Πανεπιστήμιο Μακεδονίας

Μεταπτυχιακό Εφαρμοσμένων Οικονομικών

Dynamic forecasting for exchange rates

Φιλέλης Θεμιστοκλής

MAE21001

Επιβλέπων

Καθηγητής: Δεργιαδές Θεολόγος

© Copyright συγγραφής Φιλέλης Θεμιστοκλής, 2022.

Περίληψη

Η παρούσα διπλωματική εργασία έχει ως σκοπό τη δημιουργία ενός καθολικού συστήματος το οποίο θα δέχεται σαν είσοδο οποιαδήποτε χρονοσειρά ισοτιμίας και θα προσπαθεί να προβλέψει τις μελλοντικές τις τιμές. Όλη η διαδικασία θα επιτυγχάνεται με τη χρήση νευρωνικών δικτύων και βαθιάς μάθησης.

Τα νευρωνικά δίκτυα όπως και η μηχανική μάθηση είναι ένας κλάδος στον οποίο τα τελευταία χρόνια έχει παρατηρηθεί σημαντική ανάπτυξη. Αυτό γιατί ο απλός προγραμματισμός ακολουθεί έναν αλγόριθμο για να επιλύσει το πρόβλημα ενώ τα νευρωνικά δίκτυα από μόνα τους μαθαίνουν τον τρόπο για να το λύσουν. Με αυτόν τον τρόπο επομένως τα υπολογιστικά συστήματα μπορούν να βρουν λύσεις σε προβλήματα που πολύ πιθανόν εμείς να μην ξέραμε πως. Αυτός είναι και ο σκοπός της μηχανικής μάθησης.

Abstract

The purpose of this thesis is to create a universal system which will accept as input any time series of an exchange rate and will try to predict the future prices. The whole process will be achieved using neural networks and deep learning.

Neural networks, like machine learning, is a field that has seen significant growth in recent years. This is because simple programming follows an algorithm to solve the problem while neural networks learn on their own how to solve it. In this way, therefore, computing systems can find solutions to problems that we most likely did not know how. This is also the purpose of machine learning.

Περιεχόμενα

1……7

Εισαγωγή……..7

1.1 Σημασία του προβλήματος……7

1.2 Στόχος της εργασίας……..7

1.3 Μεθοδολογία προσέγγισης……8

1.4 Ανασκόπηση στη βιβλιογραφία……8

2……9

Ισοτιμία……..9

2.1 Η πρόβλεψη ως έννοια……10

2.2 Νομίσματα και συναλλαγματική ισοτιμία……11

2.3 Η ιστορία των συναλλαγματικών ισοτιμιών……11

2.4 Η αγορά συναλλάγματος σήμερα ……12

2.5 Αντικειμενικός σκοπός της πρόβλεψης ισοτιμίας ……12

2.6 Οι παράγοντες που επηρεάζουν την ισοτιμία θεμελιώδης και τεχνική ανάλυση πρόβλεψης……13

2.7 Η κλασική προσέγγιση - Εργαλεία και μεθοδολογία πρόβλεψης……13

3……15

Νευρωνικά δίκτυα……15

3.1 Περιγραφή νευρωνικών δικτύων……15

3.2 Recurrent Neural Networks……16

3.3 Long short-term memory neural networks……18

4……20

Σχεδιασμός συστήματος……20

4.1 Αρχιτεκτονική συστήματος……20

4.2 Δημιουργία Μοντέλου……21

4.3 Simple LSTM Model – Overview & results……22

4.4 BiDirectional LSTM Model – Overview & results……27

4.5 Stacked LSTM Model – Overview & results……29

4.6 Arima Model – Overview & results……32

5……...……...……...……...……...…...……...……...……...……...35

Συμπεράσματα…….……...……....……...……...…….…...……...….……...35

5.1 Σύνοψη……...……...……...……...……...…….....……...……...…….…....35

5.2 Συμπεράσματα……...……...……...……...……...……..…...……...……....35

5.3 Μελλοντική Εργασία……...……...……...……...……...…...…...……...……....36

1

Εισαγωγή

1.1 Σημασία του προβλήματος

Τα παγκόσμια νομίσματα διαπραγματεύονται καθημερινά το ένα έναντι του άλλου στο ύψος των τριών τρισεκατομμυρίων δολαρίων την ημέρα. Μέσω αυτής της διαπραγμάτευσης κάθε νόμισμα είναι συνδεδεμένο και μετρημένο έναντι του άλλου με συναλλαγματική ισοτιμία. Μια ισοτιμία είναι η τιμή ενός νομίσματος που εκφράζεται σε όρους άλλου νομίσματος. Το ερώτημα που τίθεται είναι: τι προκαλεί οι συναλλαγματικές ισοτιμίες να αλλάξουν και πώς μπορεί κανείς να προβλέψει τη μελλοντική αξία;

Η ανάλυση χρονοσειρών περιλαμβάνει ταυτοποίηση μοντέλου και εκτίμηση παραμέτρων. Οι περισσότερες αναλύσεις θα συμφωνούσαν ότι το πρόβλημα αναγνώρισης είναι πιο δύσκολο. Μόλις η λειτουργική μορφή ενός μοντέλου καθοριστεί, η εκτίμηση των παραμέτρων του μοντέλου είναι συνήθως εύκολη.

Τι θα γινόταν επομένως εάν υπήρχε ένας τρόπος να προσδιορίσουμε ή προσεγγίσουμε μια μελλοντική τιμή της ισοτιμίας. Αυτό το ερώτημα προσπαθεί να εξετάσει η παρούσα διπλωματική εργασία.

1.2 Στόχος της εργασίας και μεθοδολογία προσέγγισης

Στόχος της εργασίας είναι η δημιουργία και η ανάπτυξη ενός υπολογιστικού συστήματος το οποίο θα καταφέρνει να προβλέψει μελλοντικές τιμές οποιαδήποτε ισοτιμίας. Πρωτοτυπία της εργασίας αυτής είναι η υλοποίηση του παραπάνω συστήματος με τη χρήση διαφόρων μεθόδων νευρωνικών δικτύων.

Για την πρόβλεψη των ενδεχόμενων μελλοντικών τιμών της ισοτιμίας, θα χρησιμοποιηθούν κυρίως σύστημα Deep Learning νευρωνικών δικτύων με διάφορες παραλλαγές, έχοντας όλες ως απώτερο σκοπό την πρόβλεψη τιμών με τη μεγαλύτερη δυνατή ακρίβεια.

Η συγκεκριμένη μελέτη προσφέρει ένα δυνατό εργαλείο τόσο για ιδιώτες που ασχολούνται με την υπάρχουσα αγορά αλλά και για μελετητές που ενδεχομένως προσπαθούν να βρουν μία προσέγγιση σε αυτό το δύσκολο πρόβλημα πρόβλεψης.

1.3 Διάρθρωση εργασίας

Η συγκεκριμένη εργασία αρχικά, στο 2ο κεφάλαιο δίνει της εισαγωγικές έννοιες για την έννοια της ισοτιμίας, την διεθνή αγορά αλλά και για τη στοχαστικότητα του προβλήματος.

Στο επόμενο κεφάλαιο δίνονται οι βασικές έννοιες των νευρωνικών δικτύων αλλά και πληροφορίες σχετικά με το είδος των νευρωνικών δικτύων που χρησιμοποιήθηκαν σε αυτή τη διπλωματική εργασία.

Στο 4ο κεφάλαιο παρουσιάζεται η αρχιτεκτονική του συστήματος που υλοποιήθηκε, τα αποτελέσματα των μοντέλων που χρησιμοποιήθηκαν και διάφορα γραφήματα για την ευρύτερη κατανόηση των αποτελεσμάτων.

Στο τελευταίο κεφάλαιο αναλύονται τα συμπεράσματα της εργασίας σε συνδυασμό με ορισμένες προτάσεις για βελτίωση του συγκεκριμένου συστήματος σε μελλοντική εργασία.

1.4 Ανασκόπηση στη βιβλιογραφία

Η πρόβλεψη χρονοσειρών είναι ένα δύσκολο έργο λόγω των μεταβολών των οικονομικών τάσεων καθώς και της έλλειψης πληροφοριών. Η αστάθεια της αγοράς τα τελευταία χρόνια έχει προκαλέσει σοβαρές ανησυχίες. Αυτό είχε ως αποτέλεσμα η προσπάθεια πρόβλεψης των αγορών αλλά και η αξιολόγηση της ακρίβειας των προβλέψεων σαν κλάδος έρευνας να εκτιναχθούν τόσο σημασία όσο και σε επιστημονικό ενδιαφέρον.

Στον χώρο των προβλέψεων χρονοσειρών, μέχρι σήμερα η πιο συνήθης μέθοδος πρόβλεψης είναι η χρήση της μεθόδου ARIMA όπου η μέθοδος αυτοπαλίνδρομου (AR) και η μέθοδος του κινητού μέσου όρου (MA) συνδυάζονται. Στον οικονομικό κλάδο παρ’όλα αυτά εμφανίζεται κατά κόρο η μέθοδος VAR.

Αντίστοιχα, η βαθιά μάθηση έχει γίνει μια από τις πιο ενεργές τεχνολογίες σε πολλούς ερευνητικούς τομείς. Σε αντίθεση με την απλή μηχανική μάθηση, η βαθιά μάθηση (ή αλλιώς deep learning) συνήθως αναφέρεται στη δημιουργία και τοποθέτηση πολλαπλών στρώσεων των νευρωνικών δικτύων.

Ένας διαφορετικός αριθμός στρωμάτων σε ένα δίκτυο μπορεί να προσφέρει ένα διαφορετικό τρόπο προσέγγισης, με απώτερο σκοπό τη βελτίωση του δικτύου αλλά και μία αναβάθμιση στην ικανότητα μάθησης και απόδοση εργασίας [1]. Ειδικά, το μακρύ-βραχυπρόθεσμης μνήμης (LSTM) επαναλαμβανόμενο νευρωνικό δίκτυο (RNN), που εισήχθη αρχικά από τους Hochreiter et al. [3], έχει λάβει τεράστια προσοχή στον κλάδο. Έχουν υπάρξει αποτελεσματικές εφαρμογές που βασίζονται σε δίκτυα LSTM σε διάφορους τομείς όπως η μετάφραση φυσικής γλώσσας [4], δημιουργία λεζάντων για εικόνες[5-7] και αναγνώριση ομιλίας [8]. Τα περισσότερα χειρίζονται την ταξινόμηση και διαχείριση δεδομένων, αλλά στις εφαρμογές παλινδρόμησης τα καθήκοντα είναι σχετικά περιορισμένα.

Αυτό είχε ως συνέπεια οι τεχνικές όλων των νευρωνικών δικτύων να ξεκινήσουν να εμφανίζονται και στον κλάδο όλων των προβλέψεων, αλλά ειδικότερα των οικονομικών προβλέψεων. Συγκεκριμένα, το LSTM έχει χρησιμοποιηθεί στην πρόβλεψη χρονοσειρών [9], [10], [11] και στα οικονομικά και χρηματοοικονομικά δεδομένα όπως ως πρόβλεψη της μεταβλητότητας του S&P 500 [12].

Αυτή η τάση επομένως γεννά πολλά ενδιφέροντα ερωτήματα. Πρώτον, πόσο εύκολα μπορούν να χρησιμοποιηθούν τεχνολογίες deep learning για την πρόβλεψη χρονοσειρών, αλλά και το ερώτημα ποια από τις δύο μεθόδους είναι πιο αποτελεσματική. Παρόμοιες εργασίες έχουν δείξει πως αν και η υλοποίηση νευρωνικών δικτύων είναι μία διαδικασία δύσκολη, τα αποτελέσματα είναι με το μέρος των νευρωνικών δικτύων καθώς φαίνεται να βελτιώνουν κατά πολύ την πρόβλεψη. [14]

Η εργασία αυτή προσπαθεί να δείξει τη διαδικασία δημιουργίας ενός συστήματος με βασικό σκοπό την πρόβλεψη οποιασδήποτε ισοτιμίας αλλά και να δείξει διάφορα αποτελέσματα χρησιμοποιώντας διάφορες τεχνικές LSTM νευρωνικών δικτύων. Ταυτόχρονα ωστόσο χρησιμοποιεί και τις πιο διαδεδομένες τεχνικές προβλέψεων οικονομικών χρονοσειρών. (ARIMA).

2

Ισοτιμία

Σήμερα στην διαρκώς εξελισσόμενη «Ψηφιακή κοινωνία» η «Κοινωνία της Πληροφορίας» η οικονομική δραστηριότητα έχει διευρυνθεί και εμπλουτιστεί σε όλες τις διαστάσεις της. Το διαδίκτυο και οι προηγμένες τεχνολογίες πληροφορικής ως εργαλεία τόσο στον τομέα της αποθήκευσης δεδομένων, όσο και στον τομέα της επεξεργασίας αυτών διαμόρφωσαν και προσδίδουν δυναμική έρευνας στην οικονομική επιστήμη και δυναμική διεύρυνσης στις εφαρμογές της.

2.1 Η πρόβλεψη ως έννοια

Η πρόβλεψη ως έννοια αποδίδει τις ενεργητικές διαδικασίας και την περιγραφή μετρήσιμης εξέλιξης και μετρήσιμης μεταβολής στο μέλλον, βασισμένες στα συνδυασμένα δεδομένα γεγονότα στο παρόν και στο παρελθόν.

Ομοίως η πρόβλεψη ως ενέργεια αποτελεί την γέφυρα διασύνδεσης αναμεσά στην πολύ-παραμετρική σημερινή ανοιχτή αγορά και στην οικονομική δραστηριότητα ατόμων, επιχειρήσεων φορέων και οργανισμών όπως και την δυναμική συσχέτιση του παρελθόντος των στατιστικών δεδομένων και του μέλλοντος των προβλεπόμενων στοιχείων. Επίσης αποτελεί την ελκυστική διασύνδεση στο επιστημονικό-ερευνητικό πεδίο των μαθηματικών, της πληροφορικής και της οικονομικής ανάλυσης, για να αποδώσουν στον σύνθετο πραγματικό κόσμο της αγοράς εργαλεία οικονομικής αποτελεσματικότητας,

Η πρόβλεψη ως εφαρμογή εξειδικεύοντας στην αγορά λειτουργεί ως ένα από τα κύρια συστατικά σε όλους τους τομείς τις αγοράς, σε κάθε επιχειρηματική δραστηριότητα, σε κάθε ατομική δραστηριότητα με οικονομικά χαρακτηριστικά.

Η παρούσα διπλωματική εργασία επιχειρεί να αναμιχθεί στον χώρο των οικονομικών προβλέψεων, αξιοποιώντας τις αρχές της οικονομικής ανάλυσης, τις αρχές της τεχνικής ανάλυσης στον απέραντο κόσμο των συναλλαγματικών ισοτιμιών και αξιοποιώντας την

επιστήμη της πληροφορικής στην ανάλυση και επεξεργασία δεδομένων με αλγορίθμους τεχνητής νοημοσύνης.

2.2 Νομίσματα και συναλλαγματική ισοτιμία

Η ανοιχτή οικονομία σε παγκόσμιο επίπεδο και οι διακρατικές σχέσεις οδηγούν και διαμορφώνουν την επιχειρηματική δραστηριότητα, τις εμπορικές συναλλαγές και τις επενδύσεις έχοντας αναφορά τα διεθνή σύνορα στην αγορά. Οι αναπόφευκτες αυτές εξελίξεις έχουν θεσπίσει ως προϋπόθεση των οικονομικών εφαρμογών την συναλλαγματική ισοτιμία, μια τιμή, που αποδίδει την σχετική τιμή δυο νομισμάτων από διαφορετικές χώρες.

Η ισοτιμία ως όρος περιέχει αφενός ποσοτικά χαρακτηριστικά που είναι η ίση συναλλακτική αξία-τιμή των δύο νομισμάτων αποδεκτή στις δύο χώρες και αφετέρου ποιοτικά χαρακτηριστικά που είναι μια συγκριτική απόδοση σε κάθε εθνική οικονομία.

2.3 Η ιστορία των συναλλαγματικών ισοτιμιών

Τα διαφορετικά νομίσματα των κρατών, τα διαφορετικά χαρακτηριστικά κάθε εθνικής οικονομίας και ο διαφορετικός ρυθμός ανάπτυξης σε κάθε κράτος ανέδειξαν την ανάγκη δημιουργίας ενός διεθνούς συστήματος συναλλαγματικών σχέσεων και την ανάγκη να επισυμβεί με ενιαίο και ρυθμισμένο τρόπο η διεθνής οικονομική ανάπτυξη.

Με βάση δημόσια ιστορικά στοιχεία στην ανάγκη αυτή ξεκίνησε να δώσει λύση η γνωστή συμφωνία του BRETTON WOODS αμέσως μετά τον δεύτερο παγκόσμιο πόλεμο, όπου ο κόσμος βίωνε την μεταπολεμική κοινωνική αναστάτωση, την ασάφεια στην οικονομία και κατά συνέπεια υπήρχε τεράστια αστάθεια στο διεθνές οικονομικό σύστημα.

Η συμφωνία στο Bretton Woods προέβλεπε την συναλλαγματική διασύνδεση του αμερικανικού δολαρίου με τον χρυσό και την κατά συνέπεια συσχέτιση των νομισμάτων των άλλων χωρών με το δολάριο.

Ταυτόχρονα με αποφάσεις του Bretton Woods ιδρύθηκαν το Διεθνές Νομισματικό Ταμείο(ΔΝΤ) και η Παγκόσμια Τράπεζα.

Η συμφωνία αρχικά βοήθησε, όμως οι υψηλού ρυθμού εξελίξεις σε όλους τους τομείς την κατέστησαν ουσιαστικά αδρανή μέχρι το 1971.

Η συνέχεια ήταν η διεθνής αποδοχή ενός κυμαινόμενου συστήματος συναλλαγματικών ισοτιμιών που βασιζόταν στην κλασική προσφορά και ζήτηση.

Οι εξελίξεις στον τραπεζικό τομέα σε συνδυασμό με τις τεχνολογικές εξελίξεις συνέβαλαν στην σταδιακή διόρθωση του εν λόγω συστήματος και την καθιέρωση του μέχρι σήμερα.

2.4 Η αγορά συναλλάγματος σήμερα

Από πολλούς συγγραφείς η διεθνής αγορά συναλλάγματος ορίζεται ως μια αγορά στην οποία συμμετέχοντες από όλο τον κόσμο αγοράζουν και πωλούν διαφορετικά νομίσματα. Οι συμμετέχοντες είναι τράπεζες, εταιρείες, κεντρικές τράπεζες, εταιρείες διαχείρισης επενδύσεων, αμοιβαία κεφάλαια κινδύνου, μεσίτες λιανικής συναλλάγματος και επενδυτές.
Η διεθνής αγορά συναλλάγματος είναι σημαντική επειδή βοηθά στη διευκόλυνση των παγκόσμιων συναλλαγών, συμπεριλαμβανομένων των δανείων, των επενδύσεων, των εταιρικών εξαγορών και του παγκόσμιου εμπορίου.

Σε αντίθεση με την ΧΡΗΜΑΤΗΣΤΗΡΙΑΚΗ ΑΓΟΡΑ, η αγορά συναλλάγματος είναι ΕΞΩΧΡΗΜΑΤΗΣΤΗΡΙΑΚΗ ΚΑΙ ΑΠΟΚΕΝΤΡΩΜΕΝΗ. Ως εξω-χρηματηστηριακή αγορά είναι σχετικά αδιαφανής. Διέπεται από κανόνες και ρυθμιστικές αρχές των διαφορετικών κρατών και όχι από τους κανόνες ενός οργανωμένου χρηματιστηρίου.

Η παγκόσμια αγορά συναλλάγματος είναι σε διαρκή ροή- δεν κλείνει δηλαδή στο τέλος της ημέρας όπως η ΧΡΗΜΑΤΗΣΤΗΡΙΑΚΗ ΑΓΟΡΑ- γεγονός που ερμηνεύει τον ανά «sec» ρυθμό αλλαγής των νομισματικών ισοτιμιών.

Το μέγεθος της αγοράς με βάση αναλύσεις σχετικών οικονομικών εκδόσεων αναφέρεται σε μερικά τρις δολάρια ημερησίως. Σε σύγκριση με τον ημερήσιο όγκο συναλλαγών του χρηματιστηρίου της Νέας Υόρκης το μέγεθος της εξω-χρηματηστηριακής αγοράς συναλλάγματος φαντάζει τεράστιο. Το βέβαιο είναι ότι δεν είναι εύκολο να μετρηθεί επακριβώς.

2.5 Αντικειμενικός σκοπός της πρόβλεψης ισοτιμίας

Η δραστηριότητα στην αγορά συναλλάγματος συνίσταται από τις ενέργειες για την σωστή πρόβλεψη της ισοτιμίας και αφορά στο ζεύγος νομισμάτων που ενδιαφέρει.

Ο συμμετέχων-φορέας, επιχείρηση, επενδυτής-στην αγορά συναλλάγματος αναζητεί για συγκεκριμένο χρονικό διάστημα στο μέλλον την εξέλιξη της ισοτιμίας και αποσκοπεί αφενός στην αποφυγή συναλλαγματικών κινδύνων και αφετέρου στην ευρύτερη εξυπηρέτηση των επιχειρηματικών, εμπορικών, επενδυτικών και οικονομικών του στόχων.

2.6 Οι παράγοντες που επηρεάζουν την ισοτιμία θεμελιώδης και τεχνική ανάλυση πρόβλεψης

Η ισοτιμία στον χρόνο όπως προαναφέραμε διαμορφώνεται στην αγορά με βάση την προσφορά κα την ζήτηση στο νόμισμα που ενδιαφέρει.

Η ισοτιμία επομένως αρχικά επηρεάζεται από την επικαιρότητα, τους επίκαιρους οικονομικούς δείκτες και επιπλέον τα γεγονότα που διαμορφώνουν το εθνικό η διεθνές οικονομικό κλίμα.

Με άλλα λόγια αναφερόμαστε στην θεμελιώδη ανάλυση για την πρόβλεψη, η οποία επιδιώκει να προσδιορίσει την «ψυχολογία» ή το «συναίσθημα» της αγοράς σήμερα και τον τρόπο που επηρεάζουν την εξέλιξη της ισοτιμίας που τελικά απεικονίζεται στην προσφορά και την ζήτηση των νομισμάτων.

Ταυτόχρονα όμως η εξέλιξη των συναλλαγματικών ισοτιμιών ακολουθεί κανόνες, είτε μαθηματικούς και «φυσικούς» νόμους που προσεγγίζουν ή σχετικά μετροποιούν την στοχαστικότητα της αγοράς και που μπορούν να αξιοποιηθούν στις προβλέψεις. Ενδεικτικά θέτουμε την θεωρία του Dow που αναφέρεται στις 3 βασικές κινήσεις στον χρόνο της αγοράς με βάση τις οποίες καθορίζεται η ανοδική ή καθοδική κίνηση, εδώ μιας ισοτιμίας, ο γνωστός αλγόριθμος Fibonacci, το οριακό θεώρημα στην πιθανοτική ανάλυση, ο κλασικός κινούμενος μέσος όρος κ. α.

Στην βάση αυτή αναδεικνύεται η αξία των οικονομικών δεδομένων, των ιστορικών στοιχείων της ισοτιμίας, η μαθηματική αξιοποίηση με την ανάλυση και στατιστική τους επεξεργασία.

Εν προκειμένω αναφερόμαστε στην τεχνική ανάλυση για την πρόβλεψη.

Στην παρούσα διπλωματική εργασία χρησιμοποιείται η τεχνολογία των νευρωνικών δικτύων και της τεχνητής νοημοσύνης, που εξ ορισμού, εμπεριέχουν παραμέτρους τόσο από την θεμελιώδη ανάλυση, όσο και παραμέτρους της τεχνικής ανάλυσης.

2.7 Η κλασική προσέγγιση - Εργαλεία και μεθοδολογία πρόβλεψης

Η κλασσική μεθοδολογία για την πρόβλεψη της ισοτιμίας περιλαμβάνει τις ενότητες:

1. Ανάλυση δεδομένων για το καθορισμό της τάσης των τιμών
2. Ανάλυση δεδομένων που αφορούν στον όγκο των συναλλαγών
3. Ανάλυση δεδομένων για τον υπολογισμό της ισχύος της τάσης και την διάρκεια της.

Τα εργαλεία στην πρόβλεψη είναι διάφορες μορφές απεικόνισης των τιμών σε σχέση με τον χρόνο και ταυτόχρονα εξειδικευμένοι δείκτες βασισμένοι σε μαθηματικές σχέσεις μέσω των οποίων εξειδικεύεται η συμπεριφορά των τιμών.

Η πιο χαρακτηριστική απεικόνιση είναι το γράφημα τιμών-Price Chart με άξονα τιμών και άξονα του χρόνου παραμετροποιημένο σε χρονικές περιόδους (Time Frames), οι οποίες μπορεί να είναι 1 μηνά, 1 εβδομάδα, 1 ημέρα, 4 ώρες, 1 ώρα κλπ.

Υπάρχουν διάφοροι τύποι γραφικής απεικόνισης της μεταβολής των τιμών με πιο χαρακτηριστικά (α) το Line Chart και το Candlestick Chart:

Το Line Chart δεν αποδίδει λεπτομέρειες της συμπεριφοράς των τιμών στο χρονικά διάστημα που επιλέγουμε. Απεικονίζει την γενικευμένη εξέλιξη και είναι καταλληλότερο για την απεικόνιση της τάσης και σημαντικών χρονικών σημείων αλλαγής της συμπεριφοράς των τιμών.

3

Νευρωνικά Δίκτυα

Νευρωνικό δίκτυο ονομάζεται ένα κύκλωμα διασυνδεδεμένων νευρώνων. Στην περίπτωση βιολογικών νευρώνων, πρόκειται για ένα τμήμα νευρικού ιστού. Στην περίπτωση τεχνητών νευρώνων, πρόκειται για ένα αφηρημένο αλγοριθμικό κατασκεύασμα το οποίο εμπίπτει στον τομέα της υπολογιστικής νοημοσύνης. [θεμης2-4]

3.1 Περιγραφή νευρωνικών δικτύων

Από τα τέλη της δεκαετίας του 1980 υπήρξε μια μεγάλη ανάπτυξη στο επίπεδο της ερευνητικής δραστηριότητας στα νευρωνικά δίκτυα. Είναι πλέον σαφές ότι τα νευρωνικά δίκτυα προσφέρουν ένα ισχυρό σύνολο εργαλείων για την επίλυση πολλών προβλημάτων καθώς επίσης μπορούν να θεωρηθούν ως συμπληρωματικά με εκείνα περισσότερων συμβατικών προσεγγίσεων.

Το νευρωνικό δίκτυο είναι ένα δίκτυο από απλούς υπολογιστικούς κόμβους (νευρώνες), διασυνδεδεμένους μεταξύ τους. Είναι εμπνευσμένο από το Κεντρικό Νευρικό Σύστημα (ΚΝΣ), το οποίο προσπαθεί να προσομοιώσει.

Οι νευρώνες είναι το βασικό στοιχείο του δικτύου. Κάθε τέτοιος κόμβος δέχεται ένα σύνολο αριθμητικών εισόδων από διαφορετικές πηγές (είτε από άλλους νευρώνες, είτε από το περιβάλλον). Η έξοδος του, είτε κατευθύνεται στο περιβάλλον, είτε τροφοδοτείται ως είσοδος σε άλλους νευρώνες του δικτύου.

Ένα νευρωνικό δίκτυο αποτελείται από τα εξής

1. Νευρώνες εισόδου
2. Νευρώνες εξόδου

3. Κρυμμένοι νευρώνες.

Τα νευρωνικά δίκτυα είναι εφαρμόσιμα σχεδόν σε κάθε κατάσταση στην οποία ισχύει μια σχέση μεταξύ μεταβλητών πρόβλεψης, ακόμα και όταν αυτή η σχέση είναι πολύ περίπλοκη.

Μερικοί από τους τομείς που εφαρμόζονται τα νευρωνικά είναι:

1. Ιατρική: Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για την αναγνώριση διαφόρων προβλέψεων, ώστε να χορηγηθεί η κατάλληλη θεραπεία.
2. Χρηματιστηριακές προβλέψεις: Τα νευρωνικά δίκτυα χρησιμοποιούνται από πολλούς τεχνικούς αναλυτές, ώστε να κάνουν προβλέψεις σχετικά με τις τιμές των μετοχών καθώς και της ισοτιμίας.
3. Πιστωτική ανάθεση: Τα νευρωνικά δίκτυα μπορούν να εκτοπίσουν τα πιο κατάλληλα και σχετικά χαρακτηριστικά και να τα χρησιμοποιήσει για την ταξινόμηση των αιτούντων ως χαμηλού ή υψηλού κινδύνου.
4. Παρακολούθηση της κατάστασης των μηχανημάτων: Τα νευρωνικά δίκτυα μπορούν να συμβάλλουν στη μείωση του κόστους με την εξασφάλιση της πρόσθετης εμπειρογνωμοσύνης για τον προγραμματισμό προληπτικής συντήρησης των μηχανημάτων.
5. Αυτοκινητοβιομηχανία: Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για την ανάλυση των εισροών που δέχονται οι αισθητήρες ενός κινητήρα.

3.2 Recurrent Neural Networks

Ο ανθρώπινος εγκέφαλος είναι ένα επαναλαμβανόμενο νευρωνικό δίκτυο (RNN): ένα δίκτυο νευρώνων με συνδέσεις ανάδρασης. Μπορεί να εκπαιδευτεί σε διάφορες εργασίες επεξεργασίας ακολουθίας, αλγόριθμους, προγράμματα που δεν μαθαίνονται με τις παραδοσιακές μεθόδους μηχανικής εκμάθησης.

Αυτό εξηγεί το ταχέως αυξανόμενο ενδιαφέρον για τεχνητά RNN για τεχνικές εφαρμογές: γενικοί υπολογιστές που μπορούν να μάθουν αλγόριθμους για να αντιστοιχίσουν ακολουθίες εισόδου σε ακολουθίες εξόδου, με ή χωρίς δάσκαλο.

Οι πρόσφατες εφαρμογές περιλαμβάνουν προσαρμοστική ρομποτική και έλεγχο, αναγνώριση γραφής, αναγνώριση ομιλίας, εντοπισμό λέξεων-κλειδιών, σύνθεση μουσικής, προσεκτική όραση, ανάλυση πρωτεϊνών, πρόβλεψη χρηματιστηρίου και πολλά άλλα προβλήματα ακολουθίας.

Τα προβλήματά τους αναλύθηκαν για πρώτη φορά στο έργο RNN του Schmidhuber. (1991). Ένα δίκτυο ανατροφοδότησης που ονομάζεται Long Short Term Memory (LSTM) ξεπερνά τα θεμελιώδη προβλήματα των παραδοσιακών RNN και μαθαίνει αποτελεσματικά να επιλύει πολλές προηγουμένως άγνωστες εργασίες που περιλαμβάνουν:

1. Αναγνώριση μοτίβων σε θορυβώδεις ακολουθίες εισόδου

2. Αναγνώριση ευρέως διαχωρισμένων συμβάντων σε θορυβώδεις ροές εισόδου

3. Εξαγωγή πληροφοριών από τη χρονική απόσταση μεταξύ των γεγονότων

4. Σταθερή παραγωγή ρυθμών

5. Αποθήκευση πραγματικών αριθμών υψηλής ακρίβειας [θεμης4 - https://people.idsia.ch/~juergen/rnn.html]

Το πρότυπο RNN περιγράφεται ως εξής:

Δεδομένου μίας σειράς διανυσμάτων εισόδου (x1, ... , xT), το RNN υπολογίζει μία ακολουθία κρυφών καταστάσεων (h1, ... , hT) και μια ακολουθία εξάγει (o1, ... , oT) επαναλαμβάνοντας τις ακόλουθες εξισώσεις για t = 1 έως T :

$$\begin{aligned} h_t &= \tanh(W_{hx}x_t + W_{hh}h_{t-1} + b_h) \\ o_t &= W_{oh}h_t + b_o \end{aligned}$$

[2]

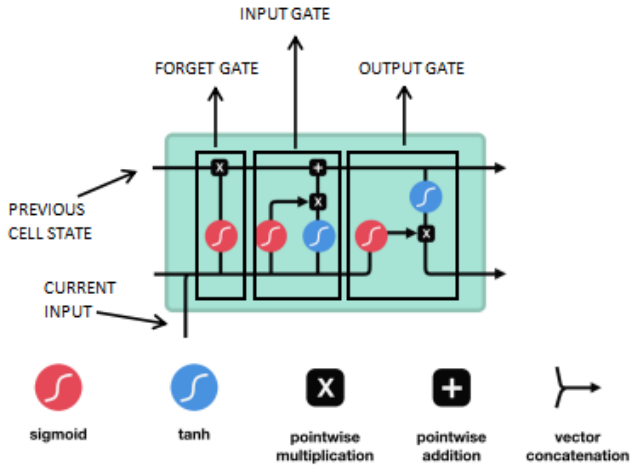
3.3 Long short-term memory neural networks

Από την αρχική εργασία LSTM του 1997, έχουν δημοσιευθεί πολυάριθμες θεωρητικές και πειραματικές εργασίες για το θέμα αυτού του τύπου ενός RNN, πολλοί από αυτούς αναφέρουν τα εκπληκτικά αποτελέσματα που επιτεύχθηκαν σε μια μεγάλη ποικιλία εφαρμογών τομείς όπου τα δεδομένα είναι διαδοχικά.

Ο αντίκτυπος του δικτύου LSTM ήταν αξιοσημείωτος στη μοντελοποίηση γλώσσας, ομιλία σε κείμενο για μετάφραση, αυτόματη μετάφραση και άλλες εφαρμογές. Εμπνευσμένο από τα εντυπωσιακά επιτεύγματα, ορισμένοι αναγνώστες σε ακαδημαϊκό και βιομηχανικό περιβάλλον αποφασίζουν να μάθουν για αυτό το είδος νευρωνικών δικτύων με σκοπό να μετρηθεί η δυνατότητα εφαρμογής του στη δική τους έρευνα ή πρακτική χρήση.

Σύμφωνα επομένως με την παραπάνω παράγραφο, με αυτόν τον τρόπο η παρούσα εργασία προσπαθεί να χρησιμοποιήσει αυτή την τεχνολογία σε έναν κλάδο που η έννοια της πρόβλεψης έχει πολύ μεγάλη σημασία. Αξιοποιώντας τεχνολογίες οι οποίες εφαρμόζονται στην καθημερινότητα γύρω στον κόσμο, έτσι η εργασία αυτή αποτελεί μία προσπάθεια στην εφαρμογή αυτών.

Το κλειδί για τα LSTM είναι η κατάσταση του κελιού, η οριζόντια γραμμή που διατρέχει την κορυφή του διαγράμματος. Το LSTM έχει την ικανότητα να αφαιρεί ή να προσθέτει πληροφορίες στην κατάσταση του κελιού, που ρυθμίζεται από δομές που ονομάζονται πύλες. Οι πύλες είναι ένας τρόπος για να περάσουν προαιρετικά πληροφορίες. Αποτελούνται από ένα στρώμα σιγμοειδούς νευρικού δικτύου και μια πράξη πολλαπλασιασμού κατά σημείο. Η συνάρτηση αυτή εξάγει αριθμούς μεταξύ μηδέν και ενός, περιγράφοντας πόσο από κάθε στοιχείο πρέπει να περάσει. Η τιμή του μηδέν σημαίνει πως το κελί δεν χρειάζεται να αλλάξει καθόλου την κατάσταση του ενώ η τιμή ενός σημαίνει πως η κατάσταση του χρειάζεται ανανέωση. Ένα LSTM έχει τρεις από αυτές τις πύλες, για την προστασία και τον έλεγχο της κατάστασης.



Το πρώτο βήμα στο LSTM είναι να αποφασίσει ποιες πληροφορίες θα αγνοήσει από την κατάσταση του. Αυτή η απόφαση λαμβάνεται από ένα σιγμοειδές στρώμα που ονομάζεται «forget gate». Εξετάζει τα h_{t-1} και x_t και βγάζει έναν αριθμό μεταξύ 0 και 1 για κάθε αριθμό στην κατάσταση κελιού C_{t-1}. Το 1 αντιπροσωπεύει την διατήρηση της πληροφορίας ενώ το 0 αντιπροσωπεύει την αγνόηση.

Το επόμενο βήμα είναι να αποφασίσει ποιες νέες πληροφορίες πρόκειται να αποθηκεύσει στην κατάσταση του νευρώνα. Αυτό έχει δύο μέρη.

1. Πρώτον, ένα σιγμοειδές στρώμα που ονομάζεται input gate αποφασίζει ποιες τιμές θα ενημερώσει.
2. Στη συνέχεια, ένα στρώμα tanh δημιουργεί ένα διάνυσμα νέων υποψήφιων τιμών.

Αμέσως μετά ενημερώνει την παλιά κατάσταση κελιού, C_{t-1}, στη νέα κατάσταση κελιού C_t . Τα προηγούμενα βήματα αποφάσισαν ήδη τι να κάνει ο νευρώνας. Πολλαπλασιάζοντας την παλιά κατάσταση επί f_t , ξεχνώντας τα πράγματα που αποφάσισε να ξεχάσει νωρίτερα. Μετά το προσθέτει στο C_t . Αυτές είναι οι νέες υποψήφιες τιμές, που κλιμακώνονται ανάλογα με το πόσο αποφάσισε να ενημερώσει κάθε τιμή κατάστασης.

Τέλος η έξοδος βασίζεται στην κατάσταση του κελιού, αλλά θα είναι μια φιλτραρισμένη έκδοση. Αρχικά, εκτελεί ένα σιγμοειδές στρώμα το οποίο αποφασίζει ποια μέρη της κατάστασης κελιού θα εξάγει. Στη συνέχεια, περνώντας την κατάσταση του κελιού μέσω tanh (για να μετατρέψει τις τιμές να είναι μεταξύ −1 και 1) και πολλαπλασιάζει με την έξοδο της σιγμοειδούς πύλης, έτσι ώστε να βγάζει μόνο τα μέρη που αποφασίσαμε.

Τα LSTM ήταν ένα μεγάλο βήμα σε σχέση με αυτό που μπορούσαν να πετύχουν τα RNN. Η ιδέα είναι το ίδιο το νευρωνικό σε κάθε του βήμα να επιλέγει πληροφορίες για να τις εξετάσει από κάποια μεγαλύτερη συλλογή πληροφοριών.

4

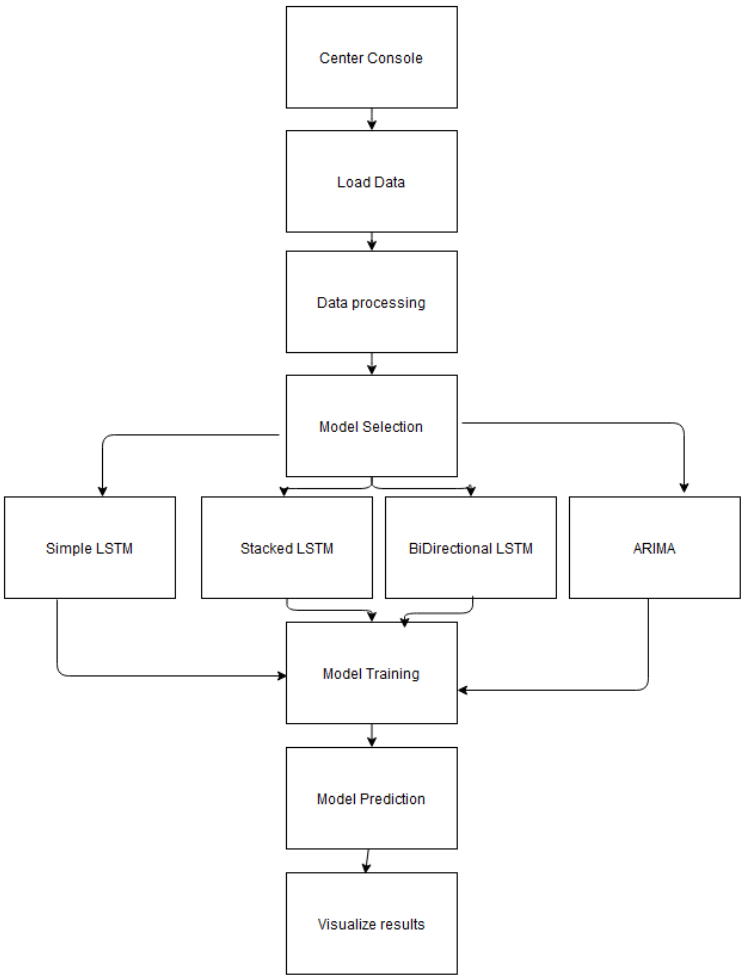
Σχεδιασμός συστήματος

Σε αυτό το κεφάλαιο παρουσιάζεται η αρχιτεκτονική του συστήματος το οποίο υλοποιήθηκε σε γλώσσα προγραμματισμού python. Αρχικά θα παρουσιαστεί η αρχιτεκτονική του συστήματος και στη συνέχεια θα αναλυθούν όλα τα αποτελέσματα με περισσότερες λεπτομέρειες.

4.1 Αρχιτεκτονική συστήματος

Παρακάτω παρουσιάζεται η αρχιτεκτονική του συστήματος ξεκινώντας από τη φόρτωση των δεδομένων μέχρι και την εμφάνιση των αποτελεσμάτων. Εκφρασμένο σε βήματα το σύστημα αποτελείται από τα εξής:

1. Κεντρική κονσόλα
2. Φόρτωση δεδομένων
3. Επεξεργασία δεδομένων
4. Προετοιμασία δεδομένων για την εκπαίδευση των νευρωνικών δικτύων
5. Επιλογή μοντέλου
6. Δημιουργία μοντέλου
7. Εκπαίδευση μοντέλου
8. Πρόβλεψη νευρωνικού δικτύου
9. Δημιουργία διαγραμμάτων

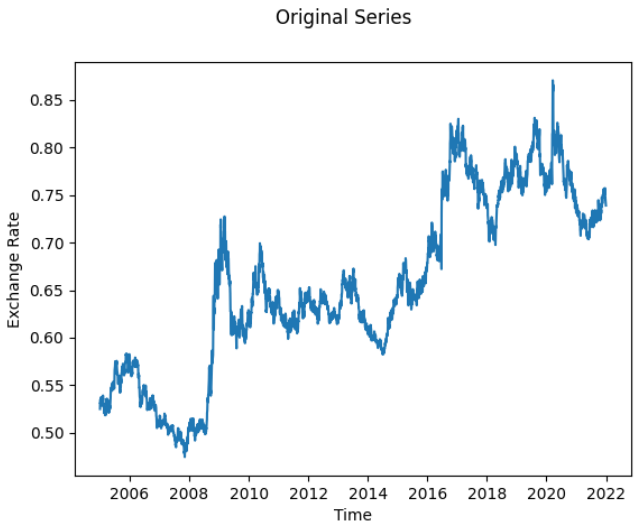


Εικόνα 4.1 – Αρχιτεκτονική συστήματος

4.2 Φόρτωση δεδομένων

Για τα παραδείγματα στην παρούσα διπλωματική εργασία χρησιμοποιήθηκε η ισοτιμία USD-GBP αλλά σύμφωνα με τον τίτλο της εργασίας, μπορούμε να χρησιμοποιήσουμε οποιαδήποτε χρονοσειρά επιθυμούμε.

Στο παρακάτω χρονοδιάγραμμα εξετάζουμε την κανονική μορφή της ισοτημίας.



Το σημαντικότερο σημείο προσοχής είναι η μορφή των δεδομένων στην οποία πρέπει να βρίσκονται έτσι ώστε να καταφέρει το σύστημα να τα διαβάσει. Συγκεκριμένα η μορφή πρέπει να είναι:

1. CSV αρχείο
2. 2 στήλες, DATE & Price

Η συχνότητα της χρονοσειράς επίσης πρέπει να είναι 1 ημέρα. Σε περίπτωση που κάποιος θελήσει να αλλάξει τη συχνότητα, θα πρέπει να προχωρήσει και σε αλλαγές στον πηγαίο κώδικα.

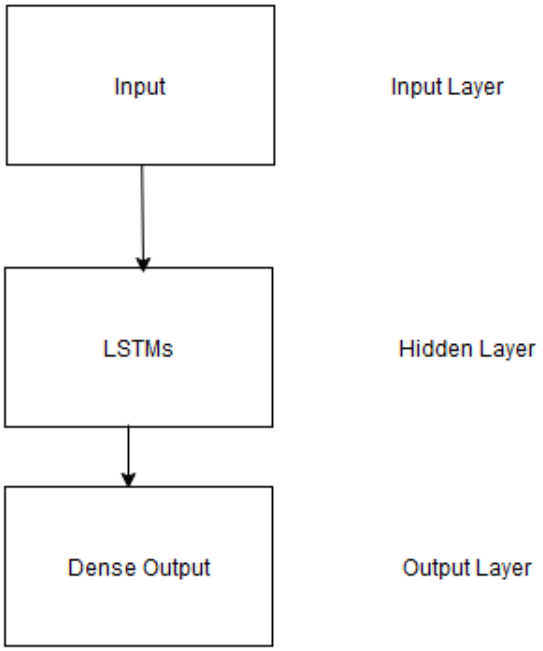
4.2 Δημιουργία Μοντέλου

Τα νευρωνικά δίκτυα, οι προβλέψεις που γίνονται καθώς και οι μετρήσεις είναι όλα εφικτά μέσα από τη χρήση της βιβλιοθήκης Keras. Το Keras, είναι ένα SDK (software development kit) το οποίο χρησιμοποιεί τεχνολογίες από προγραμματιστικές βιβλιοθήκες όπως TensorFlow, CNTK & Theano. Με τον τρόπο αυτό είναι εφικτή η γρήγορη δημιουργία νευρωνικών δικτύων και ο πειραματισμός πάνω σε αυτά.

Αναλυτικότερα θα δούμε τη μορφή κάθε μοντέλου στην κάθε υπό-ενότητά του. Αξίζει να σημειωθεί πως για την καλύτερη επίδοση αλλά και για την παρατήρηση των NN, αξιοποιήθηκαν διάφορες τεχνικές που παρέχει η βιβλιοθήκη Keras, όπως EarlyStopping, με απώτερο σκοπό την αποφυγή άσκοπης περαιτέρω εκπαίδευσης του μοντέλου.

4.3 Simple LSTM Model – Overview & results

Το πρώτο μοντέλο το οποίο υλοποιήθηκε για την παρούσα διπλωματική εργασία ήταν ένα απλό LSTM Model. Η διάρθρωση του, γραφικά φαίνεται στην παρακάτω φωτογραφία.



Εικόμα 4.2 – Αρχιτεκτονική απλού LSTM Model

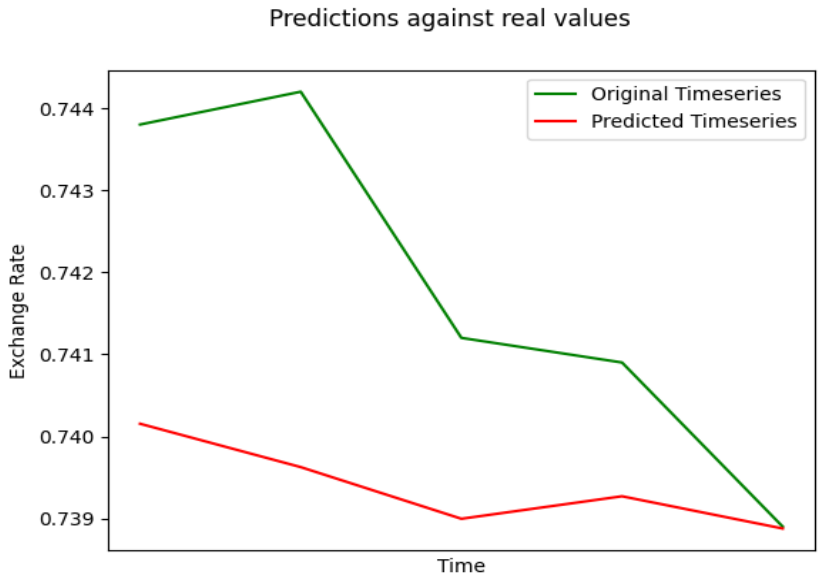
Το μοντέλο αποτελείται από τα εξής στοιχεία:

1. Ένα input layer, το οποίο είναι υπεύθυνο για την εισαγωγή των δεδομένων στο μοντέλο.
2. Το hidden layer, το οποίο για το παρόν μοντέλο αποτελείται από 100 διαφορετικούς νευρώνες LSTM.
3. Ένα Dense Output Layer, το οποίο είναι υπεύθυνο για την εξαγωγή της οποιαδήποτε πληροφορίας έχει εκπαιδευτεί το νευρωνικό δίκτυο.

Για την εκπαίδευση του μοντέλου, η χρονοσειρά της ισοτιμίας χωρίστηκε σε 2 μικρότερες χρονοσειρές με την παρακάτω μορφή:

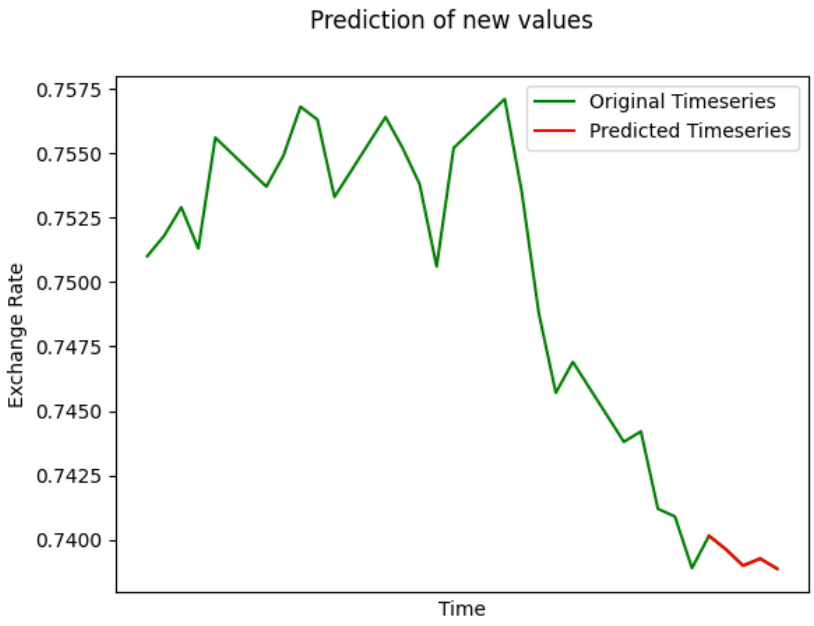
1. Χρονοσειρά εκπαίδευσης μοντέλου, με μέγεθος 90% της αρχικής
2. Χρονοσειρά επαλήθευσης μοντέλου, με μέγεθος 10% της αρχικής

Έπειτα τροφοδοτήθηκε η χρονοσειρά εκπαίδευσης 100 φορές στο μοντέλο με απώτερο σκοπό την τελική εκπαίδευση του μοντέλου. Αμέσως μετά χρησιμοποιήθηκε το μοντέλο για την πρόβλεψη 5 μελλοντικών τιμών των οποίων είχαμε τις πραγματικές με τιμές. Τα αποτελέσματα φαίνονται στο παρακάτω διάγραμμα:



Εικόνα 4.3 – Πρόβλεψη απλού LSTM σε σχέση με πραγματικές τιμές
(In sample forecasting)

Κλείνοντας, το μοντέλο χρησιμοποιήθηκε για να προβλέψει 5 μελλοντικές τιμές, των οποίων οι τιμές δεν του ήταν γνωστές. Το αποτέλεσμα φαίνεται στο παρακάτω διάγραμμα:



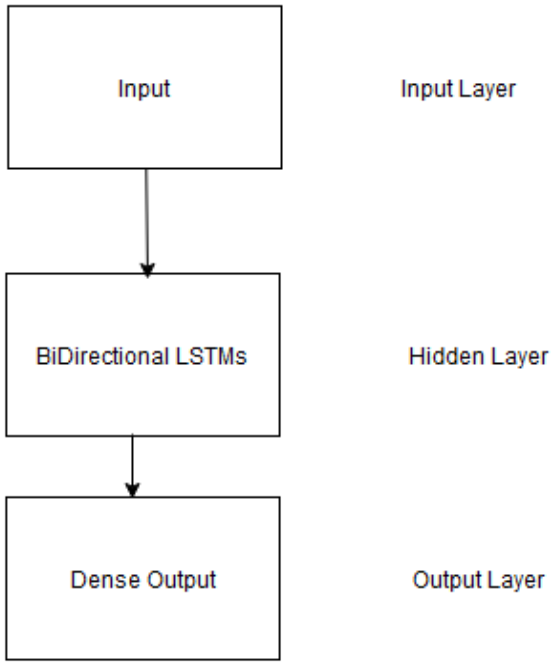
Εικόνα 4.4 – Πρόβλεψη απλού LSTM για 5 χρονικά βήματα (5 ημέρες) (Out of sample forecasting)

Παρατηρείται ότι τα αποτελέσματα που προβλέπει το νευρωνικά δεν αποκλείουν σε μεγάλο βαθμό από τις πραγματικές τιμές. Εδώ βέβαια αξίζει να σημειωθεί ότι το πεδίο τιμών είναι αρκετά μικρό. Αυτό σημαίνει ότι οι δυνατές τιμές διαφέρουν μεταξύ τους κατά πολύ λίγο. Παρ’όλα αυτά, πολύ σημαντικό φαίνεται να είναι το νευρωνικό συμβαδίζει με την τάση των πραγματικών τιμών. Πέφτει πολύ κοντά στο πότε η τιμή ανέβηκε αλλά και πότε μειώθηκε. Αυτό είναι το σημαντικότερο αποτελέσμα.

4.4 BiDirectional LSTM Model – Overview & results

Το δεύτερο μοντέλο το οποίο υλοποιήθηκε για την παρούσα διπλωματική εργασία ήταν ένα BiDirectional LSTM Model. Η διάρθρωση του, γραφικά φαίνεται στην παρακάτω φωτογραφία.

Η διαφορά ενός BiDirectional LSTM Model με ένα απλό είναι η δομή του νεύρωνα – κόμβου. Ένας κόμβος BiDirectional αποτελείτε από 2 LSTM τα οποία κάνουν την ακριβώς αντίθετη λειτουργία μεταξύ τους. Δηλαδή το ένα LSTM μεταφέρει πληροφορία προς το μπροστινό μέρος του νευρωνικού, και αντίστροφά το άλλο προς το πίσω μέρος.



Εικόμα 4.5 – Αρχιτεκτονική BiDirectional LSTM Model

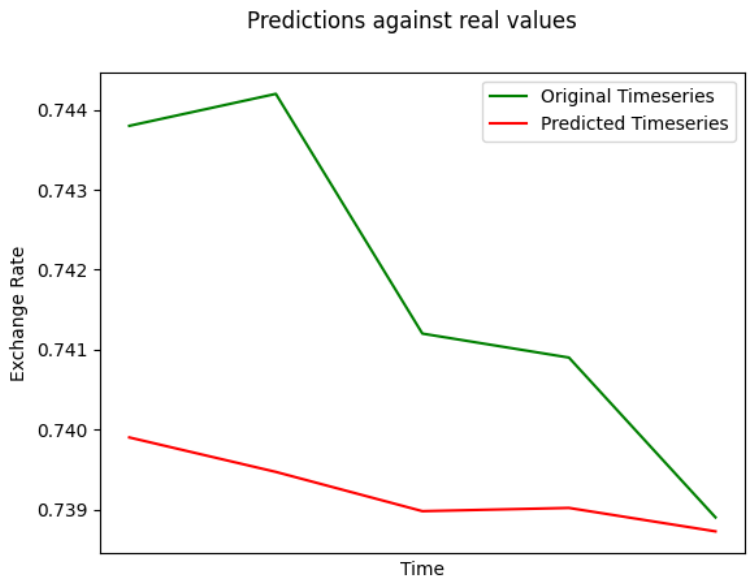
Το μοντέλο αποτελείται από τα εξής στοιχεία:

1. Ένα input layer, το οποίο είναι υπεύθυνο για την εισαγωγή των δεδομένων στο μοντέλο.
2. Το hidden layer, το οποίο για το παρόν μοντέλο αποτελείται από 50 διαφορετικούς νευρώνες BiDirectional.
3. Ένα Dense Output Layer, το οποίο είναι υπεύθυνο για την εξαγωγή της οποιαδήποτε πληροφορίας έχει εκπαιδευτεί το νευρωνικό δίκτυο.

Για την εκπαίδευση του μοντέλου, η χρονοσειρά της ισοτιμίας χωρίστηκε σε 2 μικρότερες χρονοσειρές ως εξής:

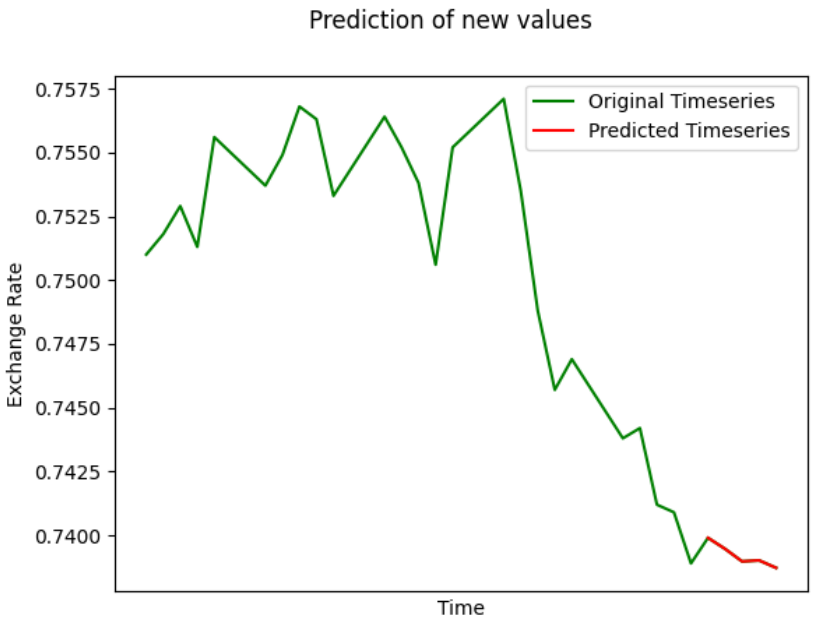
1. Χρονοσειρά εκπαίδευσης μοντέλου, με μέγεθος 90% της αρχικής
2. Χρονοσειρά επαλήθευσης μοντέλου, με μέγεθος 10% της αρχικής

Έπειτα τροφοδοτήθηκε η χρονοσειρά εκπαίδευσης 100 φορές στο μοντέλο με απώτερο σκοπό την τελική εκπαίδευση του μοντέλου. Αμέσως μετά χρησιμοποιήθηκε το μοντέλο για την πρόβλεψη 5 μελλοντικών τιμών των οποίων είχαμε τις πραγματικές με τιμές. Τα αποτελέσματα φαίνονται στο παρακάτω διάγραμμα:



Εικόνα 4.6 – Πρόβλεψη BiDirectional LSTM σε σχέση με πραγματικές τιμές (In sample forecasting)

Τελικά, στο παρακάτω γράφημα βλέπουμε τις 5 μελλοντικές τιμές, των οποίων οι τιμές δεν του ήταν γνωστές:



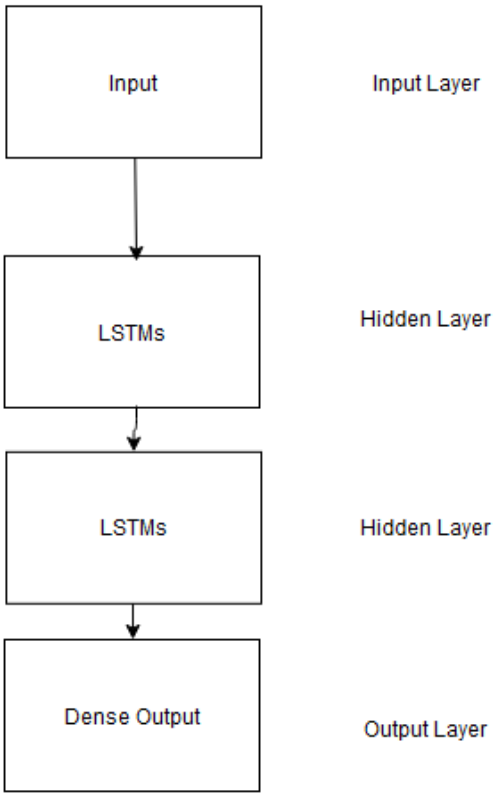
Εικόνα 4.7 – Πρόβλεψη BiDirectional LSTM για 5 χρονικά βήματα (5 ημέρες) (Out of sample forecasting)

Τα αποτελέσματα του παρατηρούμε ότι δεν διαφέρουν σχεδόν καθόλου με αυτά ενός απλού δικτύου. Όπως και προηγουμένως, το συγκεκριμένο είδος νευρωνικού έχει την δυνατότητα να πλησιάσει πολύ κοντά στις πραγματικές τιμές αλλά ταυτόχρονα καταφέρνει και ανιχνεύει την τάση της χρονοσειράς. Αυτό το καθιστά μία πολύ καλή λύση σε θέματα πρόβλεψης ισοτιμίας όπως και το προηγούμενο μοντέλο.

4.5 Stacked LSTM Model – Overview & results

Το τρίτο και τελικό νευρωνικό μοντέλο το οποίο υλοποιήθηκε για την παρούσα διπλωματική εργασία ήταν ένα Stacked LSTM Model. Η διάρθρωση του, γραφικά φαίνεται στην παρακάτω φωτογραφία.

Η διαφορά του με ένα απλό μοντέλο είναι το γεγονός ότι στο Hidden Layer αποτελείται από δύο στρώματα LSTM νευρώνων.



Εικόμα 4.8 – Αρχιτεκτονική Stacked LSTM Model

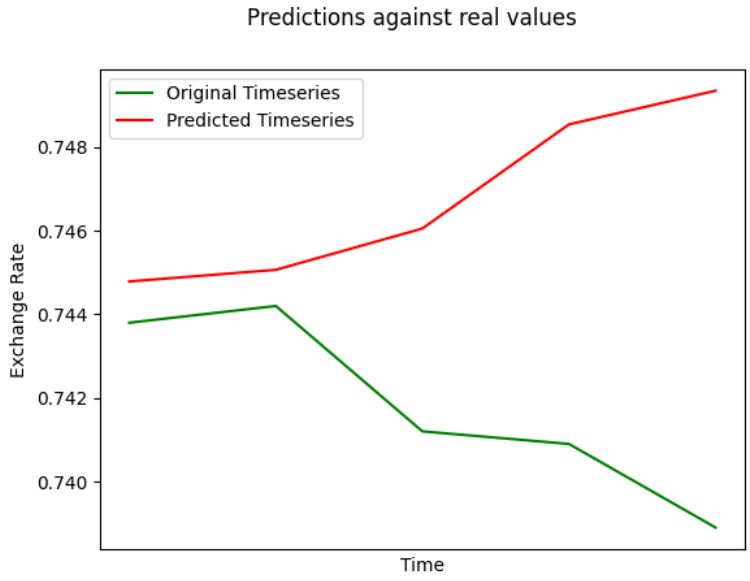
Το μοντέλο αποτελείται από τα εξής στοιχεία:

1. Ένα input layer, το οποίο είναι υπεύθυνο για την εισαγωγή των δεδομένων στο μοντέλο.
2. Το hidden layer, το οποίο για το παρόν μοντέλο αποτελείται από ένα δίκτυο 50 διαφορετικούς νευρώνες LSTM και στη συνέχεια από ένα άλλο δίκτυο 25 διαφορετικών νευρώνων.
3. Ένα Dense Output Layer, το οποίο είναι υπεύθυνο για την εξαγωγή της οποιαδήποτε πληροφορίας έχει εκπαιδευτεί το νευρωνικό δίκτυο.

Για την εκπαίδευση του μοντέλου, η χρονοσειρά της ισοτιμίας χωρίστηκε σε 2 μικροτέρες χρονοσειρές ως εξής:

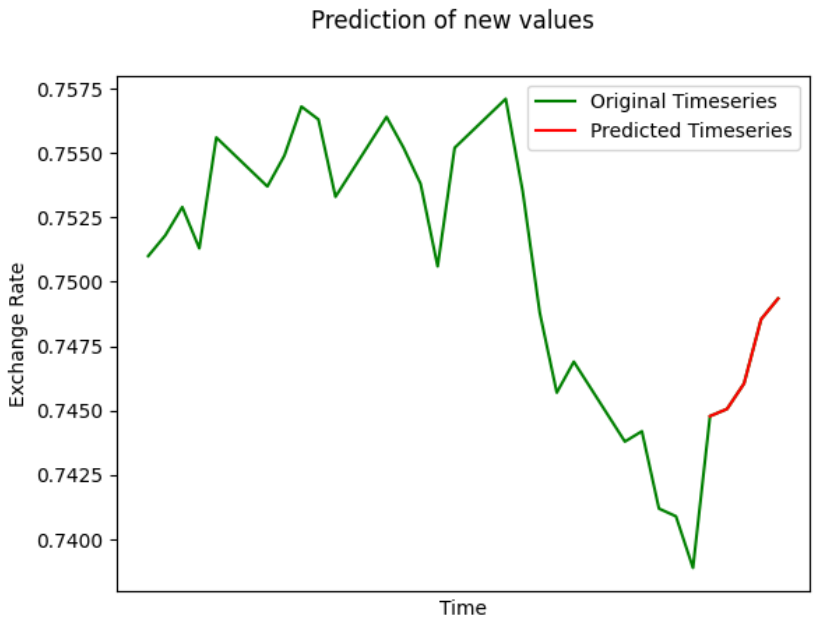
1. Χρονοσειρά εκπαίδευσης μοντέλου, με μέγεθος 90% της αρχικής
2. Χρονοσειρά επαλήθευσης μοντέλου, με μέγεθος 10% της αρχικής

Έπειτα τροφοδοτήθηκε η χρονοσειρά εκπαίδευσης 100 φορές στο μοντέλο με απότερο σκοπό την τελική εκπαίδευση του μοντέλου. Αμέσως μετά χρησιμοποιήθηκε το μοντέλο για την πρόβλεψη 5 μελλοντικών τιμών των οποίων είχαμε τις πραγματικές με τιμές. Τα αποτελέσματα φαίνονται στο παρακάτω διάγραμμα:



Εικόνα 4.9 – Πρόβλεψη Stakced LSTM σε σχέση με πραγματικές τιμές (In sample forecasting)

Τελικά, το μοντέλο κλήθηκε να προβλέψει 5 μελλοντικές τιμές, των οποίων οι τιμές δεν του ήταν γνωστές. Το αποτέλεσμα φαίνεται στο παρακάτω διάγραμμα:



Εικόνα 4.10 – Πρόβλεψη Stacked LSTM για 5 χρονικά βήματα (5 ημέρες) (Out of sample forecasting)

Παρατηρείται ότι τα αποτελέσματα του συγκεκριμένου είδους νευρωνικού απέχουν κατά πολύ από τις πραγματικές τιμές σε σχέση με τα προηγούμενα μοντέλο. Αυτό οφείλεται στο γεγονός ότι το μοντέλο είναι πολύ περισσότερο ιδιόμορφο για το πράγμα το οποίο καλείται να λύσει, δηλαδή την πρόβλεψη της ισοτιμίας. Σε περίπτωση που υπήρχαν παραπάνω δεδομένα, όχι μόνο μεγαλύτερη χρονοσειρά αλλά και περισσότερα δεδομένα τότε ίσως το μοντέλο απέδιδε καλύτερα.

4.6 Arima Model – Overview & results

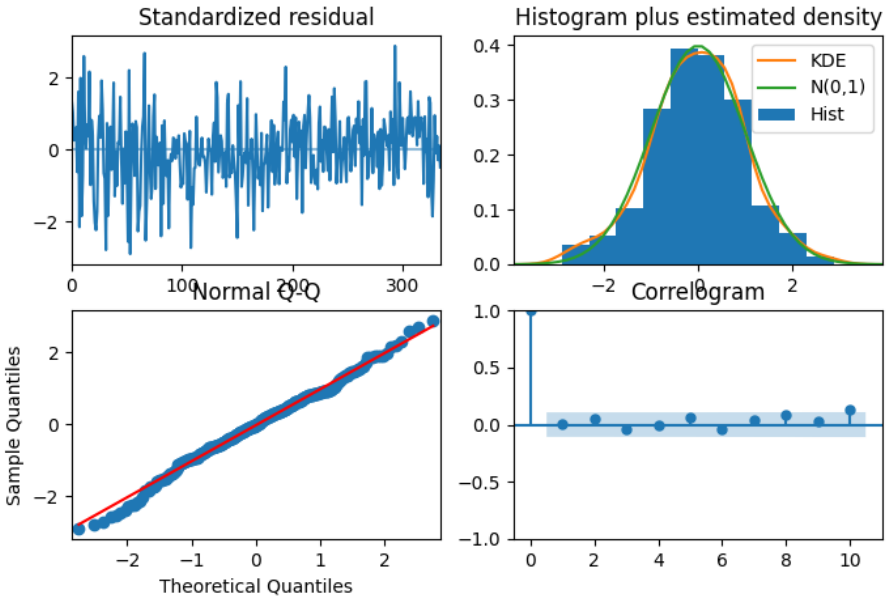
Τέλος υλοποιήθηκε και ένα μοντέλο ARIMA, χρησιμοποιώντας την βιβλιοθήκη autoarima της γλώσσας Python, με σκοπό την σύγκριση ενός γνωστού μοντέλου προβλέψεων και ενός νευρωνικού δικτύου.

Τα αποτελέσματα του Arima μοντέλου φαίνονται παρακάτω:

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                                     y   No. Observations:                365
Model:             SARIMAX(2, 0, 1)x(0, 1, [1, 2], 7)   Log Likelihood               1499.281
Date:                                Sun, 31 Jul 2022   AIC                         -2984.562
Time:                                        19:13:21   BIC                         -2957.398
Sample:                                             0   HQIC                        -2973.759
                                                - 365
Covariance Type:                                  opg
==========================================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept  -5.565e-05   8.76e-05     -0.635      0.525      -0.000       0.000
ar.L1          0.3062      0.303      1.012      0.312      -0.287       0.900
ar.L2          0.5691      0.291      1.953      0.051      -0.002       1.140
ma.L1          0.7034      0.270      2.609      0.009       0.175       1.232
ma.S.L7       -0.8423      0.066    -12.804      0.000      -0.971      -0.713
ma.S.L14       0.0875      0.060      1.458      0.145      -0.030       0.205
sigma2       1.32e-05   8.99e-07     14.690      0.000    1.14e-05     1.5e-05
===================================================================================
Ljung-Box (L1) (Q):                   2.48   Jarque-Bera (JB):                15.09
Prob(Q):                              0.12   Prob(JB):                         0.00
Heteroskedasticity (H):               0.48   Skew:                            -0.24
Prob(H) (two-sided):                  0.00   Kurtosis:                         3.88
===================================================================================
```

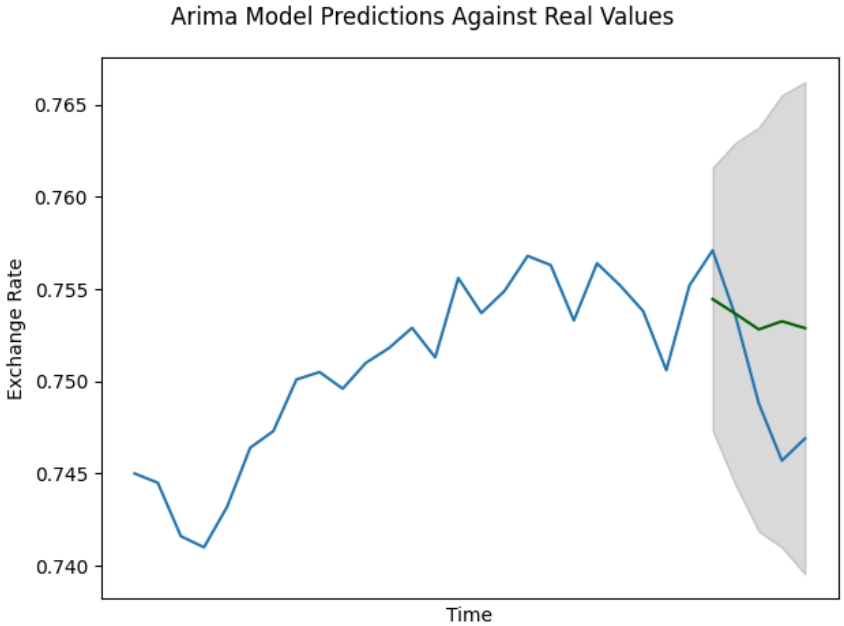
Εικόνα 4.11 – Arima model

Εξετάζοντας τα κατάλοιπα, δεν παρατηρείται κάτι το οποίο μπορεί να καταστήσει το μοντέλο μας λανθασμένο:



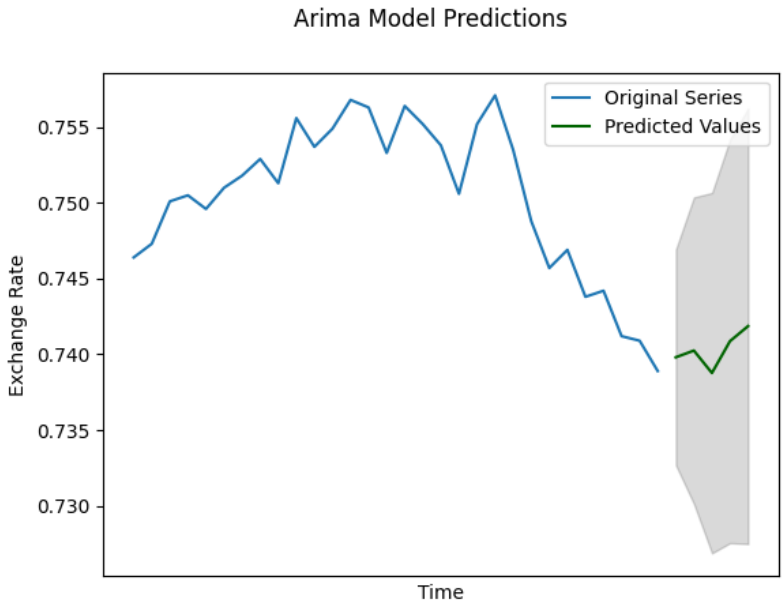
Εικόνα 4.12 – Κατάλοιπά arima model

Επομένως, έχοντας το μοντέλο, σαν πρώτο βήμα προσπαθεί να προβλέψει μελλοντικές τιμές τις οποίες οι τιμές τους βρίσκονται στα δεδομένα, με αποτελέσματα:



Εικόνα 4.13 – Πρόβλεψη Arima σε σχέση με πραγματικές τιμές (In sample forecasting)

Τέλος, το μοντέλο καλείται να προβλέψει μελλοντικές τιμές οι οποίες φαίνονται στο παρακάτω διάγραμμα:



Εικόνα 4.10 – Πρόβλεψη arima για 5 χρονικά βήματα (5 ημέρες) (Out of sample forecasting)

5

Συμπεράσματα και προοπτικές

5.1 Σύνοψη

Το σύστημα που αναπτύχθηκε στα πλαίσια της διπλωματικής εργασίας, είναι ένα σύνολο μοντέλων πρόβλεψης χρονοσειρών ισοτιμίας. Το παρόν σύστημα μπορεί να χρησιμοποιηθεί από τον οποιοδήποτε και να αποτελέσει ένα αρκετά δυνατό εργαλείο στα πλαίσια μίας έρευνας ή και ακόμα εφαρμογής.

Για να γίνει πραγματικότητα αυτό, οποιοδήποτε μοντέλο υλοποιήθηκε εκπαιδεύεται με βάση το σύνολο δεδομένων που του τροφοδοτείται. Στη συνέχεια, έχοντας αναγνωρίσει τα διάφορα μοτίβα που μπορεί να υπάρχουν, προβλέπει τόσο τιμές εντός του συνόλου δεδομένων όσο και τιμές για 5 ημέρες μετά το τέλος της χρονοσειράς.

5.2 Συμπεράσματα

Είναι επομένως τα νευρωνικά δίκτυα ικανά για να χρησιμοποιηθούν για αυτή την κατηγορία προβλέψεων;

Η απάντηση είναι ναι, αλλά σε καμία περίπτωση δεν πρέπει να είναι η μοναδική πηγή αποτελεσμάτων. Η φύση των νευρωνικών δικτύων σε συνδυασμό με μία χρονοσειρά, οδηγεί σε προβλήματα τα οποία εκ φύσεως είναι δύσκολο να λυθούν. Ακόμα και εάν οι μετρικές του δικτύου δείχνουν πολύ θετικά αποτελέσματα, αυτό δεν σημαίνει πως και το μοντέλο θα δίνει τόσο ακριβείς προβλέψεις.

Το μεγαλύτερο εμπόδιο στις έγκυρες προβλέψεις χρονοσειρών ισοτιμίας (είτε χρηματιστηρίου) είναι ο παράγοντας της στοχαστικότητας. Κανένα μοντέλο απ’ αυτά που

χρησιμοποιήθηκαν και από αυτά τα οποία είναι διαθέσιμα δεν συμπεριλαμβάνουν σε κανένα βαθμό τη στοχαστικότητα των προβλέψεων.

Ο μοναδικός κλάδος των μαθηματικών ο οποίος ασχολείται με τη στοχαστικότητα και προσπαθεί να την αντιμετωπίσει / μετριάσει είναι ο κλάδος των Σημάτων. Αυτό συμπεραίνεται από το γεγονός ότι όλες οι σύγχρονες θεωρίες που προσπαθούν να αντιμετωπίσουν προβλήματα προβλέψεων σε χρονοσειρές, χρησιμοποιούν θεωρίες σημάτων.

Επίσης η έλλειψη επιπλέον δεδομένων αποτελεί ένα δύσκολο πρόβλημα. Δεν είναι εύκολο να χρησιμοποιηθούν επιπλέον δεδομένα, καθώς αυτό θα εμπόδιζε τις προβλέψεις για μελλοντικές τιμές περνάς του Χτ + 1, οπού Χτ η τελευταία παρατήρηση.

Το μοναδικό δεδομένο το οποίο θα μπορούσαμε να χρησιμοποιήσουμε άφοβα χωρίς να μας δημιουργεί προβλήματα σε out of sample προβλέψεις είναι ο χρόνος. Το ερώτημα που τίθεται εδώ παρ' όλα αυτά, είναι πως μετατρέπουμε το χρόνο σε μαθηματική αναπαράσταση, έτσι ώστε ο αλγόριθμος να μπορέσει να τον χρησιμοποιήσει σαν πηγή πληροφορίας. Ένας τρόπος να επιτευχθεί κάτι τέτοιό είναι να μετατραπεί ο χρόνος σε ένα διάνυσμα. Αυτό είναι πλέον μια καινούργια προσέγγιση στο λύση παρόμοιων προβλημάτων όπου ο χρόνος καθιστά βασική πηγή πληροφορίας. Αυτό επιτυγχάνεται με τη χρήση μετασχηματισμών Fourier.

5.3 Μελλοντική Εργασία

Το σύστημα που υλοποιήθηκε στα πλαίσια της διπλωματικής εργασίας, μπορεί να αποτελέσει πάτημα για νέες έρευνες και εργασίες. Πιο συγκριμένα:

1) Μετατροπή του χρόνου σε διάνυσμα και τροφοδότηση του δεδομένου αυτού στο μοντέλο. Αυτό θα αποτελούσε εξαιρετική εξέλιξη αυτής της εργασίας, καθώς όπως έχει προαναφερθεί, ο χρόνος αποτελεί το μοναδικό επιπλέον δεδομένο το οποίο μπορεί να χρησιμοποιηθεί με σιγουριά.

2) Χρήση ισοτιμίας με μεγαλύτερο εύρος τιμών.

3) Χρήση νευρωνικών δικτύων Transformer. Τα Transformer αποτελούν μία καινούρια τεχνολογία, τα οποία είναι ικανά να υιοθετήσουν τον μηχανισμό της «αυτοπροσοχής» (self-attention), δίνοντας βάρος στη θέση κάθε δεδομένου εισόδου. Χρησιμοποιείται κυρίως στους τομείς της επεξεργασίας φυσικής γλώσσας.

Βιβλιογραφία – Αναφορές

[1] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, pp. 1798-1828, 2013

[2] Generating Text with Recurrent Neural Networks, Ilya Sutskever, James Martens, Geoffrey Hinton.

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, pp. 1735-1780, 1997

[4] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in Advances in neural information processing systems, 2014, pp. 3104-3112.

[5] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3156-3164.

[6] A. Karpathy and L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3128-3137

[7] J. Mao, W. Xu, Y. Yang, J. Wang, Z. Huang, and A. Yuille, "Deep captioning with multimodal recurrent neural networks (m-RNN)," arXiv preprint arXiv:1412.6632, 2014.

[8] A. Graves and N. Jaitly, "Towards End-To-End Speech Recognition with Recurrent Neural Networks," in ICML, 2014, pp. 1764-1772

[9] J. Brownlee, “Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras,” https://machinelearningmastery.com/ time-series-prediction-lstm-recurrent-neural-networks-python-keras/,

[10] F. A. Gers, J. Schmidhuber, F. Cummins, “Learning to Forget: Continual Prediction with LSTM,” in Neural Computation 12(10): 2451-2471, 2000

[11] Schmidhuber, “Deep learning in neural networks: An overview,” in Neural Networks, 61: 85-117, 2015.
bibitemSima-2018A S. Siami-Namini, D. Muhammad, F. Fahimullah, “The Short and Long Run Effects of Selected Variables on Tax Revenue - A Case Study,” in Applied Economics and Finance, 5(5), p. 23-32, 2018.

[12] N. Huck, “Pairs Selection and Outranking: An Application to the S&P 100 Index,” in European Journal of Operational Research 196(2): 819-825, 2009.

[13] N. Tavakoli, D. Dai, Y. Chen, “Log-assisted straggler-aware I/O scheduler for high-end computing,” in 45th International Conference on Parallel Processing Workshops (ICPPW), 2016.

[14] Sima Siami-Namin, A Comparison of ARIMA and LSTM in Forecasting Time Series