

2022-01

$\text{p}\ddot{\text{y}} \cdot 0 \hat{\text{A}} \pm 1 \text{ ' } \mu \acute{\text{I}} \text{ ; } \frac{1}{2} \ddot{\text{A}} \pm \hat{\text{A}} - \frac{1}{2} \pm \frac{1}{2} \hat{\text{A}} \hat{\text{A}} \text{ ; } \gg \text{ ; } 3^1$
 $\text{p}\ddot{\text{y}} \hat{\text{A}} \pm \text{ }^{-\frac{3}{4}} \mu^1 \text{ Super Mario } 3^1 \pm \ddot{\text{A}} \text{ ; } \hat{\text{A}} \hat{\text{A}}$
 $\text{p}\ddot{\text{y}} \ddot{\text{A}} \cdot \hat{\text{A}} \text{ ¶ } \acute{\text{E}} \text{ ® } \hat{\text{A}} \ddot{\text{A}} \text{ ; } \hat{\text{A}}$

$\text{p}\ddot{\text{y}} \text{ ' } \frac{1}{2} \pm \tilde{\text{A}} \ddot{\text{A}} \pm \tilde{\text{A}} \text{ }^{-} \text{ ; } \hat{\text{A}}, \text{ " } \mu \acute{\text{E}} \acute{\text{A}} 3 \text{ }^{-} \pm$

$\text{p}\ddot{\text{y}} \acute{\text{A}} \text{ }^{\text{I}^3} \acute{\text{A}} \pm \frac{1}{4} \frac{1}{4} \pm \text{ " } \cdot \frac{1}{4} \text{ }^{\text{I}} \tilde{\text{A}} 1 \pm \hat{\text{A}} \text{ " } 1 \text{ ; } \text{ }^{-0} \cdot \tilde{\text{A}} \cdot \hat{\text{A}}, \text{ £ } \text{ Ç } \text{ ; } \gg \text{ ® } \text{ }^{\text{Y}1^0} \text{ ; } \frac{1}{2} \text{ ; } \frac{1}{4} 1^0 \hat{\text{I}} \frac{1}{2} \cdot \hat{\text{A}} 1 \tilde{\text{A}} \tilde{\text{A}} \cdot \frac{1}{4} \hat{\text{I}} \frac{1}{2} \text{ }^0 \pm 1 \text{ " } 1 \text{ ; } \text{ }^{-0} \cdot$

$\text{p}\ddot{\text{y}} \pm \frac{1}{2} \mu \hat{\text{A}} 1 \tilde{\text{A}} \tilde{\text{A}} \text{ ® } \frac{1}{4} 1 \text{ ; } \cdot \mu \text{ }^{-} \hat{\text{A}} \text{ ; } \gg 1 \hat{\text{A}} \text{ }^{-} \text{ Æ } \text{ ; } \hat{\text{A}}$

<http://hdl.handle.net/11728/12135>

Downloaded from HEPHAESTUS Repository, Neapolis University institutional repository

Πανεπιστήμιο Νεάπολις Πάφου
Σχολή Οικονομικών, Διοίκησης και Πληροφορικής
Τμήμα Πληροφορικής

Εκπαιδεύοντας έναν υπολογιστή να παίξει Super Mario για το υπόλοιπο της ζωής του.

Γεωργία Αναστασίου

Επιβλέπων : Δρ. Σάββας Χατζηχριστοφής

10 Ιανουαρίου 2022

ΠΕΡΙΕΧΟΜΕΝΑ

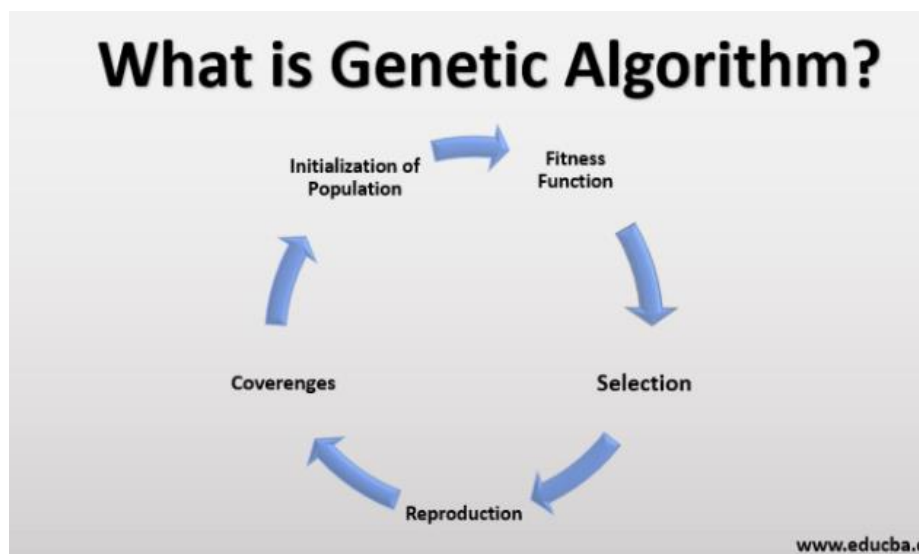
Περίληψη	3
Κεφάλαιο 1	6
1.1 Σκοπός Διπλωματικής.....	6
1.2 Θέμα Διπλωματικής	6
1.3 Δομή Διπλωματικής	6
Κεφάλαιο 2	7
2.1 Τεχνητή Νοημοσύνη (AI)	7
2.1.1 Παραδείγματα.....	7
2.1.1.1 Self – driving vehicles.....	7
2.1.1.2 Medicine.....	8
2.2 Deep Learning	8
2.2.1 Παραδείγματα.....	9
2.2.1.1 Virtual Assistants.....	9
2.2.1.2 Μεταφράσεις	9
2.2.1.3 Facial Recognition	9
2.2.1.4 Εξατομικευμένες αγορές και ψυχαγωγία	10
2.3 Reinforcement Learning	10
2.3.1 Τύποι Reinforcement Learning	10
2.3.2 Γιατί να χρησιμοποιήσουμε Reinforcement Learning.....	11
2.3.3 Πότε να μην χρησιμοποιήσουμε Reinforcement Learning.....	11
2.3.4. Παραδείγματα.....	11
2.3.4.1 Εφαρμογές στην Υγεία	11
2.3.4.2 Reinforcement Learning στην αυτοματοποίηση της Βιομηχανίας	12
2.4 Γενετικοί Αλγόριθμοι	13
2.4.1 Γιατί να χρησιμοποιήσουμε Γενετικούς Αλγορίθμους.....	13
2.4.2 Εφαρμογή Γενετικών Αλγορίθμων	13
2.5 Νευρωνικά Δίκτυα	13
2.5.1 Τρόποι με τους οποίους μπορούν να μάθουν τα Νευρωνικά Δίκτυα	14
2.5.2 Γιατί χρησιμοποιούμε Νευρωνικά Δίκτυα	14
2.5.3 Παραδείγματα.....	14
2.5.3.1 Μετατροπή φωνής σε κείμενο	14
2.5.3.2 Συστάσεις.....	15
2.6 Machine Learning	15
2.6.1 Εφαρμογές Μηχανικής Μάθησης	16

Κεφάλαιο 3	17
3.1 Τι υπάρχει στην αγορά βάση ερευνητικών άρθρων	17
Κεφάλαιο 4	19
4.1 Υλοποίηση λογισμικού	19
4.1.1 Εργαλεία που απαιτούνται για την ανάπτυξη του λογισμικού	19
Επίλογος	25

Περίληψη

Στις μέρες μας, μπορούμε να πούμε με βεβαιότητα ότι έχει αναπτυχθεί η επιστήμη της Τεχνητής Νοημοσύνης (AI). Η χρήση της έχει βοηθήσει όχι μόνο στην καθημερινότητά μας αλλά και σε πιο εξειδικευμένους τομείς. Χάρη σε αυτό έχουν αναπτυχθεί πολλοί κλάδοι της επιστήμης, όπως η Ιατρική, στις μεταφορές, στην Εκπαίδευση, στην Ασφάλεια, στην Γεωργία, στην Ρομποτική, στα Οικονομικά, στην Ενέργεια, στα παιχνίδια κλπ. Πολλοί χρησιμοποιούν τις μεθόδους του “deep learning” και του “reinforcement learning” καθώς και των νευρωνικών δικτύων (neural networks) και γενετικών αλγορίθμων (genetic algorithms) για την επίτευξη των στόχων τους.

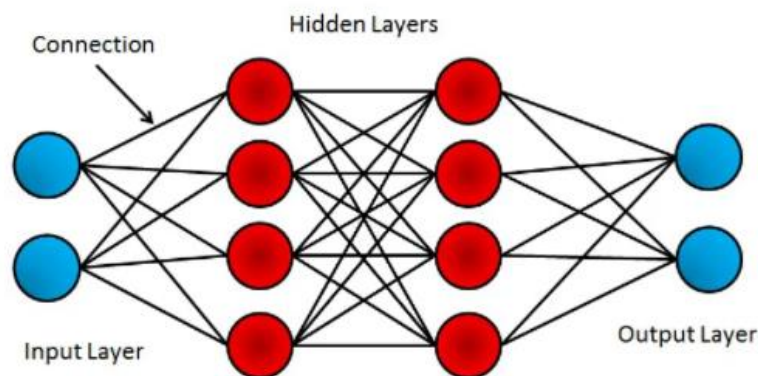
Οι γενετικοί αλγόριθμοι είναι μια μέθοδος βελτιστοποίησης που περιλαμβάνει επαναληπτικές διαδικασίες αναζήτησης που βασίζονται σε μια αναλογία με τη διαδικασία της φυσικής επιλογής και της εξελικτικής γενετικής.



Εικόνα 1. Γενετικοί Αλγόριθμοι

Τα νευρωνικά δίκτυα είναι η χρήση διαφορετικών αλγορίθμων για την κατανόηση των σχέσεων σε ένα δεδομένο σύνολο δεδομένων για την παραγωγή των καλύτερων αποτελεσμάτων από τις μεταβαλλόμενες εισόδους. Το δίκτυο εκπαιδεύεται κατάλληλα ούτως ώστε να παράγει τα επιθυμητά αποτελέσματα και χρησιμοποιούνται διαφορετικά μοντέλα για την πρόβλεψη των μελλοντικών αποτελεσμάτων με τα δεδομένα. Στην ουσία, οι κόμβοι διασυνδέονται έτσι ώστε να

λειτουργεί σαν ανθρώπινος εγκέφαλος. Για την ομαδοποίηση και την ταξινόμηση των δεδομένων χρησιμοποιούνται διαφορετικοί συσχετισμοί και κρυφά μοτίβα σε ακατέργαστα μοτίβα.



Εικόνα 2. Νευρωνικά Δίκτυα

Σύμφωνα με το θέμα της διπλωματικής μου εργασίας, πρέπει να αναβαθμίσω το παιχνίδι Super Mario ούτως ώστε να μπορεί να παίζει ο υπολογιστής για το υπόλοιπο της ζωής του χωρίς να χάνει. Αυτό θα επιτευχθεί με την χρήση του “deep learning” και ειδικότερα του “reinforcement learning”. Το πρόγραμμα πρέπει να σχεδιαστεί έτσι ώστε ο Super Mario να αποφεύγει τα εμπόδια που θα εμφανιστούν μπροστά του. Πρέπει να γίνει ο σωστός υπολογισμός ώστε να μάθει να πηδά σωστά.

Πιο συγκεκριμένα, το λογισμικό θα προσπαθεί με διάφορες κινήσεις κυρίως πηδήματα να καταφέρει να περάσει τα εμπόδια.

Για να επιτευχθεί αυτό χρειάστηκε να χρησιμοποιήσω “deep learning”, το οποίο είναι ένα υποσύνολο του “machine learning” όπως και της τεχνητής νοημοσύνης (AI). Τεχνητά νευρωνικά δίκτυα, αλγόριθμοι μαθαίνουν να κάνουν χρήση περισσότερων δεδομένων από ότι στα απλά λογισμικά. Έτσι επιτυγχάνετε η επίλυση πολύπλοκων προβλημάτων ακόμη και όταν τα δεδομένα είναι πιο περίπλοκα και όχι τόσο εύχρηστα.

Επίσης, έγινε χρήση και “reinforcement learning”, το οποίο είναι ένας τομέας του “machine learning” που έχει ως σκοπό να διδάξει τον τρόπο που πρέπει να

αναλαμβάνονται ενέργειες σε ένα περιβάλλον ώστε να μεγιστοποιείτε η έννοια της αθροιστικής ανταμοιβής.

Το αρχικό παιχνίδι που είχε γίνει τόσο γνωστό και αγαπητό στα παιδικά μας χρόνια, παίζεται με την χρήση joystick το οποίο απαιτεί να εξοικειωθείς με την χρήση πολλών πλήκτρων την ίδια ώρα.

Λέξεις κλειδιά: deep learning, reinforcement learning, evolutionary computation, machine learning, Τεχνητή Νοημοσύνη, Γενετική Αλγόριθμοι, Νευρωνικά Δίκτυα

Κεφάλαιο 1

1.1 Σκοπός Διπλωματικής

Σκοπός της παρούσας διπλωματικής εργασίας είναι η δημιουργία ενός λογισμικού του Super Mario που θα είναι εκπαιδευμένο έτσι ώστε να μην χρειάζεται οποιοσδήποτε χειρισμός από τον χρήστη, αντιθέτως θα παίζει στην ουσία μόνο του, προσπαθώντας να βρει τα κατάλληλα “πηδήματα” για να ξεπεράσει τα εμπόδια που θα βρίσκει μπροστά του.

1.2 Θέμα Διπλωματικής

Η ανάπτυξη της τεχνολογίας και ιδιαίτερα της Τεχνητής Νοημοσύνης είναι ραγδαία και πλέον εμφανείς σε πολλούς τομείς της καθημερινότητάς μας. Ένας από τους τομείς ανάπτυξης είναι και τα παιχνίδια. Πλέον εφαρμόζονται μέθοδοι όπως το deep learning και το reinforcement learning για να επιτευχθεί η δημιουργία ενός παιχνιδιού που θα είναι εκπαιδευμένο να βρίσκει μόνο του τις κατάλληλες κινήσεις για να ξεπεράσει οποιοδήποτε εμπόδιο ή αντικείμενο βρεθεί στην πορεία του παιχνιδιού.

1.3 Δομή Διπλωματικής

Η παρούσα Διπλωματική Εργασία περιλαμβάνει συνολικά 4 κεφάλαια, με πρώτο το πιο πάνω που είναι το εισαγωγικό, όπου παρατίθενται ο σκοπός, το θέμα και η δομή της εργασίας.

Στο Κεφάλαιο 2, εξηγούνται κάποιοι ορισμοί όπως της Τεχνητής Νοημοσύνης, του Deep Learning, του Reinforcement Learning, του Machine Learning, των Γενετικών Αλγορίθμων, των Νευρωνικών Δικτύων, μαζί με παραδείγματα, μειονεκτήματα, πλεονεκτήματα.

Στο Κεφάλαιο 3, βλέπουμε τι υπάρχει ήδη στην αγορά, τι έχει ήδη χρησιμοποιηθεί βάση κάποιων ερευνητικών άρθρων.

Στο Κεφάλαιο 4, εξηγείται βήμα-βήμα, πως έγινε η υλοποίηση του λογισμικού.

Κεφάλαιο 2

2.1 Τεχνητή Νοημοσύνη (AI)

Με τον όρο Τεχνητή Νοημοσύνη αναφερόμαστε στις διάφορες τεχνολογίες λογισμικού που βοηθούν ένα ρομπότ ή έναν υπολογιστή να σκέφτεται και να ενεργεί πιο ανθρώπινα. Θεωρείται ότι αυτό έχει επιτευχθεί εάν το λογισμικό αποδίδει εξίσου καλά ή ακόμη και καλύτερα από τον άνθρωπο. Η απόδοσή του υπολογίζεται από την ανθρώπινη υπολογιστική ακρίβεια, την ταχύτητα και την χωρητικότητα. Συνήθως χρησιμοποιείται για την ανάπτυξη συστημάτων υπολογιστών που μπορούν να εκτελέσουν εργασίες που συνήθως απαιτούν ανθρώπινη νοημοσύνη, όπως η αναγνώριση ομιλίας, η λήψη αποφάσεων, η οπτική αντίληψη, καθώς και η μετάφραση μεταξύ των γλωσσών.

Ο τομέας της Τεχνητής Νοημοσύνης ουσιαστικά είναι όταν οι μηχανές κάνουν εργασίες που απαιτούν ανθρώπινη νοημοσύνη. Περιλαμβάνει την μηχανική μάθηση (machine learning) όπου οι μηχανές μπορούν να αποκτήσουν δεξιότητες και εμπειρία χωρίς να χρειάζεται ανθρώπινη συμμετοχή.

2.1.1 Παραδείγματα

2.1.1.1 Self – driving vehicles

Η Τεχνητή Νοημοσύνη έχει αυξημένη εξέλιξη κάτι που έχει επηρεάσει και την αυτοκινητοβιομηχανία και πιο συγκεκριμένα αυτό φαίνεται στην δημιουργία αυτοοδηγούμενων αυτοκινήτων.

Πολλές είναι οι εταιρείες που το 2016 δοκίμασαν να εντάξουν τα αυτοοδηγούμενα αυτοκίνητα με Τεχνητή Νοημοσύνη. Κάποιες γνωστές εταιρείες, όπως η Apple, η Google και η Tesla, χρησιμοποιούν Τεχνητή Νοημοσύνη εδώ και χρόνια.

Τα αυτόνομα οχήματα είναι εκπαιδευμένα να φρενάρουν με ασφάλεια, να αλλάζουν λωρίδα, να αποτρέπουν συγκρούσεις, καθώς και να χρησιμοποιούν χάρτες.

Σε μια πόλη του Τέξας, έχουν ξεκινήσει ένα πιλοτικό πρόγραμμα όπου οι άνθρωποι μπορούν να χρησιμοποιήσουν δωρεάν αυτοοδηγούμενα οχήματα, αρκεί να έχουν μια εφαρμογή στο κινητό τους.



Εικόνα 3. Αυτόνομο αυτοκίνητο

2.1.1.2 Medicine

Επίσης, στο Τεχνολογικό Πανεπιστήμιο της Μασαχουσέτης, μια ομάδα χρησιμοποίησε την Τεχνητή Νοημοσύνη για την δημιουργία ενός μοντέλου για θεραπευτική αγωγή για τον καρκίνο του εγκεφάλου. Αυτό το μοντέλο, διαμορφώνει λιγότερο τοξικές αγωγές που πάλι καταφέρνουν να μειώσουν το μέγεθος του όγκου. Σε αρκετές περιπτώσεις, η συχνότητα της θεραπείας άλλαζε, δηλαδή μπορούσε να μειωθεί από μηνιαία σε μία ή δύο φορές το χρόνο και ακόμη κάποιες δόσεις μειώνονταν στο 50%.

2.2 Deep Learning

Το deep learning είναι ένα υποσύνολο της μηχανικής μάθησης, με τα νευρωνικά δίκτυα, τους αλγόριθμους να είναι εμπνευσμένοι από τον ανθρώπινο εγκέφαλο, μαθαίνοντας από μεγάλες ποσότητες δεδομένων. Ένας αλγόριθμος deep learning, αφού μαθαίνει από την εμπειρία ομοίως με τους ανθρώπους, θα εκτελούσε μια εργασία επανειλημμένα, κάθε φορά που την τροποποιούσε λίγο για να βελτιωθεί το αποτέλεσμα. Το deep learning μαθαίνει να λύνει σχεδόν κάθε πρόβλημα που απαιτεί σκέψη για την κατανόησή του.

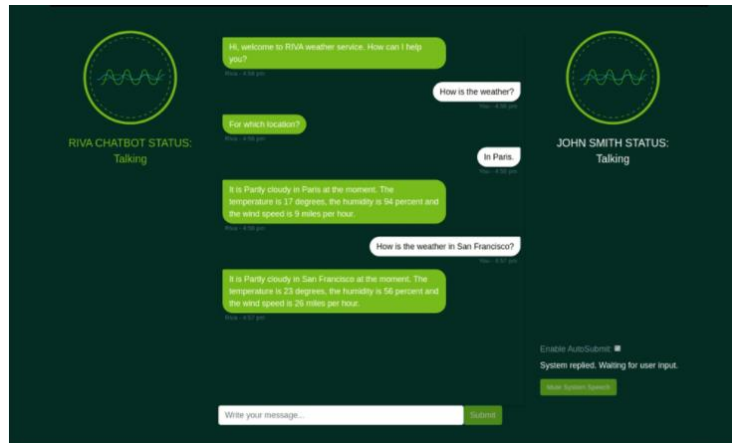
Παρατηρείται ραγδαία αύξηση τα τελευταία χρόνια στις δυνατότητες του deep learning καθώς οι αλγόριθμοι που χρησιμοποιούνται απαιτούν ένα αρκετά μεγάλο όγκο δεδομένων για μάθηση. Ο πολλαπλασιασμός της Τεχνητής Νοημοσύνης (AI) ως υπηρεσία, όπως και η ισχυρότερη υπολογιστική ισχύ που είναι διαθέσιμη σήμερα και την δημιουργία περισσότερων δεδομένων, επωφελούν τους αλγόριθμους deep learning.

Επιπρόσθετα, επιτρέπει στις μηχανές να επιλύουν πιο περίπλοκα προβλήματα, ακόμη και όταν χρησιμοποιούν ένα σύνολο δεδομένων που είναι πολύ ποικιλόμορφο, μη δομημένο και αλληλένδετο. Όσο περισσότερη μάθηση αποκτήσουν οι αλγόριθμοι deep learning, τόσο καλύτερη είναι η απόδοσή τους.

2.2.1 Παραδείγματα

2.2.1.1 Virtual Assistants

Βλέπουμε πολλές εταιρείες να χρησιμοποιούν εικονικούς βοηθούς. Μερικά παραδείγματα αυτής της τεχνολογίας είναι η Alexa της Amazon, η Siri της Apple και Cortana της Microsoft. Έχουν την δυνατότητα να κατανοούν την ομιλία και την γλώσσα των ανθρώπων και να ακολουθούν κάποιες από τις εντολές τους ακόμη και να απαντούν σε ερωτήσεις τους.



Εικόνα 4. Παράδειγμα Εικονικού βοηθού

2.2.1.2 Μεταφράσεις

Οι μεταφραστές μπορούν να μεταφράσουν μεταξύ γλωσσών με την χρήση αλγόριθμων deep learning. Αυτό είναι ιδιαίτερα χρήσιμο για πολλή κόσμο ιδιαίτερα για ταξιδιώτες, επιχειρηματίες κοκ.

2.2.1.3 Facial Recognition

Η αναγνώριση προσώπου χρησιμοποιεί deep learning καθώς συνεισφέρει στην ασφάλεια αλλά πιθανόν στο μέλλον να χρησιμεύει και στην πληρωμή προϊόντων σε καταστήματα απλά χρησιμοποιώντας τα πρόσωπά μας. Οι αλγόριθμοι deep learning καλούνται να ανταποκριθούν σε αρκετές προκλήσεις για την επίτευξη της σωστής αναγνώρισης προσώπου. Το άτομο μπορεί να έχει αλλάξει χτένισμα, να έχει ξυριστεί, να είναι μακιγιαρισμένο, να φορεί γυαλιά ή απλά να μην είναι επαρκής ο φωτισμός.



Εικόνα 5. Facial Recognition Systems

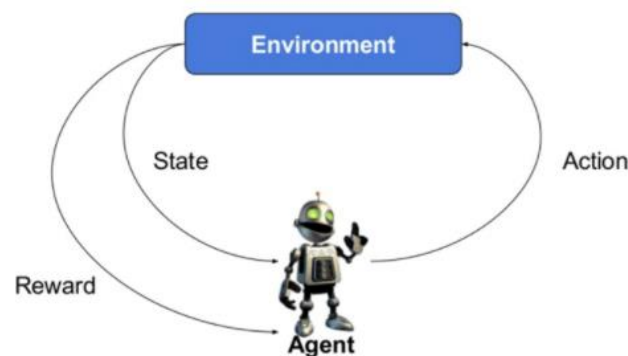
2.2.1.4 Εξατομικευμένες αγορές και ψυχαγωγία

Πολλοί έχουμε παρατηρήσει ότι υπάρχουν αρκετές εφαρμογές ψυχαγωγίας και αγορών που σου προτείνουν τι να αγοράσεις ή πια ταινία θα δεις μετά, όπως είναι η Amazon και το Netflix. Με την χρήση των αλγόριθμων deep learning εκπαιδεύονται κατάλληλα ούτως ώστε με την εμπειρία που αποκτούν με τον καιρό να μπορούν να καταλαβαίνουν τι ταινίες μας αρέσουν ή τι συνηθίζουμε να αγοράζουμε και έτσι την επόμενη φορά μας προτείνουν κάτι ανάλογο και πιο κοντά στα γούστα μας.

2.3 Reinforcement Learning

Το Reinforcement Learning είναι μία μέθοδος machine learning και ασχολείται με το πώς πρέπει να αναλαμβάνονται οι ενέργειες σε ένα λογισμικό. Βοηθά στην μεγιστοποίηση κάποιου μέρους της αθροιστικής ανταμοιβής, καθώς είναι μέρος του deep learning. Είναι μια μέθοδος εκμάθησης νευρωνικών δικτύων, που συμβάλλει στην επίτευξη σύνθετων στόχων ή στην μεγιστοποίηση μιας συγκεκριμένης διάστασης σε πολλά βήματα.

Typical RL scenario



Εικόνα 6. Reinforcement Learning Scenario

2.3.1 Τύποι Reinforcement Learning

Υπάρχουν δύο τύποι Reinforcement Learning, ο Θετικός και ο Αρνητικός.

Θετικός ορίζεται ως ένα γεγονός που συμβαίνει λόγω συγκεκριμένης συμπεριφοράς. Αυξάνει την δύναμη και τη συχνότητά της και επηρεάζει θετικά τη δράση που αναλαμβάνει ο πράκτορας. Αυτός ο τύπος βοηθά στην διατήρηση της αλλαγής για μεγαλύτερο χρονικό διάστημα και στην μεγιστοποίηση της απόδοσης.

Αρνητικός τύπος ορίζεται ως η ενίσχυση της συμπεριφοράς που εμφανίζεται λόγω μιας αρνητικής κατάστασης που θα έπρεπε να είχε σταματήσει ή να αποφευχθεί. Συμβάλλει στον ορισμό της ελάχιστης βάσης απόδοσης. Αυτή η μέθοδος έχει και ένα μειονέκτημα, παρέχει αρκετά για να καλύψει την ελάχιστη συμπεριφορά.

2.3.2 Γιατί να χρησιμοποιήσουμε Reinforcement Learning

- Βοηθά στην ανακάλυψη της ενέργειας που αποφέρει υψηλότερη ανταμοιβή για μεγαλύτερο χρονικό διάστημα.
- Παρέχει μια λειτουργία ανταμοιβής στον εκπαιδευτικό παράγοντα.
- Μπορεί να βρει την καλύτερη μέθοδο για την απόκτηση μεγάλων ανταμοιβών.
- Βοηθά να βρει ποια κατάσταση χρειάζεται δράση.

2.3.3 Πότε να μην χρησιμοποιήσουμε Reinforcement Learning

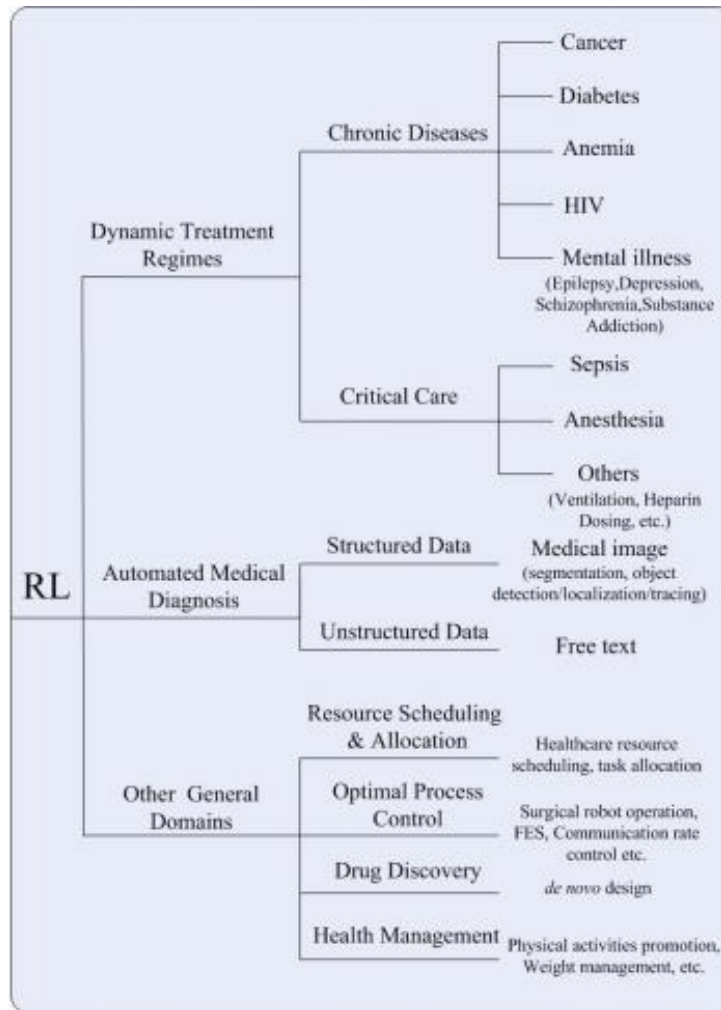
- Όταν έχετε αρκετά δεδομένα για να λύσετε το πρόβλημα τότε μπορείτε να χρησιμοποιήσετε μέθοδο supervised learning.
- Η μέθοδος Reinforcement Learning είναι βαριά-υπολογιστική και χρονοβόρα, ιδιαίτερα όταν ο χρόνος δράσης είναι μεγάλος.

2.3.4. Παραδείγματα

2.3.4.1 Εφαρμογές στην Υγεία

Στην υγειονομική περίθαλψη, οι ασθενείς μπορούν να λάβουν θεραπεία από πολιτικές που έχουν εκπαιδευτεί από συστήματα Reinforcement Learning. Με την μέθοδο αυτή, είναι δυνατή η εύρεση βέλτιστων πολιτικών με την χρήση προηγούμενων εμπειριών χωρίς να χρειάζεται προηγούμενες πληροφορίες σχετικά με το μαθηματικό μοντέλο των βιολογικών συστημάτων. Έτσι, αυτή η προσέγγιση είναι πιο εφαρμόσιμη από άλλα συστήματα που βασίζονται στον έλεγχο της υγειονομικής περίθαλψης. Αυτή η μέθοδος κατηγοριοποιείται ως δυναμικά θεραπευτικά σχήματα (DTRs) σε χρόνιες παθήσεις ή εντατικές θεραπείες, αυτοματοποιημένη ιατρική διάγνωση και άλλους τομείς.

Τα μοντέλα αυτά παίρνουν ως είσοδο ένα σύνολο κλινικών παρατηρήσεων και αξιολογήσεων ενός ασθενούς και ως αποτέλεσμα δίνει τις θεραπευτικές επιλογές για κάθε στάδιο. Η εφαρμογή του Reinforcement Learning σε DTRs είναι πλεονεκτική αφού είναι ικανό να καθορίζει αποφάσεις που εξαρτώνται από τον χρόνο για την καλύτερη θεραπεία του ασθενή την συγκεκριμένη χρονική στιγμή.



Εικόνα 7. Σχεδιάγραμμα των τομέων εφαρμογής του RL στην υγειονομική περίθαλψη.

2.3.4.2 Reinforcement Learning στην αυτοματοποίηση της Βιομηχανίας

Ρομπότ βασισμένα στη μάθηση χρησιμοποιούνται για την εκτέλεση διαφόρων εργασιών, αφού είναι πιο αποτελεσματικά από τα ανθρώπινα όντα και μπορούν να εκτελέσουν και εργασίες που είναι επικίνδυνες για τον άνθρωπο.

Η εταιρεία Deepmind που ανήκει στην Google, χρησιμοποιεί πράκτορες AI για την ψύξη των Κέντρων Δεδομένων της Google. Αυτό είχε ως αποτέλεσμα την μείωση των ενεργειακών δαπανών κατά 40%. Πλέον τα κέντρα ελέγχονται πλήρως με το σύστημα AI χωρίς να χρειάζεται ανθρώπινη παρέμβαση, χωρίς όμως αυτό να σημαίνει ότι δεν υπάρχει επίβλεψη από ειδικούς στα κέντρα δεδομένων.

Τρόπος λειτουργίας συστήματος:

- ο Λήψη στιγμιότυπων δεδομένων από τα κέντρα κάθε 5 λεπτά και τροφοδότησή τους σε deep neural networks.
- ο Πρόβλεψη για το πως διαφορετικοί συνδυασμοί θα επηρεάσουν τις μελλοντικές ενεργειακές καταναλώσεις.

- Προσδιορισμός ενεργειών που θα οδηγήσουν στην ελάχιστη κατανάλωση ενέργειας, διατηρώντας παράλληλα ένα καθορισμένο πρότυπο κριτηρίων ασφαλείας.
- Αποστολή και υλοποίηση αυτών των ενεργειών στο κέντρο δεδομένων.
- Τέλος, οι ενέργειες επαληθεύονται από το τοπικό σύστημα ελέγχου.

2.4 Γενετικοί Αλγόριθμοι

Οι Γενετικοί Αλγόριθμοι είναι προσαρμοστικοί ευρετικοί αλγόριθμοι αναζήτησης που ανήκουν στο μεγαλύτερο μέρος των εξελικτικών αλγορίθμων. Βασίζονται στις ιδέες της φυσικής επιλογής και γενετικής καθώς πρόκειται για έξυπνη εκμετάλλευση της τυχαίας αναζήτησης που παρέχεται με ιστορικά δεδομένα για να κατευθύνει την αναζήτηση στην περιοχή καλύτερης απόδοσης στο χώρο λύσεων. Προβλήματα βελτιστοποίησης και προβλήματα αναζήτησης, χρησιμοποιούν συνήθως τους γενετικούς αλγόριθμους για την δημιουργία λύσεων υψηλής ποιότητας.

Για την επίλυση ενός προβλήματος, προσομοιώνουν τη διαδικασία φυσικής επιλογής, πιο απλά προσομοιώνουν την “επιβίωση του ισχυρότερου” μεταξύ ατόμων διαδοχικής γενιάς. Κάθε γενιά αποτελείται από έναν πληθυσμό ατόμων και κάθε άτομο αντιπροσωπεύει ένα σημείο στον χώρο αναζήτησης και μια πιθανή λύση. Κάθε άτομο αναπαρίσταται ως μια συμβολοσειρά χαρακτήρων, ακέραιων, δεκαδικών, bit, η οποία είναι ανάλογη με το χρωμόσωμα.

2.4.1 Γιατί να χρησιμοποιήσουμε Γενετικούς Αλγορίθμους

Οι Γενετικοί Αλγόριθμοι αρχικά παρέχουν βελτιστοποίηση σε κατάσταση μεγάλου χώρου. Επιπρόσθετα, σε αντίθεση με την παραδοσιακή τεχνητή νοημοσύνη, δεν σπάνε όταν υπάρξει ελαφρά αλλαγή στην είσοδο ή υπάρξει παρουσία θορύβου. Επίσης είναι και στιβαροί.

2.4.2 Εφαρμογή Γενετικών Αλγορίθμων

- Επαναλαμβανόμενο Νευρωνικό Δίκτυο
- Φιλτράρισμα και επεξεργασία σήματος
- Δοκιμή μετάλλαξης
- Εκμάθηση ασαφούς βάσης κανόνων
- Σπάσιμο κώδικα

2.5 Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα είναι ένα σύνολο αλγορίθμων που έχουν μοντελοποιηθεί βάση του ανθρώπινου εγκεφάλου. Έχουν σχεδιαστεί ούτως ώστε να μπορούν να αναγνωρίζουν μοτίβα, Επίσης ονομάζονται και Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANNs). Τα νευρωνικά δίκτυα είναι πλαίσια για πολλούς διαφορετικούς αλγόριθμους μηχανικής μάθησης που συνεργάζονται, οι αλγόριθμοι αυτοί επεξεργάζονται πολύπλοκα δεδομένα. Είναι

παράδειγμα μηχανικής μάθησης, όπου το λογισμικό μπορεί να αλλάξει καθώς μαθαίνει να επιλύει ένα πρόβλημα.

Οι αλγόριθμοι σε ένα νευρωνικό δίκτυο μπορούν να μάθουν να αναγνωρίζουν φωτογραφίες που περιέχουν ένα συγκεκριμένο αντικείμενο αναλύοντας παραδείγματα εικόνων με ετικέτες πάνω τους. Μαθαίνουν σταδιακά τα χαρακτηριστικά του αντικειμένου, δημιουργώντας αυτόματα γνωρίσματα από το εκπαιδευτικό υλικό που επεξεργάζονται.

Ένα νευρωνικό δίκτυο είναι ένα σύνολο αλγορίθμων που προσπαθεί να προσδιορίσει τις υποκείμενες σχέσεις σε ένα σύνολο δεδομένων, χρησιμοποιώντας μια διαδικασία που μιμείται τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου. Μπορεί να προσαρμοστεί μόνο του, χωρίς ανθρώπινη βοήθεια στις αλλαγές όπως είναι οι διαφορετικές είσοδοι.

Η πιο πρωτοποριακή πτυχή των νευρωνικών δικτύων είναι ότι μόλις εκπαιδευτούν, μαθαίνουν μόνα τους. Έτσι, μιμούνται ανθρώπινους εγκεφάλους, οι οποίοι αποτελούνται από νευρώνες που είναι το θεμελιώδες δομικό στοιχείο της μετάδοσης πληροφοριών τόσο του ανθρώπου όσο και του νευρωνικού δικτύου.

2.5.1 Τρόποι με τους οποίους μπορούν να μάθουν τα Νευρωνικά Δίκτυα

- **Supervised Learning:** Αφού οι αλγόριθμοι εκπαιδευτούν στον τρόπο ερμηνείας των δεδομένων, τους τροφοδοτείται ένα σύνολο εισόδων και εξόδων και προβλέπουν τα αποτελέσματα.
- **Unsupervised Learning:** Οι αλγόριθμοι χρησιμοποιούν δεδομένα που δεν είναι ούτε επισημασμένα ούτε ταξινομημένα και ενεργούν βάση αυτών χωρίς καθοδήγηση. Η εκπαίδευση πραγματοποιείται χωρίς ανθρώπινη βοήθεια.
- **Reinforcement Learning:** Περιλαμβάνει τη λήψη κατάλληλων μέτρων για τη μεγιστοποίηση της ανταμοιβής σε μια συγκεκριμένη κατάσταση. Οι αλγόριθμοι μαθαίνουν ανάλογα με την ανατροφοδότηση που τους δίνετε.

2.5.2 Γιατί χρησιμοποιούμε Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα διαθέτουν ανθρώπινα χαρακτηριστικά και έχουν την ικανότητα να ολοκληρώνουν εργασίες σε άπειρες μεταθέσεις και συνδυασμούς, κάτι που τα καθιστά μοναδικά προσαρμοσμένα στις σημερινές εφαρμογές που βασίζονται σε μεγάλα δεδομένα. Με την μοναδική ικανότητά τους να κατανοούν ασαφή, αντιφατικά ή ελλιπή δεδομένα, είναι σε θέση να χρησιμοποιούν ελεγχόμενες διαδικασίες όταν δεν υπάρχουν διαθέσιμα ακριβή μοντέλα.

2.5.3 Παραδείγματα

2.5.3.1 Μετατροπή φωνής σε κείμενο

Ένα από τα κοινά χαρακτηριστικά των smartphones είναι η μετατροπή φωνής σε κείμενο. Λέγοντας μία συγκεκριμένη φράση ή πατώντας ένα κουμπί μιλάς στο τηλέφωνο, και το τηλέφωνο μετατρέπει τον ήχο σε κείμενο. Πολλές είναι οι εταιρείες που χρησιμοποιούν τα Νευρωνικά Δίκτυα για να επιτεύξουν αυτήν την λειτουργία. Η Google χρησιμοποιεί τεχνητά

νευρωνικά δίκτυα σε επαναλαμβανόμενη σύνδεση για την παροχή ισχύος φωνητικής αναζήτησης. Η Microsoft ισχυρίζεται ότι έχει αναπτύξει ένα σύστημα αναγνώρισης ομιλίας με την χρήση των νευρωνικών δικτύων, που μετατρέπει συνομιλίες με μεγαλύτερη ακρίβεια από τους ανθρώπους.

2.5.3.2 Συστάσεις

Η Amazon χρησιμοποιεί τα τεχνητά νευρωτικά δίκτυα για να εκπαιδεύσει τους αλγόριθμους της έτσι ώστε να μαθαίνουν το μοτίβο και τη συμπεριφορά των χρηστών της. Αυτό την βοηθά να παρέχει ακόμα καλύτερες και προσαρμοσμένες προτάσεις. Εμφανίζει προτάσεις χρησιμοποιώντας “οι πελάτες που είδαν αυτό το προϊόν είδαν επίσης”, “οι πελάτες που αγόρασαν αυτό το προϊόν αγόρασαν επίσης”, καθώς και μέσω επιλεγμένων προτάσεων στην αρχική σας σελίδα, καθώς και μέσω μηνυμάτων ηλεκτρονικού ταχυδρομείου.

2.6 Machine Learning

Η Μηχανική μάθηση, δίνει την δυνατότητα στις εταιρείες να επιταχύνουν τον ψηφιακό μετασχηματισμό και να περάσουν στην εποχή του αυτοματισμού, αυξάνοντας έτσι το εταιρικό ανταγωνισμό. Σήμερα, σχεδόν όλες οι εφαρμογές και τα λογισμικά, χρησιμοποιούν μηχανική μάθηση, βοηθώντας πολλές εταιρείες να επιλύσουν μια σειρά προβλημάτων. Χάρη στην μηχανική μάθηση, η τεχνητή νοημοσύνη μπόρεσε να αναπτυχθεί πέρα από την απλή εκτέλεση των εργασιών που είχε προγραμματιστεί να κάνει. Στα αρχικά της στάδια, χρησιμοποιήθηκε μόνο για την αυτοματοποίηση εργασιών χαμηλού επιπέδου σε επιχειρηματικές και εταιρικές ρυθμίσεις, όπως έξυπνη αυτοματοποίηση ή απλή ταξινόμηση βάση κανόνων. Συνεπώς, οι αλγόριθμοι τεχνητής νοημοσύνης περιορίζονταν μόνο στον τομέα για τον οποίο υποβλήθηκαν σε επεξεργασία, ωστόσο, με την μηχανική μάθηση, οι υπολογιστές άρχισαν εξελίσσονται με κάθε επανάληψη όπως είχαν προγραμματιστεί.

Αυτό που την κάνει να ξεχωρίζει από την τεχνητή νοημοσύνη είναι η ικανότητα να εξελίσσεται, χρησιμοποιώντας διάφορες τεχνικές προγραμματισμού είναι σε θέση να επεξεργάζεται μεγάλες ποσότητες δεδομένων και να εξάγει χρήσιμες πληροφορίες. Έτσι μπορούν να βελτιώσουν τις προηγούμενες επαναλήψεις τους μαθαίνοντας από τα δεδομένα που τους παρέχονται.

Η μηχανική μάθηση απαιτεί καλή ροή οργανωμένων, ποικίλων δεδομένων για μια ισχυρή λύση. Τα μεγάλα δεδομένα είναι χρονοβόρα και δύσκολο να επεξεργαστούν με τα ανθρώπινα πρότυπα, αλλά τα δεδομένα καλής ποιότητας είναι η καλύτερη τροφή για την εκπαίδευση ενός αλγόριθμου μηχανικής μάθησης. Για να επιτύχουμε αποτελεσματική εκπαίδευση του αλγόριθμου, τα δεδομένα πρέπει να είναι όσο πιο καθαρά, χρηστικά και αναγνώσιμα από μηχανή δεδομένων υπάρχουν. Οι αλγόριθμοι μηχανικής μάθησης έχουν την ικανότητα να βελτιώνονται μέσω της εκπαίδευσης χρησιμοποιώντας τρεις μεθόδους, την εποπτευόμενη μάθηση, την μάθηση χωρίς επίβλεψη και την ενισχυμένη μάθηση.

2.6.1 Εφαρμογές Μηχανικής Μάθησης

Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται σε περιπτώσεις που απαιτείται η λύση για να συνεχιστεί η βελτίωση μετά την ανάπτυξη. Οι αλγόριθμοι και οι λύσεις μηχανικής μάθησης είναι ευέλικτοι και μπορούν να χρησιμοποιηθούν ως υποκατάστατο της ανθρώπινης εργασίας μεσαίας ειδίκευσης υπό τις κατάλληλες συνθήκες.

Για παράδειγμα, τα στελέχη εξυπηρέτησης πελατών σε μεγάλες εταιρείες έχουν αντικατασταθεί από chatbots, δηλαδή αλγόριθμους μηχανικής εκμάθησης επεξεργασίας φυσικής γλώσσας. Έχουν την δυνατότητα να αναλύουν ερωτήματα πελατών και να παρέχουν υποστήριξη σε στελέχη υποστήριξης πελατών ή να συναλλάσσονται με τους πελάτες απευθείας.

Επιπρόσθετα, συμβάλλουν στην βελτίωση της εμπειρίας χρήστη και στην προσαρμογή για διαδικτυακές πλατφόρμες. Το Facebook, η Amazon, το Netflix και η Google, παρέχουν μοναδικό περιεχόμενο σε μεμονωμένους χρήστες με βάση τις προτιμήσεις τους και τις αντιπάθειές τους, αποτρέποντας την υπερβολή περιεχομένου.

Κεφάλαιο 3

3.1 Τι υπάρχει στην αγορά βάση ερευνητικών άρθρων

Μετά από συγκέντρωση κάποιων άρθρων έγινε έρευνα για τον τρόπο που χρησιμοποιούν κάποιοι το Deep Reinforcement Learning για να δημιουργήσουν το παιχνίδι του Super Mario.

Σχετικά με το πρώτο άρθρο της βιβλιογραφίας, βάση αναζητήσεων πλέγματος στις τεχνικές δημιουργίας περιβάλλοντος καθώς και τα “Kfast” “Kslow” έγινε εύρεση υπερπαραμέτρων για την μέθοδο που ακολούθησαν. Εκπαίδευσαν ένα “Cb-VAE” για την ανακατασκευή των παρατηρήσεων τους σχετικά με τα υποδείγματα τους - Kfast = 2, Kslow = 5. Στην εφαρμογή που δημιούργησαν υπάρχουν συνδυασμοί κουμπιών οι οποίοι αντιστοιχούν σε μια μοναδική ενέργεια με αποτέλεσμα 12 πιθανές κινήσεις. Έχουν προσεγγίσει το συνδυασμό τοπικής και παγκόσμιας εξερεύνησης που βασίζεται στην έννοια της περιέργειας με ανακατασκευής περιβάλλοντος. Αυτή η μέθοδος μπορεί να συνδυαστεί με οποιοδήποτε Reinforcement Learning Algorithm χωρίς προηγούμενη γνώση του περιβάλλοντος. Κατεύθυνσή τους είναι να ενσωματώσουν πολλαπλά επίπεδα εξερεύνησης ώστε να έχουν μια πιο αποτελεσματική εξερεύνηση. Επιπρόσθετα, θέλουν να αναβαθμίσουν την μέθοδο τους ώστε να μπορούν να αντιμετωπίσουν και διάφορες άλλες καταστάσεις όπως αισθητήρες ρομπότ.

Με βάση το δεύτερο άρθρο της Βιβλιογραφίας, παρατηρείτε η χρήση του Unsupervised learning σε συνδυασμό με το Deep learning όπου γίνεται εστίαση στην εκμάθηση της αναπαράστασης στο ολικό περιεχόμενο και στη συνέχεια σε ένα δοκιμαστικό νέο περιεχόμενο από αυτή την αναπαράσταση. Επίσης, μπορεί να χρησιμοποιηθεί και μετασχηματισμός δεδομένων μιας ακολουθίας και με την χρήση Unsupervised learning να μάθουμε την σχέση που έχουν τα στοιχεία μεταξύ τους. Κάποιοι τείνουν να χρησιμοποιούν evolutionary computation (EC) για να προσεγγίσουν την εκπαίδευση βαθιών νευρωνικών δικτύων. Τα περισσότερα evolutionary computation κάνουν χρήση και deep learning. Στο λογισμικό του Super Mario κάθε επίπεδο παρουσιάζεται από ένα σύνολο φωνών με το μέγεθος των πιθανών πλακιδίων σε ένα επίπεδο. Κάθε φωνή αποτελεί ένα μονοδιάστατο πίνακα του ίδιου μήκους του επιπέδου, στο οποίο κάθε στοιχείο υποδεικνύει την κατακόρυφη θέση του πλακιδίου εάν εμφανίζεται στην αντίστοιχη στήλη, αλλιώς

ισούται με 0. Για τις εισηγήσεις τοποθέτησης πλακιδίων στα επίπεδα του Super Mario, τα νευρωνικά δίκτυα εκπαιδεύτηκαν και εξελίχθηκαν μέσω της νευροεξέλιξης επαυξητικών τοπολογιών (NEAT). Για τον Super Mario, επικεντρώθηκαν ειδικά σε πράκτορες deep reinforcement learning, όπου ο ανθρώπινος έλεγχος προσομοιώνεται δίνοντας στον πράκτορα ανακριβείς ελέγχους μέσω στοχαστικών επιδράσεων στις ενέργειες.

Στο τρίτο άρθρο της Βιβλιογραφίας, προτείνεται ένας αλγόριθμος δύο σταδίων προς αντιμετώπιση του προβλήματος ενίσχυσης πολλαπλών στόχων μάθησης. Το πρώτο στάδιο αποτελεί ένας πολλαπλών πολιτικών αλγόριθμος μαλακών κριτικών-ηθοποιών που μαθαίνει συνεργατικά πολλαπλές πολιτικές με διαφορετικές προτιμήσεις σε στόχους. Το δεύτερο στάδιο είναι ένας πολυσκοπικός αλγόριθμος στρατηγικής εξέλιξης για την επίτευξη ομαλής προσέγγισης των συνόρων. Είναι μια αποτελεσματική μέθοδος στην εξερεύνηση δεδομένων και στην αναπαράσταση μοντέλων. Για την βελτίωση αυτού του μοντέλου πρέπει να διερευνηθούν περισσότερες στρατηγικές συνεργασίας για βελτίωση της αποτελεσματικότητας της δειγματοληψίας. Επίσης, μπορούν να διερευνηθούν πιο αποτελεσματικοί μέθοδοι αξιολόγησης για την επιτάχυνση του μοντέλου μάθησης στο δεύτερο στάδιο.

Κεφάλαιο 4

4.1 Υλοποίηση λογισμικού

Στην ενότητα που ακολουθεί θα περιγράψουμε τη διαδικασία που ακολουθήθηκε για την υλοποίηση του λογισμικού που απαιτείται για την εκπόνηση της διπλωματικής εργασίας.

4.1.1 Εργαλεία που απαιτούνται για την ανάπτυξη του λογισμικού

Το λογισμικό έχει αναπτυχθεί στην γλώσσα python, οπότε χρειάστηκε να την κατεβάσουμε στον υπολογιστή μας. Ακολούθως χρειάστηκε να κατεβάσουμε το Jupiter lab, ένα web-based εργαλείο, όπου παρέχει notebooks έχοντας τη δυνατότητα να εισάγεις τον κώδικα σε αυτό. Χρειάστηκε να κατεβάσουμε κάποια libraries, το gym_super_mario_bros και το nes_py .

Για να ρυθμίσουμε τον Super Mario, χρειάστηκε να εισάγουμε το παιχνίδι και να επιλέξουμε τις κινήσεις που θα μπορεί να κάνει από μια λίστα κινήσεων (RIGHT_ONLY, SIMPLE_MOVEMENT, COMPLEX_MOVEMENT). Εμείς επιλέξαμε το SIMPLE_MOVEMENT, όπου περιλαμβάνει 7 κινήσεις.

```
# Import the game
import gym_super_mario_bros
# Import the Joypad wrapper
from nes_py.wrappers import JoypadSpace
# Import the SIMPLIFIED controls
from gym_super_mario_bros.actions import SIMPLE_MOVEMENT
```

Εικόνα 8. Κομμάτι Αλγόριθμου

```
[['NOOP'],  
 ['right'],  
 ['right', 'A'],  
 ['right', 'B'],  
 ['right', 'A', 'B'],  
 ['A'],  
 ['left']]
```

Εικόνα 9. Κινήσεις SIMPLE_MOVEMENT

Ακολούθως, επιλέγουμε το περιβάλλον στο οποίο θέλουμε να διαδραματίζεται το παιχνίδι, με βάση το `gym_super_mario_bros`. από μια λίστα 6 περιβαλλόντων, εμείς επιλέξαμε το πρώτο που είναι και το πιο συνηθισμένο (`SuperMarioBros-v0`).

```
# Setup game  
env = gym_super_mario_bros.make('SuperMarioBros-v0')  
env = JoypadSpace(env, SIMPLE_MOVEMENT)
```

Εικόνα 10. Κομμάτι Αλγόριθμου

Για να αρχίσει το παιχνίδι, δημιουργούμε έναν βρόγχο ο οποίος θα περνά από κάθε πλαίσιο και ο Mario θα μπορεί να κάνει τυχαίες κινήσεις ούτως ώστε να περάσει τα εμπόδια. Αυτό επιτυγχάνετε με την εντολή `env.action_space.sample()`.

```
# Create a flag - restart or not  
done = True  
# Loop through each frame in the game  
for step in range(100000):  
    # Start the game to begin with  
    if done:  
        # Start the gamee  
        env.reset()  
        # Do random actions  
        state, reward, done, info = env.step(env.action_space.sample())  
        # Show the game on the screen  
        env.render()  
    # Close the game  
    env.close()
```

Εικόνα 10. Κομμάτι Αλγόριθμου

Ακολουθούν οι εντολές για την προεπεξεργασία του περιβάλλοντος.

Με την ακόλουθη εντολή, ***from gym.wrappers import GrayScaleObservation***, επιτρέπουμε στο λογισμικό να αποθηκεύσει κάποια στιγμιότυπα ούτως ώστε το μοντέλο AI να μπορέσει να δει πως κινήθηκε ο Mario, πως κινήθηκε ο εχθρός και πως αλληλεπιδρά με το περιβάλλον, ώστε να μπορεί να εκπαιδευτεί και να κινείται με καταλληλότερες κινήσεις.

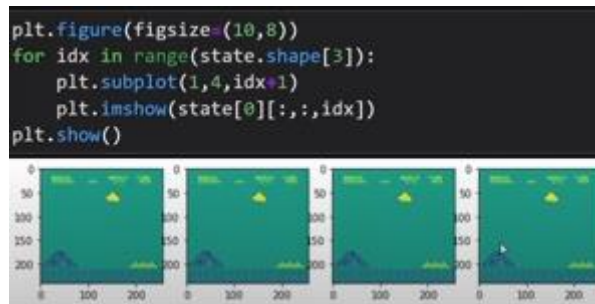
Έπειτα, δουλεύουμε με τα στοιβαγμένα πλαίσια περιβάλλοντος (VecFrameStack) και το βασικό περιβάλλον (DummyVecEnv). Και με την επόμενη εντολή θα μπορούμε να δούμε όλα τα στιγμιότυπα πλαισίων.

```
# Import Frame Stacker Wrapper and GrayScaling Wrapper
from gym.wrappers import GrayScaleObservation
# Import Vectorization Wrappers
from stable_baselines3.common.vec_env import VecFrameStack, DummyVecEnv
# Import Matplotlib to show the impact of frame stacking
from matplotlib import pyplot as plt
```

Μετά δημιουργούμε το βασικό περιβάλλον, απλοποιούμε τους ελέγχους, μετατρέπουμε το περιβάλλον σε grayscale, το περιτυλίγουμε σε ένα ψεύτικο περιβάλλον και τέλος στοιβάζουμε τα πλαίσια.

```
# 1. Create the base environment
env = gym_super_mario_bros.make('SuperMarioBros-v0')
# 2. Simplify the controls
env = JoypadSpace(env, SIMPLE_MOVEMENT)
# 3. Grayscale
env = GrayScaleObservation(env, keep_dim=True)
# 4. Wrap inside the Dummy Environment
env = DummyVecEnv([lambda: env])
# 5. Stack the frames
env = VecFrameStack(env, 4, channels_order='last')
```

Ακολούθως, για να δούμε τις τελευταίες 4 στιβάδες πλαισίων εισάγουμε τις ακόλουθες γραμμές κώδικα βλέποντας τις τελευταίες κινήσεις του.



Για να εκπαιδεύσουμε το Reinforcement μοντέλο, πρέπει να λάβουμε υπόψη τις 4 λέξεις κλειδιά: πράκτορας, περιβάλλον, ανταμοιβή, και κίνηση. Ο πράκτοράς μας σε αυτή την περίπτωση είναι ο Super Mario, ο οποίος μπορεί να κάνει κάποιες κινήσεις όπως να πηδήσει, να κινηθεί δεξιά, αριστερά κτλ., στο περιβάλλον του παιχνιδιού και αναλόγως των αποτελεσμάτων ή θα ανταμειφθεί ή θα λάβει ποινή. Το μοντέλο AI, εκπαιδεύεται ώστε να κάνει τις κατάλληλες κινήσεις κάθε φορά ώστε να μεγιστοποιήσει την ανταμοιβή.

Έγινε χρήση του PPO (proximal policy optimization algorithm) για να μπορεί να εκπαιδευτεί το μοντέλο AI ή το μοντέλο Reinforcement Learning. Αποθηκεύονται τα διάφορα μοντέλα τα οποία δημιουργήθηκαν έχοντας ένα backup κάθε ένα αριθμό βημάτων, ώστε να μπορεί να εκπαιδευτεί ο πράκτορας κατάλληλα.

```
# Import os for file path management
import os
# Import PPO for algos
from stable_baselines3 import PPO
# Import Base Callback for saving models
from stable_baselines3.common.callbacks import BaseCallback
```


Με τις ακόλουθες γραμμές κώδικα επανακαλείς κάτι το οποίο σου επιτρέπει να αποθηκεύεις το μοντέλο κάθε x βήματα. Αυτό μας επιτρέπει να εισάγουμε 'το πόσο συχνά θέλουμε να αποθηκεύετε το μοντέλο καθώς και το που.

```
class TrainAndLoggingCallback(BaseCallback):

    def __init__(self, check_freq, save_path, verbose=1):
        super(TrainAndLoggingCallback, self).__init__(verbose)
        self.check_freq = check_freq
        self.save_path = save_path

    def _init_callback(self):
        if self.save_path is not None:
            os.makedirs(self.save_path, exist_ok=True)

    def _on_step(self):
        if self.n_calls % self.check_freq == 0:
            model_path = os.path.join(self.save_path, 'best_model_{}'.format(self.n_calls))
            self.model.save(model_path)

        return True
```

Δημιουργούμε τους φακέλους που θα αποθηκεύονται τα μοντέλα καθώς και τα logs.

```
CHECKPOINT_DIR = './train/'
LOG_DIR = './logs/'
```

Στη συνέχεια, δημιουργούμε ένα προσωρινό μοντέλο AI το οποίο θα μας βοηθήσει να ξεκινήσουμε να εκπαιδεύουμε το μοντέλο.

```
# This is the AI model started
model = PPO('CnnPolicy', env, verbose=1, tensorboard_log=LOG_DIR, learning_rate=0.000001,
            n_steps=512)
```

Ακολουθούν οι εντολές για το που θα ξεκινήσει το μοντέλο AI να εκπαιδεύεται, μετά από πόσα πλαίσια.

```
# Train the AI model, this is where the AI model starts to learn
model.learn(total_timesteps=1000000, callback=callback)
```

Φτάνοντας στο τέλος, πρέπει να φορτώσουμε το μοντέλο που θέλουμε και ξεκινάμε το παιχνίδι. Πριν χρησιμοποιούσαμε τυχαίες κινήσεις, ενώ τώρα θα προβλέψει τις καταλληλότερες κινήσεις με βάση το μοντέλο που αποθηκεύσαμε και φορτώσαμε.

```
# Start the game
state = env.reset()
# Loop through the game
while True:

    action, _ = model.predict(state)
    state, reward, done, info = env.step(action)
    env.render()
```

Επίλογος

Στις μέρες μας, η ανάπτυξη της τεχνολογία βλέπουμε πως είναι ραγδαία. Κυρίως, η Τεχνητή Νοημοσύνη, έχει μπει για τα καλά στις ζωές μας, κάτι που έχει επηρεάσει σε μεγάλο βαθμό θετικά. Έχει καταφέρει να εισέλθει σε πολλούς τομείς, ιδιαίτερα στις επιχειρήσεις και στην υγεία, αλλά και σε πιο απλά πράγματα της καθημερινότητάς μας, όπως παρατηρείται. Έχουν ενταχθεί στα τηλέφωνα, στις εφαρμογές, στα αυτοκίνητα και η εξέλιξη τους είναι συνεχείς και αναπάντεχη. Πλέον έχουν αυτοματοποιηθεί σχεδόν τα πάντα. Ακόμη και τα παιχνίδια έχουν εκπαιδευτεί με την χρήση του Deep Learning και του Reinforcement Learning, να παίζουν με τις δικές τους κινήσεις, μαθαίνοντας από τα λάθη τους, χωρίς να χρειάζονται την ανθρώπινη παρέμβαση. Ας ελπίσουμε πως όλη αυτή η συνεχής ανάπτυξη θα μας βγει σε καλό.

Βιβλιογραφία

1. Bougie, Nicolas; Ichise, Ryutaro. "Fast and slow curiosity for high-level exploration in reinforcement learning." (2020).
2. Liu, Jialin; Snodgrass, Sam; Khalifa, Ahmed; Risi, Sebastian; Yannakakis, Georgios N.; Togelius, Julian. "Deep learning for procedural content generation." (2020).
3. Chen, D., Wang, Y., & Gao, W. "Combining a gradient-based method and an evolution strategy for multi-objective reinforcement learning." (2020).
4. <https://youtu.be/2eeYqJ0uBKE>
5. <https://www.educba.com/what-is-neural-networks/>
6. <https://www.sciencedirect.com/topics/engineering/genetic-algorithm>
7. <https://becominghuman.ai/where-is-artificial-intelligence-used-today-3fd076d15b68>
8. <https://www.educba.com/what-is-reinforcement-learning/>
9. <https://marketbusinessnews.com/financial-glossary/artificial-intelligence/>
10. <https://bernardmarr.com/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/>
11. <https://www.guru99.com/reinforcement-learning-tutorial.html>
12. <https://www.google.com/search?q=reinforcement+learning+and+examples&oq=reinforcement+learning+and+examples&aqs=chrome..69i57j0i22i30l9.20057j0j4&sourceid=chrome&ie=UTF-8>
13. <https://www.geeksforgeeks.org/genetic-algorithms/>

14. <https://marketbusinessnews.com/financial-glossary/neural-networks/>
15. <https://www.upgrad.com/blog/neural-networks-applications-in-the-real-world/#:~:text=For%20example%2C%20a%20neural%20network,Pattern%20Recognition%2C%20Classification%20and%20Optimization.>
16. <https://www.potentiaco.com/what-is-machine-learning-definition-types-applications-and-examples/>