1220	0.1050	0.1183	0.1060	0.1214	-	-
35	0.1191	0.1275	0.1144	0.1242	0.1158	0.1252	0.1101	0.1217	0.1047	0.1181	0.1057	0.1213	-	-
36	0.1187	0.1273	0.1141	0.1241	0.1155	0.1250	0.1099	0.1215	0.1043	0.1179	0.1053	0.1211	-	-
37	0.1185	0.1272	0.1139	0.1240	0.1152	0.1249	0.1097	0.1213	0.1041	0.1177	0.1051	0.1210	-	-
38	0.1181	0.1271	0.1136	0.1239	0.1149	0.1248	0.1095	0.1211	0.1038	0.1176	0.1048	0.1209	-	-
39	0.1179	0.1270	0.1134	0.1238	0.1146	0.1247	0.1093	0.1209	0.1036	0.1174	0.1046	0.1208	-	-
40	0.1177	0.1269	0.1132	0.1237	0.1144	0.1246	0.1091	0.1208	0.1034	0.1173	0.1044	0.1207	-	-
41	0.1175	0.1268	0.1131	0.1236	0.1142	0.1245	0.1090	0.1206	0.1033	0.1172	0.1042	0.1206	-	-
42	0.1173	0.1267	0.1128	0.1236	0.1140	0.1244	0.1088	0.1205	0.1031	0.1171	0.1041	0.1205	-	-
43	0.1172	0.1266	0.1128	0.1235	0.1139	0.1243	0.1087	0.1204	0.1030	0.1170	0.1039	0.1204	-	-
44	0.1170	0.1266	0.1126	0.1234	0.1138	0.1243	0.1086	0.1202	0.1028	0.1169	0.1038	0.1204	-	-
45	0.1169	0.1265	0.1125	0.1234	0.1136	0.1242	-	-	0.1027	0.1168	0.1036	0.1203	-	-
46	0.1168	0.1265	0.1124	0.1233	0.1135	0.1241	-	-	0.1026	0.1167	0.1035	0.1203	-	-
47	0.1167	0.1264	0.1123	0.1233	0.1134	0.1241	-	-	0.1025	0.1167	0.1034	0.1202	-	-
48	0.1166	0.1264	0.1122	0.1233	0.1133	0.1241	-	-	0.1024	0.1166	0.1034	0.1202	-	-
49	0.1165	0.1263	0.1122	0.1232	0.1132	0.1240	-	-	0.1024	0.1166	-	-	-	-
50	0.1164	0.1263	0.1121	0.1232	0.1131	0.1240	-	-	-	-	-	-	-	-
51	-	-	-	-	0.1131	0.1240	-	-	-	-	-	-	-	-

Πίνακας Π. 10: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το LSTM_ed_attention

Χρονικά Βήματα	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
1	0.119	0.000	0.270	0.000	0.119	0.000	0.049	0.000
2	0.115	0.017	0.181	0.009	0.136	0.015	0.065	0.001
3	0.103	0.024	0.130	0.007	0.136	0.027	0.059	0.003
4	0.102	0.030	0.114	0.008	0.138	0.037	0.069	0.004
5	0.096	0.037	0.081	0.011	0.137	0.045	0.072	0.006
6	0.105	0.042	0.143	0.010	0.135	0.052	0.071	0.007
7	0.098	0.048	0.078	0.011	0.133	0.058	0.068	0.009
8	0.106	0.053	0.161	0.012	0.132	0.064	0.061	0.011
9	0.106	0.059	0.145	0.015	0.133	0.069	0.064	0.012
10	0.110	0.065	0.186	0.017	0.132	0.073	0.066	0.014
11	0.101	0.071	0.099	0.020	0.132	0.078	0.064	0.015
12	0.105	0.075	0.153	0.020	0.140	0.082	0.087	0.017
13	0.116	0.079	0.237	0.021	0.140	0.087	0.085	0.019
14	0.114	0.084	0.238	0.024	0.137	0.092	0.077	0.021
15	0.116	0.088	0.245	0.024	0.133	0.096	0.067	0.023
16	0.114	0.094	0.190	0.029	0.136	0.099	0.067	0.024
17	0.128	0.099	0.253	0.032	0.140	0.103	0.075	0.026
18	0.132	0.103	0.285	0.033	0.143	0.106	0.066	0.028
19	0.140	0.108	0.262	0.037	0.158	0.109	0.092	0.029
20	0.150	0.112	0.310	0.036	0.170	0.114	0.109	0.032
M.O.	0.114	0.064	0.188	0.019	0.138	0.070	0.072	0.015
M.O. χωρίς χαρακτηριστικά	0.127	0.077	0.195	0.025	0.159	0.091	0.087	0.028

Πίνακας Π. 11: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το CNN_LSTM_ed_attention

Χρονικά Βήματα	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
1	0.120	0.000	0.269	0.000	0.123	0.000	0.000	0.000
2	0.117	0.018	0.183	0.009	0.135	0.016	0.001	0.001
3	0.104	0.025	0.130	0.007	0.135	0.028	0.003	0.003
4	0.101	0.030	0.114	0.008	0.137	0.036	0.004	0.004
5	0.095	0.037	0.082	0.010	0.136	0.044	0.005	0.005
6	0.105	0.042	0.144	0.010	0.135	0.051	0.007	0.007
7	0.098	0.048	0.078	0.011	0.134	0.057	0.009	0.009
8	0.107	0.053	0.160	0.012	0.134	0.063	0.011	0.011
9	0.107	0.059	0.144	0.014	0.135	0.068	0.012	0.012
10	0.111	0.065	0.184	0.017	0.133	0.073	0.013	0.013
11	0.101	0.071	0.100	0.020	0.135	0.077	0.015	0.015
12	0.105	0.075	0.156	0.020	0.142	0.082	0.017	0.017
13	0.117	0.079	0.242	0.021	0.142	0.087	0.019	0.019
14	0.115	0.085	0.242	0.025	0.138	0.092	0.021	0.021
15	0.117	0.088	0.242	0.025	0.136	0.096	0.023	0.023
16	0.115	0.095	0.191	0.030	0.139	0.100	0.024	0.024
17	0.128	0.099	0.258	0.032	0.143	0.104	0.026	0.026
18	0.133	0.104	0.292	0.034	0.146	0.108	0.028	0.028
19	0.140	0.109	0.263	0.038	0.160	0.111	0.030	0.030
20	0.151	0.113	0.303	0.037	0.172	0.116	0.033	0.033
M.O.	0.114	0.065	0.189	0.019	0.139	0.070	0.073	0.015
M.O. χωρίς χαρακτηριστικά	0.122	0.071	0.194	0.022	0.152	0.083	0.081	0.022

Πίνακας Π. 12: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το BiLSTM_ed

Χρονικά Βήματα	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
1	0.116	0.000	0.272	0.000	0.121	0.000	0.055	0.000
2	0.112	0.017	0.184	0.009	0.133	0.015	0.062	0.001
3	0.099	0.024	0.128	0.007	0.132	0.027	0.055	0.002
4	0.096	0.029	0.114	0.008	0.133	0.035	0.065	0.004
5	0.089	0.036	0.082	0.010	0.131	0.042	0.068	0.005
6	0.099	0.040	0.141	0.010	0.128	0.048	0.067	0.006
7	0.091	0.045	0.075	0.010	0.127	0.053	0.064	0.008
8	0.098	0.050	0.159	0.012	0.125	0.058	0.057	0.009
9	0.097	0.055	0.142	0.014	0.125	0.062	0.061	0.010
10	0.101	0.060	0.183	0.016	0.124	0.066	0.062	0.011
11	0.091	0.065	0.097	0.019	0.125	0.069	0.059	0.012
12	0.095	0.068	0.151	0.018	0.132	0.072	0.082	0.013
13	0.105	0.072	0.238	0.020	0.131	0.077	0.078	0.014
14	0.102	0.077	0.238	0.023	0.127	0.080	0.072	0.015
15	0.103	0.080	0.241	0.022	0.124	0.084	0.063	0.017
16	0.099	0.085	0.188	0.027	0.127	0.087	0.062	0.018
17	0.111	0.089	0.249	0.029	0.129	0.089	0.068	0.019
18	0.113	0.092	0.281	0.029	0.130	0.092	0.058	0.020
19	0.117	0.096	0.253	0.033	0.144	0.095	0.083	0.021
20	0.125	0.098	0.294	0.031	0.157	0.098	0.100	0.023
M.O.	0.103	0.059	0.185	0.017	0.13	0.063	0.067	0.011
M.O. χωρίς χαρακτηριστικά	0.104	0.058	0.183	0.018	0.137	0.069	0.072	0.016

Πίνακας Π. 13: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το BiLSTM_ed_attention

Χρονικά Βήματα	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
1	0.110	0.000	0.240	0.000	0.117	0.000	0.048	0.000
2	0.108	0.017	0.172	0.009	0.130	0.015	0.059	0.001
3	0.097	0.023	0.129	0.007	0.129	0.026	0.053	0.002
4	0.094	0.028	0.111	0.007	0.130	0.035	0.061	0.003
5	0.088	0.034	0.079	0.010	0.129	0.042	0.065	0.005
6	0.098	0.039	0.141	0.009	0.126	0.047	0.064	0.006
7	0.089	0.043	0.075	0.010	0.124	0.051	0.061	0.007
8	0.098	0.048	0.158	0.011	0.122	0.056	0.056	0.008
9	0.098	0.053	0.141	0.013	0.122	0.059	0.058	0.009
10	0.101	0.058	0.181	0.015	0.121	0.063	0.059	0.010
11	0.091	0.063	0.095	0.018	0.123	0.066	0.058	0.010
12	0.095	0.066	0.151	0.017	0.130	0.069	0.081	0.011
13	0.106	0.070	0.235	0.018	0.130	0.072	0.076	0.013
14	0.103	0.075	0.236	0.021	0.127	0.076	0.071	0.014
15	0.105	0.078	0.241	0.021	0.123	0.079	0.061	0.015
16	0.103	0.083	0.188	0.025	0.125	0.082	0.059	0.016
17	0.114	0.088	0.248	0.028	0.126	0.084	0.065	0.017
18	0.115	0.092	0.277	0.029	0.128	0.087	0.055	0.018
19	0.119	0.095	0.251	0.032	0.142	0.090	0.079	0.019
20	0.128	0.098	0.294	0.030	0.155	0.094	0.098	0.022
M.O.	0.103	0.058	0.182	0.016	0.128	0.060	0.064	0.010
M.O. χωρίς χαρακτηριστικά	0.115	0.068	0.19	0.021	0.146	0.080	0.078	0.021

Πίνακας Π. 14: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το CNN_BiLSTM_ed

Χρονικά Βήματα	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
1	0.113	0.000	0.267	0.000	0.124	0.000	0.057	0.000
2	0.110	0.017	0.178	0.009	0.130	0.015	0.062	0.001
3	0.096	0.023	0.127	0.007	0.127	0.026	0.053	0.002
4	0.094	0.028	0.113	0.007	0.129	0.034	0.061	0.003
5	0.088	0.035	0.080	0.010	0.127	0.041	0.065	0.005
6	0.096	0.039	0.140	0.010	0.126	0.047	0.066	0.006
7	0.088	0.044	0.075	0.010	0.125	0.052	0.062	0.007
8	0.096	0.049	0.158	0.011	0.124	0.056	0.056	0.008
9	0.096	0.053	0.142	0.013	0.124	0.060	0.059	0.009
10	0.100	0.058	0.183	0.015	0.123	0.063	0.061	0.010
11	0.090	0.064	0.096	0.018	0.124	0.066	0.059	0.011
12	0.093	0.066	0.151	0.018	0.131	0.070	0.082	0.012
13	0.103	0.070	0.239	0.019	0.130	0.074	0.078	0.013
14	0.100	0.075	0.239	0.022	0.127	0.077	0.072	0.014
15	0.101	0.077	0.241	0.021	0.123	0.080	0.062	0.016
16	0.099	0.082	0.188	0.026	0.125	0.083	0.061	0.016
17	0.110	0.087	0.247	0.029	0.128	0.086	0.067	0.017
18	0.112	0.090	0.277	0.029	0.130	0.088	0.058	0.018
19	0.117	0.094	0.253	0.033	0.142	0.091	0.081	0.019
20	0.125	0.096	0.298	0.031	0.153	0.094	0.097	0.021
M.O.	0.101	0.057	0.185	0.017	0.128	0.060	0.066	0.01
M.O. χωρίς χαρακτηριστικά	0.112	0.066	0.189	0.02	0.144	0.077	0.077	0.019

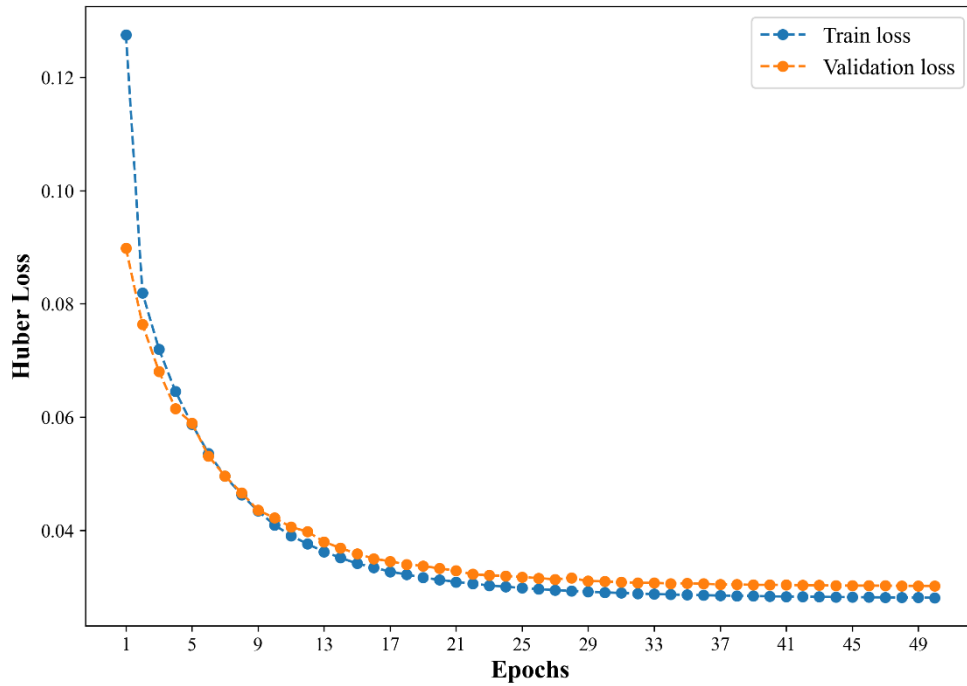
Πίνακας Π. 15: Αριθμητική εξέλιξη μετρικών ανά κατηγορία και πράκτορα για το Stacking Ensemble

Χρονικά Βήματα	Αυτόνομο Όχημα				Πεζός			
	MAE		MSE		MAE		MSE	
	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
1	0.093	0.000	0.250	0.000	0.110	0.000	0.047	0.000
2	0.097	0.015	0.173	0.009	0.124	0.014	0.057	0.001
3	0.087	0.021	0.126	0.007	0.123	0.025	0.051	0.002
4	0.085	0.025	0.110	0.007	0.125	0.033	0.060	0.003
5	0.079	0.030	0.077	0.009	0.124	0.039	0.063	0.004
6	0.089	0.034	0.138	0.009	0.121	0.044	0.063	0.005
7	0.080	0.038	0.073	0.009	0.120	0.048	0.060	0.007
8	0.089	0.042	0.156	0.010	0.118	0.052	0.054	0.007
9	0.089	0.047	0.140	0.012	0.118	0.055	0.056	0.008
10	0.092	0.051	0.180	0.014	0.117	0.058	0.058	0.009
11	0.082	0.056	0.094	0.017	0.119	0.061	0.056	0.009
12	0.086	0.058	0.149	0.016	0.126	0.064	0.078	0.010
13	0.096	0.062	0.234	0.017	0.125	0.067	0.075	0.011
14	0.093	0.066	0.234	0.020	0.122	0.070	0.069	0.012
15	0.094	0.068	0.238	0.019	0.119	0.073	0.059	0.013
16	0.091	0.073	0.185	0.023	0.120	0.076	0.057	0.014
17	0.102	0.077	0.245	0.026	0.122	0.078	0.063	0.015
18	0.103	0.080	0.274	0.025	0.123	0.081	0.053	0.016
19	0.106	0.083	0.248	0.028	0.136	0.083	0.076	0.017
20	0.113	0.085	0.291	0.026	0.148	0.087	0.094	0.018
M.O.	0.092	0.051	0.181	0.015	0.123	0.056	0.062	0.009
M.O. χωρίς χαρακτηριστικά	0.101	0.058	0.184	0.018	0.136	0.07	0.072	0.016

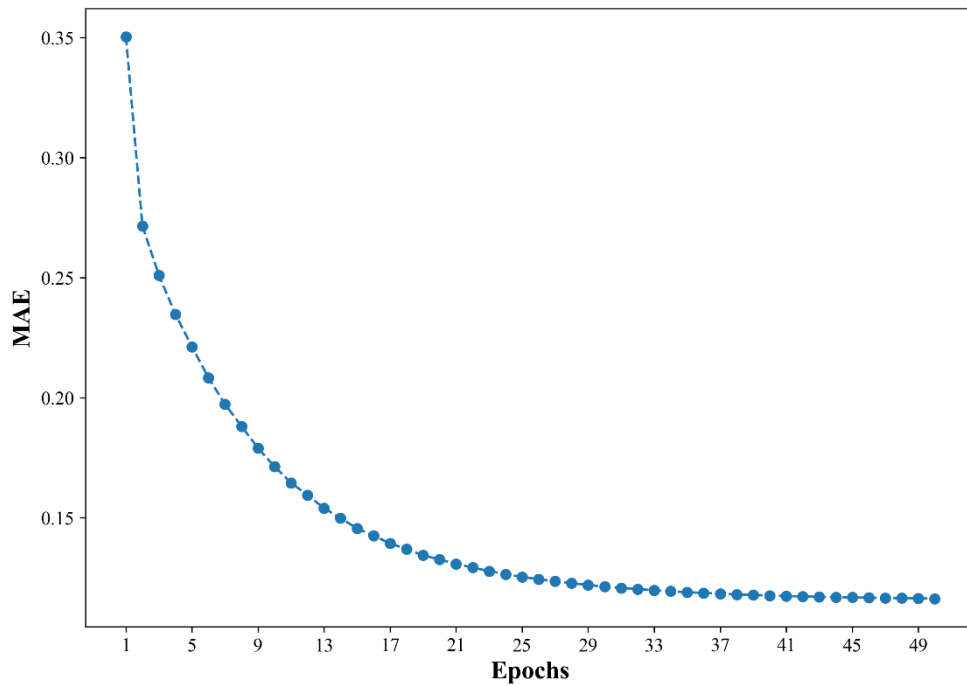
Πίνακας Π. 16: MAE και MSE ανά πράκτορα, κατηγορία και πρότυπο

Πρότυπα	+/- features	Αυτόνομο Όχημα				Πεζός			
		MAE		MSE		MAE		MSE	
		Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση	Ταχύτητα	Θέση
LSTM_ed	+	0.117	0.068	0.190	0.020	0.142	0.074	0.075	0.016
LSTM_ed_attention	+	0.114	0.064	0.188	0.019	0.138	0.070	0.072	0.015
CNN_LSTM_ed_attention	+	0.114	0.065	0.189	0.019	0.139	0.070	0.073	0.015
BiLSTM_ed	-	0.104	0.058	0.183	0.018	0.137	0.069	0.072	0.016
BiLSTM_ed_attention	+	0.103	0.058	0.182	0.016	0.128	0.060	0.064	0.011
CNN_BiLSTM_ed	+	0.101	0.057	0.185	0.017	0.128	0.060	0.066	0.010
Stacking Ensemble	+	0.092	0.051	0.181	0.015	0.123	0.056	0.062	0.009

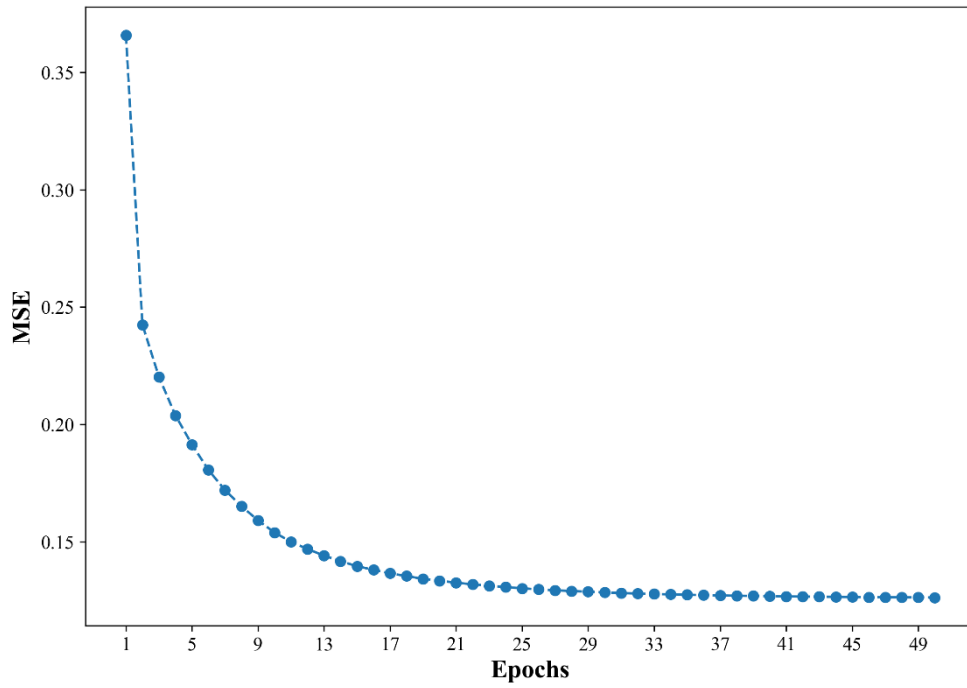
Μέρος 2: Διαγράμματα



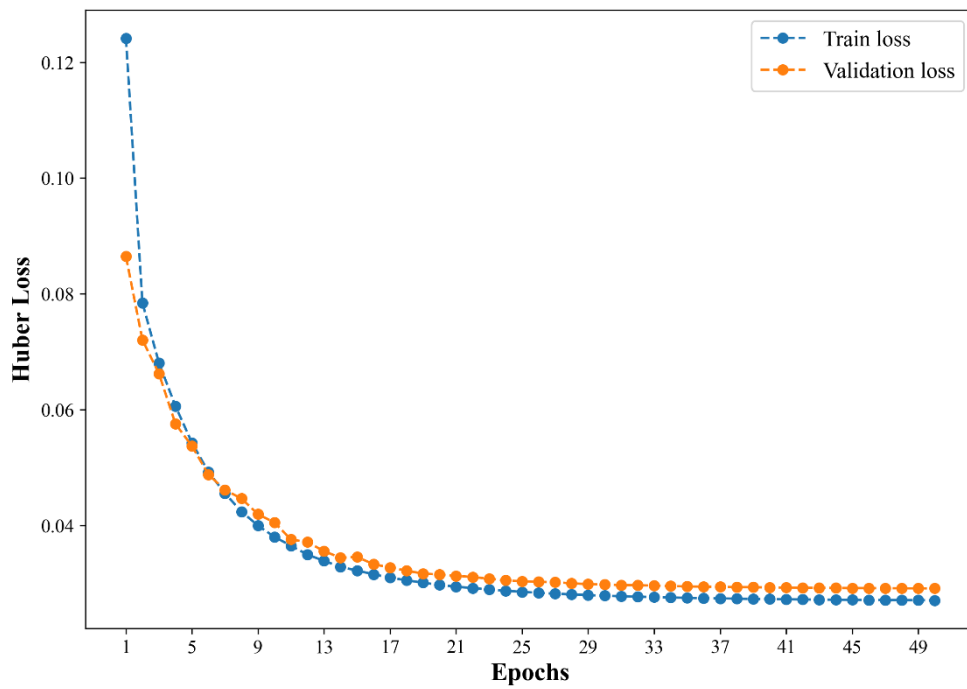
Διάγραμμα Π. 1: Εξέλιξη απώλειας Huber ανά εποχή για το LSTM_ed_wf



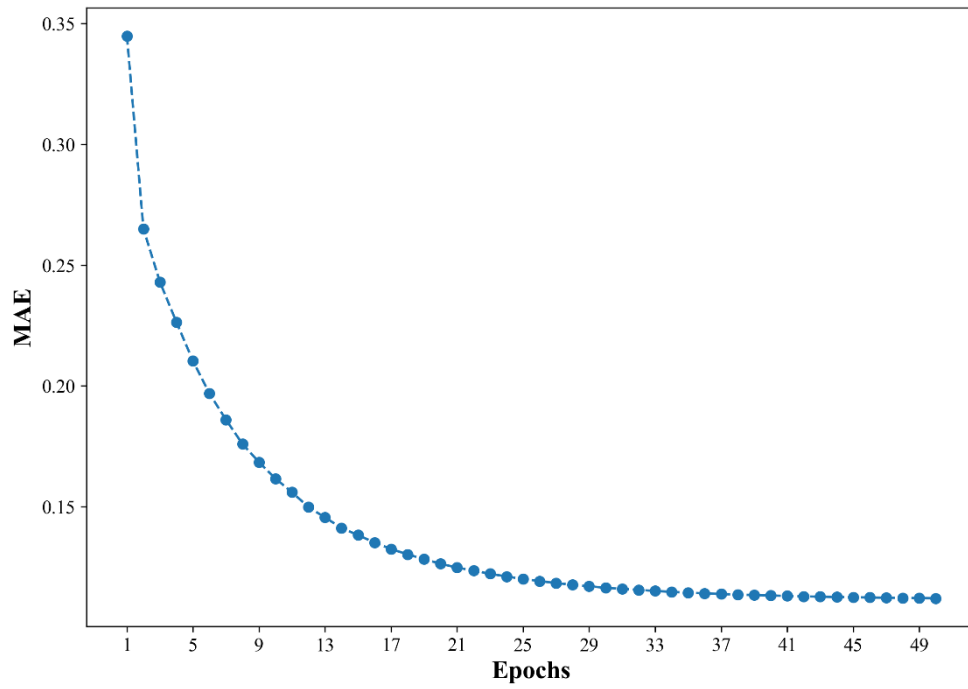
Διάγραμμα Π. 2: Εξέλιξη MAE ταχύτητας ανά εποχή για το LSTM_ed_wf



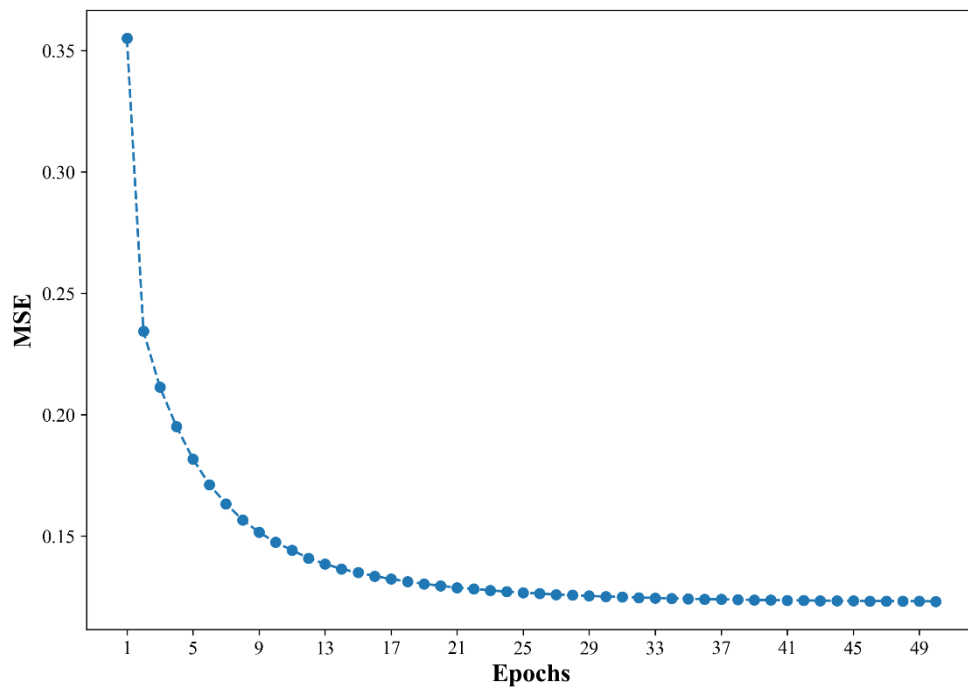
Διάγραμμα Π. 3: Εξέλιξη MSE ανά εποχή για το LSTM_ed_wf



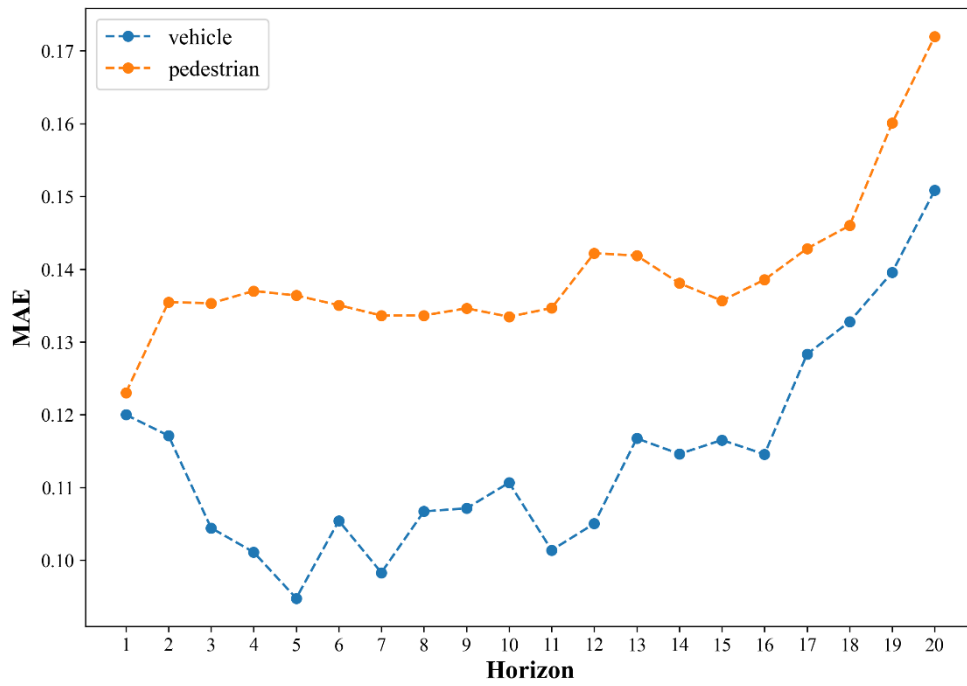
Διάγραμμα Π. 4: Εξέλιξη απώλειας Huber ανά εποχή για το LSTM_ed_attention_wf



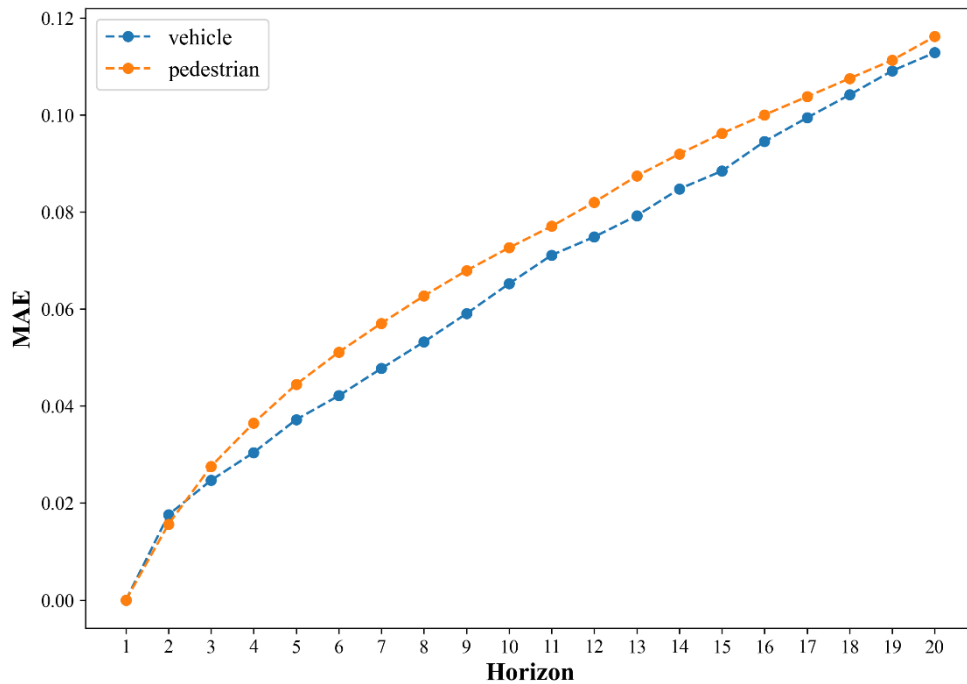
Διάγραμμα Π. 5: Εξέλιξη MSE ανά εποχή για το LSTM_ed_attention_wf



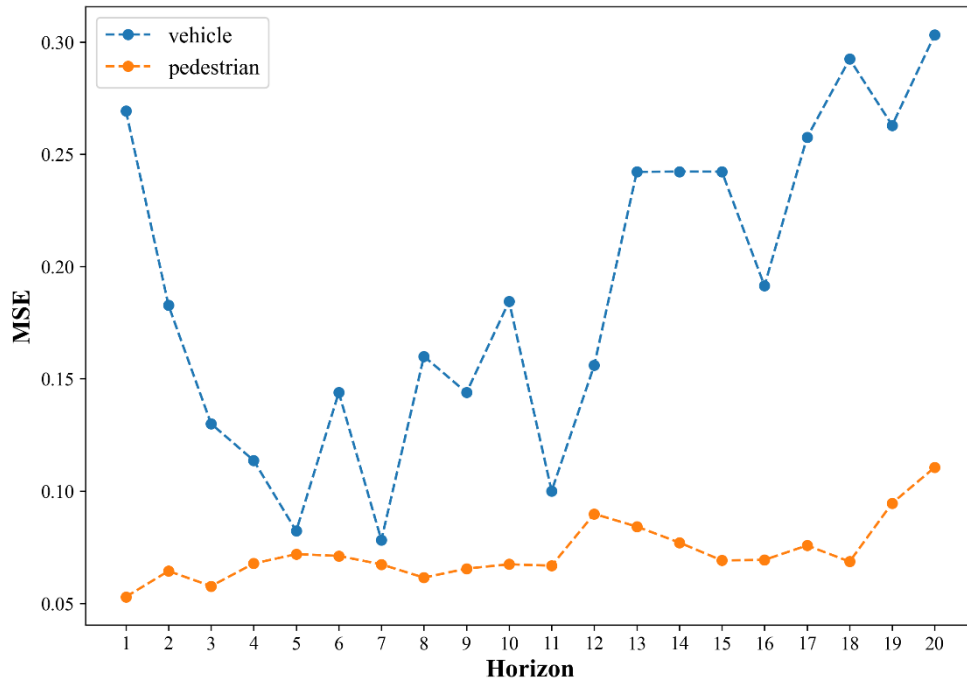
Διάγραμμα Π. 6: Εξέλιξη MSE ανά εποχή για το LSTM_ed_attention_wf



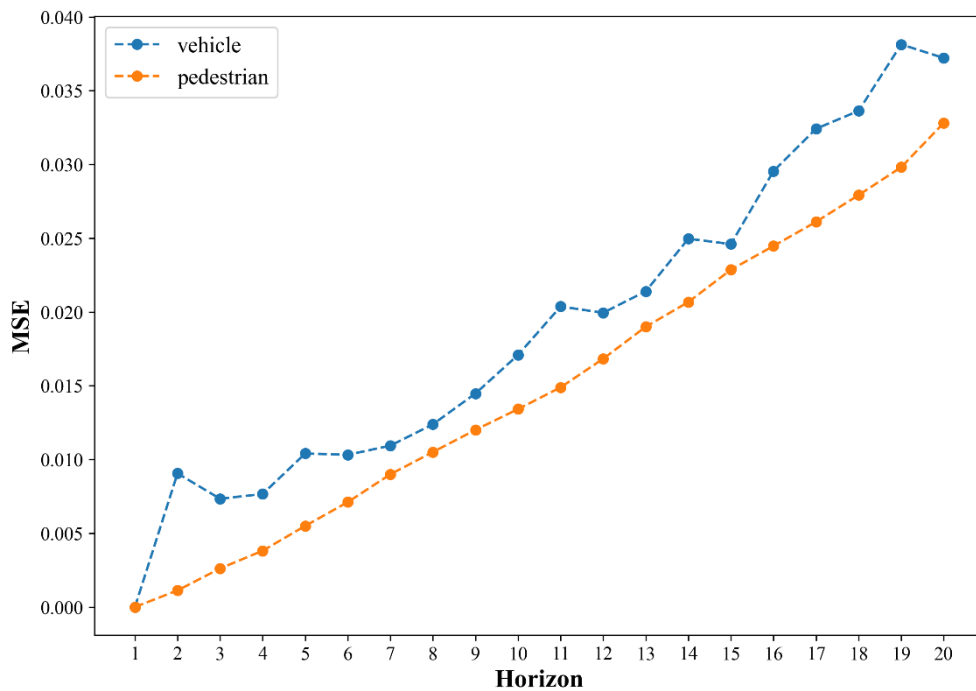
Διάγραμμα Π. 7: MAE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το CNN_LSTM_ed_attention_wf



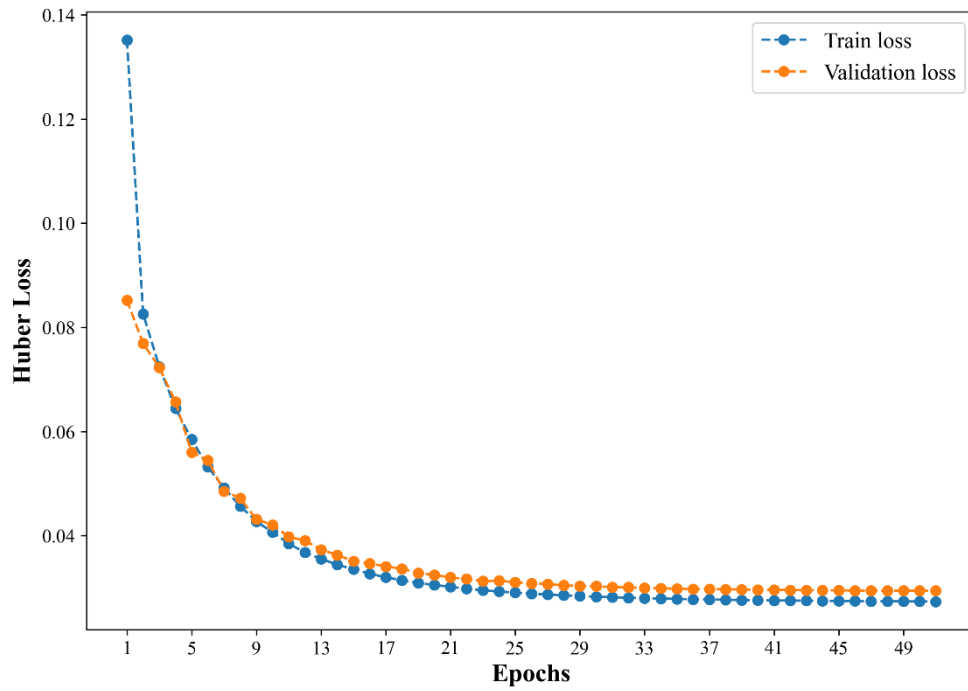
Διάγραμμα Π. 8: MAE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το CNN_LSTM_ed_attention_wf



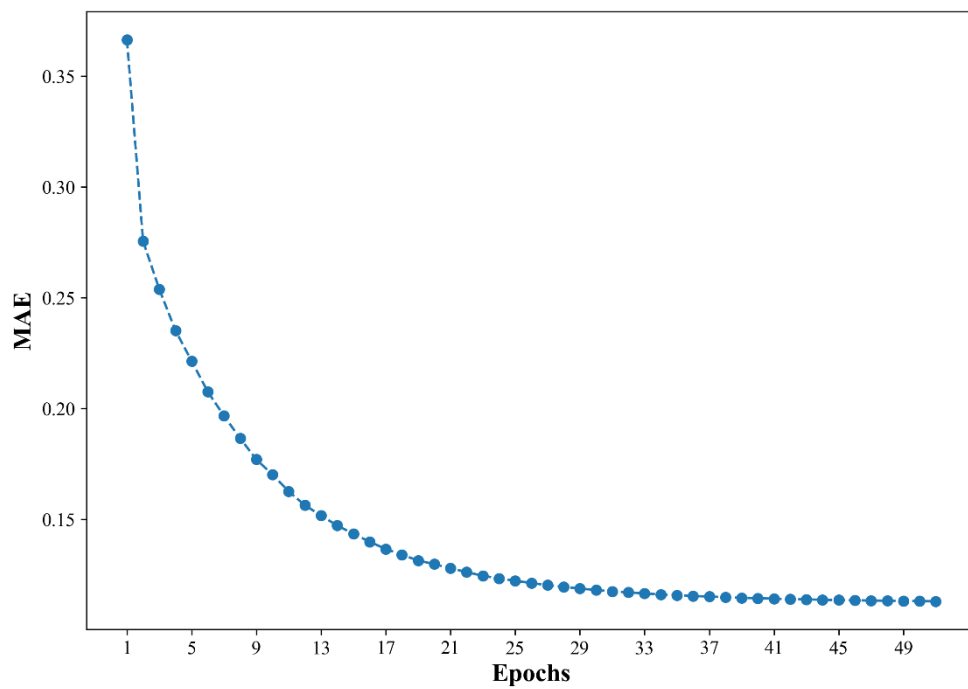
Διάγραμμα Π. 9: MSE ταχυτήτων στον ορίζοντα 20 χρονικών βημάτων για το CNN_LSTM_ed_attention_wf



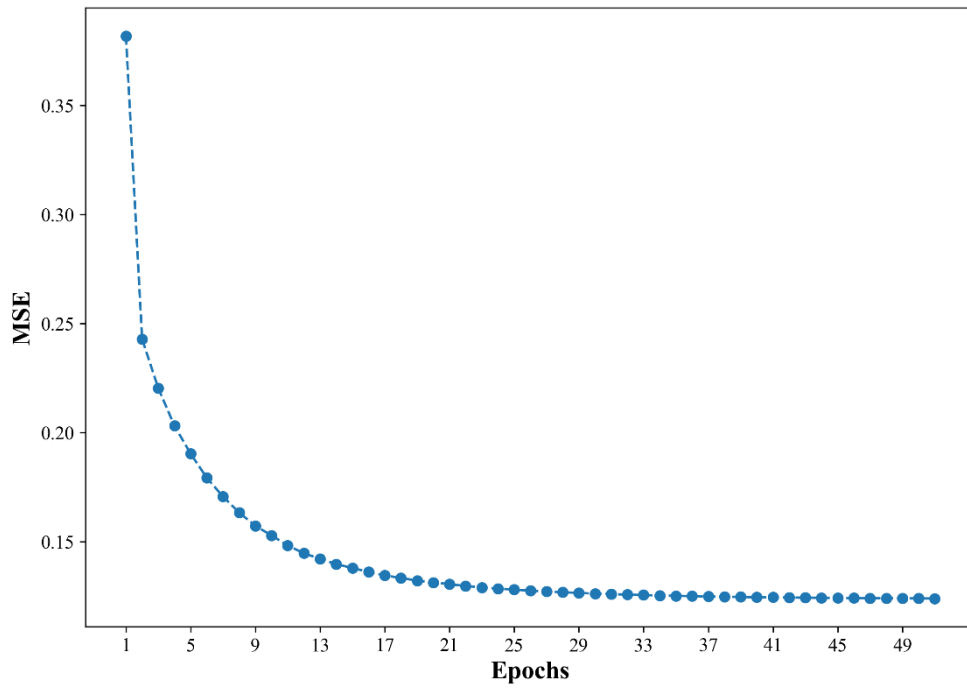
Διάγραμμα Π. 10: MSE θέσεων στον ορίζοντα 20 χρονικών βημάτων για το CNN_LSTM_ed_attention_wf



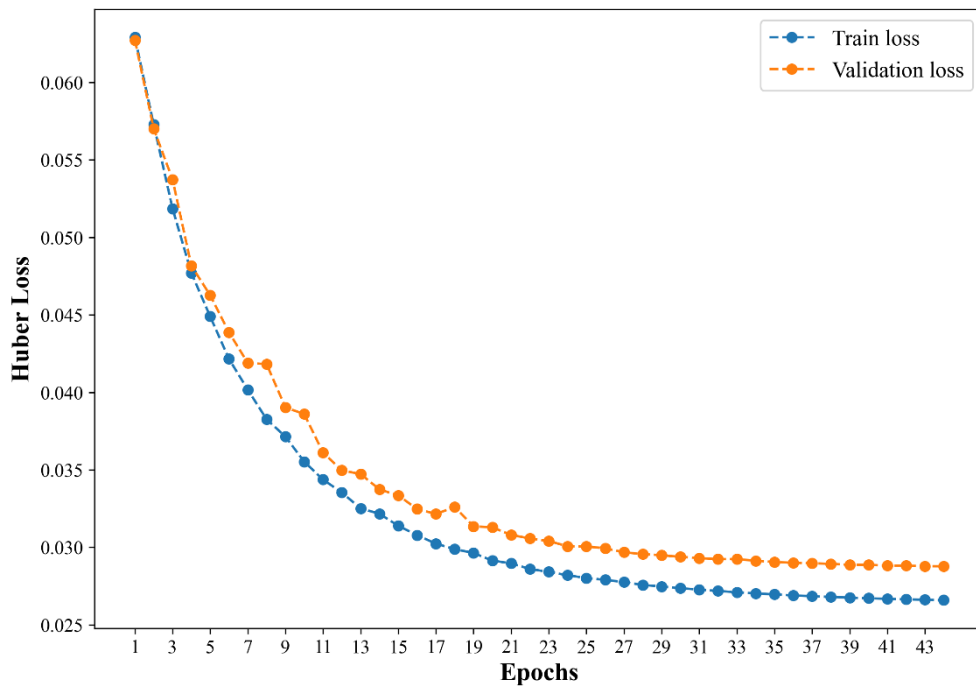
Διάγραμμα Π. 11: Εξέλιξη απώλειας Huber ανά εποχή για το CNN_LSTM_ed_wf



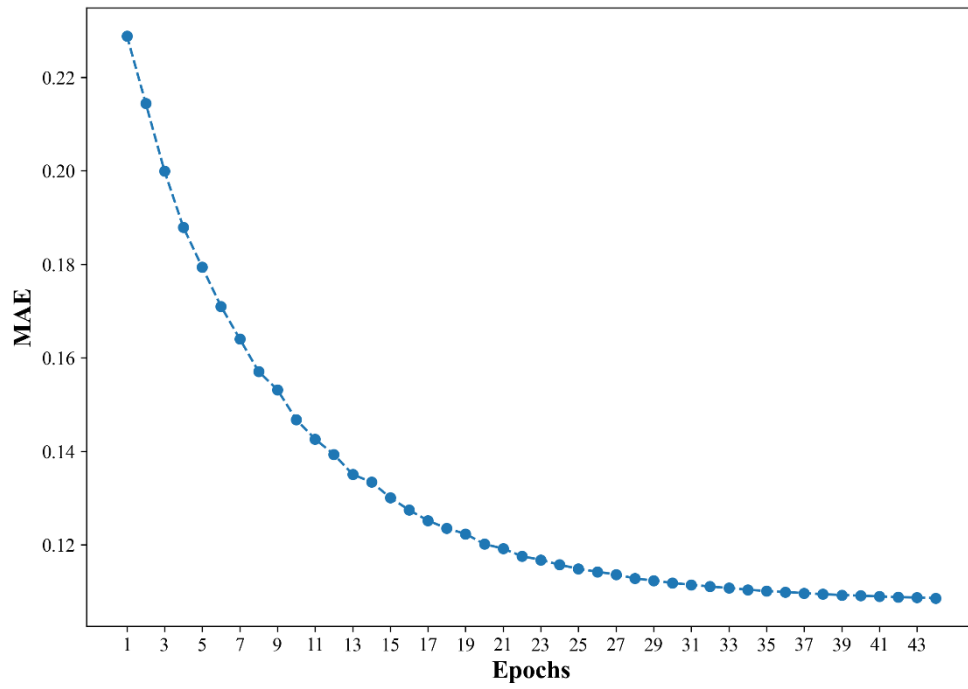
Διάγραμμα Π. 12: Εξέλιξη MSE ανά εποχή για το CNN_LSTM_ed_wf



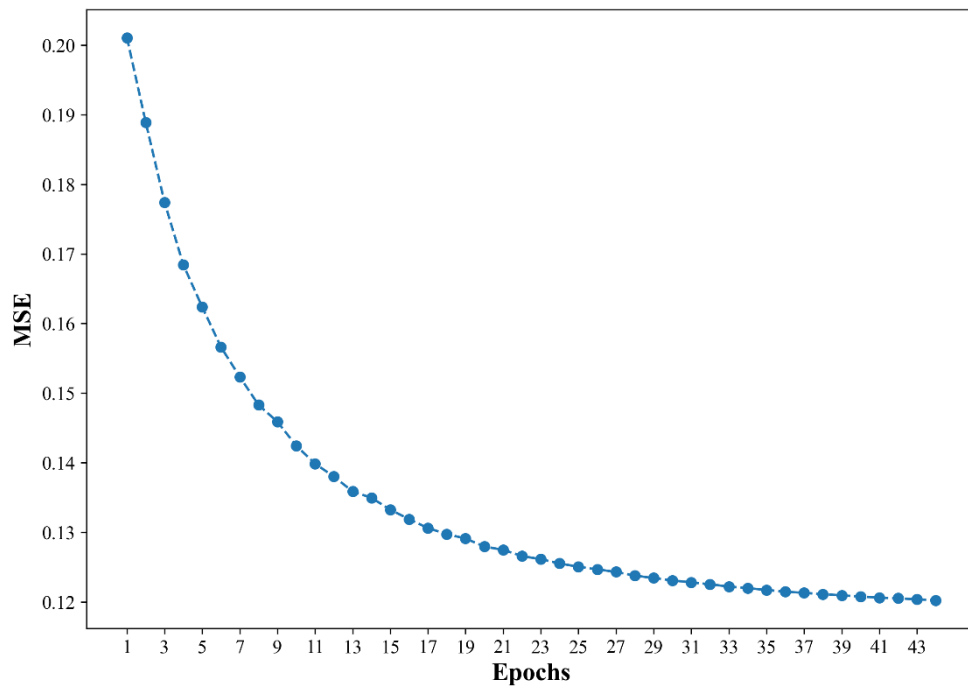
Διάγραμμα Π. 13: Εξέλιξη MSE ανά εποχή για το LSTM_ed_attention_wf



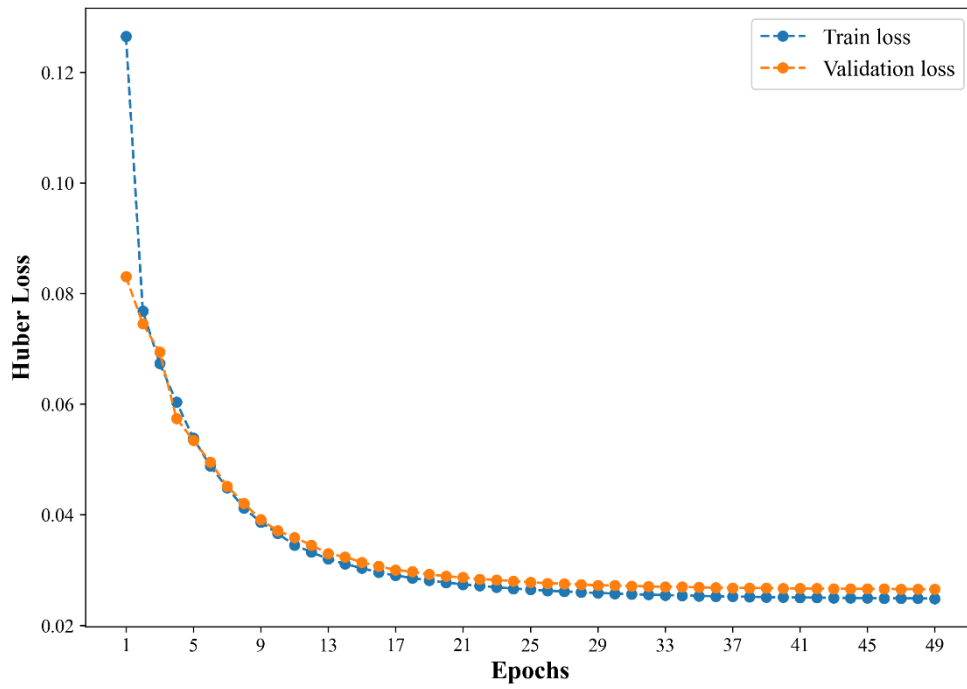
Διάγραμμα Π. 14: Εξέλιξη απώλειας Huber ανά εποχή για το BiLSTM_ed_nf



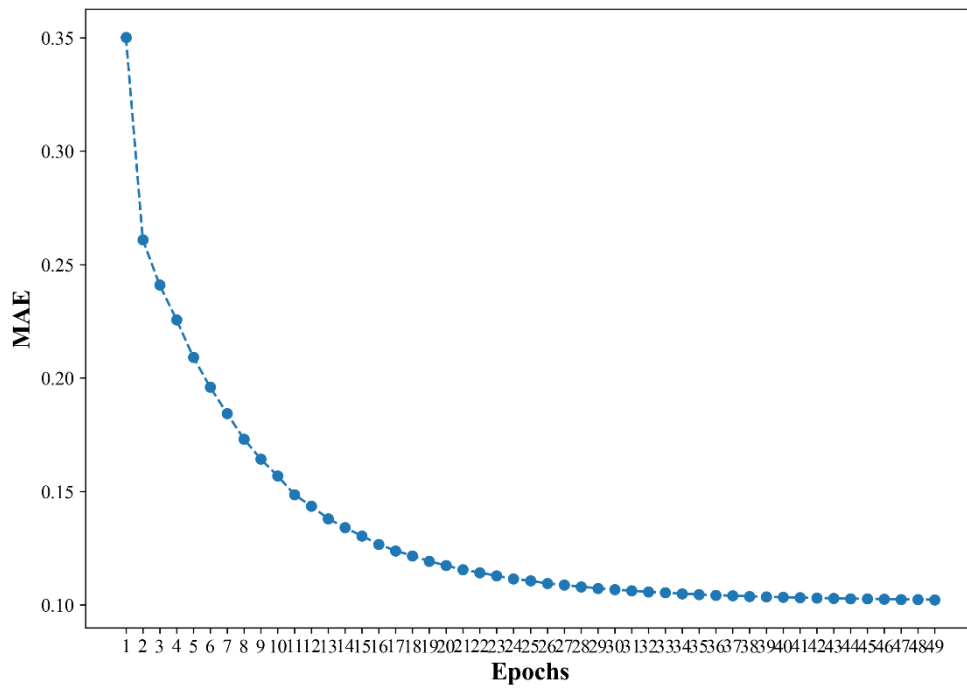
Διάγραμμα Π. 15: Εξέλιξη MSE ανά εποχή για το BiLSTM_ed_nf



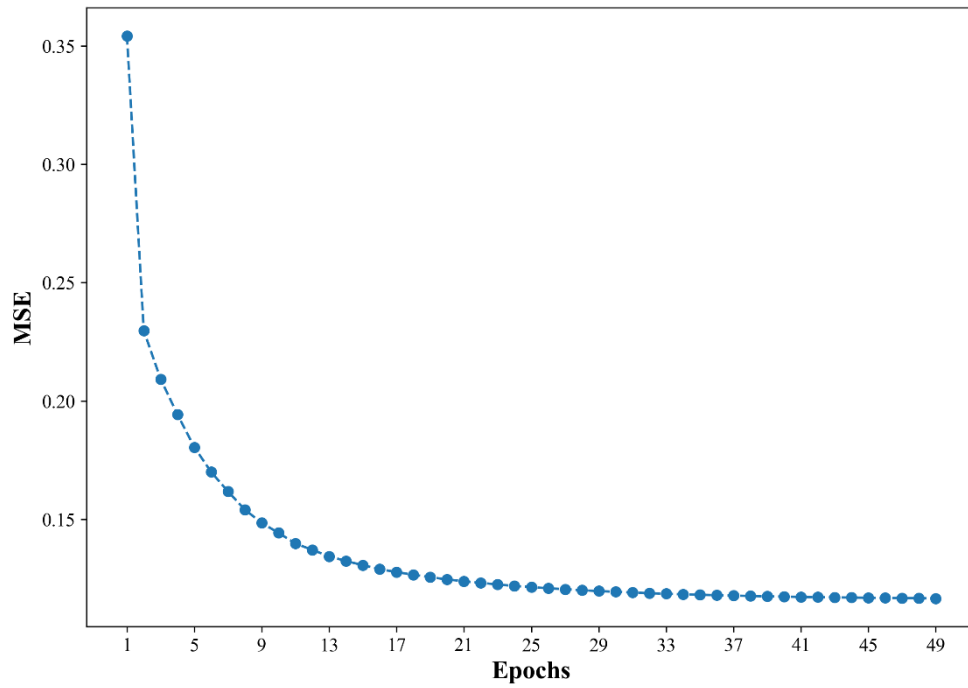
Διάγραμμα Π. 16: Εξέλιξη MSE ανά εποχή για το BiLSTM_ed_nf



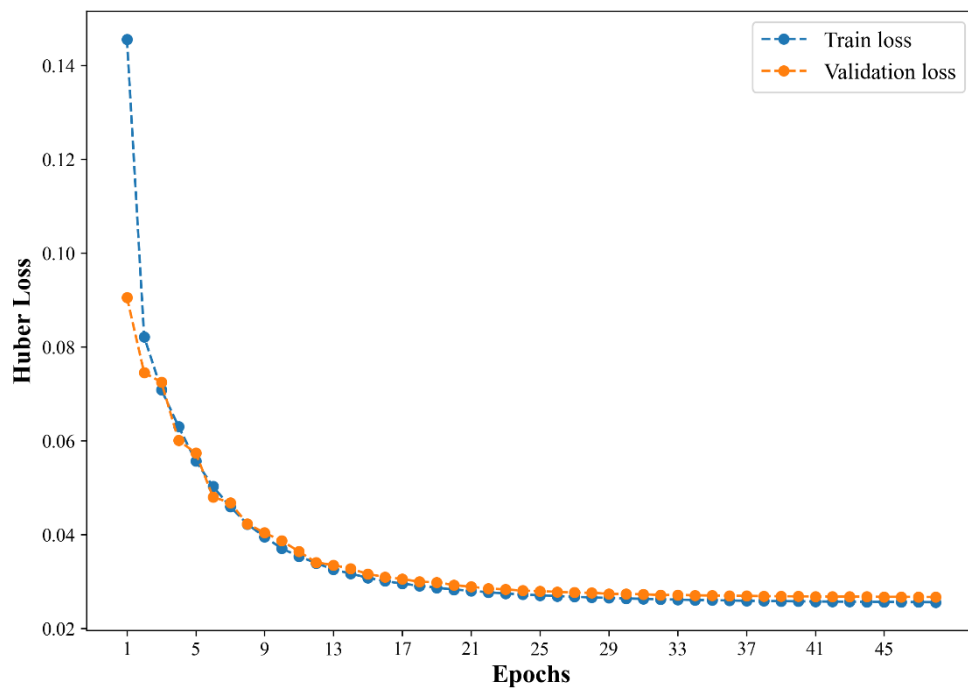
Διάγραμμα Π. 17: Εξέλιξη απώλειας Huber ανά εποχή για το BiLSTM_ed_attention_wf



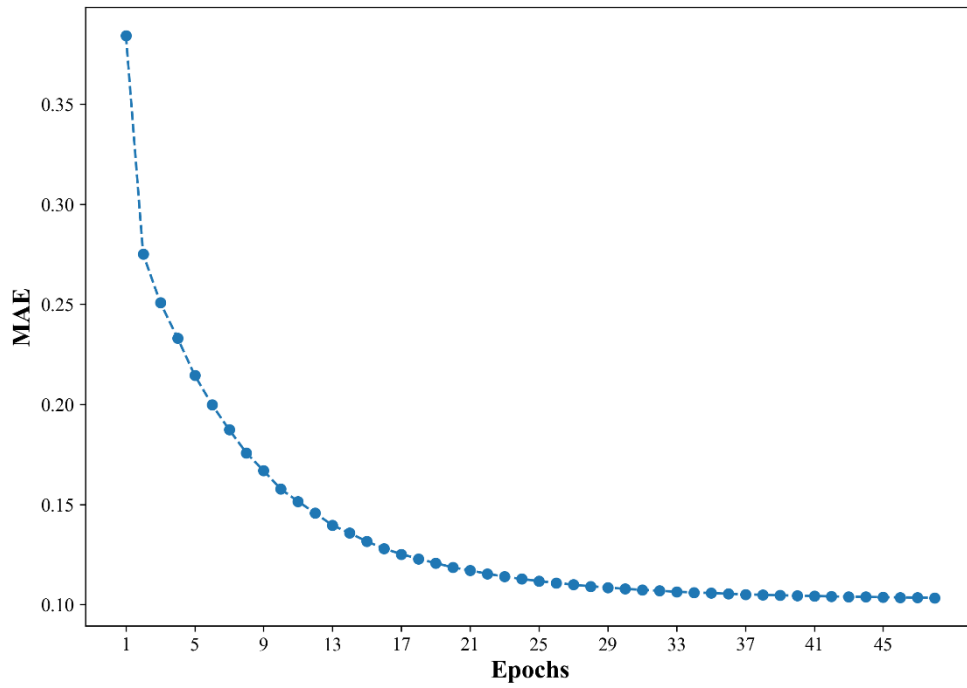
Διάγραμμα Π. 18: Εξέλιξη MSE ανά εποχή για το BiLSTM_ed_attention_wf



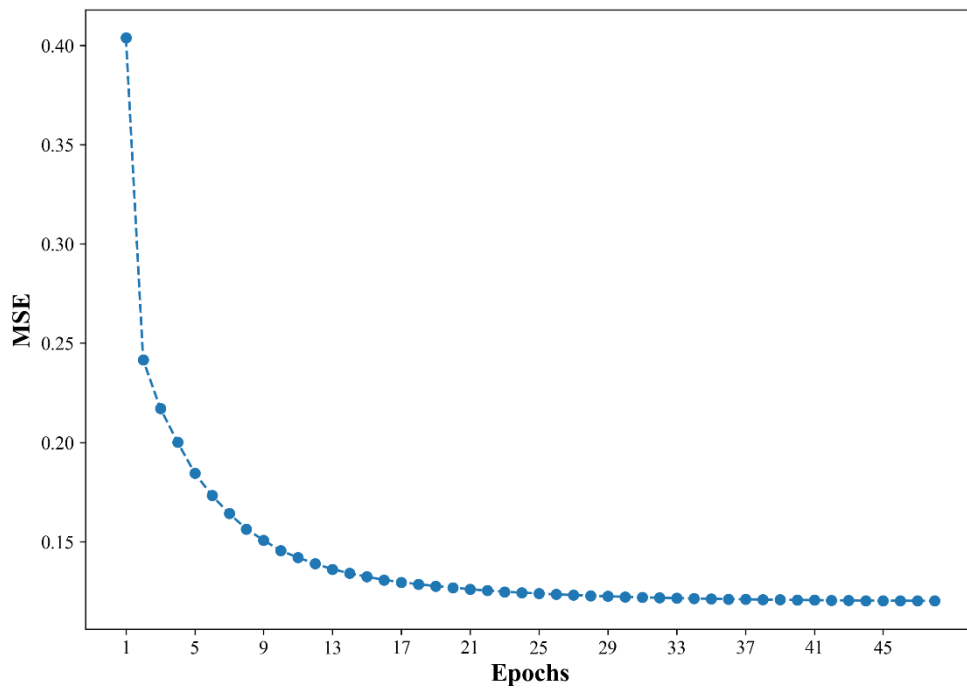
Διάγραμμα Π. 19: Εξέλιξη MSE ανά εποχή για το BiLSTM_ed_attention_wf



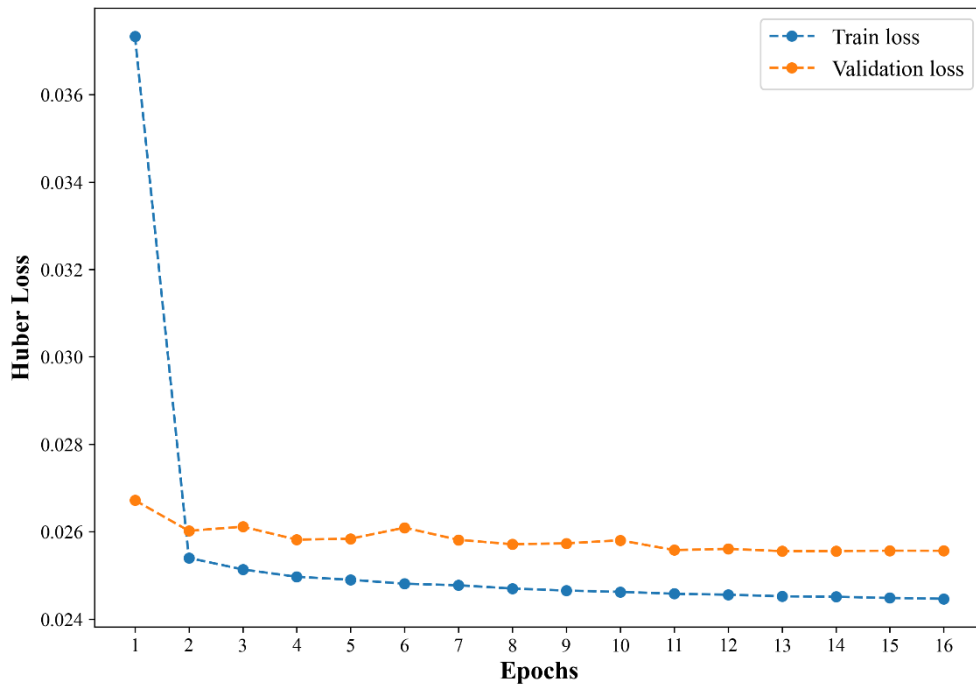
Διάγραμμα Π. 20: Εξέλιξη απώλειας Huber ανά εποχή για το CNN_BiLSTM_ed_wf



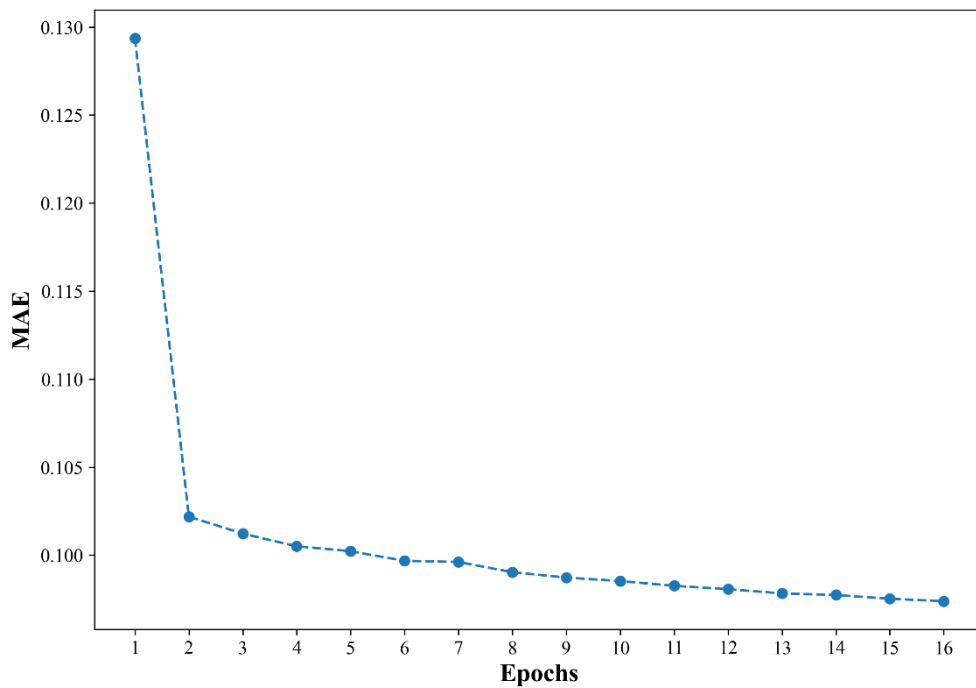
Διάγραμμα Π. 21: Εξέλιξη MSE ανά εποχή για το CNN_BiLSTM_ed_wf



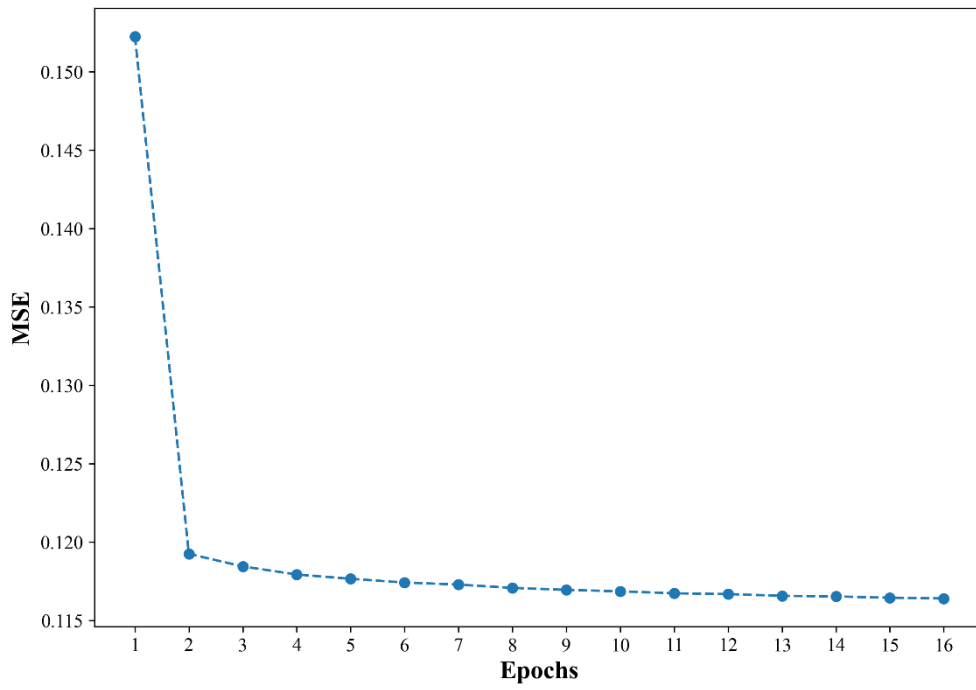
Διάγραμμα Π. 22: Εξέλιξη MSE ανά εποχή για το CNN_BiLSTM_ed_wf



Διάγραμμα Π. 23: Εξέλιξη απώλειας Huber ανά εποχή για το Stacking Ensemble



Διάγραμμα Π. 24: Εξέλιξη MSE ανά εποχή για το Stacking Ensemble



Διάγραμμα Π. 25: Εξέλιξη MSE ανά εποχή για το Stacking Ensemble