



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

UNIVERSITY OF PIRAEUS

Τμήμα Ψηφιακών Συστημάτων  
Πληροφοριακά Συστήματα και Υπηρεσίες  
Μεγάλα Δεδομένα και Αναλυτική

**Χρήση τεχνητών νευρωνικών δικτύων για τον σχεδιασμό  
στρατηγικής στην αγορά συναλλάγματος**

Μιλτιάδης Μαγκώνης  
Επιβλέπων Καθηγητής: Δημοσθένης Κυριαζής

Πειραιάς  
2022



## Περίληψη

Τη σύγχρονη εποχή, το πεδίο της τεχνητής νοημοσύνης γνωρίζει σημαντική άνθηση σε κάθε κλάδο της καθημερινότητας. Ένας από τους κλάδους είναι οι χρηματοοικονομικές αγορές. Η χρήση τεχνητής νοημοσύνης έχει συνεισφέρει σημαντικά στην πρόβλεψη τιμών στην αγορά συναλλάγματος και στο χρηματιστήριο. Η παρούσα εργασία ερευνά και αναλύει τη χρήση αλγορίθμων μηχανικής μάθησης για την επίτευξη μιας κερδοφόρας στρατηγικής στην αγορά συναλλάγματος επικεντρώνοντας στην ισοτιμία Ευρώ – αμερικανικού Δολαρίου.

# Περιεχόμενα

1. Εισαγωγή.....	4
2. Μοντέλα Χρονοσειρών για την Πρόβλεψη Ισοτιμιών.....	5
2.1. Περιγραφή.....	5
2.2. Απλός κινητός μέσος.....	6
3. Μηχανική μάθηση.....	8
3.1. Εισαγωγικά στοιχεία.....	8
3.2. Μηχανική Μάθηση, Βαθιά Μάθηση και Νευρωνικά Δίκτυα.....	10
3.3. Η λειτουργία της μηχανικής μάθησης.....	12
3.4. Μέθοδοι μηχανικής μάθησης.....	12
4. Τεχνητά Νευρωνικά Δίκτυα.....	14
4.1. Εισαγωγικά στοιχεία.....	14
4.2. Μοντέλο McCulloch Pitts.....	15
4.3. Ο αλγόριθμος Perceptron.....	17
4.4. Ο αλγόριθμος Multi-layer Perceptron.....	19
5. Χρήση Νευρωνικών Δικτύων σε Πραγματικά Δεδομένα.....	28
5.1. Καθορισμός Στρατηγικής Βάσει Απλού Κινητού Μέσου.....	28
5.2. Καθορισμός Στρατηγικής Βάσει Διαφορετικών Τεχνικών Δεικτών.....	38
6. Αποτελέσματα.....	46
7. Συμπεράσματα.....	48
8. Βιβλιογραφία.....	49

# 1. Εισαγωγή

Στο πεδίο των χρηματοοικονομικών, οι ερευνητές, προσπαθούσαν για δεκαετίες να εκτιμήσουν με μαθηματικά μοντέλα τη συμπεριφορά των επενδυτών, και συνακόλουθα των αγορών, με τον ίδιο τρόπο που οι φυσικοί, παραδείγματος χάρη, αναλύουν περίπλοκα φυσικά φαινόμενα. Η επιστήμη, για αιώνες, είναι ταυτόσημη με την παραγωγή και την ανάλυση δεδομένων. Από την άλλη, το οικονομικό πεδίο χαρακτηριζόταν στο παρελθόν από θεωρίες οι οποίες είναι βασισμένες σε σχετικά απλά μαθηματικά μοντέλα, και στηριζόταν εν πολλοίς σε υποθέσεις γραμμικών σχέσεων.

Πλέον, λόγω της υπερπληθώρας των διαθέσιμων δεδομένων, έχει μετατοπιστεί το ενδιαφέρον που υπήρχε παλαιότερα για μια προσέγγιση κατά πρώτο λόγο θεωρητική, προς μια κατεύθυνση όπου η ανάλυση δεδομένων θα έχει την πρωτοκαθεδρία. Έχουν υπάρξει πολλά παραδείγματα άλλωστε όπου πολλά γνωστά οικονομικά μοντέλα και θεωρίες δεν κατάφεραν να εξηγήσουν πολλά από τα οικονομικά φαινόμενα που έχουν εμφανιστεί. Αυτό παρατηρείται γιατί λειτουργούν με απλουστεύσεις και αδυνατούν να ερμηνεύσουν την πολυπλοκότητα και τη συνεχή μεταβλητότητα των χρηματαγορών. Το θεωρητικό υπόβαθρο της ανάλυσής μας βασίζεται σε μια υπόθεση που έχει ισχυρή εμπειρική βάση. Αυτό το θεωρητικό υπόδειγμα ονομάζεται υπόθεση της αποτελεσματικής αγοράς (Efficient Market Hypothesis).

Σε βασικές γραμμές, το θεώρημα αναφέρει ότι οι αγορές εμπεριέχουν το σύνολο της πληροφορίας, δηλαδή ότι οι τιμές των χρηματοπιστωτικών προϊόντων αντανακλούν το σύνολο των πληροφοριών την κάθε δεδομένη στιγμή. Συνακόλουθα μπορούμε να παράγουμε το συμπέρασμα ότι κάποιος επενδυτής δεν μπορεί να νικήσει την αγορά χρησιμοποιώντας κάποια πρόσθετη πληροφορία, καθώς όπως προαναφέραμε όλες οι πληροφορίες έχουν ενσωματωθεί στην τιμή του χρηματοπιστωτικού προϊόντος. Με άλλα λόγια, αν υποθέσουμε ότι η προαναφερθείσα θεωρία ανταποκρίνεται στην πραγματικότητα, τότε οποιαδήποτε μορφή ανάλυσης δεδομένων και έρευνας για την πραγματοποίηση υπερβολικών αποδόσεων/κερδών, θεωρείται άσκοπη.

Για παράδειγμα η τιμή μιας μετοχής ή οποιουδήποτε άλλου χρηματοπιστωτικού προϊόντος έχει τόσες πιθανότητες να σημειώσει άνοδο, όσες έχει και να σημειώσει πτώση. Με αυτό ως δεδομένο, μπορούμε να συμπεράνουμε ότι ο καλύτερος εκτιμητής της αυριανής τιμής μιας μετοχής, είναι η τιμή της στον παρόντα χρόνο, δηλαδή η κατανομή της μελλοντικής τιμής της είναι ανεξάρτητη από τη παρελθούσα διακύμανσή της. Εξαρτάται μόνο από τη σημερινή τιμή. Ως εκ τούτου, σε αυτό το πλαίσιο, η ανάλυση των παρελθουσών τιμών για την πρόβλεψη των μελλοντικών καθίσταται αδύνατη. Η εν λόγω υπόθεση χρονολογείται από τις δεκαετίες του 1960 και 1970, σε περιόδους δηλαδή όπου τα δεδομένα στο τέλος της κάθε ημέρας ήταν τα μόνα διαθέσιμα.

Η παρούσα εργασία πραγματεύεται την επιβεβαίωση της προαναφερθείσας θεωρίας. Συγκεκριμένα θα ερευνήσουμε αν οι αλγόριθμοι μηχανικής μάθησης όπως είναι τα τεχνητά νευρωνικά δίκτυα μπορούν να αποτελέσουν αποδοτικό εργαλείο για την εύρεση μοτίβων που θα μας οδηγήσουν σε μια κερδοφόρα στρατηγική σε σχέση με παλαιότερες οικονομετρικές μεθόδους όπως αυτή της γραμμικής παλινδρόμησης.

## **2. Μοντέλα Χρονοσειρών για την Πρόβλεψη Ισοτιμιών**

### **2.1. Περιγραφή**

Η συναλλαγματική ισοτιμία αντιπροσωπεύει την τιμή ενός νομίσματος σε σχέση με την τιμή ενός άλλου. Δηλαδή την αναλογία με την οποία ένα νόμισμα μπορεί να ανταλλαχθεί με κάποιο άλλο. Η πρόβλεψη της ισοτιμίας είναι ένα πολύ σημαντικό και απαιτητικό αντικείμενο στον κλάδο των χρηματοπιστωτικών αγορών, ιδιαίτερα μετά την κατάρρευση του συστήματος σταθερών ισοτιμιών Bretton Woods το 1973. Πλέον η απόκτηση αξιόπιστων προβλέψεων των ισοτιμιών κρίνεται ιδιαίτερα σημαντική για την επίτευξη κερδών, για τον καθορισμό της βέλτιστης νομισματικής πολιτικής από την πλευρά των κυβερνήσεων, όπως επίσης για η λήψη αποφάσεων από τις ιδιωτικές εταιρείες.

Κυρίως δηλαδή για τους οργανισμούς οι οποίοι επιχειρούν σε διάφορα μέρη του πλανήτη και συγκεντρώνουν κεφάλαια σε μια πληθώρα νομισμάτων. Αυτοί οι οργανισμοί χρησιμοποιούν τις προβλέψεις συναλλαγματικών ισοτιμιών για ένα πλήθος λειτουργιών όπως είναι η επιλογή του νομίσματος για την τιμολόγηση, ο δανεισμός, η διαχείριση ανοικτών συναλλαγματικών θέσεων. Η αγορά στην οποία συντελούνται αγοραπωλησίες νομισμάτων ονομάζεται αγορά Forex. Οι συμμετέχοντες είναι ιδιωτικές και κεντρικές τράπεζες κρατών, ιδιωτικές εταιρείες, hedge funds οργανισμοί και μεμονωμένοι επενδυτές. Η πρόβλεψη των βραχυπρόθεσμων διακυμάνσεων και η κατεύθυνση μεταβολής της ισοτιμίας είναι ιδιαιτέρως σημαντική για όλους τους συμμετέχοντες.

Η πρόβλεψη της ισοτιμίας μπορεί να γίνει χρησιμοποιώντας μοντέλα στα οποία η τιμή της συνδέεται με μεταβλητές όπως είναι η προσφορά χρήματος, η εισροή και εξαγωγή κεφαλαίου, ο πληθωρισμός, το τραπεζικό επιτόκιο, το εμπορικό ισοζύγιο κ.α. Όμως υπάρχουν δεδομένοι περιορισμοί κατά τη διαδικασία αυτή και το μοντέλο δεν είναι ιδιαίτερα αποδοτικό σε αυτό τον κλάδο της αγοράς. Αυτό συμβαίνει γιατί τα μακροοικονομικά δεδομένα αυτών των μεταβλητών είναι διαθέσιμα, στην καλύτερη περίπτωση, μόνο σε μηνιαία βάση ενώ το ενδιαφέρον για την πρόβλεψη της ισοτιμίας αφορά πολύ μικρότερα χρονικά πλαίσια (πχ ημερήσια ή ωριαία βάση). Για την αποφυγή τέτοιων εμποδίων χρησιμοποιούμε μοντέλα που επιχειρούν να προβλέψουν χρηματοοικονομικές μεταβλητές χρησιμοποιώντας ως μοναδική πληροφορία τις παρελθούσες τιμές τους. Τέτοια μοντέλα χρονοσειρών είναι ο απλός κινητός μέσος (SMA), η εκθετική εξομάλυνση (ES), και τα αυτοπαλινδρομικά μοντέλα κινητού μέσου (ARIMA ).

## **2.2. Απλός κινητός μέσος**

Η μέθοδος του απλού κινητού μέσου υπολογίζει τη μέση τιμή ενός καθορισμένου αριθμού τιμών.

$$\hat{Y}_{t+1} = \frac{\sum_{k=1}^r Y_{t-k+1}}{r}$$

Όπου:

$\hat{Y}_{t+1}$  Η τιμή της πρόβλεψης για την περίοδο  $t + 1$

$Y_t$  Η τιμή της περιόδου  $t$

$r$  ο αριθμός των περιόδων

Μπορούμε να εφαρμόσουμε δύο τύπους απλών κινητών μέσων. Έναν βραχυπρόθεσμο και έναν μακροπρόθεσμο. Ένας συνδυασμός τους μπορεί να καθορίσει την στρατηγική μας σχετικά με το αν θα προχωρήσουμε στην αγορά ή την πώληση ενός νομίσματος. Ένα κρίσιμο χαρακτηριστικό του παραπάνω μοντέλου είναι ότι οι εκτιμήσεις του είναι, σχεδόν πάντα, πίσω χρονικά από την παρούσα αξία. Όταν η ισοτιμία έχει την τάση να μειωθεί, τα δεδομένα του απλού κινητού μέσου έχουν υψηλότερη τιμή. Αναλόγως όταν η ισοτιμία ανεβαίνει, τότε ο SMA θα δίνει χαμηλότερες τιμές.



## 3. Μηχανική μάθηση

### 3.1. Εισαγωγικά στοιχεία

Η μηχανική μάθηση είναι ένας κλάδος της τεχνητής νοημοσύνης και της επιστήμης των υπολογιστών ο οποίος εστιάζει στη χρήση δεδομένων και αλγορίθμων για την προσομοίωση της διαδικασίας και του τρόπου με τον οποίο ένας ανθρώπινος εγκέφαλος εκπαιδεύεται. Εγκολπώνει στοιχεία από τις νευροεπιστήμες, τη βιολογία, τη στατιστική, τα μαθηματικά, τη φυσική. Σημαντικό κομμάτι της εκμάθησης είναι η βελτιστοποίηση στην επίλυση προβλημάτων. Ειδικά κατά τη διάρκεια των τελευταίων δύο δεκαετιών, η τεχνολογική πρόοδος που οδήγησε στην αύξηση του αποθηκευτικού χώρου για τα δεδομένα και η αύξηση στην υπολογιστική ισχύ των υπολογιστών έχουν βοηθήσει σημαντικά στη δημιουργία προϊόντων τα οποία είναι βασισμένα στον κλάδο της μηχανικής μάθησης, όπως είναι τα “έξυπνα” οχήματα χωρίς οδηγό.

Η μηχανική μάθηση είναι συστατικό στοιχείο του πεδίου της επιστήμης δεδομένων. Με τη χρήση στατιστικών μεθόδων, συντελείται η εκπαίδευση των αλγορίθμων με σκοπό την πρόβλεψη και την κατάταξη αναλόγως με τον τύπο του προβλήματος προς επίλυση. Επίσης το πεδίο της μηχανικής μάθησης έχει πολύ σημαντικό ρόλο στη διαδικασία εξόρυξης δεδομένων, δηλαδή τη διαδικασία εξαγωγής χρήσιμης πληροφορίας από σύνολα δεδομένα τεράστιου όγκου.

Κάποιες από τις χρήσεις της μηχανικής μάθησης στην καθημερινότητα είναι οι ακόλουθες. Η αυτόματη αναγνώριση ομιλίας όπου χρησιμοποιεί την επεξεργασία φυσικής γλώσσας έτσι ώστε να μεταφράσει την ανθρώπινη ομιλία σε γραπτό κείμενο. Πολλές ηλεκτρονικές συσκευές χρησιμοποιούν τη συγκεκριμένη δυνατότητα στα συστήματά τους.

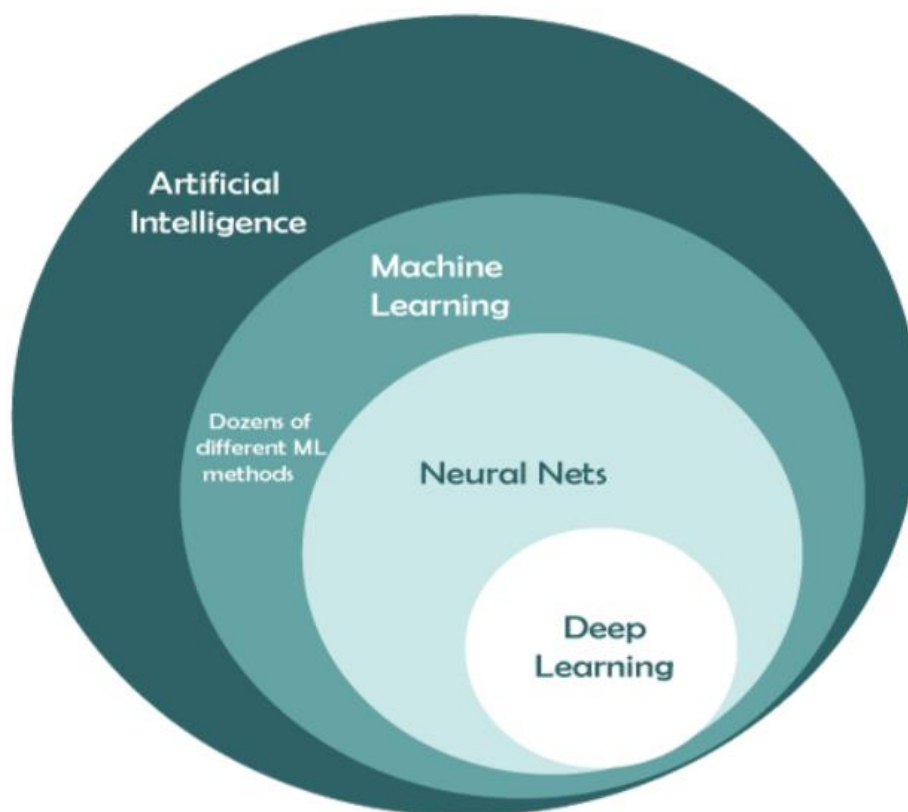
Μια άλλη χρήση είναι η εξυπηρέτηση πελατών όπου chatbots αντικαθιστούν τον άνθρωπο. Τα συγκεκριμένα chatbots απαντούν στις πιο συχνές ερωτήσεις των πελατών ή τους βοηθούν δίνοντας προσωποποιημένη ενημέρωση σχετικά με το χρόνο παράδοσης των προϊόντων ή συμβουλεύοντάς τους σχετικά με τα πιο συμβατά και κατάλληλα γι' αυτούς προϊόντα, αναλόγως τις προτιμήσεις τους. Άλλη χρήση είναι η υπολογιστική όραση η οποία επιτρέπει στους υπολογιστές να εξάγουν χρήσιμες πληροφορίες από ψηφιακές εικόνες, βίντεο, και άλλες οπτικοακουστικές πηγές. Συγκεκριμένα μέσω των συνελκτικών νευρωνικών δικτύων, η υπολογιστική όραση μπορεί να βοηθήσει σε προβλήματα ιατρικής φύσεως (αξιολόγηση ακτινογραφιών), ή στην ανάπτυξη έξυπνων οχημάτων στην αυτοκινητοβιομηχανία.

Άλλη χρήση είναι στην παρελθούσα συμπεριφορά των καταναλωτών όπου χρησιμοποιείται από τους αλγόριθμους τεχνητής νοημοσύνης για τη εξαγωγή στοιχείων καταναλωτικών τάσεων και την ανάπτυξη πιο αποδοτικών στρατηγικών για δημιουργία κατάλληλων συστάσεων αγοράς προς αυτούς όπως επίσης για τη βελτιστοποίηση των προϊόντων. Μια άλλη σημαντική χρήση είναι ο σχεδιασμός χαρτοφυλακίου μετοχών και άλλων χρηματοπιστωτικών προϊόντων και η συνακόλουθη διαχείριση εκατομμυρίων συναλλαγών καθημερινώς χωρίς την ανθρώπινη παρέμβαση με σκοπό τη μεγιστοποίηση των αποδόσεων και του κέρδους.

Επιπροσθέτως, όσον αφορά τον τραπεζικό τομέα, οι τράπεζες και άλλοι χρηματοπιστωτικοί οργανισμοί χρησιμοποιούν τη μηχανική μάθηση για την ανίχνευση ύποπτων-έκνομων συναλλαγών. Οι αλγόριθμοι επιβλεπόμενης μάθησης, που θα αναλύσουμε παρακάτω, μπορούν να εκπαιδεύσουν ένα μοντέλο χρησιμοποιώντας δεδομένα ύποπτων συναλλαγών για την περαιτέρω διερεύνηση μελλοντικών παράταιρων και μη συνηθισμένων συναλλαγών.

### 3.2. Μηχανική Μάθηση, Βαθιά Μάθηση και Νευρωνικά Δίκτυα

Η μηχανική μάθηση, η βαθιά μάθηση και τα τεχνητά νευρωνικά δίκτυα αποτελούν υποκατηγορίες του πεδίου της τεχνητής νοημοσύνης. Συγκεκριμένα τα τεχνητά νευρωνικά δίκτυα αποτελούν μια υποκατηγορία της τεχνητής μάθησης, και η βαθιά μάθηση αποτελεί υποκατηγορία των τεχνητών νευρωνικών δικτύων. Η διαφορά ανάμεσα στη βαθιά και τη μηχανική μάθηση βασίζεται στον τρόπο με τον οποίο κάθε αλγόριθμος εκπαιδεύεται. Η βαθιά μάθηση μπορεί να χρησιμοποιήσει κατηγοριοποιημένα αλλά και μη κατηγοριοποιημένα σύνολα δεδομένων για την εκπαίδευση του αλγόριθμου.



Εικόνα 1. Υποκατηγορίες της Τεχνητής Νοημοσύνης

Ένας τέτοιος αλγόριθμος μπορεί να τροφοδοτηθεί με μη δομημένα δεδομένα (πχ εικόνες ή κείμενο) και αυτομάτως να μπορεί να προσδιορίσει εκείνο το σύνολο στοιχείων στο οποίο οφείλεται η διαφοροποίηση των δεδομένων σε διακριτές κατηγορίες. Αυτό μπορεί να περιορίσει την ανθρώπινη παρέμβαση ή βοήθεια για την επίλυση του προβλήματος και είναι η κατάλληλη διαδικασία για τη διαχείριση ενός μεγάλου όγκου δεδομένων.

Η κλασική “μη-βαθιά” μηχανική μάθηση είναι περισσότερο εξαρτώμενη από την ανθρώπινη παρέμβαση κατά την οποία καθορίζεται το πλήθος των χαρακτηριστικών από τα οποία θα κατανοηθούν οι διαφορές ανάμεσα στα στοιχεία του συνόλου δεδομένων προς επεξεργασία. Τα τεχνητά νευρωνικά δίκτυα αποτελούνται από επίπεδα κόμβων ή αλλιώς νευρώνων. Συγκεκριμένα ένα επίπεδο εισόδου των δεδομένων, ένα ή περισσότερα κρυφά επίπεδα και τέλος ένα επίπεδο εξόδου. Κάθε νευρώνας-κόμβος ενός επιπέδου συνδέεται με όλους τους νευρώνες του προηγούμενου επιπέδου. Κάθε σύνδεση χαρακτηρίζεται από ένα συγκεκριμένο βάρος (weight) και από μια συγκεκριμένη τιμή-κατώφλι (threshold). Αν το αποτέλεσμα ενός οποιουδήποτε κόμβου είναι πάνω από την προαναφερθείσα τιμή, τότε εκείνος ο νευρώνας ενεργοποιείται, στέλνοντας δεδομένα στο επόμενο επίπεδο του τεχνητού νευρωνικού δικτύου. Στην αντίθετη περίπτωση, τα δεδομένα δεν διαπερνούν το επίπεδο από τον συγκεκριμένο νευρώνα.

Συνεπώς, η λέξη βαθιά που χαρακτηρίζει το συγκεκριμένο είδος της μηχανικής μάθησης έχει να κάνει με το πόσο βαθύ είναι το νευρωνικό δίκτυο, κάτι που καθορίζεται από το βαθμό των επιπέδων από το οποίο αποτελείται. Ένα νευρωνικό δίκτυο που αποτελείται από παραπάνω από τρία επίπεδα μπορεί να θεωρηθεί ότι αποτελεί έναν αλγόριθμο βαθιάς μηχανικής μάθησης. Το νευρωνικό δίκτυο που αποτελείται από μόλις τρία επίπεδα χαρακτηρίζεται ως απλό.

Στη βαθιά μάθηση και στα τεχνητά νευρωνικά δίκτυα βασίζεται η ολοένα και μεγαλύτερη πρόοδος σε πεδία όπως η μηχανική ή υπολογιστική όραση (computer vision), η επεξεργασία φυσικής γλώσσας (natural language processing) και η αναγνώριση ομιλίας (speech recognition).

### 3.3. Η λειτουργία της μηχανικής μάθησης

Η διαδικασία της λειτουργίας ενός αλγορίθμου μηχανικής μάθησης χωρίζεται σε τρία στάδια. Πρώτον είναι η διαδικασία της απόφασης (decision process). Σε γενικές γραμμές οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται για την πρόβλεψη ή για την κατάταξη δεδομένων σε κατηγορίες. Με βάση τα τροφοδοτούμενα δεδομένα, ο αλγόριθμος θα παράγει μια πρώτη εκτίμηση για το πως μπορούν αυτά να κατηγοριοποιηθούν. Στη συνέχεια μια συνάρτηση κόστους-λάθους (error function) κάνει εκτίμηση στην πρόβλεψη που παράγει το μοντέλο. Μετέπειτα αυτή η συνάρτηση προχωρά σε σύγκριση της πρόβλεψης με το επιθυμητό αποτέλεσμα, για την εκτίμηση της αποδοτικότητας του μοντέλου.

Τέλος είναι η διαδικασία της παραμετροποίησης-βελτιστοποίησης του μοντέλου. Αν το μοντέλο ανταποκρίνεται σωστά στα δεδομένα εισόδου του συνόλου δεδομένων εκπαίδευσης, τότε τα βάρη παραμετροποιούνται έτσι ώστε να μειώσουν τη διάσταση-διαφορά ανάμεσα στην ορθή κατηγοριοποίηση ή πρόβλεψη και στην εκτίμηση του μοντέλου. Ο αλγόριθμος επαναλαμβάνει συνεχώς αυτή τη διαδικασία της αξιολόγησης-παραμετροποίησης, παραμετροποιώντας τα βάρη αυτόματα μέχρι το επιθυμητό επίπεδο αποδοτικότητας-ακρίβειας.

### 3.4. Μέθοδοι μηχανικής μάθησης

Τα μοντέλα μηχανικής μάθησης χωρίζονται σε διάφορες κατηγορίες. Πρώτα θα μιλήσουμε για την επιβλεπόμενη μηχανική μάθηση (supervised machine learning). Η συγκεκριμένη διαδικασία χρησιμοποιεί κατηγοριοποιημένα δεδομένα για την εκπαίδευση του αλγορίθμου για την κατάταξή τους ή για την αποδοτική πρόβλεψη. Κατά την είσοδο των δεδομένων στο μοντέλο, αυτό παραμετροποιεί τα βάρη.

Δηλαδή έχουμε ένα σύνολο δεδομένων προς εκπαίδευση μαζί με τις σωστές απαντήσεις-αποτελέσματα, και βάσει αυτών των δεδομένων, ο αλγόριθμος παραμετροποιείται έτσι ώστε μπορεί να ανταποκριθεί αποδοτικά στο να προβλέπει τα αντίστοιχα ορθά αποτελέσματα ανεξαρτήτως των νέων δεδομένων εισόδου.

Γνωστοί αλγόριθμοι μηχανικής μάθησης είναι τα τεχνητά νευρωνικά δίκτυα, ο Naive Bayes, η γραμμική παλινδρόμηση, η λογιστική παλινδρόμηση, οι ταξινομητές δένδρων αποφάσεων, (μηχανές διανυσμάτων υποστήριξης) support vector machines κ.α. Κατά τη διαδικασία μη επιβλεπόμενης μάθησης, οι αλγόριθμοι αναλύουν και κατηγοριοποιούν μη κατηγοριοποιημένα δεδομένα. Οι συγκεκριμένοι αλγόριθμοι εντοπίζουν κρυφά, μη ορατά, μοτίβα στα δεδομένα χωρίς την ανθρώπινη παρέμβαση. Χρησιμοποιούνται επίσης για τη μείωση των χαρακτηριστικών σε ένα μοντέλο μέσω μιας διαδικασίας που ονομάζεται μείωση διαστάσεων. Ο αλγόριθμος PCA (Principal Component Analysis), ο SVD (Singular Value Decomposition) και οι αλγόριθμοι συσταδοποίησης όπως ο K-means Clustering είναι χαρακτηριστικά δείγματα αλγορίθμων μη επιβλεπόμενης μάθησης.

Η ημι-επιβλεπόμενη μηχανική μάθηση θεωρείται η μέση λύση ανάμεσα στις δύο προαναφερθείσες κατηγορίες. Κατά τη διάρκεια της εκπαίδευσης, χρησιμοποιεί ένα μικρότερο μοντέλο κατηγοριοποιημένων δεδομένων για να συνδράμει στην ταξινόμηση και την εξαγωγή των κατάλληλων χαρακτηριστικών από ένα ογκωδέστερο σύνολο μη κατηγοριοποιημένων δεδομένων. Οι αλγόριθμοι αυτού του είδους είναι ικανοί να λύσουν προβλήματα όταν δεν υπάρχει ικανό πλήθος κατηγοριοποιημένων δεδομένων για να τροφοδοτηθεί και να εκπαιδευτεί με αυτά ο αλγόριθμος. Χρησιμοποιούνται επίσης όταν το υπολογιστικό κόστος για την κατηγοριοποίηση των μοντέλων εισόδου κρίνεται πολύ μεγάλο.

Το μοντέλο της ενισχυτικής μηχανικής μάθησης είναι παρόμοιο με εκείνα της επιβλεπόμενης μηχανικής μάθησης με τη διαφορά ότι οι εν λόγω αλγόριθμοι δεν εκπαιδεύονται με τη χρήση δείγματος δεδομένων. Το μοντέλο αυτό εκπαιδεύεται μέσω της διαδικασίας δοκιμής-λάθους. Δηλαδή γνωστοποιείται στον αλγόριθμο ότι κάνει λάθος, αλλά δεν γνωστοποιείται πως μπορεί να το διορθώσει.

Θα πρέπει να εξερευνήσει και να δοκιμάσει μια πληθώρα επιλογών μέχρι να μάθει πως θα βελτιστοποιηθεί. Μια αλληλουχία αποδοτικών λύσεων ενισχύονται έτσι ώστε να δομηθεί η καταλληλότερη στρατηγική για την επίλυση ενός προβλήματος.

## **4. Τεχνητά Νευρωνικά Δίκτυα**

### **4.1. Εισαγωγικά στοιχεία**

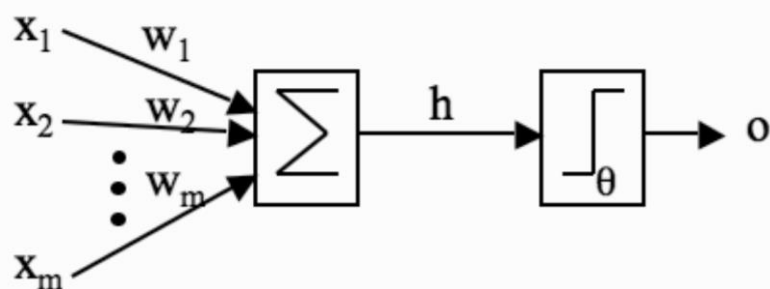
Τα τεχνητά νευρωνικά δίκτυα αποτελούν τεχνικές μηχανικής μάθησης οι οποίες προσομοιώνουν το μηχανισμό της μάθησης στους βιολογικούς οργανισμούς. Το ανθρώπινο νευρικό σύστημα αποτελείται από κύτταρα-νευρώνες. Οι νευρώνες συνδέονται μεταξύ τους με νευροάξονες και δενδρίτες, και τα μέρη που βρίσκονται ανάμεσα στους άξονες και τους δενδρίτες ονομάζονται συνάψεις. Οι συνάψεις είναι μικροσκοπικές δομές που επιτρέπουν σε ένα νευρώνα να μεταδώσει ένα ηλεκτρικό ή χημικό σήμα σε έναν άλλον νευρώνα. Οι τάσεις μεταξύ των συνάψεων “ανταποκρίνονται” σε εξωτερικά ερεθίσματα και αυτές οι αλλαγές αποτελούν τον τρόπο με τον οποίο συντελείται η εκπαίδευση στα έμβια όντα. Ο βιολογικός αυτός μηχανισμός προσομοιώνεται στα τεχνητά νευρωνικά δίκτυα τα οποία αποτελούνται από υπολογιστικές μονάδες οι οποίες ονομάζονται νευρώνες-κόμβοι.

Οι συνδέσεις αυτών των υπολογιστικών μονάδων χαρακτηρίζονται από βάρη (weights) που λειτουργούν αντίστοιχα με τις τάσεις (διαφορές δυναμικού) μεταξύ των συνάψεων. Κάθε τιμή-πληροφορία που εισέρχεται σε έναν νευρώνα παραμετροποιείται βάσει του βάρους και στη συνέχεια περνά στη συνάρτηση που υπάρχει σε κάθε υπολογιστική μονάδα. Ένα τεχνητό νευρωνικό δίκτυο υπολογίζει μια συνάρτηση μεταξύ των τιμών εισόδου και στη συνέχεια η πληροφορία μεταδίδεται στους νευρώνες εξόδου χρησιμοποιώντας τα ενδιάμεσα βάρη ως παραμέτρους. Η διαδικασία της εκπαίδευσης συντελείται ακριβώς κατά την παραμετροποίηση των βαρών.

Όπως στους βιολογικούς οργανισμούς απαιτούνται εξωτερικά ερεθίσματα για την εκπαίδευσή τους, έτσι και στις μηχανές, χρησιμοποιούνται σύνολα δεδομένων προς εκπαίδευση αποτελούμενα από παραδείγματα ζευγαριών δεδομένων εισόδου-εξόδου. Αυτά τα δεδομένα αποτελούν ένα μέτρο για την αξιολόγηση των τιμών των βαρών όσον αφορά στο πόσο καλά το αποτέλεσμα της πρόβλεψης, για κάθε δεδομένο εισόδου, ταιριάζει με την πραγματική τιμή του όπως μας δίνεται κατά την είσοδο.

## 4.2. Μοντέλο McCulloch Pitts

Ένα μαθηματικό μοντέλο για το πως λειτουργεί ο νευρώνας περιέγραψαν το 1943 οι McCulloch και Pitts.



Εικόνα 2. Το μοντέλο McCulloch - Pitts

Όπως βλέπουμε στο παραπάνω σχήμα τα δεδομένα εισόδου ( $x$ ) πολλαπλασιάζονται με τα βάρη ( $w$ ) και ο νευρώνας υπολογίζει το αποτέλεσμα. Αν αυτό το αποτέλεσμα είναι μεγαλύτερο από την τιμή-κατώφλι ( $\theta$ ) τότε μόνο ο νευρώνας ενεργοποιείται. Έχουμε δηλαδή ένα σύνολο βαρών που αντιστοιχεί στις συνάψεις, έναν αθροιστή που υπολογίζει το άθροισμα των σημάτων εισόδου και μια συνάρτηση ενεργοποίησης που αποφασίζει αν ο νευρώνας ενεργοποιείται όπως περιγράφεται από την παρακάτω σχέση.



$$h = \sum_{i=1}^m w_i x_i$$

Το μοντέλο νευρώνα του McCulloch και Pitts παράγει δυαδικά αποτελέσματα. Δηλαδή το άθροισμα του γινομένου δεδομένων εισόδου και αντίστοιχων βαρών είτε παράγει το αποτέλεσμα 1 και ενεργοποιείται ο νευρώνας, είτε παράγει αποτέλεσμα 0 και δεν ενεργοποιείται.

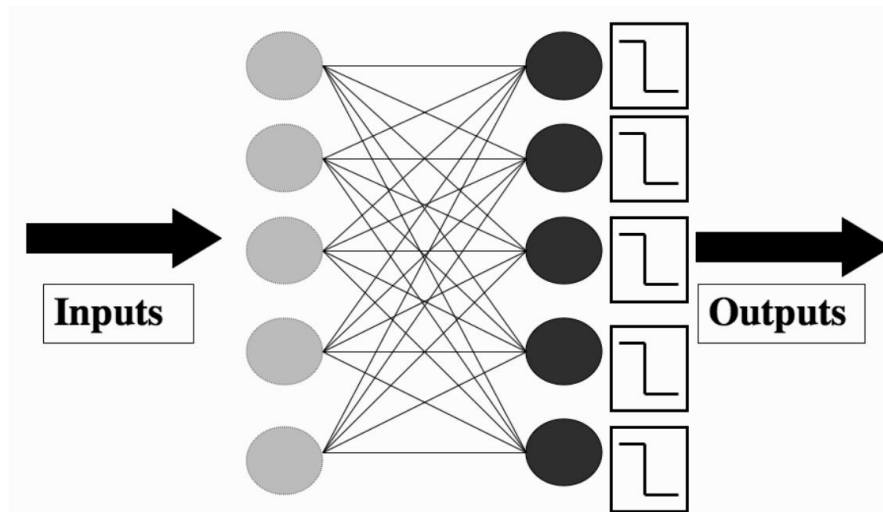
$$o = g(h) = \begin{cases} 1 & \text{if } h > \theta \\ 0 & \text{if } h \leq \theta. \end{cases}$$

Η παραπάνω περιγραφή παρουσιάζει μια πολύ απλή εκδοχή του μοντέλου. Στην πραγματικότητα ένα δίκτυο νευρώνων είναι αρκετά πολύπλοκο και μπορεί να εκτελέσει σύνθετους υπολογισμούς όταν ένας απλός υπολογιστής αδυνατεί να το πράξει. Επίσης τα βάρη μπορούν να πάρουν θετικές ή αρνητικές τιμές. Η εκπαίδευση του νευρώνα καθορίζεται από την παραμετροποίηση των βαρών και των συναρτήσεων ενεργοποίησης έτσι ώστε το δίκτυο να γίνει αποδοτικό και να κάνει σωστές προβλέψεις. Επίσης τα βάρη μπορούν να πάρουν θετικές ή αρνητικές τιμές. Επιπροσθέτως οι τιμές εισόδου δεν είναι απαραίτητο να αθροιστούν με γραμμικό τρόπο. Η πλειοψηφία των τεχνητών νευρωνικών δικτύων χρησιμοποιούν μη γραμμικές συναρτήσεις ενεργοποίησης γιατί στην πλειοψηφία των εφαρμογών, οι σχέσεις μεταξύ των τιμών εισόδου και εξόδου δεν είναι γραμμικές. Οι μη γραμμικές συναρτήσεις επιτρέπουν στο νευρωνικό δίκτυο να εκφράσει περίπλοκα μοτίβα μεταξύ εισόδων και εξόδων.

Επιτρέπουν επίσης την εκπαίδευση του αλγόριθμου χρησιμοποιώντας περίπλοκες διαδικασίες που περιγράφουν σύνθετα σύνολα δεδομένων με μεγάλο αριθμό χαρακτηριστικών όπως είναι οι ψηφιακές εικόνες και τα ηχητικά δεδομένα.

### 4.3. Ο αλγόριθμος Perceptron

Ένα από τα πιο γνωστά είδη τεχνητών νευρωνικών δικτύων είναι ο αλγόριθμος Perceptron. Ο Perceptron είναι ένα σύνολο νευρώνων McCulloch-Pitts οι οποίοι συνδέονται μεταξύ τους με μια σειρά εισόδων και βαρών τα οποία ενώνουν τις εισόδους με τους νευρώνες.



Εικόνα 3. Perceptron

Οι νευρώνες είναι τελείως ανεξάρτητοι μεταξύ τους όσον αφορά τη λειτουργία τους. Το μόνο που έχουν κοινό είναι οι τιμές εισόδου. Ο αριθμός των εισόδων ( $m$ ) δεν είναι απαραίτητο να είναι ίδιος με τον αριθμό των νευρώνων ( $n$ ). Ο αριθμός των εισόδων καθορίζεται από τα χαρακτηριστικά του συνόλου δεδομένων. Κατά τον ίδιο τρόπο καθορίζεται και ο αριθμός των εξόδων από τη στιγμή που επιθυμούμε ο αλγόριθμος να εκπαιδεύεται να παράγει ένα συγκεκριμένο αποτέλεσμα συμβατό με τα δεδομένα εισόδου. Όταν ένας νευρώνας δεν παράγει το ορθό αποτέλεσμα τότε πρέπει να συντελεστεί η παραμετροποίηση των βαρών ( $w$ ). Για παράδειγμα αν ένας νευρώνας ( $k$ ) παράγει λάθος αποτέλεσμα, τότε θα πρέπει να παραμετροποιηθούν όλα τα βάρη που αντιστοιχούν σε αυτόν τον νευρώνα.

Αν ο νευρώνας που λανθασμένα ενεργοποιήθηκε ενώ δεν έπρεπε, τότε τα βάρη πρέπει να πάρουν μικρότερες τιμές, και αν δεν ενεργοποιήθηκε, ενώ θα έπρεπε, τότε θα πρέπει να λάβουν υψηλότερες τιμές. Η παραπάνω διαδικασία περιγράφει μια συνάρτηση σφάλματος (error function). Επίσης θα πρέπει να δώσουμε προσοχή στο ενδεχόμενο τα βάρη να παίρνουν αρνητικές τιμές. Σε αυτή την περίπτωση η συνάρτηση γίνεται ως εξής, όπου  $w$  αντιπροσωπεύει την τιμή του βάρους,  $y$  η τιμή της εξόδου,  $t$  η τιμή του αποτελέσματος που επιθυμούμε:

$$\Delta w_{ik} = -(y_k - t_k) \times x_i$$

Κατά αυτόν τον τρόπο η νέα τιμή που πρέπει να λάβει το βάρος είναι το άθροισμα της τιμής  $\Delta$  με την παλαιά τιμή. Μια σημαντική έννοια που πρέπει να προσθέσουμε είναι ο ρυθμός μάθησης ( $\eta$ ). Ο ρυθμός μάθησης καθορίζει πόσο γρήγορα το νευρωνικό δίκτυο εκπαιδεύεται.

$$w_{ij} \leftarrow w_{ij} - \eta(y_j - t_j) \cdot x_i$$

Η παραπάνω σχέση μας δείχνει με ποιον τρόπο θα παραμετροποιήσουμε τις τιμές των βαρών. Ο ρυθμός μάθησης ελέγχει το πόσο πολύ θα παραμετροποιηθούν αυτά. Αν χρησιμοποιήσουμε μεγάλη τιμή για τον ρυθμό μάθησης τότε τα βάρη θα αλλάξουν κατά ένα μεγάλο βαθμό κάθε φορά που ο αλγόριθμος θα παράγει ένα λανθασμένο αποτέλεσμα και αυτό θα προκαλέσει την αστάθεια του νευρωνικού δικτύου. Αν χρησιμοποιήσουμε μια μικρή τιμή τότε θα χρειαστεί πολύ περισσότερος χρόνος για την εκπαίδευση του νευρωνικού δικτύου. Παρόλα αυτά θα είναι πιο σταθερό και δεν θα επηρεάζεται τόσο πολύ από τον θόρυβο των δεδομένων.

Συμπερασματικά ένας αλγόριθμος Perceptron μπορεί να παρουσιαστεί ως εξής: Πρώτα αρχικοποιούμε τα βάρη χρησιμοποιώντας μικρές τυχαίες τιμές. Στη συνέχεια εκπαιδεύουμε το νευρωνικό δίκτυο για  $T$  επαναλήψεις μέχρι όλα τα αποτελέσματα-έξοδοι να είναι σωστά.

Για κάθε διάνυσμα εισόδου υπολογίζουμε την ενεργοποίηση του κάθε νευρώνα  $j$  χρησιμοποιώντας τη συνάρτηση  $g$ :

$$y_j = g \left( \sum_{i=0}^m w_{ij} x_i \right) = \begin{cases} 1 & \text{if } \sum_{i=0}^m w_{ij} x_i > 0 \\ 0 & \text{if } \sum_{i=0}^m w_{ij} x_i \leq 0 \end{cases}$$

Στη συνέχεια συντελείται η παραμετροποίηση των βαρών ως εξής:

$$w_{ij} \leftarrow w_{ij} - \eta(y_j - t_j) \cdot x_i$$

Τέλος υπολογίζουμε την ενεργοποίηση του κάθε νευρώνα  $j$  με την ακόλουθη σχέση.

$$y_j = g \left( \sum_{i=0}^m w_{ij} x_i \right) = \begin{cases} 1 & \text{if } w_{ij} x_i > 0 \\ 0 & \text{if } w_{ij} x_i \leq 0 \end{cases}$$

#### 4.4. Ο αλγόριθμος Multi-layer Perceptron

Όπως προαναφέραμε, τα γραμμικά μοντέλα είναι σχετικά απλά στην κατανόηση και τη χρήση τους, όμως μπορούν να ανιχνεύσουν μόνο ευθείες γραμμές διαχωρισμού, επίπεδα ή υπερεπίπεδα. Αυτό δεν είναι αρκετό γιατί η πλειοψηφία των προβλημάτων δεν σχετίζονται με τον γραμμικό διαχωρισμό.

Αυτό το πρόβλημα αντιμετωπίζεται με τη χρήση του αλγορίθμου Multi-layer Perceptron (MLP) ο οποίος είναι από τους πιο δημοφιλείς μεθόδους μηχανικής μάθησης. Είναι ένα νευρωνικό δίκτυο που αποτελείται από τρία ή περισσότερα επίπεδα. Δηλαδή ανάμεσα στα επίπεδα εισόδου και εξόδου υπάρχουν και άλλα κρυφά επίπεδα.

Όπως ακριβώς και στον απλό Perceptron, η εκπαίδευση του (MLP) αποτελείται από δύο μέρη. Κατά το πρώτο μέρος υπολογίζονται οι τιμές εξόδου δεδομένων των τιμών εισόδου και των τιμών των βαρών. Κατά το δεύτερο μέρος συντελείται η παραμετροποίηση των βαρών βάσει της συναρτήσεως σφάλματος που υπολογίζει τη διαφορά ανάμεσα στα αποτελέσματα του αλγορίθμου και τις πραγματικές τιμές.

Η πληροφορία κινείται με πρόσθιο τρόπο από το επίπεδο εισόδου, διαπερνώντας τα κρυφά επίπεδα, υπολογίζοντας τις συναρτήσεις ενεργοποίησης κάθε νευρώνα ενός επιπέδου, περνώντας εν συνεχεία στο επόμενο επίπεδο νευρώνων, μέχρι την έξοδο. Κατά τη διαδικασία υπολογισμού των λαθών, τα πράγματα γίνονται πιο πολύπλοκα. Η μέθοδος που ακολουθείται για την αντιμετώπιση αυτού το προβλήματος ονομάζεται μέθοδος οπισθοδιάδοσης του λάθους (back-propagation of error). Πρόκειται για μια μορφή καθόδου βασισμένη στην κλίση (επικλινής κάθοδος - gradient descent) όπως θα περιγράψουμε παρακάτω. Αυτό που επιθυμούμε δηλαδή είναι η ελαχιστοποίηση του λάθους έτσι ώστε οι νευρώνες να ενεργοποιούνται μόνο όταν πρέπει.

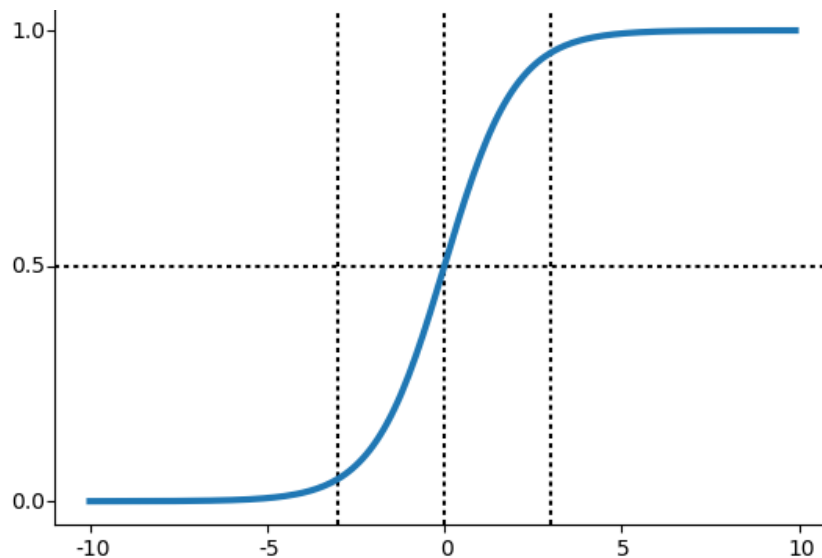
Η συνάρτηση σφάλματος που χρησιμοποιείται είναι εκείνη του αθροίσματος των τετραγώνων.

$$E(\mathbf{t}, \mathbf{y}) = \frac{1}{2} \sum_{k=1}^N (y_k - t_k)^2.$$

Αν υπολογίσουμε τη μερική παράγωγο μιας συνάρτησης σφάλματος ως προς τα βάρη τότε θα διαπιστώσουμε την κλίση του σφάλματος.

Τα βάρη λοιπόν παραμετροποιούνται με τέτοιον τρόπο ώστε το σφάλμα να φτάσει ένα τοπικό ελάχιστο.

Επίσης χρησιμοποιούμε μη γραμμική συνάρτηση ενεργοποίησης (threshold function). Ένας τέτοιος τύπος συνάρτησης είναι η σιγμοειδής συνάρτηση.



Εικόνα 4. Η σιγμοειδής συνάρτηση

$$a = g(h) = \frac{1}{1 + \exp(-\beta h)}$$

Αυτή η συνάρτηση καθορίζει αν ένας νευρώνας ενεργοποιείται. Βάσει του αποτελέσματος της παραγώγου της, παραμετροποιούμε τα βάρη στην κατεύθυνση του τοπικού ελαχίστου. Με αυτόν τον τρόπο βελτιστοποιούμε τη συνάρτηση σφάλματος του νευρωνικού δικτύου. Στο επίπεδο της εξόδου έχουμε υπολογίσει τα σφάλματα ως το άθροισμα της διαφοράς των τετραγώνων των αποτελεσμάτων εξόδου με τις επιθυμητές τιμές.

Στη συνέχεια θέλουμε να υπολογίσουμε την κλίση αυτών των σφαλμάτων και αυτή την πληροφορία να την αξιοποιήσουμε για τη βέλτιστη παραμετροποίηση των βαρών. Ξεκινάμε αυτή τη διαδικασία πρώτα από τους νευρώνες του επιπέδου εξόδου και έπειτα κινούμαστε προς τα πίσω διαπερνώντας, προς την αντίθετη κατεύθυνση, το νευρωνικό δίκτυο μέχρι να φτάσουμε στο επίπεδο της εισόδου.

Το πρόβλημα είναι ότι για τους νευρώνες του επιπέδου εξόδου δεν γνωρίζουμε τις τιμές που εισάγονται σε αυτούς και ότι για τους νευρώνες των κρυφών επιπέδων δεν γνωρίζουμε ούτε τα αποτελέσματα που αυτοί εξάγουν αλλά ούτε και τα δεδομένα που εισάγονται σε αυτούς. Συνακόλουθα, ενώ μπορούμε να υπολογίσουμε το σφάλμα στην έξοδο, επειδή δεν γνωρίζουμε τα δεδομένα εισαγωγής που το προκάλεσαν, δεν μπορούμε να παραμετροποιήσουμε τα βάρη των ενδιάμεσων επιπέδων. Αυτό λύνεται χρησιμοποιώντας τον κανόνα αλυσίδας.

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx},$$

Ο κανόνας αλυσίδας μας βοηθάει να κατανοήσουμε ότι αν θέλουμε να υπολογίσουμε με ποιον τρόπο το σφάλμα αλλάζει τιμές όσο εμείς παραμετροποιούμε τα βάρη, μπορούμε να σκεφτούμε πως ακριβώς το σφάλμα αλλάζει, όταν παραμετροποιούμε τις εισόδους των βαρών, και στη συνέχεια να πολλαπλασιάσουμε αυτό με το πόσο αλλάζουν αυτές οι τιμές των εισόδων όταν παραμετροποιούμε τα βάρη.

Το παραπάνω είναι ιδιαίτερος χρήσιμο γιατί με αυτόν τον τρόπο μπορούμε να υπολογίσουμε όλες τις παραγώγους που επιθυμούμε, δηλαδή να εκφράσουμε τις ενεργοποιήσεις των νευρώνων εξόδου μέσω των ενεργοποιήσεων των νευρώνων των κρυφών επιπέδων και στη συνέχεια να στείλουμε τους υπολογισμούς του σφάλματος προς τα πίσω για να καθοριστούν ποιοι είναι οι νευρώνες που είναι υπεύθυνοι για τα αποτελέσματα εξόδου. Δηλαδή υπολογίζουμε μέσω των μερικών παραγώγων της συνάρτησης σφάλματος ως προς τα βάρη, την κλίση των σφαλμάτων βάσει των βαρών, έτσι ώστε να τα ελαχιστοποιήσουμε.

Συμπερασματικά ο αλγόριθμος λειτουργεί ως εξής: Ένα διάνυσμα δεδομένων εισόδου εισέρχεται στο πρώτο επίπεδο. Τα δεδομένα εισόδου διαπερνούν το νευρωνικό δίκτυο και συνυπολογίζοντας τα βάρη αποφασίζεται αν οι νευρώνες του πρώτου κρυφού επιπέδου θα ενεργοποιηθούν ή όχι. Η συνάρτηση ενεργοποίησης  $g$  είναι η σιγμοειδής συνάρτηση. Οι έξοδοι αυτών των νευρώνων μαζί με τα βάρη τους στο δεύτερο κρυφό επίπεδο καθορίζουν αν θα ενεργοποιηθούν οι νευρώνες του δεύτερου κρυφού επιπέδου. Έπειτα το σφάλμα υπολογίζεται ως το άθροισμα της διαφοράς των τετραγώνων των τιμών εξόδου και των ορθών αποτελεσμάτων που είναι ήδη γνωστά από το σύνολο δεδομένων εκπαίδευσης. Στη συνέχεια το σφάλμα τροφοδοτείται προς της αντίθετη κατεύθυνση για να παραμετροποιήσει τα βάρη του τελευταίου επιπέδου πρώτα, και συνεχίζει παραμετροποιώντας τα βάρη του αμέσως επόμενου επιπέδου.

Πιο αναλυτικά, με μαθηματικούς όρους, πρώτα αρχικοποιούνται τα βάρη ( $u$ ) θέτοντας σε αυτά αρνητικές ή θετικές μικρές τυχαίες τιμές. Κατά την εκπαίδευση για κάθε διάνυσμα εισόδου υπολογίζουμε την τιμή που εισέρχεται σε κάθε νευρώνα αλλά και την συνάρτηση ενεργοποίησης.

$$h_{\zeta} = \sum_{i=0}^L x_i v_{i\zeta}$$

$$a_{\zeta} = g(h_{\zeta}) = \frac{1}{1 + \exp(-\beta h_{\zeta})}$$

Διαπερνώντας το νευρωνικό δίκτυο φτάνουμε στο επίπεδο εξόδου το οποίο έχει και αυτό συναρτήσεις ενεργοποίησης.

$$h_{\kappa} = \sum_j a_j w_{j\kappa}$$

$$y_{\kappa} = g(h_{\kappa}) = \frac{1}{1 + \exp(-\beta h_{\kappa})}$$



Στη συνέχεια με την τεχνική της οπισθοδιάδοσης υπολογίζουμε το σφάλμα στο επίπεδο εξόδου.

$$\delta_o(\kappa) = (y_\kappa - t_\kappa) y_\kappa (1 - y_\kappa)$$

Έπειτα υπολογίζεται το σφάλμα στα κρυφά ενδιάμεσα επίπεδα.

$$\delta_h(\zeta) = a_\zeta (1 - a_\zeta) \sum_{k=1}^N w_{\zeta k} \delta_o(k)$$

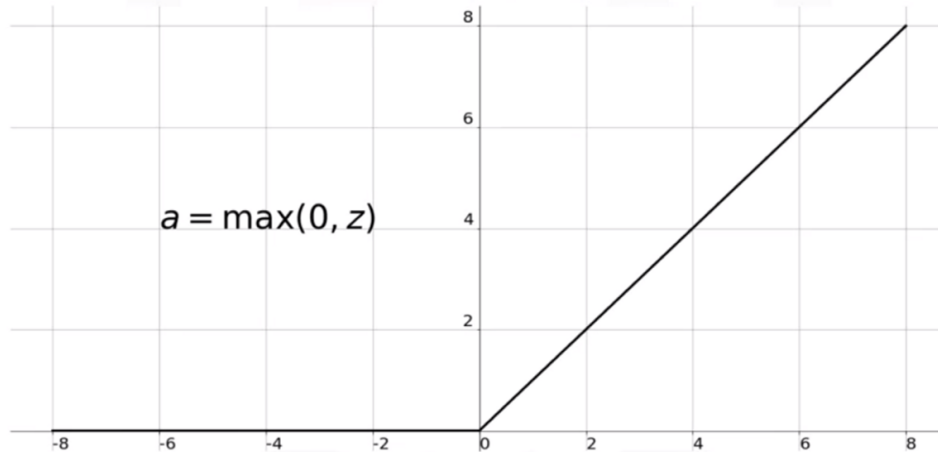
Τα βάρη του επιπέδου εξόδου υπολογίζονται ως εξής:

$$w_{\zeta \kappa} \leftarrow w_{\zeta \kappa} - \eta \delta_o(\kappa) a_\zeta^{\text{hidden}}$$

Τέλος παραμετροποιούνται τα βάρη των κρυφών επιπέδων με τον ακόλουθο τύπο.

$$v_l \leftarrow v_l - \eta \delta_h(\kappa) x_l$$

Πρέπει να σημειώσουμε ότι εκτός από τη σιγμοειδή συνάρτηση που αναλύσαμε πιο πάνω, άλλο ένα είδος συνάρτησης ενεργοποίησης είναι η ανορθωμένη γραμμική συνάρτηση (ReLU – Rectified Linear Unit).



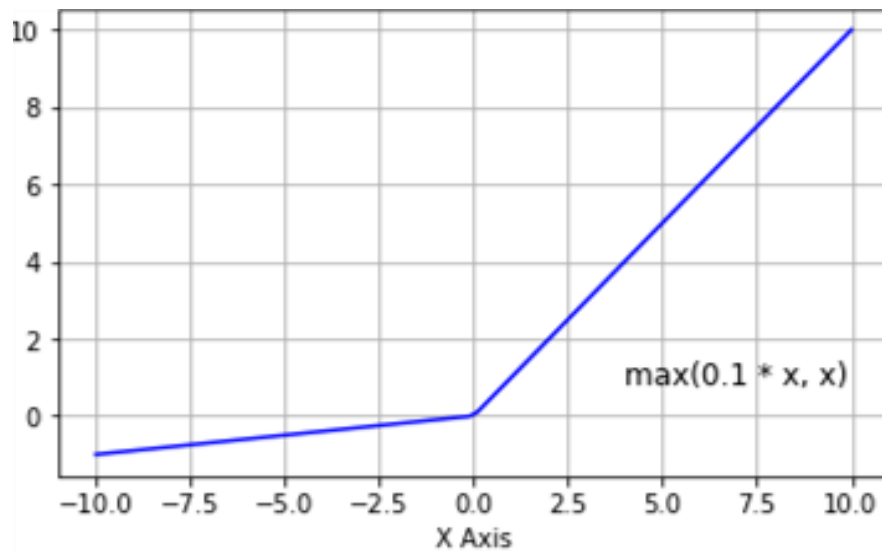
Εικόνα 5. ReLU

Πρόκειται για ιδιαιτέρως αποδοτική συνάρτηση η οποία βγάζει ένα αποτέλεσμα  $X$  αν το αποτέλεσμα εξόδου είναι  $X$  είναι θετικό αλλιώς παράγει το αποτέλεσμα 0. Η συγκεκριμένη συνάρτηση μπορεί να πάρει τιμές από το 0 έως το άπειρο.

Το μειονέκτημά της είναι ότι η παράγωγός της είναι μηδενική όταν οι τιμές εισόδου είναι αρνητικές. Κατά συνέπεια δεν μπορεί να υπολογιστεί η οπισθοδιάδοση του σφάλματος και το νευρωνικό δίκτυο αδυνατεί να εκπαιδευτεί παρά μόνο αν τα δεδομένα εισόδου είναι μεγαλύτερα του μηδενός.

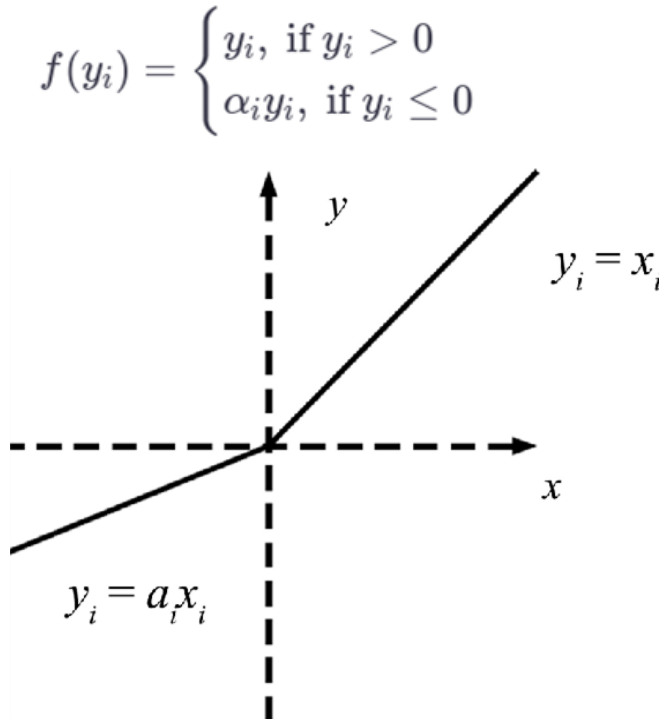
Για την επίλυση αυτού του προβλήματος υπάρχουν δυο εκδοχές της συνάρτησης ReLU. Μία είναι η εκδοχή Leaky η οποία επιστρέφει ως αποτέλεσμα το γινόμενο του  $X$  με την τιμή 0,1 όταν το  $X$  είναι μικρότερο του μηδενός.

Αυτή η συνάρτηση έχει μικρή θετική κλίση ίση με 0,1 όταν η τιμή του  $X$  παίρνει αρνητικές τιμές συνεπώς μπορεί πλέον να υπολογιστεί η οπισθοδιάδοση του σφάλματος.



Εικόνα 6. Leaky ReLU

Η δεύτερη εκδοχή της συνάρτησης ReLU ονομάζεται παραμετρική. Είναι παρόμοια με την Leaky ReLU και δέχεται ως παράμετρο-συντελεστή την κλίση (gradient) ( $\alpha$ ).



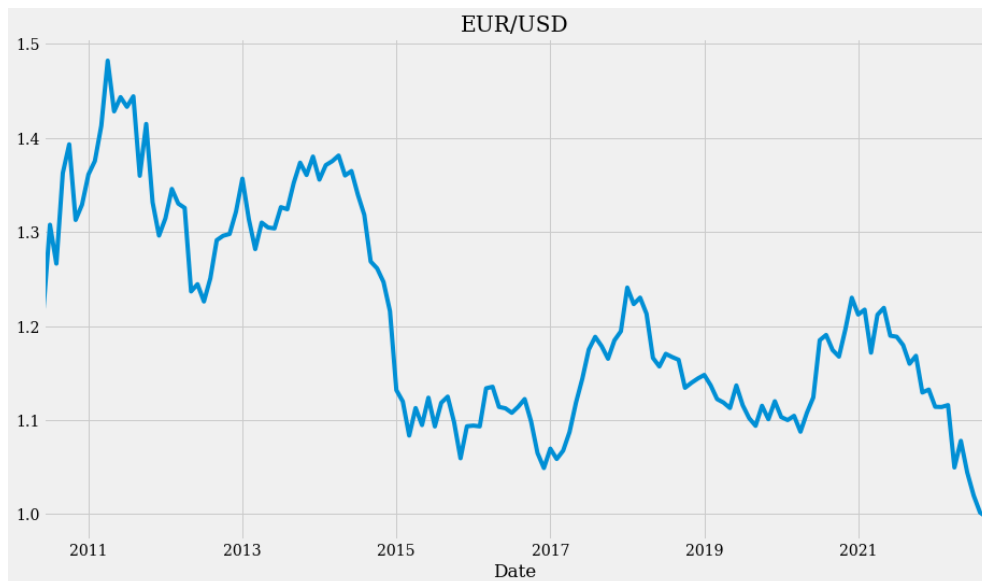
Εικόνα 7. Παραμετρική ReLU

Η εκπαίδευση του Multi-Layer Perceptron απαιτεί συνεχείς επαναλήψεις της διαδικασίας που αναλύσαμε πιο πάνω, με τα βάρη να παραμετροποιούνται συνεχώς βάσει των σφαλμάτων εξόδου σε κάθε επανάληψη. Ο αριθμός των επαναλήψεων μπορεί να καθοριστεί από το σφάλμα του αθροίσματος των τετραγώνων κατά τη διάρκεια της εκπαίδευσης. Συνήθως η τιμή του μειώνεται με γρήγορο ρυθμό κατά τις πρώτες επαναλήψεις της διαδικασίας εκπαίδευσης, όμως στη συνέχεια αυτός ο ρυθμός ελαχιστοποιείται όταν βρεθεί το τοπικό ελάχιστο. Σε αυτό ακριβώς το σημείο εισάγουμε στον αλγόριθμο ένα ξεχωριστό σύνολο δεδομένων (validation set) και επαναλαμβάνουμε εκ νέου την εκπαίδευση έτσι ώστε να εκτιμήσουμε το πόσο αποδοτικά ανταποκρίνεται ο αλγόριθμος στα νέα δεδομένα. Σε κάποιο σημείο το σφάλμα θα αρχίσει πάλι να αυξάνεται γιατί πλέον ο αλγόριθμος θα λαμβάνει υπόψη και τον θόρυβο από τα νέα δεδομένα. Σε αυτό το σημείο μπορούμε να σταματήσουμε την εκπαίδευση.

## 5. Χρήση Νευρωνικών Δικτύων σε Πραγματικά Δεδομένα

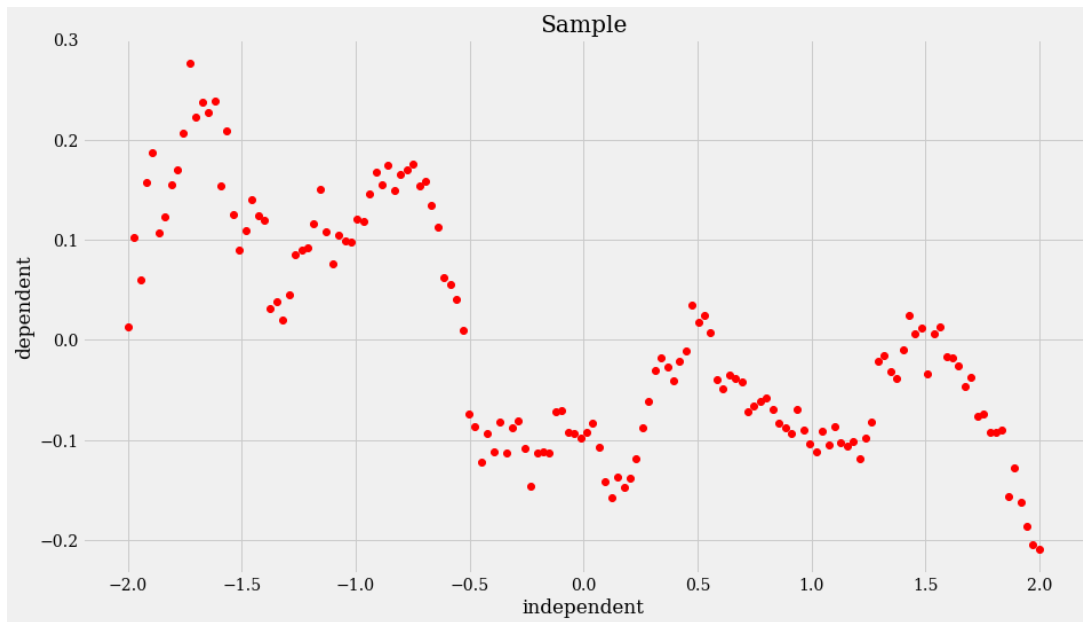
### 5.1. Καθορισμός Στρατηγικής Βάσει Απλού Κινητού Μέσου

Η αγορά συναλλάγματος (foreign exchange market) είναι μια παγκόσμια αγορά που καθορίζει την ισοτιμία των νομισμάτων σε παγκόσμιο επίπεδο. Οι συμμετέχοντες σε αυτές τις αγορές μπορούν να προχωρήσουν σε πωλήσεις, αγορές, ανταλλαγές πάνω στην ισοτιμία ζεύγους νομισμάτων. Σε αυτή τη διαδικασία συμμετέχουν ιδιωτικές τράπεζες, ιδιωτικές εταιρείες, κεντρικές τράπεζες κρατών, και κάθε είδους επενδυτικός οργανισμός. Αποτελεί τη μεγαλύτερη χρηματοοικονομική αγορά παγκοσμίως η οποία λειτουργεί μέσω ενός δικτύου χρηματοοικονομικών κέντρων οι οποίες συναλλάσσονται μεταξύ τους 24 ώρες την ημέρα. Ο σκοπός της παρούσας εργασίας είναι η ανάλυση του τρόπου λειτουργίας συγκεκριμένων αλγόριθμων για την πρόβλεψη τιμών της ισοτιμίας Ευρώ/Δολαρίου. Τα δεδομένα που χρησιμοποιήθηκαν βασίζονται σε τιμές χρονοσειρών που περιγράφουν την ισοτιμία Ευρώ/Δολαρίου. Τα δεδομένα αφορούν το διάστημα από τον Ιανουάριο του 2005 μέχρι τον Σεπτέμβρη του 2022. Για το πείραμα έγινε επαναδειγματοληψία (reshampling) των δεδομένων σε μηνιαία βάση.



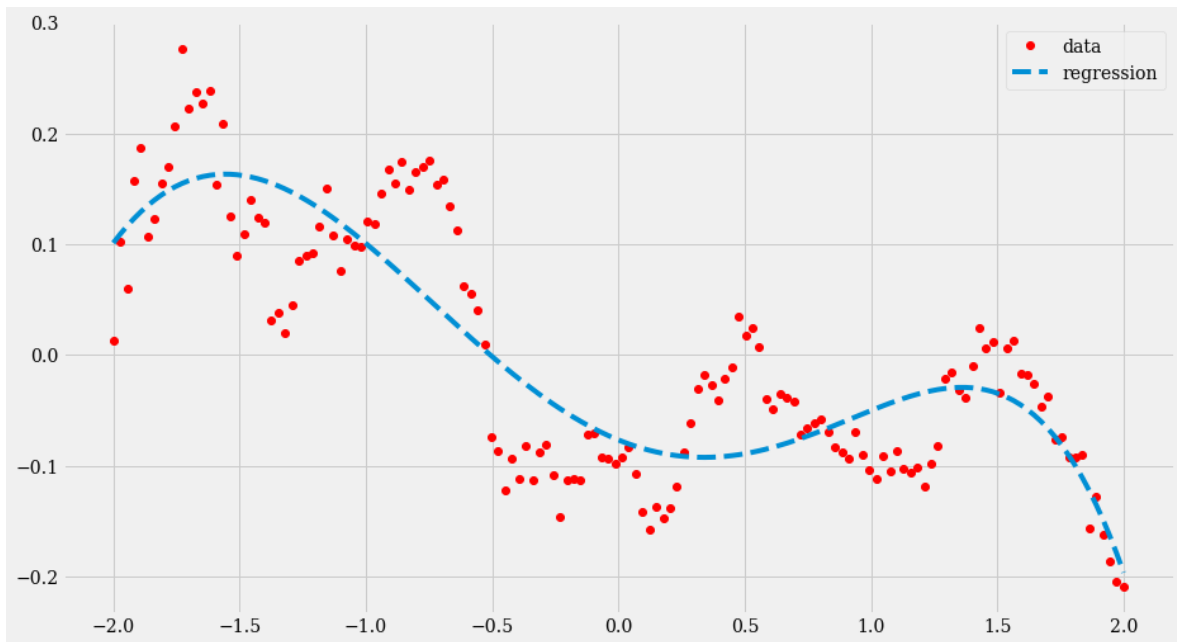
Εικόνα 8. Εξέλιξη ισοτιμίας Ευρώ/Δολαρίου

Για να έχουμε μια πιο κατανοητή οπτικοποίηση του πειράματος σε δύο διαστάσεις δημιουργήσαμε ένα τεχνητό χαρακτηριστικό ως την ανεξάρτητη μεταβλητή μας η οποία βέβαια δεν έχει κάποια επεξηγηματική σημασία σχετικά με την εξαρτημένη μεταβλητή η οποία είναι η ισοτιμία EUR/USD.



Εικόνα 9. Οι παρατηρήσεις βάσει εξαρτημένης και ανεξάρτητης μεταβλητής

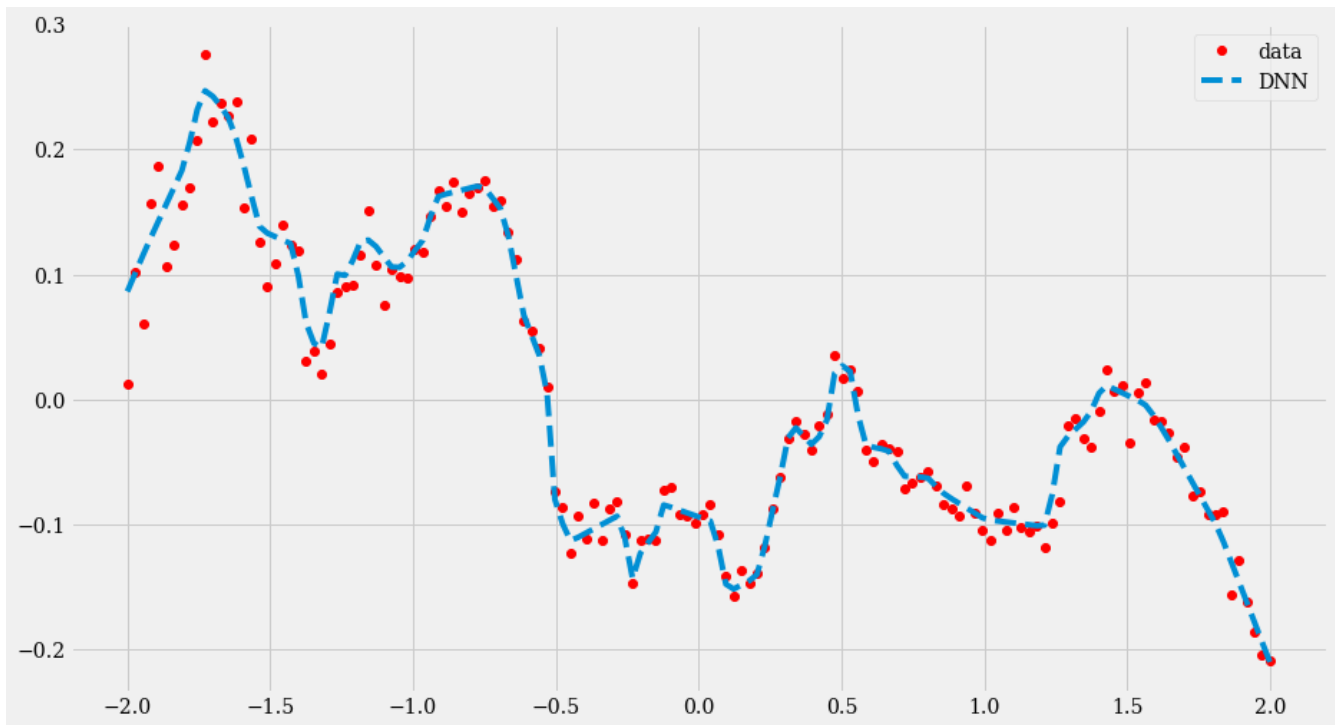
Ως μέτρο επιτυχίας της πρόβλεψης χρησιμοποιήσαμε το Μέσο Τετραγωνικό Σφάλμα (MSE). Χρησιμοποιήσαμε τον αλγόριθμο της πολυωνμικής παλινδρόμησης ελαχίστων τετραγώνων και υπολογίστηκε το μέσο τετραγωνικό σφάλμα.



Εικόνα 10. Πολυωνυμική παλινδρόμηση

Όπως είναι λογικό όσο περισσότερα δεδομένα εισάγουμε στον αλγόριθμο για την εκμάθηση τόσο περισσότερο μειώνεται η τιμή του MSE. Εν συνεχεία χρησιμοποιήσαμε για την πρόβλεψη των τιμών της ισοτιμίας νευρωνικά δίκτυα. Αυτό έγινε με τη βοήθεια του keras framework.

Παρακάτω μπορούμε να διαπιστώσουμε πως το νευρωνικό δίκτυο που χρησιμοποιήσαμε προσεγγίζει τα δεδομένα μας. Για τη συγκεκριμένη εφαρμογή χρησιμοποιήσαμε μόνο ένα hidden layer. Εάν το παραμετροποιήσουμε χρησιμοποιώντας παραπάνω επίπεδα (layers) παρατηρούμε ότι το μέσο τετραγωνικό σφάλμα μειώνεται σημαντικά. Αυτό μπορούμε να το διαπιστώσουμε με την παρακάτω οπτικοποίηση.

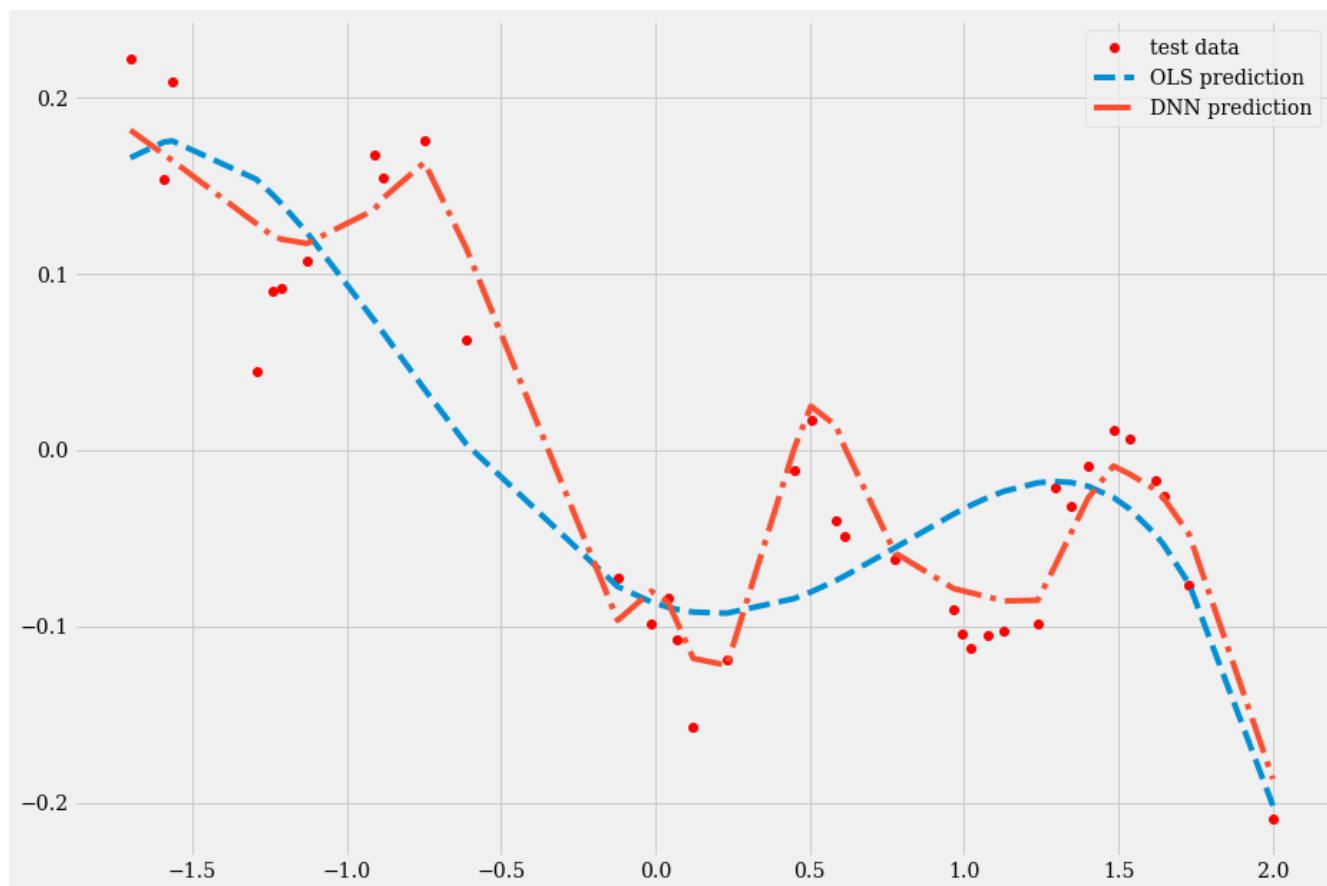


Εικόνα 11. Χρήση νευρωνικού δικτύου

Για το παραπάνω παράδειγμα η ανάλυση της λειτουργίας των συγκεκριμένων αλγορίθμων έγινε χρησιμοποιώντας το ίδιο σύνολο δεδομένων για την εκπαίδευση των αλγορίθμων αλλά και για την πρόβλεψη. Όμως στην πραγματικότητα το ζήτημα είναι η πρόβλεψη των μελλοντικών τιμών των ισοτιμιών. Γι' αυτή την περίπτωση τα δεδομένα μας χωρίζονται σε τρία διαφορετικά σύνολα. Το σύνολο δεδομένων που θα χρησιμοποιηθεί για την εκπαίδευση του αλγόριθμου (training data set), ένα διαφορετικό σύνολο δεδομένων για να διαπιστωθεί η απόδοση του αλγόριθμου κατά τη διάρκεια της εκπαίδευσής του και ένα άλλο σύνολο δεδομένων (test data set) όπου θα τροφοδοτήσουμε στον αλγόριθμο αφού ολοκληρωθεί η διαδικασία της εκπαίδευσης.

Χρησιμοποιήσαμε ξανά για την συγκεκριμένη περίπτωση τον αλγόριθμο της πολυωνυμικής παλινδρόμησης ελαχίστων τετραγώνων και το νευρωνικό δίκτυο και παρακάτω βλέπουμε τη σύγκριση στην απόδοση των δυο αλγορίθμων.





Εικόνα 12. Σύγκριση νευρωνικού - παλινδρόμησης

Τα παραπάνω αποτελούν μια εισαγωγή στην έννοια του algorithmic trading δηλαδή τη χρησιμοποίηση και εκμετάλλευση οικονομικών μοτίβων και τάσεων μέσω μιας αυτοματοποιημένης εκτέλεσης εντολών, βασισμένες σε συγκεκριμένες στρατηγικές με σκοπό την όσο δυνατόν καλύτερη πρόβλεψη.

Αυτή η διαδικασία πρόβλεψης τιμών της αγοράς βασίζεται σε μια τεχνική προσέγγιση που ονομάζεται διανυσματικό backtesting (vectorized backtesting). Ο εννοιολογικός πυρήνας του backtesting είναι ουσιαστικά η προσομοίωση μιας συγκεκριμένης χρηματοοικονομικής στρατηγικής σε παρελθόντα χρόνο, δηλαδή σε παρελθούσες τιμές. Στη συνέχεια τα παραγόμενα οφέλη συγκρίνονται με ένα συγκεκριμένο σταθερό μέτρο (benchmark) έτσι ώστε να αξιολογηθεί η απόδοση της στρατηγικής (Κάποια γνωστά benchmarks είναι το S&P 500 Index και ο βιομηχανικός μέσος του Dow Jones).

Με τη μέθοδο του vectorized backtesting συλλέγουμε δεδομένα όπου αυτά οργανώνονται σε διανύσματα και πίνακες τα οποία στη συνέχεια επεξεργάζονται για την ταχεία προσομοίωση παρελθουσών επιδόσεων της επενδυτικής στρατηγικής.

Η αποδοτικότητα κάθε στρατηγικής μπορεί να υπολογιστεί με το άθροισμα των λογαρίθμων των κερδών (log returns) όπως δίνεται από τον παρακάτω τύπο.

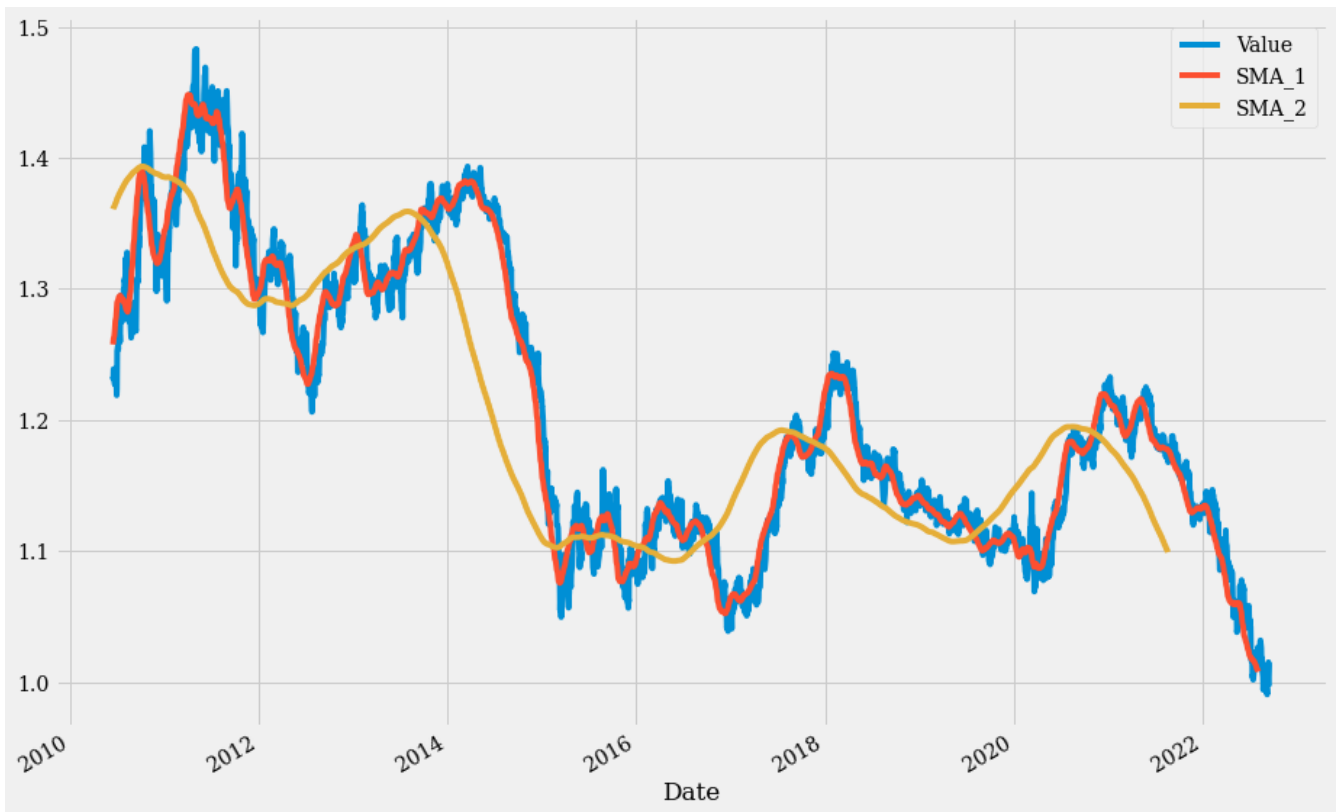
$$r_t = \log\left(\frac{p_t}{p_{t-1}}\right) = \log(p_t) - \log(p_{t-1})$$

Ο παραπάνω τύπος περιγράφει όπως είναι εμφανές τη διαφορά των λογαριθμικών τιμών στο διάστημα (t-1, t). Με τη χρήση του μπορούμε να υπολογίσουμε το αθροιστικό κέρδος μιας στρατηγικής. Στη συνέχεια κάνουμε χρήση της εκθετικής συνάρτησης.

Για την εφαρμογή του backtesting θα χρησιμοποιήσουμε τα ίδια δεδομένα που χρησιμοποιήσαμε και πριν. Η ιδέα της στρατηγικής είναι η ακόλουθη. Θα χρησιμοποιήσουμε για την πρόβλεψη τη μέθοδο του απλού κινητού μέσου (SMA). Πρόκειται για μια απλή μέθοδο πρόβλεψης τιμών που χρησιμοποιεί την τιμή του μέσου όρου των  $n$  πιο πρόσφατων παρατηρήσεων μιας χρονοσειράς. Αυτό γιατί οι πιο πρόσφατες παρατηρήσεις της χρονοσειράς πιστεύεται ότι είναι αντιπροσωπευτικότερες για την πρόβλεψη από τις πιο απομακρυσμένες χρονικά. Ονομάζεται κινητός γιατί η τιμή του δεν έχει σταθερή τιμή αλλά αλλάζει όταν είναι διαθέσιμη μια νέα παρατήρηση της χρονοσειράς. Παρακάτω δίνεται ο τύπος.

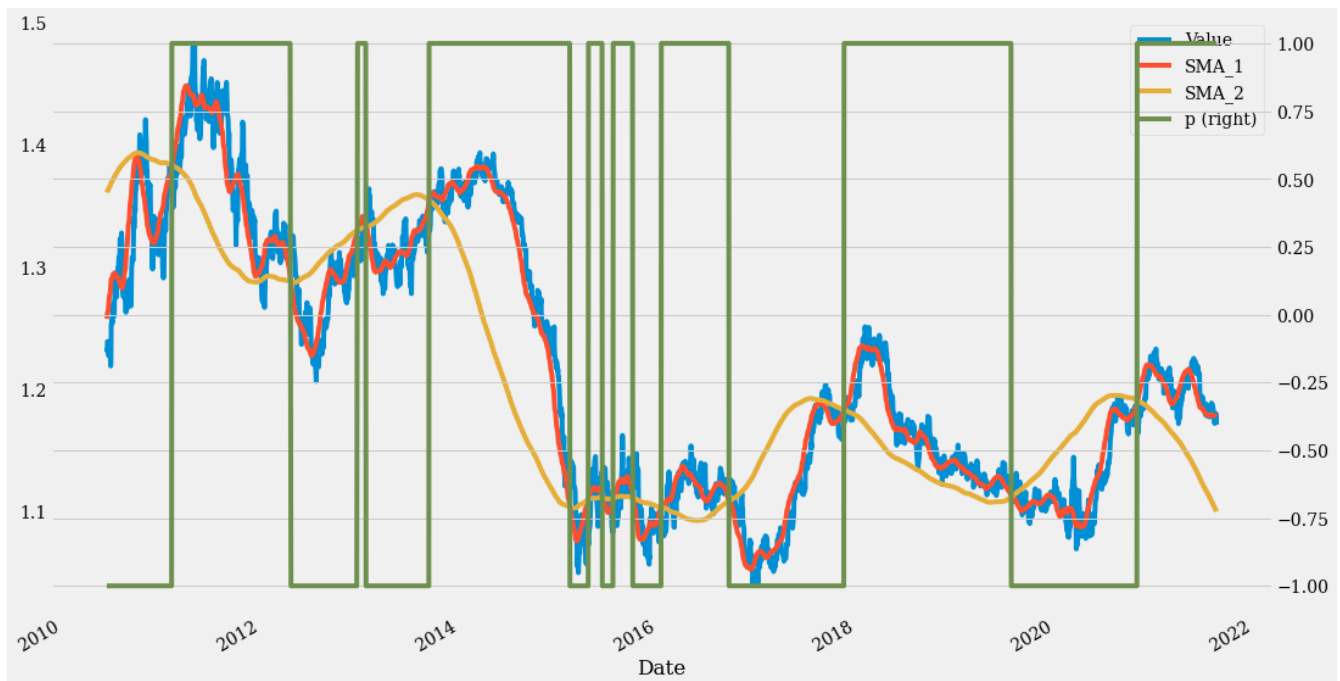
$$SMA = \frac{A_1 + A_2 + \dots + A_n}{n}$$

Συγκεκριμένα υπολογίζουμε δύο τιμές του απλού κινητού μέσου. Έναν βραχύ για 30 ημέρες (SMA1) και έναν πιο μεγάλο για 280 ημέρες (SMA2).



Εικόνα 13. Εφαρμογή δύο κινητών μέσων

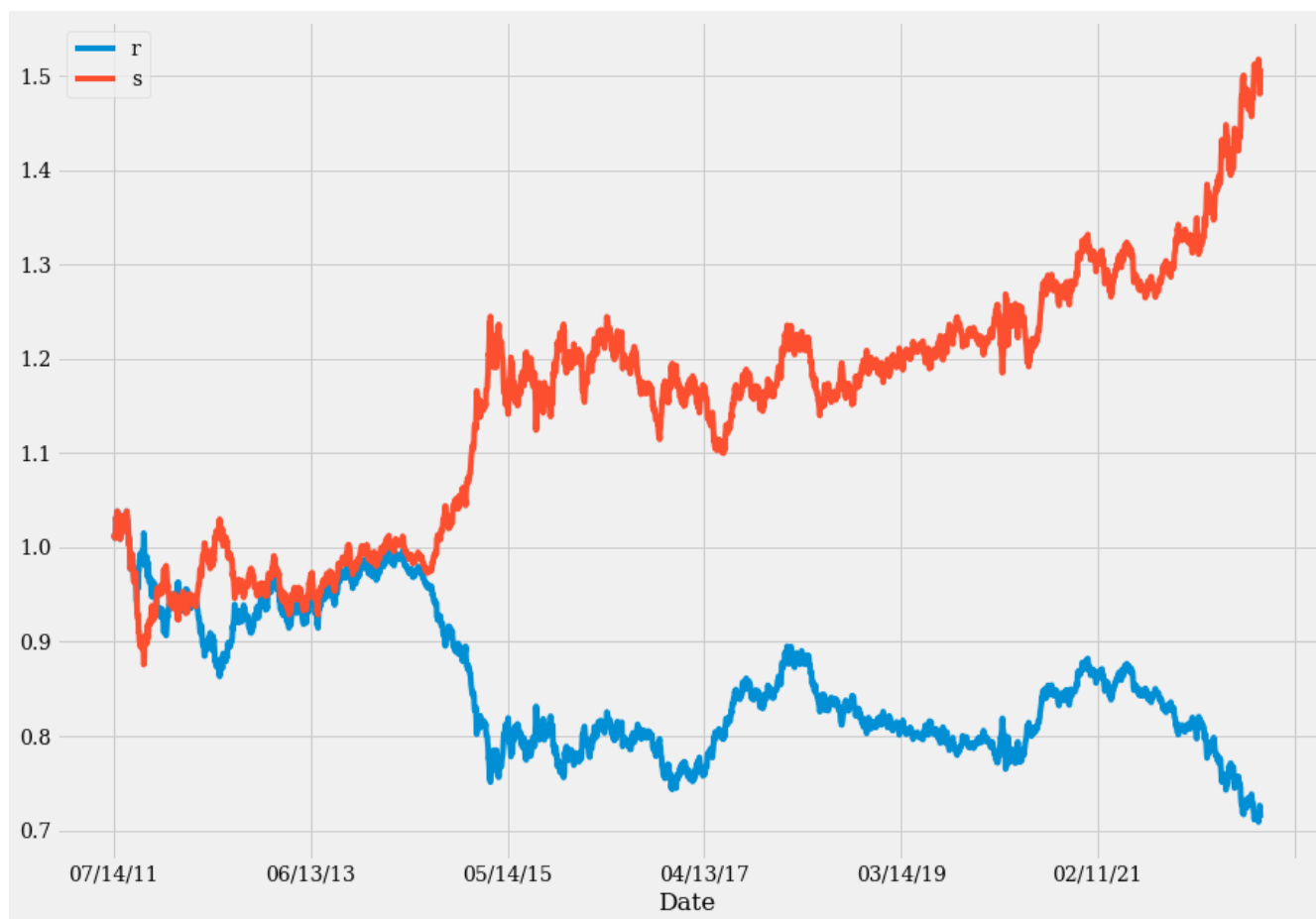
Κάθε φορά που ο βραχύς είναι πιο πάνω από τον μακρύ τότε θα αγοράσουμε μονάδες του χρηματοπιστωτικού προϊόντος (going long) καθώς αναμένουμε την άνοδο της αξίας του προϊόντος και αν ο βραχύς είναι πιο χαμηλά από τον μακροχρόνιο μέσο κινητό μέσο τότε θα προχωρήσουμε σε πωλήσεις του (going short) καθώς περιμένουμε πτώση της τιμής του. Παρακάτω μπορούμε να δούμε την οπτικοποίηση της προαναφερθείσας στρατηγικής.



Εικόνα 14. Καθορισμός στρατηγικής

Βλέπουμε ότι όταν ο βραχύς απλός κινητός μέσος είναι μεγαλύτερος από τον μακροχρόνιο παίρνουμε θέση long (τιμή +1) και όταν ο μακροχρόνιος είναι μεγαλύτερος παίρνουμε θέση short (τιμή -1).

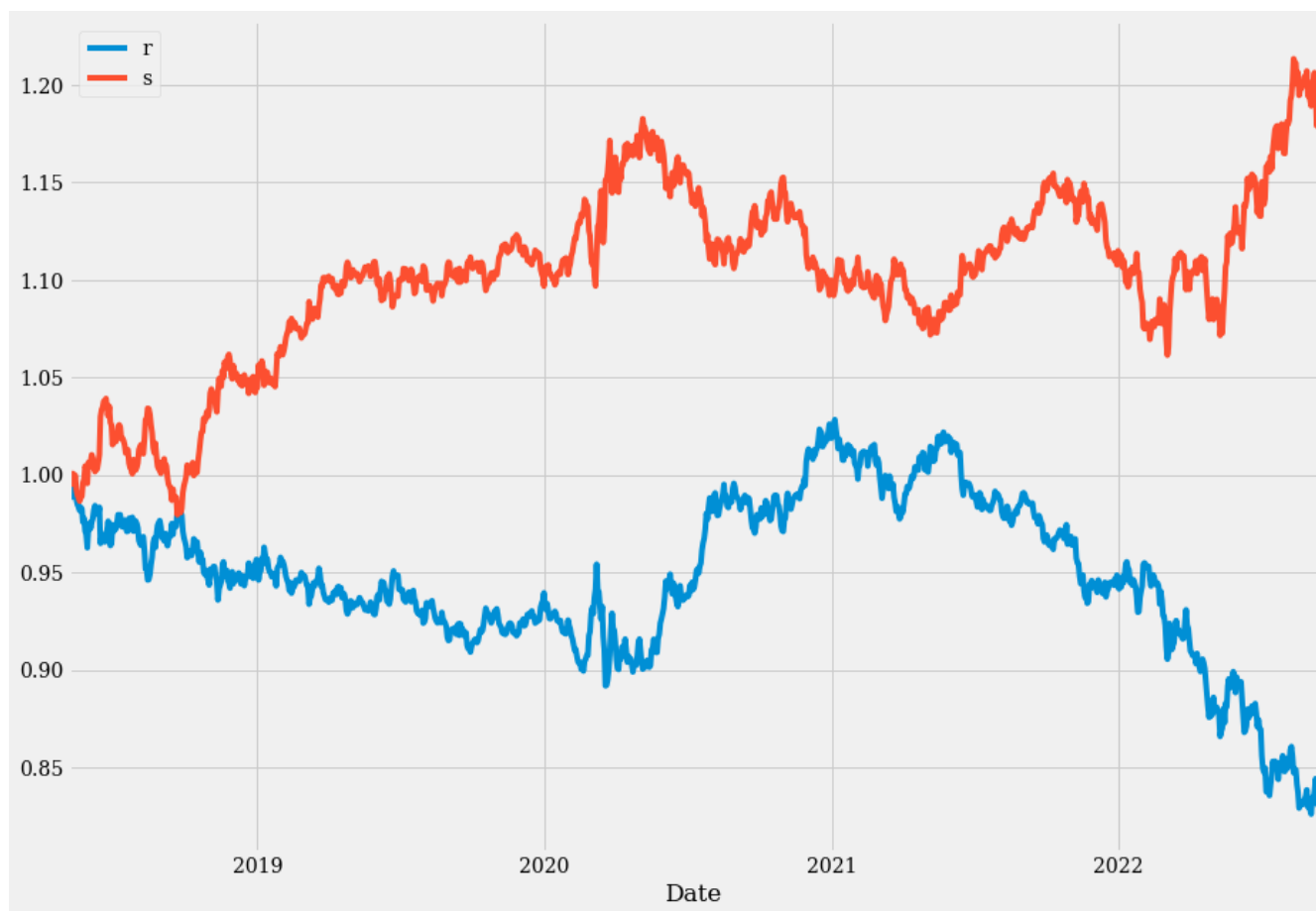
Στη συνέχεια θα συνδυάσουμε τις θέσεις μας (long ή short) με τις αποδόσεις του προϊόντος. Αυτό γίνεται με τον πολλαπλασιασμό δύο στηλών του dataframe. Όπως βλέπουμε στο παρακάτω διάγραμμα η στρατηγική που επιλέξαμε μας αποφέρει πολύ μεγαλύτερες αποδόσεις (s) από το να μην είχαμε καθόλου στρατηγική (παθητικό benchmark – r)



Εικόνα 15. Αποτελέσματα στρατηγικής βάσει αποδόσεων

Παρακάτω θα χρησιμοποιήσουμε τα ίδια δεδομένα για να εξετάσουμε την επενδυτική στρατηγική μας με τη χρήση νευρωνικών δικτύων. Σε αυτή την περίπτωση πέρα από τη μέθοδο του απλού κινητού μέσου θα χρησιμοποιήσουμε τις (κυλιόμενες) μέγιστες και ελάχιστες τιμές, τον λογαριθμικό μέσο όρο των αποδόσεων (ένδειξη τάσης ανόδου ή καθόδου), τη μεταβλητότητα των αποδόσεων.

Όλα τα παραπάνω στοιχεία θα χρησιμοποιήσουμε την εκπαίδευση του νευρωνικού δικτύου αφού τα διαχωρίσουμε για την εκπαίδευση και τον έλεγχο και τα κανονικοποιήσουμε. Μια πρόβλεψη προς τα πάνω μας κατευθύνει προς μια στρατηγική αγοράς του προϊόντος και μια πρόβλεψη προς τα κάτω μας κατευθύνει προς μια στρατηγική πώλησής του.



Εικόνα 16. Αποδόσεις με χρήση νευρωνικού δικτύου

Βλέπουμε ότι με τη χρήση του νευρωνικού δικτύου για τον καθορισμό της θέσης μας οι αποδόσεις μας (s) είναι πιο υψηλές από εκείνες του παθητικού benchmark (r).

36/36 [=====] - 0s 2ms/step - loss: 0.7141 - accuracy: 0.5224

Επίσης βλέπουμε ότι η ακρίβεια του μοντέλου φτάνει το 52%.

Όμως στην πραγματική αγορά όπου υπάρχουν τα κόστη (προμήθειες) των συναλλαγών και όπου εκεί οι συναλλαγές είναι πολυπληθέστερες, καταλήγουμε στο συμπέρασμα ότι οι συγκεκριμένες στρατηγικές δεν είναι εντέλει οικονομικά βιώσιμες μακροχρόνια.

Ουσιαστικά στην παρούσα διεργασία βασιζόμαστε μόνο σε αριθμητικά δεδομένα χωρίς να έχουμε ως οδηγό κάποια οικονομική θεωρία ή υποθέσεις σχετικά με τη συμπεριφορά των συμμετεχόντων στη συγκεκριμένη αγορά. Δηλαδή η προσπάθεια εστιάζει κυρίως στην εύρεση κάποιων στατιστικών φαινομένων από την ανάλυση των χρονοσειρών, χωρίς να λαμβάνει υπόψη της την ανθρώπινη συμπεριφορά και οικονομικές υποθέσεις.

## **5.2. Καθορισμός Στρατηγικής Βάσει Διαφορετικών Τεχνικών Δεικτών**

Πέρα από τον απλό κινητό μέσο, για την παρούσα ερευνα, χρησιμοποιήσαμε και άλλους τεχνικούς δείκτες για τον καθορισμό μιας αποδοτικής στρατηγικής. Σε γενικές γραμμές οι τεχνικοί δείκτες χωρίζονται σε δύο κατηγορίες. Στους lagging indicators και τους leading indicators. Ο απλός κινητός μέσος είναι ένας lagging indicator. Οι δείκτες που ανήκουν σε αυτή την κατηγορία χρησιμοποιούν παλαιές τιμές και παράγουν αποτελέσματα για την ανάλυση της αγοράς στον παρόντα χρόνο έτσι ώστε να δομηθεί μια μελλοντική στρατηγική. Ονομάζονται lagging γιατί ενώ ανιχνεύουν την κατεύθυνση της αγοράς τείνουν να έρχονται “καθυστερημένοι” καθώς η καθοδική ή η ανοδική τάση έχει ήδη ξεκινήσει πριν την εντοπίσει ο δείκτης. Συνήθως λειτουργούν ικανοποιητικά σε αγορές με ιδιαίτερα ευκρινείς τάσεις. Από την άλλη οι leading indicators, όπως ο δείκτης RSI που θα δούμε παρακάτω, ανιχνεύουν τη δυναμική της αγοράς και της πώλησης του προϊόντος πριν την έναρξη της ανοδικής ή καθοδικής τάσης. Στους επενδυτές υπάρχει πάντα ο προβληματισμός για το ποιος είναι ο πιο αποδοτικός συνδυασμός lagging και leading δεικτών.

Αν βασιστούμε μόνο στους leading δείκτες τότε είναι προφανές ότι θα αστοχήσουμε πολλές φορές στην πρόβλεψη της κατεύθυνσης της αγοράς. Αν βασιστούμε μόνο στους lagging δείκτες τότε είναι πολύ πιθανό να ενεργήσουμε με μεγάλη καθυστέρηση και να απωλέσουμε πολλά κέρδη. Συνακόλουθα η χρήση των δεικτών έγκειται στο είδος της στρατηγικής του επενδυτή και προφανώς στο είδος του χρηματοπιστωτικού προϊόντος που πραγματεύεται.

Αυτή τη φορά χρησιμοποιήσαμε για κάθε παρατήρηση, την ημερήσια τιμή ανοίγματος της ισοτιμίας, την ημερήσια τιμή κλεισίματος, αλλά και τη μέγιστη και ελάχιστη τιμή που είχε μέσα στην ημέρα.

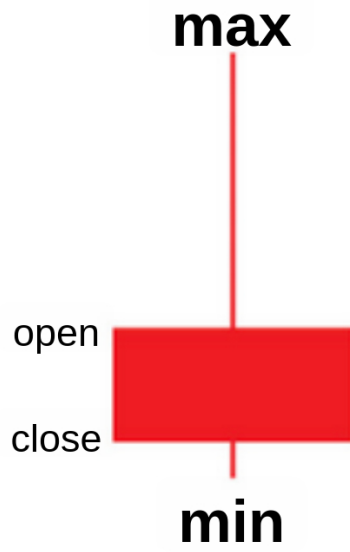
Ο σκοπός μας ήταν η εύρεση μοτίβων από την οπτικοποίηση σε μορφή candlesticks (leading indicator) των αριθμητικών δεδομένων.



Εικόνα 17. Παρατηρήσεις υπό μορφή candlesticks

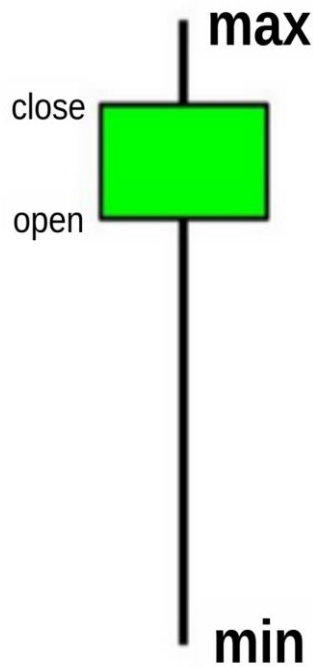
Κάθε παρατήρηση οπτικοποιείται μέσω ενός κόκκινου ή πράσινου παραλληλόγραμμου και από μια κάθετη γραμμή που το διαπερνά. Το κύριο σώμα του candlestick αντιπροσωπεύει το εύρος της μεταξύ της τιμής ανοίγματος της ισοτιμίας και της τιμής κλεισίματος. Όταν το κύριο σώμα είναι πράσινο τότε η τιμή ανοίγματος είναι χαμηλότερη από την τιμή κλεισίματος, και όταν είναι κόκκινο τότε αυτό σημαίνει ότι η τιμή κλεισίματος είναι χαμηλότερη από την τιμή ανοίγματος. Η κάθετη γραμμή αντιπροσωπεύει το εύρος ανάμεσα στη μέγιστη και την ελάχιστη ημερήσια τιμή της κάθε παρατήρησης.





Εικόνα 18. Παρατήρηση μοτίβου διάττοντος αστέρα

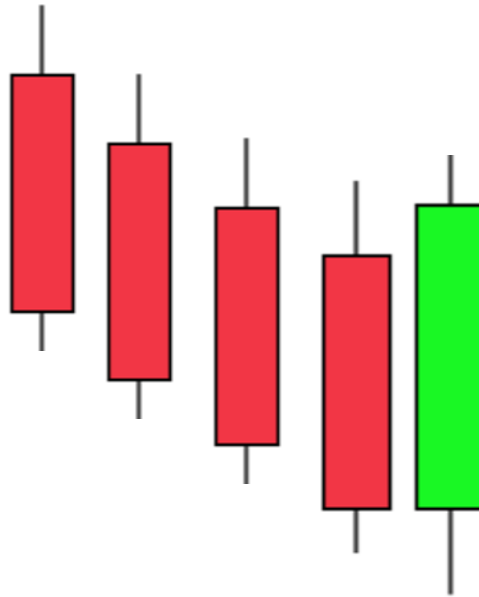
Ένα μοτίβο που χρησιμοποιήσαμε ονομάζεται διάττων αστέρας (shooting star). Στη συγκεκριμένη παρατήρηση βλέπουμε μια πολύ υψηλή μέγιστη τιμή σε σχέση με την τιμή ανοίγματος, και την ελάχιστη τιμή να βρίσκεται πολύ κοντά στη τιμή κλεισίματος. Δηλαδή ψάχνουμε μοτίβα των οποίων το εύρος ανάμεσα στη μέγιστη τιμή και την τιμή ανοίγματος, να είναι τουλάχιστον διπλάσιο από το κύριο σώμα της παρατήρησης. Επίσης θα πρέπει το εύρος ανάμεσα στην ελάχιστη τιμή και την τιμή κλεισίματος να είναι το πολύ το μισό του εύρους της μέγιστης τιμής από την τιμή ανοίγματος. Το συγκεκριμένο μοτίβο είναι ένδειξη για πτώση της ισοτιμίας συνεπώς εμείς θα κινηθούμε προς πώληση. Έχοντας τα παραπάνω ως δεδομένα σχηματίζουμε την αντίστοιχη συνάρτηση.



Εικόνα 19. Παρατήρηση μοτίβου ένδειξης αύξησης ισοτιμίας

Με αντίστοιχο τρόπο ψάχνουμε για μοτίβα τα οποία κρύβουν ενδεχόμενη αύξηση της ισοτιμίας όπως βλέπουμε στο παραπάνω σχήμα. Κατά αυτόν τον τρόπο παράγουμε τις θέσεις της στρατηγικής μας. Δηλαδή αν θα αγοράσουμε ή αν θα πουλήσουμε. Βέβαια υπάρχει και η περίπτωση να μην παρατηρηθεί κάποιο αντίστοιχο μοτίβο από τα δεδομένα, συνεπώς σε αυτήν την περίπτωση δεν θα κινηθούμε στην αγορά. Αυτά ακριβώς τα δεδομένα που περιγράφουν την στρατηγική μας θα τροφοδοτηθούν στον αλγόριθμο για την εκπαίδευσή του. Για τη συγκεκριμένη μελέτη επιλέξαμε τη χρήση κι άλλων μοτίβων για τον καθορισμό της στρατηγικής μας.

Συγκεκριμένα επιλέξαμε τα μοτίβα Bullish Engulfing (όταν αναμένουμε αύξηση της ισοτιμίας) και Bearish (όταν αναμένουμε πτώση της).



Εικόνα 20. Bullish μοτίβο

Στο πάνω σχήμα βλέπουμε ένα Bullish μοτίβο. Αυτό παρατηρείται όταν έχουμε συνεχείς πτώσεις της ισοτιμίας ακολουθούμενο από candlestick ανόδου. Συνήθως η τιμή κλεισίματος της τελευταίας καθοδικής παρατήρησης είναι περίπου στο ίδιο επίπεδο με την τιμή ανοίγματος της ανοδικής, και η τιμή κλεισίματος της ανοδικής είναι υψηλότερη από την τιμή κλεισίματος της προηγούμενης καθοδικής παρατήρησης.

Έχουμε τις αντίστοιχες παρατηρήσεις για το Bearish Engulfing μοτίβο, όπου εκεί μετά την ανοδική τάση εμφανίζεται πτωτικό candlestick με τιμή κλεισίματος χαμηλότερη από την τιμή ανοίγματος της αμέσως προηγούμενης ανοδικής παρατήρησης. Χρησιμοποιώντας τις παραπάνω παραδοχές σχηματίζουμε τη συνάρτηση που θα εντοπίσει τα συγκεκριμένα μοτίβα.

Στην τεχνική ανάλυση που κάνουμε θα λάβουμε υπόψη επίσης δύο ακόμη σημαντικές έννοιες-μοτίβα που πιστεύεται ότι καθορίζουν την κατεύθυνση της αγοράς. Αυτές οι δυο έννοιες αποτελούν τα επίπεδα στήριξης και αντίστασης (support and resistance) που παρατηρούνται στα δεδομένα μας. Σε μια καθοδική τάση της αγοράς, οι τιμές υποχωρούν γιατί υπάρχει υπερβάλλουσα προσφορά σε σχέση με τη ζήτηση. Όσο πιο πολύ πέφτουν οι τιμές, τόσο πιο θελκτικές είναι για τους επενδυτές που επιθυμούν να αγοράσουν το προϊόν.

Σε κάποια χρονική στιγμή όμως, και αφού παρατηρηθεί σταδιακή αύξηση της ζήτησης, η ζήτηση θα φτάσει στα ίδια επίπεδα με την προσφορά. Σε αυτό ακριβώς το σημείο η ισοτιμία θα σταματήσει να υποχωρεί. Αυτό ακριβώς το σημείο το ονομάζουμε επίπεδο στήριξης (support level). Το επίπεδο αντίστασης (resistance) αποτελεί ακριβώς το αντίθετο. Η ισοτιμία κινείται προς τα πάνω γιατί υπάρχει μεγαλύτερη ζήτηση από προσφορά.

Όσο η ισοτιμία αυξάνεται, θα υπάρξει κάποιο χρονικό σημείο όπου η πρόθεση των επενδυτών για πώληση θα υπερिσχύσει εκείνης για αγορά. Συνακόλουθα θα παρατηρηθεί πτώση της ισοτιμίας.



Εικόνα 20. Μοτίβα Αντίστασης - Στήριξης

Για τη δημιουργία της αντίστοιχης συνάρτησης και συγκεκριμένα για τον εντοπισμό του support level, θα ανιχνεύσουμε στα δεδομένα, μια αλληλουχία πτωτικών παρατηρήσεων με ολοένα και περισσότερο πτωτικές ελάχιστες τιμές, τις οποίες θα ακολουθήσουν ανοδικές παρατηρήσεις με αύξοντα ρυθμό των ελάχιστων τιμών τους. Το support level είναι η ελάχιστη τιμή της τελευταία χρονικά πτωτικής παρατήρησης. Ακριβώς το αντίστροφο θα ανιχνεύσουμε στα resistance levels.

Δηλαδή θα εντοπίσουμε αλληλουχίες ανοδικών παρατηρήσεων με ανοδική μέγιστη τιμή, ακολουθούμενες από αλληλουχία πτωτικών παρατηρήσεων των οποίων ο ρυθμός των μέγιστων τιμών θα είναι κι αυτός πτωτικός. Το resistance level είναι η μέγιστη τιμή της τελευταίας χρονικά ανοδικής παρατήρησης.

Για το συγκεκριμένο παράδειγμα θα χρησιμοποιήσουμε συνδυαστικά τα προαναφερθέντα μοτίβα ώστε να εντοπιστούν στα δεδομένα ανοδικές και καθοδικές τάσεις. Εκτός από τις συναρτήσεις των μοτίβων θα χρησιμοποιήσουμε ως είσοδο στον αλγόριθμο και τον δείκτη RSI (Relative Strength Index).

Ο συγκεκριμένος δείκτης αναδεικνύει περιπτώσεις όπου υπάρχουν παροδικές στρεβλώσεις στην αγορά. Δηλαδή είτε όταν ένα προϊόν έχει δεχτεί πιέσεις για να ανέβει η τιμή του λόγω κάποιων συγκυριών αλλά αναμένεται σύντομα η επιστροφή του σε λογικά επίπεδα. Αντίστοιχα και στην περίπτωση της μη λογικής πτώσης της τιμής του όταν αυξάνεται ο ρυθμός της πώλησής του, ενώ η πραγματική του τιμή, κανονικά, θα έπρεπε να είναι υψηλότερη. Συνήθως παρατηρείται η πλασματική άνοδος της ισοτιμίας όταν ο δείκτης βρίσκεται πάνω από τις 70 μονάδες και η πλασματική πτώση όταν ο δείκτης βρίσκεται κάτω από τις 30.

Στη συγκεκριμένη περίπτωση θα χρησιμοποιήσουμε ένα feedforward τεχνητό νευρωνικό δίκτυο έτσι ώστε να κατατάξουμε τις παρατηρήσεις μας σε 3 τάξεις/κατηγορίες. Θα χρησιμοποιήσουμε τον ταξινομητή Multilayer Perceptron ο οποίος αποτελείται από πολλαπλά επίπεδα τα οποία συνδέονται μεταξύ τους. Οι κόμβοι των επιπέδων αποτελούν τους νευρώνες οι οποίοι ενεργοποιούνται με μη γραμμικές συναρτήσεις (εκτός των κόμβων του επιπέδου εισόδου). Ανάμεσα στα επίπεδα της εξόδου και της εισόδου μπορούν να υπάρχουν ένα ή περισσότερα κρυμμένα επίπεδα.

Ως δεδομένα εισαγωγής έχουμε ουσιαστικά έναν πίνακα με τέσσερις στήλες. Τα δεδομένα του RSI, και άλλες 3 κατηγορικές στήλες όπου θα υποδεικνύουν στον αλγόριθμο το είδος του σήματος που θα καθορίσει την στρατηγική μας, δηλαδή τις στήλες signal\_0, signal\_1, signal\_2, όπου θα παίρνουν μόνο τις τιμές 0 ή 1. Όταν δεν παρατηρείται μια κάποια συγκεκριμένη κατεύθυνση στην αγορά τότε στη στήλη signal\_0 θα έχουμε την τιμή 1 και στις υπόλοιπες στήλες την τιμή 0.

Όταν έχουμε καθοδική κατεύθυνση η στήλη `signal_1` θα παίρνει την τιμή 1 και οι υπόλοιπες στήλες την τιμή 0, και τέλος όταν έχουμε ανοδική κατεύθυνση στην αγορά τότε η στήλη `signal_2` θα παίρνει την τιμή 1 και υπόλοιπες την τιμή 0. Χωρίζουμε τα δεδομένα 70/30 για την εκπαίδευση και τον για τον έλεγχο. Θα χρησιμοποιήσουμε έξι hidden layers και relu για τη συνάρτηση ενεργοποίησης.

## 6. Αποτελέσματα

Train Results	Accuracy: 49.6529%
Test Results	Accuracy: 29.6986%

Όπως βλέπουμε τα αποτελέσματα δεν είναι καλά και αυτός ο τρόπος πρόβλεψης δεν είναι ιδιαίτερα αποδοτικός. Δεν μπορούμε να χρησιμοποιήσουμε αυτή την τεχνική πρόβλεψης στην αγορά. Βγάζουμε το συμπέρασμα ότι το τεχνητό νευρωνικό δίκτυο δεν είναι το κατάλληλο εργαλείο για να προβλέψουμε την αλλαγή της κατεύθυνσης μιας αγοράς. Είναι πολύ πιθανό να λειτουργήσει πιο αποδοτικά όταν έχει να προβλέψει μόνο ένα είδος κίνησης (πτωτικό ή ανοδικό).

	precision	recall	f1-score	support
0.0	0.00	0.00	0.00	252
1.0	0.55	0.28	0.37	1204
2.0	0.49	0.82	0.62	1281
accuracy			0.51	2737
macro avg	0.35	0.37	0.33	2737
weighted avg	0.47	0.51	0.45	2737

	precision	recall	f1-score	support
0.0	0.00	0.00	0.00	670
1.0	0.25	0.16	0.20	571
2.0	0.31	0.77	0.44	584
accuracy			0.30	1825
macro avg	0.19	0.31	0.21	1825
weighted avg	0.18	0.30	0.20	1825

Εικόνα 20. Confusion Matrix

Όπως βλέπουμε στο παραπάνω σχήμα (confusion matrix) για την κατηγορία 0 το μοντέλο δεν έχει προβλέψει κάτι. Αυτό συμβαίνει γιατί η πρώτη κατηγορία εμφανίζεται τις πιο λίγες φορές σε σχέση με τις υπόλοιπες δύο. Αυτό δεν μας πειράζει καθώς ενδιαφερόμαστε κυρίως για τις κατηγορίες που δείχνουν την ανοδική ή καθοδική κίνηση της αγοράς. Όσον αφορά τα δεδομένα εκπαίδευσης, μπορούμε να δούμε ότι για την πρώτη κατηγορία το μοντέλο ανιχνεύει το 28% των περιπτώσεων που υπάρχουν στα δεδομένα και έχει ακρίβεια 55%.

Το 28% υποδεικνύει ότι είναι πολύ δύσκολο για το μοντέλο να ανιχνεύσει την πλειοψηφία των τάσεων στην αγορά. Για την κατηγορία 2 βλέπουμε ότι το μοντέλο είναι πολύ πιο ευαίσθητο στην ανίχνευση των τάσεων της αγοράς. Η ακρίβειά του είναι στο 49%.

Για το test dataset βλέπουμε ότι το μοντέλο είναι ακόμη λιγότερο αποδοτικό. Συμπερασματικά το παρόν μοντέλο από μόνο του, για το παρόν πείραμα, είναι ακατάλληλο για την πρόβλεψη τάσεων στην αγορά και για τη κατασκευή μιας πραγματικά κερδοφόρας στρατηγικής.



## 7. Συμπεράσματα

Στην παρούσα εργασία χρησιμοποιήσαμε ως εργαλεία διάφορους τύπους τεχνικών δεικτών (SMA, μοτίβα candlesticks, δείκτης RSI), τα δεδομένα των οποίων εισαγάγαμε σε τεχνητά νευρωνικά δίκτυα έτσι ώστε να ερευνήσουμε αν μπορεί να δομηθεί μια κερδοφόρα στρατηγική στην αγορά συναλλάγματος.

Συμπερασματικά, η επιλογή της στρατηγικής μας μέσω της πρόβλεψης των τάσεων της αγοράς, δεν είναι αρκετή. Κάθε πρόβλεψη θα πρέπει να ελεγχθεί υπό το πρίσμα λεπτομερών κανόνων οι οποίοι καθορίζουν την τελική απόφαση. Πολύ σημαντικά είναι τα στοιχεία της συνεχούς παρακολούθησης της αγοράς, και η εφαρμογή μέτρων αποφυγής κινδύνου και απώλειας κεφαλαίου.

Για μια πραγματικά μακροχρόνια κερδοφόρα στρατηγική επιβάλλεται η χρήση ανάλυσης κλασικού τύπου, όπου μεταβλητές όπως η οικονομική ανάπτυξη μιας χώρας, ο πληθωρισμός, τα τραπεζικά επιτόκια, το εμπορικό ισοζύγιο, οι συνθήκες στην αγορά εργασίας, η γεωπολιτική συγκυρία, θα συνδυαστούν με μια σειρά από τεχνικούς δείκτες όπως αυτοί που είδαμε στην παρούσα εργασία. Οι ενδείξεις των ανωτέρω τεχνικών δεικτών θα χρησιμοποιηθούν ως δεδομένα εισόδου σε αλγόριθμους μηχανικής μάθησης έτσι ώστε να έχουμε μια συνολική εικόνα για τη δόμηση μιας στρατηγικής που θα αποφέρει σημαντικές αποδόσεις σε βάθος χρόνου.

## 8. Βιβλιογραφία

1. Marcos López de Prado, 2018, 'Advances in Financial Machine Learning', John Wiley & Sons, Inc.
2. Laurent Bernut, 2021, 'Algorithmic Short Selling with Python', Packt Publishing Ltd.
3. Damon Lee, 2021, 'Build Your Own AI Investor with Machine Learning and Python', Damon Lee
4. Fumiya Akashi, Masanobu Taniguchi, Anna Clara Monti, Tomoyuki Amano, 2021, 'Diagnostic Methods in Time Series', Springer Nature Singapore Pte Ltd. <https://doi.org/10.1007/978-981-16-2264-9>
5. Yves Hilpisch, 2022, 'Financial Theory with Python', O'Reilly Media, Inc.
6. Jiri Pik, Sourav Ghosh, 2021, 'Hands-On Financial Trading with Python', Packt Publishing Ltd.
7. Jeffrey Ng, CFA, Subhash Shah, 2020, 'Hands-On Artificial Intelligence for Banking', Packt Publishing Ltd.
8. Luigi Troiano, Arjun Bhandari, Elena Mejuto Villa, 2020, 'Hands-On Deep Learning for Finance', Packt Publishing Ltd.
9. Stefan Jansen, 2020, 'Hands-On Machine Learning for Algorithmic Trading', Packt Publishing Ltd.
10. David Easley, Marcos López de Prado, Maureen O'Hara, 2013, 'High-Frequency Trading - New Realities for Traders, Markets and Regulators', Risk Books, a Division of Incisive Media Investments Ltd
11. Irene Aldridge, 2010, 'High-Frequency Trading - A Practical Guide to Algorithmic Strategies and Trading Systems', John Wiley & Sons, Inc.
12. Sebastien Donadio, Sourav Ghosh, 2019, 'Learn Algorithmic Trading', Packt Publishing Ltd.
13. Jannes Klaas, 2019, 'Machine Learning for Finance', Packt Publishing Ltd.
14. Ernest P. Chan, 2021, 'Quantitative Trading - How to Build Your Own Algorithmic Trading Business', John Wiley & Sons, Inc.
15. Yves Hilpisch, 2021, 'Artificial Intelligence in Finance', O'Reilly Media, Inc.

16. Stephen Marsland, 2015, 'Machine Learning – An Algorithmic Perspective', Taylor & Francis Group, LLC
17. Charu C. Aggarwal, 2018, 'Neural Networks and Deep Learning', Springer International Publishing AG, <https://doi.org/10.1007/978-3-319-94463-0>
18. Ian Goodfellow, Yoshua Bengio, Aaron Courville, 2016, 'Deep Learning', Massachusetts Institute of Technology
19. Sudharsan Ravichandiran, 2019, 'Hands-On Deep Learning Algorithms with Python', Packt Publishing Ltd.