



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Πληροφορική»

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Χρήση τεχνικών μηχανικής μάθησης στις αυτοματοποιημένες συναλλαγές: Μια ανασκόπηση Machine Learning techniques in automated trading: A review
Όνοματεπώνυμο Φοιτητή	Νικόλαος Γκούρας
Πατρώνυμο	Γεώργιος
Αριθμός Μητρώου	ΜΠΠΛ17010
Επιβλέπων	Διονύσιος Σωτηρόπουλος, Επίκουρος Καθηγητής

Τριμελής Εξεταστική Επιτροπή

(υπογραφή)

(υπογραφή)

(υπογραφή)

Δ. Σωτηρόπουλος
(Επίκ. Καθηγητής)

Ε. Σακκόπουλος
(Ανάπλ. Καθηγητής)

Γ. Τσιχριντζής
(Καθηγητής)

Περιεχόμενα

Περίληψη	4
Abstract	4
1.Εισαγωγή	5
1.1 Βιβλίο Οριακών Εντολών - Limit Order Book	5
1.2 Εντολή Αγοράς - Market Order	5
1.3 Οριακή Εντολή - Limit Order	5
1.4 Long and Short Position	6
2.Μηχανική Μάθηση	6
2.1 Νευρωνικά δίκτυα (Neural Networks – NN)	6
2.2 Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines – SVM)	7
2.3 Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Network – CNN)	8
2.4 Επαναλαμβανόμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNN)	9
2.5 Δίκτυα Μακράς Βραχυπρόθεσμης Μνήμης (Long Short – term Memory – LSTM)	9
3. Οι μελέτες.....	10
3.1 M.A.H. Dempster, V. Leemans (2006).....	10
Το βασικό μοντέλο	10
Δεδομένα εισόδου	10
Αποτελέσματα	11
3.2 Kercheval, & Zhang (2013)	12
Βασικό μοντέλο	12
Δεδομένα εισόδου	12
Αποτελέσματα	13
3.3 Tsantekidis et al (2017)	14
Βασικό μοντέλο	14
Δεδομένα εισόδου	14
Αποτελέσματα	16
3.4 Selvin et al (2017).....	16
Βασικό μοντέλο	16
Δεδομένα εισόδου	16
Αποτελέσματα	17
3.5 Nousi et al (2019)	19
Βασικό Μοντέλο	19
Δεδομένα Εισόδου	20
	2

Μεταπτυχιακή Διατριβή	Γκούρας Νικόλαος
Αποτελέσματα	21
3.6 Tsantekidis et al, (2020)	22
Βασικό μοντέλο	22
Δεδομένα Είσοδου	23
Αποτελέσματα	24
4. Σύγκριση Αποτελεσμάτων	25
Βιβλιογραφία	26

Περίληψη

Στην παρούσα πτυχιακή εργασία γίνεται μια προσπάθεια ανασκόπησης των εφαρμογών τεχνικών μηχανικής μάθησης στη χρηματιστηριακή αγορά. Οι συνολικά έξι μελέτες που παρουσιάζονται αξιοποιούν έναν μεγάλο όγκο πραγματικών χρηματοοικονομικών δεδομένων, όπως το «βιβλίο οριακών εντολών» (limit order book) με μικρό χρονικό βήμα, ως δεδομένα εκπαίδευσης σε συστήματα μηχανικής μάθησης, με σκοπό την πρόβλεψη συναλλακτικών ευκαιριών στη χρηματιστηριακή αγορά σε μια μελλοντική χρονική στιγμή. Χρησιμοποιήθηκαν αρχιτεκτονικές πολυεπίπεδων νευρωνικών δικτύων (MLP), μηχανών διανυσμάτων υποστήριξης πολλαπλών κλάσεων (SVM), επαναλαμβανόμενων νευρωνικών δικτύων (RNN), συστημάτων μακράς βραχυπρόθεσμης μνήμης (LSTM) και συνελκτικών νευρωνικών δικτύων (CNN) ή και συνδυασμός των παραπάνω σε ποικίλα σύνολα δεδομένων. Σαν αποτέλεσμα, η αρχιτεκτονική των CNN φαίνεται να υπερσχύει των άλλων τεχνικών στην πρόβλεψη των τιμών των χρηματιστηριακών προϊόντων, λόγω της δυναμικής φύσης του χρηματιστηριακού περιβάλλοντος.

Abstract

This project is an attempt to review some machine learning applications in the field of the stock market. Six such studies were taken under consideration, each of which utilizes a large amount of market data, such as the limit order books of stocks with small time steps, as training data in Machine Learning systems, aiming to predict stock exchange opportunities in the short future. Various machine learning techniques and architectures are being tested including Multi – Layer Perceptrons (MLP), Support Vector Machines (SVM), Recurrent Neural Networks (RNN), Long Short – Term Memory RNNs (LSTM), Convolutional Neural Networks (CNN) or combinations of the above. Based on the studies' results, the CNN architecture seems to be a more appropriate solution on the given problem, due to the dynamic nature of the stock market environment.

1. Εισαγωγή

Η εύκολη πρόσβαση σε μεγάλο όγκο και ποικιλία δεδομένων, όπως και η εξέλιξη των τεχνολογικών μέσων σε υπολογιστική ισχύ, αποθηκευτική δυνατότητα και ταχύτητα διασύνδεσης, έχουν καταστήσει ραγδαία την πρόοδο στον τομέα της τεχνητής νοημοσύνης. Η πιο σύγχρονη μορφή της, η μηχανική μάθηση, αξιοποιώντας τα παραπάνω έχει βρει εφαρμογή ήδη σε μεγάλο εύρος επαγγελματικών και επιστημονικών δραστηριοτήτων, παρέχοντας αξιόπιστες και ταχείες λύσεις σε άλλοτε χρονοβόρα ή δυσεπίλυτα προβλήματα.

Στην εργασία αυτή γίνεται μια ανασκόπηση της επιστημονικής έρευνας που έχει διεξαχθεί και διερευνά τη χρήση μηχανικής μάθησης για την πραγματοποίηση χρηματοοικονομικών συναλλαγών, με τη χρήση μεγάλου όγκου πραγματικών δεδομένων. Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή στην χρηματοοικονομική ορολογία που θα χρησιμοποιηθεί παρακάτω, ενώ στη συνέχεια παρατίθενται μερικές πληροφορίες για τον κλάδο της μηχανικής μάθησης και των τεχνικών της που χρησιμοποιήθηκαν στην υπό ανασκόπηση βιβλιογραφία. Το τρίτο κεφάλαιο περιέχει τη συνοπτική παρουσίαση έξι πειραματικών μοντέλων μηχανικής μάθησης που δοκιμάστηκαν, καθώς και τα αποτελέσματά τους, τα οποία παρουσιάζονται συγκριτικά στο τελευταίο κεφάλαιο.

1.1 Βιβλίο Οριακών Εντολών - Limit Order Book

Το βιβλίο οριακών εντολών (Limit Order Book – LOB) μιας μετοχής αποτελείται από το σύνολο των εντολών αγοράς και πώλησής της σε όλα τα επίπεδα τιμών. Πρακτικά, περιγράφει την εικόνα της πραγματικής προσφοράς και ζήτησης για κάθε μετοχή, αφού σε αυτό «δηλώνεται» η πρόθεση αγοράς ή πώλησης από τον κάθε ενδιαφερόμενο, τόσο σε πλήθος μετοχών όσο και σε επίπεδο τιμής μετοχής. Ο μηχανισμός αυτός είναι, στην ουσία, ο δίαυλος για τη σύνδεση αγοραστών και πωλητών μετοχών. Το πότε και πώς εκτελούνται οι εντολές, καθώς και η κίνηση της τιμής της μετοχής, ορίζονται από αυτόν. (Sirignano, 2018).

Η μοντελοποίηση των βιβλίων οριακών εντολών είναι δύσκολη, λόγω της πολυπλοκότητας και της πολυδιάστασής τους. Η δυναμική ενός μη γραμμικού βιβλίου οριακών εντολών περιλαμβάνει εκατοντάδες επίπεδα τιμών, στα οποία μπορούν να υποβληθούν οι παραγγελίες (Sirignano, 2018).

Τα αποθέματα μπορούν να αγοραστούν ή να πωληθούν χρησιμοποιώντας μια οριακή εντολή που ορίζεται σε μια συγκεκριμένη τιμή. Εάν υποβάλει κάποιος μια οριακή εντολή σε αυτήν την τιμή, θα παραμείνει στο βιβλίο οριακών εντολών μέχρι να την εκτελεστεί ή να ακυρωθεί. Στο "limit order book" συγκεντρώνονται όλες οι οριακές εντολές, ανεξάρτητα από την τιμή στην οποία τοποθετούνται. Οι «προσφορές» και οι «ζητήσεις» είναι οι εντολές ορίου αγοράς και πώλησης, αντίστοιχα (Sirignano, 2018).

1.2 Εντολή Αγοράς - Market Order

Είναι μια εντολή αγοράς ή πώλησης μιας συγκεκριμένης ποσότητας ενός περιουσιακού στοιχείου στην καλύτερη τιμή στο βιβλίο οριακών παραγγελιών αμέσως. Μπορεί να χρειαστεί πολύς χρόνος για να εκτελεστεί μια εντολή ορίου, εάν η απαιτούμενη τιμή απέχει από την καλύτερη τιμή (Tsantekidis et al, 2017). Σε αυτήν την περίπτωση, η παραγγελία μπορεί να ακυρωθεί από τον έμπορο. Η πλευρά της προσφοράς (αγορά) και της ζήτησης (πώληση) είναι εκεί όπου οι παραγγελίες χωρίζονται. Η πλευρά της προσφοράς και η πλευρά της ζήτησης περιλαμβάνουν εντολές ταξινομημένες ανά τιμή σε αύξουσα και φθίνουσα σειρά, αντίστοιχα.

1.3 Οριακή Εντολή - Limit Order

Η οριακή εντολή είναι μια μορφή εντολής αγοράς ή πώλησης συγκεκριμένου αριθμού μετοχών σε προκαθορισμένη τιμή στις χρηματοοικονομικές αγορές μετοχών. Η πώληση 100 μετοχών στα 10 \$ η καθεμία, υποδηλώνει ότι ο πωλητής έχει κατά νου μια ελάχιστη τιμή για κάθε μετοχή. Οι εντολές ορίου αγοράς

(γνωστές και ως προσφορές) των 10 \$ υποδεικνύουν ότι ο αγοραστής θέλει να αγοράσει έναν συμφωνημένο αριθμό μετοχών, καθεμία από τις οποίες δεν θα κοστίζει περισσότερο από 10 \$ (Tsantekidis et al, 2017).

Υποβάλλοντας μια εντολή ορίου αγοράς, ο επενδυτής διασφαλίζει ότι η τιμή που πληρώνει δεν θα υπερβεί την τιμή ορίου. Η πλήρωση μιας παραγγελίας δεν είναι εγγυημένη και οι οριακές εντολές δεν θα εκτελεστούν εκτός εάν η τιμή ασφάλειας ταιριάζει με τις απαιτήσεις της παραγγελίας. Εάν η τιμή του περιουσιακού στοιχείου δεν αυξηθεί στο προκαθορισμένο επίπεδο, η παραγγελία δεν θα ολοκληρωθεί και ο επενδυτής θα χάσει την ευκαιρία να πραγματοποιήσει συναλλαγές.

Αυτό έρχεται σε αντίθεση με μια παραγγελία αγοράς, στην οποία μια συμφωνία ολοκληρώνεται στην τρέχουσα τιμή αγοράς χωρίς να έχει οριστεί περιορισμός τιμής.

1.4 Long and Short Position

Όταν μιλούν για μετοχές και δικαιώματα προαίρεσης, οι αναλυτές και οι διαπραγματευτές αγοράς αναφέρονται συχνά σε έναν επενδυτή που έχει long ή short position. Ενώ το long και το short σε χρηματοοικονομικά θέματα μπορούν να αναφέρονται σε πολλά πράγματα, σε αυτό το πλαίσιο, οι θέσεις long και short περιγράφουν πόσο απέχει το χαρτοφυλάκιο των μετοχών που κατέχει ένας επενδυτής χαρτοφυλάκιο που θα έπρεπε να κατέχει (Kramer, 2021). Πιο συγκεκριμένα:

Εάν ένας επενδυτής έχει long θέση, σημαίνει ότι ο επενδυτής έχει αγοράσει και κατέχει κάποιες μετοχές. Αντίθετα, εάν ο επενδυτής έχει short θέση, σημαίνει ότι ο επενδυτής οφείλει αυτές τις μετοχές σε κάποιον, αλλά δεν τις κατέχει ακόμη.

Συχνά, ο “short” επενδυτής δανείζεται τις μετοχές από μια χρηματιστηριακή εταιρεία (με λογαριασμό περιθωρίου) για να πραγματοποιήσει την παράδοση. Στη συνέχεια, με την ελπίδα ότι η τιμή της μετοχής θα πέσει, ο επενδυτής αγοράζει τις μετοχές σε χαμηλότερη τιμή για να αποπληρώσει τον έμπορο που τις δάνεισε. Εάν η τιμή δεν πέσει και συνεχίσει να ανεβαίνει, ο επενδυτής υφίσταται ζημιά.

Αντίθετα, στη long position ο επενδυτής που αγοράζει μια μετοχή σε χαμηλή τιμή με την προσδοκία ότι η τιμή της θα ανέβει μακροπρόθεσμα και θα μπορέσει να βγάλει κέρδος από την πώλησή της αργότερα.

Οι θέσεις long και short χρησιμοποιούνται από τους επενδυτές για την επίτευξη διαφορετικών αποτελεσμάτων και συχνά οι θέσεις long και short δημιουργούνται ταυτόχρονα από έναν επενδυτή για τη μόχλευση ή την παραγωγή εισοδήματος από έναν τίτλο.

2. Μηχανική Μάθηση

Ένας γενικός ορισμός της μηχανικής μάθησης (Machine Learning – ML) είναι ο εξής:

Ένα πρόγραμμα υπολογιστή λέμε ότι μαθαίνει από την εμπειρία E ως προς κάποια κλάση εργασιών T και μέτρο απόδοσης P , αν η απόδοσή του σε εργασίες από το T , όπως μετριέται από το P , βελτιώνεται μέσω της εμπειρίας E (Mitchell, 1997).

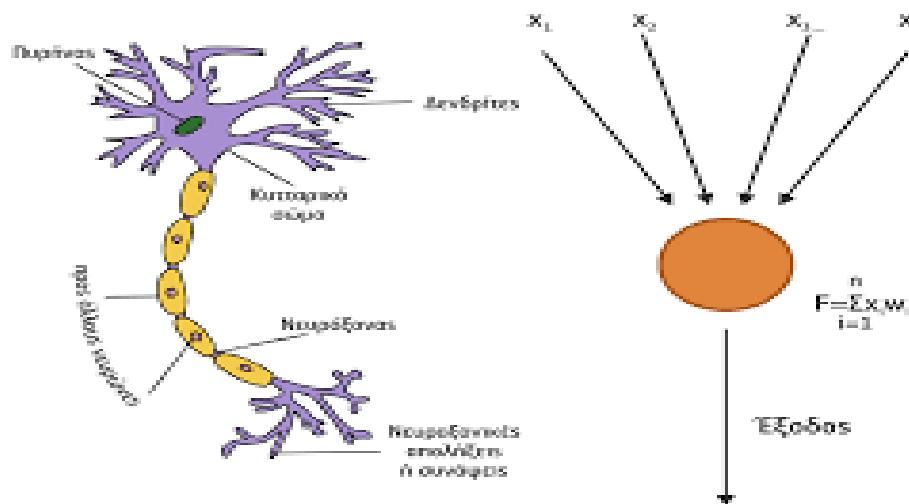
Πρακτικά, τα προγράμματα που χρησιμοποιούν μηχανική μάθηση περνούν από μια διαδικασία εκπαίδευσης, κατά τη διάρκεια της οποίας αξιολογούν και επαναπροσδιορίζουν ανάλογα τις παραμέτρους τους, με τρόπο που δεν έχει προγραμματιστεί αυστηρά.

2.1 Νευρωνικά δίκτυα (Neural Networks – NN)

Τα νευρωνικά δίκτυα είναι μια τεχνική μηχανικής μάθησης που βασίζεται στην ανθρώπινη βιολογία, προσπαθώντας να αναπαραστήσει σε αλγοριθμικό επίπεδο τη συμπεριφορά του ανθρώπινου νευρικού συστήματος κατά τη διαδικασία της μάθησης.

Τα ανθρώπινα νευρικά κύτταρα, ή αλλιώς νευρώνες, είναι υπεύθυνα για την πρόσληψη ενός σήματος από το περιβάλλον και τη μεταφορά του, μέσω του νευρικού συστήματος, στον εγκέφαλο. Η πρόσληψη του σήματος γίνεται μέσω των «δενδριτών» και, ύστερα από επεξεργασία στον νευρώνα, το σήμα

προωθείται μέσω του νευροάξονα στον επόμενο νευρώνα μέσω των νευρικών συνάψεων, των συνδέσεων δηλαδή μεταξύ νευρώνων. Η μεταφορά της πληροφορίας διεξάγεται με ηλεκτρικά σήματα, τα οποία ενεργοποιούν τους δενδρίτες των νευρώνων (και επομένως και τις συνάψεις τους) εφόσον η ένταση των σημάτων ξεπερνά μια ορισμένη τιμή. Η τιμή αυτή για κάθε νευρώνα, καθώς και ο αριθμός των συνάψεων του, μεταβάλλεται κατά τη διάρκεια της ζωής ενός ανθρώπου, ανάλογα με τα εξωτερικά ερεθίσματα τα οποία δέχεται. Αυτή η διαδικασία της ρύθμισης των συνάψεων είναι η διαδικασία της μάθησης για έναν βιολογικό οργανισμό (Aggarwal, 2018). Στην Εικόνα 1 απεικονίζεται μια αναπαράσταση του ανθρώπινου νευρώνα με τα τμήματά του, ενώ σε αντιπαραβολή φαίνεται μια σχηματική απεικόνιση ενός ψηφιακού νευρώνα.



Εικόνα 1: Μοντέλο νευρώνα (Γεωργούλη, 2015)

Στην απεικόνιση αυτή, οι δενδρίτες αντικαθίστανται με τις εισόδους στον ψηφιακό νευρώνα, ενώ ο νευροάξονας είναι ουσιαστικά η έξοδος του συστήματος. Οι τιμές εισόδου (x_i) ζυγίζονται με ένα βάρος (w_i) και στη συνέχεια αθροίζονται, ενώ η έξοδος παράγεται μόνο όταν το άθροισμα αυτό ξεπερνά μια ορισμένη τιμή. Κατά τη διαδικασία της μάθησης, το perceptron (όπως ονομάζεται το παραπάνω μοντέλο) τροφοδοτείται με μεγάλο αριθμό εισόδων και αναμενόμενων εξόδων, που χρησιμεύουν στην αξιολόγηση της απόδοσης. Αν το ποσοστό αστοχιών είναι σε υψηλότερα από το επιθυμητό πλαίσιο, αναθεωρούνται οι τιμές των βαρών (w_i) και επαναλαμβάνεται η διαδικασία, μέχρις ότου το επίπεδο σφάλματος να είναι εντός αποδεκτών ορίων (Aggarwal, 2018).

Συχνά, για την ανάπτυξη μιας εφαρμογής μηχανικής μάθησης χρησιμοποιούνται δίκτυα από perceptrons, τα οποία οργανώνονται σε επίπεδα (layers) με τέτοιο τρόπο, ώστε η έξοδος του ενός να είναι είσοδος του επόμενου. Η αρχιτεκτονική αυτή ονομάζεται Multi-Layer Perceptron (MLP).

2.2 Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines – SVM)

Οι μηχανές διανυσμάτων υποστήριξης (SVM) είναι ένα σύνολο εποπτευόμενων μεθόδων εκμάθησης που χρησιμοποιούνται για ταξινόμηση, παλινδρόμηση και ανίχνευση ακραίων τιμών. Στην πράξη, οι SVM εκπαιδεύονται στον προσδιορισμό της επιφάνειας – συνόρου μεταξύ των κλάσεων που καλείται να διαχωρίσει, και μάλιστα εκείνης που απέχει όσο το δυνατόν περισσότερο.

Τα πλεονεκτήματα των SVM είναι (Pradhan, 2012):

- Αποτελεσματικό σε χώρους υψηλών διαστάσεων.
- Εξακολουθεί να είναι αποτελεσματικό σε περιπτώσεις όπου ο αριθμός των διαστάσεων είναι μεγαλύτερος από τον αριθμό των δειγμάτων.
- Χρησιμοποιεί ένα υποσύνολο σημείων εκπαίδευσης στη συνάρτηση απόφασης (που ονομάζονται διανύσματα υποστήριξης), επομένως είναι επίσης αποδοτικό στη μνήμη.
- Ευέλικτο: μπορούν να καθοριστούν διαφορετικές συναρτήσεις πυρήνα για τη συνάρτηση απόφασης. Παρέχονται κοινοί πυρήνες, αλλά είναι επίσης δυνατός ο καθορισμός προσαρμοσμένων πυρήνων.

Τα μειονεκτήματα των μηχανών διανυσμάτων υποστήριξης περιλαμβάνουν:

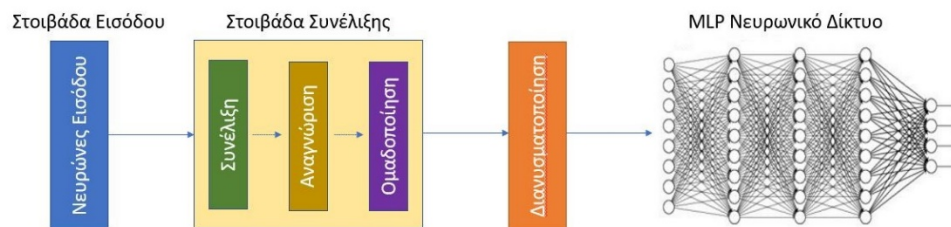
- Εάν ο αριθμός των χαρακτηριστικών είναι πολύ μεγαλύτερος από τον αριθμό των δειγμάτων, ο χρήστης πρέπει να αποφύγει την υπερβολική προσαρμογή στην επιλογή των συναρτήσεων πυρήνα και ο όρος τακτοποίησης είναι ζωτικής σημασίας.
- Τα SVM δεν παρέχουν άμεσα εκτιμήσεις πιθανοτήτων, αυτές υπολογίζονται χρησιμοποιώντας μια ακριβή πενταπλάσια διασταυρούμενη επικύρωση.

2.3 Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Network – CNN)

Τα συνελικτικά νευρωνικά δίκτυα (CNN) είναι σχεδιασμένα, ώστε να εκμεταλλεύονται τον χωρικό (η χρονικό) συσχετισμό των δεδομένων εισόδου, εφόσον αυτός υπάρχει. Για το λόγο αυτό είναι μια κατηγορία τεχνητών νευρωνικών δικτύων που έχουν γίνει κυρίαρχα σε διάφορες εργασίες υπολογιστικής όρασης και προσελκύουν το ενδιαφέρον σε διάφορους τομείς, συμπεριλαμβανομένης της ακτινολογίας (Goodfellow, 2016).

Η τροφοδοσία των δεδομένων γίνεται με την τεχνική του κυλιόμενου παράθυρου, γίνεται δηλαδή μια σταδιακή σάρωση των δεδομένων με βάση μια συνάρτηση βήματος, η οποία ουσιαστικά ορίζει το μέγεθος του παραθύρου και την «ταχύτητα» κίνησής του. Από το επίπεδο εισόδου, τα δεδομένα προωθούνται (αφού τους δοθούν τα απαραίτητα βάρη) σε ένα επίπεδο συνέλιξης, όπου γίνεται η αναγνώριση των «σημαντικών» χαρακτηριστικών του δείγματος και η ομαδοποίηση των δεδομένων με βάση τα χαρακτηριστικά αυτά. Το αποτέλεσμα της διεργασίας αυτής είναι η μείωση των εξεταζόμενων διαστάσεων των δεδομένων εισόδου (Goodfellow, 2016).

Στη συνέχεια, τα αποτελέσματα του επιπέδου συνέλιξης τροφοδοτούν συνήθως ένα MLP, ενώ συχνά περνούν έπειτα και από άλλο ένα επίπεδο συνέλιξης μέχρι να παράξουν την τελική έξοδο. Με τη μέθοδο αυτή, το σύστημα γίνεται πιο ευσταθές και ανθεκτικό σε μικρές παραλλαγές της εισόδου, και συνεπώς πιο αξιόπιστο, ενώ παράλληλα μειώνονται οι απαιτήσεις σε μνήμη και υπολογιστική ισχύ (Goodfellow, 2016). Παρατίθεται μια σχηματική αναπαράσταση ενός απλού CNN:

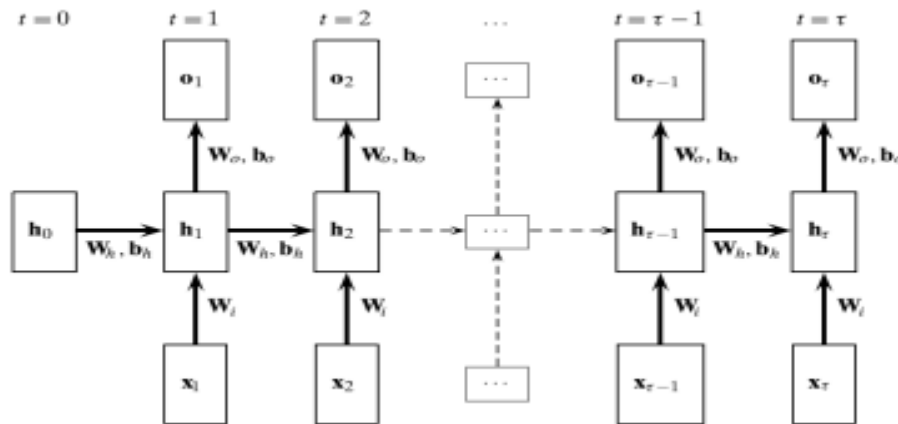


Εικόνα 2: Μοντέλο CNN

2.4 Ανατροφοδοτούμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNN)

Τα ανατροφοδοτούμενα νευρωνικά δίκτυα (RNN), βάσει σχεδιασμού, είναι κατάλληλα για επεξεργασία δεδομένων συσχετιζόμενων σε μία διάσταση (π.χ. χρονοσειρές), σε αντίθεση με τα CNN παραπάνω που αφορούν δύο διαστάσεις.

Η βασική δομή ενός RNN αποτελείται από το επίπεδο εισόδου, την κρυφή κατάσταση (hidden state) και τη στοιβάδα εξόδου. Στο επίπεδο εισόδου γίνεται η εισαγωγή των δεδομένων της χρονοσειράς και, αφού τους ανατεθεί ένα βάρος, προωθούνται στο επόμενο επίπεδο. Στην κρυφή κατάσταση υπάρχει η συνάρτηση ενεργοποίησης του νευρώνα, της οποίας ωστόσο οι παράμετροι τροποποιούνται με βάση την κρυφή κατάσταση του προηγούμενου (ή του επόμενου) στιγμιότυπου δεδομένων. Έτσι, το αποτέλεσμα στη στοιβάδα εξόδου είναι στην ουσία επηρεασμένα από την «ανάμνηση» των δεδομένων που προηγούνται (ή έπονται) (Zhang et al., 2018). Παρακάτω ακολουθεί σχηματικό μοντέλο:



Εικόνα 2: Μοντέλο RNN, όπου x_i : στοιβάδα εισόδου h_i : κρυφή κατάσταση o_i : στοιβάδα εξόδου

Τα βαθιά νευρωνικά δίκτυα, όπως το παραπάνω, αντιμετωπίζουν προβλήματα στη διαδικασία εκπαίδευσής τους, που αφορούν την κλίση της συνάρτησης απώλειας που εφαρμόζεται στον καθορισμό της συνάρτησης ενεργοποίησης. Όταν η κλίση παίρνει πολύ μεγάλες ή πολύ μικρές τιμές κοντά στη στοιβάδες εξόδου, από τις πρώτες εποχές της εκπαίδευσης, τα αποτελέσματα του συστήματος δεν συγκλίνουν, αφού η επιρροή από τη συνάρτηση απώλειας στις πρώτες στοιβάδες του μοντέλου είναι αναξιόπιστη (Geron, 2019).

2.5 Δίκτυα Μακράς Βραχυπρόθεσμης Μνήμης (Long Short – term Memory – LSTM)

Για την επίλυση του προβλήματος αυτού έχουν κατασκευαστεί πολυάριθμες εκδοχές των RNN, που διαφοροποιούνται ως προς τη δομή (Deep RNN), την κατεύθυνση (Bidirectional RNN) και το περιεχόμενο της κρυφής κατάστασης (Goodfellow, 2016).

Η αρχιτεκτονική των δικτύων μακράς βραχυπρόθεσμης μνήμης (LSTM) προσπαθεί να λύσει το παραπάνω πρόβλημα εισάγοντας περισσότερες λειτουργίες στο κρυφό επίπεδο του RNN, που εμποδίζουν την εμφάνιση εξαφανιζόμενων ή εκρηγνυόμενων κλίσεων. Πιο συγκεκριμένα, στο κρυφό επίπεδο εισάγεται μια πύλη λήθης (forget gate), στην οποία ουσιαστικά διοχετεύονται τιμές δεδομένων που αποσταθεροποιούν το μοντέλο. Έτσι στην έξοδο του LSTM, έχουμε δεδομένα που δεν μπορούν να επηρεάσουν με αρνητικό τρόπο την ευστάθεια του συστήματος, αλλά ταυτόχρονα δεν επηρεάζουν την ποιότητα της τελικής πρόβλεψης.

3. Οι μελέτες

3.1 M.A.H. Dempster, V. Leemans (2006)

Το βασικό μοντέλο

Με σκοπό ένα πλήρως αυτοματοποιημένο σύστημα συναλλαγών, αυτή η εργασία παρουσιάζει την προσαρμοστική ενισχυμένη μάθηση (ARL). Ένας αλγόριθμος μηχανικής μάθησης, ένα επίπεδο διαχείρισης κινδύνων και ένα επίπεδο δυναμικής βελτιστοποίησης βοηθητικών προγραμμάτων συνθέτουν την πολυεπίπεδη δομή του συστήματος, η οποία χρησιμοποιείται για τις συναλλαγές σε αγορές συναλλάγματος. Η επαναλαμβανόμενη ενισχυμένη μάθηση (RRL), μια καθιερωμένη προσέγγιση μηχανικής μάθησης, επιλέχθηκε ως ο βασικός αλγόριθμος για το ARL.

Ο αλγόριθμος RRL (Moody & Saffell, 1999) βρίσκεται στην καρδιά του συστήματος συναλλαγών που περιγράφεται στην έρευνα. Απαιτήθηκαν, ωστόσο, ορισμένες τροποποιήσεις για να βελτιωθεί η απόδοσή του και να καταστεί κατάλληλο για μια στρατηγική διαχείρισης κινδύνου που βασίζεται σε ανάλυση διαρθρωτικού κινδύνου.

Όπως αναφέρεται στους Zhang & Mariner (2015), οι ερευνητές μηχανικής μάθησης έχουν δείξει ότι η επαναλαμβανόμενη ενισχυμένη μάθηση (RRL) είναι μια αποτελεσματική μέθοδος για την ανάπτυξη αλγορίθμων χρηματοοικονομικών συναλλαγών.

Η αλγοριθμική διαπραγμάτευση εξακολουθεί να μελετάται. Το RRL, μια διαδικτυακή προσέγγιση εκμάθησης για την προσέγγιση λύσεων σε προβλήματα στοχαστικού δυναμικού προγραμματισμού, χρησιμοποιείται από ερευνητές σε αυτό το πλαίσιο για τη βελτιστοποίηση των συστημάτων νομισματικών συναλλαγών με σκοπό τη μεγιστοποίηση της χρησιμότητας στις συναλλαγές FX υψηλής συχνότητας. Οι ερευνητές έχουν χρησιμοποιήσει την προσέγγιση RRL για να εμπορευούνται άλλα χρηματοοικονομικά περιουσιακά στοιχεία, όπως μετοχές και ομόλογα. Στη βιβλιογραφία, για παράδειγμα, μπορούν να εντοπιστούν εφαρμογές RRL για καθημερινές συναλλαγές χρηματιστηριακών δεικτών.

Οι Moody et al. (1998) χρησιμοποίησαν επαναλαμβανόμενη ενισχυμένη μάθηση για να τελειοποιήσουν τους αλγόριθμους χρηματοοικονομικών συναλλαγών, προκειμένου να μεγιστοποιήσουν τη χρησιμότητα. Η μέθοδος ανόδου στοχαστικής κλίσης RRL (Moody et al., 1998) βελτιστοποιεί συνεχώς ένα μέτρο χρησιμότητας χρησιμοποιώντας πληροφορίες αγοράς. Παρά το γεγονός ότι τα συστήματα συναλλαγών RRL μπορούν εύκολα να περιλαμβάνουν τεχνικούς δείκτες και θεμελιώδη χρηματοοικονομικά στοιχεία στις εισροές του συστήματος, οι καθυστερημένες αποδόσεις χρησιμοποιούνται συχνά στις περισσότερες συζητήσεις για συστήματα συναλλαγών τύπου RRL.

Οι γραμμικοί συνδυασμοί αποδόσεων χρησιμοποιούνται στο σύστημα διαπραγμάτευσης RRL (Moody et al., 1998), το οποίο προορίζεται για τη διαπραγμάτευση ενός περιουσιακού στοιχείου με μια ενέργεια δύο θέσεων (long/short).

Ο στόχος του συστήματος συναλλαγών είναι να μεγιστοποιήσει τη συνάρτηση χρησιμότητας U_t αλλάζοντας συνεχώς το σύνολο παραμέτρων t , όπου το U_t συμβολίζεται ως συνάρτηση που βασίζεται στην πραγματοποιηθείσα απόδοση R_t τη στιγμή t .

Δεδομένα εισόδου

Το σύστημα συναλλαγών που συζητείται σε αυτή την εργασία αποτελείται από τρία επίπεδα: το βασικό σύστημα συναλλαγών, το επίπεδο διαχείρισης κινδύνου και απόδοσης και το επίπεδο βελτιστοποίησης παραμέτρων. Το μοντέλο συναλλαγών λαμβάνει ως εισροές τυποποιημένες αλλαγές προηγούμενων χρονοσειρών και εξάγει την προτιμώμενη θέση (long ή short) σε ένα από τα δύο εμπλεκόμενα νομίσματα. Περαιτέρω, το μοντέλο προσδιορίζεται ως ένα επαναλαμβανόμενο νευρωνικό δίκτυο μονού επιπέδου της τιμής ενός νομίσματος σε σχέση με ένα άλλο, δηλ.

$$F_t = \text{sign} \left(\sum_{i=0}^M w_{i,t} r_{t-i} + w_{M+1,t} F_{t-1} + v_t \right)$$

όπου F_t είναι η θέση που πρέπει να ληφθεί τη χρονική στιγμή t , ενώ w_i και v_t είναι, αντίστοιχα, το διάνυσμα βάρους και το κατώφλι του νευρώνα τη χρονική στιγμή t και r_t είναι οι αποδόσεις ως προς μεταβολές τιμών της χρονοσειράς FX .

Δεν καθορίστηκαν περιθώρια προσφοράς-ζήτησης κατά τη φάση εκπαίδευσης, αλλά ο παράγοντας κόστους συναλλαγής παρέμεινε μια παράμετρος συντονισμού. Το αυξημένο κόστος χρειάζεται μεγαλύτερη προσδοκία ακατέργαστου κέρδους πριν από τη σύναψη μιας συμφωνίας. Κατά συνέπεια, η παράμετρος d θα έχει αντίκτυπο στην απόδοση και το προφίλ κινδύνου της στρατηγικής συναλλαγών.

Ο αλγόριθμος RRL μπορεί να γίνει ασταθής, εάν συμβούν μεγάλες διακυμάνσεις στις αποδόσεις συναλλάγματος, όπως όταν παρεμβαίνουν οι κεντρικές τράπεζες. Όλα τα βάρη επανακλιμακώθηκαν κατά ένα συντελεστή $f1$ όταν έφτασαν σε μια συγκεκριμένη τιμή κατωφλίου για να αποτραπούν από το να λάβουν ασυνήθιστα τεράστιες τιμές. Για να μην κινούνται μεμονωμένα προς τα πάνω, τα βάρη είναι επίσης δεμένα μεταξύ τους. Μπορεί να προκύψουν ανακριβείς εκτιμήσεις μοντέλων, λόγω της αριθμητικής αστάθειας που σχετίζεται με τα υπερβολικά βάρη. Αυτό το φαινόμενο είναι ανεπιθύμητο.

Ένα καλύτερο σύστημα ενημέρωσης θέσης βελτίωσε επίσης την απόδοση του μοντέλου συναλλαγών RRL. Η εναλλαγή μεταξύ διαφορετικών σημείων από το ένα tick στο άλλο είναι σύνηθες φαινόμενο με τον παλιό μηχανισμό ενημέρωσης. Μπορεί να προκύψουν μεγάλα κόστη συναλλαγής, γι' αυτό είναι εξαιρετικά ανεπιθύμητο.

Με βάση ιστορικά δεδομένα, το σύστημα συναλλαγών ARL χρησιμοποιήθηκε για τις συναλλαγές στις αγορές ξένου νομίσματος (FX). Για το σκοπό αυτό συλλέχθηκαν και αξιολογήθηκαν οι επιδόσεις εκτός δείγματος του συστήματος σε ιστορικά δεδομένα FX . Η δυναμική βελτιστοποίηση χρησιμότητας δοκιμάστηκε σε μερικές πρόσθετες δοκιμές.

Αποτελέσματα

Διεξήχθη ένας αριθμός μελετών που συγκρίνουν την απόδοση και τη χρησιμότητα της απόδοσης με και χωρίς δυναμική βελτιστοποίηση για να καταδείξουν τα πλεονεκτήματα της δομικής μεθόδου και της βελτιστοποίησης χρησιμότητας ανώτατου επιπέδου. Λόγω της αυξημένης πιθανότητας και σοβαρότητας του κινδύνου σε αυτές τις δοκιμές, η χρησιμότητα μειώνεται καθώς αυξάνεται η αποστροφή κινδύνου. Αυτό μειώνει την ικανότητα του συστήματος να επιλέξει μια στρατηγική που μεγιστοποιεί τη χρησιμότητα. Λόγω των εξαιρετικά χαμηλών παραμέτρων προσαρμογής κινδύνου συνάρτησης χρησιμότητας από μηδέν έως δύο, δεν μπορεί να επιτευχθεί μια λιγότερο επικίνδυνη προσέγγιση. Το συνολικό κέρδος και ο αριθμός των συμφωνιών αυξάνονται όταν η αποστροφή κινδύνου αυξάνεται στο 0,4, αλλά το μέσο κέρδος ανά συναλλαγή μειώνεται. Η τάση του συστήματος να επιλέγει λιγότερο προσοδοφόρες συμφωνίες είναι πιθανότατα αποτέλεσμα της επιθυμίας του να ελαχιστοποιήσει την έκθεσή του στον κίνδυνο. Σε αυτό το στάδιο, υιοθετείται μια νέα προσέγγιση που αυξάνει τον αριθμό των συναλλαγών προκειμένου να διατηρηθεί το σωρευτικό κέρδος σε υψηλά επίπεδα ακόμη και όταν το μέσο κέρδος και η χρησιμότητα μειώνονται.

Η αντιστάθμιση του συστήματος μεταξύ κινδύνου στρατηγικής και απόδοσης ελέγχεται από την παράμετρο αποφυγής κινδύνου. Το ARL καταργεί την απαίτηση για βαθμονόμηση ή προσαρμογή ανθρώπινου μοντέλου ρυθμίζοντας αυτόματα τις υπερ-παραμέτρους, μειώνοντας έτσι σημαντικά τον κίνδυνο κατασκοπείας δεδομένων. Χρησιμοποιώντας το ιστορικό δεδομένων υψηλής συχνότητας FX , το σύστημα απέδωσε καλά.

Ευρετικές πληροφορίες εκτός δείγματος στις προσομοιώσεις συναλλαγών: αποφυγή κινδύνου n , σωρευτικό κέρδος WN , κίνδυνος S , συνολικός αριθμός συναλλαγών N , μέσο κέρδος ανά συναλλαγή R , ποσοστό συναλλαγών προς τη σωστή κατεύθυνση (μετά την αφαίρεση του κόστους), βοηθητικό πρόγραμμα με δυναμική βελτιστοποίηση επιπέδου 3 και το αντίστοιχο βοηθητικό πρόγραμμα χωρίς βελτιστοποίηση επιπέδου 3 ενεργοποιημένο.

r	W_N (pips)	Σ	N	R (pips)	Direction correct (%)	U with layer 3	U without layer 3
0	4779	0.5334	2854	1.67	62	2.07	1.92
0.1	4717	0.5369	2852	1.65	62	1.79	1.67
0.2	4717	0.5369	2852	1.65	62	1.53	1.43
0.3	4663	0.5596	3004	1.55	62	1.25	1.18
0.4	5144	0.5225	3353	1.53	62	1.13	0.93
0.5	5104	0.4880	3337	1.53	62	0.86	0.69
0.6	5083	0.4962	3220	1.58	62	0.58	0.44
0.7	4780	0.4254	3129	1.53	61	0.32	0.19
0.8	4635	0.3901	2657	1.74	62	0.09	-0.05
0.9	4669	0.3761	2639	1.77	63	-0.14	-0.30

Εικόνα 3: Αποτελέσματα

3.2 Kercheval, & Zhang (2013)

Βασικό μοντέλο

Σε αυτή την έρευνα οι ερευνητές χρησιμοποίησαν μεθόδους «πολλαπλών κλάσεων SVM» με έναν νέο τρόπο για να καταγράψουν τη δυναμική στα δεδομένα του βιβλίου παραγγελιών ορίων υψηλής συχνότητας και τις προβλέψεις κινήσεων της μέσης τιμής και της κατεύθυνσης των διασταυρώσεων προσφοράς-ζήτησης σε μικρά χρονικά διαστήματα. Πειράματα με πραγματικά δεδομένα από το NASDAQ δείχνουν ότι τα μοντέλα SVM πολλαπλών κατηγοριών που έχουν δημιουργηθεί στην έρευνα αυτή όχι μόνο προβλέπουν διάφορες μετρήσεις με υψηλή ακρίβεια, αλλά παρέχουν επίσης προβλέψεις αποτελεσματικά. Στην εποπτευόμενη μηχανική εκμάθηση, τα δεδομένα εκπαίδευσης επισημαίνονται εκ των προτέρων ανά τάξη και χρησιμοποιείται ένα μοντέλο για την εκχώρηση νέων δεδομένων σε αυτές τις κλάσεις. Αυτό περιλαμβάνει τις μεθόδους λογιστικής παλινδρόμησης και SVM.

Δεδομένα εισόδου

Για να δημιουργηθεί ένα μοντέλο εκμάθησης για μια δεδομένη μέτρηση, όπως η κίνηση στη μέση τιμή, θα πρέπει να προετοιμαστεί ένα σύνολο δειγμάτων με ετικέτα, που ονομάζονται δεδομένα εκπαίδευσης, στο οποίο κάθε σημείο δεδομένων χαρακτηρίζεται με ένα διάνυσμα χαρακτηριστικών, κατασκευάζεται ένα μοντέλο SVM και στη συνέχεια υπόκειται σε διαδικασία επικύρωσης για την επαλήθευση της ευρωστίας και της αξιοπιστίας. Για να συμβαδίζει με την ταχέως μεταβαλλόμενη δυναμική του limit order book, τα δεδομένα εκπαίδευσης ενημερώνονται συχνά.

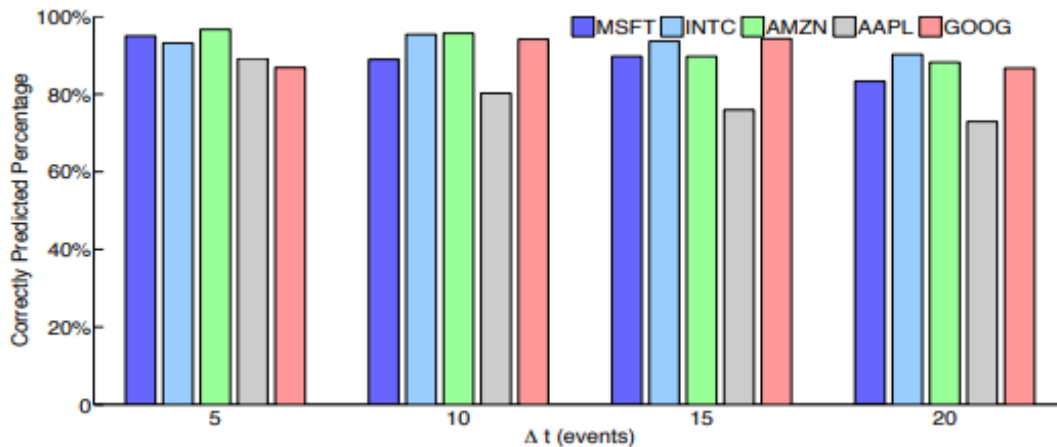
Για τη δημιουργία ενός μοντέλου μηχανικής μάθησης για μια συγκεκριμένη μέτρηση, χρησιμοποιείται στην έρευνα η ακόλουθη διαδικασία τεσσάρων φάσεων.

- Αναπαράσταση δυνατοτήτων: τα δεδομένα στο βιβλίο παραγγελιών και στο βιβλίο μηνυμάτων μετατρέπονται σε μορφή κατάλληλη για χειρισμό μεθόδων μηχανικής μάθησης.
- Κατασκευή μαθησιακού μοντέλου: κατασκευάζεται ένα μοντέλο SVM.
- Επικύρωση μαθησιακού μοντέλου: το μοντέλο αξιολογείται και επικυρώνεται χρησιμοποιώντας συγκεκριμένες μετρήσεις απόδοσης.
- Ταξινόμηση μη ορατών δεδομένων: το κατασκευασμένο μοντέλο μάθησης αυτοματοποιεί την πρόβλεψη της επιλεγμένης μέτρησης σε πραγματικό χρόνο.

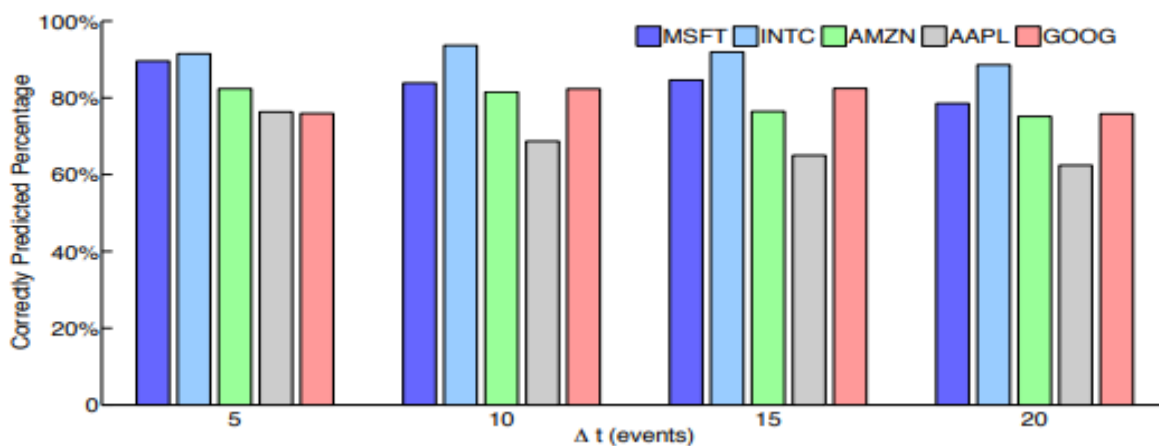
Πολλά τυχαία σύνολα 1500 σημείων δεδομένων εκπαίδευσης κατασκευάζονται με την αναπαραγωγή της κατανομής κάθε μέτρησης στα αρχικά δεδομένα. Οι αριθμοί που απεικονίζονται είναι ο μέσος χρόνος που χρησιμοποιούν τα μοντέλα για την αναγνώριση κάθε αόρατου παραδείγματος.

Αποτελέσματα

Η αποστολή του μοντέλου είναι να προβλέψει σε ποια από τις τρεις πιθανές τάξεις θα βρίσκεται η κατάσταση του βιβλίου παραγγελιών LOB τη φορά Δt στο μέλλον. Οι δυνατότητες είναι: η μέση τιμή θα είναι υψηλότερη, χαμηλότερη ή η ίδια (μετρική μέσης τιμής), ή θα υπάρχει μια ανοδική διασαύρωση διαφοράς, μια διασαύρωση προς τα κάτω ή χωρίς διασαύρωση (μέτρηση διασαύρωσης διασποράς).



Εικόνα 4: Ποσοστό σωστής πρόβλεψης για διαφορετικούς ορίζοντες πρόβλεψης. Δεδομένα εκπαίδευσης μέγεθος συνόλου = 1200, μετρικό = κίνηση στη μέση τιμή. $\Delta t = 5, 10, 15, 20$ (συμβάντα)



Εικόνα 5: Ποσοστό σωστής πρόβλεψης για διαφορετικούς ορίζοντες πρόβλεψης. Δεδομένα εκπαίδευσης μέγεθος σετ = 1200, μετρικό = διασαύρωση εξάπλωσης. $\Delta t = 5, 10, 15, 20$ (συμβάντα)

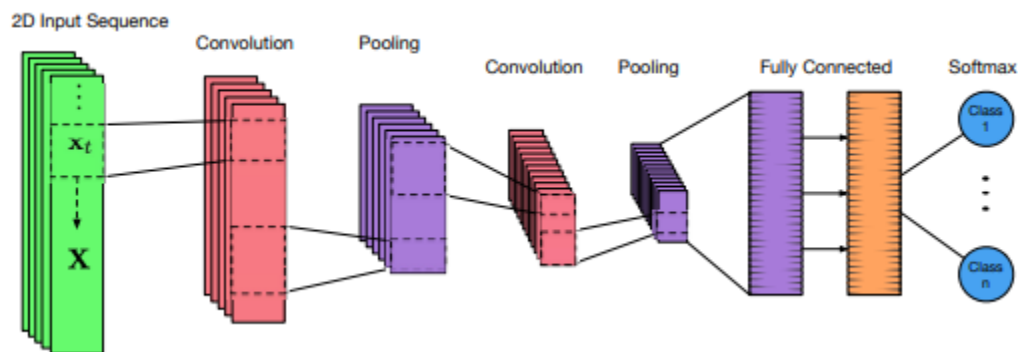
Για να διερευνηθεί το ερώτημα εάν τα αποτελέσματα οφείλονται απλώς στην τύχη, χρησιμοποιήθηκε έπειτα το ζευγαρωμένο t-test για να εξεταστεί η στατιστική σημασία των μετρήσεων της απόδοσής του μοντέλου των ερευνητών.

Ως τελική δοκιμή της αποτελεσματικότητας του μοντέλου SVM, το μοντέλο δοκιμάστηκε με μια απλή στρατηγική συναλλαγών έναντι των δεδομένων που συλλέχθηκαν. Τα πειράματά με πραγματικά δεδομένα δείχνουν ότι τα σύνολα χαρακτηριστικών που έχουν σχεδιαστεί στο προτεινόμενο πλαίσιο είναι αποτελεσματικά. Για παράδειγμα, τα μοντέλα που κατασκευάστηκαν για κίνηση στη μέση τιμή χρησιμοποιώντας μόνο τα βασικά χαρακτηριστικά μπορούν να επιτύχουν κατά μέσο όρο 79,6%, 82,6% και 80,3% για ακρίβεια, ανάκληση και μέτρηση F1, αντίστοιχα. Η απόδοση μπορεί να βελτιωθεί περαιτέρω κατά περίπου 6,3% και 3,2% όταν περιλαμβάνονται χαρακτηριστικά που δεν είναι ευαίσθητα στο χρόνο.

3.3 Tsantekidis et al (2017)

Βασικό μοντέλο

Σε αυτή την εργασία προτάθηκε μια μεθοδολογία βαθιάς μάθησης, βασισμένη στα Συνελκτικά Νευρωνικά Δίκτυα (CNN), που προβλέπει τις κινήσεις των τιμών των μετοχών, χρησιμοποιώντας ως είσοδο μεγάλης κλίμακας χρονοσειρές υψηλής συχνότητας που προέρχονται από το βιβλίο παραγγελιών των χρηματοοικονομικών ανταλλαγών.



Εικόνα 6: Βασικό CNN

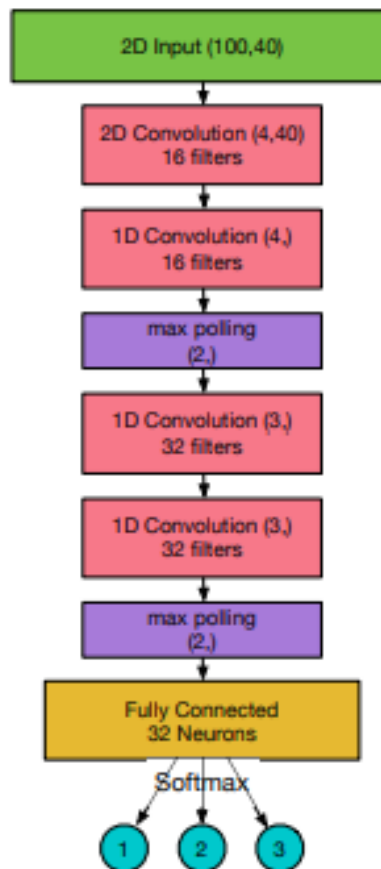
Δεδομένα εισόδου

Το σύνολο δεδομένων που χρησιμοποιήθηκε περιέχει περισσότερα από 4 εκατομμύρια συμβάντα οριακής παραγγελίας και τη σύγκριση με άλλες μεθόδους, όπως τα πολυεπίπεδα νευρωνικά δίκτυα και τις μηχανές διανυσμάτων υποστήριξης. Τα δεδομένα εισαγωγής αποτελούνται από 10 παραγγελίες για κάθε πλευρά του LOB (προσφορά και ζήτηση). Κάθε παραγγελία περιγράφεται από 2 μεγέθη, την τιμή και τον όγκο. Συνολικά χρησιμοποιήθηκαν 40 τιμές για κάθε χρονικό βήμα. Τα στοιχεία των μετοχών, που παρέχονται από τη Nasdaq Nordic, προέρχονται από τις φινλανδικές εταιρείες Kesko Oyj, Outokumpu Oyj, Sampo, Rautaruukki και Wartsila Oyj.

Η χρονική περίοδος που χρησιμοποιήθηκε για τη συλλογή αυτών των δεδομένων κυμαίνεται από την 1η έως τις 14 Ιουνίου 2010 (συμπεριλαμβάνονται μόνο εργάσιμες ημέρες), ενώ τα δεδομένα παρέχονται από τις ροές δεδομένων Nasdaq Nordic. Το σύνολο δεδομένων αποτελείται από 10 ημέρες για 5 διαφορετικές μετοχές και ο συνολικός αριθμός μηνυμάτων είναι 4,5 εκατομμύρια με εξίσου πολλά διαφορετικά βάθη. Δεδομένου ότι το εύρος τιμών και όγκου είναι πολύ μεγαλύτερο από το εύρος των τιμών της συνάρτησης ενεργοποίησης του νευρωνικού δικτύου, τα δεδομένα έπρεπε να κανονικοποιηθούν πριν τροφοδοτηθούν στο δίκτυο.

Η αρχιτεκτονική του προτεινόμενου μοντέλου CNN αποτελείται από τα ακόλουθα επίπεδα:

- 1) 2D Convolution με 16 φίλτρα μεγέθους (4, 40)
- 2) 1D Convolution με 16 φίλτρα μεγέθους (4,) και max pooling με μέγεθος (2,)
- 3) 1D Convolution με 32 φίλτρα μεγέθους (3,)
- 4) 1D Convolution με 32 φίλτρα μεγέθους e (3,) και max pooling με μέγεθος (2,)
- 5) Πλήρως συνδεδεμένο στρώμα με 32 νευρώνες
- 6) Πλήρως συνδεδεμένο στρώμα με 3 νευρώνες.



Εικόνα 7: Εξεταζόμενο μοντέλο

Αποτελέσματα

Για να μετρήσουν οι ερευνητές την απόδοση του μοντέλου χρησιμοποίησαν το κάπα του Kohen, το οποίο χρησιμοποιείται για τη μέτρηση της συμφωνίας μεταξύ των συνόλων των δοσμένων απαντήσεων, λαμβάνοντας υπόψη την πιθανότητα να συμβούν τυχαίες συμφωνίες.

Model	Recall	Precision	F1	Cohen's κ
Prediction Horizon $k = 10$				
SVM	39.62%	44.92%	35.88%	0.068
MLP	47.81%	60.78%	48.27%	0.226
CNN	50.98%	65.54%	55.21%	0.35
Prediction Horizon $k = 20$				
SVM	45.08%	47.77%	43.20%	0.139
MLP	51.33%	65.20%	51.12%	0.255
CNN	54.79%	67.38%	59.17%	0.39
Prediction Horizon $k = 50$				
SVM	46.05 %	60.30%	49.42%	0.243
MLP	55.21%	67.14%	55.95%	0.324
CNN	55.58%	67.12%	59.44%	0.38

Εικόνα 8: Αποτελέσματα του CNN σε σχέση με άλλα μοντέλα που δοκιμάστηκαν

Η προτεινόμενη μέθοδος βρέθηκε να υπερέχει σημαντικά όλων των άλλων αξιολογημένων μοντέλων στις παρουσιαζόμενες μετρήσεις, δείχνοντας ότι το συνελκτικό νευρωνικό δίκτυο μπορεί να χειριστεί καλύτερα τη διαδοχική φύση των δεδομένων του LOB και να προσδιορίσει καλύτερα τη μικροδομή της αγοράς προκειμένου να ανιχνεύσει τυχόν αλλαγές στη μέση τιμή.

3.4 Selvin et al (2017)

Βασικό μοντέλο

Σε αυτή τη μελέτη αξιολογούνται τρεις αρχιτεκτονικές βαθιάς μάθησης (DL): τα επαναλαμβανόμενα δίκτυα μακροπρόθεσμης βραχυπρόθεσμης μνήμης (LSTM), τα συνελκτικά νευρωνικά δίκτυα (CNN) και τα ανατροφοδοτούμενα νευρωνικά δίκτυα (RNN). Στόχος είναι η πρόβλεψη της τιμής ορισμένων μετοχών σε χρονικό βάθος 10 λεπτών, με την αξιοποίηση εισερχόμενων δεδομένων ανά λεπτό. Χρησιμοποιήθηκε η μέθοδος του κυλιόμενου παραθύρου με αλληλοκαλυπτόμενα δεδομένα για την εκπαίδευση του κάθε μοντέλου, με το μέγεθος του παραθύρου στα 100 λεπτά και το μέγεθος αλληλοκάλυψης στα 90 λεπτά. Η επιλογή έγινε με βάση διαδοχικές δοκιμές για τις τιμές μεγεθών και σύγκριση των τελικών αποτελεσμάτων.

Δεδομένα εισόδου

Τα δεδομένα που χρησιμοποιήθηκαν αφορούν τιμές δυο μετοχών από τον τομέα του IT (Infosys, TCS) και μιας μετοχής από το φαρμακευτικό τομέα (CIPLA) του εθνικού Χρηματιστηρίου της Ινδίας. Λήφθηκαν δεδομένα ανά λεπτό για την χρονική περίοδο από την 1/7/2014 μέχρι τη 14/10/2014, που αφορούσαν σε χαρακτηριστικά όπως η ημέρα, η ώρα, η ταυτότητα της συναλλαγής, το ύψος της τιμής της μετοχής και όγκος μετοχών που μεταβιβάστηκαν.

Λόγω της απόκλισης μεταξύ των μεγεθών για κάθε μετοχή, τα δεδομένα υφίστανται κανονικοποίηση, ώστε να φτάσουν στο επίπεδο $[0, 1]$ και να μπορούν να τροφοδοτήσουν την είσοδο της κάθε αρχιτεκτονικής. Στην έξοδο ακολουθείται διαδικασία αποκανονικοποίησης, ώστε τα αποτελέσματα να είναι χρησιμοποιήσιμα.

Το κάθε μοντέλο εκπαιδεύτηκε για 1000 εποχές, ενώ δοκιμάστηκαν διαφορετικά πλήθη επιπέδων, μέχρι να επιλεγούν οι κατάλληλες παράμετροι, με κριτήριο τη μέγιστη δυνατή απόδοση. Για τον υπολογισμό της απόδοσης επιλέχθηκε η ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Square Error – RMSE). Έτσι, στην περίπτωση που το RMSE της τρέχουσας εποχής είναι μικρότερο από αυτό της προηγούμενης, ο πίνακας των βαρών αποθηκεύεται ως ο βέλτιστος μέχρι εκείνη τη στιγμή.

$$ep = \frac{abs [X_{real}^i - X_{predicted}^i]}{X_{real}^i} \times 100$$

Μετά το πέρας της εκπαίδευσης, αξιολογήθηκε η ακρίβεια του κάθε μοντέλου σε πραγματικά δεδομένα με τον υπολογισμό του ποσοστιαίου σφάλματος, όπως περιγράφεται στον παραπάνω τύπο, όπου:

X_{real}^i : η πραγματική τιμή της μετοχής τη χρονική στιγμή i

$X_{predicted}^i$: η προβλεπόμενη τιμή της μετοχής σύμφωνα με το μοντέλο τη χρονική στιγμή i

Αποτελέσματα

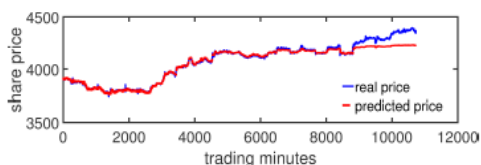
Έπειτα από την εκπαίδευση και την αξιολόγηση των τριών μοντέλων, τα αποτελέσματα αποτυπώνονται στον παρακάτω πίνακα:

COMPANY	RNN	LSTM	CNN
Infosys	3.90	4.18	2.36
TCS	7.65	7.82	8.96
Cipla	3.83	3.94	3.63

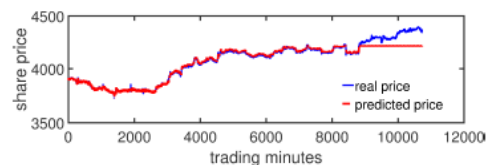
Εικόνα 9: Συγκριτικός πίνακας ποσοστιαίων σφαλμάτων των RNN, LSTM, CNN

Αναλυτικότερα παρατίθενται γραφήματα για τη σύγκλιση της κάθε μεθόδου για το κάθε χρηματιστηριακό προϊόν:

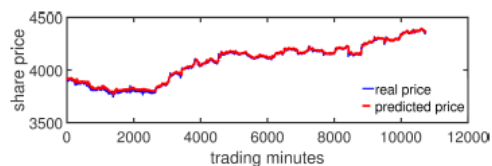
- Μετοχή INFOSYS:



Εικόνα 10: Μοντέλο RNN



Εικόνα 11: Μοντέλο LSTM

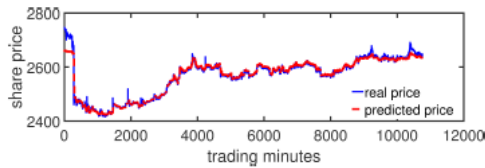


Εικόνα 12: Μοντέλο CNN

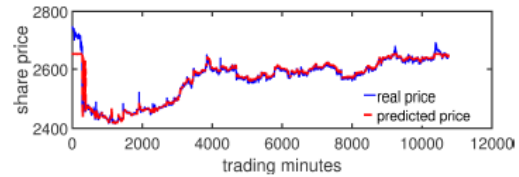
Μεταπτυχιακή Διατριβή

Γκούρας Νικόλαος

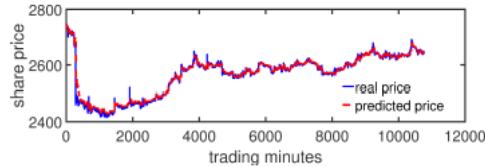
- Μετοχή TCS:



Εικόνα 13: Μοντέλο RNN

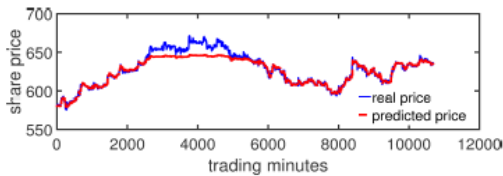


Εικόνα 14: Μοντέλο LSTM

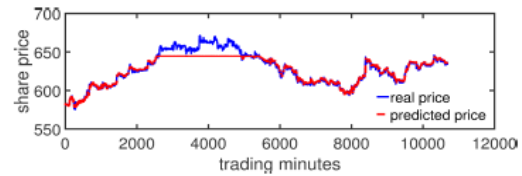


Εικόνα 15: Μοντέλο CNN

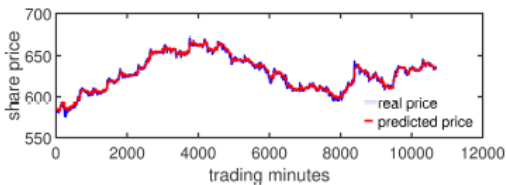
- Μετοχή CIPLA:



Εικόνα 16: Μοντέλο RNN



Εικόνα 17: Μοντέλο LSTM



Εικόνα 18: Μοντέλο CNN

Είναι προφανές από τον πίνακα και τα γραφήματα, ότι το συνελκτικό νευρωνικό δίκτυο αποδίδει αισθητά καλύτερα από τα άλλα δύο. Το CNN παράγει πρόβλεψη που εξαρτάται μόνο από το τρέχον χρονικό παράθυρο δεδομένων, αγνοώντας οποιαδήποτε πληροφορία από τα προηγούμενα παράθυρα. Αντίθετα, τόσο το επαναλαμβανόμενο νευρωνικό δίκτυο (RNN), όσο και το μοντέλο μακράς βραχυπρόθεσμης μνήμης (LSTM), τροποποιούνται με βάση την εμπειρία που αποκτούν ενδεχομένως από το πρώτο χρονικό παράθυρο μέχρι και το τελευταίο.

Η χρηματιστηριακή αγορά αποτελεί ένα έντονα δυναμικό περιβάλλον, που διέπεται από διαρκείς αλλαγές σε πολύ μικρά χρονικά διαστήματα. Έτσι, είναι αναμενόμενο η συνολική εικόνα του συστήματος να μεταβάλλεται δραστικά από ώρα σε ώρα και, ως εκ τούτου, τα παλαιότερα δεδομένα να καθίστανται άσχετα με την αναμενόμενη απόκριση του συστήματος στο βραχυπρόθεσμο μέλλον. Για το λόγο αυτό, εκτιμάται ότι η πρόβλεψη του CNN είναι μάλλον πιο αξιόπιστη από τις άλλες δύο μεθόδους που εφαρμόστηκαν.

Στα πλαίσια της μελέτης, αναπτύχθηκε και ένα γραμμικό στατιστικό μοντέλο ARIMA, ώστε να συγκριθούν τα αποτελέσματα των παραπάνω μοντέλα με μια πιο «παραδοσιακή» μέθοδο πρόβλεψης. Το μοντέλο ARIMA έχει χρησιμοποιηθεί ευρέως στην πρόβλεψη ενός μεγέθους χρησιμοποιώντας την κινητή μέση τιμή του μεγέθους αυτού κατά μήκος μιας δεδομένης χρονοσειράς (Li et al, 2015). Στον παρακάτω πίνακα αποτυπώνονται τα αντίστοιχα ποσοστά σφαλμάτων:

COMPANY	Error Percentage
Infosys	31.91
TCS	21.16
Cipla	36.53

Εικόνα 19: Πίνακας ποσοστιαίων σφαλμάτων μοντέλου ARIMA

Όπως φαίνεται από τα ποσοστά σφαλμάτων από τους δύο πίνακες, οι μη γραμμικές μέθοδοι μηχανικής μάθησης υπερτερούν σημαντικά της γραμμικής μεθόδου ARIMA στην πρόβλεψη της τιμής των παραπάνω μετοχών, ακόμα και οι λιγότερο επιτυχημένες.

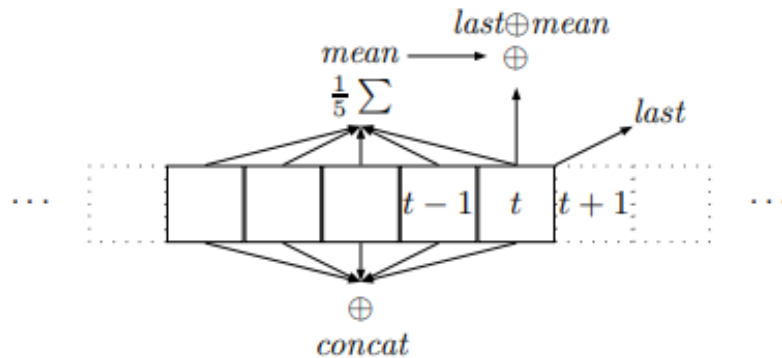
3.5 Nousi et al (2019)

Βασικό Μοντέλο

Σε αυτό το άρθρο, οι ερευνητές χρησιμοποιούν αλγόριθμους Μηχανικής Μάθησης (ML) για την πρόβλεψη μελλοντικών κινήσεων τιμών, χρησιμοποιώντας δεδομένα του limit order book. Συνδυάζονται και αξιολογούνται δύο διαφορετικά σύνολα χαρακτηριστικών: χαρακτηριστικά που βασίζονται στα ακατέργαστα δεδομένα του βιβλίου παραγγελιών και χαρακτηριστικά που εξάγονται από αλγόριθμους ML, με αποτέλεσμα διανύσματα χαρακτηριστικών με εξαιρετικά ποικίλες διαστάσεις.

Τα ακατέργαστα δεδομένα του βιβλίου παραγγελιών υποβάλλονται πρώτα σε προεπεξεργασία αφαιρώντας τα περιττά μηνύματα από την ανταλλαγή, π.χ. μηνύματα συμβάντων, και στη συνέχεια εξάγονται τα χαρακτηριστικά που προτείνονται. Αρχικά, εξάγεται ένα βασικό σύνολο χαρακτηριστικών, που περιλαμβάνει τις τιμές και τους όγκους για κάθε επίπεδο της πλευράς ζήτησης και προσφοράς του βιβλίου παραγγελιών. Αυτές οι πληροφορίες αποδίδουν 40 τιμές σε κάθε χρονικό βήμα. Στη συνέχεια, εξάγονται χαρακτηριστικά που δεν είναι ευαίσθητα στο χρόνο που περιγράφουν το spread, τη μέση τιμή, την τιμή και τις συσσωρευμένες διαφορές τιμής μεταξύ των παραγγελιών προσφοράς και ζήτησης κάθε επιπέδου βάθους και τα spread τιμής και όγκου. Τέλος, εξάγονται χρονικά ευαίσθητα χαρακτηριστικά που αντιστοιχούν στη μέση ένταση για συναλλαγές, εντολές, ακυρώσεις, διαγραφή, εκτέλεση ορατών οριακών εντολών και εκτέλεση κρυφών οριακών εντολών.

Αυτό το σύνολο χαρακτηριστικών περιλαμβάνει επίσης τις μέσες τιμές τιμής και όγκου σε κάθε επίπεδο του LOB, τη μέση ένταση ανά τύπο συναλλαγών καθώς και συγκρίσεις μεταξύ των εντάσεων και της οριακής επιτάχυνσης δραστηριότητας (παράγωγα μέσες εντάσεις). Λόγω της μη γραμμικής φύσης του χρόνου στα δεδομένα LOB, οι ερευνητές ακολούθησαν μια εισροή που βασίζεται σε συμβάντα.



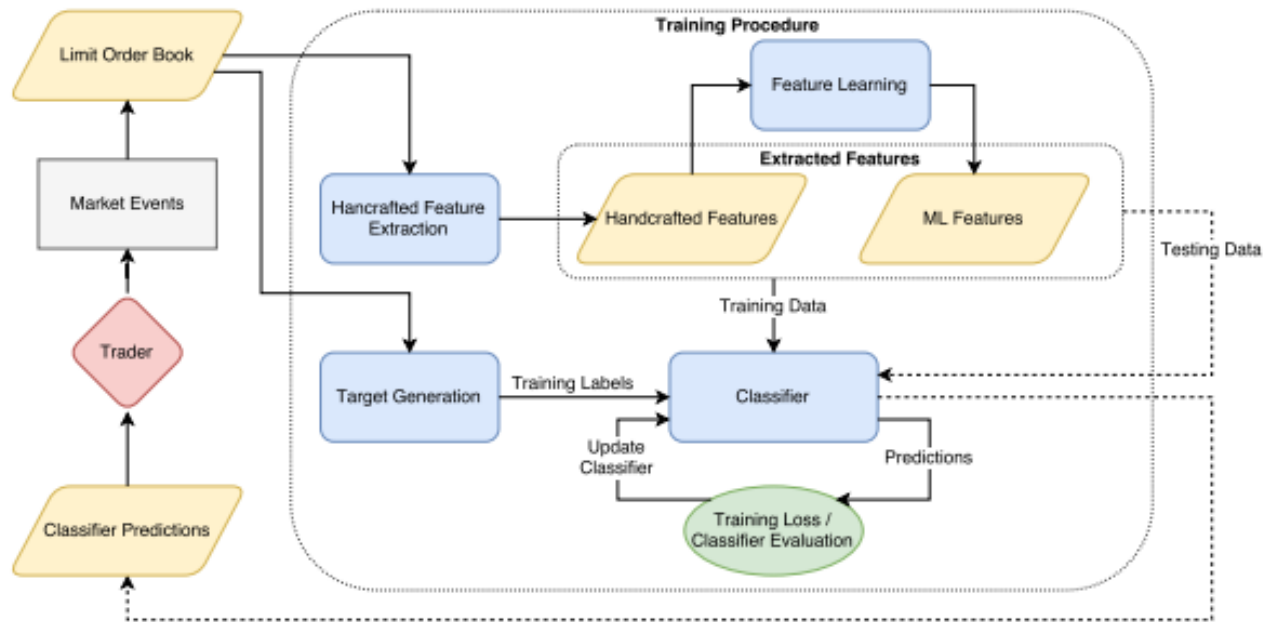
Εικόνα 20: Μοντελοποίηση των Nousi et al (2019)

Δεδομένα Εισόδου

Σε αυτό το άρθρο οι ερευνητές χρησιμοποίησαν διάφορους αλγόριθμους Machine Learning και τεχνικές προεπεξεργασίας δεδομένων για την πρόβλεψη μελλοντικών κινήσεων των τιμών των μετοχών. Το σύνολο δεδομένων που χρησιμοποιήθηκε προέρχεται από τη χρονική εξέλιξη του βιβλίου οριακών εντολών (LOB), που είναι η υψηλότερη δυνατή ανάλυση που μπορεί κανείς να παρατηρήσει τις χρηματιστηριακές αγορές.

Τα ακόλουθα τέσσερα διαφορετικά διανύσματα χαρακτηριστικών παράγονται και χρησιμοποιούνται ως εισδοί στα αξιολογούμενα μοντέλα, χρησιμοποιώντας ένα κυλιόμενο χρονικό παράθυρο:

1. Ένα μεμονωμένο διάνυσμα χαρακτηριστικών με 144 τιμές όπως περιγράφεται παραπάνω, που αντιστοιχεί στο τελευταίο δείγμα στο κυλιόμενο παράθυρο (στην εικόνα ως $last$).
2. Ο μέσος όρος των 5 δειγμάτων που βρίσκονται αυτή τη στιγμή στο παράθυρο, ο οποίος είναι επίσης ένα διάνυσμα 144 διαστάσεων.
3. Η συνένωση του τελευταίου δείγματος και του μέσου όρου και των 5 δειγμάτων στο παράθυρο, αποδίδοντας ένα διάνυσμα χαρακτηριστικών 288 διαστάσεων (στην εικόνα ως $last \oplus mean$ - το κυκλικό σύμβολο συν χρησιμοποιείται για να υποδηλώσει τη λειτουργία συνένωσης).
4. Η συνένωση και των 5 δειγμάτων, αποδίδοντας ένα διάνυσμα χαρακτηριστικών 720 διαστάσεων (στην εικόνα ως $concat$).



Εικόνα 21: Σύστημα πρόβλεψης μεσαίας τιμής μετοχών

Τα γεγονότα οριακής παραγγελίας που δημιουργούνται από γεγονότα της αγοράς χρησιμοποιούνται για την εξαγωγή διανυσμάτων χαρακτηριστικών που περιγράφουν τις μετοχές που υπάρχουν στο LOB και ένας ταξινομητής εκπαιδεύεται να αναγνωρίζει την κίνηση της μέσης τιμής αυτών των μετοχών. Οι διακεκομμένες γραμμές υποδεικνύουν τη φάση ανάπτυξης, όπου ένας έμπορος λαμβάνει προβλέψεις ταξινόμησης και ενδεχομένως ενεργεί βάσει αυτών επηρεάζοντας έτσι την αγορά.

Αποτελέσματα

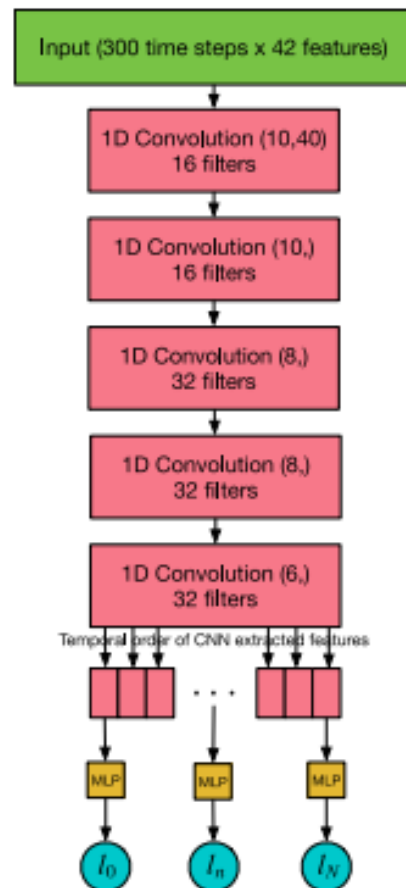
Τα αποτελέσματα που επιτεύχθηκαν υποδεικνύουν ότι το LOB περιέχει πολύτιμες πληροφορίες που, σε συνδυασμό με διάφορους αλγόριθμους Machine Learning, μπορούν να δώσουν ουσιαστική εικόνα για την τάση της χρηματιστηριακής αγοράς και να οδηγήσουν τα μοντέλα να κάνουν ακριβείς προβλέψεις, χωρίς εξωτερική ανθρώπινη παρέμβαση.

Ο ταξινομητής MLP αποδίδει καλύτερα όταν η αναπαράσταση εισόδου προκύπτει από τα προεπεξεργασμένα χαρακτηριστικά, καθώς τα κρυφά του στρώματα εξάγουν μια αναπαράσταση της εισόδου του βάσει ταξινόμησης. Η χαμηλή διάσταση των αναπαραστάσεων AE και BoF σε συνδυασμό με τη μη εποπτευόμενη φύση τους οδηγεί το MLP να κάνει λιγότερες σωστές προβλέψεις. Η χαμηλή διάσταση αυτών των μαθησιακών αναπαραστάσεων επιτρέπει ταχύτερους υπολογισμούς, αν και με κόστος την επίτευξη ελαφρώς υποβαθμισμένων επιδόσεων. Η απόδοση του ταξινομητή SVM αυξάνεται με τη διάσταση της αναπαράστασης εισόδου, αλλά υπολείπεται της απόδοσης του MLP, καθώς είναι λιγότερο ανθεκτικός στον εγγενή θόρυβο των δεδομένων. Τέλος, η απόδοση του SLFN μαστίζεται από τον υψηλό αριθμό ψευδώς αρνητικών προβλέψεων. Η ενσωμάτωση προηγούμενων πληροφοριών μετριάξει κάπως αυτό το μειονέκτημα, αλλά η συνολική απόδοσή της υποχωρεί σε σύγκριση με τους άλλους ταξινομητές.

3.6 Tsantekidis et al, (2020)

Βασικό μοντέλο

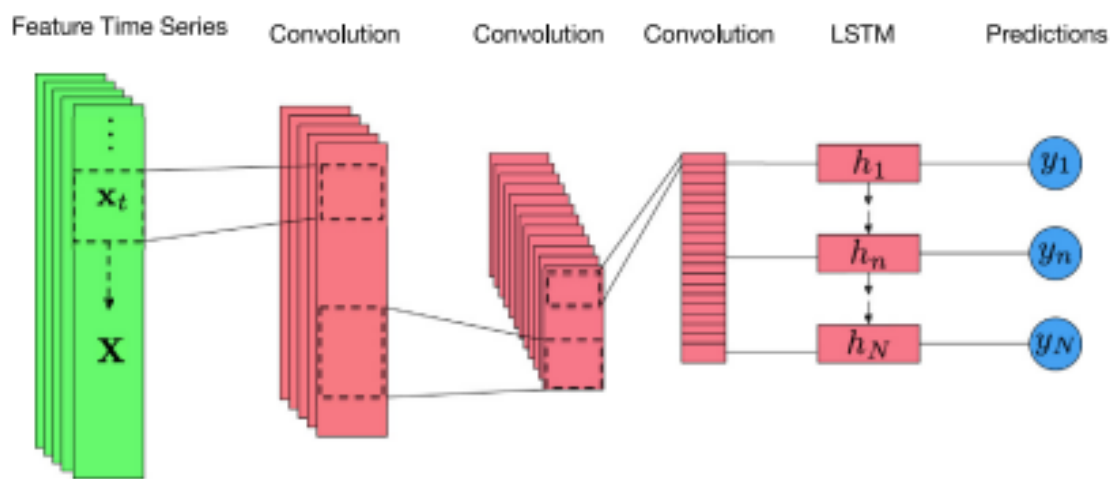
Σε αυτή την εργασία προτείνεται μια νέα μέθοδος για την κατασκευή σταθερών χαρακτηριστικών έτσι ώστε να επιτρέπεται η αποτελεσματική εφαρμογή των μοντέλων DL. Αυτά τα χαρακτηριστικά έχουν δοκιμαστεί διεξοδικά για την πρόβλεψη των κινήσεων στη μέση τιμή του Βιβλίου Παραγγελιών ορίων. Αξιολογούνται αρκετά μοντέλα DL, όπως τα επαναλαμβανόμενα δίκτυα μακροπρόθεσμης βραχυπρόθεσμης μνήμης (LSTM) και τα συνελκτικά νευρωνικά δίκτυα (CNN). Τέλος, προτείνεται και αξιολογείται ένα νέο μοντέλο που συνδυάζει την ικανότητα του CNN να εξάγει χρήσιμα χαρακτηριστικά και την ικανότητα των LSTM να αναλύουν χρονοσειρές.



Εικόνα 22: Μοντέλο CNN που αξιολογείται

Οι ερευνητές παρουσίασαν, επίσης, έναν ισχυρό συνδυασμό των δύο μοντέλων που περιεγράφηκαν προηγουμένως. Το μοντέλο CNN εφαρμόζεται χρησιμοποιώντας αιτιώδεις συνελίξεις και χρονική ομαδοποίηση για την παραγωγή ενός συνόλου χαρακτηριστικών για κάθε χρονικό βήμα. Στην ουσία το CNN λειτουργεί ως ο εξολκέας χαρακτηριστικών της χρονοσειράς βάθους LOB, η οποία παράγει μια νέα χρονοσειρά χαρακτηριστικών με το ίδιο μήκος με την αρχική, με καθένα από αυτά να έχει χρονικά βήματα που αντιστοιχούν το ένα στο άλλο. Στη συνέχεια εφαρμόζεται ένα στρώμα LSTM στις χρονοσειρές που παράγονται από το CNN και με τη σειρά του παράγει μια ετικέτα για κάθε χρονικό βήμα. Αυτό λειτουργεί με παρόμοιο τρόπο με το πλήρως συνδεδεμένο για προσωρινή ομαδοποίηση, αλλά αντί για το επίπεδο που

είναι πλήρως συνδεδεμένο, το LSTM επιτρέπει στο μοντέλο να ενσωματώνει τα χαρακτηριστικά από προηγούμενα βήματα. Η αρχιτεκτονική του μοντέλου απεικονίζεται παρακάτω:



Εικόνα 23: Μοντέλο CNN-LSTM που αξιολογείται

Δεδομένα Εισόδου

Χρησιμοποιήθηκε το κύριο οικονομικό σύνολο δεδομένων που αποτελείται από τα δεδομένα του βιβλίου παραγγελιών ορίων (LOB) για να παρουσιαστεί το προτεινόμενο χαρακτηριστικό σταθερής τιμής και την ικανότητα του μοντέλου μας να προβλέπει κινήσεις στη μέση τιμή. Συμπεριλήφθηκε επίσης ένα δεύτερο απλούστερο σύνολο δεδομένων τιμών FOREX για να προβλέψουν οι ερευνητές την κίνηση των τιμών και για να ενισχύσουν τους ισχυρισμούς της προτεινόμενης προσέγγισής τους.

Τα δεδομένα LOB που χρησιμοποιούνται σε αυτήν την εργασία παρέχονται από τη Nasdaq Nordic και αποτελούνται από εκδηλώσεις LOB διάρκειας 10 ημερών για 5 διαφορετικές μετοχές της Φινλανδικής εταιρείας, συγκεκριμένα Kesko Oyj, Outokumpu Oyj, Sampo, Rautaruukki και Wärtsilä Oyj. Η ακριβής χρονική περίοδος των συγκεντρωθέντων στοιχείων ξεκινά από την 1η Ιουνίου 2010 έως την 14η Ιουνίου 2010.

Τα δεδομένα αποτελούνται από διαδοχικά στιγμιότυπα της κατάστασης LOB μετά από κάθε συμβάν αλλαγής κατάστασης. Αυτό το συμβάν μπορεί να είναι μια εισαγωγή, εκτέλεση ή ακύρωση παραγγελίας και αφού αλληλοεπιδράσει με το LOB και αλλάξει την κατάστασή του, λαμβάνεται ένα στιγμιότυπο της νέας κατάστασης. Το βάθος LOB των δεδομένων που χρησιμοποιούνται είναι 10 για κάθε πλευρά του Βιβλίου Παραγγελιών, το οποίο καταλήγει να είναι 10 ενεργές παραγγελίες (που αποτελούνται από τιμή και όγκο) για κάθε πλευρά, προσθέτοντας συνολικά 40 τιμές για κάθε στιγμιότυπο LOB. Αν και το σύνολο δεδομένων που δοκιμάστηκε δεν εκτείνεται σε μεγάλο χρονικό διάστημα, η δειγματοληψία υψηλής συχνότητας που χρησιμοποιείται επιτρέπει τη συλλογή 4,5 εκατομμυρίων στιγμιότυπων που μπορούν να χρησιμοποιηθούν για την εκπαίδευση και την αξιολόγηση των προτεινόμενων μοντέλων.

Οι τιμές Open – High – Low - Close (OHLC) είναι μια τεχνική δειγματοληψίας, που επιτρέπει την απεικόνιση των κινήσεων των τιμών σε διάφορες χρονικές κλίμακες. Τα δεδομένα που παρέχονται από τις χρηματοοικονομικές ανταλλαγές ξεκινούν από τα δεδομένα LOB όπου είναι η υψηλότερη διαθέσιμη ανάλυση. Τα δεδομένα συμβάντος LOB μπορούν να υποδειγματιστούν για να συμπεριλάβουν μόνο τα συμβάντα στα οποία έχει εκτελεστεί μια παραγγελία και έχουν γίνει συναλλαγές περιουσιακών στοιχείων μεταξύ των συμμετεχόντων. Οι εμπορικές εκδηλώσεις αποτελούνται κυρίως από τρεις αξίες, την τιμή στην οποία έγινε η συναλλαγή, το ποσό του διαπραγματευόμενου περιουσιακού στοιχείου που ανταλλάχθηκε και τον ακριβή χρόνο που έλαβε χώρα.

Δεδομένου ενός χρονικού πλαισίου (π.χ. 1 ώρα), τα δεδομένα συναλλαγών δεσμεύονται στα αντίστοιχα χρονικά πλαίσια με βάση τον χρόνο εκτέλεσής τους και, στη συνέχεια, για κάθε κάδο υπολογίζεται η πρώτη (ανοιχτή), η υψηλότερη, η χαμηλότερη και η τελευταία (κλείσιμο) τιμή. Αυτό παρέχει μια χρήσιμη απεικόνιση της τιμής ενός περιουσιακού στοιχείου σε βάθος χρόνου. Το σύνολο δεδομένων που περιέχει 66 ζεύγη FOREX συμπεριλαμβάνει ευρέως χρησιμοποιούμενα νομίσματα όπως το ευρώ, το δολάριο ΗΠΑ και η λίρα Αγγλίας.

Αποτελέσματα

Χρησιμοποιώντας διαφορετικά μοντέλα ML, π.χ. SVM, MLP, CNN και LSTM, αποδείχθηκε πειραματικά ότι τα προτεινόμενα χαρακτηριστικά ξεπερνούν σημαντικά τα χαρακτηριστικά της πρωτογενούς τιμής. Τα προτεινόμενα σταθερά χαρακτηριστικά το επιτυγχάνουν αυτό καθιστώντας τη διαφορά μεταξύ των τιμών στο βάθος LOB ως κύρια μέτρηση αντί για την ίδια την τιμή, η οποία συνήθως κυμαίνεται πολύ περισσότερο με την πάροδο του χρόνου από το επίπεδο τιμής εντός του LOB. Ένα νέο συνδυασμένο μοντέλο CNN-LSTM προτάθηκε επίσης για προβλέψεις χρονοσειρών και αποδείχθηκε ότι παρουσιάζει πιο σταθερή συμπεριφορά και οδηγεί σε καλύτερα αποτελέσματα από τα μοντέλα CNN και LSTM

Forex data resulting performance. The prediction horizon k is used to calculate the $m_a(t)$ average filter based on the future k OHLC price samples.

Model	Mean recall	Mean precision	Mean F1	Cohen's κ
Prediction horizon $k = 5$				
CNN	0.441	0.444	0.417	0.182
LSTM	0.446	0.448	0.438	0.191
CNN-LSTM	0.455	0.454	0.454	0.202
Prediction horizon $k = 10$				
CNN	0.448	0.449	0.447	0.199
LSTM	0.448	0.449	0.445	0.202
CNN-LSTM	0.452	0.452	0.451	0.207
Prediction horizon $k = 15$				
CNN	0.444	0.443	0.443	0.198
LSTM	0.441	0.443	0.439	0.199
CNN-LSTM	0.447	0.446	0.446	0.203
Prediction horizon $k = 20$				
CNN	0.437	0.435	0.435	0.193
LSTM	0.444	0.445	0.444	0.204
CNN-LSTM	0.441	0.441	0.439	0.201

Εικόνα 24: Σύγκριση Μοντέλων

Είναι αρκετά σαφές από τον πίνακα ότι το LSTM έχει παρόμοια απόδοση με το προτεινόμενο μοντέλο CNN-LSTM, αν και το μοντέλο που προτάθηκε κατάφερε να το ξεπεράσει ελαφρώς, ενώ στο προηγούμενο σύνολο δεδομένων LOB το CNN-LSTM ξεπέρασε σταθερά τα μοντέλα με μεγαλύτερα περιθώρια σε σημαντικές μετρήσεις όπως το κ του Cohen. Αυτή η σχετική απόκλιση απόδοσης μεταξύ των συνόλων δεδομένων μπορεί να αποδοθεί στη μεγαλύτερη δυνατότητα εξαγωγής χαρακτηριστικών που υπάρχει στα δεδομένα LOB υψηλής ανάλυσης έναντι των υποδειγμάτων δεδομένων τιμών OHLC όπου έχουν αφαιρεθεί πολλές πληροφορίες.

4. Σύγκριση Αποτελεσμάτων

Από τους Dempster & Leemans (2006) εξετάστηκε ένα αυτοματοποιημένο σύστημα συναλλαγών που βασίζεται στην προσαρμοστική μάθηση ενίσχυσης. Το σύστημα απέδωσε πολύ καλά στο ιστορικό δεδομένων FX υψηλής συχνότητας που εξετάστηκε.

Επίσης επιτυχείς, οι Kercheval, & Zhang (2015) αναφέρουν ότι η προσέγγιση SVM πολλαπλών κατηγοριών έχει χρησιμοποιηθεί με επιτυχία στο παρελθόν για την πρόβλεψη τιμών συναλλάγματος υψηλής συχνότητας, αλλά η δική τους είναι η πρώτη απόπειρα να αποδειχθεί ότι αυτή η προσέγγιση είναι αποτελεσματική και για τα ίδια κεφάλαια και την πρόβλεψη βιβλίου παραγγελιών ορίων (LOB). Οι τεχνικές μοντελοποίησης SVM πολλαπλών κατηγοριών, που προτείνονται σε αυτή την έρευνα, μειώνουν το κόστος εκπαίδευσης σε σύγκριση με τα τρία SVM πολλαπλών πυρήνων σε άλλες μελέτες, αλλά εξακολουθούν να επιτυγχάνουν ισχυρή απόδοση όσον αφορά την ακρίβεια και την ανάκληση.

Οι Tsantekidis et al (2017) είναι η πρώτη εργασία που χρησιμοποιεί ένα σύνολο δεδομένων μεγάλης κλίμακας με περισσότερες από 4 εκατομμύρια οριακές εντολές για την εκπαίδευση των CNN για την πρόβλεψη της κίνησης των τιμών των μετοχών. Διαπίστωσαν ότι το CNN αποδίδει σημαντικά καλύτερα από άλλες τεχνικές, όπως τα Γραμμικά SVM και τα MLP, όταν προσπαθούν να προβλέψουν τις βραχυπρόθεσμες κινήσεις των τιμών.

Εντυπωσιακά μπορούν να χαρακτηριστούν τα αποτελέσματα της μελέτης των Selvin et al (2017), οι οποίοι δοκίμασαν τρία διαφορετικά μοντέλα βαθιάς μάθησης: RNN, LSTM και CNN. Αποδείχθηκε ότι η επιλογή ενός συνελκτικού νευρωνικού δικτύου είναι προτιμώμενη από τις άλλες δύο, λόγω του δυναμικού περιβάλλοντος της χρηματιστηριακής αγοράς, ενώ διαπιστώνεται η υπεροχή των μη γραμμικών μοντέλων έναντι των γραμμικών στην πρόβλεψη της τιμής των χρηματιστηριακών προϊόντων

Στους Nousi et al (2019) τα απαιτητικά χαρακτηριστικά των δεδομένων LOB, όπως η υψηλή ταχύτητα, η διακύμανση, ο όγκος και οι αυστηροί περιορισμοί σε πραγματικό χρόνο κατέστησαν δυσμενή την πρόβλεψη της μέσης κίνησης της τιμής. Ωστόσο, παρατηρήθηκαν σημαντικές διακυμάνσεις στην αξιοπιστία του μοντέλου, όταν ως δεδομένα εισόδου χρησιμοποιήθηκαν προεπεξεργασμένα δεδομένα από το LOB, αντί για χρονοσειρές που προέκυψαν από τεχνικές μηχανικής μάθησης.

Τέλος, στους Tsantekidis et al (2020) έγινε μια προσπάθεια βελτίωσης του προηγούμενου μοντέλου (2018) και παρατηρήθηκε ότι το LSTM έχει παρόμοια απόδοση με το προτεινόμενο μοντέλο CNN-LSTM, αν και τα μοντέλα τους καταφέρνουν να το ξεπεράσουν ελαφρώς.

Βιβλιογραφία

- Mitchell, T. M., & Mitchell, T. M. (1997). *Machine learning* (Vol. 1, No. 9). New York: McGraw-hill.
- Moody, J., & Saffell, M. (2001). Learning to trade via direct reinforcement. *IEEE transactions on neural Networks*, 12(4), 875-889.
- Dempster, M., & Leemans, V. (2006). An automated FX trading system using adaptive reinforcement learning. *Expert Systems with Applications*, 30(3), 543-552. <https://doi.org/10.1016/j.eswa.2005.10.012>
- Pradhan, A. (2012). Support vector machine-a survey. *International Journal of Emerging Technology and Advanced Engineering*, 2(8), 82-85.
- Li, P., Jing, C., Liang, T., Liu, M., Chen, Z., & Guo, L. (2015, October). Autoregressive moving average modeling in the financial sector. In *2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)* (pp. 68-71). IEEE.
- Zhang, J., & Maringer, D. (2015). Using a genetic algorithm to improve recurrent reinforcement learning for equity trading. *Computational Economics*, 47(4), 551-567. <https://doi.org/10.1007/s10614-015-9490-y>
- Kercheval, A. N., & Zhang, Y. (2015). Modelling high-frequency limit order book dynamics with support vector machines. *Quantitative Finance*, 15(8), 1315-1329. <https://doi.org/10.1080/14697688.2015.1032546>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Tsantekidis, A., Passalis, N., Tefas, A., Kannianen, J., Gabbouj, M., & Iosifidis, A. (2017). Forecasting stock prices from the limit order book using Convolutional neural networks. *2017 IEEE 19th Conference on Business Informatics (CBI)*. <https://doi.org/10.1109/cbi.2017.23>
- Aggarwal, C. C. (2018). *Neural networks and deep learning: A textbook*. Springer.
- Sirignano, J. A. (2018). Deep learning for limit order books. *Quantitative Finance*, 19(4), 549-570. <https://doi.org/10.1080/14697688.2018.1546053>
- Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc."
- Nousi, P., Tsantekidis, A., Passalis, N., Ntakaris, A., Kannianen, J., Tefas, A., Gabbouj, M., & Iosifidis, A. (2019). Machine learning for forecasting mid-price movements using limit order book data. *IEEE Access*, 7, 64722-64736. <https://doi.org/10.1109/access.2019.2916793>
- Tsantekidis, A., Passalis, N., Tefas, A., Kannianen, J., Gabbouj, M., & Iosifidis, A. (2020). Using deep learning for price prediction by exploiting stationary limit order book features. *Applied Soft Computing*, 93, 106401. <https://doi.org/10.1016/j.asoc.2020.106401>
- Kramer L. (2022, July 4). *Long position vs. short position: What's the difference?* Investopedia. <https://www.investopedia.com/ask/answers/100314/whats-difference-between-long-and-short-position-market.asp>
- Γεωργούλη Κ. (2015) *Τεχνητή Νοημοσύνη : Μια εισαγωγική προσέγγιση*. ISBN: 978-960-603-031-4