ΣΤΡΑΤΙΩΤΙΚΗ ΣΧΟΛΗ ΕΥΕΛΠΙΔΩΝ
Τμήμα Στρατιωτικών Επιστημών

ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

ΔΙΙΔΡΥΜΑΤΙΚΟ ΔΙΑΤΜΗΜΑΤΙΚΟ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ
ΣΠΟΥΔΩΝ
Ευφυή Συστήματα
(Intelligent Systems)

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
Σχολή Μηχανικών Παραγωγής & Διοίκησης

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

Ευφυείς τεχνικές δημιουργίας συνθετικών δεδομένων για εκπαίδευση μοντέλων μηχανικής μάθησης

Του:

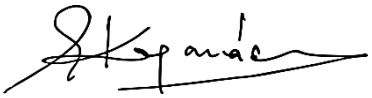
ΑΓΙΟΜΑΥΡΙΤΗ
ΦΩΤΙΟΥ

ΑΜ: 2021018001

Αθήνα 2024

ΤΡΙΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Καθηγήτρια Ειρήνη Καρανάσιου (Επιβλέπουσα),



Καθηγητής Παναγιώτης Τραχανιάς ,

Αν. Καθηγητής Ελευθέριος Δοϊτσίδης

ELEFTHERIOS DOITSIDIS
18.10.2024 19:13

ΠΕΡΙΛΗΨΗ

Η απόκτηση επαρκών δεδομένων για την εκμάθηση αλγορίθμων μηχανικής ή βαθιάς μάθησης είναι πάντα ένα ζητούμενο το οποίο απασχολεί τον ερευνητικό αλλά και τον εταιρικό χώρο, παρόλα αυτά δεν είναι πάντοτε δυνατό να συμβεί. Υπάρχουν τομείς όπως ο στρατιωτικός, ο ιατρικός και άλλοι όπου είναι πιθανό να μην έχουμε την δυνατότητα συλλογής επαρκών δεδομένων, είτε εξαιτίας αδειών χρήσης και προσωπικών δεδομένων είτε λόγω γενικότερης έλλειψης πραγματικών μετρήσεων σε αυτούς τους νευραλγικούς χώρους. Στην παρούσα μεταπτυχιακή διατριβή αναλύονται τεχνικές δημιουργίας συνθετικών δεδομένων που μπορούν να έχουν σημαντικά αποτελέσματα στην εκμάθηση αλγορίθμων όταν δεν διαθέτουμε αρκετά πραγματικά δεδομένα. Ο χειρισμός των τεχνικών διαφοροποιείται όταν θέλουμε να δημιουργήσουμε συνθετικές εικόνες ή συνθετικές συμβολοσειρές αντίστοιχα, με την πρώτη κατηγορία να έχει μεγάλες υπολογιστικές απαιτήσεις και σύνθετες μεθόδους. Για την δημιουργία συνθετικών εικόνων παρουσιάζονται μοντέλα διάχυσης (Diffusion) τα οποία είναι ήδη προ-εκπαιδευμένα (pre-trained) με σκοπό να γίνει μεταφορά μάθησης (transfer learning) πάνω σε συγκεκριμένα δεδομένα που θέλουμε για τα δικά μας αποτελέσματα. Βασικό στοιχείο της εργασίας είναι οι τεχνικές που χρησιμοποιούνται να εμπίπτουν στις αρχές του frugal learning και να μην χρειάζονται μεγάλο όγκο δεδομένων και υπολογιστική ισχύ. Επιπλέον, αναλύονται τεχνικές για την δημιουργία συνθετικών δεδομένων συμβολοσειρών και συγκεκριμένα συμβολοσειρών GPS με σκοπό την εκπαίδευση και τη δοκιμή συστημάτων πρόβλεψης παραπλάνησης (spoofing). Στα πλαίσια της παρούσας εργασίας, έχει δημιουργηθεί πειραματικό σύστημα πρόβλεψης spoofing με σκοπό να αποκτήσουμε μια πιο ρεαλιστική οπτική πάνω στο θέμα και να γίνει testing των συνθετικών δεδομένων που δημιουργήθηκαν.

ABSTRACT

The acquisition of sufficient data for training machine learning or deep learning algorithms is always of interest for both the research and corporate sectors. However, it is not always possible to achieve such a task. There are fields such as the military, medical, and others where it may not be feasible to collect adequate data, either due to usage permissions and personal data restrictions or due to a general lack of real measurements in these critical areas. This master's thesis analyzes techniques for creating synthetic data that can significantly impact the training of algorithms when sufficient real data are not available. The handling of these techniques varies depending on whether the aim is to create synthetic images or synthetic strings, with the former category having high computational requirements and complex methods. For the creation of synthetic images, Diffusion models that are already pre-trained are presented, with the aim of performing transfer learning on specific data we need for our results. A key element of this work is that the techniques used fall under the principles of frugal learning and do not require a large volume of data and computational power. Additionally, techniques for creating synthetic string data, specifically GPS strings, are analyzed for the purpose of training and testing spoofing prediction systems. As part of this work, an experimental spoofing prediction system has been created to gain a more realistic perspective on the subject and to test the synthetic data generated.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω την επιβλέπουσα Καθηγήτρια μου κα. Ειρήνη Καρανάσιου για την στήριξη της κατά την διάρκεια του μεταπτυχιακού προγράμματος, τον Καθηγητή κ. Παναγιώτη Τραχανιά για την άριστη συνεργασία, καθώς και τον επιτηρητή της μεταπτυχιακής διατριβής μου Δρ. Μάρκο Σιγάλα για την άψογη καθοδήγηση του όλων αυτό τον καιρό. Επίσης, θέλω να ευχαριστήσω τους γονείς και την αδερφή μου που πάντα βρίσκονται δίπλα μου, βοηθούν στην εξέλιξη μου και με στηρίζουν. Τέλος, θέλω να ευχαριστήσω τους φίλους μου που έχουν πάντα θετική επίδραση σε μένα με τις συμβουλές τους.

Πίνακας περιεχομένων

Περίληψη 4

Abstract 5

Ευχαριστίες 6

Κατάλογος εικόνων 9

Πίνακας συντομεύσεων όρων 10

ΚΕΦΑΛΑΙΟ 1 11
Εισαγωγή 11
1.1. Related work 11
1.2. Μηχανική μάθηση 11
1.2.1. Επιβλεπόμενη μάθηση 13
1.2.2. Μη επιβλεπόμενη μάθηση 16
1.2.3. Ενισχυτική μάθηση 18
1.2.4. Transfer learning 20
1.3. Βαθιά μάθηση 22
1.3.1. MLP classifier 23
1.3.2. CNN classifier 25
1.3.3. RNN classifier 26
1.4. Frugal learning 28

ΚΕΦΑΛΑΙΟ 2 29
Δημιουργία συνθετικών δεδομένων 29
2.1. Εισαγωγή 29
2.2. Συνθετικά δεδομένα εικόνων 30
2.2.1. Diffusion models 30
2.2.1.1. Stable diffusion model 32
2.2.2. GANs 33
2.2.3. VAEs 34
2.3. Συνθετικά δεδομένα συμβολοσειρών 35
2.3.1. LLMs 36
2.3.2. Random walk 37

ΚΕΦΑΛΑΙΟ 3 40

Υλοποίηση 40
3.1. Δημιουργία συνθετικών εικόνων με stable diffusion 40
3.2. Τροποποίηση εικόνων με stable diffusion 42
3.3. GPS Spoofing detection για UAVs 47
3.3.1. Dataset 48
3.3.2. Επιλογή & εκμάθηση μοντέλου 49
3.3.3. Δημιουργία συνθετικών συμβολοσειρών με random walk 50
3.3.4. Εξομοίωση & Testing 51

ΚΕΦΑΛΑΙΟ 4 53
Συμπεράσματα - Μελλοντικές Εφαρμογές 53
4.1. Συμπεράσματα 53
4.2. Συνεισφορά της διατριβής 54
4.3. Μελλοντικές εφαρμογές 55

Βιβλιογραφία 56
Ιστότοποι 58

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1.1 : Διαδικασία εκμάθησης με επίβλεψη

Εικόνα 1.2 : Σχηματική αναπαράσταση δέντρου απόφασης

Εικόνα 1.3 : LDA input/output workflow

Εικόνα 1.4 : Ενισχυτική μάθηση

Εικόνα 1.5 : Υποδιαιρέσεις ενισχυτικής μάθησης

Εικόνα 1.6 : Απεικόνιση του Transfer Learning

Εικόνα 1.7 : Σχηματική Απεικόνιση ενός MLP classifier

Εικόνα 1.8 : Τυπική αρχιτεκτονική ενός CNN

Εικόνα 1.9 : Σχηματική Απεικόνιση ενός RNN classifier

Εικόνα 2.1 : Διαδικασία αποθορυβοποίησης ενός diffusion model

Εικόνα 2.2 : Γενική αρχιτεκτονική ενός GAN

Εικόνα 2.3 : Παράδειγμα δομής ενός Variational Autoencoder

Εικόνα 2.4 : Πρόβλεψη για το μέγεθος του market των LLMs μέχρι το 2030

Εικόνα 2.5 : Απεικόνιση τριών random walk σε τρεις διαστάσεις

Εικόνα 3.1 : $1^{η}$ εικόνα αποτελεσμάτων

Εικόνα 3.2 : $2^{η}$ εικόνα αποτελεσμάτων

Εικόνα 3.3 : $3^{η}$ εικόνα αποτελεσμάτων

Εικόνα 3.4 : Παράδειγμα εικόνας σε διαφορετικές εποχές που χρησιμοποιήθηκε στο transfer learning

Εικόνα 3.5 : Εικόνα πριν την τροποποίηση

Εικόνα 3.6 : Εικόνα μετά την τροποποίηση

Εικόνα 3.7 : Εικόνα μετά την τροποποίηση (2)

Εικόνα 3.8 : Σχέδιο του μοντέλου συλλογής σημάτων GPS

Εικόνα 3.9 : Τιμές accuracy των αλγορίθμων

Εικόνα 3.10 : Παράδειγμα συνθετικών δεδομένων συμβολοσειρών GPS (κανονικών και spoofed)

Εικόνα 3.11 : Στιγμιότυπο από την εξομοίωση στο Webots

ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΕΥΣΕΩΝ ΟΡΩΝ

Συντομεύσεις	Απόδοση Όρου
TN / AI	Τεχνητή Νοημοσύνη / Artificial Intelligence
EM	Ενισχυτική Μάθηση
ML	Machine Learning
TL	Transfer Learning
NLP	Natural Language Processing / Επεξεργασία Φυσικής Γλώσσας
MLP	Multilayer perceptron
CNN	Convolutional neural network
RNN	Recurrent neural network
LLMs	Large language models
GANs	Generative adversarial networks
VAEs	Variational autoencoders
UAVs	Unmanned Aerial Vehicles
SSIM	Structural Similarity Index
NDVI	Normalized Difference Vegetation Index
CLIP	Contrastive Language-Image Pretraining

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

1.1 Related work

Υπάρχουν πολλά συστήματα που έχουν σχεδιαστεί για να αντιμετωπίσουν το spoofing στα drones. Ένα από αυτά είναι τα συστήματα αναχαίτισης των μη επανδρωμένων αεροσκαφών (C-UAV), τα οποία χρησιμοποιούν διάφορες μεθόδους για να εντοπίσουν και να αναχαιτίσουν τα UAV. Έχουν διεξαχθεί ερευνητικές μελέτες για τις απαιτήσεις της επίθεσης spoofing σε πολλαπλά drones, τα οποία χρησιμοποιούν σήματα GPS και έχουν τη δυνατότητα να επικοινωνούν μεταξύ τους. Πλέον, εξετάζονται τεχνικές για την εύρεση και την αντιμετώπιση του spoofing σε drones και με τη χρήση της τεχνητής νοημοσύνης κάτι το οποίο μέχρι στιγμής βρίσκεται μόνο σε ερευνητικό επίπεδο.

Σχετικά με τα συνθετικά δεδομένα, έχουν γίνει διάφορες υλοποιήσεις για την παραγωγή συνθετικών δεδομένων με την εφαρμογή της τεχνητής νοημοσύνης. Τα συνθετικά δεδομένα μπορεί να περιλαμβάνουν κείμενο, πίνακες και εικόνες. Η χρήση συνθετικών δεδομένων αξίζει να σημειωθεί ότι γίνεται ολοένα και πιο δημοφιλής και ευρέως αποδεκτή, καθώς μπορεί να προσφέρει πολλά πλεονεκτήματα σε σχέση με τα πραγματικά δεδομένα. Η Gartner προέβλεψε ότι, μέχρι το 2030, τα περισσότερα δεδομένα που χρησιμοποιούνται για AI θα παράγονται τεχνητά μέσω κανόνων, στατιστικών μοντέλων, προσομοιώσεων ή άλλων τεχνικών. Μερικά παραδείγματα τέτοιων υλοποιήσεων είναι το MDClone για ιατρικά συνθετικά δεδομένα, το Mostly AI για συνθετικά δεδομένα γενικής φύσεως καθώς και το DALL·E που είναι εξειδικευμένο στην δημιουργία συνθετικών εικόνων.

1.2 Μηχανική μάθηση

Η Μάθηση είναι μια κεντρική ιδιότητα της ανθρώπινης νοητικής συμπεριφοράς. Παρόλο που έχουν γίνει πολλές έρευνες και μελέτες από επιστήμονες στον τομέα της Γνωστικής Ψυχολογίας και από φιλοσόφους, η κατανόηση της μάθησης δεν είναι ακόμη πλήρης. Με ποιον τρόπο τότε θα μπορούσαν οι επιστήμονες στον τομέα της Τεχνητής Νοημοσύνης (ΤΝ) να κατασκευάσουν υπολογιστικά συστήματα που μπορούν να μάθουν, δηλαδή, να επιτύχουν τη λεγόμενη Μηχανική Μάθηση (Machine Learning).

Αυτή μπορεί να οριστεί ως το φαινόμενο κατά το οποίο ένα σύστημα βελτιώνει την απόδοσή του κατά την εκτέλεση μιας συγκεκριμένης εργασίας, χωρίς να χρειάζεται να προγραμματιστεί εκ νέου. Βάσει αυτού του ορισμού, η Μηχανική Μάθηση στοχεύει στη δημιουργία μηχανών που μπορούν να μάθουν, δηλαδή, να βελτιώνουν την απόδοσή τους σε κάποιους τομείς μέσω της αξιοποίησης προηγούμενης γνώσης και εμπειρίας. Ένας γενικός ορισμός της Μηχανικής Μάθησης παρέχεται από το [1] :«Ένα πρόγραμμα υπολογιστή λέμε ότι μαθαίνει από την εμπειρία Ε ως προς κάποια κλάση εργασιών Τ και μέτρο απόδοσης Ρ, αν η απόδοσή του σε εργασίες από το Τ, όπως μετριέται από το Ρ, βελτιώνεται μέσω της εμπειρίας Ε.»

Στην Επαγωγική Μάθηση, με τη διαδικασία της επαγωγής, ο άνθρωπος μαθαίνει κατανοώντας το περιβάλλον του μέσω παρατηρήσεων και δημιουργεί μια απλοποιημένη (αφαιρετική) εκδοχή του που ονομάζεται νοητικό μοντέλο. Επιπλέον, ο άνθρωπος έχει τη δυνατότητα να οργανώνει και να συσχετίζει τις εμπειρίες και τις παρατηρήσεις του δημιουργώντας νέες δομές που ονομάζονται νοητικά πρότυπα, με αξιοποίηση και του επαγωγικού και του απαγωγικού συλλογισμού.

Στη δημιουργία νέων προτύπων από παλαιά βασίζονται οι τρόποι μάθησης που εξαρτώνται σε μεγαλύτερο ή μικρότερο βαθμό από την προϋπάρχουσα γνώση για ένα πρόβλημα, όπως είναι η μάθηση από επεξηγήσεις και η μάθηση από περιπτώσεις. Σε σχέση με την ανθρώπινη ικανότητα προς μάθηση, οι φιλόσοφοι θέτουν το ερώτημα: «Πώς μπορεί ένας επαγωγικός συλλογισμός που οδηγεί στη μάθηση να αξιολογηθεί ως προς την ορθότητά του;». Αντίστοιχα, οι ψυχολόγοι ρωτούν: «Πώς αποθηκεύει ο εγκέφαλος τα αποτελέσματα της διαδικασίας της μάθησης, δηλαδή τα νοητικά μοντέλα και τα πρότυπα;». Στο χώρο της Τεχνητής Νοημοσύνης απλώς ρωτούν: «Πώς μπορεί μία μηχανή να δημιουργήσει νέα μοντέλα και πρότυπα μάθησης από συγκεκριμένα παραδείγματα και πόσο αξιόπιστα είναι αυτά τα μοντέλα και πρότυπα στην πράξη;».

Βάσει των προηγούμενων, μπορούμε να προσφέρουμε έναν εναλλακτικό ορισμό για τη Μηχανική Μάθηση: Η Μηχανική Μάθηση αναφέρεται στη δυνατότητα ενός υπολογιστικού συστήματος να κατασκευάζει μοντέλα ή πρότυπα από ένα σύνολο δεδομένων. Ως επιστημονικός κλάδος της Τεχνητής Νοημοσύνης, η Μηχανική Μάθηση ασχολείται με την εξέταση αλγορίθμων που βελτιώνουν την απόδοσή τους σε μια δεδομένη εργασία, χρησιμοποιώντας την εμπειρία τους. Σχετικά με τη σχεδίαση των συστημάτων Μηχανικής Μάθησης, για τα συστήματα που ανήκουν στη συμβολική ΤΝ, η ικανότητα μάθησης καθορίζεται ως η ικανότητα απόκτησης πρόσθετης γνώσης, που προκαλεί αλλαγές στην υπάρχουσα καταχωρημένη γνώση, είτε αλλάζοντας τα χαρακτηριστικά της είτε με αυξομείωσή της. Στην περίπτωση των συστημάτων ΤΝ που ανήκουν στη Μη Συμβολική ΤΝ (όπως η περίπτωση των Τεχνητών Νευρωνικών Δικτύων), η μάθηση καθορίζεται ως η δυνατότητα που έχουν τα συστήματα να μετασχηματίζουν την εσωτερική τους δομή, παρά να τροποποιούν κατάλληλα τη γνώση που έχει καταχωρηθεί μέσα σε αυτά κατά το σχεδιασμό τους.

Παρόλο που είμαστε ακόμη μακριά από την κατασκευή μηχανών που μπορούν να μάθουν τόσο αποτελεσματικά όσο ο άνθρωπος, για συγκεκριμένες περιοχές μάθησης έχουν αναπτυχθεί αλγόριθμοι που έχουν επιτρέψει την εμφάνιση σύγχρονων εμπορικών εφαρμογών με σημαντική επιτυχία. Επιπλέον, τα αποτελέσματα από τις εφαρμογές της Τεχνητής Νοημοσύνης έχουν ήδη αρχίσει να είναι ορατά και να παρέχουν απαντήσεις σε αναπάντητα, μέχρι τώρα, ερωτήματα των άλλων επιστημονικών κλάδων που εξετάζουν την ικανότητα του ανθρώπου να μαθαίνει. Ο τομέας της Μηχανικής Μάθησης αναπτύσσει

επιτυχώς την Εξελικτική Μάθηση, η οποία αντιγράφει διαδικασίες φυσικής αναπαραγωγής σε ζωντανά όντα. Χρησιμοποιείται κυρίως σε προβλήματα βελτιστοποίησης.

Πέρα από την Τεχνητή Νοημοσύνη (ΤΝ) καθαυτή, υπάρχουν πολλοί επιστημονικοί κλάδοι που επωφελούνται από τις επιτυχίες στον τομέα της Μηχανικής Μάθησης. Αυτοί περιλαμβάνουν την Εξόρυξη Δεδομένων, τις Πιθανότητες και τη Στατιστική, τη Θεωρία της Πληροφορίας, την Αριθμητική Βελτιστοποίηση, τη Θεωρία της Πολυπλοκότητας, τη Θεωρία Ελέγχου (προσαρμοστική), την Ψυχολογία (εξελικτική, γνωστική), τη Νευροβιολογία και τη Γλωσσολογία. Σχετικά με την επεξεργασία φυσικής γλώσσας, υπάρχουν πολλοί τομείς που μπορούν να επωφεληθούν από τη χρήση της, με τον πιο σημαντικό να είναι, πρώτα απ' όλα, η επικοινωνία ανθρώπου-μηχανής. Σε αυτόν τον τομέα, η χρήση φυσικής γλώσσας επιτρέπει στους χρήστες να χρησιμοποιούν τον φυσικό τρόπο επικοινωνίας τους αντί για τεχνητές γλώσσες (προγραμματισμού, μηχανής, κ.ά.) ή δομημένα μενού. Αυτή η προσέγγιση έχει τόσο πλεονεκτήματα όσο και αδυναμίες. Ενώ δεν απαιτείται εκπαίδευση για τη χρήση της γλώσσας, αυτό διευκολύνει περισσότερο τους περιστασιακούς χρήστες και λιγότερο τους εξειδικευμένους, όπως για παράδειγμα οι προγραμματιστές ή οι υπάλληλοι γραφείου που εισάγουν δεδομένα σε φόρμες.

Μια περιοχή που βλέπουμε αισθητά την επίδραση της NLP είναι η διαχείριση πληροφοριών, όπου η επεξεργασία φυσικής γλώσσας θα μπορούσε να ενεργοποιήσει διαδικασίες αυτόματης διαχείρισης και επεξεργασίας της πληροφορίας με βάση την ερμηνεία της. Για παράδειγμα, εάν ένα σύστημα μπορούσε να κατανοήσει το νόημα ενός εγγράφου, θα μπορούσε να το αρχειοθετήσει μαζί με άλλα σχετικά έγγραφα.

Μια ακόμη περιοχή στην οποία διακρίνεται η επιδράση της NLP είναι η αναζήτηση σε βάσεις δεδομένων. Οι συνηθισμένοι τρόποι για να εκφράσει κανείς μια επιθυμητή πληροφορία είναι μέσω της επιλογής από λίστες, της συμπλήρωσης μενού ή της σύνταξης του αιτήματος σε τεχνητή γλώσσα (όπως τη γλώσσα ερωτημάτων SQL). Η χρήση τεχνητής γλώσσας επιτρέπει την ανάπτυξη απλών μηχανισμών αναζήτησης, αλλά ο χρήστης πρέπει να έχει κάποια γνώση σχετικά με τη δομή της βάσης. Από την άλλη πλευρά, ο χρήστης είναι πιο εξοικειωμένος με το περιεχόμενο ή την περιοχή ενδιαφέροντος της βάσης παρά με τη δομή της. Με τη χρήση φυσικής γλώσσας, τα αιτήματα μπορούν να περιοριστούν σε όρους που σχετίζονται με το περιεχόμενο και την περιοχή ενδιαφέροντος.

1.2.1 Επιβλεπόμενη μάθηση

Η εκμάθηση με επίβλεψη είναι μια διαδικασία των μηχανών εκμάθησης, κατά την οποία κάθε δείγμα αποτελείται από ένα ζεύγος που περιλαμβάνει την επεξηγηματική μεταβλητή x και μια "ετικέτα" (label), την τιμή απόκρισης, την μεταβλητή y. Ο αλγόριθμος εκμάθησης με επίβλεψη αναλύει ένα γνωστό σύνολο δεδομένων, το οποίο αποκαλείται σύνολο εκπαίδευσης, και παράγει μια συνάρτηση f, η οποία μπορεί να χρησιμοποιηθεί για να χαρτογραφηθούν (ταξινομηθούν) νέα άγνωστα σύνολα δεδομένων. Με άλλα λόγια, η f καθορίζει την κατηγορία στην οποία θα αντιστοιχιστούν τα νέα άγνωστα δεδομένα. Για να γίνει πιο κατανοητή η διαδικασία, παρατίθενται κάποιοι ορισμοί.

Το πρόβλημα της ταξινόμησης στον τομέα της μηχανικής μάθησης αναφέρεται στην κατάσταση όπου απαιτείται ο αυτοματοποιημένος καθορισμός της κατηγορίας (κλάσης) στην οποία ανήκει μια νέα παρατήρηση, για παράδειγμα, εάν ένα e-mail είναι spam ή όχι.

Το σύνολο εκπαίδευσης αποτελείται από γνωστά ζευγάρια δεδομένων (x_{ij}, y_i), όπου τα x_{ij}=(x_{i1} , x_{i2} ,…x_{in}) είναι οι επεξηγηματικές μεταβλητές (τα χαρακτηριστικά των δεδομένων) και τα y_i είναι οι κατηγορίες στις οποίες ανήκουν τα αντίστοιχα x_{ij} (η κατηγορία που αντιστοιχεί στο κάθε x_{ij}). Το σύνολο εκπαίδευσης παρέχεται ως είσοδος στον αλγόριθμο εκμάθησης, ώστε ο αλγόριθμος να εκπαιδευτεί πάνω σε αυτό και να παράγει μοτίβα ή κανόνες αντιστοίχισης των x_{ij} με τα y_i. Το σύνολο επικύρωσης είναι ένα σύνολο που αποτελείται επίσης από γνωστά ζευγάρια δεδομένων (x_{ij}, y_i). Αυτό το σύνολο χρησιμοποιείται μετά την εκπαίδευση του αλγορίθμου για να ελεγχθεί η απόδοσή του.

Τα βήματα που ακολουθούνται κατά την εκμάθηση με επίβλεψη είναι τα εξής:

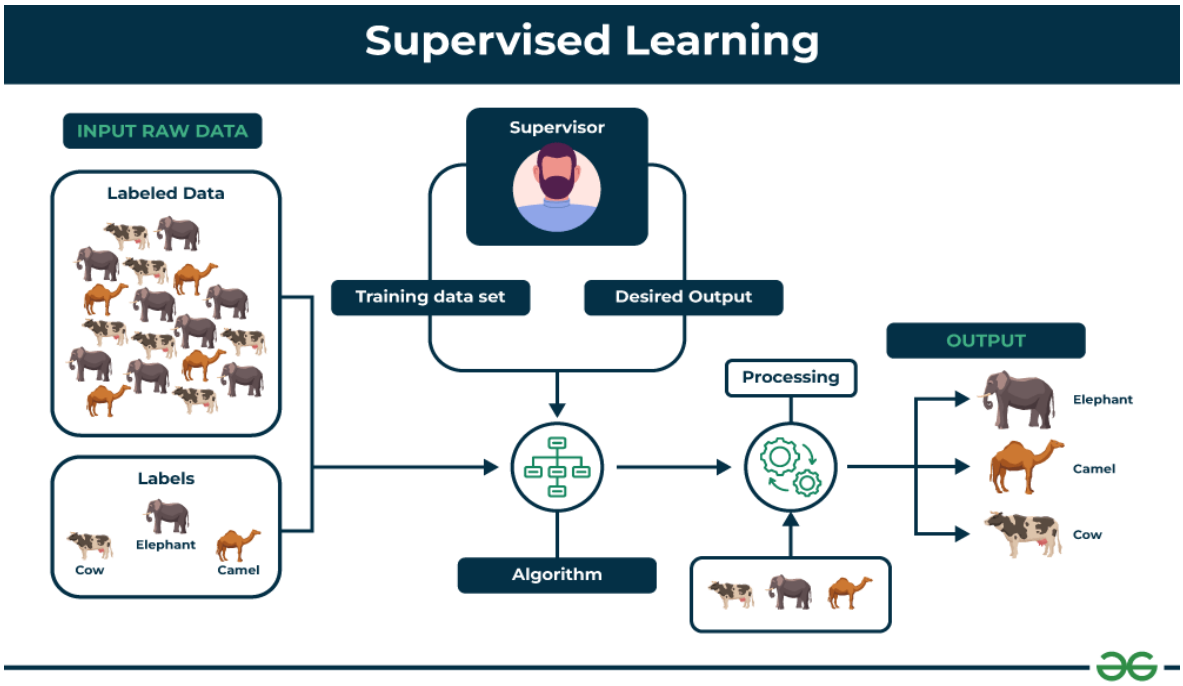
1. Προσδιορισμός του τύπου των δειγμάτων που είναι διαθέσιμα.
2. Διαχωρισμός του συνόλου εκπαίδευσης και του συνόλου επικύρωσης.
3. Επιλογή του κύριου αλγορίθμου μηχανικής μάθησης (π.χ., δέντρα αποφάσεων).
4. Εκπαίδευση του αλγορίθμου με βάση το σύνολο εκπαίδευσης.
5. Έλεγχος της ακρίβειας του μοντέλου με βάση το σύνολο επικύρωσης.

Με αυτόν τον τρόπο, ο αλγόριθμος εκμάθησης με επίβλεψη λειτουργεί ως εξής:

- Τροφοδότηση του αλγορίθμου με ένα σύνολο εκπαίδευσης S που περιέχει m ζευγάρια παραδειγμάτων: S={(x_{1j} , y_1),(x_{2j} , y_2),……,(x_{mj} , y_m)}.
- Εύρεση μιας προσεγγιστικής συνάρτησης h:X →Y από τον αλγόριθμο, όπου X είναι ο χώρος εισόδου και Y ο χώρος εξόδου. Η h καλείται συνάρτηση πρόβλεψης.

Οι μέθοδοι για εκμάθηση με επίβλεψη μπορούν να αντιμετωπίσουν προβλήματα ταξινόμησης και παλινδρόμησης και συνήθως αυτές οι μέθοδοι είναι γρήγορες και ακριβείς.

Παραδείγματα αλγορίθμων εκμάθησης με επίβλεψη περιλαμβάνουν τα τυχαία δάση (random forest), τα δέντρα αποφάσεων (προαιρετικά με τη χρήση τεχνικών bagging ή boosting, τη μέθοδο των k πλησιέστερων γειτόνων (k nearest neighbor - kNN), , τη λογιστική παλινδρόμηση, τη γραμμική παλινδρόμηση και τις “Μηχανές Διανυσμάτων Υποστήριξης” (Support Vector Machines – SVM).



Εικόνα 1.1 : Διαδικασία εκμάθησης με επίβλεψη [33]

Δέντρα Απόφασης

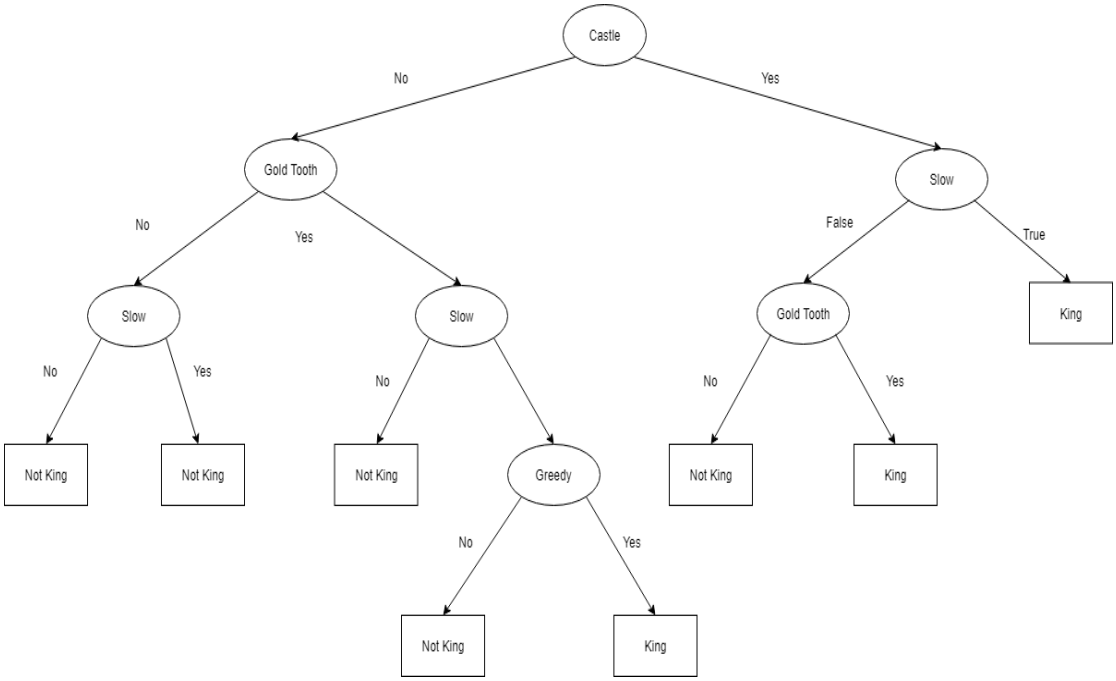
Τα δέντρα αποφάσεων είναι συναρτήσεις ταξινόμησης που αναπαρίστανται ως δέντρα και αποτελούν μία από τις πιο βασικές μεθόδους κατηγοριοποίησης και πρόβλεψης. Τα δέντρα αποφάσεων έχουν τις εξής ιδιότητες:

- Οι κόμβοι του δέντρου αναπαριστούν τα χαρακτηριστικά και ο κόμβος που βρίσκεται στο υψηλότερο επίπεδο ονομάζεται ρίζα του δέντρου.
- Οι κόμβοι του δέντρου συνδέονται μεταξύ τους με κλαδιά, τα οποία αναπαριστούν τις δυνατές συνδέσεις του κόμβου “πατέρα” με τους κόμβους “παιδιά”.
- Οι κόμβοι του δέντρου από τους οποίους δεν ξεκινά κάποιο κλαδί λέγονται φύλλα και παίρνουν το όνομα κάποιας κλάσης.

Μια νέα παρατήρηση κατηγοριοποιείται ξεκινώντας από τη ρίζα του δέντρου, η οποία αναπαριστά ένα από τα χαρακτηριστικά. Αφού ελέγξουμε ποια τιμή έχει το αντίστοιχο χαρακτηριστικό της νέας παρατήρησης, προχωράμε στο κατάλληλο κλαδί. Συνεχίζουμε επαναληπτικά την ίδια διαδικασία για κάθε επόμενο κόμβο στον οποίο καταλήγει το κάθε κλαδί που ακολουθήσαμε. Η παραπάνω διαδικασία σταματάει όταν φτάσουμε σε κάποιο φύλλο, το οποίο μας υποδεικνύει και την κατηγορία του παραδείγματος.

Για την κατασκευή ενός δέντρου αποφάσεων ακολουθούνται τα παρακάτω βήματα:

1. Δίνεται ένα σύνολο εκπαίδευσης, ξεκινάμε με έναν κόμβο που περιέχει όλα τα παραδείγματα του συνόλου εκπαίδευσης.
2. Στη συνέχεια, διασπάμε τον κόμβο βάσει μιας συνθήκης σε ένα από τα χαρακτηριστικά.
3. Το βήμα (2) εκτελείται επαναληπτικά για κάθε κόμβο, μέχρι οι εγγραφές ενός τελικού κόμβου να ανήκουν σε μία μόνο κατηγορία.
4. Αφού κατασκευαστεί το δέντρο, μπορούν αν χρειαστεί να χρησιμοποιηθούν κάποιες τεχνικές βελτιστοποίησης.



Εικόνα 1.2 : Σχηματική αναπαράσταση δέντρου απόφασης [31]

1.2.2 Μη επιβλεπόμενη μάθηση

Στη μηχανική μάθηση, η εκμάθηση χωρίς επίβλεψη αναφέρεται σε προβλήματα όπου τα δεδομένα δεν είναι κατηγοριοποιημένα, δηλαδή δεν έχουν "ετικέτα" ή τιμή απόκρισης Y (unlabeled data). Έτσι, η εκμάθηση χωρίς επίβλεψη προσπαθεί να βρει κρυφά μοτίβα ή τυχαίες συσχετίσεις μεταξύ των δεδομένων. Αυτή η περίπτωση δεν είναι καθόλου σπάνια και τότε δίνεται σαν είσοδος στον αλγόριθμο μόνο οι μεταβλητές $\{x_{1j}, x_{2j}, \ldots, x_{nj}\}$. Οι είσοδοι αυτοί για παράδειγμα μπορεί να αντιστοιχούν στα pixels μιας φωτογραφίας, στα αντικείμενα ενός καλαθιού αγορών κ.α… Αυτό έχει ως αποτέλεσμα να μην μπορεί να χρησιμοποιηθεί κάποια συνάρτηση σφάλματος ή ανταμοιβής κατά την εκπαίδευση, με βάση την οποία θα αξιολογηθεί η ενδεχόμενη λύση. Αυτός ακριβώς είναι και ο λόγος που η εκμάθηση χωρίς

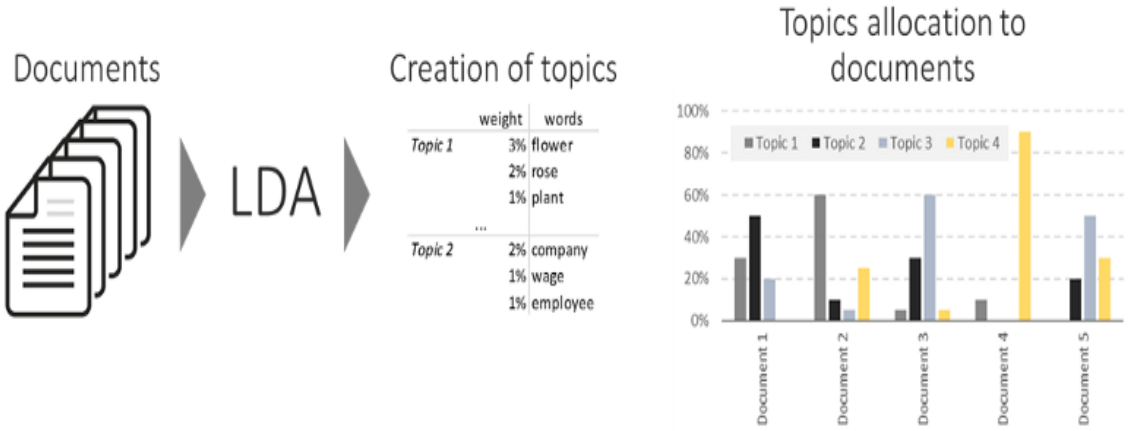
επίβλεψη συνδέεται άμεσα με μια πολύ διαδεδομένη τεχνική, την εκτίμηση πυκνότητας πιθανότητας (probability density estimation) μιας μη παρατηρήσιμης συνάρτησης βάσει ενός συνόλου στοιχείων παρατήρησης. Απλά παραδείγματα Μη-Επιβλεπόμενης Μάθησης είναι η Μοντελοποίηση Θεμάτων (Topic Modeling), η Συσταδοποίηση (Clustering) και η Μείωση Διαστάσεων (Dimensionality Reduction).

Μοντελοποίηση Θεμάτων με Λανθάνουσα Κατανομή Dirichlet (LDA)

Η Λανθάνουσα Κατανομή Dirichlet (Latent Dirichlet Allocation - LDA) είναι ένα παραγωγικό πιθανοτικό μοντέλο για συλλογές διακριτών δεδομένων, όπως κείμενα. Τα δεδομένα μοντελοποιούνται ως μια κατανομή Θεμάτων (Topics), και κάθε Θέμα με τη σειρά του ως κατανομή Λέξεων (Words). Σήμερα, δεν γνωρίζουμε με ακρίβεια σε ποιο βαθμό οι μηχανές αναζήτησης χρησιμοποιούν ήδη τη μοντελοποίηση θεμάτων στους αλγορίθμους τους, αλλά γνωρίζουμε ότι οι μηχανές αναζήτησης χρειάζονται μοντελοποίηση θεμάτων και ότι η κατάταξη των μηχανών αναζήτησης συσχετίζεται σημαντικά με την ανάλυση μοντελοποίησης θεμάτων. Γι’ αυτό, σε αυτή την πτυχιακή, το πρώτο καθήκον ήταν να δημιουργηθεί ένα μοντέλο LDA. Πολλοί εργαζόμενοι σε μεγάλες μηχανές αναζήτησης έχουν κάνει παρόμοια δουλειά επιλέγοντας το LDA για εξαγωγή θεμάτων, ένα παράδειγμα αυτών είναι ο πρώην Διευθυντής Ερευνών της Google, Edward Y. Chang. Ήταν σε μια ομάδα που δημιούργησε το Latent Dirichlet Allocation (LDA) που χρησιμοποιεί μια παράλληλη αρχιτεκτονική υπολογιστών που επιτρέπει στην Google να αξιοποιήσει τη τεράστια υπολογιστική δύναμη που έχει στη διάθεσή της.

Το LDA αναπαράγει τη διαδικασία εύρεσης θεμάτων. Για να το κάνει αυτό, κάνει τα εξής για κάθε έγγραφο m:

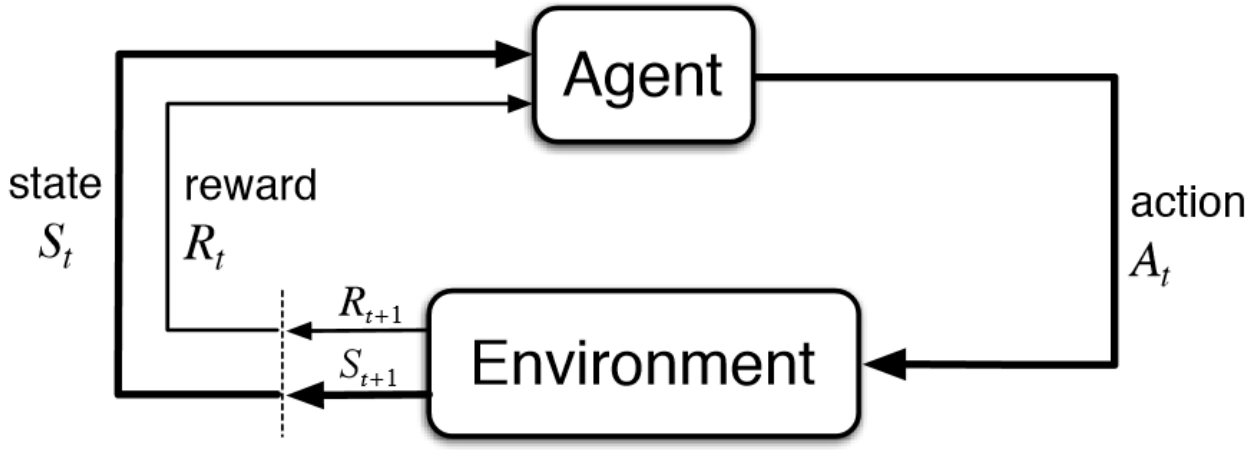
1. Υποθέτουμε ότι υπάρχουν k θέματα σε όλα τα έγγραφα.
2. Διανέμουμε αυτά τα θέματα k στο έγγραφο m (αυτή η κατανομή είναι γνωστή ως α και μπορεί να είναι συμμετρική ή ασύμμετρη), εκχωρώντας σε κάθε λέξη ένα θέμα.
3. Για κάθε λέξη w στο έγγραφο m, υποθέτουμε ότι το θέμα της είναι λάθος, αλλά σε κάθε άλλη λέξη αντιστοιχεί το σωστό θέμα.
4. Πιθανοτικά εκχωρούμε μια λέξη σε θέμα βασισμένο σε δύο πράγματα:
    - ποια θέματα βρίσκονται στο έγγραφο m
    - πόσες φορές η λέξη w έχει εκχωρηθεί σε ένα συγκεκριμένο θέμα σε όλα τα έγγραφα (αυτή η διανομή ονομάζεται β)
5. Επαναλαμβάνουμε αυτήν τη διαδικασία πολλές φορές για κάθε έγγραφο και ολοκληρώνουμε.



Εικόνα 1.3 : LDA input/output workflow [31]

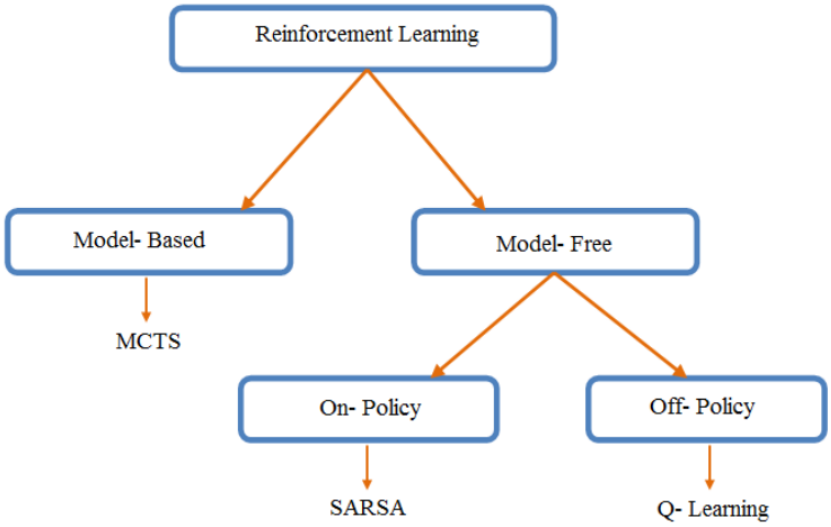
1.2.3 Ενισχυτική μάθηση

Η ενισχυτική μάθηση (EM) [2] αποτελεί ένα συνεπές μαθηματικό πλαίσιο για την εμπειρική, στοχοθετημένη μάθηση και λήψη αποφάσεων. Η ΕΜ ξεκινά με την αλληλεπίδραση μεταξύ πράκτορα και περιβάλλοντος, όπου ο πράκτορας παίρνει δράσεις στο περιβάλλον, που τον οδηγούν στην επόμενη κατάσταση και λαμβάνει ανταμοιβές βάσει της επιτυχίας αυτής της δράσης. Όπως φαίνεται στην εικόνα 1.4, σε οποιοδήποτε χρονικό βήμα t, ο πράκτορας βρίσκεται στην κατάσταση s_t, παίρνει δράση a_t στο περιβάλλον του, λαμβάνει ανταμοιβή r_t και παρατηρεί την επόμενη κατάσταση $s_t + 1$. Οι αλγόριθμοι ενισχυτικής μάθησης, αυξάνουν τις πιθανότητες λήψης καλών ενεργειών για την επίτευξη των επιθυμητών στόχων.



Εικόνα 1.4 : Ενισχυτική μάθηση [30]

Η ενισχυτική μάθηση μπορεί γενικά να υποδιαιρεθεί σε model-free και model-based, όπως φαίνεται στην εικόνα 1.5. Στην ΕΜ βασισμένη σε μοντέλο, χρησιμοποιείται ένα δυναμικό μοντέλο του περιβάλλοντος, ενώ στην ΕΜ ελεύθερου μοντέλου, μαθαίνεται μια πολιτική ή συνάρτηση αξίας.



Εικόνα 1.5 : Υποδιαιρέσεις ενισχυτικής μάθησης [31]

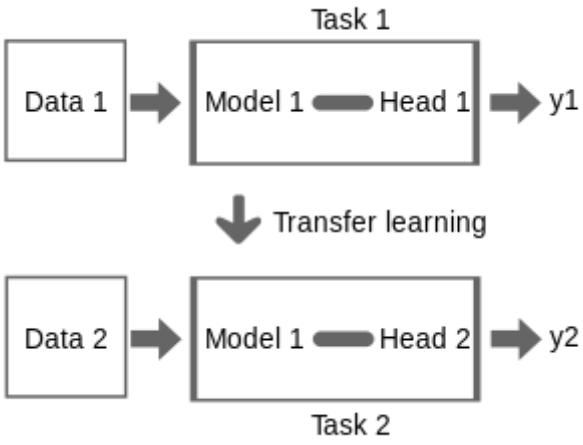
Η ΕΜ ελεύθερου μοντέλου χωρίζεται σε δύο ευρείες κατηγορίες, μάθηση εκτός πολιτικής και μάθηση εντός πολιτικής. Στις μεθόδους εκτός πολιτικής, η πολιτική που χρησιμοποιείται για τη δημιουργία συμπεριφοράς, ονομάζεται πολιτική συμπεριφοράς, μπορεί να μην έχει σχέση με την πολιτική που αξιολογείται και βελτιώνεται η οποία ονομάζεται πολιτική εκτίμησης. Ένα πλεονέκτημα αυτού του διαχωρισμού είναι ότι η πολιτική εκτίμησης μπορεί να είναι ντετερμινιστική (π.χ. άπληστη), ενώ η πολιτική συμπεριφοράς μπορεί να συνεχίσει να δειγματοληπτεί όλες τις δυνατές ενέργειες. Ο πράκτοράς μας θα μπορούσε ακόμη και να επιδεικνύει τυχαία συμπεριφορά και παρά αυτό, οι μέθοδοι εκτός πολιτικής μπορούν ακόμη να βρουν τη βέλτιστη πολιτική. Κυρίως, οι μέθοδοι εκτός πολιτικής συλλέγουν πληροφορίες από (εν μέρει) τυχαίες κινήσεις, αξιολογούν καταστάσεις σαν να χρησιμοποιούσαν μια άπληστη πολιτική και τελικά, μειώνουν σταδιακά την τυχαιότητα.

Η ενισχυτική μάθηση βασισμένη σε μοντέλο αναφέρεται στην εκμάθηση της βέλτιστης συμπεριφοράς έμμεσα, μαθαίνοντας ένα μοντέλο του περιβάλλοντος μέσω της λήψης δράσεων και της παρατήρησης των αποτελεσμάτων που περιλαμβάνουν την επόμενη κατάσταση και την άμεση ανταμοιβή. Αν ο πράκτορας μπορεί να μάθει κάνοντας προβλέψεις για τις συνέπειες των ενεργειών του, τότε πρόκειται για ΕΜ βασισμένη σε μοντέλο. Το προτέρημα αυτής της προσέγγισης είναι ότι ο πράκτορας μπορεί να κάνει προβλέψεις για τις πιθανές ανταμοιβές που σχετίζονται με ορισμένες ενέργειες. [3]

1.2.4 Transfer learning

Η μεταφορά μάθησης (TL), είναι μια μέθοδος στη μηχανική μάθηση (ML) που εκμεταλλεύεται τη γνώση που αποκτήθηκε από μια εργασία για να ενισχύσει την απόδοση σε μια παρόμοια εργασία. Για παράδειγμα, στην ταξινόμηση εικόνων, η κατανόηση που αναπτύχθηκε από την αναγνώριση αυτοκινήτων μπορεί να χρησιμοποιηθεί για να βελτιώσει την αναγνώριση των φορτηγών. Αυτή η έννοια έχει συνδέσεις με τις ψυχολογικές μελέτες για τη μεταφορά της μάθησης, αν και πρακτικά οι δεσμοί μεταξύ αυτών των δύο περιοχών είναι ελάχιστοι. Η πρακτική της επανεφαρμογής της γνώσης από παλαιότερες εργασίες σε νέες υπόσχεται ως τομέας να ενισχύσει σημαντικά την αποδοτικότητα της μάθησης.

Καθώς το transfer learning περιλαμβάνει την εκπαίδευση με διάφορες συναρτήσεις στόχου, συνδέεται με τη μηχανική μάθηση ευαίσθητη στο κόστος και την πολλαπλή βελτιστοποίηση. Αυτό οφείλεται στο γεγονός ότι όλες αυτές οι τεχνικές στοχεύουν στη βελτιστοποίηση της απόδοσης με βάση πολλαπλά κριτήρια.



Εικόνα 1.6 : Απεικόνιση του Transfer Learning

Όπως αναφέρθηκε και προηγουμένως, κατά τη διάρκεια του transfer learning, η γνώση που προέρχεται από μια αρχική εργασία χρησιμοποιείται για να ενισχύσει τη μάθηση σε μια παρεμφερή εργασία. Όταν η τεχνική μεταφοράς όμως οδηγεί σε μείωση της απόδοσης της νέας εργασίας, αυτό ονομάζεται αρνητική μεταφορά. Ένα από τα σημαντικά εμπόδια στη διαμόρφωση μεθόδων μεταφοράς είναι η εξασφάλιση θετικής μεταφοράς μεταξύ σχετικών εργασιών και ταυτόχρονα να αποφευχθεί αρνητική μεταφορά μεταξύ λιγότερο σχετικών εργασιών.

Όταν χρησιμοποιούμε γνώση από μια εργασία για μια άλλη, τα χαρακτηριστικά της αρχικής εργασίας συνήθως προβάλλονται στην άλλη εργασία για να καθοριστεί ένας συσχετισμός μεταξύ τους. Αυτός ο συσχετισμός γίνεται συνήθως από έναν άνθρωπο, αλλά υπάρχουν εξελισσόμενες μέθοδοι που μπορούν να αυτοματοποιήσουν αυτήν τη διαδικασία.

Η αποτελεσματικότητα των τεχνικών του transfer learning μπορούν να αξιολογηθούν με τρεις κοινές μετρικές:

- Η πρώτη μετρική αξιολογεί εάν η εκτέλεση της στοχευμένης εργασίας είναι εφικτή μόνο με τη μεταφερόμενη γνώση.

- Η δεύτερη μετρική μετρά τον χρόνο που απαιτείται για να μάθει ο αλγόριθμος τη στοχευμένη εργασία με τη βοήθεια της μεταφερόμενης μάθησης, σε σύγκριση με τον χρόνο που θα χρειαζόταν χωρίς αυτήν.

- Η τρίτη μετρική εξετάζει αν η απόδοση της εργασίας που έχει μάθει μέσω της μεταφοράς μάθησης είναι συγκρίσιμη με την ολοκλήρωση της αρχικής εργασίας χωρίς τη χρήση της μεταφοράς μάθησης.

Η μεταφορά μάθησης μπορεί να επιτευχθεί με διάφορους τρόπους. Ένας τρόπος είναι να βρούμε μια σχετική εργασία μάθησης, που ονομάζεται Εργασία Β, η οποία διαθέτει πολλά μεταφερόμενα επισημασμένα δεδομένα. Το νέο μοντέλο εκπαιδεύεται στην Εργασία Β. Μετά από αυτή την εκπαίδευση, το μοντέλο έχει ένα σημείο εκκίνησης για την επίλυση της αρχικής του εργασίας, δηλαδή της Εργασίας Α.

Ένας άλλος τρόπος για να επιτευχθεί η μεταφορά μάθησης είναι η χρήση ενός προ-εκπαιδευμένου μοντέλου. Αυτή η διαδικασία είναι πιο εύκολη, καθώς περιλαμβάνει τη χρήση ενός ήδη εκπαιδευμένου μοντέλου. Το προ-εκπαιδευμένο μοντέλο θα πρέπει να έχει εκπαιδευτεί χρησιμοποιώντας ένα μεγάλο σύνολο δεδομένων για να επιλύσει μια παρόμοια εργασία όπως η εργασία Α. Τα μοντέλα αυτά δημοσιεύονται από προγραμματιστές στο διαδίκτυο για να μπορεί κάποιος τρίτος να τα χρησιμοποιήσει για την δική του υλοποίηση.

Μια τρίτη μέθοδος, γνωστή ως εξαγωγή χαρακτηριστικών ή μάθηση αναπαράστασης, χρησιμοποιεί τη βαθιά μάθηση για να εντοπίσει τα πιο σημαντικά χαρακτηριστικά για την Εργασία Α, τα οποία στη συνέχεια λειτουργούν ως αναπαράσταση της εργασίας. Αντί να δημιουργούνται χειροκίνητα, τα χαρακτηριστικά εξάγονται αυτόματα μέσω της βαθιάς μάθησης. Οι επιστήμονες δεδομένων πρέπει στη συνέχεια να επιλέξουν ποια από αυτά τα χαρακτηριστικά θα συμπεριλάβουν στο μοντέλο. Αυτή η τελική αναπαράσταση μπορεί επίσης να χρησιμοποιηθεί για άλλες μελλοντικές εργασίες.

Η μεταφορά μάθησης μπορεί να εφαρμοστεί σε τομείς όπως τα νευρωνικά δίκτυα, η επεξεργασία φυσικής γλώσσας (NLP) και η υπολογιστική όραση.

Στη μηχανική μάθηση, η γνώση ή τα δεδομένα που αποκτώνται κατά την επίλυση ενός προβλήματος αποθηκεύονται, επισημαίνονται και στη συνέχεια εφαρμόζονται σε ένα διαφορετικό αλλά σχετικό πρόβλημα. Για παράδειγμα, η γνώση που αποκτά ένας αλγόριθμος μηχανικής μάθησης για την αναγνώριση αυτοκινήτων μπορεί αργότερα να μεταφερθεί για χρήση σε ένα άλλο μοντέλο μηχανικής μάθησης που αναπτύσσεται για την αναγνώριση άλλων τύπων οχημάτων.

Το transfer learning είναι επίσης χρήσιμο κατά την ανάπτυξη υψηλότερης τεχνολογίας, όπως ένα chatbot. Εάν ο νέος τομέας είναι αρκετά παρόμοιος με προηγούμενες εφαρμογές, η μεταφορά μάθησης μπορεί να αξιολογήσει ποια γνώση πρέπει να μεταφερθεί. Χρησιμοποιώντας τη μεταφορά μάθησης, οι προγραμματιστές μπορούν να αποφασίσουν ποια γνώση και δεδομένα από προηγούμενες υλοποιήσεις

είναι επαναχρησιμοποιήσιμα και να τα μεταφέρουν για χρήση στην ανάπτυξη μιας αναβαθμισμένης έκδοσης της εκάστοτε εφαρμογής.

Στην επεξεργασία φυσικής γλώσσας (NLP), για παράδειγμα, ένα σύνολο δεδομένων από ένα παλιό μοντέλο που κατανοεί το λεξιλόγιο μιας περιοχής μπορεί να χρησιμοποιηθεί για την εκπαίδευση ενός νέου μοντέλου που στοχεύει στην κατανόηση διαλέκτων σε πολλές περιοχές. Μια εταιρεία θα μπορούσε στη συνέχεια να το εφαρμόσει αυτό σε ένα έργο της για την ανάλυση των συναισθημάτων.

Ένα νευρωνικό δίκτυο χρησιμοποιείται συνήθως για την ανάλυση ιατρικών εικόνων με σκοπό την αναγνώριση πιθανών ασθενειών. Σε αυτή την περίπτωση, η μεταφορά μάθησης μπορεί να βοηθήσει στην αναγνώριση αυτών των ασθενειών χρησιμοποιώντας προ-εκπαιδευμένα μοντέλα, ειδικά όταν υπάρχουν ανεπαρκή δεδομένα για την εκπαίδευση του δικτύου.

1.3 Βαθιά μάθηση

Η βαθιά μάθηση έχει μακρά ιστορία στον ερευνητικό χώρο και πολλές φιλοδοξίες ως τομέας. Αρκετές προτεινόμενες προσεγγίσεις της όμως δεν έχουν ακόμη καρποφορήσει πλήρως, όπως επίσης και αρκετοί φιλόδοξοι στόχοι της δεν έχουν ακόμη πραγματοποιηθεί. Η σύγχρονη βαθιά μάθηση παρέχει ένα ισχυρό πλαίσιο για την εποπτευόμενη μάθηση. Προσθέτοντας περισσότερα επίπεδα και περισσότερες μονάδες εντός ενός επιπέδου, ένα βαθύ δίκτυο μπορεί να αναπαραστήσει συναρτήσεις αυξανόμενης πολυπλοκότητας. Οι περισσότερες εργασίες που αποτελούνται από την χαρτογράφηση ενός διανύσματος εισόδου σε ένα διάνυσμα εξόδου, και που είναι εύκολο για ένα άτομο να το κάνει γρήγορα, μπορούν να επιτευχθούν μέσω της βαθιάς μάθησης, δεδομένου ότι υπάρχουν επαρκώς μεγάλα μοντέλα και επαρκώς μεγάλα σύνολα δεδομένων με επισημασμένα παραδείγματα εκπαίδευσης. Αντιθέτως, άλλες εργασίες που δεν μπορούν να περιγραφούν ως συσχέτιση ενός διανύσματος με ένα άλλο, ή που είναι αρκετά δύσκολες ώστε ένα άτομο να χρειάζεται χρόνο για να σκεφτεί και να τις αναλύσει για να ολοκληρώσει την εργασία, παραμένουν εκτός στόχων της βαθιάς μάθησης για τώρα.

Διάφορες αρχιτεκτονικές της βαθιάς μάθησης, όπως τα βαθιά νευρωνικά δίκτυα, deep belief networks, επαναλαμβανόμενα νευρωνικά δίκτυα, συνελικτικά νευρωνικά δίκτυα και οι transformers έχουν εφαρμοστεί σε πεδία που περιλαμβάνουν την υπολογιστική όραση, την αναγνώριση ομιλίας, την επεξεργασία φυσικής γλώσσας, τη μηχανική μετάφραση, τη βιοπληροφορική, τον σχεδιασμό φαρμάκων, την ανάλυση ιατρικών εικόνων, την κλιματολογία, την επιθεώρηση υλικών και τα προγράμματα επιτραπέζιων παιχνιδιών, όπου έχουν παράγει αποτελέσματα συγκρίσιμα και σε ορισμένες περιπτώσεις υπερβαίνοντας την απόδοση των ανθρώπων.

Τα περισσότερα σύγχρονα μοντέλα βαθιάς μάθησης βασίζονται σε πολυεπίπεδα τεχνητά νευρωνικά δίκτυα, όπως τα convolutional neural networks και οι transformers, αν και μπορούν να περιλαμβάνουν επίσης προτασιακές φόρμουλες ή κρυφές μεταβλητές οργανωμένες σε επίπεδα σε deep generative models, όπως οι κόμβοι στα deep belief networks και βαθιές μηχανές Boltzmann. [4]

1.3.1 MLP classifier

Τα MLPs, αποτελούν μια σημαντική τεχνική στη μηχανική μάθηση που ανήκει στην κατηγορία των τεχνητών νευρωνικών δικτύων, χρησιμοποιούνται συνήθως για εργασίες ταξινόμησης. Η προσαρμόσιμη και αποτελεσματική προσέγγιση τους μπορεί να χειριστεί μια ευρεία γκάμα ζητημάτων ταξινόμησης, όπως η ταξινόμηση κειμένου και η αναγνώριση εικόνων. Σε προβλήματα όπου οι παραδοσιακοί γραμμικοί ταξινομητές μπορεί να μην ανταποκρίνονται, τα MLPs είναι γνωστά για την ικανότητά τους να αναπαριστούν πολύπλοκες, μη γραμμικές συσχετίσεις στα δεδομένα.

Ένα MLP είναι ένας τύπος τεχνητού νευρωνικού δικτύου που χαρακτηρίζεται από πολλαπλά επίπεδα διασυνδεδεμένων κόμβων, γνωστών επίσης ως νευρώνες. Χρησιμοποιείται συχνά για διάφορες εργασίες μηχανικής μάθησης, συμπεριλαμβανομένων της ταξινόμησης και της παλινδρόμησης. Παρακάτω παρουσιάζεται μια σύνοψη της δομής και της λειτουργίας ενός MLP:

Δομή

Input Layer: Το επίπεδο εισόδου αποτελείται από νευρώνες που λαμβάνουν άμεσα τα χαρακτηριστικά του συνόλου δεδομένων. Κάθε νευρώνας σε αυτό το επίπεδο αντιστοιχεί σε ένα χαρακτηριστικό, και ο συνολικός αριθμός των νευρώνων ισούται με τον συνολικό αριθμό των χαρακτηριστικών στο σύνολο δεδομένων.

Hidden Layer: Μπορεί να υπάρχει ένα ή περισσότερα κρυφά επίπεδα που βρίσκονται μεταξύ του επιπέδου εισόδου και του επιπέδου εξόδου. Ο αριθμός των νευρώνων σε κάθε κρυφό επίπεδο είναι μία παράμετρος που μπορεί να προσαρμοστεί και διαφέρει ανά επίπεδο. Τα κρυφά επίπεδα είναι ζωτικής σημασίας για την αναγνώριση πολύπλοκων μοτίβων στα δεδομένα.

Output Layer: Το επίπεδο εξόδου παράγει τις τελικές προβλέψεις ή εξόδους, χρησιμοποιώντας τα δεδομένα που επεξεργάστηκαν στα κρυφά επίπεδα. Ο αριθμός των νευρώνων στο επίπεδο εξόδου καθορίζεται από τις απαιτήσεις της εργασίας: Για δυαδική ταξινόμηση, υπάρχει συνήθως ένας νευρώνας που παράγει έναν βαθμό πιθανότητας. Στην ταξινόμηση πολλαπλών κλάσεων, ο αριθμός των νευρώνων είναι ίσος με τον αριθμό των κλάσεων, με κάθε νευρώνα να παράγει έναν βαθμό πιθανότητας για μια συγκεκριμένη κλάση. Για προβλήματα παλινδρόμησης, ένας νευρώνας παράγει τη συνεχή προβλεπόμενη τιμή.

Λειτουργία

Initialization: Τα βάρη (W) και οι προκαθορισμένες τιμές (B) όλων των νευρώνων στο δίκτυο ορίζονται στις αρχικές τους τιμές. Συνήθως, αυτές οι παράμετροι αρχικοποιούνται με μικρούς τυχαίους αριθμούς.

Forward Propagation: Κατά τη διάρκεια της εκπαίδευσης, τα δεδομένα εισόδου διέρχονται διαρκώς από το δίκτυο. Κάθε νευρώνας σε ένα επίπεδο λαμβάνει το σταθμισμένο άθροισμα των εισόδων από

το προηγούμενο επίπεδο, εφαρμόζει μια συνάρτηση ενεργοποίησης και προωθεί το αποτέλεσμα στο επόμενο επίπεδο. Οι συναρτήσεις ενεργοποίησης εισάγουν μη γραμμικότητα στο μοντέλο, επιτρέποντάς του να μάθει πολύπλοκες σχέσεις.

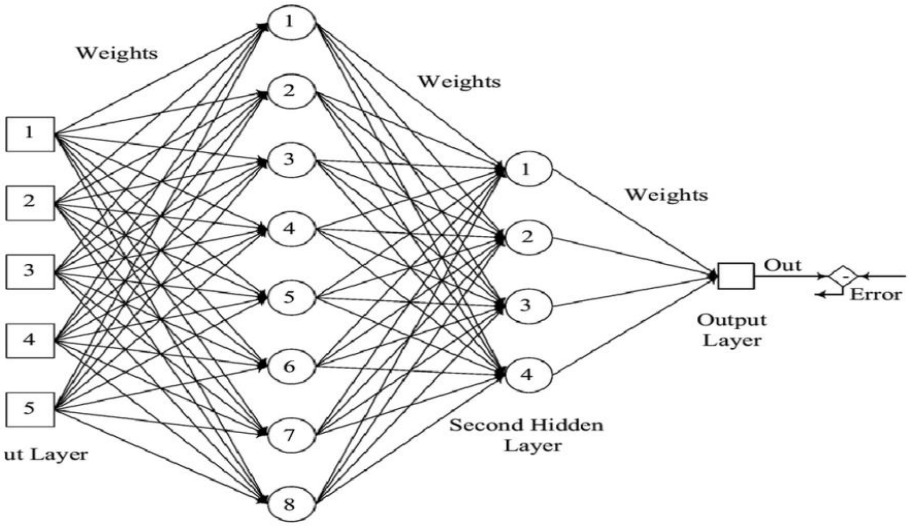
Loss Calculation: Υπολογίζεται μια απώλεια (ή σφάλμα) συγκρίνοντας την έξοδο του δικτύου με τις τιμές που στόχευε. Οι κοινές συναρτήσεις απώλειας περιλαμβάνουν το Μέσο Τετραγωνικό Σφάλμα (MSE) για εργασίες παλινδρόμησης και την Εντροπία Διασταύρωσης για εργασίες ταξινόμησης.

Backpropagation: Για να ελαχιστοποιήσει την απώλεια, το δίκτυο προσαρμόζει τα βάρη και τις προκαθορισμένες τιμές του. Αυτό επιτυγχάνεται με τον αλγόριθμο αντίστροφης διάδοσης, ο οποίος υπολογίζει τις κλίσεις της απώλειας σε σχέση με κάθε παράμετρο του δικτύου. Στη συνέχεια, αυτές οι κλίσεις χρησιμοποιούνται για την ενημέρωση των βαρών και των προκαθορισμένων τιμών με τη χρήση μεθόδων βελτιστοποίησης, όπως η Gradient Descent.

Training: Οι διαδικασίες της προώθησης, του υπολογισμού της απώλειας και της αντίστροφης διάδοσης επαναλαμβάνονται κατά τη διάρκεια πολλών επαναλήψεων (εποχές) μέχρι το μοντέλο να συγκλίνει σε μια λύση. Ο ρυθμός μάθησης και ο αριθμός των επαναλήψεων είναι παραδείγματα παραμέτρων που μπορούν να προσαρμοστούν.

Prediction: Μόλις τα βάρη και οι προκαθορισμένες τιμές έχουν προσαρμοστεί μέσω της προώθησης, το MLP μπορεί να χρησιμοποιηθεί για να κάνει προβλέψεις σε μη παρατηρήσιμα δεδομένα.

Παρά το γεγονός ότι τα MLPs είναι ιδιαίτερα ικανά στο να απεικονίζουν πολύπλοκες σχέσεις στα δεδομένα, μπορούν να είναι ευαίσθητα σε συγκεκριμένες παραμέτρους. Αυτές περιλαμβάνουν τον αριθμό των κρυφών επιπέδων και νευρώνων, την επιλογή των συναρτήσεων ενεργοποίησης και τις τεχνικές κανονικοποίησης. Για την αποτελεσματική λειτουργία των MLPs, είναι ουσιώδης η κατάλληλη προσαρμογή αυτών των παραμέτρων.



Εικόνα 1.7 : Σχηματική Απεικόνιση ενός MLP classifier [44]

1.3.2 CNN classifier

Ένα Συνελικτικό Νευρωνικό Δίκτυο (CNN) είναι ένας εξειδικευμένος τύπος νευρωνικού δικτύου προώθησης που μαθαίνει αυτόνομα την μηχανική των χαρακτηριστικών μέσω της βελτιστοποίησης των φίλτρων (ή πυρήνων). Μία βασική χρήση των CNNs ως classifiers είναι αυτή της κατηγοριοποίησης εικόνων.

Η ταξινόμηση των εικόνων περιλαμβάνει την ανάθεση μιας ετικέτας ή κατηγορίας σε μια δεδομένη εικόνα. Αυτό είναι ένα πρόβλημα επιβλεπόμενης μάθησης, όπου ένα μοντέλο εκπαιδεύεται χρησιμοποιώντας ένα σύνολο δεδομένων εικόνων που έχουν τους ανατεθεί συγκεκριμένες ετικέτες κλάσης. Στη συνέχεια, το μοντέλο χρησιμοποιείται για την πρόβλεψη της ετικέτας κλάσης των νέων εικόνων που τροφοδοτούμε.

Τα CNNs είναι ιδιαίτερα αποτελεσματικά στην ταξινόμηση εικόνων λόγω της ικανότητάς τους να μάθουν αυτόνομα τις χωρικές ιεραρχίες των χαρακτηριστικών. Αυτά τα χαρακτηριστικά περιλαμβάνουν στοιχεία όπως άκρες, υφές και σχήματα, τα οποία είναι κρίσιμα για την αναγνώριση αντικειμένων σε εικόνες.

Τα CNNs αποτελούνται από μια σειρά συνδεδεμένων επιπέδων που επεξεργάζονται τα δεδομένα εισόδου. Το πρώτο κρυφό επίπεδο ενός CNN είναι συνήθως ένα συνελικτικό επίπεδο που εφαρμόζει ένα σύνολο φίλτρων στα δεδομένα εισόδου για να ανιχνεύσει συγκεκριμένα μοτίβα. Κάθε φίλτρο δημιουργεί έναν χάρτη χαρακτηριστικών κινούμενο πάνω από τα δεδομένα εισόδου και εκτελώντας πολλαπλασιασμό στοιχείο προς στοιχείο με τις καταχωρήσεις στο φίλτρο. Αυτοί οι χάρτες χαρακτηριστικών στη συνέχεια συνδυάζονται και περνούν μέσα από μη γραμμικές συναρτήσεις ενεργοποίησης, όπως η συνάρτηση ReLU. Αυτό εισάγει μη γραμμικότητες στο μοντέλο, επιτρέποντάς του να μάθει πιο περίπλοκα μοτίβα που βρίσκονται στα δεδομένα.

Τα επόμενα επίπεδα σε ένα CNN μπορεί να περιλαμβάνουν περισσότερα συνελικτικά επίπεδα, επίπεδα pooling και πλήρως συνδεδεμένα επίπεδα. Τα επίπεδα pooling μειώνουν το μέγεθος των χαρτών χαρακτηριστικών, το οποίο βοηθά στη μείωση του συνολικού αριθμού των παραμέτρων στο μοντέλο, καθιστώντας το πιο αποδοτικό από υπολογιστική άποψη. Τα πλήρως συνδεδεμένα επίπεδα βρίσκονται συνήθως μετά τα συνελικτικά και τα επίπεδα pooling ενός CNN. Ένα πλήρως συνδεδεμένο επίπεδο συνδέει όλους τους νευρώνες σε ένα επίπεδο με όλους τους νευρώνες στο επόμενο επίπεδο, επιτρέποντας στο μοντέλο να μάθει πιθανούς μη γραμμικούς συνδυασμούς των χαρακτηριστικών που έχουν αποκτηθεί από τα συνελικτικά επίπεδα.

Το τελευταίο επίπεδο ενός CNN είναι συνήθως ένα επίπεδο softmax. Αυτό το επίπεδο δημιουργεί μια κατανομή πιθανοτήτων πάνω στις πιθανές ετικέτες κλάσης για τα δεδομένα εισόδου. Η πρόβλεψη του μοντέλου είναι η κλάση που έχει την υψηλότερη πιθανότητα.

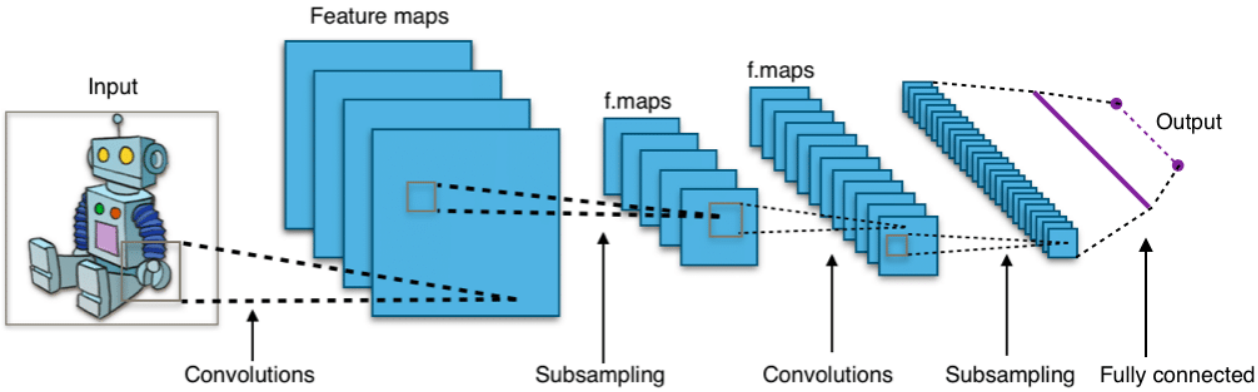
Παρακάτω παρατίθενται μερικά κρίσιμα σημεία για την κατανόηση των CNNs classifiers:

- Τα CNNs έχουν τη δυνατότητα να μάθουν να αναγνωρίζουν μοτίβα και χαρακτηριστικά σε εικόνες χρησιμοποιώντας συνελικτικά επίπεδα. Αυτά τα επίπεδα εφαρμόζουν μια συλλογή φίλτρων στα δεδομένα εισόδου για να αναγνωρίσουν συγκεκριμένα μοτίβα.

- Τα CNNs μπορούν να μάθουν αυτόνομα χωρικές ιεραρχίες χαρακτηριστικών, ξεκινώντας από απλά μοτίβα όπως οι ακμές και προχωρώντας σε πιο περίπλοκα μοτίβα καθώς τα επίπεδα βαθαίνουν. Αυτή η ιεραρχική εκμάθηση χαρακτηριστικών είναι ιδιαίτερα κατάλληλη για την ταξινόμηση εικόνων, όπου τα οπτικά χαρακτηριστικά μιας εικόνας μπορούν να διαφέρουν σημαντικά.

- Ορισμένες αρχιτεκτονικές των CNN έχουν τη δυνατότητα να επεξεργάζονται εικόνες σε πραγματικό χρόνο, καθιστώντας τα ιδανικά για εφαρμογές όπου η γρήγορη ταξινόμηση είναι κρίσιμη, όπως για παράδειγμα σε αυτόνομα αυτοκίνητα ή συστήματα ασφαλείας.

- Τα CNNs έχουν επιτύχει κορυφαίες επιδόσεις σε πολλά benchmarks ταξινόμησης εικόνων και χρησιμοποιούνται ευρέως στη βιομηχανία και την έρευνα.



Εικόνα 1.8 : Τυπική αρχιτεκτονική ενός CNN [44]

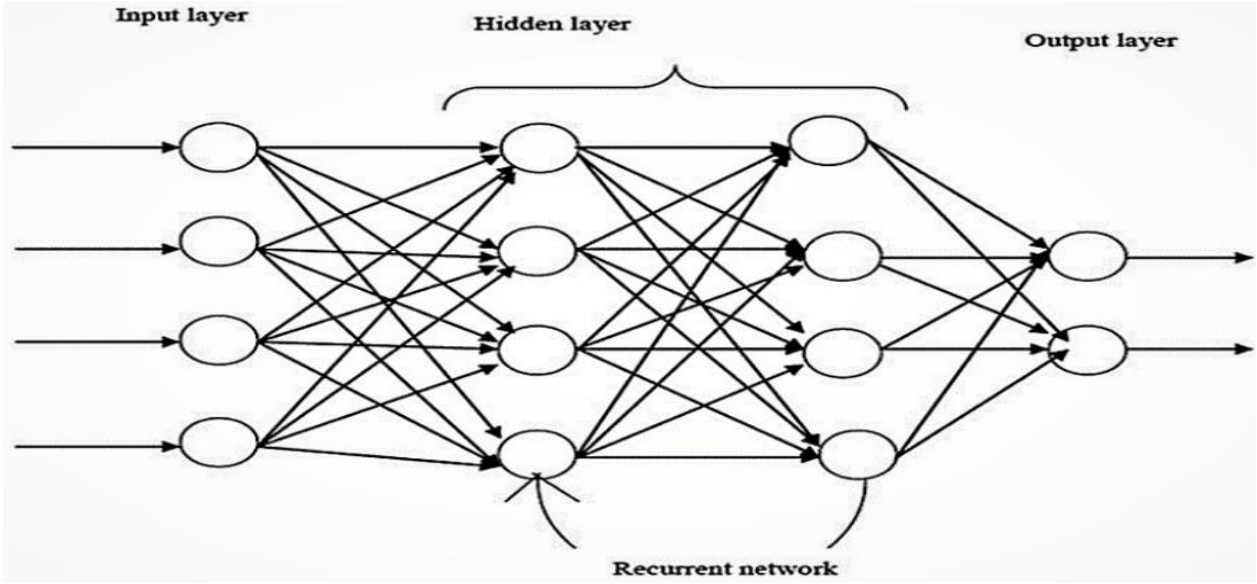
1.3.3 RNN classifier

Ένα Επαναλαμβανόμενο Νευρωνικό Δίκτυο (RNN) είναι μια δομή δικτύου βαθιάς μάθησης που κάνει προβλέψεις σε δεδομένα ακολουθίας ή χρονοσειρών.

Τα RNNs είναι ιδιαίτερα αποτελεσματικά όταν τροφοδοτούνται με δεδομένα ακολουθίας που διαφέρουν σε μήκος και είναι ικανά να λύσουν προβλήματα όπως η ταξινόμηση φυσικών σημάτων, η επεξεργασία γλώσσας και η ανάλυση βίντεο.

Ένα RNN θεωρείται μια αρχιτεκτονική βαθιάς μάθησης που εκμεταλλεύεται προηγούμενα δεδομένα και πληροφορίες για να βελτιώσει την απόδοση του δικτύου σε τρέχουσες και μελλοντικές εισόδους. Η διαφορετικότητα ενός RNN σε σχέση με άλλα βρίσκεται στο γεγονός ότι το δίκτυο περιέχει μία κρυφή κατάσταση και βρόχους. Αυτή η δομή βρόχου επιτρέπει στο δίκτυο να διατηρεί παλαιότερες πληροφορίες στην κρυφή κατάσταση και να λειτουργεί πάνω σε ακολουθίες.

Όταν ένα RNN χρησιμοποιείται ως classifier λαμβάνει μια ακολουθία εισόδων και τις διαβιβάζει μέσα από ένα δίκτυο νευρώνων. Το δίκτυο περιέχει βρόχους που επιτρέπουν τη μετάδοση πληροφοριών από το ένα βήμα στην ακολουθία στο επόμενο. Αυτό επιτρέπει στο RNN να χρησιμοποιεί όπως αναφέρθηκε και προηγουμένως πληροφορίες από το παρελθόν για να επηρεάσει την τρέχουσα έξοδο. Κατά τη διάρκεια της εκπαίδευσης, τα βάρη του δικτύου προσαρμόζονται για να ελαχιστοποιήσουν τη διαφορά μεταξύ των predicted και πραγματικών εξόδων. Μόλις εκπαιδευτεί, το RNN μπορεί να ταξινομήσει νέες ακολουθίες διαβιβάζοντάς τες μέσα από το δίκτυο και εξάγοντας την κλάση με την υψηλότερη πιθανότητα.



Εικόνα 1.9 : Σχηματική Απεικόνιση ενός RNN classifier [44]

Τα μοντέλα γλώσσας που βασίζονται σε RNNs αξίζει να σημειωθεί ότι μπορούν να χρησιμοποιηθούν για τη δημιουργία κειμένου. Η δημιουργία κειμένου έχει σημαντική πρακτική αξία και εμπλέκεται σε εργασίες όπως η απάντηση ερωτήσεων, η μηχανική μετάφραση, η περίληψη κειμένου, η διόρθωση γραμματικής, η δημιουργία ιστοριών και ο διάλογος συνομιλίας. Ουσιαστικά, οποιαδήποτε εργασία που απαιτεί από ένα σύστημα να παράγει κείμενο με βάση κάποιο άλλο κείμενο όπου έχει εισαχθεί. Η χρήση ενός μοντέλου γλώσσας για τη δημιουργία κειμένου είναι ένας από τους τομείς όπου η επίδραση των νευρωνικών μοντέλων γλώσσας στην NLP ήταν η πιο έντονη. Η δημιουργία κειμένου, μαζί με τη δημιουργία εικόνων και κώδικα, αποτελούν έναν νέο τομέα της AI που συχνά αποκαλείται generative AI. Τέτοιου είδους μοντέλα θα αναφερθούν στο 2^{ο} κεφάλαιο όταν θα μιλήσουμε για δημιουργία συνθετικών δεδομένων. [6],[7]

1.3.4 Frugal learning

Η μηχανική μάθηση πλέον είναι ένας βασικός τομέας ο οποίος χρησιμοποιείται στην έρευνα, στην βιομηχανία αλλά σιγά σιγά εισχωρεί και στην ζωή όλων μας, με πολλές εφαρμογές που απευθύνονται στον μέσο άνθρωπο. Η άνθηση αυτή τα τελευταία χρόνια έφερε στο προσκήνιο σημαντικά θέματα που αντιμετωπίζουν τα συστήματα όπου την χρησιμοποιούν, όπως η τεράστια υπολογιστική ισχύ που χρειάζεται καθώς και τα πάρα πολλά σε αριθμό δεδομένα που χρησιμοποιούνται για την εκμάθηση τέτοιον αλγορίθμων. Με αυτόν τον τρόπο μελετώνται τρόποι με τους οποίους μπορούμε να βελτιώσουμε και να αποφύγουμε τέτοιες δυσκολίες, μία πρόταση που στηρίζεται αρκετά πάνω στο θέμα τα τελευταία χρόνια είναι το λεγόμενο frugal learning.

Η έννοια του frugal learning επισημαίνει το κόστος που σχετίζεται με τη χρήση δεδομένων και υπολογιστικών πόρων. Το frugality, δηλαδή η ιδέα της εργασίας με περιορισμένους πόρους, παρουσιάζεται σε διάφορες μορφές:

- Το input frugality επισημαίνει το κόστος που σχετίζεται με τα δεδομένα, ειδικά με την απόκτηση των δεδομένων εκπαίδευσης, την εκμετάλλευση των περιγραφικών χαρακτηριστικών, ή και τα δύο. Οι frugal είσοδοι μπορεί να περιλαμβάνουν λιγότερα δεδομένα εκπαίδευσης ή λιγότερα χαρακτηριστικά από ό,τι απαιτείται για την καλύτερη ποιότητα πρόβλεψης που είναι εφικτή με τους παραδοσιακούς κανόνες της μηχανικής μάθησης. Το input frugality μπορεί να είναι χρήσιμο όταν έχουμε περιορισμούς πόρων ή περιορισμούς λόγω ιδιωτικότητας (για παράδειγμα σε στρατιωτικές ή ιατρικές εφαρμογές).

- Το frugality της διαδικασίας μάθησης επισημαίνει το κόστος που σχετίζεται με τη διαδικασία μάθησης, ειδικά τους υπολογιστικούς λόγους και τους πόρους μνήμης. Το frugality μάθησης μπορεί να παράγει ένα μοντέλο με χαμηλότερη ποιότητα πρόβλεψης από ό,τι είναι εφικτό με τις παραδοσιακές τεχνικές μηχανικής μάθησης, αλλά το κάνει πολύ πιο αποτελεσματικά. Το frugality της διαδικασίας μάθησης χρειάζεται κυρίως όταν έχουμε περιορισμούς πόρων, συμπεριλαμβανομένης της περιορισμένης υπολογιστικής ισχύος και της περιορισμένης χωρητικότητας της μπαταρίας.

- Το frugality του μοντέλου επισημαίνει το κόστος που σχετίζεται με την αποθήκευση ή τη χρήση ενός μοντέλου μηχανικής μάθησης, όπως ένα ταξινομητής ή ένα μοντέλο παλινδρόμησης. Για την εποπτευόμενη μάθηση, τα frugal μοντέλα μπορούν να απαιτούν λιγότερη μνήμη και να παράγουν προβλέψεις με λιγότερη υπολογιστική προσπάθεια από ό,τι απαιτείται για την βέλτιστη ποιότητα πρόβλεψης. Το frugality του μοντέλου χρειάζεται κυρίως όταν έχουμε περιορισμούς πόρων όπως περιορισμένη μνήμη ή περιορισμένες δυνατότητες επεξεργασίας.

Στην παρούσα διατριβή όλες οι υλοποιήσεις έχουν γίνει με βάση τις αρχές του frugal learning ώστε να διαπιστωθεί κατά πόσο μπορούμε να έχουμε βέλτιστο αποτέλεσμα με τα λιγότερα δυνατά δεδομένα και κατά πόσο μπορούμε να αποφύγουμε απαιτητικές σε υπολογιστικούς πόρους τεχνικές που είναι ένα μεγάλο ζητούμενο στις μέρες μας. [9]

ΚΕΦΑΛΑΙΟ 2

ΔΗΜΙΟΥΡΓΙΑ ΣΥΝΘΕΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

2.1 Εισαγωγή

Τα συνθετικά δεδομένα αναφέρονται σε δεδομένα που δεν προέρχονται από πραγματικά γεγονότα, αλλά παράγονται τεχνητά. Οι αλγόριθμοι χρησιμοποιούνται για τη δημιουργία αυτών των δεδομένων, τα οποία στη συνέχεια χρησιμοποιούνται για τη δοκιμή συνόλων δεδομένων λειτουργικών. Οι κύριες χρήσεις τους περιλαμβάνουν την επικύρωση μαθηματικών μοντέλων και την εκπαίδευση μοντέλων βαθιάς μάθησης σε συνθετικά δεδομένα.

Η χρήση των συνθετικών δεδομένων έχει σημαντικά οφέλη. Μειώνει τους περιορισμούς όταν ασχολούμαστε με ρυθμιζόμενα ή ευαίσθητα δεδομένα και επιτρέπει τη δημιουργία δεδομένων προσαρμοσμένων σε συγκεκριμένες ανάγκες που δεν μπορούν να ικανοποιηθούν με πραγματικά δεδομένα. Τα συνθετικά σύνολα δεδομένων παράγονται συνήθως για σκοπούς διασφάλισης της ποιότητας και δοκιμών λογισμικού.

Ωστόσο, τα συνθετικά δεδομένα παρουσιάζουν επίσης μειονεκτήματα. Αυτές περιλαμβάνουν ασυνέπειες που προκύπτουν όταν προσπαθούμε να αντιγράψουμε την πολυπλοκότητα που είναι ενσωματωμένη στα αρχικά δεδομένα, και το γεγονός ότι δεν μπορεί απλά να αντικαταστήσει τα πραγματικά δεδομένα επειδή θα κάνει χρήση των πραγματικών για να παράγει χρήσιμα αποτελέσματα.

Τα συνθετικά δεδομένα μπορούν να αποτελέσουν πολύτιμο πόρο για τις επιχειρήσεις λόγω των θεμάτων απορρήτου, των ταχύτερων κύκλων δοκιμών προϊόντων και της εκπαίδευσης των αλγορίθμων μηχανικής μάθησης. Πολλοί νόμοι για την ιδιωτικότητα των δεδομένων περιορίζουν τον τρόπο με τον οποίο οι επιχειρήσεις μπορούν να χειρίζονται ευαίσθητα δεδομένα.

Οποιαδήποτε διαρροή ή κοινοποίηση προσωπικά αναγνωρίσιμων πληροφοριών πελατών μπορεί να οδηγήσει σε δαπανηρές αγωγές που μπορούν επίσης να βλάψουν την εικόνα της επιχείρησης. Επομένως, η μείωση των ανησυχιών για την ιδιωτικότητα είναι ο κύριος λόγος για τον οποίο οι εταιρείες επενδύουν σε μεθόδους παραγωγής συνθετικών δεδομένων. Επίσης, αξίζει να σημειωθεί, σε κάποιους τομείς όπως ο στρατιωτικός και ο ιατρικός υπάρχει περίπτωση να μην δίνεται πρόσβαση εξαρχής στα δεδομένα λόγω της ιδωτικότητας οπότε και εκεί η δημιουργία και χρήση συνθετικών είναι αναπόφευκτη.

Για εντελώς νέα προϊόντα, τα πραγματικά δεδομένα συνήθως δεν είναι αρκετά ή διαθέσιμα οπότε και εκεί βλέπουμε την χρήση των συνθετικών. Επιπλέον, τα δεδομένα που έχουν δημιουργηθεί από ανθρώπους είναι μια δαπανηρή και χρονοβόρα διαδικασία. Αυτό μπορεί να αποφευχθεί εάν οι εταιρείες

επενδύουν σε συνθετικά δεδομένα, τα οποία μπορούν να παραχθούν γρήγορα και να βοηθήσουν στην ανάπτυξη αξιόπιστων μοντέλων μηχανικής μάθησης. [10]

2.2 Συνθετικά δεδομένα εικόνων

Τα συνθετικά δεδομένα εικόνας αναφέρονται σε εικόνες που δεν προέρχονται από πραγματικά γεγονότα, αλλά παράγονται τεχνητά. Διάφοροι αλγόριθμοι χρησιμοποιούνται για τη δημιουργία αυτών των δεδομένων, τα οποία στη συνέχεια χρησιμοποιούνται για την εκπαίδευση μοντέλων μηχανικής μάθησης.

Οι πρόσφατες μεθοδολογίες για τη δημιουργία συνθετικών δεδομένων εικόνας περιλαμβάνουν τη χρήση μοντέλων που παράγουν τεχνητά δεδομένα, όπως Diffusion models, τα Generative Adversarial Networks (GANs) και τα Variational Autoencoders (VAEs), τα οποία έχουν αποδείξει την ικανότητά τους να παράγουν ρεαλιστικές εικόνες. Ωστόσο, δεν ήταν παρά μόνο τα τελευταία χρόνια με την εμφάνιση των LLMs που το Generative AI άρχισε πραγματικά να ανθεί ώστε να δώσει μία ενοποιημένη λύση σχετικά με τα συνθετικά δεδομένα. Τα LLMs, που εκπαιδεύονται σε τεράστια σύνολα δεδομένων και παρουσιάζουν μια πρωτοφανή ικανότητα να παράγουν συνεκτικό και πληροφοριακά σχετικό κείμενο, ωθώντας με αυτό τον τρόπο τα όρια του τι μπορεί να επιτύχει η τεχνητή νοημοσύνη. Για τα LLMs θα μιλήσουμε αναλυτικά αργότερα σε αυτό το κεφάλαιο μιας και είναι ένα από τα πιο σημαντικά ερευνητικά θέματα του κλάδου αυτή την εποχή. [11],[12]

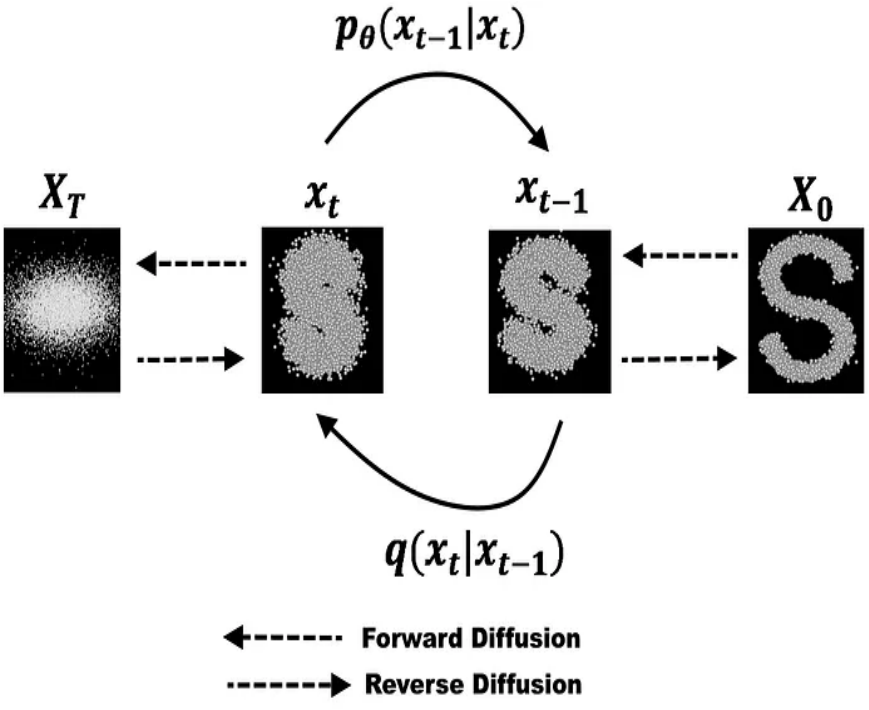
2.2.1 Diffusion models

Τα diffusion models, τα οποία εμπνέονται από τη μοντελοποίηση της θερμοδυναμικής, έχουν κάνει σημαντικές προόδους τα τελευταία χρόνια, υπερβαίνοντας τις δυνατότητες των προηγούμενων μοντέλων που προτάθηκαν, τα Generative Adversarial Networks (GANs) και τα Variational AutoEncoders (VAEs). Αυτά τα μοντέλα χρησιμοποιούνται εκτενώς σε τομείς όπως η υπολογιστική όραση και η παραγωγή ήχου, και έχουν βρει περαιτέρω εφαρμογές στη δημιουργία κειμένου, στη μοντελοποίηση δεδομένων ακολουθίας, στην ενισχυτική μάθηση και τον έλεγχο, και ακόμη και στις επιστήμες ζωής.

Η εξαιρετική απόδοση των diffusion models είναι ζωτικής σημασίας για μια σειρά μεθοδολογικών προόδων που διευρύνουν σημαντικά το πεδίο τους και ενισχύουν τη λειτουργικότητά τους. Αυτό περιλαμβάνει την παραγωγή υψηλής πιστότητας, την αποτελεσματική δειγματοληψία και τον ευέλικτο έλεγχο της παραγωγής δειγμάτων. Για παράδειγμα, τα diffusion models έχουν επεκταθεί για την παραγωγή διακριτών δεδομένων, ενώ τα αρχικά μοντέλα σχεδιάστηκαν για συνεχή δεδομένα. Ταυτόχρονα, υπάρχει μια ενεργή γραμμή έρευνας που στοχεύει στην επίταχυνση της ταχύτητας παραγωγής

δειγμάτων των diffusion models. Τέλος, αλλά όχι λιγότερο σημαντικό, ένας πρόσφατος κλάδος της έρευνας επικεντρώνεται στην λεπτομερή ρύθμιση των diffusion models προς την παραγωγή δειγμάτων με επιθυμητές ιδιότητες, όπως η παραγωγή εικόνων με ιδιαίτερες αισθητικές ποιότητες. Αυτές οι ειδικές για την εργασία ιδιότητες συχνά ενσωματώνονται ως καθοδήγηση στο diffusion model, αποτελούντας συνθήκες και σήματα ελέγχου για την κατεύθυνση της παραγωγής δειγμάτων. Είναι αξιοσημείωτο ότι η καθοδήγηση επιτρέπει τη δημιουργία διαφορετικού και σχετικού περιεχομένου σε μια ευρεία γκάμα εφαρμογών, τονίζοντας την πολυμορφία και την προσαρμοστικότητα των diffusion models. Ονομάζουμε τα diffusion models με καθοδήγηση ως conditional diffusion models.

Συνοπτικά, ένα diffusion model αποτελείται από μια διαδικασία προς τα εμπρός και μια διαδικασία προς τα πίσω. Κατά τη διαδικασία προς τα εμπρός, ένα καθαρό δείγμα από την κατανομή δεδομένων διαταράσσεται σταδιακά από τυχαίο Gaussian θόρυβο και, στο όριο του άπειρου χρόνου, η κατανομή δεδομένων μεταμορφώνεται σε καθαρό θόρυβο. Στη διαδικασία προς τα πίσω, εκπαιδεύεται ένα νευρωνικό δίκτυο αποθορυβοποίησης για να αφαιρέσει διαδοχικά την προστιθέμενη κατανομή θορύβου στα δεδομένα και να αποκαταστήσει μια νέα καθαρή κατανομή δεδομένων. Τα diffusion models έχουν επιτύχει κορυφαίες επιδόσεις σε εργασίες παραγωγής εικόνας & ήχου και αποτελούν ένα από τα βασικά στοιχεία των συστημάτων σύνθεσης εικόνας & ήχου, όπως το DALL-E, το stable diffusion και το Diffwave. [13]

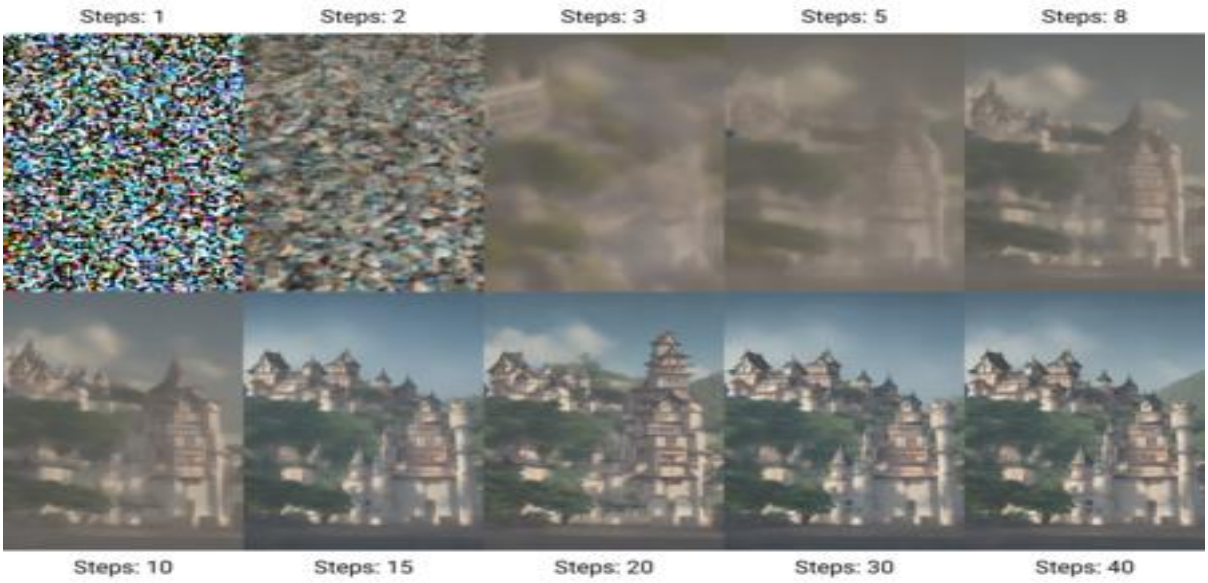


Εικόνα 2.1 : Διαδικασία αποθορυβοποίησης ενός diffusion model [45]

2.2.1.1 Stable diffusion model

Το Stable Diffusion είναι ένα μοντέλο βαθιάς μάθησης που επικεντρώνεται κυρίως στη μετατροπή κειμένου σε εικόνες. Κυκλοφόρησε το 2022 και θεωρείται σημαντικό μέρος της τρέχουσας επανάστασης που φέρνει στις μέρες μας η τεχνητή νοημοσύνη. Το μοντέλο είναι ικανό να δημιουργεί λεπτομερείς εικόνες βασισμένες σε περιγραφές κειμένου, και μπορεί επίσης να εφαρμοστεί σε άλλες εργασίες, όπως η επιδιόρθωση εικόνων, η επέκταση εικόνων και η δημιουργία μεταφράσεων εικόνας προς εικόνα με την καθοδήγηση ενός κειμένου. Η ανάπτυξη του Stable Diffusion περιλάμβανε ερευνητές από την Ομάδα CompVis στο Πανεπιστήμιο Ludwig Maximilian του Μονάχου και την Runway με μια υπολογιστική δωρεά από την Stability και δεδομένα εκπαίδευσης από μη κερδοσκοπικές οργανώσεις. Ο κώδικας του μοντέλου και τα βάρη του μοντέλου έχουν δημοσιευτεί δημόσια, και μπορεί να τρέξει στους περισσότερους σύγχρονους υπολογιστές εξοπλισμένους με μια μέτρια GPU με τουλάχιστον 4 GB VRAM και πάνω.

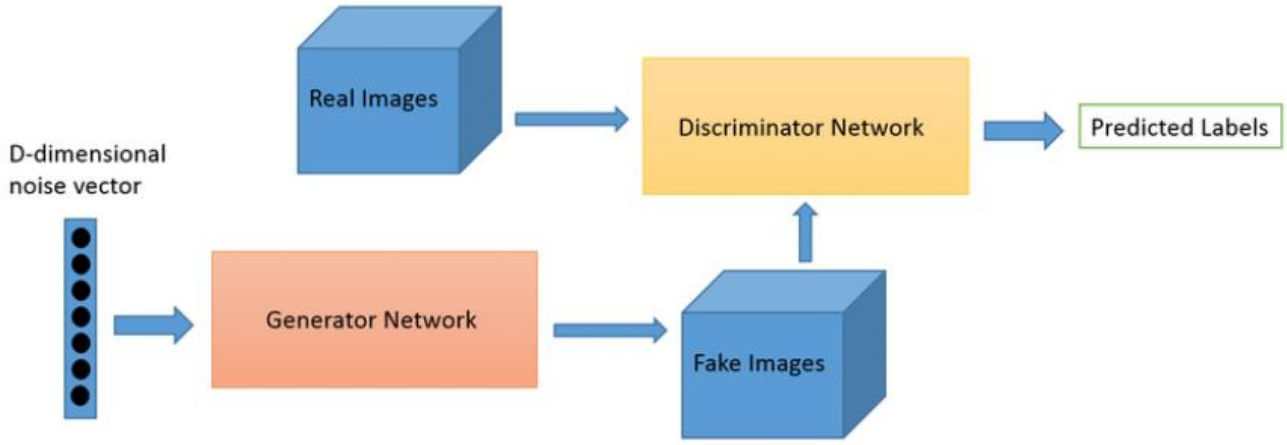
Στην παρούσα διατριβή το stable diffusion είναι αυτό που τελικά επιλέχθηκε για την υλοποίηση συνθετικών δεδομένων εικόνας που χρειαζόμαστε καθώς υπερείχε των άλλων τεχνικών. Η ανάλυση της επιλογής αυτής και τα αποτελέσματα θα αναπτυχθούν σε μεταγενέστερο κεφάλαιο.



Εικόνα 2.1 : Διαδικασία αποθορυβοποίησης ενός stable diffusion model

2.2.2 GANs

Τα Generative adversarial networks (GANs) είναι μια ισχυρή κατηγορία νευρωνικών δικτύων που ακολουθούν μία έξυπνη προσέγγιση της μη εποπτευόμενης μάθησης. Τα GANs μπορούν να δημιουργήσουν δείγματα που μοιάζουν στενά με την κατανομή των πραγματικών δεδομένων, εξερευνώντας την υποκείμενη δομή τους και μάθοντας από τα υπάρχοντα πρότυπα των πραγματικών δεδομένων. Το πλαίσιο των GAN υιοθετεί μια προοπτική παιχνιδιού, με δύο βασικούς παίκτες: τον generator και τον discriminator. Αυτά τα δίκτυα εμπλέκονται σε συνεχή αντιπαράθεση κατά τη διάρκεια της εκπαίδευσης, με τον generator να προσπαθεί να δημιουργήσει ρεαλιστικά δείγματα ενώ αποφεύγει την ανίχνευση από τον discriminator που προσπαθεί να προβλέπει ποια είναι τα αληθινά δεδομένα και ποια όχι. Τα δύο αυτά δίκτυα βελτιώνουν τα αποτελέσματα τους σε κάθε επόμενη διαδικασία μάθησης και θεωρούμε πως φτάνουμε σε ικανοποιητικά αποτελέσματα μόνο όταν ο generator πλέον μπορεί να "εξαπατήσει" τον discriminator και να θεωρήσει ότι τα δεδομένα που παράγει είναι αληθινά.



Εικόνα 2.2 : Γενική αρχιτεκτονική ενός GAN [45]

Πιο αναλυτικά σε μια τυπική αρχιτεκτονική όπως αναφέρθηκε και πριν ένα Generative adversarial network (GAN) αποτελείται από δύο τύπους δικτύων: τον discriminator (D) και τον generator (G).

Ο generator δημιουργεί εικόνες χρησιμοποιώντας τυχαίο θόρυβο (Z), και αυτές οι δημιουργημένες εικόνες σημειώνονται ως G(z). Ο τυχαίος θόρυβος είναι συνήθως Γκαουσιανός και αντιπροσωπεύει ένα τυχαίο σημείο στον κρυφό χώρο. Κατά τη διάρκεια της εκπαίδευσης του GAN, οι παράμετροι και των δύο δικτύων G και D ενημερώνονται επαναληπτικά.

Ο discriminator (D) λειτουργεί ως δίκτυο διάκρισης για να προσδιορίσει εάν μια δεδομένη εικόνα ανήκει σε μια πραγματική κατανομή. Παίρνει μια εικόνα εισόδου X και παράγει την έξοδο D(x), που αντιπροσωπεύει την πιθανότητα ότι το X ανήκει σε μια πραγματική κατανομή. Η τιμή εξόδου 1 υποδηλώνει μια πραγματική κατανομή εικόνας, ενώ η τιμή 0 υποδηλώνει μια ψεύτικη κατανομή εικόνας.

Υπάρχουν πολλές υποκατηγορίες των GANs όπως επίσης και ο συνδυασμός τους με άλλες τεχνικές, κάποιες από τις πιο γνωστές κατηγορίες είναι τα Conditional GANs, τα Deep convolutional GANs, το StyleGAN και το Adaptive GAN. [14]-[17]

2.2.3 VAEs

Στον τομέα της μηχανικής μάθησης, ένα variational autoencoder (VAE) είναι μια τεχνητή αρχιτεκτονική νευρωνικού δικτύου που αναπτύχθηκε από τους Diederik P. Kingma και Max Welling. Ανήκει στην κατηγορία των πιθανοτικών γραφικών μοντέλων και των μεθόδων του variational Bayesian. Εκτός από τον ρόλο του ως αρχιτεκτονική νευρωνικού δικτύου autoencoder, τα variational autoencoders μπορούν επίσης να μελετηθούν μέσα στο μαθηματικό πλαίσιο των variational Bayesian μεθόδων. Αυτό περιλαμβάνει τη σύνδεση ενός νευρωνικού δικτύου κωδικοποιητή με τον αποκωδικοποιητή του μέσω ενός πιθανοτικού χώρου κρυφής μεταβλητής, ο οποίος μπορεί να περιγραφεί ως μια πολυδιάστατη κανονική κατανομή. Αυτός ο χώρος κρυφής μεταβλητής αντιστοιχεί στις παραμέτρους μιας variational κατανομής.

Η βασική αξία των Variational Autoencoders (VAEs) βρίσκεται στην αρχιτεκτονική τους, που τους διαφοροποιεί από τους παραδοσιακούς autoencoders. Λειτουργούν γενικά ως εξής:

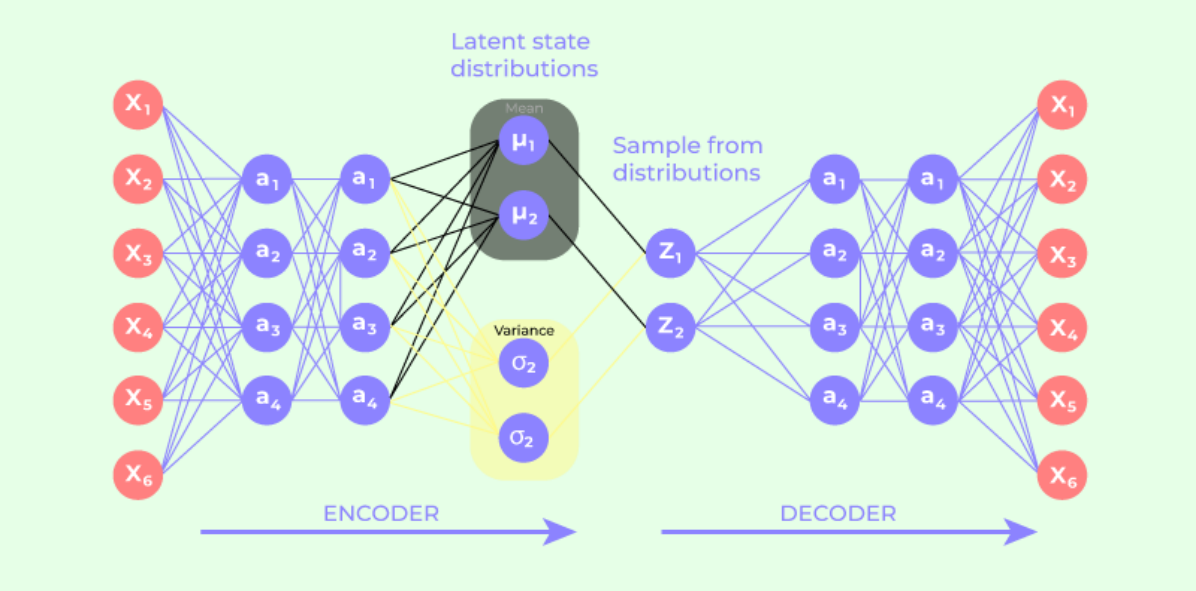
- Το δίκτυο κωδικοποιητή παίρνει ακατέργαστα δεδομένα εισόδου και τα μετατρέπει σε μια κατανομή πιθανότητας στον χώρο κρυφής μεταβλητής.

- Ο κρυφός κώδικας, που παράγεται από τον κωδικοποιητή, αντιπροσωπεύει μια πιθανοτική κωδικοποίηση. Αυτό επιτρέπει στο VAE να εκφράσει όχι μόνο ένα μοναδικό σημείο, αλλά μια κατανομή πιθανών αναπαραστάσεων στον χώρο κρυφής μεταβλητής.

- Το δίκτυο αποκωδικοποιητή ανακατασκευάζει τα δεδομένα από ένα δείγμα στην κρυφή κατανομή πίσω στον αρχικό χώρο δεδομένων.

- Κατά τη διάρκεια της εκπαίδευσης, το μοντέλο ρυθμίζει τόσο τις παραμέτρους του κωδικοποιητή όσο και του αποκωδικοποιητή για να ελαχιστοποιήσει την απώλεια ανακατασκευής, η οποία μετρά τη διαφορά μεταξύ των εισαγόμενων δεδομένων και της αποκωδικοποιημένης έξοδου.

- Ο στόχος δεν είναι μόνο η ακριβής ανακατασκευή, αλλά και η ρύθμιση του κρυφού χώρου, εξασφαλίζοντας ότι διαμορφώνεται με μια συγκεκριμένη κατανομή. [18]



Εικόνα 2.3 : Παράδειγμα δομής ενός Variational Autoencoder [33]

2.3 Συνθετικά δεδομένα συμβολοσειρών

Μετά την ανάλυση για τα συνθετικά δεδομένα εικόνων και τις τεχνικές δημιουργίας τους σε αυτό το σημείο γίνεται αναφορά και στην άλλη κατηγορία που είναι τα συνθετικά δεδομένα συμβολοσειρών. Τα δεδομένα συμβολοσειρών όπως είναι αυτά του GPS ή σήματα ραντάρ, χρησιμοποιούνται συχνά τεχνικές όπως η προσομοίωση και η παρεμβολή. Η προσομοίωση περιλαμβάνει τη δημιουργία δεδομένων με βάση γνωστά μοντέλα ή νόμους, όπως οι φυσικές ιδιότητες των σημάτων GPS. Η παρεμβολή, από την άλλη πλευρά, περιλαμβάνει τη δημιουργία νέων σημείων δεδομένων εντός του εύρους των υπαρχόντων σημείων δεδομένων, τα οποία μπορεί να είναι χρήσιμα για την κάλυψη κενών στα δεδομένα GPS ή ραντάρ. Άλλες τεχνικές περιλαμβάνουν μοντέλα μηχανικής εκμάθησης που έχουν εκπαιδευτεί σε δεδομένα πραγματικού κόσμου, τα οποία μπορούν στη συνέχεια να δημιουργήσουν νέα συνθετικά δεδομένα. Αυτές οι μέθοδοι συμβάλλουν στη διατήρηση του απορρήτου, αυξάνουν τη διαθεσιμότητα των δεδομένων και επιτρέπουν διαφορετικά και ολοκληρωμένα σενάρια δοκιμών. Για κάποιες εφαρμογές η δημιουργία τέτοιων δεδομένων μπορεί να είναι πιο απλή ενώ σε άλλες μπορεί να χρειάζεται βαθύτερη γνώση και κατανόηση της θεματολογίας από τον ερευνητή ώστε να υπάρξει το επιθυμητό αποτέλεσμα. Στην παρούσα διατριβή για λόγους testing χρειάστηκε η δημιουργία τέτοιων συνθετικών συμβολοσειρών και συγκεκριμένα GPS για τη χρήση τους σε προσομοιωτή για να δούμε την συμπεριφορά του μοντέλου spoofing detection του drone σε νέα άγνωστα σήματα. Αξίζει να σημειωθεί ότι για όχι τόσο εξειδικευμένες συμβολοσειρές και πιο απλά προβλήματα που χρειάζονται συνθετικά text δεδομένα μπορεί να γίνει χρήση των LLMs που σε δευτερόλεπτα μπορούν να παράγουν υψηλής ποιότητας data.

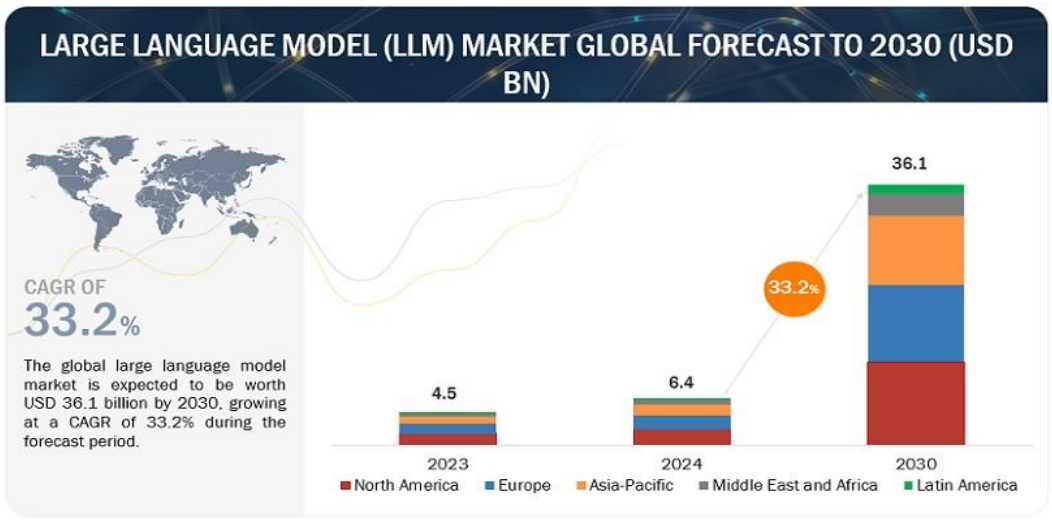
2.3.1 LLMs

Όπως αναφέρθηκε και πριν η χρήση των LLMs για την δημιουργία συνθετικών δεδομένων σε πιο απλά προβλήματα είναι πλέον η πιο συνηθισμένη τεχνική και θα καθιερωθεί όσο διεισδύει η τεχνητή νοημοσύνη στην καθημερινότητα μας.

Γενικότερα, τα Large Language Models (LLMs) ανήκουν στο κλάδο του generative AI και είναι προηγμένα συστήματα τεχνητής νοημοσύνης που σχεδιάστηκαν για να δημιουργούν κείμενο παρόμοιο με αυτό που λαμβάνουν ως είσοδο. Αυτά τα μοντέλα εκπαιδεύονται σε μεγάλα σύνολα δεδομένων, επιτρέποντάς τους να μάθουν πρότυπα, δομές και πληροφορίες που περιέχονται στο κείμενο. Ως αποτέλεσμα, τα LLMs μπορούν να δημιουργήσουν συνεκτικές και συναφείς περιεχομένου απαντήσεις, καθιστώντας τα χρήσιμα εργαλεία για διάφορες εφαρμογές, όπως η υποστήριξη πελατών, η εκπαίδευση και η δημιουργία περιεχομένου & δεδομένων.

Η αρχιτεκτονική ενός Large Language Model (LLM) συνήθως περιλαμβάνει πολλά επίπεδα συνδεδεμένων νευρωνικών δικτύων που επεξεργάζονται και δημιουργούν κείμενο, αναλύοντας πρότυπα σε μεγάλα σύνολα δεδομένων. Επιπλέον η αρχιτεκτονική τους περιλαμβάνει στοιχεία όπως input embeddings, transformer layers, και output layers, επιτρέποντας στο μοντέλο να μάθει πολύπλοκα γλωσσικά πρότυπα και σχέσεις. Ο self-attention μηχανισμός του LLM διευκολύνει την κατανόηση του πλαισίου κάθε λέξης στην είσοδο, καθιστώντας το υψηλά αποτελεσματικό σε διάφορες εργασίες επεξεργασίας φυσικής γλώσσας. Η σχεδίαση της αρχιτεκτονικής ενός LLM αντικατοπτρίζει έναν συμβιβασμό μεταξύ υπολογιστικής αποδοτικότητας και απόδοσης του μοντέλου.

Τα πιο γνωστά LLMs προέρχονται από την εταιρεία OpenAI και είναι το GPT-3, το GPT-4 και το ChatGPT. Αυτά τα μοντέλα έχουν καταπλήξει τους ερευνητές καθώς και γενικότερα την ανθρωπότητα με τις δυνατότητες τους και για τις μελλοντικές τους επεκτάσεις. [19]-[21]



Εικόνα 2.4 : Πρόβλεψη για το μέγεθος του market των LLMs μέχρι το 2030 [40]

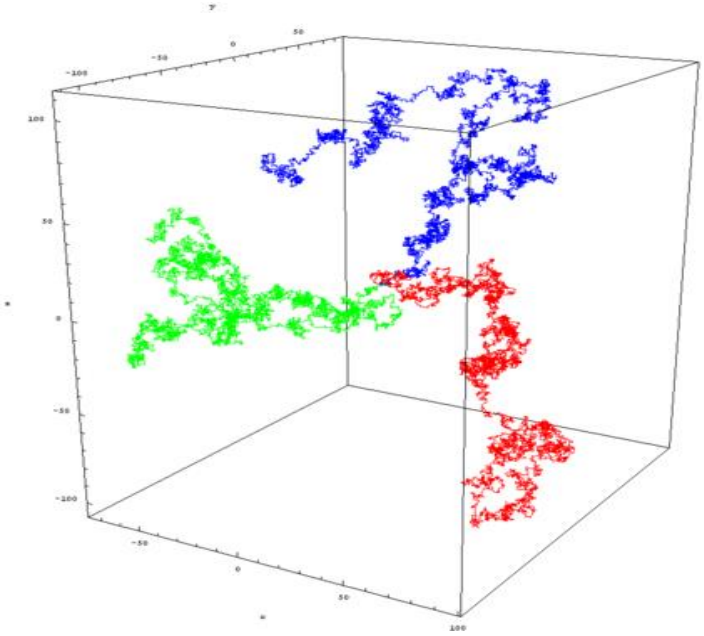
Λόγω των προαναφερθέντων πλέον οι ερευνητές έχουν αρχίσει να εξερευνούν τη δυνατότητα χρήσης των LLMs για τη δημιουργία συνθετικών δεδομένων που προσαρμόζονται σε κάποια συγκεκριμένη θεματολογία εκπαίδευσης μοντέλων για διάφορες εργασίες ταξινόμησης κειμένου. Ένα σημαντικό σημείο για την σωστή δημιουργία συνθετικών δεδομένων με LLMs είναι ο σαφής καθορισμός και η λεπτομέρεια του input text το οποίο θα δοθεί από τον χρήστη. Τα μειονεκτήματα που παρατηρούνται λόγο αυτού του ανθρώπινου καθορισμού που εκφράζεται διαφορετικά από τον καθένα λόγω της διαφορετικότητας του σκεπτικού, είναι πως τα αποτελέσματα μπορεί να διαφέρουν πολύ ακόμα και αν επεξηγείτε ένα πρόβλημα ίδιας θεματολογία. Έχει παρατηρηθεί ότι ακόμα και με ένα ολόιδιο input τα αποτελέσματα υπάρχει πιθανότητα να είναι πολύ διαφορετικά κάτι που μας δείχνει ότι είναι inconsistent ως τεχνική. Επιπλέον, προφανώς σε θέματα τα οποία δεν διακατέχονται από κάποια αντικειμενικότητα και η απάντηση είναι υποκειμενική, η ποιότητα των δεδομένων που θα παραχθούν πολύ πιθανών να μην έχει κάποια αξία. Παρόλα αυτά, όπως αναφέρθηκε σε απλά προβλήματα δημιουργίας συνθετικών δεδομένων κειμένου/συμβολοσειρών τα αποτελέσματα των LLMs είναι πολύ χρήσιμα και ο προγραμματιστής ή ο οποιοσδήποτε χρήστης πλέον δεν χρειάζεται να αναπτύξει κάποιον μακροσκελή κώδικα για να φτιάξει ένα ολοκληρωμένο dataset όπως γινόταν παλαιότερα.

2.3.2 Random Walk

Για πιο σύνθετα προβλήματα παραγωγής συνθετικών δεδομένων συμβολοσειρών το πιο σημαντικό είναι η βαθιά κατανόηση και γνώση γύρω από το αντικείμενο (π.χ παραγωγή συμβολοσειρών GPS, radar) και σε δεύτερο χρόνο οι αλγόριθμοι και οι τεχνολογίες που θα χρησιμοποιηθούν. Ένα αλγόριθμός που βοηθά στην παραγωγή δεδομένων χρησιμοποιώντας όρια και λογική είναι random walk.

Πιο συγκεκριμένα, τυχαίος περίπατος (random walk) στη θεωρία πιθανοτήτων είναι μια διαδικασία που καθορίζει την πιθανή θέση ενός σημείου υπό τυχαίες κινήσεις, δεδομένων των πιθανοτήτων (ίδιες σε κάθε βήμα) να μετακινηθεί κάποια απόσταση σε κάποια κατεύθυνση. Οι τυχαίοι περίπατοι αποτελούν παράδειγμα των διαδικασιών Markov, στις οποίες τη μελλοντική συμπεριφορά επηρεάζει μόνο το παρελθόν. Ένα τυπικό παράδειγμα είναι ο περίπατος του μεθυσμένου, όπου ένα σημείο που ξεκινά από την αρχή του Ευκλείδειου επιπέδου μετακινείται σε απόσταση μιας μονάδας για κάθε στιγμή του χρόνου, με την κατεύθυνση της κίνησης να είναι τυχαία σε κάθε βήμα. Το πρόβλημα είναι να βρεθεί, μετά από ένα σταθερό χρόνο, η συνάρτηση κατανομής των πιθανοτήτων της απόστασης του σημείου από την αρχή. Πολλοί οικονομολόγοι πιστεύουν ότι οι διακυμάνσεις του χρηματιστηρίου, τουλάχιστον για ένα σύντομο χρονικό διάστημα, είναι τυχαίοι περίπατοι.

Ο πιο συνηθισμένος αλγόριθμος βασισμένος σε τυχαίους περιπάτους στην επιστήμη των υπολογιστών είναι ο PageRank. Ο PageRank υπολογίζει τη σημασία των ιστοσελίδων περπατώντας τυχαία από αυτές. Οι ερευνητές έχουν δημιουργήσει επίσης αρκετές παραλλαγές του τυχαίου περιπάτου, όπως ο εξατομικευμένος PageRank [22], ο random walk with restart (RWR) [23] και ο lazy random walk (LRW) [24].



Εικόνα 2.5 : Απεικόνιση τριών random walk σε τρεις διαστάσεις

Μια άλλη κατηγορία είναι οι κβαντικοί περίπατοι που προτάθηκαν αρχικά από τον Aharonov το 1993 [25]. Μπορούν να θεωρηθούν ως ο κβαντικός αντίστοιχος των κλασικών τυχαίων περιπάτων που συναντάμε στην κβαντική μηχανική. Η κύρια διαφορά μεταξύ των κλασικών τυχαίων περιπάτων και των κβαντικών περιπάτων είναι ότι οι κβαντικοί περίπατοι δεν συγκλίνουν σε κάποιες περιορισμένες κατανομές. Επιπλέον, λόγω της κβαντικής παρεμβολής, μπορούν να εξαπλωθούν σημαντικά πιο γρήγορα ή πιο αργά από τους κλασικούς τυχαίους περιπάτους. Σε ό,τι αφορά την χρονική πολυπλοκότητα, οι αλγόριθμοι βασισμένοι σε κβαντικούς περιπάτους έχουν χαμηλότερη πολυπλοκότητα σε σύγκριση με τους αλγόριθμους βασισμένους σε κλασικούς τυχαίους περιπάτους.

Όσον αφορά για την θεματολογία που ενδιαφερόμαστε, δηλαδή τη χρήση του random walk για δημιουργία συνθετικών δεδομένων βλέπουμε την εφαρμογή του στους εξής τομείς:

Επέκταση Δεδομένων

- Οι τυχαίοι περίπατοι μπορούν να χρησιμοποιηθούν για να επεκτείνουν υπάρχοντα σύνολα δεδομένων δημιουργώντας επιπλέον δείγματα. Για παράδειγμα, στα δεδομένα εικόνας, οι τυχαίοι περίπατοι μπορούν να προσομοιώσουν ποικίλες διακυμάνσεις στις θέσεις, τις προσανατολίσεις ή τις συνθήκες φωτισμού των αντικειμένων.

- Με την εισαγωγή ελεγχόμενης τυχαιότητας, τα συνθετικά δεδομένα που παράγονται μέσω τυχαίων περιπάτων βοηθούν στη βελτίωση της ανθεκτικότητας και της γενίκευσης των μοντέλων μηχανικής μάθησης.

Προσομοίωση και Δοκιμή

Σε πεδία όπως η υπολογιστική γραφική, η φυσική και η επιδημιολογία, οι τυχαίοι περίπατοι χρησιμοποιούνται για την προσομοίωση πολύπλοκων διεργασιών. Για παράδειγμα:

- Κίνηση Σωματιδίων: Οι τυχαίοι περίπατοι προσομοιώνουν την κίνηση σωματιδίων σε ένα μέσο, όπως η διάχυση ή η κίνηση Brownian.

- Χρηματιστηριακές Αγορές: Τα συνθετικά δεδομένα τιμών μετοχών μπορούν να παραχθούν χρησιμοποιώντας τυχαίους περιπάτους για τη δοκιμή στρατηγικών συναλλαγών ή αλγορίθμων διαχείρισης κινδύνου.

- Επιδημιολογία: Οι τυχαίοι περίπατοι προσομοιώνουν την εξάπλωση της νόσου, βοηθώντας στην κατανόηση των εστιών και στην αξιολόγηση των στρατηγικών παρέμβασης.

Τεχνικές Διατήρησης Ιδιωτικότητας

- Τα συνθετικά δεδομένα που παράγονται από τυχαίους περιπάτους μπορούν να χρησιμοποιηθούν για την προστασία της ιδιωτικότητας. Αντί να κοινοποιούν ευαίσθητα πραγματικά δεδομένα, οι οργανισμοί μπορούν να δημοσιεύουν συνθετικά δεδομένα που διατηρούν τις στατιστικές ιδιότητες.

- Οι τεχνικές διαφορικής ιδιωτικότητας συχνά περιλαμβάνουν την προσθήκη θορύβου μέσω τυχαίων περιπάτων για να εξασφαλίσουν την ιδιωτικότητα ενώ διατηρούν την χρησιμότητα των δεδομένων.

Benchmarking και Αξιολόγηση

- Οι ερευνητές χρησιμοποιούν συνθετικά δεδομένα για το benchmarking αλγορίθμων και την αξιολόγηση της απόδοσής τους. Οι τυχαίοι περίπατοι επιτρέπουν τη δημιουργία ποικίλων συνόλων δεδομένων με γνωστές ιδιότητες.
- Για παράδειγμα, στη δοκιμή αλγορίθμων γράφων, τα συνθετικά γραφήματα που παράγονται μέσω τυχαίων περιπάτων παρέχουν τη βάση σύγκρισης.

Διερευνητική Ανάλυση

- Τα συνθετικά δεδομένα που παράγονται από τυχαίους περιπάτους μπορούν να χρησιμοποιηθούν για διερευνητική ανάλυση. Οι ερευνητές μπορούν να εξερευνήσουν διαφορετικά σενάρια, να αποκαλύψουν μοτίβα και να αποκτήσουν γνώσεις χωρίς να βασίζονται αποκλειστικά σε δεδομένα πραγματικού κόσμου. [26]

ΚΕΦΑΛΑΙΟ 3

ΥΛΟΠΟΙΗΣΗ

3.1 Δημιουργία συνθετικών εικόνων με stable diffusion

Η πρώτη υλοποίηση που εφαρμόστηκε αφορούσε τη δημιουργία συνθετικών εικόνων, οι οποίες προσομοίαζαν λήψεις στιγμιότυπων από UAV (Unmanned Aerial Vehicles) σε συγκεκριμένο υψόμετρο και περιοχή. Αυτή η διαδικασία περιλάμβανε διάφορα στάδια και τεχνικές για την επίτευξη ρεαλιστικών και χρήσιμων αποτελεσμάτων. Αρχικά, επιλέχθηκε η περιοχή ενδιαφέροντος, λαμβάνοντας υπόψη τις γεωγραφικές και περιβαλλοντικές συνθήκες που θα επηρέαζαν τις λήψεις από τα UAV. Στη συνέχεια, καθορίστηκε το υψόμετρο από το οποίο θα γίνονταν οι λήψεις, ώστε να εξασφαλιστεί η βέλτιστη ανάλυση και κάλυψη της περιοχής. Η ανάγκη για τη δημιουργία συνθετικών δεδομένων σε αυτό το παράδειγμα είναι ιδιαίτερα σημαντική, καθώς αντιμετωπίζει ένα κλασικό πρόβλημα που προκύπτει όταν τα UAVs χάνουν την επικοινωνία τους με το GPS. Σε τέτοιες περιπτώσεις, η πλοήγηση των UAVs μπορεί να διαταραχθεί σοβαρά, οδηγώντας σε απώλεια ελέγχου και πιθανές επικίνδυνες καταστάσεις. Για να αντιμετωπιστεί αυτό το πρόβλημα, δημιουργήθηκαν συνθετικές εικόνες που θα μπορούσαν να χρησιμοποιηθούν ως εναλλακτική πηγή δεδομένων πλοήγησης. Αυτές οι εικόνες δημιουργήθηκαν με τη χρήση προηγμένων αλγορίθμων και τεχνικών επεξεργασίας εικόνας, που προσομοίωναν τις πραγματικές συνθήκες λήψης από τα UAVs. Οι εικόνες αυτές ενσωματώθηκαν σε ένα συνθετικό χάρτη, ο οποίος θα μπορούσε να χρησιμοποιηθεί σε περίπτωση απώλειας του σήματος GPS. Ο σκοπός αυτής της υλοποίησης είναι να υπάρχει πάντα διαθέσιμος ένας συνθετικός χάρτης, ο οποίος θα επιτρέπει τη συνέχιση της πλοήγησης των UAVs χωρίς προβλήματα, ακόμα και σε περιπτώσεις έκτακτης ανάγκης. Με αυτόν τον τρόπο, εξασφαλίζεται η ασφάλεια και η αποτελεσματικότητα των αποστολών των UAVs, ανεξάρτητα από τις εξωτερικές συνθήκες και τις πιθανές απώλειες επικοινωνίας.

Στην αρχή, εξετάστηκε η πιθανότητα δημιουργίας και εκμάθησης ενός μοντέλου GAN (Generative Adversarial Network) από την αρχή. Αυτή η διαδικασία περιλάμβανε την ανάπτυξη ενός νέου μοντέλου που θα μπορούσε να παράγει συνθετικές εικόνες υψηλής ποιότητας. Ωστόσο, γρήγορα διαπιστώθηκε ότι η εκπαίδευση τέτοιων μοντέλων απαιτεί τεράστιο όγκο δεδομένων και σημαντική υπολογιστική ισχύ. Συγκεκριμένα, για να επιτευχθεί ένα ικανοποιητικό αποτέλεσμα, θα χρειαζόταν να συλλεχθούν και να επεξεργαστούν εκατοντάδες χιλιάδες εικόνες, καθώς και να χρησιμοποιηθούν ισχυροί υπολογιστές με εξειδικευμένο υλικό. Λόγω αυτών των απαιτήσεων, αποφασίστηκε να χρησιμοποιηθεί ένα προεκπαιδευμένο μοντέλο stable diffusion από την Huggin Face. Η Huggin Face είναι μια γαλλοαμερικάνικη εταιρεία που παρέχει μοντέλα τεχνητής νοημοσύνης έτοιμα για χρήση, καθώς και για transfer learning. Το συγκεκριμένο μοντέλο stable diffusion επιλέχθηκε επειδή είχε ήδη εκπαιδευτεί σε μεγάλο όγκο δεδομένων και μπορούσε να παράγει συνθετικές εικόνες υψηλής ποιότητας με λιγότερη υπολογιστική ισχύ.

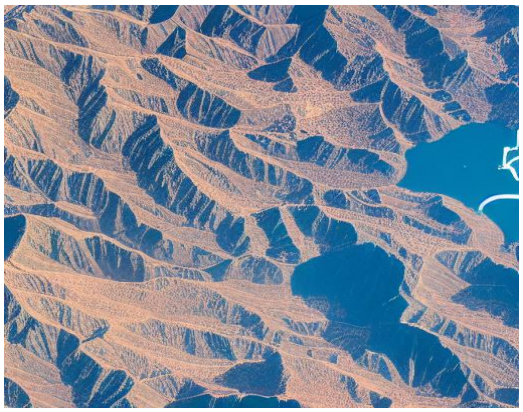
Αξίζει να σημειωθεί ότι, ακόμα και με τη χρήση αυτής της τεχνικής, η οποία θεωρήθηκε η βέλτιστη, ο χρόνος εκτέλεσης για τη δημιουργία πολλαπλών συνθετικών εικόνων ήταν σχετικά μεγάλος. Αυτό οφείλεται στην πολυπλοκότητα των αλγορίθμων και την ανάγκη για επεξεργασία μεγάλου όγκου δεδομένων σε κάθε εκτέλεση. Παρόλο αυτά, η χρήση του προεκπαιδευμένου μοντέλου από την Huggin Face επέτρεψε την επίτευξη των επιθυμητών αποτελεσμάτων με λιγότερους πόρους και σε μικρότερο χρονικό διάστημα από ό,τι θα απαιτούσε η εκπαίδευση ενός νέου μοντέλου από την αρχή.

Όπως γνωρίζουμε τα stable diffusion μοντέλα χρειάζονται ένα σαφές text input καθώς με βάση αυτό δημιουργούν τις συνθετικές εικόνες. Παρακάτω δίνετε ένα input που τροφοδοτήθηκε στο μοντέλο και τα αποτελέσματα που είχαμε:

Text input = {Images taken from a UAV at 3,000 meters altitude, showing a region of Crete from a top-down perspective.}

Αποτελέσματα:

Εικόνα 3.1



Εικόνα 3.2

Εικόνα 3.3

Όπως βλέπουμε με το input που επιλέξαμε έχουμε κάποια αρκετά ρεαλιστικά αποτελέσματα τα οποία μπορούν να ενταχθούν σε ένα συνθετικό map. Παρόλο αυτά είναι καλό για ένα καλύτερο feedback να δούμε την αξία των αποτελεσμάτων μέσα από τη μετρική που θα αναφερθεί παρακάτω. Επίσης, με διαφορετική είσοδο και με ανάλυση πολλών λεπτομερειών μπορούμε να λάβουμε πολύ διαφορετικά αποτελέσματα για την ίδια θεματολογία κάτι που δείχνει γιατί το stable diffusion υπερέχει αισθητά έναντι άλλων μοντέλων.

Μετρική CLIP Score

Η μετρική **CLIP score** είναι ένα εργαλείο που χρησιμοποιείται για την αξιολόγηση της ποιότητας των εικόνων που δημιουργούνται από μοντέλα τεχνητής νοημοσύνης, όπως τα μοντέλα διάχυσης (diffusion models). Η μετρική αυτή βασίζεται στο μοντέλο CLIP (Contrastive Language-Image Pretraining), το οποίο αναπτύχθηκε από την OpenAI και εκπαιδεύτηκε σε ένα μεγάλο σύνολο δεδομένων εικόνων και κειμένων.

Το CLIP μοντέλο έχει την ικανότητα να αντιστοιχίζει εικόνες με περιγραφές κειμένου, επιτρέποντας την αξιολόγηση της ποιότητας μιας εικόνας με βάση το πόσο καλά ταιριάζει με μια δεδομένη περιγραφή. Η μετρική CLIP score υπολογίζεται ως η πιθανότητα που αποδίδει το μοντέλο στην αντιστοιχία μεταξύ της εικόνας και της περιγραφής της.

Στην παρούσα διατριβή, ενδεικτικά αναφέρονται τα CLIP score των παραπάνω τριών συνθετικών εικόνων (3.1, 3.2, 3.3). Τα αποτελέσματα έδειξαν ότι όλες οι εικόνες είχαν CLIP score ίσο με 1.0. Αυτό σημαίνει ότι οι εικόνες αυτές ταιριάζουν απόλυτα με τις περιγραφές τους, υποδεικνύοντας ότι το μοντέλο διάχυσης που χρησιμοποιήθηκε για τη δημιουργία των εικόνων είναι εξαιρετικά αποτελεσματικό στην παραγωγή εικόνων που αντιστοιχούν ακριβώς στις περιγραφές τους. Η επίτευξη CLIP score 1.0 είναι ιδιαίτερα σημαντική, καθώς υποδηλώνει ότι το μοντέλο έχει μάθει να δημιουργεί εικόνες με υψηλή ακρίβεια και συνέπεια

3.2 Τροποποίηση εικόνων με stable diffusion

Ένας εναλλακτικός τρόπος για τη δημιουργία συνθετικών εικόνων είναι η τροποποίηση των πραγματικών εικόνων που μπορεί να έχουμε σε ένα dataset. Αυτή η μέθοδος μπορεί να εφαρμοστεί και σε συνθετικές εικόνες που ήδη διαθέτουμε, υπό ορισμένες προϋποθέσεις. Στην περίπτωσή μας, για τη δημιουργία ενός συνθετικού χάρτη πλοήγησης, είναι σημαντικό να μπορούμε να δημιουργήσουμε εικόνες που απεικονίζουν πώς θα έμοιαζε ένα τοπίο σε διαφορετικές εποχές του χρόνου. Για παράδειγμα, αν έχουμε μια εικόνα ενός τοπίου κατά τη διάρκεια του καλοκαιριού, μπορούμε να την τροποποιήσουμε ώστε να δείχνει πώς θα έμοιαζε το ίδιο τοπίο το χειμώνα, με χιόνι και διαφορετικό φωτισμό. Αυτή η διαδικασία μας επιτρέπει να επεκτείνουμε το υπάρχον σετ δεδομένων μας με νέες,

συνθετικές εικόνες που προσομοιώνουν διαφορετικές συνθήκες. Για να επιτευχθεί αυτό, χρησιμοποιούμε ένα stable diffusion model, και συγκεκριμένα ένα μοντέλο Img2Img. Το Img2Img είναι μια τεχνική που επιτρέπει τη μετατροπή μιας εικόνας σε μια άλλη, διατηρώντας τη βασική δομή της αρχικής εικόνας αλλά τροποποιώντας συγκεκριμένα χαρακτηριστικά της. Η κύρια διαφορά του Img2Img από άλλα μοντέλα είναι ο μηχανισμός αποθορυβοποίησης που χρησιμοποιεί, ο οποίος επιτρέπει την παραγωγή καθαρών και ρεαλιστικών συνθετικών εικόνων. Με τη χρήση του μοντέλου Img2Img, μπορούμε να δημιουργήσουμε συνθετικές εικόνες που επεκτείνουν το υπάρχον σετ δεδομένων μας, παρέχοντας έτσι περισσότερες επιλογές για την πλοήγηση των UAVs σε διαφορετικές συνθήκες και εποχές. Αυτός ο τρόπος δημιουργίας συνθετικών εικόνων είναι ιδιαίτερα χρήσιμος όταν οι πραγματικές εικόνες είναι περιορισμένες ή όταν θέλουμε να προσομοιώσουμε συνθήκες που δεν είναι εύκολα διαθέσιμες στο πραγματικό περιβάλλον [27].

Για την υλοποίηση αυτής της μεθόδου, επιλέχθηκε το προεκπαιδευμένο μοντέλο Img2Img stable diffusion από την Huggin Face. Αυτό το μοντέλο χρησιμοποιήθηκε ως βάση για μεταφορά εκπαίδευσης (transfer learning) με εικόνες που έχουν τραβηχτεί από δορυφόρους σε διάφορες εποχές του χρόνου. Η επιλογή του συγκεκριμένου μοντέλου έγινε λόγω της ικανότητάς του να μετατρέπει εικόνες διατηρώντας τη βασική τους δομή, ενώ τροποποιεί συγκεκριμένα χαρακτηριστικά τους. Ο σκοπός του transfer learning με αυτές τις δορυφορικές εικόνες είναι να επιτευχθεί καλύτερη προσαρμογή και πιο ακριβή αποτελέσματα στη θεματολογία του προβλήματός μας. Στην προκειμένη περίπτωση, το πρόβλημα αφορά εικόνες που έχουν τραβηχτεί από δορυφόρους ή UAVs (Unmanned Aerial Vehicles) από ψηλά (top-down) σε διαφορετικές εποχές του χρόνου. Με τη χρήση του transfer learning, το μοντέλο μπορεί να προσαρμοστεί καλύτερα στις συγκεκριμένες συνθήκες και να παράγει πιο ρεαλιστικές και χρήσιμες συνθετικές εικόνες. Το transfer learning είναι πάντα προτεινόμενο σε τέτοιες περιπτώσεις, καθώς τα προεκπαιδευμένα μοντέλα είναι συνήθως εκπαιδευμένα σε γενικής φύσεως δεδομένα. Αυτό σημαίνει ότι, χωρίς περαιτέρω εκπαίδευση, τα αποτελέσματα μπορεί να είναι μέτρια και να μην ανταποκρίνονται πλήρως στις συγκεκριμένες απαιτήσεις της θεματολογίας μας. Με την εφαρμογή του transfer learning, το μοντέλο προσαρμόζεται καλύτερα στα ειδικά δεδομένα που του παρέχονται, βελτιώνοντας έτσι την απόδοσή του και τα αποτελέσματα που παράγει.

Συνοψίζοντας, η χρήση του προεκπαιδευμένου μοντέλου Img2Img stable diffusion από την Huggin Face, σε συνδυασμό με το transfer learning με ίδιες δορυφορικές εικόνες σε διαφορετικές εποχές, επιτρέπει την παραγωγή συνθετικών εικόνων υψηλής ποιότητας που ανταποκρίνονται στις συγκεκριμένες ανάγκες του προβλήματός μας. Αυτός ο συνδυασμός τεχνικών εξασφαλίζει ότι οι συνθετικές εικόνες θα είναι ρεαλιστικές και χρήσιμες για την πλοήγηση των UAVs σε διαφορετικές εποχές του χρόνου.

Εικόνα 3.4 : Παράδειγμα εικόνας σε διαφορετικές εποχές που χρησιμοποιήθηκε στο transfer learning

Στη συνέχεια, όπως συμβαίνει με κάθε μοντέλο stable diffusion, έχουμε ένα text input, το οποίο, στην περίπτωση του Img2Img, αφορά την τροποποίηση που θέλουμε να γίνει στην εικόνα μας. Είναι σημαντικό το input αυτό να είναι σαφές και να περιγράφει με λεπτομέρεια το επιθυμητό αποτέλεσμα της τροποποίησης. Η ακρίβεια και η λεπτομέρεια στο text input είναι κρίσιμες για την επιτυχία της διαδικασίας, καθώς καθορίζουν τον τρόπο με τον οποίο το μοντέλο θα επεξεργαστεί την εικόνα. Εκτός από το text input, το μοντέλο μας δέχεται και μια εικόνα ως input, την οποία έχουμε επιλέξει να τροποποιηθεί. Η εικόνα αυτή αποτελεί τη βάση πάνω στην οποία θα εφαρμοστούν οι αλλαγές που περιγράφονται στο text input. Η διαδικασία αυτή επιτρέπει τη δημιουργία συνθετικών εικόνων που ανταποκρίνονται στις συγκεκριμένες απαιτήσεις μας.

Σε περίπτωση που το αποτέλεσμα της τροποποίησης δεν είναι ικανοποιητικό, πριν προβούμε σε κάποια δραστική ενέργεια, είναι καλό να δοκιμάσουμε να αλλάξουμε ή να παραφράσουμε το κείμενο που έχουμε δώσει ως είσοδο. Μικρές συντακτικές ή νοηματικές αλλαγές στο text input μπορούν να οδηγήσουν σε αρκετά διαφορετικά αποτελέσματα. Αυτό συμβαίνει επειδή το μοντέλο επεξεργάζεται το text input με συγκεκριμένο τρόπο και ακόμα και μικρές αλλαγές μπορούν να επηρεάσουν σημαντικά την τελική εικόνα. Για παράδειγμα, αν το αρχικό text input δεν παράγει το επιθυμητό αποτέλεσμα, μπορούμε να δοκιμάσουμε να προσθέσουμε περισσότερες λεπτομέρειες ή να χρησιμοποιήσουμε διαφορετικές λέξεις που περιγράφουν καλύτερα αυτό που θέλουμε να πετύχουμε. Αυτή η διαδικασία δοκιμής και σφάλματος είναι συχνά απαραίτητη για την επίτευξη του βέλτιστου αποτελέσματος και αποτελεί μέρος της διαδικασίας βελτιστοποίησης του μοντέλου. Συνοψίζοντας, η χρήση του text input και της εικόνας ως εισόδου στο μοντέλο Img2Img stable diffusion απαιτεί προσοχή στη λεπτομέρεια και ακρίβεια στις περιγραφές. Η δυνατότητα να τροποποιούμε το text input και να δοκιμάζουμε διαφορετικές προσεγγίσεις είναι κρίσιμη για την επίτευξη των επιθυμητών αποτελεσμάτων και την παραγωγή συνθετικών εικόνων υψηλής ποιότητας.

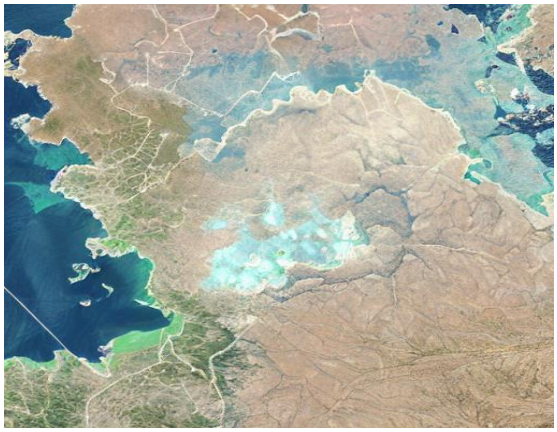
Παρακάτω παρατίθενται δύο παραδείγματα τροποποίησης με την χρήση του μοντέλου για συγκριμένο text input και εικόνα εισόδου.

Παράδειγμα 1:

Text input = {Keep the same satellite image and show how it will look in winter.}

Αποτελέσματα:

Εικόνα 3.5 : Εικόνα πριν την τροποποίηση



Εικόνα 3.6 : Εικόνα μετά την τροποποίηση

Παράδειγμα 2:

Για την ίδια εικόνα εισόδου με νέο text input έχουμε το παρακάτω αποτέλεσμα.

Text input = {keep the same satellite image and enhance the details.}

Αποτελέσματα:

Εικόνα 3.7 : Εικόνα μετά την τροποποίηση (2)

Όπως βλέπουμε ακόμα και με ένα πολύ απλό αλλά σαφές text input μπορούμε να πετύχουμε ένα πολύ ικανοποιητικό αποτέλεσμα και να εντάξουμε τις νέες συνθετικές εικόνες στο dataset μας. Παρόλο αυτά είναι καλό για ένα καλύτερο feedback να δούμε την αξία των αποτελεσμάτων μέσα από τις μετρικές.

Μετρικές

Δείκτης Δομικής Ομοιότητας (SSIM)

Ο Δείκτης Δομικής Ομοιότητας (SSIM) είναι μια μετρική που χρησιμοποιείται για τη μέτρηση της ομοιότητας μεταξύ δύο εικόνων. Ο SSIM λαμβάνει υπόψη τρεις βασικούς παράγοντες: τη φωτεινότητα, την αντίθεση και τη δομή των εικόνων. Ο τύπος για τον υπολογισμό του SSIM είναι πιο περίπλοκος, αλλά βασικά συγκρίνει τα εξής:

1. **Φωτεινότητα (Luminance)**: Πόσο φωτεινές είναι οι εικόνες.
2. **Αντίθεση (Contrast)**: Η διαφορά μεταξύ των φωτεινών και σκοτεινών περιοχών.
3. **Δομή (Structure)**: Η διάταξη των pixel και τα μοτίβα στις εικόνες.

Ο SSIM κυμαίνεται από -1 έως 1:

- **1** υποδεικνύει τέλεια ομοιότητα.
- **0** υποδεικνύει καμία ομοιότητα.
- **Αρνητικές τιμές** υποδεικνύουν διαφορές.

Στην παρούσα εργασία, ενδεικτικά ο SSIM μεταξύ της αρχικής εικόνας 3.5 και της συνθετικής εικόνας 3.6 ήταν **0.2406** ενώ η σχέση της αρχικής 3.5 με την συνθετική εικόνα 3.7 ήταν **0.2064**. Τα αποτελέσματα υποδεικνύουν ότι υπάρχουν κάποιες ομοιότητες, αλλά και σημαντικές διαφορές μεταξύ των εικόνων, κάτι το οποίο είναι αναμενόμενο μιας και γίνεται προσπάθεια αναπαράστασης της αρχικής εικόνας σε διαφορετικές συνθήκες.

Δείκτης Κανονικοποιημένης Διαφοράς Βλάστησης (NDVI)

Ο Δείκτης Κανονικοποιημένης Διαφοράς Βλάστησης (NDVI) είναι ένας δείκτης που χρησιμοποιείται για την εκτίμηση της υγείας και της πυκνότητας της βλάστησης σε μια περιοχή. Ο NDVI υπολογίζεται χρησιμοποιώντας τις τιμές των φασματικών ζωνών του κοντινού υπέρυθρου (NIR) και του κόκκινου (Red) φάσματος.

Ο τύπος για τον υπολογισμό του NDVI είναι:

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$$

Οι τιμές του NDVI κυμαίνονται από -1 έως 1:

- Τιμές κοντά στο 1 υποδεικνύουν υγιή και πυκνή βλάστηση.
- Τιμές κοντά στο 0 υποδεικνύουν περιοχές με λίγη ή καθόλου βλάστηση.
- Τιμές κοντά στο -1 υποδεικνύουν περιοχές με νερό ή άλλες επιφάνειες χωρίς βλάστηση.

Στην παρούσα εργασία, ενδεικτικά οι τιμές του NDVI για τις παραπάνω εικόνες ήταν:

- **NDVI Αρχικής Εικόνας 3.5**: 0.2322
- **NDVI Συνθετικής Εικόνας 3.6**: 0.2425
- **NDVI Συνθετικής Εικόνας 3.7**: 0.2450

Αυτές οι τιμές υποδεικνύουν ότι οι συνθετικές εικόνες διατηρούν παρόμοια επίπεδα βλάστησης με την αρχική εικόνα, με ελαφρώς υψηλότερη τιμή NDVI.

3.3 GPS Spoofing detection για UAVs

Το τελευταίο ζήτημα που εξετάζουμε στην παρούσα διατριβή είναι η ανίχνευση GPS spoofing για τα UAVs (Unmanned Aerial Vehicles). Η ανίχνευση GPS spoofing είναι ένα κρίσιμο μέτρο ασφαλείας, καθώς το GPS spoofing είναι μια μέθοδος όπου ένας κακόβουλος παράγοντας επιχειρεί να εξαπατήσει έναν δέκτη GPS μεταδίδοντας ψευδείς πληροφορίες. Αυτό μπορεί να οδηγήσει σε λανθασμένη πλοήγηση των UAVs, με δυνητικά σοβαρές συνέπειες. Για την ανίχνευση GPS spoofing, τα UAVs μπορούν να χρησιμοποιήσουν μια ποικιλία τεχνικών. Μια προσέγγιση είναι η χρήση πολλαπλών δεκτών GPS και ο διασταυρούμενος έλεγχος των δεδομένων που λαμβάνουν. Εάν οι ενδείξεις από διαφορετικούς δέκτες είναι σημαντικά διαφορετικές, αυτό μπορεί να υποδηλώνει spoofing. Επιπλέον, τα UAVs μπορούν να χρησιμοποιούν άλλα συστήματα πλοήγησης, όπως συστήματα αδρανειακής πλοήγησης ή τοπικά συστήματα εντοπισμού θέσης, ως εφεδρικό μέτρο ή για επαλήθευση. Αυτά τα συστήματα μπορούν να παρέχουν επιπλέον ασφάλεια και να βοηθήσουν στην επαλήθευση της ακρίβειας των δεδομένων GPS. Ορισμένα συστήματα, όπως αυτό που προτείνεται στη διατριβή, χρησιμοποιούν αλγόριθμους μηχανικής μάθησης για τον εντοπισμό τυπικών μοτίβων του GPS spoofing. Τα βήματα που ακολουθήθηκαν και θα αναφερθούν σε παρακάτω ενότητες με λεπτομέρεια, περιλαμβάνουν πρώτα τη συλλογή έμπιστων δεδομένων. Το GPS spoofing για UAVs αποτελεί ένα πολύ εξειδικευμένο πρόβλημα που συναντάται πιο συχνά σε στρατιωτικές επιχειρήσεις, καθιστώντας τη συλλογή δεδομένων μια πρόκληση. Στη συνέχεια, σκοπός ήταν να βρεθεί ο καταλληλότερος αλγόριθμος για το θέμα και να γίνει η εκμάθησή του πάνω σε αυτό το dataset, επιδιώκοντας το υψηλότερο δυνατό accuracy. Εφόσον τα παραπάνω έχουν επιτευχθεί, είναι σημαντικό να γίνει ένα testing του συστήματος. Δεδομένου ότι η εύρεση δεδομένων GPS spoofing για UAVs είναι αρκετά δύσκολη (λόγω ιδιωτικότητας) και τα δεδομένα που είχαμε χρησιμοποιήθηκαν στο training του αλγορίθμου, έγινε συνθετική δημιουργία τέτοιων συμβολοσειρών με εμπειρική γνώση και χρήση του random walk αλγορίθμου που αναφέρθηκε σε προηγούμενο κεφάλαιο. Με αυτόν τον τρόπο, μπορούμε πλέον στο τέλος να πραγματοποιήσουμε μια μικρή εξομοίωση τέτοιου σεναρίου με τη βοήθεια του Webots, ώστε να δούμε κατά πόσο ανταποκρίνεται το σύστημα σε ένα ενδεχόμενο spoofing.

3.3.1 Dataset

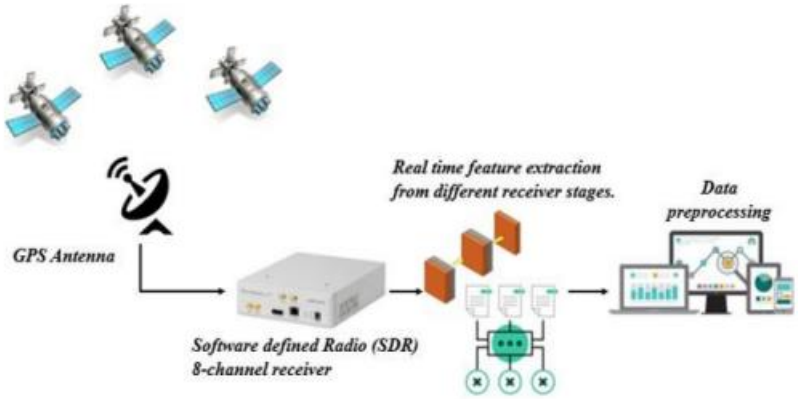
Όπως αναφέρθηκε το να βρεθούν αξιόπιστα δεδομένα GPS spoofing και μάλιστα για UAVs είναι κάτι εξαιρετικά δύσκολο. Επίσης, αρκετά δύσκολο είναι για κάποιον να δημιουργήσει τέτοια δεδομένων καθώς χρειάζεται εξειδικευμένο εξοπλισμό για να πραγματοποιήσει simulation ενός spoofing. Για την διατριβή αυτή χρησιμοποιήθηκε το dataset που δίνετε από το paper "A DATASET for GPS Spoofing Detection on Autonomous Vehicles" [28] και συγκεκριμένα η απλοποιημένη 2D έκδοση του η οποία είναι αρκετή για την υλοποίηση μας. Αναλυτικότερα σύμφωνα με το paper, το σύνολο δεδομένων περιλαμβάνει δεδομένα που εξάγονται από αυθεντικά σήματα GPS που συλλέγονται από διάφορες τοποθεσίες για την μίμηση τόσο κινούμενων όσο και στατικών αυτόνομων οχημάτων. Αυτό εξηγούν οι συγγραφείς πως επιτεύχθηκε χρησιμοποιώντας μια περιφερειακή μονάδα ραδιοφώνου γενικού λογισμικού (USRP) που έχει διαμορφωθεί ως δέκτης GPS. Κατά τη διαδικασία συλλογής δεδομένων, εξήχθησαν 13 χαρακτηριστικά από οκτώ παράλληλα κανάλια σε διαφορετικά στάδια της λειτουργίας του δέκτη, συμπεριλαμβανομένης της απόκτησης, της παρακολούθησης και της αποκωδικοποίησης πλοήγησης.

Εκτός από τα αυθεντικά σήματα GPS, το σύνολο δεδομένων περιλαμβάνει τρεις προσομοιωμένους τύπους επίθεσης GPS spoofing:

1. Simplistic attacks
2. Intermediate attacks
3. Sophisticated attacks

Αυτές οι προσομοιώσεις οδήγησαν σε ένα σύνολο δεδομένων που περιέχει συνολικά 158.170 δείγματα, με ισορροπημένη κατανομή 55% κανονικών δειγμάτων και 45% δειγμάτων που αντιστοιχούν στους τρεις τύπους επιθέσεων spoofing.

Τα δεδομένα αυτά μπορούν να χρησιμοποιηθούν για τη διερεύνηση του αντίκτυπου των επιθέσεων GPS spoofing στα εξαγόμενα χαρακτηριστικά. Μπορούν επίσης να συμβάλουν στην ανάπτυξη τεχνικών ανίχνευσης τέτοιων επιθέσεων, αξιοποιώντας μεθόδους μηχανικής εκμάθησης τόσο με επίβλεψη όσο και χωρίς επίβλεψη, άρα είναι κατάλληλα για την δική μας εφαρμογή.



Εικόνα 3.8 : Σχέδιο του μοντέλου συλλογής σημάτων GPS

3.3.2 Επιλογή & εκμάθηση μοντέλου

Το δεύτερο βήμα μετά την επιλογή κατάλληλου συνόλου δεδομένων είναι η επιλογή αλγορίθμου μηχανικής μάθησης που θα έχει την καλύτερη απόδοση για το ζητούμενο μας. Από την στιγμή που διαθέτουμε την μεταβλητή του output στο dataset μας το οποίο μπορεί να πάρει τέσσερις τιμές και υποδηλώνει το αν έχει υποστεί κάποιου είδους spoofing (απλό, μεσαίας κατηγορίας, ευφυές) το σήμα μας ή όχι, μπορούμε να θεωρήσουμε πως το πρόβλημα μας ανήκει σε αυτά της κατηγοριοποίησης. Για τον λόγο αυτό πειραματικά γίνεται χρήση των decision tree classifier, neural network classifier και random forest classifier για να διαπιστώσουμε ποιος αλγόριθμος έχει την καλύτερη απόδοση και να πορευθούμε με αυτόν στην φάση της εξομοίωσης. Οι μεταβλητές που έχουμε διαθέσιμές προς εκμάθηση στα δεδομένα μας είναι οι: Satellite Vehicle Number (PRN), Carrier Doppler in Hz (DO), Pseudorange in meter (PD), Receiver Time (RX) , Time of the Week in seconds (TOW), Carrier Phase Cycles (CP), Magnitude of the Early Correlator (EC), Magnitude of the Late Correlator (LC), The Magnitude of the Prompt Correlator (PC), Prompt in phase correlator (PIP), Prompt Quadrature Component (PQP), Carrier Doppler in Tracking loop in Hz (TCD) και Carrier to Noise Ratio in dB-Hz (C/N◦). Ενώ πρέπει να γίνεται predict η μεταβλητή του output για την οποία αναφερθήκαμε και πριν. Μετά από εκμάθηση & εφαρμογή των τριών αλγορίθμων καταλήγουμε στα εξής αποτελέσματα σχετικά με το accuracy:

```
Decision Tree Accuracy:  0.9203474820284802
Neural Network Accuracy:  0.7568409300139071
Random Forest Accuracy:  0.8988795957142577
```

Εικόνα 3.9 : Τιμές accuracy των αλγορίθμων

Άρα, λόγω καλύτερης ακρίβειας έγινε η επιλογή του decision tree classifier για την συνέχεια της υλοποίησης και πλέον διαθέτουμε έναν αξιόπιστο αλγόριθμο που μπορεί να προβλέψει αν υπάρχει spoofing και τι είδος συγκεκριμένα.

3.3.3 Δημιουργία συνθετικών συμβολοσειρών με random walk

Στο σημείο αυτό έγινε η διερεύνηση τεχνικών δημιουργίας συνθετικών συμβολοσειρών και συγκεκριμένα για GPS, καθώς πρόκειται για δεδομένα τα οποία συνήθως δεν βρίσκονται σε αφθονία λόγω αδειών και η κατάσταση γίνεται ακόμα πιο δύσκολη όταν ψάχνουμε για spoofed GPS data. Για τον λόγο αυτό αναπτύχθηκε στρατηγική που θα δώσει λύση στο πρόβλημα αυτό αλλά και σε παρόμοια προβλήματα που δεν έχουμε καθόλου πρόσβαση ή έχουμε μερική πρόσβαση στα δεδομένα. Η πρώτη κίνηση είναι να οριοθετήσουμε τις τιμές των μεταβλητών που έχουμε γνωρίζοντας την ελάχιστη και την μέγιστη τιμή που μπορούν να πάρουν. Στο δικό μας θέμα από την στιγμή που διαθέτουμε dataset με πραγματικές μετρήσεις βγάζουμε το min και το max value μέσα από αυτές αλλά και να μην διαθέταμε θα προσπαθούσαμε με λογική να ορίσουμε εμείς αυτές τις τιμές. Στην συνέχεια μαζί με την βοήθεια του random walk αλγόριθμου ξεκινάμε από μία τυχαία τιμή μέσα στα όρια μας και ο αλγόριθμος γεμίζει τις επόμενες με τυχαίες κοντινές τιμές ώστε να έχουμε συνοχή και λογική με σκοπό να μοιάζουν όσο το δυνατό πιο αληθοφανής. Με αυτόν τον τρόπο έχουμε πλέον στην διάθεση μας ένα σύνολο συνθετικών δεδομένων συμβολοσειρών GPS που δεν έχουν υποστεί spoofing. Για να δημιουργήσουμε spoofed συμβολοσειρές κρατάμε τις αρχές του αρχικού dataset που αναφέραμε πριν και μεταβάλουμε συγκεκριμένες μεταβλητές με τέτοιο τρόπο ώστε να καταφέρουμε να προσομοιάσουμε τις τρεις κατηγορίες επιθέσεων simplistic, intermediate και sophisticated. Συγκεκριμένα, για simplistic επιθέσεις πρέπει να μεταβάλουμε σύμφωνα με το paper [28] την τιμή του Carrier to Noise Ratio in dB-Hz (C/N◦) με μεγάλες διαφορές για να καταφέρουμε να έχουμε ρεαλιστικές spoofed συμβολοσειρές αυτού του είδους. Για intermediate σήματα πρέπει να εισάγουμε άλματα στις τιμές των Magnitude of the Early Correlator (EC), Magnitude of the Late Correlator (LC) και Magnitude of the Prompt Correlator (PC). Στην περίπτωση των sophisticated έχουμε κάποια delays στην αποστολή του σήματος από των transmitter που προσπαθούμε να αποτυπώσουμε στην μεταβλητή Time of the Week in seconds (TOW) και κάποιες μεταβολές που φανερώνονται στην τιμές Prompt Quadrature Component (PQP). Συνθέτοντας αυτούς τους κανόνες είμαστε πλέον σε θέση να δημιουργήσουμε ένα επαρκές συνθετικό σύνολο δεδομένων το οποίο θα χρησιμοποιήσουμε και στην φάση του testing του συστήματος μας.

	PRN	DO	PD	RX	TOW	CP	EC	LC	\
0	1	-4245.1	-6.8e+06	37121.1	3.7e+04	-1.0e+06	-10140.7	-10589.8	
1	1	-4345.4	-7.0e+06	37121.1	3.7e+04	-1.0e+06	-11047.4	-7098.8	
2	1	-4394.7	-6.7e+06	37121.1	3.7e+04	-1.0e+06	-9109.9	-7233.0	
3	1	-4311.1	-6.7e+06	37121.1	3.7e+04	-1.0e+06	-9455.3	-3877.1	
4	1	-4288.7	-6.8e+06	37121.1	3.7e+04	-1.0e+06	-12228.8	-3028.9	
..	...	...	...	...	...	...	...	...	
195	1	-2273.7	9.9e+08	37122.7	1.0e+09	-9.4e+05	-58870.4	-50662.8	
196	1	-2342.3	9.9e+08	37122.7	1.0e+09	-9.3e+05	-54888.9	-53448.6	
197	1	-2446.5	9.9e+08	37122.7	1.0e+09	-9.4e+05	-57833.0	-56084.0	
198	1	-2537.4	9.9e+08	37122.7	1.0e+09	-9.3e+05	-53580.1	-59196.3	
199	1	-2642.2	9.9e+08	37122.7	1.0e+09	-9.3e+05	-51728.6	-60846.8	

	PC	PIP	PQP	TCD	CN0	Output
0	2430.4	-273009.3	-217310.8	-3310.6	9.0e-02	0
1	526.7	-265510.8	-225933.1	-3274.1	9.8e-02	0
2	1719.9	-258082.1	-228480.7	-3147.2	7.8e-01	0
3	1617.7	-259745.7	-231448.7	-3051.3	3.1e-02	0
4	1992.5	-264417.0	-231332.7	-3199.8	3.2e-01	0
..	...	...	...	...	...	...
195	10933.2	-250941.9	-213356.9	-3443.9	9.4e-02	3
196	10508.7	-252485.1	-205577.7	-3484.2	8.3e-01	3
197	11557.3	-247678.5	-202990.1	-3274.4	1.0e+00	3
198	14677.1	-242456.6	-205138.8	-3143.1	1.0e+00	3
199	18131.6	-239068.2	-207411.8	-3150.0	1.1e+00	3

Εικόνα 3.10 : Παράδειγμα συνθετικών δεδομένων συμβολοσειρών GPS (κανονικών και spoofed)

3.3.4 Εξομοίωση & Testing

Έχοντας πλέον στην κατοχή μας ένα μοντέλο ανίχνευσης spoofing και αξιόπιστα συνθετικά δεδομένα, μπορούμε να προχωρήσουμε στο βήμα της εξομοίωσης για να αξιολογήσουμε την ποιότητα των αποτελεσμάτων. Το πρώτο βήμα είναι να προσθέσουμε και να αντιστοιχίσουμε σε κάθε γραμμή του συνθετικού dataset ένα ζευγάρι συντεταγμένων που θα υποδηλώνει ένα path πτήσης του UAV μας. Αυτό μπορούμε να το υλοποιήσουμε εύκολα με τη βοήθεια του αλγορίθμου random walk, αν και υπάρχουν και άλλοι τρόποι για να δημιουργήσουμε τα paths. Για να δημιουργήσουμε ένα ρεαλιστικό σενάριο, επιλέξαμε το Webots για την εξομοίωση. Το Webots είναι ένας προσεκτικά σχεδιασμένος simulator για ρομπότ, που δημιουργήθηκε από την ελβετική εταιρεία Cyberbotics Ltd. Στην εξομοίωση χρησιμοποιήθηκε το drone Mavic 2 Pro, το οποίο διαθέτει λογισμικό και είναι εξοπλισμένο με έναν controller αυτόνομης πλοήγησης που επίσης διατίθεται από το Webots. Στο δικό μας σενάριο, το Mavic 2 Pro έπρεπε να πλοηγηθεί σε ένα συγκεκριμένο σημείο στόχο μέσα στον κόσμο που βρίσκεται. Όταν το decision tree μοντέλο μας προβλέψει spoofing σε real-time, το drone θα επιστρέψει από άλλο path στην αφετηρία του. Για να γίνει αυτό, φορτώθηκαν στον εξομοιωτή ένα συνθετικό GPS dataset, το οποίο αναπαριστά τα δεδομένα που λαμβάνουμε την ώρα της πτήσης, και το μοντέλο για πρόβλεψη του spoofing.

Η διαδικασία της εξομοίωσης περιλαμβάνει τα εξής βήματα:

1. **Δημιουργία του περιβάλλοντος εξομοίωσης**: Ρύθμιση του Webots για να αναπαραστήσει το περιβάλλον πτήσης του UAV.

2. **Εισαγωγή των συνθετικών δεδομένων**: Φόρτωση του συνθετικού dataset στο σύστημα και αντιστοίχιση των paths πτήσης.
3. **Εκτέλεση της εξομοίωσης**: Παρακολούθηση της πτήσης του UAV και ανίχνευση τυχόν spoofing.
4. **Αξιολόγηση των αποτελεσμάτων**: Ανάλυση των δεδομένων που συλλέχθηκαν κατά τη διάρκεια της εξομοίωσης για να αξιολογηθεί η απόδοση του συστήματος ανίχνευσης.

Με αυτόν τον τρόπο, μπορούμε να διαπιστώσουμε την αποτελεσματικότητα του μοντέλου ανίχνευσης spoofing και να κάνουμε τις απαραίτητες βελτιώσεις πριν από την εφαρμογή του σε πραγματικές συνθήκες. Θα μπορούσαμε επίσης, αντί το UAV να επιστρέφει, να αντιμετωπίζαμε διαφορετικά την κατάσταση, αλλά για αποφυγή της πολυπλοκότητας του σεναρίου μείναμε σε αυτό το παράδειγμα.

Εικόνα 3.11 : Στιγμιότυπο από την εξομοίωση στο Webots

ΚΕΦΑΛΑΙΟ 4

ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

4.1 Συμπεράσματα

Η δημιουργία συνθετικών δεδομένων, και συγκεκριμένα συνθετικών εικόνων, είναι μια διαδικασία που απαιτεί μεγάλα αποθέματα υπολογιστικής ισχύος για να επιτύχουμε υψηλής ποιότητας αποτελέσματα. Η καλύτερη λύση για την αντιμετώπιση αυτής της πρόκλησης είναι η χρήση pre-trained μοντέλων και η εφαρμογή της τεχνικής transfer learning. Με αυτόν τον τρόπο, μπορούμε να προσαρμόσουμε τα μοντέλα στις συγκεκριμένες ανάγκες του προβλήματος που έχουμε να αντιμετωπίσουμε, εξοικονομώντας χρόνο και πόρους. Για τη δημιουργία συνθετικών δεδομένων συμβολοσειρών, είναι απαραίτητο να έχουμε βαθιά γνώση του αντικειμένου για το οποίο σκοπεύουμε να τα δημιουργήσουμε. Αυτό συμβαίνει επειδή χρειάζεται να συνθέσουμε κανόνες που θα καθοδηγούν τη διαδικασία δημιουργίας των δεδομένων. Σε αυτή την περίπτωση, δεν έχουμε κάποιο ολοκληρωμένο μοντέλο που να αυτοματοποιεί πλήρως τη διαδικασία, όπως συμβαίνει με τα συνθετικά δεδομένα εικόνων. Αντίθετα, χρησιμοποιούμε αλγορίθμους που παίζουν υποστηρικτικό ρόλο στην υλοποίηση. Παρά τις προκλήσεις, η δημιουργία συνθετικών δεδομένων κατέχει μεγάλη αξία και διευκολύνει την επίλυση πολλών προβλημάτων που αντιμετωπίζουν εταιρείες και ερευνητές στον χώρο της τεχνολογίας. Τα συνθετικά δεδομένα μπορούν να χρησιμοποιηθούν για την εκπαίδευση και τη δοκιμή μοντέλων μηχανικής μάθησης, την ανάπτυξη νέων αλγορίθμων και την αξιολόγηση της απόδοσης συστημάτων σε διάφορα σενάρια. Η χρήση συνθετικών δεδομένων προσφέρει επίσης τη δυνατότητα δημιουργίας μεγάλων και ποικίλων datasets, τα οποία μπορεί να είναι δύσκολο ή ακριβό να συλλεχθούν στον πραγματικό κόσμο. Αυτό είναι ιδιαίτερα σημαντικό σε περιπτώσεις όπου τα πραγματικά δεδομένα είναι σπάνια ή ευαίσθητα, όπως στην ιατρική ή σε στρατιωτικές εφαρμογές. Επιπλέον, η δημιουργία συνθετικών δεδομένων μπορεί να βοηθήσει στην αντιμετώπιση προβλημάτων που σχετίζονται με την ιδιωτικότητα και την ασφάλεια των δεδομένων. Χρησιμοποιώντας συνθετικά δεδομένα, μπορούμε να προστατεύσουμε την ιδιωτικότητα των ατόμων και να μειώσουμε τον κίνδυνο διαρροής ευαίσθητων πληροφοριών. Συνοψίζοντας, η δημιουργία συνθετικών δεδομένων είναι ένα πολύτιμο εργαλείο που μπορεί να ενισχύσει την έρευνα και την ανάπτυξη στον τομέα της τεχνολογίας. Παρά τις προκλήσεις που αντιμετωπίζουμε, οι δυνατότητες που προσφέρουν είναι τεράστιες και μπορούν να οδηγήσουν σε σημαντικές καινοτομίες και βελτιώσεις σε διάφορους τομείς.

4.2 Συνεισφορά της διατριβής

Η παρούσα διατριβή συμβάλλει σημαντικά στο επιστημονικό πεδίο της ανίχνευσης spoofing και της ασφάλειας των UAVs, προσφέροντας μεθόδους και τεχνικές που μπορούν να βελτιώσουν την απόδοση και την αξιοπιστία των συστημάτων αυτών. Επίσης, η δημιουργία συνθετικών δεδομένων, και συγκεκριμένα συνθετικών εικόνων, αποτελεί μια κρίσιμη πτυχή της έρευνάς μου. Η διαδικασία αυτή απαιτεί σημαντική υπολογιστική ισχύ και η χρήση pre-trained μοντέλων με τεχνικές transfer learning αποδείχθηκε η πιο αποδοτική προσέγγιση. Με αυτόν τον τρόπο, κατάφερα να προσαρμόσω τα δεδομένα στις συγκεκριμένες ανάγκες του προβλήματος, εξοικονομώντας χρόνο και πόρους.

Η δημιουργία συνθετικών εικόνων περιλάμβανε τη χρήση diffusion models τα οποία είναι ικανά να παράγουν υψηλής ποιότητας συνθετικές εικόνες που μπορούν να χρησιμοποιηθούν για την εκπαίδευση και τη δοκιμή μοντέλων μηχανικής μάθησης. Αυτή η τεχνική επιτρέπει τη δημιουργία μεγάλων και ποικίλων datasets, τα οποία μπορεί να είναι δύσκολο ή ακριβό να συλλεχθούν στον πραγματικό κόσμο. Επιπλέον, η χρήση συνθετικών δεδομένων συμβολοσειρών απαιτεί βαθιά γνώση του αντικειμένου, καθώς οι κανόνες που καθοδηγούν τη διαδικασία πρέπει να είναι ακριβείς και αντιπροσωπευτικοί. Η ανάπτυξη αλγορίθμων που υποστηρίζουν αυτή τη διαδικασία ήταν κρίσιμη για την επιτυχία του έργου.

Το σύστημα ανίχνευσης spoofing που ανέπτυξα δοκιμάστηκε σε διάφορα σενάρια εξομοίωσης, χρησιμοποιώντας το Webots και το drone Mavic 2 Pro. Τα αποτελέσματα έδειξαν ότι το σύστημα μπορεί να ανιχνεύσει αποτελεσματικά τις επιθέσεις spoofing και να καθοδηγήσει το UAV πίσω στην αφετηρία του μέσω ασφαλών διαδρομών. Η επιτυχία αυτών των δοκιμών ανοίγει τον δρόμο για περαιτέρω έρευνα και ανάπτυξη, όπως η αυτοματοποίηση της δημιουργίας συνθετικών δεδομένων συμβολοσειρών με τη χρήση εξειδικευμένων μοντέλων LLM, και η μελέτη πιο πολύπλοκων σεναρίων ανίχνευσης spoofing, όπως η συνέχιση της αποστολής του UAV παρά την ύπαρξη spoofing ή η χρήση σμήνους UAVs.

Η διατριβή συμβάλλει στην ανάπτυξη ενός ολοκληρωμένου συστήματος ανίχνευσης spoofing που μπορεί να βελτιώσει την ασφάλεια και την αξιοπιστία των UAVs. Η έρευνά μου συνεισφέρει στις λύσεις σε προβλήματα που αντιμετωπίζουν εταιρείες και ερευνητές στον χώρο της τεχνολογίας, διευκολύνοντας την ανάπτυξη νέων αλγορίθμων και την αξιολόγηση της απόδοσης συστημάτων σε διάφορα σενάρια. Επιπλέον, η χρήση συνθετικών δεδομένων προσφέρει τη δυνατότητα δημιουργίας μεγάλων και ποικίλων datasets, προστατεύοντας παράλληλα την ιδιωτικότητα των ατόμων και μειώνοντας τον κίνδυνο διαρροής ευαίσθητων πληροφοριών.

4.3 Μελλοντικές εφαρμογές

Από την παρούσα διατριβή, ένα ενδιαφέρον θέμα που προκύπτει και όπου υπάρχει χώρος εξέλιξης είναι η υλοποίηση έξυπνης αυτοματοποιημένης διαδικασίας δημιουργίας συνθετικών δεδομένων συμβολοσειρών χρησιμοποιώντας ένα εξειδικευμένο μοντέλο LLM (Large Language Model). Με αυτόν τον τρόπο, δεν θα χρειάζεται πλέον βαθιά γνώση από τον χρήστη για το θέμα στο οποίο χρειάζεται τις συνθετικές συμβολοσειρές. Η επεξεργασία και ο καθορισμός των κανόνων θα γίνεται μέσω του μοντέλου, καθιστώντας τη διαδικασία πιο προσιτή και εύχρηστη για τους χρήστες. Η χρήση ενός εξειδικευμένου μοντέλου LLM μπορεί να αυτοματοποιήσει τη διαδικασία δημιουργίας συνθετικών δεδομένων, μειώνοντας την ανάγκη για ανθρώπινη παρέμβαση και επιταχύνοντας τη διαδικασία. Αυτό θα επιτρέψει στους ερευνητές και τις εταιρείες να δημιουργούν μεγάλα και ποικίλα datasets με μεγαλύτερη ευκολία, βελτιώνοντας την απόδοση των μοντέλων μηχανικής μάθησης και διευκολύνοντας την ανάπτυξη νέων αλγορίθμων. Επιπλέον, μια χρήσιμη μελλοντική επέκταση είναι να μελετηθούν στο πειραματικό σύστημα ανίχνευσης spoofing και στην εξομοίωση περαιτέρω σενάρια εκτός από την επιστροφή του UAV στη βάση. Για παράδειγμα, θα μπορούσαμε να εξετάσουμε σενάρια όπου το UAV συνεχίζει την επιχείρηση παρά την ύπαρξη spoofing. Αυτό θα απαιτούσε την ανάπτυξη στρατηγικών αντιμετώπισης των επιθέσεων spoofing, επιτρέποντας στο UAV να συνεχίσει την αποστολή του με ασφάλεια. Ένα άλλο ενδιαφέρον σενάριο είναι η χρήση ενός σμήνους από UAVs. Σε αυτή την περίπτωση, θα μπορούσαμε να εξετάσουμε πώς το σύστημα ανίχνευσης spoofing θα λειτουργούσε σε ένα περιβάλλον με πολλαπλά UAVs που συνεργάζονται μεταξύ τους. Αυτό θα απαιτούσε την ανάπτυξη αλγορίθμων που θα επιτρέπουν στα UAVs να επικοινωνούν και να συνεργάζονται για την αντιμετώπιση των επιθέσεων spoofing. Επίσης, η μελέτη της επίδρασης διαφορετικών τύπων επιθέσεων spoofing και η ανάπτυξη προσαρμοσμένων στρατηγικών αντιμετώπισης για κάθε τύπο επίθεσης θα μπορούσε να βελτιώσει την ανθεκτικότητα του συστήματος. Αυτό θα περιλάμβανε την ανάλυση των χαρακτηριστικών των επιθέσεων και την ανάπτυξη μοντέλων που μπορούν να ανιχνεύσουν και να αντιδράσουν σε κάθε τύπο επίθεσης με τον πιο αποτελεσματικό τρόπο. Τέλος, η συνεργασία με ρυθμιστικές αρχές και οργανισμούς για την ανάπτυξη προτύπων και κανονισμών που θα διασφαλίζουν την ασφάλεια των UAVs από επιθέσεις spoofing είναι κρίσιμη. Αυτό θα βοηθήσει στην ευρύτερη υιοθέτηση των τεχνολογιών ανίχνευσης spoofing και θα διασφαλίσει ότι τα UAVs θα μπορούν να λειτουργούν με ασφάλεια σε διάφορα περιβάλλοντα. Με αυτές τις βελτιώσεις και επεκτάσεις, μπορούμε να διασφαλίσουμε ότι το σύστημα ανίχνευσης spoofing θα είναι πιο αποτελεσματικό και αξιόπιστο, συμβάλλοντας στην ασφάλεια και την αξιοπιστία των UAVs σε διάφορες εφαρμογές.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Tom M. Mitchell, "Machine Learning", McGraw-Hill, 1997.
[2] Richard S. Sutton and Andrew G. Barto, "Introduction to Reinforcement Learning", *The MIT Press, second edition*, 2018.
[3] Thomas M. Moerland , J. Broekens, A. Plaat and Catholijn M. Jonker, "Model-based Reinforcement Learning: A Survey", *arXiv*, 2022.
[4] I. Goodfellow, Y. Bengio and A. Courville, "Deep Learning", *MIT Press*, 2016.
[5] I. Kavdır, "Evaluation of different pattern recognition techniques for apple sorting", *Biosystems Engineering*, 2008.
[6] D. Jurafsky and James H. Martin, "Speech and Language Processing", *Draft of February 3*, 2024.
[7] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network", *Elsevier*, 2020.
[8] J. Alanya-Beltran , R. Shankar , P. Krishna , S. Kumar S, "Investigation of Bi-Directional LSTM deep learning-based ubiquitous MIMO uplink NOMA detection for military application considering Robust channel conditions", *The Journal of Defense Modeling and Simulation*, 2023.
[9] M. Evchenko, J. Vanschoren, Holger H. Hoos, M. Schoenauer, M. Sebag, "Frugal Machine Learning", *arXiv*, 2021.
[10] X. Guo and Y. Chen, "Generative AI for Synthetic Data Generation: Methods, Challenges and the Future", *arXiv*, 2024.
[11] R. He, S. Sun, X. Yu, C. Xue, W. Zhang, P. Torr, S. Bai and X. Qi, "Is synthetic data from generative models ready for image recognition?", *arXiv*, 2023.
[12] Z. Yang, F. Zhan, K. Liu, M. Xu and S. Lu, "AI-Generated Images as Data Source: The Dawn of Synthetic Era", *arXiv*, 2023.
[13] M. Chen, S. Mei, J. Fan and M. Wang, "An Overview of Diffusion Models: Applications, Guided Generation, Statistical Rates and Optimization", *arXiv*, 2024.
[14] P. Salehi, A. Chalechale and M. Taghizadeh, "Generative Adversarial Networks (GANs): An Overview of Theoretical Model, Evaluation Metrics, and Recent Developments", *arXiv*, 2020.
[15] K. Ganguly, "Learning Generative Adversarial Networks: Next-generation deep learning simplified", *Packt Publishing*, 2017.
[16] I. Goodfellow, "NIPS 2016 tutorial: Generative adversarial networks", arXiv, 2016.
[17] H. Alqahtani, M. Kavakli-Thorne and G. Kumar, "Applications of Generative Adversarial Networks (GANs): An Updated Review", *Archives of Computational Methods in Engineering*, 2019.
[18] S. Chen and W. Guo, "Auto-Encoders in Deep Learning—A Review with New Perspectives", *Mathematics*, 2023.
[19] H. Naveed, A. Ullah Khan, S. Qiu, M. Saqib,S. Anwar, M. Usman, N. Akhtar, N. Barnes and A. Mian, "A Comprehensive Overview of Large Language Models", *arXiv*, 2023.
[20] Y. Liu, H. He, T. Han, X. Zhang, M. Liu, J. Tian, Y. Zhang, J. Wang, X. Gao, T. Zhong, Y. Pan, S. Xu, Z. Wu, Z. Liu, X. Zhang, S. Zhang, X. Hu, T. Zhang, N. Qiang, T. Liu and B. Ge, "Understanding LLMs: A Comprehensive Overview from Training to Inference", *arXiv*, 2024.
[21] Z. Li1 , H. Zhu , Z. Lu , M. Yin, "Synthetic Data Generation with Large Language Models for Text Classification: Potential and Limitations", *arXiv*, 2023.

[22] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: Bringing order to the web.” *Stanford InfoLab, Tech. Rep.*, 1999.
[23] D. Fogaras, B. R´acz, K. Csalog´any, and T. Sarl ´os, “Towards scaling fully personalized pagerank: Algorithms, lower bounds, and experiments.”, *Internet Mathematics*, 2005.
[24] T. H. Haveliwala, “Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search,” *IEEE Transactions on Knowledge and Data Engineering*, 2003.
[25] Y. Aharonov, L. Davidovich, and N. Zagury, “Quantum random walks.” *Physical Review A*, 1993.
[26] F. Xia, J. Liu, H. Nie, Y. Fu, L. Wan and X. Kong, “Random Walks: A Review of Algorithms and Applications”, *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2019.
[27] C. Meng, Y. He, Y. Song, J. Song, J. Wu, J. Zhu and S. Ermon, “SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations”, *arXiv*, 2021.
[28] G. Aissou, S. Benouadah, H. El Alami and N. Kaabouch, “A DATASET for GPS Spoofing Detection on Autonomous Vehicles”, *IEEE Dataport*, 2022.
[29] O. Michael, “Webots: Professional Mobile Robot Simulation”, *Journal of Advanced Robotics Systems*, 2004.
[30] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction," *in IEEE Transactions on Neural Networks*, 1998.

ΙΣΤΟΤΟΠΟΙ

[31] towardsdatascience.com

[32] www.techtarget.com

[33] www.geeksforgeeks.org

[34] datage.tech

[35] commons.wikimedia.org

[36] www.mathworks.com

[37] www.turing.com

[38] github.com

[39] stability.ai

[40] www.marketsandmarkets.com

[41] www.britannica.com

[42] huggingface.co

[43] www.nasa.gov

[43] www.cyberbotics.com

[44] www.researchgate.net

[45] medium.com